IBM b-type Data Center Networking

Product Introduction and Initial Setup

- Learn about the products in the IBM b-type portfolio
- Discover how to quickly perform an initial setup
- Read about advanced features

Jon Tate
Andrew Bernoth
Ivo Gomilsek
Peter Mescher
Steven Tong

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM b-type Data Center Networking: Product Introduction and Initial Setup

June 2010

SG24-7785-00

Note: Before using this information and the product it supports, read the information in "Notices" on page xi.

First Edition (June 2010)

This edition applies to the supported products in the IBM b-type portfolio in September 2009.

Note: This book is based on a pre-GA version of a product and might not apply when the product becomes generally available. Consult the product documentation or follow-on versions of this book for more current information.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

IBM®
Redbooks®
Redbooks (logo) ®
System Storage™
Tivoli®
xSeries®

The following terms are trademarks of other companies:

Microsoft, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

As organizations drive to transform and virtualize their IT infrastructures to reduce costs, and manage risk, networking is pivotal to success. Optimizing network performance, availability, adaptability, security, and cost is essential to achieving the maximum benefit from your infrastructure.

In this IBM® Redbooks® publication, we address the requirements:

- Expertise to plan and design networks with holistic consideration of servers, storage, application performance and manageability
- Networking solutions that enable investment protection with a range of performance and cost options that match your environment
- Technology and expertise to design and implement and manage network security and resiliency
- Robust network management software to provide integrated, simplified management that lowers operating costs of complex networks

IBM and Brocade have entered into an agreement to provide expanded network technology choices with the new IBM b-type Ethernet Switches and Routers, to provide an integrated end-to-end resiliency and security framework.

Combined with the vast data center design experience of IBM and the networking expertise of Brocade, this portfolio represents the ideal convergence of strength and intelligence. For organizations striving to transform and virtualize their IT infrastructure, such a combination can help you reduce costs, manage risks, and prepare for the future.

In this book, we introduce the products and the highlights of the IBM b-type portfolio from a viewpoint of design and suggested practices.

This book is meant to be used in conjunction with *IBM b-type Data Center Networking: Design and Best Practices Introduction*, SG24-7786.

We realize that the scope of this subject is enormous, so we have concentrated on certain areas only. However, as our portfolio matures, we intend to update this book to include new areas, and revisit other areas.

Be sure to let us know of any additions you want to see in this book, because we always welcome fresh ideas.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center, and at Brocade Communications, San Jose.

Jon Tate is a Project Manager for IBM System Storage™ Networking and Virtualization Solutions at the International Technical Support Organization, San Jose Center. Before joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2 support for IBM storage products. Jon has 25 years of experience in storage software and management, services, and support, and is both an IBM Certified IT Specialist and an IBM SAN Certified Specialist. He is also the UK Chairman of the Storage Networking Industry Association.

Andrew Bernoth is the IBM Network Services Lead Architect for the Asia Pacific region based out of Melbourne, Australia. Prior to this, Andrew worked on global architecture and security standards for the IBM services extranet environment. He has 20 years of experience in computing, over 15 years of which has been focused on network and security. Andrew holds GSEC and CISSP security certifications as well as IBM Certified IT Architect. His work on a security checking program for communication between networks was awarded a patent in 2008.

Ivo Gomilsek is a Certified IT Architect working in IBM Austria as a solution architect for STG Sales in CEE/CEEMEA region. His responsibilities include architecting, consulting, deploying, and supporting infrastructure solutions. His areas of expertise include SAN, storage, networking, HA systems (GDPS/GDOC), cross platform server and storage consolidation, and open platform operating systems. Ivo holds several certifications from various vendors (IBM, Red Hat, Microsoft®, Symantec, and VMware). Ivo has contributed to various other Redbooks publications on Tivoli® products, SAN, storage subsystems, Linux/390, xSeries®, and Linux®.

Peter Mescher is a Product Engineer on the SAN Central team within the IBM Systems and Technology Group in Research Triangle Park, North Carolina. He has seven years of experience in SAN Problem Determination and SAN Architecture. Before joining SAN Central, he performed Level 2 support for network routing products. He is a co-author of the SNIA Level 3 FC Specialist Exam. This is his sixth Redbooks publication.

Steven Tong is a corporate Systems Engineer for Brocade focused on qualification and solutions development of IBM Data Center Networking products. His areas of expertise include Storage Area Networks (SAN) as well as Ethernet and IP based networks.

Thanks to the following people for their contributions to this project:

Brian Steffler
Marcus Thordal
Kamron Hejazi
Mike Saulter
Jim Baldyga
Brocade

Holger Mueller
Pete Danforth
Casimer DeCusatis
Doris Konieczny
Aneel Lakhani
Mark Lewis
Tom Parker
Steve Simon
IBM

Emma Jacobs
International Technical Support Organization, San Jose Center

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author - all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks publications form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks publications weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Product introduction

In this chapter we discuss the IBM Networking b-type family of IBM networking products. We describe various products of the b-type family, including their functions and components, including IBM versus Brocade naming conventions.

We cover the following products:

- IBM m-series Ethernet/IP Routers
- IBM r-series Ethernet Switches
- IBM x-series Ethernet Switches
- IBM c-series Ethernet Switches
- IBM s-series Ethernet Switches
- IBM g-series Ethernet access switches

The b-type family is shown in Figure 1-1.

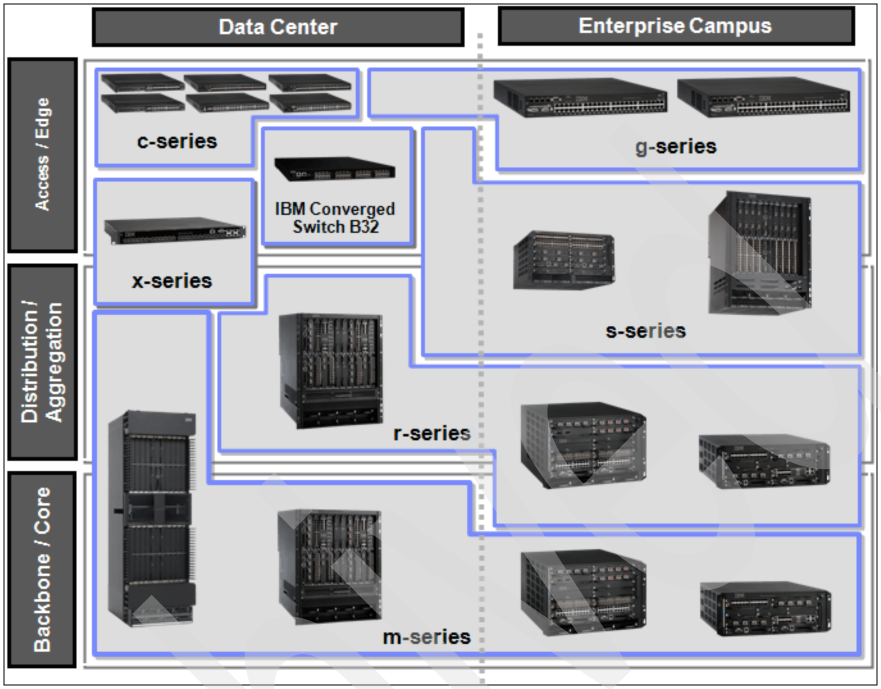


Figure 1-1 IBM b-type Data Center Networking family

1.1 Product overview

In the sections that follow, we describe the IBM Networking b-type family of IBM networking products. For the most up to date information, see the website:

http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/

1.1.1 Product features

IBM Networking b-type Ethernet routers and switches provide all the functionality required to build an efficient dynamic infrastructure and transform physical and digital assets into more relevant services. The b-type family was engineered to provide performance, better manage risk, improve service and reduce cost from the end user to the server connectivity:

- IBM m-series Ethernet/IP Routers are designed to meet the most demanding requirements of data center core and border, as well as enterprise campus backbone and border layer solutions.
- IBM r-series Ethernet Switches are designed to meet the most demanding requirements of the data center aggregation and distribution, as well as campus backbone and border layer solutions.
- IBM x-series Ethernet Switches are designed to meet the most demanding requirements of the data center Top-of-Rack (TOR) server access, cost-efficient aggregation, and high performance computing solutions.
- IBM c-series Ethernet Switches are designed to meet the most demanding requirements of data center TOR server access, metro, and border edge layer solutions.
- IBM s-series Ethernet Switches are designed to meet critical cost requirements of data center aggregation and enterprise campus Power over Ethernet (PoE) access and distribution layer solutions.
- IBM g-series Ethernet Switches are designed to meet critical cost requirements of enterprise campus stackable and PoE converged access layer solutions.

1.1.2 Naming convention: IBM versus Brocade

In Table 1-1 we list the IBM Networking b-type family products, along with their equivalent Brocade names. Note that we reference the products by their standard IBM names as well as the IBM type/model throughout this text.

Table 1-1 IBM Networking b-type family

IBM family name	Brocade family name	IBM product name	IBM machine type and model	Brocade name
IBM m-series Ethernet/IP Routers	Brocade NetIron MLX	IBM Ethernet/IP Router B04M	4003-M04	Brocade NetIron MLX-4
IBM m-series Ethernet/IP Routers	Brocade NetIron MLX	IBM Ethernet/IP Router B08M	4003-M08	Brocade NetIron MLX-8
IBM m-series Ethernet/IP Routers	Brocade NetIron MLX	IBM Ethernet/IP Router B16M	4003-M16	Brocade NetIron MLX-16
IBM m-series Ethernet/IP Routers	Brocade NetIron MLX	IBM Ethernet/IP Router B32M	4003-M32	Brocade NetIron MLX-32
IBM r-series Ethernet Switches	Brocade BigIron RX	IBM Ethernet Switch B04R	4003-R04	Brocade BigIron RX-4
IBM r-series Ethernet Switches	Brocade BigIron RX	IBM Ethernet Switch B08R	4003-R08	Brocade BigIron RX-8
IBM r-series Ethernet Switches	Brocade BigIron RX	IBM Ethernet Switch B08R	4003-R16	Brocade BigIron RX-16
IBM x-series Ethernet Switches	Brocade TurboIron	IBM Ethernet Switch B24X	4002-X2A 4002X2A	Brocade TurboIron 24X
IBM c-series Ethernet Switches	Brocade NetIron CES	IBM Ethernet Switch B24C (C)	4002-C2A 4002C2A	Brocade NetIron CES 2024C
IBM c-series Ethernet Switches	Brocade NetIron CES	IBM Ethernet Switch B24C (F)	4002-C2B 4002C2B	Brocade NetIron CES 2024F

IBM family name	Brocade family name	IBM product name	IBM machine type and model	Brocade name
IBM c-series Ethernet Switches	Brocade NetIron CES	IBM Ethernet Switch B48C (C)	4002-C4A 4002C4A	Brocade NetIron CES 2048C
IBM c-series Ethernet Switches	Brocade NetIron CES	IBM Ethernet Switch B48C (F)	4002-C4B 4002C4B	Brocade NetIron CES 2048F
IBM c-series Ethernet Switches	Brocade NetIron CES	IBM Ethernet Switch B50C (C)	4002-C5A 4002C5A	Brocade NetIron CES 2048CX
IBM c-series Ethernet Switches	Brocade NetIron CES	IBM Ethernet Switch B50C (F)	4002-C5B 4002C5B	Brocade NetIron CES 2048FX
IBM s-series Ethernet Switches	Brocade FastIron SX	IBM Ethernet Switch B08S	4003-S08	Brocade FastIron SX 800
IBM s-series Ethernet Switches	Brocade FastIron SX	IBM Ethernet Switch B16S	4003-S16	Brocade FastIron SX 1600
IBM g-series Ethernet Switches	Brocade FastIron GS/GS-STK	IBM Ethernet Switch B48G	4002-G4A 4002-G4A	Brocade FastIron GS 648P
IBM g-series Ethernet Switches	Brocade FastIron GS/GS-STK	IBM Ethernet Switch B50G	4002-G5A 4002-G5A	Brocade FastIron GS 648P-STK

1.2 Product description

In this section we provide descriptions of various products and their components.

1.2.1 IBM m-series Ethernet/IP Routers

IBM m-series Ethernet/IP Routers are designed to enable reliable converged infrastructures and support mission-critical applications. The m-series features an advanced N+1 redundant switch fabric architecture designed for very high availability, even in the case of a switch fabric card failure. The redundant fabric architecture is complemented by comprehensive hardware redundancy for the management modules, power supplies, and cooling system.

With their superior 1 GbE and 10 GbE port densities, m-series switching routers are well suited for large-scale high performance cluster computing. By combining superior data capacity with ultra-low latency, m-series switching routers can accelerate application performance in high performance computing clusters, thereby increasing processing power and productivity.

Following are the m-series key features:

- High performance IPv4/IPv6/MPLS/L2/VRF-enabled routers
- State-of-the-art Clos fabric offering up to 7.68 Tbps raw fabric capacity, 3.2 Tbps data capacity, and 100 Gbps FDX per full slot
- Exceptional density for 10/100/1000 Copper, 100/1000 Fiber, 10 GbE, and POS interfaces
- Up to 2 million BGP routes and 512,000 IPv4 routes in hardware (scalability limits dependent on configured system parameters, system profile selected, and routing database complexity)
- Wire-speed edge (PE), core (P) Label Switching Router functions support IP over MPLS, Virtual Leased Lines (VLLs), Virtual Private LAN Services (VPLS), and BGP/MPLS VPNs
- Comprehensive MPLS traffic engineering based on either OSPF-TE or IS-IS/TE
- MPLS Fast Reroute and hot standby paths for highly resilient service delivery
- Advanced Layer 2 metro switching including Super Aggregated VLANs (Q-in-Q) and rapid converging MRP, VSRP, RSTP, and MSTP protocols
- Carrier Ethernet Service Delivery (MEF9 and MEF14 certified) with support for up to 1 million MAC addresses

- Comprehensive hardware redundancy with hitless management failover and hitless software upgrades for Layer 2/Layer 3 with BGP and OSPF graceful restart

The m-series Ethernet routers are available in the following model configurations:

- IBM Ethernet/IP Router B04M (4003-M04) - 4-slot switching router, 400 Gbps data capacity, and up to 16 10GbE and 128 1GbE ports per system
- IBM Ethernet/IP Router B08M (4003-M08) - 8-slot switching router, 800 Gbps data capacity, and up to 32 10GbE and 384 1GbE ports per system
- IBM Ethernet/IP Router B16M (4003-M16) - 16-slot switching router, 1.6 Tbps data capacity, and up to 64 10GbE and 768 1GbE ports per system
- IBM Ethernet/IP Router B32M (4003-M32) - 32-slot switching router, 3.2 Tbps data capacity, and up to 128 10GbE and 1,536 1GbE ports per system.

Al four models are shown in Figure 1-2.

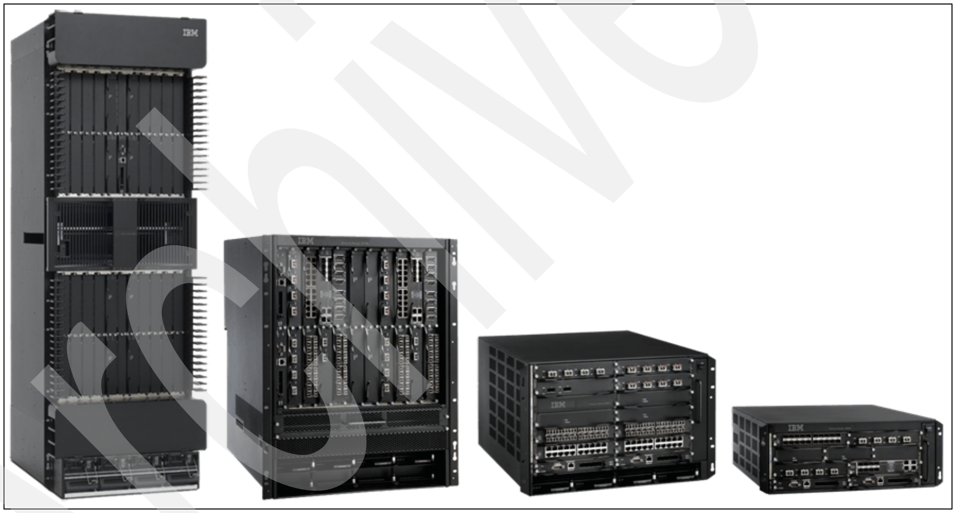


Figure 1-2 IBM m-series Ethernet/IP Routers

All m-series models can only be installed in the rack. Non-rack installation is not supported.

Operating system

All m-series systems run Brocade Multi-Service IronWare R4.0.00 or higher operating system.

Exceptional density

The m-series is scalable to one of the industry leading densities for the occupied space of 128 10 Gigabit Ethernet ports or 1,536 Gigabit Ethernet ports in a single chassis. The m-series also supports up to or 64 OC-192 or 256 OC-12/48 ports in a single chassis.

Scalable Clos fabric architecture

The m-series Ethernet/IP Routers are using a Clos fabric architecture that provides a high level of scalability, redundancy and performance. As shown in Figure 1-3, there are multiple switch fabric modules (SFMs) in the system. A switch fabric module has multiple fabric elements, each of which has multiple connections to every interface slot.

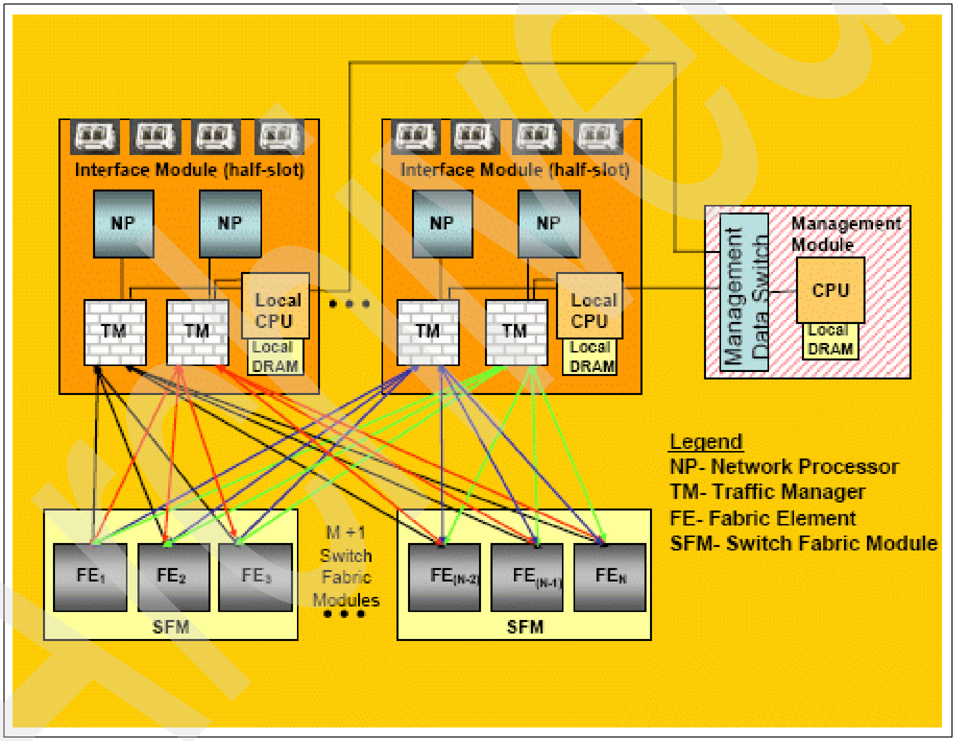


Figure 1-3 m-series Clos architecture

Note: The Clos architecture is named after the ground-breaking work by the researcher Charles Clos. The Clos architecture has been the subject of much research over several years. A multi-stage Clos architecture has been mathematically proven to be non-blocking. The resiliency of this architecture makes it the ideal building block in the design of high availability, high performance systems.

The Clos architecture uses data striping technology to ensure optimal utilization of fabric interconnects. This mechanism always distributes the load equally across all available links between the input and output interface modules. By using fixed-size cells to transport packets across the switch fabric, the m-series switching architecture ensures predictable performance with very low and deterministic latency and jitter for any packet size. The presence of multiple switching paths between the input and output interface modules also provides an additional level of redundancy.

Here are several advantages of a Clos architecture over a traditional architecture:

- Common architecture across the product family, because the same fabric elements are used on all chassis of the m-series product range. This demonstrates the superior scalability of the architecture from a small 4-slot system to a large 16-slot system.
- No head-of-line blocking at any point irrespective of traffic pattern, packet size, or type of traffic.
- Optimal utilization of switch fabric resources at all times. The data striping capability ensures that there is fair utilization of the switch fabric elements at all times without overloading of any single switch fabric element.
- “Intra-SFM” redundancy: An SFM can withstand the failure of some of the fabric elements and yet continue to operate with the remaining fabric elements. This unique capability provides a very high level of redundancy even within an SFM.
- Exceptional high availability: The m-series SFMs have (N+1) redundancy allowing m-series models to gracefully adapt to the failure of multiple switch fabric elements. Moreover, because there are multiple fabric elements within an SFM, the failure of a fabric element does not bring down the entire SFM.

High availability

Both the hardware and software architecture of the m-series are designed to ensure very high Mean Time Between Failures (MTBF) and low Mean Time To Repair (MTTR). Cable management and module insertion on the same side of the chassis allows ease of serviceability when a failed module needs to be replaced or a new module needs to be inserted.

The ability to handle the failure of not only an SFM, but also elements within an SFM, ensures a robust, redundant system ideal for non-stop operation. The overall system redundancy is further bolstered by redundancy in other active system components such as power supplies, fans, and management modules. The passive backplane on the m-series chassis increases the reliability of the system.

The modular architecture of the Multi-Service IronWare operating system has several distinguishing characteristics that differentiate it from legacy operating systems that run on routers:

- Industry-leading cold restart time of less than a minute
- Support for hitless software upgrade
- Hitless Layer 2 and Layer 3 failovers
- Sub-second switchover to the standby management module if a communication failure occurs between active and standby management modules

Distributed queuing for fine-grained Quality of Service (QoS)

A unique characteristic of the m-series is the use of a distributed queuing scheme that maximizes the utilization of buffers across the whole system during congestion. This scheme marries the benefits of input-side buffering (Virtual Output Queuing) with those of an output port driven scheduling mechanism. Input queuing using virtual output queues ensures that bursty traffic from one port does not hog too many buffers on an output port. An output-port driven scheduling scheme ensures that packets are sent to the output port only when the port is ready to transmit a packet.

Each interface module maintains multiple, distinct priority queues to every output port on the system. Packets are “pulled” by the outbound interface module when the output port is ready to send a packet. Switch fabric messaging is used to ensure that there is tight coupling between the two stages. This closed loop feedback between the input and output stages ensures that no information is lost between the two stages. The use of such “virtual output queues” maximizes the efficiency of the system by storing packets on the input module until the output port is ready to transmit the packet. In all, there are 512k virtual output queues on the m-series chassis.

Congestion avoidance is handled by applying Weighted Random Early Discard (WRED) or taildrop policy. On the output ports, a variety of scheduling mechanisms such as strict priority, weighted fair queuing, or a combination of these approaches can be applied to deliver tiered QoS guarantees for several applications.

The QoS subsystem on the m-series has extensive classification and packet marking capabilities that can be configured:

- Prioritization based on Layer 2 (802.1p), TOS, DSCP, or MPLS EXP bit of an input packet
- Mapping of packet/frame priority from ingress encapsulation to Egress encapsulation
- Remarking of a packet's priority based on the result of the 2-rate, 3-color policer

Traffic policers and ACLs

All interface modules support a large number of both inbound and outbound traffic policers in hardware. Up to 512k traffic policers can be concurrently configured in the system. The 2-rate, 3-color policers meter subscriber flows by classifying them into compliant (CIR) rates or excess (EIR) rates. This capability is especially useful when mixing traffic flows with different characteristics on the same port.

For security purposes, both input ACLs (Access Control Lists) and output ACLs are supported by the system on every interface module. Up to 114,688 input ACL entries and 131,072 output ACL entries for ACL rules can be applied to local interfaces on every interface module.

Denial of Service (DoS) guards

Layer 2 services such as VPLS require support for efficient replication of packets to the entire broadcast domain. For example, traditional architectures handle Ethernet frames with unknown MAC address by sending them to a processor to replicate the packet to the broadcast domain. The involvement of the CPU makes the system vulnerable to a potential denial of service attack. In contrast, the m-series handles this scenario very efficiently by performing the flooding in hardware.

The m-series has a dedicated out-of-band management link between each interface module and the management module to isolate control traffic from data traffic. Multiple queues to the management module allow different types of control traffic to be prioritized. These capabilities, together with secure management and ACLs, are immensely useful in protecting the system from potential DoS attacks in the network.

Spatial multicast support

The m-series architecture has native support for spatial multicast, a critical requirement for offering video services in a network. The input interface module sends one copy of an incoming multicast packet to the switch fabric. The switch fabric then replicates the packet within itself to multiple output interface modules in the system, which in turn replicate the multicast packet to the destination ports.

Industry-leading multi-service feature set

In contrast to some systems that limit the capabilities that can be concurrently enabled, the m-series architecture allows both Layer 2 and Layer 3 services to be offered on the same device and the same port concurrently. This ability gives unprecedented flexibility to the service provider in tailoring the system to meet end user needs.

Scalability

The m-series of routers is a highly scalable family of routers. Some examples of its industry-leading scalability include:

- Up to 4k VPLS/VLL instances and up to 256k VPLS MAC addresses
- Support for 4094 VLANs and up to 1 million MAC addresses
- 512k IPv4 routes in hardware FIB
- 112k IPv6 routes in hardware FIB
- 2 million BGP routes
- 400 BGP/MPLS VPNs and up to 256k VPN routes

Investment protection

The m-series chassis uses a half slot design for interface modules. The divider between two adjacent half slots can be removed in future to combine them into a full slot. All chassis have 100 Gbps of full-duplex bandwidth per full slot. In addition, with the ability to offer multiple services including dual-stack IPv4/IPv6 and MPLS services in hardware, the m-series offers excellent investment protection.

Physical and thermal parameters

The physical and thermal parameters are shown in Table 1-2.

Table 1-2 m-series physical and thermal parameters

Component	IBM Ethernet Router B04M	IBM Ethernet Router B08M	IBM Ethernet Router B16M	IBM Ethernet Router B32M
Chassis type	Modular 4-slot chassis	Modular 8-slot chassis	Modular 16-slot chassis	Modular 32-slot chassis

Component	IBM Ethernet Router B04M	IBM Ethernet Router B08M	IBM Ethernet Router B16M	IBM Ethernet Router B32M
H/W/D (cm)	17.68 x 44.32 x 57.15	31.01 x 44.32 x 57.15	62.15 x 44.32 x 64.77	146.58 x 44.32 x 61.21
Rack units (RUs)	4	7	14	33
Max. weight	35 kg	60 kg	107 kg	217 kg
Max. power draw	1289 W	2560 W	5191 W	10781 W
Op. temperature (°C)	0 - 40 °C	0 - 40 °C	0 - 40 °C	0 - 40 °C
Heat emission (BTU/hr)	4389	8737	17717	36787
Airflow	Side-to-side	Side-to-side	Front-to-back	Front-to-back
Fan assemblies	1	1	3	10

The cooling system of the m-series is configured as follows:

- IBM Ethernet Router B04M (4003-M04): Is equipped with a fan module containing two 4-speed fans and two fan controllers to support redundancy.
- IBM Ethernet Router B08M (4003-M08): Is equipped with a fan module containing four 4-speed fans and four fan controllers to support redundancy.
- IBM Ethernet Router B16M (4003-M16): Is equipped with three fan assemblies. The fan tray located in the lower front of the chassis contains six 4-speed fans. There are two fan assemblies located in the rear of the chassis.
- IBM Ethernet Router B32M (4003-M32): Is equipped with ten fan assemblies all located at the back of the chassis. Two smaller fan assemblies are located at the bottom to cool the power supplies while the remaining eight cool the rest of the chassis.

All the fans are hot swappable and self adjusting based on sensor readings.

Power parameters

All m-series models provide redundant and removable power supplies with AC power options. Power supplies can be exchanged between B08M and B16M models but not between the B04M or B32M. None of the m-series models provide the Power over Ethernet (PoE) option.

The power parameters are shown in Table 1-3.

Table 1-3 m-series power parameters

Component	IBM Ethernet Router B04M	IBM Ethernet Router B08M	IBM Ethernet Router B16M	IBM Ethernet Router B32M
Power supplies	AC - 100-120V, 200-240V 1200 W	AC - 100-120V, 200-240V 1200 W	AC - 100-120V, 200-240V 1200 W	AC - 200-240V DC - -60. -40V 2400 W
Power supply bays	3	4	8	8
Number of power supply bays required for fully loaded chassis	1	2	4	4

Slots, ports, memory, and performance

All m-series models provide a *Store & Forward* switching engine. Each model has two management module slots that are exchangeable between all m-series models. Fabric modules can be exchanged between B08M and B16M models but not between the B04M and B32M. All models have a passive backplane.

All m-series models have maximum 1 GB RAM and FLASH of 32 MB.

The number of slots, ports, and performance metrics are shown in Table 1-4.

Table 1-4 Slots, ports, and performance metrics

Component	IBM Ethernet Router B04M	IBM Ethernet Router B08M	IBM Ethernet Router B16M	IBM Ethernet Router B32M
Slots	9	13	22	42
Payload slots	4	8	16	32
Max. number of slots for fabric modules	3	3	4	8
Min. number of switch fabric modules required for fully-loaded chassis at line-rate	2	2	3	8
10/100/1000 copper ports per module (MRJ21)	48	48	48	48

Component	IBM Ethernet Router B04M	IBM Ethernet Router B08M	IBM Ethernet Router B16M	IBM Ethernet Router B32M
10/100/1000 max. copper ports per system (MRJ21)	192	384	768	1536
10/100/1000 copper ports per module (RJ-45)	20	20	20	20
10/100/1000 max. copper ports per system (RJ-45)	80	160	320	640
1 GbE SFP ports per module	20	20	20	20
1 GbE max. SFP ports per module	80	160	320	640
10 GbE ports per module	4	4	4	4
10 GbE max. ports per module	16	32	64	128
POS-OC12	32	64	128	256
POS-OC48	32	64	128	256
POS-OC192	8	16	32	64
Fabric switching capacity	960 Gbps	1.92 Tbps	3.84 Tbps	7.68 Tbps
Data switching capacity	400 Gbps	800 Gbps	1.6 Tbps	3.2 Tbps
L2 throughput	240 Mbps	480 Mbps	960 Mbps	~2 bpps
L3 throughput	240 Mbps	480 Mbps	960 Mbps	~2 bpps
Wirespeed forwarding rate	240 Mbps	480 Mbps	960 Mbps	~2 bpps

All slots have half-slot line module design. Slots have removable dividers to support future full slot modules.

With such design, m-series models are providing exceptional density of usable ports.

Each half-slot provides 50 Gbps full-duplex user bandwidth. With full-slot configurations, 100 Gbps full-duplex user bandwidth will be available per slot.

All modules are hot-swappable and do not require power-off to be replaced.

Interface modules

Table 1-5 shows which modules can be installed in the m-series chassis payload slots.

Table 1-5 Interface modules

Speed	Number of ports	Connector type
10/100/1000MbE	48	MRJ21
10/100/1000MbE	20	RJ45
100/1000MbE	20	SFP
10GbE	2	XFP
10GbE	4	XFP
OC-192 POS/SDH	2	SFP
OC-12/48 POS/SDH	2	SFP
OC-12/48 POS/SDH	4	SFP
OC-12/48 POS/SDH	8	SFP

Interface types

Following are the available interface types:

- 10/100/1000 Mbps Ethernet port with MRJ21 connector
- 10/100/1000 Mbps Ethernet port with RJ45 connector
- 100/1000 Mbps Ethernet port with SFP connector
- 10 Gbps Ethernet port with XFP connector
- OC-192 (STM-64) port with SFP connector
- OC-12/48 (STM-4/STM-16) port with SFP connector

Transceivers

In Table 1-6, Table 1-7, and Table 1-8, we show the available transceivers to be used in interface modules.

Table 1-6 Transceivers for 100/1000 Mbps Ethernet ports

Type	Connector	Speed	Distance
1000BASE-T SFP copper	RJ-45	1Gbps	Up to 100 m with CAT5 or higher
1000BASE-SX 850 nm SFP optics	LC	1 Gbps	Up to 550 m over multi-mode fiber
1000BASE-LX 1310 nm SFP optics	LC	1 Gbps	Up to 10 km over single-mode fiber
1000BASE-LHA 1550 nm SFP optics	LC	1 Gbps	Up to 70 km over single-mode fiber
100BASE-FX 1310 nm SFP optics	LC	100 Mbps	Up to 2 km over multi-mode fiber

Table 1-7 Transceivers for 10 Gbps Ethernet ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm XFP optics	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm XFP optics	LC	10 Gbps	Up to 10 km over single-mode fiber
10GBASE-ER 1550 nm XFP optics	LC	10 Gbps	Up to 40 km over single-mode fiber
10GBASE-CX4 XFP copper	LC	10 Gbps	Up to 15 m over CX4 grade copper

Table 1-8 Transceivers for the OC-12/48 (STM-4/STM-16)

Type	Connector	Speed	Distance
POS OC-12 (STM-4) SFP optics	LC	OC-12 (STM-4)	Up to 500 m over multi-mode fiber
POS-12 (STM-4) SR-1/IR-1 SFP optics	LC	OC-12 (STM-4)	Up to 15 km over single-mode fiber
POS-12 (STM-4) LR-1 SFP optics	LC	OC-12 (STM-4)	Up to 40 km over single-mode fiber
POS-12 (STM-4) LR-2 SFP optics	LC	OC-12 (STM-4)	Up to 80 km over single-mode fiber
POS-48 (STM-16) SR-1 SFP optics	LC	OC-48 (STM-16)	Up to 2 km over single-mode fiber
POS-48 (STM-16) IR-1 SFP optics	LC	OC-48 (STM-16)	Up to 15 km over single-mode fiber
POS-48 (STM-16) LR-1 SFP optics	LC	OC-48 (STM-16)	Up to 40 km over single-mode fiber
POS-48 (STM-16) LR-2 SFP optics	LC	OC-48 (STM-16)	Up to 80 km over single-mode fiber

Table 1-9 Transceivers for OC-192 (STM-64)

Type	Connector	Speed	Distance
POS-192 (STM-64) SR-1 SFP optic	LC	OC-192 (STM-64)	Up to 2 km over single-mode fiber
POS-192 (STM-64) IR-2 optic	LC	OC-192 (STM-64)	Up to 40 km over single-mode fiber
POS-192 (STM-64) LR-2 optic	LC	OC-192 (STM-64)	Up to 80 km over single-mode fiber

Cables for MRJ21 must be ordered separately. One distributor of such cables is Tyco Electronics:

http://www.ampnetconnect.com/brocade/

Services, protocols, and standards

IBM m-series Ethernet routers support these services, protocols, and standards.

The suite of Layer 2 Metro Ethernet technologies with advanced services is based on the following features:

- IEEE 802.1Q
- Rapid Spanning Tree Protocol (RSTP)
- Metro Ring Protocol (MRP)
- Virtual Switch Redundancy Protocol (VSRP)
- Up to 1 million MAC addresses per system

MPLS complements Layer 2 Metro Ethernet capabilities with these features:

- MPLS-TE
- Fast Reroute (FRR)
- MPLS Virtual Leased Line (VLL)
- Virtual Private LAN Service (VPLS)
- Border Gateway Protocol (BGP)
- MPLS VPNs
- MPLS L3VPNs

In addition to a rich set of Layer 2 and MPLS based capabilities, routers provide creation of scalable resilient services with the Metro Ethernet Forum (MEF) specifications for these features:

- Ethernet Private Line (EPL)
- Ethernet Virtual Private Line (EVPL)
- Ethernet LAN (E-LAN)

For Internet edge and aggregation routing, these capabilities are provided:

- Dual stack IPv4/IPv6 wire-speed routing performance
- Up to 512000 IPv4 routes in the hardware Forwarding Information Base (FIB)
- Up to 1 million BGP routes in the BGP Routing Information Base (RIB)
- RIPv1/v2
- RIPng
- OSPFv2/v3
- IS-IS
- IS-IS for IPv6
- Foundry Direct Routing (FDR)
- BGP-4
- BGP-4+

To enable use in advanced converged enterprise backbones, providing reliable transport of Voice over IP (VoIP), video services and mission-critical data, the following capabilities are utilized:

- Quality of Service (QoS)

- Wire-speed unicast/multicast routing for IPv4 and IPv6
- IPv4 multicasting routing protocols:
  - IGMPv1/v2/v3
  - PIM-DM/-SM/-SSM
  - MSDP
- IPv6 multicasting routing protocols:
  - MLDv1/v2
  - PIM-SM/-SSM
  - Anycast RP

Provided IPv4 and IPv6 multicast protocols help to make the most efficient use of the network bandwidth.

Multi-VRF virtual routing allows enterprises to create multiple security zones and simplified VPNs for different applications and business units while streamlining overall network management.

The routers also provide intrinsic wire-speed sFlow scalable network-wide monitoring of flows for enhanced security management by malicious traffic detection and intrusion detection as well as for proactive management of network bandwidth through traffic trend analysis and capacity upgrade planning.

For the whole list of supported standard and RFC compliance, see the website:

`http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/`

We describe some of the service implementations in more detail in the following sections.

Link aggregation

Following are the main characteristics of link aggregation implementation on m-series models:

- 802.3ad/LCAP support
- 256 servers trunks supported
- Up to 32 ports per trunk group
- Cross module trunking
- Ports in the group do not need to be physically consecutive
- Tagged ports support in trunk group
- Compatibility with Cisco EtherChannel
- Ports can be dynamically added or deleted from the group, except for the primary port

Layer 2 switching

Following are the main characteristics of Layer 2 switching implementation on m-series models:

- Up to 1 million MAC addresses per system
- Up to 288000 MAC entries per network processor
- 9216 byte jumbo frames
- L2 MAC filtering
- MAC authentication
- MAC port security
- 4090 VLANs
- Port and protocol based VLANs
- VLAN tagging:
  - 802.1q
  - Dual Mode
  - SAV/Q-in-Q
- STP (Spanning Tree Protocol) per VLAN
- Compatibility with CISCO PVST (Per VLAN Spanning Tree)
- STP fast forwarding (fast port, fast uplink), root guard, Bridge Protocol Data Unit (BPDU) guard
- Up to 128 Spanning-Tree instances
- Rapid STP (802.1w compatible)
- MSTP (Multiple Spanning Tree Protocol) (802.1s)
- MRP Phase I&II
- Q-in-Q/SAV support with unique tag-type per port
- VSRP (Virtual Switch Redundancy Protocol)
- Up to 255 topology groups
- Hitless OS with 802.1ag (Connectivity Fault Management) and UDLD (Uni-directional Link Detection)

Multicast

Following are the main characteristics of multicast implementation on m-series models:

- IGMP/IGMPv3 (Internet Group Management Protocol)
- IGMP/IGMPv3 snooping
- IGMP proxy
- PIM (Protocol-Independent Multicast) proxy/snooping
- Multicast routing PIM/DVMRP (Distance Vector Multicast Routing Protocol)
- Up to 153600 multicast routes
- IPv4 PIM modes - Sparse, Dense, Source-Specific
- IPv6 PIM modes - Sparse, Source-Specific
- Up to 4096 IPv4/IPv6 multicast cache entries

- Multi-VRF (Virtual Routing and Forwarding) multicast - PIM SM, SSM, DM, IGMP, Mroute, Anycast RP, MSDP
- Anycast RP for IPv4/IPv6

1.2.2 IBM r-series Ethernet Switches

The IBM r-series Ethernet Switches are a modular switch series designed to provide improved service with industry leading levels of performance; reduced cost with a high density and energy efficient design; and a commonly used management interface which can leverage existing staff training and experience to help manage risk when integrating new technology.

Doing this allows network designers to standardize on a single product family for end-of-row, aggregation, and backbone switching, and is ideal for data center and enterprise deployment. In addition, the switches, with their high-density and compact design, are an ideal solution for High-Performance Computing (HPC) environments and Internet Exchanges and Internet Service Providers (IXPs and ISPs) where non-blocking, high-density Ethernet switches are needed.

Following are the r-series key features:

- Powerful suite of unicast and multicast IPv4 and IPv6 protocol support
- Interchangeable half-height line modules reduce sparing costs, TCO, and provide cost-effective modular growth
- Highly density chassis design supports up to 256 10 GbE or 1,536 wire-speed 1 GbE ports in a single 32-slot chassis
- High availability design features redundant and hot-pluggable hardware, hitless Layer 2 software upgrades, and graceful BGP (Border Gateway Protocol) and OSPF (Open Shortest Path First) restart
- Advanced non-blocking Clos fabric features adaptive self-routing with graceful system degradation in the event of two or more module failures
- End to End Quality of Service (QoS) supported with hardware based honoring and marking and congestion management
- Scalable hardware-based IP routing to 512,000 IPv4 routes per line module
- High-capacity 80 Gbps cross-module link aggregation supports high-bandwidth inter-switch trunking
- Embedded sFlow per port supports scalable hardware-based traffic monitoring across all switch ports without impacting performance

The r-series Ethernet Switches are available in the following model configurations:

- IBM Ethernet/IP Router B04R (4003-R04) - 4-slot switching router, 400 Gbps data capacity, and up to 64 10GbE and 128 1GbE ports per system
- IBM Ethernet/IP Router B08R (4003-R08) - 8-slot switching router, 800 Gbps data capacity, and up to 128 10GbE and 384 1GbE ports per system
- IBM Ethernet/IP Router B16R (4003-R16) - 16-slot switching router, 1.6 Tbps data capacity, and up to 256 10GbE and 768 1GbE ports per system

All r-series models can only be installed in the rack. Non-rack installation is not supported.

All three models are shown in Figure 1-4.

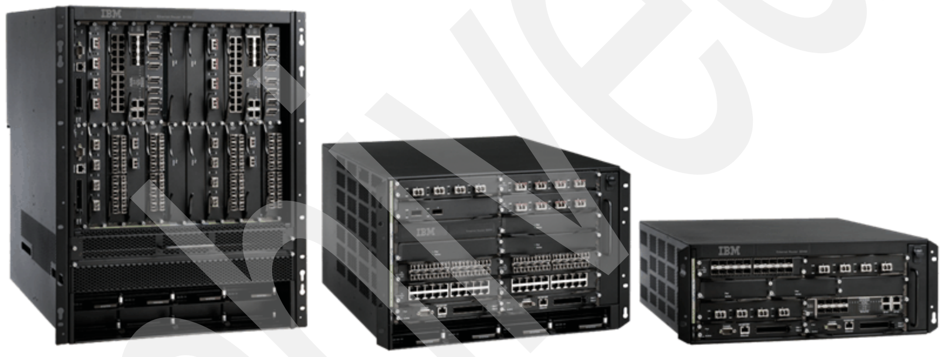


Figure 1-4 IBM r-series Ethernet Switches

Operating system

All r-series systems run Brocade Multi-Service IronWare for BigIron RX R2.7.02 or higher operating system.

Exceptional density

The r-series is scalable to one of the industry leading densities for the occupied space of 256 10 Gigabit Ethernet ports or 768 Gigabit Ethernet ports in a single chassis.

High availability

Both the hardware and software architecture of the m-series are designed to ensure very high Mean Time Between Failures (MTBF) and low Mean Time To Repair (MTTR). Cable management and module insertion on the same side of the chassis allows ease of serviceability when a failed module needs to be replaced or a new module needs to be inserted.

The ability to handle the failure of not only an SFM but also elements within an SFM ensures a robust, redundant system ideal for non-stop operation. The overall system redundancy is further bolstered by redundancy in other active system components such as power supplies, fans, and management modules. The passive backplane on the m-series chassis increases the reliability of the system.

The modular architecture of the Multi-Service IronWare operating system has several distinguishing characteristics that differentiate it from legacy operating systems that run on routers:

- Industry-leading cold restart time of less than a minute
- Support for hitless software upgrade
- Hitless Layer 2 and Layer 3 failovers
- Sub-second switchover to the standby management module if a communication failure occurs between active and standby management modules

Scalability

The m-series of routers is a highly scalable family of routers. Here are a few examples of its industry-leading scalability:

- Support for 4094 VLANs and up to 1 million MAC addresses
- 512k IPv4 routes in hardware FIB
- 65k IPv6 routes in hardware FIB
- 1 million BGP routes

Investment protection

The r-series chassis uses a half slot design for interface modules. The divider between two adjacent half slots can be removed in future to combine them into a full slot. All chassis have 100 Gbps of full-duplex bandwidth per full slot. In addition, with the ability to offer multiple services including dual-stack IPv4/IPv6 in hardware.

Physical and thermal parameters

The physical and thermal parameters are shown in Table 1-10.

Table 1-10 m-series physical and thermal parameters

Component	IBM Ethernet Switch B04R	IBM Ethernet Switch B08R	IBM Ethernet Switch B16R
Chassis type	Modular 4-slot chassis	Modular 8-slot chassis	Modular 16-slot chassis
H/W/D (cm)	17.68 x 44.32 x 57.15	31.01 x 44.32 x 57.15	62.15 x 44.32 x 64.77
Rack units (RUs)	4	7	14
Max. weight	35 kg	60 kg	107 kg
Max. power draw	1217 W	2417 W	4905 W
Op. temperature (°C)	0 - 40 °C	0 - 40 °C	0 - 40 °C
Heat emission (BTU/hr)	4155	8249	16741
Airflow	Side-to-side	Side-to-side	Front-to-back
Fan assemblies	1	1	3

The cooling system of the m-series is configured as follows:

- IBM Ethernet Switch B04R (4003-R04): Is equipped with a fan module containing two 4-speed fans and two fan controllers to support redundancy.
- IBM Ethernet Switch B04R (4003-R08): Is equipped with a fan module containing four 4-speed fans and four fan controllers to support redundancy.
- IBM Ethernet Switch B04R (4003-R16): Is equipped with three fan assemblies. The fan tray located in the lower front of the chassis contains six 4-speed fans. There are two fan assemblies located in the rear of the chassis.

All the fans are hot swappable and self adjusting based on sensor readings.

Power parameters

All r-series models provide redundant and removable power supplies with AC power options. Power supplies can be exchanged between B08R and B16R models but not between the B04R. None of the r-series models provide the Power over Ethernet (PoE) option.

The power parameters are shown in Table 1-11.

Table 1-11 r-series power parameters

Component	IBM Ethernet Switch B04R	IBM Ethernet Switch B08R	IBM Ethernet Switch B16R
Power supplies	AC - 100-120V, 200-240V 1200 W	AC - 100-120V, 200-240V 1200 W	AC - 100-120V, 200-240V 1200 W
Power supply bays	3	4	8
Number of power supply bays required for fully loaded chassis	1	2	4

Slots, ports, memory, and performance

All r-series models provide a *Store & Forward* switching engine. Each model has two management module slots that are exchangeable between all m-series models. Fabric modules can be exchanged between B08R and B16R models but not between the B04R. All models have passive backplane.

The number of slots, ports, and performance metrics are shown in Table 1-12.

Table 1-12 Slots, ports, and performance metrics

Component	IBM Ethernet Switch B04R	IBM Ethernet Switch B08R	IBM Ethernet Switch B16R
Slots	9	13	22
Payload slots	4	8	16
Max. number of slots for fabric modules	3	3	4
Min. number of switch fabric modules required for fully-loaded chassis at line-rate	2	2	3
10/100/1000 copper ports per module (MRJ21)	48	48	48

Component	IBM Ethernet Switch B04R	IBM Ethernet Switch B08R	IBM Ethernet Switch B16R
10/100/1000 max. copper ports per system (MRJ21)	192	384	768
10/100/1000 copper ports per module (RJ-45)	24	24	24
10/100/1000 max. copper ports per system (RJ-45)	96	192	384
1 GbE SFP ports per module	24	24	24
1 GbE Max. SFP ports per module	96	192	384
10 GbE ports per module (SFP+)	16	16	16
10 GbE max. ports per module (SFP+)	16	128	256
10 GbE ports per module (XFP)	4	4	4
10 GbE max. ports per module (XFP)	16	32	64
Fabric switching capacity	960 Gbps	1.92 Tbps	3.84 Tbps
Data switching capacity	400 Gbps	800 Gbps	1.6 Tbps
Wirespeed forwarding rate	286 Mbps	571 Mbps	1142 Mbps

All slots have a half-slot line module design, and the slots have removable dividers to support future full slot modules.

With such a design, m-series models provide exceptional density of usable ports.

Each half-slot provides 50 Gbps full-duplex user bandwidth. With full-slot configurations, 100 Gbps full-duplex user bandwidth will be available per slot.

All modules are hot-swappable and do not require a power-off to be replaced.

Interface modules

Table 1-13 shows which modules can be installed in the r-series chassis payload slots.

Table 1-13 Interface modules

Speed	Number of ports	Connector type
10/100/1000MbE	48	MRJ21
10/100/1000MbE	24	RJ45
100/1000MbE	24	SFP
10GbE	16	SFP+
10GbE	4	XFP

Interface types

Following are the available interface types:

- 10/100/1000 Mbps Ethernet port with MRJ21 connector
- 10/100/1000 Mbps Ethernet port with RJ45 connector
- 100/1000 Mbps Ethernet port with SFP connector
- 10 Gbps Ethernet port with SFP+ connector
- 10 Gbps Ethernet port with XFP connector

Transceivers

In Table 1-14, Table 1-15 and Table 1-16, we show the available transceivers to be used in interface modules.

Table 1-14 Transceivers for 100/1000 Mbps Ethernet ports

Type	Connector	Speed	Distance
1000BASE-T SFP copper	RJ-45	1 Gbps	Up to 100 m with CAT5 or higher
1000BASE-SX 850 nm SFP optics	LC	1 Gbps	Up to 550 m over multi-mode fiber
1000BASE-LX 1310 nm SFP optics	LC	1 Gbps	Up to 10 km over single-mode fiber
1000BASE-LHA 1550 nm SFP optics	LC	1 Gbps	Up to 70 km over single-mode fiber

Type	Connector	Speed	Distance
100BASE-FX 1310 nm SFP optics	LC	100 Mbps	Up to 2 km over multi-mode fiber

Table 1-15 Transceivers for 10 Gbps SFP+ Ethernet ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm SFP+ optics	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm SFP+ optics	LC	10 Gbps	Up to 10 km over single-mode fiber

Table 1-16 Transceivers for 10 Gbps XFP Ethernet ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm XFP optics	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm XFP optics	LC	10 Gbps	Up to 10 km over single-mode fiber
10GBASE-ER 1550 nm XFP optics	LC	10 Gbps	Up to 40 km over single-mode fiber
10GBASE-CX4 XFP copper	LC	10 Gbps	Up to 15 m over CX4 grade copper

Cables for the MRJ21 must be ordered separately. One of the distributors of such cables is Tyco Electronics:

http://www.ampnetconnect.com/brocade/

Services, protocols, and standards

IBM r-series Ethernet routers support various services, protocols, and standards.

The suite of Layer 2 technologies is based on the following standards:

- IEEE 802.1Q
- Rapid Spanning Tree Protocol (RSTP)
- Metro Ring Protocol (MRP)
- Virtual Switch Redundancy Protocol (VSRP)

For advanced data center aggregation or enterprise core, the following capabilities are provided:

- Dual stack IPv4/IPv6 wire-speed routing performance
- Up to 512000 IPv4 routes in the hardware Forwarding Information Base (FIB)
- Up to 1 million BGP routes in the BGP Routing Information Base (RIB)
- RIPv1/v2
- RIPng
- OSPFv2/v3
- Foundry Direct Routing (FDR)
- BGP-4
- BGP-4+

To enable use in advanced converged enterprise backbones, providing reliable transport of Voice over IP (VoIP), video services and mission-critical data, the following capabilities are utilized:

- Quality of Service (QoS)
- Wire-speed unicast/multicast routing for IPv4 and IPv6
- IPv4 multicasting routing protocols:
  - IGMPv1/v2/v3
  - PIM-DM/-SM/-SSM
  - MSDP
- IPv6 multicasting routing protocols:
  - MLDv1/v2
  - PIM-SM/-SSM
  - Anycast RP

The IPv4 and IPv6 multicast protocols help to make the most efficient use of the network bandwidth.

The switches also provide intrinsic wire-speed sFlow scalable network-wide monitoring of flows for enhanced security management by malicious traffic detection and intrusion detection as well as for proactive management of network bandwidth through traffic trend analysis and capacity upgrade planning.

For the whole list of supported standard and RFC compliance, see the website:

http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/

In the following sections, we describe various service implementations in more detail.

Link aggregation

Following are the main characteristics of link aggregation implementation on m-series models:

- 802.3ad/LCAP support
- 31 servers trunks supported
- Up to 8 ports per trunk group
- Cross module trunking
- Ports in the group do not need to be physically consecutive
- Tagged ports support in trunk group
- Compatibility with Cisco EtherChannel
- Ports can be dynamically added or deleted from the group, except primary port

Layer 2 switching

Following are the main characteristics of Layer 2 switching implementation on r-series models:

- 9216 byte jumbo frames
- L2 MAC filtering
- MAC authentication
- MAC port security
- 4090 VLANs
- Port and protocol based VLANs
- VLAN tagging:
  - 802.1q
  - Dual Mode
- STP (Spanning Tree Protocol) per VLAN
- Compatibility with CISCO PVST (Per VLAN Spanning Tree)
- STP fast forwarding (fast port, fast uplink), root guard, BPDU (Bridge Protocol Data Unit) guard
- Up to 128 Spanning-Tree instances
- Rapid STP (802.1w compatible)
- MSTP (Multiple Spanning Tree Protocol) (802.1s)
- MRP Phase I&II
- VSRP (Virtual Switch Redundancy Protocol)
- Up to 255 topology groups
- Hitless OS with UDLD (Uni-directional Link Detection)

Multicast

Following are the main characteristics of multicast implementation on m-series models:

- IGMP/IGMPv3 (Internet Group Management Protocol)

- IGMP/IGMPv3 snooping
- IGMP proxy
- PIM (Protocol-Independent Multicast) proxy/snooping
- Multicast routing PIM/DVMRP (Distance Vector Multicast Routing Protocol)
- Up to 153600 multicast routes
- IPv4 PIM modes - Sparse, Dense, Source-Specific
- IPv6 PIM modes - Sparse, Source-Specific
- Up to 4096 IPv4/IPv6 multicast cache entries
- Anycast RP for IPv4/IPv6

1.2.3 IBM x-series Ethernet Switches

The IBM x-series Ethernet Switch is a compact, high-performance, high-availability, and high-density 1 RU switch specifically designed for mission-critical data centers and High-Performance Computer (HPC) requirements. This switch provides twenty-four 10/1 GbE (SFP+) ports plus four 10/100/1000 MbE (RJ45) ports of connectivity in an ultra-low-latency, cut-through, non-blocking architecture.

This switch is an ideal cost-effective solution for server or compute-node connectivity. It can support 1 GbE servers until they are upgraded to 10 GbE capable Network Interface Cards (NICs), simplifying migration to 10 GbE server farms. In addition, the switch can be positioned as a 10 GbE aggregation switch behind 1 GbE access switches.

In any deployment scenario, this switch is designed to save valuable rack space, power, and cooling in the data center while delivering 24x7 service through its high-availability design.

Following are the key features:

- Flexibility to mix 10 GbE and 1 GbE servers, protecting investments and streamlining migration to 10 GbE-capable server farms
- Wire-speed performance with an ultra-low-latency, cut-through, non-blocking architecture that is ideal for HPC environments
- Highly efficient power and cooling with front-to-back airflow, automatic fan speed adjustment, and use of SFP+ and direct attached SFP+ copper (Twinax) for maximum flexibility
- High availability with redundant, load-sharing, hot-swappable, auto-sensing/switching power supplies and a resilient triple-fan assembly
- End-to-end Quality of Service (QoS) with hardware-based marking, queuing, and congestion management

- Embedded per-port sFlow capabilities to support scalable hardware-based traffic monitoring

These switches are available in the following model configurations:

- IBM Ethernet Switch B24X (4002-X2A;4002AX2) - 24-port 10/1 GbE dual-speed ports plus 4 10/100/1000 MbE RJ45 ports

IBM x-series can be installed in an EIA-310D compliant rack.

The IBM Ethernet Switch B24X is shown in Figure 1-5.

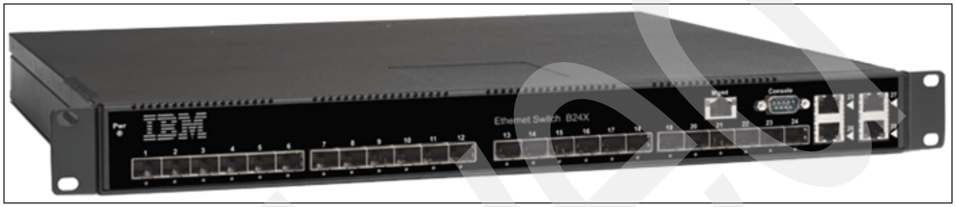


Figure 1-5 IBM x-series Ethernet Switch

Operating system

IBM x-series supports running Brocade IronWare OS R04.1.00 or higher.

Flexible data center deployment and future-proofing

Each dual-speed port on the IBM Ethernet B24X can function as a 1 GbE port by plugging in a 1 GbE SFP transceiver, making it a flexible solution for environments where some servers have not yet been upgraded to 10 GbE capable NICs. In data center environments where many servers still utilize 1 GbE links, organizations can initially deploy the B24X switch as a compact and cost-effective 10 GbE aggregation switch and later move it to the access layer as servers are upgraded to 10 GbE.

When organizations upgrade a server's NICs to 10 GbE, they will only need to replace the 1 GbE SFPs with 10 GbE SFP+ transceivers or direct attached 10 GbE SFP+ copper (Twinax) transceivers. This approach protects Ethernet-based investments and streamlines migration to 10 GbE. The switch also includes four 10/100/1000 MbE RJ45 ports for additional server connectivity or separate management network connectivity.

The high density of dual-speed ports in a 1U space enables organizations to design highly flexible and cost-effective networks. In addition, organizations can utilize various combinations of short-range and long-range transceivers for a variety of connectivity options.

High availability hardware features

The IBM Ethernet Switch B24X supports two 300 W AC power supplies for 1+1 redundancy. The AC power supplies are hot-swappable and load-sharing with auto-sensing and auto-switching capabilities which are critical components of any highly available system. In addition, efficient front-to-back cooling is achieved with a hot-swappable, resilient triple-fan assembly that automatically and intelligently adjusts the speed of the fan. A single fan can fail without any impact on system operation for a 2+1 fan redundancy within the fan assembly.

The hot-swappable power supplies and fan assembly are designed to enable organizations to replace components without service disruption. In addition, several high-availability and fault-detection features are designed to help in failover of critical data flows, enhancing overall system availability and reliability. Organizations can use sFlow-based network monitoring and trending to proactively monitor risk areas and optimize network resources to avoid many network issues altogether.

Physical and thermal parameters

The physical and thermal parameters are shown in Table 1-17.

Table 1-17 x-series physical and thermal parameters

Component	IBM Ethernet Switch B24X
Chassis type	Fixed form factor
H/W/D (cm)	4.28 x 43.5x 39.378
Rack units (RUs)	1
Max. weight	7.4 kg
Max. power draw	176 W
Op. temperature (°C)	0 - 40 °C
Heat emission (BTU/hr)	600
Airflow	Front/side-to-back
Number of fans	1 fan tray (3 fans in fan tray)

All the fans are hot-swappable and have fixed speeds.

Power parameters

All x-series models provide redundant and removable power supplies with the AC power option.

The x-series models do not support Power over Ethernet (PoE).

The power supplies are auto-sensing and auto-switching, and provide up to 300 watts of total output power. The power supplies are hot swappable and can be removed and replaced without powering down the system.

The system power parameters are shown in Table 1-18.

Table 1-18 x-series power parameters

Component	IBM Ethernet Switch B24X
Power supplies	100 - 240 VAC / 50 - 60 Hz
Number of power supply bays	2
Number of power supply bays required for fully loaded chassis	1

Ports, memory, and performance

The x-series models provide a *Cut & Forward* switching engine and 512 MB of memory.

The number of ports and performance metrics are shown in Table 1-19.

Table 1-19 Ports and performance metrics

Component	IBM Ethernet Switch B48G
10 GbE ports per system SFP+	24
10/100/1000 Mbps RJ45 ports per system	4
Data switching capacity	488 Gbps
Packet routing capacity	363 Mpps

Interface types

Following are the available interface types:

- 10 Gbps Ethernet port with SFP+ connector
- 10/100/1000 Mbps Ethernet port with RJ45 connector

Transceivers

Table 1-20 shows the available transceivers that can be used.

Table 1-20 Transceivers for 10/1 Gbps Ethernet SFP+ ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm optic	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm optic	LC	10 Gbps	Up to 10 km over single-mode fiber
10GBASE active TwinAx 1m	LC	10 Gbps	Up to 1 m
10GBASE active TwinAx 3m	N/A - SFP+ transceiver on either end	10 Gbps	Up to 3 m
10GBASE active TwinAx 5m	N/A - SFP+ transceiver on either end	10 Gbps	Up to 5 m
1000BASE-SX 850 nm optic	N/A - SFP+ transceiver on either end	1 Gbps	Up to 550 m over multi-mode fiber
1000BASE-LX 1310 nm optic	LC	1 Gbps	Up to 10 km over single-mode fiber
1000BASE-LHA 1550 nm optic	LC	1 Gbps	Up to 70 km over single-mode fiber

Services, protocols, and standards

IBM x-series Ethernet Switches support these services, protocols, and standards.

The following Layer 2 protocols are supported:

- Protected link groups
- Link aggregation (IEEE 802.3ad, LACP)
- UDLD
- STP/RSTP/MSTP
- Root guard
- BPDU guard
- Up to 32,000 MAC addresses
- Up to 4096 VLANs
- Up to 512 STP groups

- Up to 8 ports per trunk, up to 28 trunk groups

Quality of Service:

- MAC address mapping to priority queue
- ACL mapping to priority queue
- ACL mapping to ToS/DSCP
- Honoring DSCP and 802.1p
- ACL mapping and marking of ToS/DSCP
- DHCP assist
- QoS queue management using weighted round robin (WRR), strict priority (SP), and a combination of WRR and SP

Traffic management:

- Inbound rate limiting per port
- ACL-based inbound rate limiting and traffic policies
- Outbound rate limiting per port and per queue
- Broadcast, multicast, and unknown unicast

The whole list of supported standards and RFC compliance can be found at:

http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/

1.2.4 IBM c-series Ethernet Switches

Network planners today must expand and extend the range of services offered further into the edge of the network. This requires extending the intelligence and high-touch processing capabilities to the network edge, whether in a metro network, a campus network, or in a data center. The challenge at the network edge is compounded by the need to flexibly define and easily manage customer services in an intuitive manner. Further, the expanding role of the converged network makes Quality of Service (QoS), resiliency, and security crucial to the success of many rollouts.

Whether deployed from a central or remote location, availability of space often determines the feasibility of deploying new equipment and services within any environment. To meet these challenges, IBM c-series Ethernet Switches are purpose-built to offer flexible, resilient, secure and advanced services in a compact form factor.

IBM c-series Ethernet Switches are compact 1 RU, multi-service edge/aggregation switches with a powerful set of capabilities that combine performance with rich functionality at the network edge. These switches offer network planners a broad set of high-performance IPv4 Full Layer 2 and Base Layer 3 functionality with flexible software upgrade options in the same device.

Following are key features of the c-series models:

- Compact 1 RU Layer 3 switch that is purpose-built for advanced Ethernet applications
- MEF 9 and MEF 14 certified
- Comprehensive OAM capabilities based on IEEE 802.1ag-2007 and MEF Service OAM Framework provide rapid troubleshooting of Layer 2 networks and Ethernet services
- Innovative Ethernet Service Instance (ESI) framework provides complete flexibility in separating, combining and managing customer service instances
- MEF standard E-LINE, E-LAN and E-TREE services
- Full IPv4 unicast and multicast capabilities
- Powered by field-proven Multi-Service IronWare OS that also runs on m-series models
- Available in a variety of configurations (48-port) in both Hybrid Fiber (HF) and RJ45 versions
- Wire-speed, non-blocking performance in all configurations
- Flexible 100M/1G SFP support and 10G XFP support with advanced optical monitoring
- Provider Backbone Bridge (IEEE 802.1ah) and Provider Bridge (IEEE 802.1ad) support

The c-series Ethernet Switches are available in these model configurations:

- IBM Ethernet Switch B24C (4002AC2) - 24 x 10/100/1000 MbE RJ-45 ports including 4 x 100/1000 MbE SFP combination ports with optional 2-port 10 GbE XFP upgrade slot
- IBM Ethernet Switch B24C (4002BC2) - 24x 100/1000 MbE SFP ports including 4 x 10/100/1000 MbE RJ-45 combination ports with optional 2-port 10 GbE XFP upgrade slot
- IBM Ethernet Switch B48C (4002-C4A; 4002AC4) - 48 x 10/100/1000 Mbps Ethernet RJ45 ports including 4 x 100/1000 MbE combination SFP ports
- IBM Ethernet Switch B48C (4002-C4B;4002BC4) - 48 x 100/1000 Mbps Ethernet hybrid fiber SFP ports
- IBM Ethernet Switch B50C (4002-C5A;4002AC5) - 48 x 10/100/1000 Mbps Ethernet RJ45 ports and 2 x 10 Gbps Ethernet XFP uplink ports
- IBM Ethernet Switch B50C (4002-C5B; 4002BC5) - 48 x 100/1000 Mbps Ethernet hybrid fiber SFP ports and 2 x 10 Gbps Ethernet XFP uplink ports

All c-series models support rack and non-rack installation.

All six models are shown in Figure 1-6.

Figure 1-6 IBM c-series Ethernet Switches

Operating system

All c-series systems run Brocade Multi-Service IronWare R3.8.00 or a higher operating system.

Carrier-class resiliency with Multi-Service IronWare

The c-series is built on Multi-Service IronWare, the same operating system software that powers widely deployed Brocade equivalents of m-series of routers, thereby allowing ease of integration with existing networks. These capabilities include support for robust routing protocols, advanced Layer 2 protocols, industry-standard user interface, a broad range of OAM protocols, security and simplified management capabilities. Multi-Service IronWare on c-series includes all these capabilities and additionally supports Provider Bridge and Provider Backbone Bridge functionality.

Enabling true Carrier Grade Ethernet services

Carrier Grade Ethernet, or Carrier Ethernet for short, is a ubiquitous, standardized service that is defined by five attributes:

- Standardized services
- Scalability
- Service management
- Reliability
- Quality of Service (QoS)

A Carrier Ethernet service can be delivered over any transport technology as long as it satisfies the standards and attributes associated with the service. Underlying transport mechanisms that can be used include native Ethernet using 802.1Q VLANs, MPLS-based Layer 2 VPNs, IEEE 802.1ad Provider Bridges, IEEE 802.1ah Provider Backbone Bridges, Ethernet over SONET, and so on.

The c-series models supports all five key attributes of Carrier Ethernet.

Standardized services

The c-series is compliant with both the MEF 9 and MEF 14 specifications. Using the c-series models, a provider can offer E-LINE, E-LAN and E-TREE services, the standardized service names for point-to-point, multipoint, and rooted multipoint services. These services can be offered using 802.1Q VLANs, Provider Bridges or Provider Backbone Bridges.

Scalability

The c-series supports up to 128k MAC addresses per system. Support for 100/1000 Mbps SFP ports or 10/100/1000 Mbps RJ45 ports, with wire-speed performance even at full load, ensures that abundant capacity is available on user facing ports to accommodate a provider's customers who want to upgrade to a higher bandwidth service. Additionally, the use of Link Aggregation Groups (LAGs) allows multiple links to be aggregated and offer even higher bandwidth services at the user network interface (UNI) to the end-user.

Service management

Recently developed specifications such as IEEE 802.1ag-2007 (Connectivity Fault Management) and MEF 17 (Service OAM Framework and Specifications) allow the rapid and proactive identification and isolation of faults in the network or service, thereby maintaining service uptime and maximizing the ability to meet customer SLAs. The c-series supports all the capabilities in IEEE 802.1ag, including Connectivity Check Messages, Loopback Message/Response and LinkTrace Message/Response. It allows flexible association and definition of both Maintenance End Points (MEP) and Maintenance Intermediate Points (MIP) within a network. Fault management functions of MEF 17 Service OAM are also supported.

Reliability

To provide a high level of reliability in the Carrier Ethernet service, the c-series supports Foundry's innovative Metro Ring Protocol (MRP/MRP-II), the ring resiliency protocol of choice on several metro networks worldwide. Standard Layer 2 protocols such as MSTP, RSTP and STP are also supported. Foundry's MRP/MRP-II allows Carrier Ethernet services to be delivered over ring-based topologies, including overlapping rings that help optimize the use of fiber in metro rings and provide fast recovery from node/link failures in milliseconds. Foundry MRP/MRP-II can also be used within a PB/PBB network.

Hard QoS

The c-series supports up to eight queues per port, each with a distinct priority level. Advanced QoS capabilities such as the use of 2-rate, 3-color traffic policers, Egress shaping, and priority remarking can also be applied to offer deterministic “hard QoS” capability to customers of the service. The c-series can be configured with Ingress and Egress bandwidth profiles per UNI that are in compliance with the rigid traffic management specifications of MEF 10/MEF 14.

Multicast support

Multicast transport is a key enabler of next-generation services like IPTV. It is also typically a major consumer of capacity in many multiservice networks. It is therefore critical for next-generation edge switches to efficiently handle multicast traffic. The c-series has comprehensive support for multicast switching and routing by a variety of protocols, including PIM-SM, PIM-DM, PIM-SSM, IGMP v2/v3 and other platform-independent multicast capabilities built in Multi-Service IronWare.

Multicast traffic within c-series is handled with a very high degree of efficiency by avoiding unnecessary replications and conserving bandwidth within the system. By performing Egress interface based replication, switch performance and buffer usage are optimally used within the system thereby maximizing network performance when running multicast traffic.

Routing capabilities

Based on Multi-Service IronWare, the operating system software that successfully powers thousands of m-series routers deployed around the world, the c-series offers routing capabilities that are commonly required in edge aggregation and other applications within a provider’s domain.

These routing capabilities include Brocade advanced hardware-based routing technology, which ensures secure and robust wire-speed routing performance. Multi-Service IronWare on the c-series includes support for IPv4 unicast protocols—RIP, OSPF, IS-IS and BGP. Further, to increase overall service availability, the c-series supports Graceful Restart helper mode for both OSPF and BGP to support hitless management failover and hitless OS upgrades on adjacent modular routers with these functions.

The powerful feature set of the c-series makes it an ideal candidate for applications beyond Carrier Ethernet service delivery. For example, data center networks and edge/aggregation routing within ISP networks often require a compact Layer 3 switch with sufficient scalability in IPv4 routes. The comprehensive support for IPv4 routing protocols, when complemented with VRRP, and VRRP-E makes the c-series ideally suited for such applications.

Physical and thermal parameters

The physical and thermal parameters are shown in Table 1-21.

Table 1-21 c-series physical and thermal parameters

Component	IBM Ethernet Switch B24C (C&F)	IBM Ethernet Switch B48C (C&F)	IBM Ethernet Switch B50C (C&F)
Chassis type	Fixed form factor	Fixed form factor	Fixed form factor
H/W/D (cm)	4.4 x 44.3 x 44.8	4.4 x 44.3 x 44.8	4.4 x 44.3 x 44.8
Rack units (RUs)	1	1	1
Max. weight	7.5 kg	7.5 kg	8 kg
Max. power draw	170 W B48C (Cooper) 195 W B48C (Fiber)	205 W B48C (Cooper) 245 W B48C (Fiber)	255 W B50C (Cooper) 295 W B50C (Fiber)
Op. temperature (°C)	0 - 40 °C	0 - 40 °C	0 - 40 °C
Heat emission (BTU/hr)	580 - B48C (Cooper) 666 - B48C (Fiber)	700 - B48C (Cooper) 836 - B48C (Fiber)	870 - B50C (Cooper) 1007 - B50C (Fiber)
Airflow	Front-to-back	Front-to-back	Front-to-back
Number of fans	6 (1 fan tray)	6 (1 fan tray)	6 (1 fan tray)

All the fans are hot swappable and self adjusting based on sensor readings.

Power parameters

All c-series models provide redundant and removable power supplies with AC power options. Power supplies can be exchanged between various c-series models. None of the c-series models provide Power over Ethernet (PoE) option.

The power parameters are shown in Table 1-22.

Table 1-22 c-series power parameters

Component	IBM c-series Ethernet Switch
Power supplies	AC - 100-120V, 200-240V
Power supply bays	2
Number of power supply bays required for fully loaded system	1

Ports, memory, and performance

All c-series models provide a *Store & Forward* switching engine.

All c-series models have maximum 512 GB RAM and FLASH of 32 MB.

The number of ports and performance metrics are shown in Table 1-23.

Table 1-23 Ports and performance metrics

Component	IBM Ethernet Switch B24C (C)	IBM Ethernet Switch B24C (F)	IBM Ethernet Switch B48C (C)	IBM Ethernet Switch B48C (F)	IBM Ethernet Switch B50C (C)	IBM Ethernet Switch B50C (F)
10/100/1000 RJ45 copper ports per system	24	N/A	48	N/A	48	N/A
100/1000 SFP ports per system	N/A	24	N/A	48	N/A	48
10G XFP uplinks	2 (optional module)	2 (optional module)	N/A	N/A	2 (built-in)	2 (built-in)
Combination ports	4 (100/1000 SFP ports)	4 (10/100/ 1000 RJ45 ports)	4 (100/1000 SFP Ports)	N/A	N/A	N/A
Forwarding performance	88 Gbps (w/2x 10 GbE)	88 Gbps (w/2x 10 GbE)	96 Gbps	96 Gbps	136 Gbps	136 Gbps
Packet forwarding performance	65 Mpps (w/2x 10 GbE)	65 Mpps (w/2x 10 GbE)	71 Mpps	71 Mpps	101 Mpps	101 Mpps
Buffering	128 MB	128 MB	128 MB	128 MB	192 MB	192 MB

Interface types

Following are the available interface types:

- 10/100/1000 Mbps Ethernet port with RJ45 connector
- 100/1000 Mbps Ethernet port with SFP connector
- 10 Gbps Ethernet port with XFP connector

Transceivers

Table 1-24 and Table 1-25 show the available transceivers that can be used.

Table 1-24 Transceivers for 100/1000 Mbps Ethernet ports

Type	Connector	Speed	Distance
1000BASE-T SFP copper	RJ-45	1 Gbps	Up to 100 m with CAT5 or higher
1000BASE-SX 850 nm SFP optics	LC	1 Gbps	Up to 550 m over multi-mode fiber
1000BASE-LX 1310 nm SFP optics	LC	1 Gbps	Up to 10 km over single-mode fiber
1000BASE-LHA 1550 nm SFP optics	LC	1 Gbps	Up to 70 km over single-mode fiber
100BASE-FX 1310 nm SFP optics	LC	100 Mbps	Up to 2 km over multi-mode fiber

Table 1-25 Transceivers for 10 Gbps Ethernet ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm XFP optics	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm XFP optics	LC	10 Gbps	Up to 10 km over single-mode fiber
10GBASE-ER 1550 nm XFP optics	LC	10 Gbps	Up to 40 km over single-mode fiber
10GBASE-CX4 XFP copper	LC	10 Gbps	Up to 15 m over CX4 grade copper

Optional features

Following are the optional features for c-series models:

- Full Layer 3 Premium Activation:

  Enables OSPFv2, IS-IS, IGMPv1/v2/v3, PIM-DM/-SM/-SSM, MSDP, Anycast RP, MPLS, VPLS, Multi-VRF, Ethernet Service Instance (ESI), IEEE 802.1ag Connectivity Fault Management (CFM), 802.1ad (Provider Bridges), and 802.1ah (Provider Backbone Bridges)

- Metro Edge Premium Activation:

  Enables OSPFv2, BGP-4, IS-IS, IGMPv1/v2/v3, PIM-DM/-SM/-SSM

Services, protocols, and standards

IBM c-series Ethernet Switches support these services, protocols, and standards.

Advanced Carrier-Grade Ethernet services are provided by:

- Up to 128k MAC addresses
- 4k VLANs/S-VLANs/B-VLANs
- Ability to reuse VLAN-ID on each port using Brocade innovative “Ethernet Service Instance” (ESI) framework
- IEEE 802.1ad Provider Bridges
- IEEE 802.1ah Provider Backbone Bridges
- IEEE 802.1ag Connectivity Fault Management
- Comprehensive set of Layer 2 control protocols: Foundry MRP/MRP-II, VSRP, RSTP, MSTP
- MEF 9 and MEF 14 certified
- E-LINE (EPL and EVPL), E-LAN and E-TREE support
- Protocol tunneling of customer BPDUs

Comprehensive IPv4 unicast routing support based on the rich feature set of Multi-Service IronWare:

- High performance, robust routing by Foundry Direct Routing (FDR) for complete programming of Forwarding Information Base (FIB) in hardware
- RIP, OSPF, IS-IS, BGP-4 support
- Support for VRRP and VRRP-E
- 8-path Equal Cost Multipath (ECMP)
- Up to 32k IPv4 unicast routes in FIB

Support for trunks (link aggregation groups) using either IEEE 802.3ad LACP or static trunks:

- Up to 12 links per trunk
- Support for single link trunk

Rich multicast support:

- Supported IPv4 multicast protocols include PIM-DM, PIM-SM, PIM-SSM
- IGMP v2/v3 routing and snooping support
- IGMP static groups support

- Multicast boundaries facilitate admission control
- Up to 4k multicast groups in hardware
- Multicast traffic distribution over LAGs
- Efficient Egress interface based replication maximizes performance and conserves buffer usage

Deep Egress buffering for handling transient bursts in traffic:

- 128 MB to 192 MB of buffering, based on configuration

Advanced QoS:

- Inbound and outbound two rate three color traffic policers with accounting
- 8 queues per port, each with a distinct priority level
- Multiple queue servicing disciplines: Strict Priority, Weighted Fair Queuing and hybrid
- Advanced remarking capabilities based on port, VLAN, PCP, DSCP, or IPv4 flow
- Egress port and priority-based shaping

Comprehensive hardware-based security and policies:

- Hardware-based Layer 3 and Layer 2 ACLs (both inbound and outbound) with logging
- Ability to bind multiple ACLs to the same port
- Hardware-based receive ACLs

Additional security capabilities:

- Port-based network access control using 802.1x or MAC port security
- Root guard and BPDU guard
- Broadcast, multicast and unknown unicast rate limits
- ARP Inspection for static entries

Advanced monitoring capabilities:

- Port and ACL-based mirroring allows traffic to be mirrored based on incoming port, VLAN-ID, or IPv4/TCP/UDP flow
- Hardware-based sFlow sampling allows extensive Layer 2-7 traffic monitoring for IPv4 and Carrier Ethernet services
- ACL-based sFlow support

Interface capabilities:

- Jumbo frame support up to 9,216 bytes
- Optical monitoring of SFP and XFP optics for rapid detection of fiber faults
- UDLD, LFS/RFN support

The whole list of supported standards and RFC compliance can be found at:

http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/

1.2.5 IBM s-series Ethernet Switches

IBM s-series Ethernet Switches are designed to extend control from the network edge to the backbone. The switches provide intelligent network services, including superior Quality of Service (QoS), predictable performance, advanced security, comprehensive management and integrated resiliency. A common operating system and shared interface and power supply modules between the Ethernet Switch B08S and B16S help reduce the cost of ownership by minimizing operational expenses and improving return on investment (ROI). A highly dense, resilient, and flexible architecture allows scaling up to 384 10/100/1000 Mbps Class 3 (15.4 watts) PoE capable ports or 36 ports of high-speed 10 GbE.

IBM s-series Ethernet Switches have an extensive feature set, making them well suited for real-time collaborative applications, IP telephony, IP video, e-learning and wireless LANs to raise an organization’s productivity. With wire-speed performance and ultra low latency, these systems are ideal for converged network applications such as VoIP and video conferencing. Providing one of the industry’s most scalable and resilient PoE designs, the 1 GbE PoE capable ports support the IEEE 802.1AB LLDP and ANSI TIA1057 LLDP-MED standards, enabling organizations to build advanced multi-vendor networks.

LLDP enables discovery of accurate physical network topologies, including those that have multiple VLANs where all subnets might not be known. LLDP-MED advertises media and IP telephony specific messages, providing exceptional interoperability, IP telephony troubleshooting, and automatic deployment of policies, advanced PoE power negotiation, and location/emergency call service. These features make converged network services easier to install, manage, and upgrade, significantly reducing operational costs.

Following are the s-series key features:

- Industry-leading, chassis-based convergence solution provides a scalable, secure, low-latency and fault-tolerant infrastructure for cost-effective deployment of Voice over IP (VoIP), wireless, and high-capacity data services throughout the enterprise.
- N+1 power redundancy design to enhance power operation and simplify system configuration.
- A rich suite of security features including IP source guard, dynamic Address Resolution Protocol (ARP) inspection, and DHCP snooping shields the enterprise from internal and external threats.

- Highest Class 3 PoE capacity in the industry; the s-series B16S scales to 36 10-GE and 384 PoE ports of 10/100/1000 Mbps, each capable of delivering 15.4 watts to provide customers with a convergence-ready infrastructure that will scale to support future growth.
- Combined SP (Strict Priority) /WRR (Weighted Round Robin) queuing and cell-based switch fabric ensure low latency and jitter for voice and video traffic.
- Intelligent PoE and configuration management with LLDP Link Layer Discovery Protocol), LLDP-MED (Link Layer Discovery Protocol-Media Endpoint Discovery) and PoE Prioritization for IP Phones.
- Redundant architecture and resilient protocols ensure business continuity in the event of network or equipment failure(s).
- Embedded, hardware-based sFlow traffic monitoring enables network-wide accounting, utilization reporting, capacity planning, intrusion detection, and more.
- Advanced IronWare Layer 2 Ethernet switching with robust suite of security capabilities including extended ACLs, MAC filters, TCP and IGMP denial of service protection, spanning tree BPDU Guard, Root Guard, unicast and multicast rate limiting, Metro Ring Protocol, Virtual Switch Redundancy Protocol, and more.
- Flexibility option to upgrade the software to full Layer 3, including support for IP routing protocols such as RIPv1/v2, OSPF, BGP, and support for multicast routing.
- IronShield 360 intrusion protection delivers dynamic and real-time protection from network and host-based attacks.
- IPv6 capable blades to future proof the network for IPv6 application migration.

The s-series Ethernet Switches are available in the following model configurations:

- IBM Ethernet Switch B08S (4003-S08): Switch with redundant management and switch fabric modules for enhanced system resilience; 464 Gbps data switching capacity, PoE over tri-speed 10/100/1000 Mbps interfaces
- IBM Ethernet Switch B16S (4003-S16): Switch with redundant management and switch fabric modules for enlaced system resilience; 848 Gbps data switching capacity, up to 384 Class 3 PoE ports with N+1 power redundancy making it the most powerful PoE solution in the industry, PoE over tri-speed 10/100/1000 Mbps interfaces

Both configurations are shown in Figure 1-7.

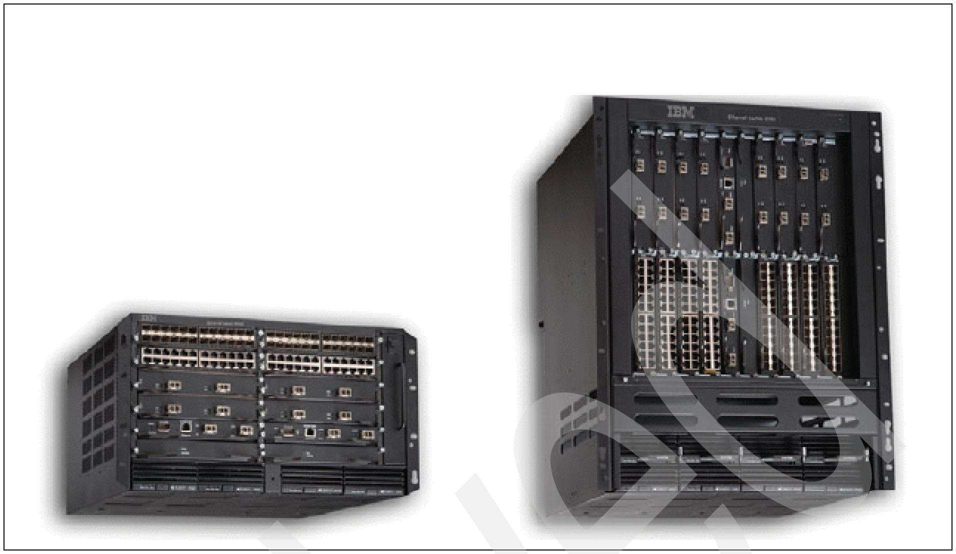


Figure 1-7   IBM s-series Ethernet Switches

All s-series models can only be installed in the rack. Non-rack installation is not supported.

Operating system

All s-series systems run Brocade IronWare R5.0.00 or higher operating system.

Future-proofing the network with IPv6

Migration to IPv6 is inevitable, but by starting with the deployment of IPv6-capable hardware the transition can be more controlled and less disruptive to the network. Japan and Europe are aggressively deploying IPv6, and deployment in North America is on the rise. In fact, some government agencies are mandating the purchase of IPv6-capable switches and routers. Therefore, it is important that enterprises and service providers plan to deploy IPv6-capable devices to capitalize on this inevitable change.

Configuration alternatives

The s-series family of switches is optimized for flexibility with upgradeability for PoE, redundant management, switch fabric and power, and 10 Gigabit Ethernet. Available in three chassis models, the scalable s-series family helps enterprises and service providers reduce costs and gain the operational benefits of a common operating system, a shared interface, and common power supply modules.

High-quality and reliable network convergence

The s-series family provides a scalable, secure, low latency, and fault-tolerant infrastructure for cost-effective integration of VoIP, video, wireless access, and high-performance data onto a common network. The system architecture features a scalable and resilient PoE design and a low-latency, cell-based switch fabric with intelligent traffic management to ensure reliable and high-quality VoIP service.

A rich suite of security features, including policy-based access control, IP source guard, dynamic ARP inspection, and DHCP snooping, work in unison to control network access and shield the network from internal and external threats. The s-series family establishes a new class of convergence-ready solutions, enabling organizations to implement a secure, reliable, scalable, and high-quality infrastructure for total network convergence.

Resilient green power distribution and consumption

The s-series family features a unique power distribution design for the system and PoE power. The chassis are designed with independent systems and PoE power subsystems. This design achieves optimal power operation and configuration, reducing the equipment and ongoing costs, in comparison to modular systems that use a common power supply for both the systems and the PoE equipment. In the s-series family, the power consumption of a line module's PoE circuitry does not impact the system power.

Similarly, the power consumption of the line modules, switch modules, and management modules does not impact the PoE power. Power consumption for the system and PoE are calculated, provisioned, and managed independently of one another. As more PoE devices are added to a switch, a simple power budget calculation determines whether another PoE power supply needs to be added to the switch.

The system power distribution and the PoE power distribution subsystems are each designed for M+N load-sharing operation. This dual-distribution power design simplifies the power configuration of the system while enhancing system reliability. The chassis can be configured for a wide range of power environments including: 110V/220V AC power, -48V DC power and mixed AC/DC power configurations. To scale PoE configurations, PoE power supplies are available in two ratings of 1250W and 2500W. When configured with four 2500W PoE supplies, the s-series supports up to 384 10/100/1000 Mbps Class 3 PoE ports and still maintains N+1 power redundancy. This resiliency is unmatched in the industry.

Intelligent and scalable Power over Ethernet

Power over Ethernet (PoE) is a key enabler of applications such as VoIP, IEEE 802.11 wireless LANs, and IP video. The s-series is Brocade's third-generation PoE-capable switch family and incorporates the latest advances in PoE provisioning and system design, delivering scalable and intelligent PoE to the enterprise. The PoE power distribution subsystem is independent of the system power, eliminating system disruption in the event of PoE over-subscription or a PoE power failure.

Customers have the choice of purchasing PoE-ready line modules or upgrading 10/100/1000 Mbps line modules when needed with field-installable PoE daughter modules. PoE power per port can be manually or dynamically configured. Dynamic configuration is supported using standards-based auto discovery or legacy Layer 2 discovery protocols. Port priorities are also configurable and are used to prioritize PoE power in over-subscribed configurations.

Advanced QoS and low latency for enterprise convergence

The s-series family offers superior Quality of Service (QoS) features that enable network administrators to prioritize high-priority and delay-sensitive services throughout the network. S-series switches can classify, re-classify, police, mark, and re-mark an Ethernet frame or an IP packet prior to delivery. This flexibility lets network administrators discriminate among various traffic flows and enforce packet-scheduling policies based on Layer 2 and Layer 3 QoS fields.

After being classified, the traffic is queued and scheduled for delivery. Three configured queuing options provide the network administrator with flexible control over how the system services the queues. Weighted Round Robin (WRR) queuing applies user-configured weighting for servicing multiple queues, ensuring that even low priority queues are not starved for bandwidth. With Strict Priority (SP) queuing, queues are serviced in priority order ensuring that the highest-priority traffic is serviced ahead of lower priority queues. Combined SP and WRR queuing ensures that packets in the SP queue are serviced ahead of the WRR queues. Combined queuing is often used in VIP networks where the VIP traffic is assigned to the SP queue and data traffic to the WRR queues.

In addition, the switch management modules are available with integrated Gigabit Ethernet or 10-Gigabit Ethernet ports. These modules provide cost-effective system configurations supporting high-capacity connections to upstream switches. The management modules utilize high-performance system processors with high-capacity memory for scalable networking up to a routing capacity of 1 million BGP routes and 20 BGP peers.

The s-series switches utilize an advanced cell-based switch fabric with internal flow-control, ensuring very low latency and jitter performance for converged applications.

Ease of use: plug and play

The s-series family supports the IEEE 802.1AB LLDP and ANSI TIA 1057 LLDP-MED standards, enabling organizations to build open convergence, advanced multi-vendor networks. LLDP greatly simplifies and enhances network management, asset management, and network troubleshooting. For example, it enables discovery of accurate physical network topologies, including those that have multiple VLANs where all subnets might not be known.

LLDP-MED addresses the unique needs that voice and video demand in a converged network by advertising media and IP telephony specific messages that can be exchanged between the network and the endpoint devices. LLDP-MED provides exceptional interoperability, IP telephony troubleshooting, and automatic deployment of policies, inventory management, advanced PoE power negotiation, and location/emergency call service. These sophisticated features make converged network services easier to install, manage, and upgrade and significantly reduce operations costs.

Flexible bandwidth management

The s-series switches support a rich set of bandwidth management features, allowing granular control of bandwidth utilization. On ingress, extended ACLs can be used in combination with traffic policies to control bandwidth by user, by application, and by VLAN. On egress, outbound rate limiting can control bandwidth per port and per priority queue. These features allow the network operator fine-grained control of bandwidth utilization based on a wide range of application and user criteria.

Complete solution for multicast and broadcast video

The use of video applications in the workplace requires support for scalable multicast services from the edge to the core. IGMP and PIM snooping improves bandwidth utilization in Layer 2 networks by restricting multicast flows to only those switch ports that have multicast receivers. In Layer 3 networks, support for IGMP (v1, v2, and v3), IGMP Proxy, PIM-SM, PIM-SSM, and PIM-DM multicast routing optimizes traffic routing and network utilization for multicast applications.

Advanced full Layer 2/3 wire-speed IP routing solution

Advanced IronWare supports a full complement of unicast and multicast routing protocols, enabling users to build fully featured Layer 2/Layer 3 networks. Supported routing protocols include RIPv1/v2, OSPF, PIM-SM/DM, BGP, and Equal Cost Multi-path (ECMP) for improved network performance. M2, M3, and M4 management modules can support routing table capacity of up to 1,000,000 BGP routes and 20 BGP peers. s-series switches can be upgraded with Advanced IronWare routing software (a Layer 3 upgrade).

To achieve wire-speed Layer 3 performance, the s-series switches support Foundry Direct Routing (FDR), in which the Forwarding Information Base (FIB) is maintained in local memory on the line modules. The hardware forwarding tables are dynamically populated by system management with as many as 256,000 routes.

Comprehensive bulletproof security suite

Security is a concern for today’s network managers, and the s-series switches support a powerful set of network management solutions to help protect the switch. Multilevel access security on the console and a secure Web Management interface prevent unauthorized users from accessing or changing the switch configuration. Using Terminal Access Controller Access Control Systems (TACACS/TACACS+) and RADIUS authentication, network managers can enable considerable centralized control and restrict unauthorized users from altering network configurations.

The s-series family includes Secure Shell (SSHv2), Secure Copy, and SNMPv3 to restrict and encrypt communications to the management interface and system, thereby ensuring highly secure network management access. For an added level of protection, network managers can use ACLs to control which ports and interfaces have TELNET, web, and/or SNMP access.

After the user is permitted access to the network, protecting the user’s identity and controlling where the user connects becomes a priority. To prevent “user identity theft” (spoofing), the s-series switches support DHCP snooping, Dynamic ARP inspection, and IP source guard. These three features work together to deny spoofing attempts and to defeat man-in-the-middle attacks. To control where users connect, the s-series switches support private VLANs, quarantine VLANs, policy-based routing, and extended ACLs, all of which can be used to control a user’s access to the network.

In addition, s-series switches feature embedded sFlow packet sampling, which provides system-wide traffic monitoring for accounting, troubleshooting, and intrusion detection.

Resilient design for business continuity

A s-series networking solution is built for high-value environments. Featuring redundant management modules, redundant fans, redundant load-sharing switch fabrics, and power supply modules, the s-series switches are designed for maximum system availability. Switch fabric failover preserves network connectivity in the event of a switch module failure. Automatic management failover quickly restores network connectivity in the event of a management module failure.

In the event of a topology change due to a port or facility failure, Layer 1 and Layer 2 protocols, such as Protected Link, Metro Ring Protocol (MRP), IEEE 802.3ad, UDLD, VSRP, and Rapid Spanning Tree Protocol, can restore service in sub-second time (tens to hundreds of milliseconds, depending on the protocol), protecting users from costly service disruption. Enhanced spanning tree features such as Root Guard and BPDU Guard prevent rouge hijacking of spanning tree root and maintain a contention and loop-free environment especially during dynamic network deployments. These high availability capabilities enable network deployments of a highly reliable network infrastructure that is resilient to, and tolerant of, network and equipment failures.

Investment protection through IPv6 capable hardware

Networks are in the early stages of large-scale IPv6 production deployment, however few IPv6 innovative applications are currently on the market. Although the success of IPv6 will ultimately depend on the new applications that run over IPv6, a key part of the IPv6 design is the ability to integrate into and coexist with existing IPv4 switches within the network and across networks during the steady migration from IPv4 to IPv6.

Following are benefits of the IPv6-capable modules:

- The IPv6-capable s-series management modules are non-blocking, with a built-in switch fabric module and 12 combination Gigabit Ethernet copper or fiber ports that provide connectivity to your existing management network.
- The IPv6-capable s-series management modules have a console port and a 10/100/1000 port for out-of-band management. The management modules optionally support 2-port 10-GbE ports.
- The IPv6-capable s-series management modules are interchangeable between devices.
- Redundant management modules on the IPv6-capable s-series provide 100% redundancy.
- The crossbar (xbar) architecture enables the management module to switch 30 Gbps between each interface module and within the management module.
- The IPv6-capable interface modules and power supplies are interchangeable among s-series switches.
- The IPv6-capable s-series management, switch fabric, and interface modules are hot swappable, which means a module can be removed and replaced while the chassis is powered on and running.

Physical and thermal parameters

The physical and thermal parameters are shown in Table 1-26.

Table 1-26 s-series physical and thermal parameters

Component	IBM Ethernet Switch B08S	IBM Ethernet Switch B16S
Chassis type	Modular 8-Slot chassis	Modular 16-Slot chassis
H/W/D (cm)	26.3 x 44.5 x 43.8	62.2 x 44.5 x 43.8
Rack units (RUs)	6	14
Max. weight	31 kg	88.6 kg
Max. power draw	1428 W	2440 W
Max. power draw with PoE using 1250 W PoE PS	4203 W	7990 W
Max. power draw with PoE using 2500 W PoE PS	5227 W	10037 W
Op. temperature (°C)	0 - 40 °C	0 - 40 °C
Heat emission (BTU/hr)	4874	8326
Heat emission (BTU/hr) with PoE using 1250 W PoE PS	6986	12551
Heat emission (BTU/hr) with PoE using 2500 W PoE PS	7747	14073
Airflow	Side-to-side	Front-to-back
Fan tray/assemblies	1	2

The cooling system of the s-series is configured as follows:

- IBM Ethernet Switch B08S (4003-S08): Is equipped with six fans.
- IBM Ethernet Switch B16S (4003-S16): Is equipped with two fans in the rear of the chassis.

All the fans are hot swappable and have adjustable speeds.

Power parameters

All s-series models provide redundant and removable power supplies with AC power options. Power supplies can be exchanged between B08S and B16S models.

The s-series models provide Power over Ethernet (PoE) option.

There are separate power supplies for system power (SYS) and PoE power (PoE). Power consumption between PoE and SYS power supplies is not shared, meaning that the loss of a System power supply does not impact a PoE power supply, and vice versa.

System power supplies have internal power of 12V and PoE power supplies have internal power of 48V.

All power supplies are auto-sensing and auto-switching. All are hot swappable and can removed and replaced without powering down the system.

The system power parameters are shown in Table 1-27.

Table 1-27 s-series SYS power parameters

Component	**IBM Ethernet Switch B08S**	**IBM Ethernet Switch B16S**
1200W system (SYS) Power Supplies	100 - 240 VAC / 50 - 60 Hz	100 - 240 VAC / 50 - 60 Hz
Number of SYS power supply bays	2	4
Number of SYS power supply bays required for fully loaded chassis	1	2

The system (SYS) power supplies provide power to the management module, all non-PoE interface modules, and all ports on PoE modules that do not require PoE power or to which no power-consuming devices are attached. The installed SYS power supplies provide power to all chassis components, sharing the workload equally. If a SYS power supply fails or overheats, the failed power supply’s workload is redistributed to the redundant power supply.

The PoE power parameters are shown in Table 1-28.

Table 1-28 s-series PoE power parameters

Component	IBM Ethernet Switch B08S	IBM Ethernet Switch B16S
1250W PoE power supplies	100 - 240 VAC / 50 - 60 Hz	100 - 240 VAC / 50 - 60 Hz
2500W PoE power supplies	200 - 240 VAC / 50 - 60 Hz	200 - 240 VAC / 50 - 60 Hz
Number of PoE power supply bays	2	4
Number of PoE power supply bays required for fully loaded chassis	1	3

The PoE Power Supplies provide power to the PoE daughter card, and ultimately to PoE power consuming devices. The installed PoE power supplies share the workload equally. If a PoE power supply fails or overheats, the failed power supply's workload is redistributed to the redundant power supply. The number of PoE power-consuming devices that one PoE power supply can support depends on the number of watts (Class) required by each power-consuming device (PD).

The number of PoE power-consuming devices that one 1250W PoE power supply can support depends on the number of watts required by each power-consuming device. Each supply can provide a maximum of 1080 watts of PoE power, and each PoE port supports a maximum of 15.4 watts of power per PoE power-consuming device. For example, if each PoE power-consuming device attached to the s-series consumes 15.4 watts of power, one power supply will power up to 70 PoE ports. You can install additional power supply for additional PoE power.

Each 2500W PoE power supply can provide a maximum of 2160 watts of PoE power, and each PoE port supports a maximum of 15.4 watts of power per PoE power-consuming device. For example, if each PoE power-consuming device attached to the s-series consumes 15.4 watts of power, it will supply power up to 140 PoE ports.

Note: The system powers on as many PoE ports as each PoE power supplies can handle. The system calculates the maximum number of PoE ports it can support based on the number of PoE power supplies installed. PoE ports are enabled based on their priority settings. Keep in mind that the system will reserve the maximum configured power per PoE-enabled port, even if the PoE power-consuming device is drawing less power.

In the B08S chassis, the system power supplies occupy slot numbers 3 and 4 on the right, with the redundant supply in slot 4. The PoE power supplies occupy slot numbers 1 and 2 on the left. Figure 1-8 shows power supply placement.

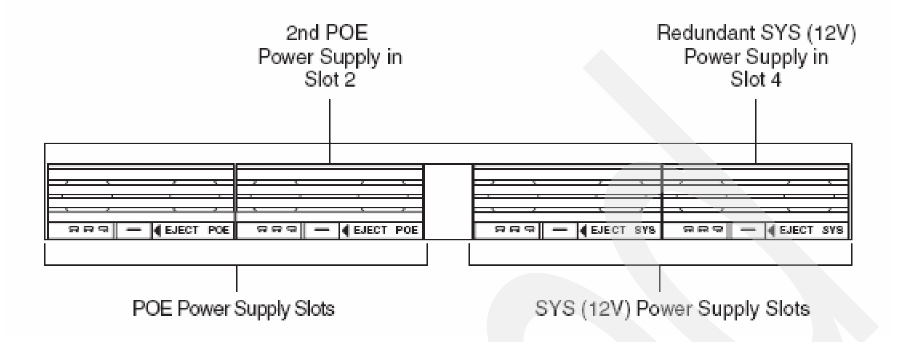


Figure 1-8 BS08S power supply placement

In the B16S chassis, the system power supplies occupy slot numbers 1 – 4 in the top row with the redundant supplies in slot numbers 3 and 4. The PoE power supplies occupy slot numbers 5 – 8 in the bottom row. Figure 1-9 shows power supply placement.

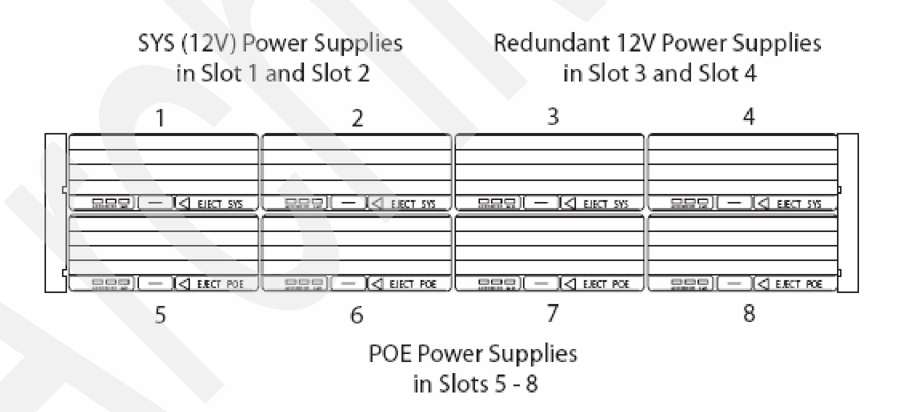


Figure 1-9 BS16S power supply placement

What happens when one or more system power supplies fail:

If one or more system power supplies fail and the system is left with less than the minimum number of power supplies required for normal operation, the power supplies will go into overload and the system will start to shut down.

Several things can happen with a system power supply failure. The output voltage of the remaining good power supplies will likely drop as they try unsuccessfully to generate more power than they are capable of. The system will react to a drop in voltage by increasing the current draw. The hardware will shut down due to over-current protection or under-voltage protection, whichever takes place first. One by one, the interface modules will shut down until the power is within the power budget of the remaining power supplies. There is no particular order in which the interface modules will shut down, as this will occur in hardware and not in software. The management CPU requires power as well, and can also shut down during a power supply failure.

What happens when one or more PoE power supplies fail:

If one or more PoE power supplies fail and the system is left with less than the minimum number of PoE power supplies, the PoE power supplies will go into overload. Non-PoE functions will not be impacted, provided the System power supplies are still up and running.

Several things can happen with a PoE power supply failure. The output voltage of the remaining good power supplies will likely drop as they try unsuccessfully to generate more power than they are capable of. The system will react to a drop in voltage by increasing the current draw. The hardware will shut down PoE function due to over-current protection or under-voltage protection, whichever occurs first. The interface modules will start to shut down its PoE ports one by one until the over-power is within the power budget of the remaining power supplies. There is no particular order in which the PoE ports will shut down, as this occurs in hardware and not in software.

After a power loss, if the system is left with less than the minimum number of power supplies required for normal operation, the system will be left in an unknown state. At this point, manual recovery is required (that is, restore power and power cycle the chassis).

Slots, ports, memory, and performance

The s-series family employs a distributed switching design to deliver high-speed forwarding across the platform. Switching between ports on the same module is performed locally, and switching between ports across modules is performed across the crossbar switch fabric.

Each model has two management modules and two fabric module slots that are exchangeable between all s-series models. Two management and fabric modules provide redundancy as s-series can also operate with only one management and/or one fabric module. In case of only one operational fabric module, performance can be degraded in certain traffic patterns. All models have passive backplanes.

All s-series models have maximum 512MB RAM and management processor with 667 MHz.

The number of slots, ports, and performance metrics are shown in Table 1-29.

Table 1-29 Slots, ports, and performance metrics

Component	IBM Ethernet Switch B08S	IBM Ethernet Switch B16S
Interface slots	8	16
Max. number of slots for management modules	2	2
Min. number of management modules required for operations	1	1
Max. number of slots for fabric modules	2	2
Min. number of switch fabric modules required for fully-loaded chassis at wire-speed	2	2
Min. number of switch fabric modules required for operations	1	1
10/100/1000 copper ports per module	24	24
1 GbE SFP ports per module	24	24
10 GbE ports per module	2	2
Backplane switching capacity	600 Gbps	1080 Gbps
Data switching capacity	464 Gbps	848 Gbps
Packet forwarding capacity	384 Mpps	636 Mpps

All modules are hot-swappable and do not require power-off to be replaced.

Table 1-30 and Table 1-31 show the available port density.

Table 1-30 Non-PoE port density

Port type	IBM Ethernet Switch B08S	IBM Ethernet Switch B16S
100 Base FX (SFP)	200	392
1000 Base X (SFP)	200	392
1000 Base T (RJ45)	200	392
10/100/1000 Base Total (SFP+RJ45)	200	392
10 Base X (XFP)	20	36

Table 1-31 PoE port density

Port type	IBM Ethernet Switch B08S	IBM Ethernet Switch B16S
IEEE 802.3af Class-1,2,3 10/100/1000	192	384
IEEE 802.3af Class-1,2,3 10/100/1000 with N+1 power supply redundancy	140	384

Management modules

The following types of management modules are available:

- IPv4 management module
- IPv4 management module with 2-port 10GbE (XFP)
- IPv6 management module with 2-port 10GbE (XFP)

Interface modules

Table 1-32 shows which modules can be installed in the s-series chassis interface slots.

Table 1-32 Interface modules

IP version and speed	Number of ports	Connector type
IPv4 10/100/1000MbE	24	RJ45
IPv4 100/1000MbE	24	SFP
IPv4 10GbE	2	XFP
IPv6 10/100/1000MbE	24	RJ45
IPv6 100/1000MbE	24	SFP
IPv6 10GbE	2	XFP

Interface types

Following are the available interface types:

- 10/100/1000 Mbps Ethernet port with RJ45 connector
- 100/1000 Mbps Ethernet port with SFP connector
- 10 Gbps Ethernet port with XFP connector

Transceivers

Table 1-33 and Table 1-34 show the available transceivers to be used in interface modules.

Table 1-33 Transceivers for 100/1000 Mbps Ethernet ports

Type	Connector	Speed	Distance
1000BASE-T SFP Copper	RJ-45	1Gbps	Up to 100 m with CAT5 or higher
1000BASE-SX 850 nm SFP optics	LC	1 Gbps	Up to 550 m over multi-mode fiber
1000BASE-LX 1310 nm SFP optics	LC	1 Gbps	Up to 10 km over single-mode fiber
1000BASE-LHA 1550 nm SFP optics	LC	1 Gbps	Up to 70 km over single-mode fiber
100BASE-FX 1310 nm SFP optics	LC	100 Mbps	Up to 2 km over multi-mode fiber

Table 1-34 Transceivers for 10 Gbps Ethernet ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm XFP optics	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm XFP optics	LC	10 Gbps	Up to 10 km over single-mode fiber
10GBASE-ER 1550 nm XFP optics	LC	10 Gbps	Up to 40 km over single-mode fiber
10GBASE-CX4 XFP copper	LC	10 Gbps	Up to 15 m over CX4 grade copper

Optional features

The s-series is capable of providing Layer 3 functions. Following are the optional features:

- IPv4 Full Layer 3 Premium Activation:

  Enables RIPv1/v2, OSPFv2, BGP-4, IGMPv1/v2/v3, PIM-SM/-DM/-SSM, VRRP-E

- IPv6 Full IPv4 Layer 3 Premium Activation:

  Enables RIPv1/v2, OSPFv2, BGP-4, IGMPv1/v2/v3, PIM-SM/-DM/-SSM, VRRP-E

- IPv6 Full IPv6 Layer 3 Premium Activation:

  Enables RIPv1/v2, RIPng, OSPFv2, OSPFv3, BGP-4, IGMPv1/v2/v3, PIM-SM/-DM, DVMRP, VRRP-E

Services, protocols, and standards

IBM s-series Ethernet Switches support these services, protocols, and standards.

PoE capable ports support the following standards:

- IEEE 802.1AB
- LLDP
- ANSI TIA 1057 LLDP-MED

Various Layer 2 and Layer 3 protocols are supported:

- Protected Link Groups
- Link Aggregation (IEEE 802.3ad, LACP)
- UDLD
- Virtual Switch Redundancy Protocol (VSRP)
- STP/RSTP/MSTP
- Root Guard
- BPDU Guard
- MRP
- VLAN stacking

Layer 3 Premium Activation adds support for these features:

- IGMP (v1,2,3)
- IGMP Proxy
- PIM-SM/PIM-SSM/PIM-DM multicast routing
- RIP v1/2
- OSPFv2
- BGP-4
- Equal Cost Multi Path (ECMP)

Layer 3 IPv4+IPv6 Premium activation is adding support for these features:

- OSPFv3
- RIPng
- IPv6 over IPv4 tunnels
- IPv6 ACLs

QoS is provided with these features:

- Weighted round robin (WVR) queuing
- Strict priority (SR) queuing
- Extended ACLs with traffic policies on ingress traffic for controlling the bandwidth per user, application or VLAN

Powerful sets of network management solutions are supported:

- Multilevel access security on the console
- Secure Web Management
- Terminal Access Controller Access Control (TACACS/TACACS+)
- Remote Authentication Dial in User Service (RADIUS) authentication

Access to management interface can be restricted and encrypted; this can be achieved by using these features:

- Secure Shell (SSHv2) access
- Secure Copy (SCPv2)
- SNMPv3
- HTTPS
- ACLs to define which ports and interfaces have CLI, web, and/or SNMP access

To prevent “user identity theft” (spoofing), the s-series supports these features:

- DHCP snooping
- Dynamic ARP inspection
- IP source guard

For a complete list of supported standard and RFC compliant features, see the website:

http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/

1.2.6 IBM g-series Ethernet Switches

IBM g-series Ethernet access switches provide enterprise organizations with a flexible and feature-rich solution for building a secure and converged network edge. The switches support 48 x 1 GbE RJ45 ports including 4x 1 GbE SFP combination ports. The B48G is upgradeable with two 10 GbE uplink ports to consolidate connections into the enterprise aggregation point, campus LANs, or metro area networks. The B50G comes with 2x 10 GbE CX4 stacking ports, providing the flexibility of a “pay-as-you-grow” architecture.

Both models enable a converged solution for vital network applications such as VoIP, wireless access, WebTV, video surveillance, building management systems, triple play (voice + video + data) services and remote video kiosks in a cost-effective, high-performance compact design.

Following are the g-series key features:

- Compact 48-port 10/100/1000 Mbps access switch models; field upgradeable with Power over Ethernet (PoE), 10 Gigabit Ethernet, and IronStack stacking for scalable and secure network access
- Hot-swappable, load-sharing AC and DC power supply options
- Industry leading IEEE 802.3af PoE Class 3 port capacity in a compact form factor delivers a scalable and cost-effective solution for unified communications at the network edge

- Advanced IronWare Layer 2 Ethernet switching with robust suite of security capabilities including ACLs, MAC filters, TCP SYN and ICMP denial of service (DoS) protection, Spanning Tree BPDU Guard, Root Guard, unicast, broadcast and multicast rate limiting, 802.1X authentication, and enhanced lawful intercept features
- Base Layer 3 capabilities enable routed topologies to the network edge; supported features include: RIP v1/v2 route announcement, static routes, virtual and routed interfaces, DHCP relay, and VRRP (Virtual Router Redundancy Protocol)
- Open and standards-based network access control features multi-host 802.1x access control, multi-device MAC authentication, and policy controlled MAC-based VLANs
- Low packet latency and advanced Quality of Service (QoS) with eight hardware-based priority queues and combined strict priority and weighted round robin scheduling ensure dependable and high-quality network convergence
- Embedded hardware-based sFlow packet sampling enables network wide traffic monitoring for traffic accounting, intrusion detection, 802.1x identity monitoring, link utilization, and fault isolation
- IronShield 360 intrusion protection delivers dynamic, real-time protection from network and host-based attacks
- Brocade's IronStack stacking technology provides cost-effective expansion at network edge with operational simplicity of a single switch
- IronStack supports up to eight B50G units in a logical chassis scaling to 384 PoE ports, and it features automatic failover in event of link fault or active controller failure and hot insertion and cross-unit trunking for increased resilience

The g-series Ethernet Switches are available in these model configurations:

- IBM Ethernet Switch B48G (4002-G4A;4002-AG4): 48-port 10/100/1000 with 4-port combo ports that support 10/100/1000 or 100/1000 SFP connections and one redundant, removable power supply; field upgradeable to include a 2-port 10GbE module for either XFP or CX4 connectivity and another redundant removable power supply; PoE upgradeable
- IBM Ethernet Switch B50G (4002-G5A;4002-AG5): The same connectivity, availability and PoE features as B48G plus advanced IronStack stacking technology over 2 x 10 GbE CX4 ports

Both models are shown in Figure 1-10.

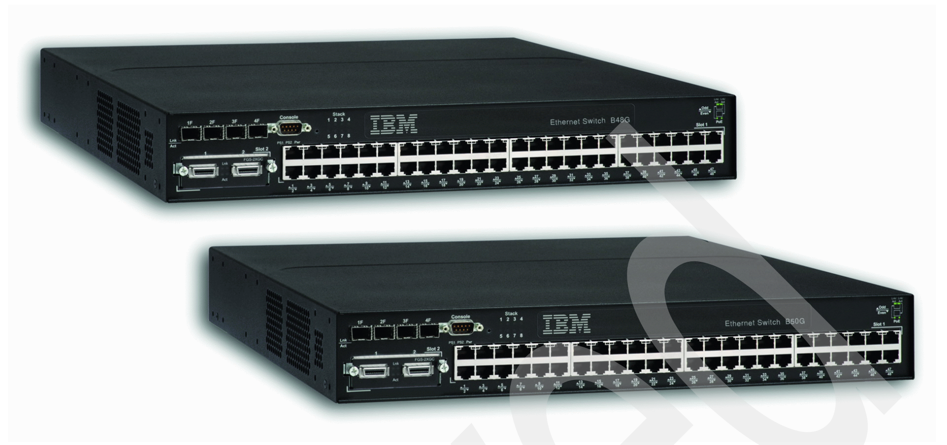


Figure 1-10 IBM g-series Ethernet Switches

All g-series models can only be installed in the rack. Non-rack installation is not supported.

Operating system

G-series B48G is running Brocade IronWare R4.3.01 or higher and B50G is running R5.0.01 or higher version of operating system.

Performance and scalability

Today’s enterprise organizations require cost-effective, flexible, and secure solutions for delivering data and unified communication services on a network architecture that can scale and evolve to meet their ever-changing needs. The g-series is designed to meet these requirements. Its wire-speed architecture delivers non blocking performance for high-speed Gigabit Ethernet desktops while providing QoS assurances at VoIP endpoints.

For cost-effective and rapid scaling at the network edge, the g-series is equipped with IronStack stacking technology, which supports stacking up to eight units in a virtual chassis. The IronStack system supports 40-Gbps switching capacity between stacked units providing a high-capacity interconnect across the stack. g-series IronStack supports stacking over copper and fiber cables. This provides for flexible stack configurations in which stacked units can be separated by more than several hundred meters of fiber.

Convergence planning and deployment can occur over an extended period, and budget constraints might require phased deployments. The g-series models make it easy to deploy a solution today that can be upgraded later to support PoE,10-GbE, and stacking as needed.

Each power supply within a g-series delivers up to 480 watts of PoE power. In a dual power supply configuration, up to 48 10/100/1000 Mbps PoE ports of 15.4 watts per port (full Class 3) can be supported. This scalability enables the network manager to size the installation to meet current needs and have room for future growth.

As network traffic increases, network managers can easily upgrade to 10-GbE to provide high-capacity connectivity to the network backbone and/or high-performance server. The B48G can be upgraded in the field with a two-port 10-GbE XFP/CX4 module.

High availability hardware features

Convergence solutions such as VoIP require high availability, especially for the power supplies that power the PoE interfaces. G-series switches fulfill this requirement with dual, hot-swappable AC or DC power supplies. Both redundant AC and redundant DC power configurations are included.

The g-series features 1+1 power redundancy, using hot-swappable and field replaceable power modules, which install into the rear of the unit. The power modules are load-sharing supplies providing full 1+1 redundancy for as many as 48 Class 1and Class 2 PoE ports and 31 Class 3 (15.4 watts) PoE ports.

Additional design features include intake and exhaust temperature sensors and fan spin detection to aid in rapid detection of abnormal or failed operating conditions to help minimize mean time to repair.

IronStack solution

IronStack is advanced stacking technology that supports stacked configurations in which as many as eight g-series switches can be interconnected and maintain the operational simplicity of a single switch. Each IronStack enabled g-series model can support up to 40Gbps of stacking bandwidth per unit. IronStack configurations can be built using 10-GbE CX4 copper or XFP-based fiber connections. When XFP-based fiber connections are used, an IronStack configuration can be extended between racks, floors, and buildings with fiber lengths up to several hundred meters.

The B50G models are pre-configured with a two-port 10-GbE CX4 module, expanded CPU memory, and IronStack license (IronStack PROM) and software.

An IronStack system operates as a single logical chassis (with a single IP management address) and supports cross-member trunking, mirroring, switching, static routing, sFlow, multicast snooping, and other switch functions across the stack. An IronStack stack has a single configuration file and supports remote console access from any stack member. Support for active-standby controller failover, stack link failover, and hot insertion/removal of stack members delivers the resilience that is typical of higher end modular switches.

High density and full class 3 Power over Ethernet

When configured with Power over Ethernet (PoE), the g-series switches support IEEE 802.3af standards-based PoE on all ports. The g-series switches capability to deliver high-density, full-power PoE on all ports reduces the need to purchase additional hardware to support the higher power requirements.

When configured with dual power supplies, the 48-port g-series switch supports up to 48 10/100/1000 Class 3 (15.4 watts) PoE ports, which is one of the highest Class 3 PoE port density for a compact switch in the industry. These capacities are a significant advantage for environments that require full Class 3 power for devices such as surveillance cameras, color LCD phones, point-of-service terminals, and other powered endpoints.

An IronStack configuration of eight g-series switches can support as many as 384 PoE ports supporting full Class 3 PoE power without the need for external power supplies. Other solutions require external power supplies adding installation and operational complexity.

Ease of use: plug and play

The g-series supports the IEEE 802.1AB LLDP and ANSI TIA 1057 LLDP-MED standards that enable organizations to deploy interoperable multi-vendor solutions for unified communications. Configuring IP endpoints, such as VoIP, stations can be a complex task requiring manual and time-consuming configuration.

LLDP and LLDP-MED address these challenges, providing organizations with a standard and open method for configuring, discovering, and managing their network infrastructure. The LLDP protocols help reduce operations costs by simplifying and automating network operations. For example, LLDP-MED provides an open protocol for configuring QoS, security policies, VLAN assignments, PoE power levels, and service priorities. Additionally, LLDP-MED provides for the discovery of device location and asset identity, information that is used for inventory management and by emergency response services. These sophisticated features make converged networks services easier to deploy and operate while enabling new and critical services.

Comprehensive enterprise-class security

The g-series switches are powered by IronWare operating software, which offers a rich set of Layer 2 switching services, Base Layer 3 functionality, an advanced security suite for network access control (NAC), and DoS protection. IronWare security features include protection against TCP SYN and ICMP DoS attacks, Spanning Tree Root Guard and BPDU Guard to protect network spanning tree operation, and broadcast and multicast packet rate limiting.

Network access control

Network managers can rely on features such as multi-device and 802.1X authentication with dynamic policy assignment to control network access and perform targeted authorization on a per-user level. Additionally, the g-series supports enhanced MAC policies with the ability to deny traffic to and from a MAC address on a per-VLAN basis. This powerful tool allows network administrators to control access policies per endpoint device.

Standards-based NAC enables network operators to deploy best-of-breed NAC solutions for authenticating network users and validating the security posture of a connecting device. Support for policy-controlled MAC-based VLANs provides additional control of network access, allowing for policy-controlled assignments of devices to Layer 2 VLANs.

Traffic monitoring and lawful Intercept

Organizations might need to set up traffic intercept, that is, lawful intercept, due to today’s heightened security environment. For example, in the United States, the Communications Assistance for Law Enforcement Act (CALEA) requires businesses be able to intercept and replicate data traffic directed to a particular user, subnet, port, and so on. This capability is particularly essential in networks implementing IP phones. The g-series provides the capability necessary to support this requirement through ACL-Based Mirroring, MAC filter-Based Mirroring, and VLAN-Based Mirroring.

Network managers can apply a “mirror ACL” on a port and mirror a traffic stream based on IP source/destination address, TCP/UDP source/destination ports, and IP protocols such as ICMP, IGMP, TCP, and UDP. A MAC filter can be applied on a port and mirror a traffic stream based on a source/destination MAC address. VLAN-Based mirroring is another option for CALEA compliance. Many enterprises have service-specific VLANs, such as voice VLANs. With VLAN mirroring, all traffic on an entire VLAN within a switch can be mirrored or specific VLANs can be transferred to a remote server.

Threat detection and mitigation

Support for embedded, hardware-based sFlow traffic sampling extends Brocade IronShield 360 security shield to the network edge. This unique and powerful closed loop threat mitigation solution uses best-of-breed intrusion detection systems to inspect sFlow traffic samples for possible network attacks. In response to a detected attack, network management can apply a security policy to the compromised port. This automated threat detection and mitigation stops network attacks in real time, without human intervention. This advanced security capability provides a network-wide security umbrella without the added complexity and cost of ancillary sensors.

Advanced multicast features

g-series switches support a rich set of Layer 2 multicast snooping features that enable advanced multi-cast services delivery. Internet Group Management Protocol (IGMP) snooping for IGMP version 1, 2, and 3 is supported. Support for IGMPv3 source-based multicast snooping improves bandwidth utilization and security for multicast services. To enable multicast service delivery in IPv6 networks, the g-series supports Multicast Listener Discovery (MLD) version 1 and 2 snooping, the multicast protocols used in IPv6 environments.

Building resilient networks

Software features such as Virtual Switch Redundancy Protocol, Foundry's Metro Ring Protocol, Rapid Spanning Tree Protocol, protected link groups, and 802.3ad Link Aggregation, and trunk groups provide alternate paths for traffic in the event of a link failure. Sub-second fault detection utilizing Link Fault Signaling and Remote Fault Notification ensures rapid fault detection and recovery.

Enhanced Spanning Tree features such as Root Guard and BPDU Guard prevent rogue hijacking of Spanning Tree root and maintain a contention and loop free environment especially during dynamic network deployments. Additionally, the g-series supports Port Loop Detection on edge ports that do not have spanning tree enabled. This capability protects the network from broadcast storms and other anomalies that can result from Layer 1 or Layer 2 loopbacks on Ethernet cables or endpoints.

Base Layer 3 functionality enhances the capability of the g-series as an edge platform. Base Layer 3 allows enterprises to use simple Layer 3 features such as IPv4 static routes, virtual interfaces (VE), routing between directly connected subnets, RIPv1/v2 announce, VRRP, DHCP Relay, and routed interfaces. Network managers can remove complexity from an end- to-end Layer 3 network design and eliminate the cost required for a full Layer 3 edge switch.

Fault detection

The g-series switches support logical fault detection through software features such as Link Fault Signaling (LFS), Remote Fault Notification (RFN), Protected Link Groups, and Unidirectional Link Detection (UDLD).

- Link Fault Signaling (LFS) is a physical layer protocol that enables communication on a link between two 10-GbE switches. When configured on a 10-GbE port, the port can detect and report fault conditions on transmit and receive ports.
- Remote Fault Notification (RFN) enabled on 1Gb transmit ports notifies the remote port whenever the fiber cable is either physically disconnected or has failed. When this occurs the device disables the link and turns OFF both LEDs associated with the ports.
- Protected Link Groups minimize disruption to the network by protecting critical links from loss of data and power. In a protected link group, one port in the group acts as the primary or active link, and the other ports act as secondary or standby links. The active link carries the traffic. If the active link goes down, one of the standby links takes over.
- UDLD monitors a link between two g-series switches and brings the ports on both ends of the link down if the link goes down at any point between the two devices.

In addition, the g-series supports stability features such as Port Flap Dampening, single link LACP, and Port Loop Detection. Port Flap Dampening increases the resilience and availability of the network by limiting the number of port state transitions on an interface. This reduces the protocol overhead and network inefficiencies caused by frequent state transitions occurring on misbehaving ports.

Single link LACP provides a fast detection scheme for unidirectional or bi-directional faults. This standards based solution works with other switch vendors. The Port Loop Detection feature enables network managers to detect and prevent Layer 2 loops without using STP. Enterprises that do not enable a Layer 2 Protocol, such as STP to detect physical loops at the edge, can use Port Loop detection. Port Loop detection can be used to detect loops occurring on a port and within an entire network.

Physical and thermal parameters

The physical and thermal parameters are shown in Table 1-35.

Table 1-35 g-series physical and thermal parameters

Component	IBM Ethernet Switch B48G	IBM Ethernet Switch B50G
Chassis type	Fixed form factor	Fixed form factor
H/W/D (cm)	6.68 x 44.45 x 49.78	6.68 x 44.45 x 49.78
Rack units (RUs)	1.5	1.5
Max. weight	11.36 kg	11.36 kg
Max. power draw	1200 W	1200 W
Op. temperature (°C)	0 - 40 °C	0 - 40 °C
Heat emission (BTU/hr)	4094	4094
Airflow	Front-to-back	Front-to-back
Number of fans	2	2

All the fans are not swappable and have fixed speeds.

Power parameters

All g-series models provide redundant and removable power supplies with AC power options. Power supplies can be exchanged between B48G and B50G models.

The g-series models provide Power over Ethernet (PoE) option.

Both power supplies provide power for the system and PoE ports.

Power supplies

The power supplies are auto-sensing and auto-switching, and provide 600 watts of total output power, including +12VDC @ 10A to the system and -48VDC@ 10A for Power over Ethernet applications. The power supplies provide 100-240 VAC input, 50-60Hz @ 8A to 3.2A. All are hot swappable and can removed and replaced without powering down the system.

Power parameters for PoE

The system power parameters are shown in Table 1-36.

Table 1-36 g-series SYS power parameters

Component	IBM Ethernet Switch B48G	IBM Ethernet Switch B50G
Power supplies	100 - 240 VAC / 50 - 60 Hz	100 - 240 VAC / 50 - 60 Hz
Number of power supply bays	2	2
Number of power supply bays required for fully loaded chassis	1	1

The PoE port density is shown in Table 1-37.

Table 1-37 s-series PoE power parameters

Component	IBM Ethernet Switch B48G	IBM Ethernet Switch B50G
10/100/1000 Mbps PoE density with 15.4W each	48 (with 2 Power Supplies)	48 (with 2 power supplies)
10/100/1000 Mbps PoE density with 10W each	48 (with 1 Power Supply)	48 (with 1 power supply)

Power specifications for PoE

The implementation of the 802.3af standard limits power to 15.4W (44V to 57V) from the power sourcing device. This limit complies with safety standards and existing wiring limitations. Though limited by the 802.3af standard, 15.4 watts of power is ample, as most powered devices consume an average of 5 to 12 watts of power. IP phones, wireless LAN access points, and network surveillance cameras each consume an average of 3.5 to 9 watts of power.

Foundry 48-volt power supplies provide power to the PoE daughter card, and ultimately to PoE power-consuming devices. The number of PoE power-consuming devices that one 48-volt power supply can support depends on the number of watts required by each device. Each 48-volt power supply provides 480 watts of power for PoE, and each PoE port supports a maximum of 15.4 watts of power per PoE power-consuming device. For example, if each PoE power-consuming device attached to the g-series consumes 12 watts of power, one 48-volt supply will power up to 40 PoE ports. You can install a second power supply for additional PoE power.

Note: If your g-series device has 48 ports and only one power supply, and each PoE enabled port needs 15.4 watts, then a maximum of 31 ports can supply power to connected devices.

Ports, memory, and performance

The g-series models provide a *Store & Forward* switching engine.

All g-series models have a maximum of 256 MB RAM.

The number of ports and performance metrics are shown in Table 1-38.

Table 1-38 Ports and performance metrics

Component	IBM Ethernet Switch B48G	IBM Ethernet Switch B50G
10/100/1000 Mbps RJ45 Ports per System	44 plus 4 combo ports	44 plus 4 combo ports
100/1000 Mbps SFP Ports per System	4 combo ports	4 combo ports
10 GbE ports per system	2	2
Data switching capacity	136 Gbps	136 Gbps
Packet routing capacity	101 Mpps	101 Mpps

Interface types

Following are the available interface types:

- 10/100/1000 Mbps Ethernet port with RJ45 connector
- 100/1000 Mbps Ethernet port with SFP connector
- 10 Gbps Ethernet port with XFP connector
- 10 Gbps Ethernet port with CX4 connector

Transceivers

Table 1-39 and Table 1-40 show the available transceivers that can be used.

Table 1-39 Transceivers for 100/1000 Mbps Ethernet ports

Type	Connector	Speed	Distance
1000BASE-T SFP Copper	RJ-45	1Gbps	Up to 100 m with CAT5 or higher
1000BASE-SX 850 nm SFP optics	LC	1 Gbps	Up to 550 m over multi-mode fiber
1000BASE-LX 1310 nm SFP optics	LC	1 Gbps	Up to 10 km over single-mode fiber
1000BASE-LHA 1550 nm SFP optics	LC	1 Gbps	Up to 70 km over single-mode fiber
100BASE-FX 1310 nm SFP optics	LC	100 Mbps	Up to 2 km over multi-mode fiber

Table 1-40 Transceivers for 10 Gbps Ethernet ports

Type	Connector	Speed	Distance
10GBASE-SR 850 nm XFP optics	LC	10 Gbps	Up to 300 m over multi-mode fiber
10GBASE-LR 1310 nm XFP optics	LC	10 Gbps	Up to 10 km over single-mode fiber
10GBASE-ER 1550 nm XFP optics	LC	10 Gbps	Up to 40 km over single-mode fiber
10GBASE-CX4 XFP copper	LC	10 Gbps	Up to 15 m over CX4 grade copper

Optional features

The g-series is capable of providing Layer 3 functions. Following are the optional features:

- Edge Layer 3 Premium Activation:

  Enables RIPv1/v2, OSPFv2

Services, protocols, and standards

IBM g-series Ethernet Switches support various services, protocols, and standards.

The following Layer 2 protocols are supported:

- Protected Link Groups
- Link Aggregation (IEEE 802.3ad, LACP)
- UDLD
- STP/RSTP/MSTP
- Root Guard
- BPDU Guard
- Up to 16000 MAC addresses (valid also for 8 unit stack)
- Up to 4096 VLANs
- Up to 253 STPs
- Up to 8 ports per trunk, up to 25 trunk groups

The following Layer 2 Metro features are supported:

- VLAN Stacking
- Metro Ring Protocol (MRP I)
- Virtual Switch Redundancy Protocol
- Topology Groups
- Super Aggregated VLANs (SAV)

The following Base Layer 3 features are supported:

- Virtual Interfaces (VE)
- Routed Interfaces
- IPv4 Static Routes
- Routing between directly connected subnets
- RIP v1/v2 Announce
- Virtual Route Redundancy Protocol

The following Quality of Service features are supported:

- MAC Address Mapping to Priority Queue
- ACL Mapping to Priority Queue
- ACL Mapping to ToS/DSCP
- Honoring DSCP and 802.1p
- ACL Mapping and Marking of ToS/DSCP
- DiffServ Support
- Classifying and Limiting Flows based on TCP flags
- DHCP Relay
- QoS Queue Management Using Weighted
- Round Robin (WRR), Strict Priority (SP), and a combination of WRR and SP

The following traffic management features are supported:

- Inbound rate limiting per port
- ACL-based inbound rate limiting and traffic policies
- outbound rate limiting per port and per queue
- Broadcast, multicast, and unknown unicast

The whole list of supported standards and RFC compliance can be found at:

http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/

2

Command Line Interface and Element Manager basics

In this chapter we introduce you to the basic user interface for the network products covered in this book. We discuss both the Command Line Interface (CLI) and the Element Manager (the web interface).

2.1 Basic CLI access

There are three main ways to access the CLI, by using the serial port, a telnet connection, and an SSH connection.

Although all configuration operations are possible through telnet or SSH, it is always useful to have serial port access available as a backup, in case of a catastrophic problem with the configuration.

2.1.1 Serial port access

Unlike many brands of networking equipment, the serial console for these units is a simple straight-through DB-9 cable. No RJ-45 or null-modem adapter is necessary. Configure your terminal program as follows:

- Type: VT-100
- Speed: 9600 baud
- Data bits: 8
- Parity: None
- Stop bits: 1
- Flow Control: None

> **Note:** Users with access to the serial port can perform any configuration operation, including the bypass of password protection. Therefore, it is vital that physical access to the serial port be restricted.

On models with multiple management modules, connect to the active module.

2.1.2 Telnet access

Telnet is used by many users because it is simple to implement, and requires no special operations on the client.

The Telnet server might not be enabled by default on some models; to ensure that it is enabled, enter the **`telnet server`** command to your configuration. (See 2.2.3, “CONFIG level commands” on page 83 for information about adding commands to a configuration).

By default, telnet is available through any IP address defined on the unit. Most users choose to assign an IP address to the management port of the network device and use that address for telnet operations.

Important: All users are strongly encouraged to restrict access to all management access methods by the use of Access Control Lists. Allowing any user with network access to the product to attempt login is extremely insecure.

2.1.3 Secure Shell (SSH) access

Secure Shell (SSH) provides the same access as telnet, but it is far more secure. SSH sessions cannot be “sniffed” and decoded while the data is in transit. Access through this method requires users that want to login have a file called a “private key” loaded on their administration workstation. A SSH configuration also enables the use of Secure Copy (SCP). This allows the transfer of files to and from the product in a secure fashion.

For more information about configuring the products for Secure Shell, see Chapter 13, “Manageability and monitoring” on page 361.

2.2 Command Line Interface basics

The Command Line Interface (CLI) for the products is very similar to the CLI for many other brands of network equipment, so administrators with previous networking experience will find it very familiar.

2.2.1 CLI prompt

Immediately after logging in, you are presented with the “User” prompt, for example:

```
telnet@m_Series>
```

The beginning of the prompt `telnet` informs you that you are logging into a Telnet session; this might also say “SSH” if you were logged in that way. Users using the console cable will not be notified. The `m_Series` is the host name that we have assigned to the switch. (The host name can be set with the **hostname** ***hostname*** configuration command).

For (g-series only): If the product that you are using is running switching code and there is no host name set, there will be Switch after the default host name (for example, FGS648P-STK Switch>). Products running routing code will have Router instead of Switch.

The > at the end of the prompt indicates that you are running in “User” mode. In this mode, little can be done except to run **show** commands, which are used to display information about the status of the product, the configuration, and so on.

For most operations, administrators will immediately go into “Enable” mode, which allows you to actually configure and maintain the product. To enter this mode, simply enter the **enable** command. (This command can be restricted only to authorized users, which is why it is not run by default upon login.)

After you enter Enable mode, the prompt will have a # at the end, instead of a >, for example:

```
telnet@m_Series#
```

From Enable mode, you can perform any operation possible on the switch, so commands here must be typed with care. Most (but not all) commands take effect immediately, so an error can lead to catastrophic results.

From Enable mode, the commands are either “EXEC” level, or “CONFIG” level.

2.2.2 EXEC level commands

EXEC commands are used for directly changing the status of the product immediately. For instance, there are commands to clear logs, tables, or counters, copy firmware, reset the product, and so on.

The show commands

The most common set of commands you will run in EXEC mode are probably the **show** commands. These commands display the current status or configuration of the various parts of the firmware the runs on the product.

We introduce many of the various **show** commands throughout this book. One of the most useful **show** commands is **show running-config**. This command outputs the current running configuration to the terminal. If you capture the output of this command, you can keep it as a reference to how your product is configured.

One feature of note with the **show** commands is that they can be run at any time, including while the product is in the CONFIG Level. This is different from some other vendors of network equipment, which require you to exit the CONFIG level before running **show** commands.

Terminal length (c-series and m-series only)

Many **show** commands have lengthy output. By default, the terminal session will pause every 24 lines to let you examine the result. However, if you are capturing the output of a lengthy command, or prefer to use the buffer in your terminal program, this can be inconvenient. To turn off this pause, enter the **terminal length 0** command at the EXEC prompt, and/or add it to your configuration.

2.2.3 CONFIG level commands

CONFIG commands are for all configuration operations. Most, but not all, of these commands take effect immediately and are added to a configuration file, the *running-config*. We introduce these commands throughout the book as we cover the various operations that you can perform on these products.

Using configuration commands

To enter CONFIG Level, simply enter the **configure terminal** command at the Enable prompt. Your prompt will change to have (config) before the #, for example:

```
telnet@m_series(config)#
```

As mentioned before, you can display the complete current configuration at any time by running the **show running-config** command.

Important: Configuration changes are not saved persistently or across reboots by default. Run **write memory** on a routine basis to commit configuration changes to flash into the *startup-config*.

After you have made configuration changes to the product, the changes are *not* automatically saved to flash memory. Unless you save them to flash, they will *not* be persistent across reboots of the product. For this reason, make a habit of running the **write memory** command on a regular basis to commit the current running configuration to flash. Note that this command is different from the command used by some other vendors of network equipment.

Note: Unless otherwise noted, all commands in this book are run from the CONFIG level, instead of the EXEC level.

For information about loading a configuration from a file, see 14.5, “Configuration maintenance” on page 377.

Removing configuration commands

Any line of a configuration can be removed by putting a **`no`** at the beginning of a command. For instance, if you have assigned an IP address to an interface with IP address 1.2.3.4/24, you can remove the IP address with the **`no ip address 1.2.3.4/24`** command. You must enter the complete command; even if an interface only has a single IP address; the **`no ip address`** command will have no effect.

2.2.4 CLI navigation

The CLI has many features that make using it more simple.

Command editing

From a command prompt, you can use the left and right arrow keys to navigate within a command.

Repetition

If you want to repeat an earlier command, simply press the up arrow on your keyboard to scroll through a buffer of previously run commands. Press the Enter key when the command you want to run is displayed. You can also use the left and right arrow keys to edit the command before running it.

Prompting

At any time, you can press the Tab key on your keyboard to list the current options available for the command you are running. This can also be accomplished by putting a `?` at the end of the current command.

Abbreviation

You can abbreviate commands instead of typing out the entire command. For lengthy commands, this can be a great time-saver for experienced users. For instance, the **`show running-configuration`** command can be entered as simply **`sh ru`**. You only need enter as many letters as are necessary to make the command unambiguous. If you enter insufficient letters, the CLI will return an error message. For clarity, we always use the complete commands within this book.

Changing levels

To move up one level, for instance, from CONFIG to EXEC, simply type the **`exit`** command. To go back to EXEC, no matter how deep you are in CONFIG, you can also press ctrl-z.

Output filtering

When running `show` commands that produce lengthy repetitive, output, you can filter the output so only lines that do or do not contain certain text are displayed. To do this, append a `| include texttoshow` or `| exclude texttonotshow`. (Note that the vertical character is a “pipe”, not a lower-case L, upper-case i, or a number 1. The pipe symbol is created by shift-backslash on U.S. keyboards; it might vary on non-U.S. keyboards).

The text that you filter with `include` or `exclude` commands is case-sensitive.

The products have a full regular expression engine built-in for advanced users that want to filter the output that way. Syntax for regular expressions is beyond the scope of this book.

2.3 CONFIG level

Some config commands within the products are executed right from the (config) prompt. These are commands that usually effect the product as a whole; these are called “Global” config commands, such as the `hostname` ***`yourhostnamehere`*** command. Other commands are executed from a deeper level than the (config) prompt. Different levels require different commands to get in that level. For instance, to enter the ospf level, you invoke the config `router ospf` command. (This command also adds router ospf to your configuration, if it is not already there). This might be referred to as the “OSPF Config level.” To enter interface-level commands, you can use a command like `interface ethernet 1`.

To explicitly move back up a level, enter the `exit` command. However, you can execute global-level commands, or change contexts, from any level.

Examples of navigating up and down levels are given in Example 2-1. In the example, we also show the practice of creating configuration entries during the context switch.

Example 2-1 Config level navigation (truncated for brevity)

```
telnet@m_Series#show running-config
Current configuration:
!
interface management 1
 ip address 10.64.210.181/20
 enable
end
```

```
telnet@m_Series#config
telnet@m_Series(config)#
telnet@m_Series(config)#router ospf
telnet@m_Series(config-ospf-router)#
telnet@m_Series(config-ospf-router)#area 0
telnet@m_Series(config-ospf-router)#
telnet@m_Series(config-ospf-router)#exit
telnet@m_Series(config)#
telnet@m_Series(config)#interface ethernet 1/1
telnet@m_Series(config-if-e10000-1/1)#
telnet@m_Series(config-if-e10000-1/1)#exit
telnet@m_Series(config)#
telnet@m_Series(config)#vlan 300
telnet@m_Series(config-vlan-300)#
telnet@m_Series(config-vlan-300)#sntp server 1.2.3.4 [Here we execute a
global-level config command from a config context. Since the command
does not switch levels, we stay in the vlan 300 level.
telnet@m_Series(config-vlan-300)#exit
telnet@m_Series(config)#vlan 45
telnet@m_Series(config-vlan-45)#
telnet@m_Series(config-vlan-45)#router ospf [here we switch from
vlan-45 context to router ospf level without an “exit”]
telnet@m_Series(config-ospf-router)#
telnet@m_Series(config-ospf-router)#show running-config [here we run a
show command from a config level]

Current configuration:
!
vlan 45
vlan 300
sntp server 1.2.3.4
router ospf
 area 0
interface management 1
 ip address 10.64.210.181/20
 enable
end
```

Notice that switching the interface context also created those commands within the configuration, even if we did not actually configure anything under that context.

2.4 Specifying interfaces in the CLI

There are special circumstances in the CLI when it comes to abbreviating interface names.

2.4.1 Interface abbreviating: s-series and m-series

There are multiple ways to specify an interface name within the CLI. The two ways of specifying the Ethernet interface slot 1, port 1, are: **`e1/1`** or **`ethernet 1/1`**. (SONET interfaces on the m-series use a “p” instead of an “e”.) We can see the two methods in Example 2-2.

Example 2-2 Interface abbreviating

```
telnet@m_Series(config)#
telnet@m_Series(config)#interface e1/1
telnet@m_Series(config-if-e10000-1/1)#
telnet@m_Series(config-if-e10000-1/1)#exit
telnet@m_Series(config)#interface ethernet 1/1
telnet@m_Series(config-if-e10000-1/1)#
telnet@m_Series(config-if-e10000-1/1)#
telnet@m_Series(config-if-e10000-1/1)#exit
```

The management port is **`management 1`**. or **`m1`**.

2.4.2 Interface abbreviating: c-series

For the c-series, the following considerations apply.

IBM Ethernet Switch B24C and B50C

In the case of the c-series, there are two models that either have built in XFP ports or an optional slot for a 2-port 10 GbE XFP module. On these models, the GigE ports are treated as “slot 1”, and the XFP ports are treated as “slot 2”. So, to configure the 24th ethernet port, we can use **`interface ethernet 1/24`** or **`interface e1/24`**. To configure the 1st XFP port, we can use interface **`ethernet 2/1`** or **`interface e2/1`**.

IBM Ethernet Switch B48C

For the 48-port c-series switch without 10 GbE uplinks, the four GigE “combo” ports on the left-hand side of the product override the respective UTP GigE port if they are active. In either case, the ports are considered part of “slot 1”, so, for instance, the 23rd port is addressed as **`ethernet 1/23`** or **`e1/23`**.

The management port is **`management 1`** or **`m1`**.

2.4.3 Interface abbreviating: g-series

Even though the IBM Ethernet Switch B48G cannot be stacked, the g-series uses a `stacknum/slot/port` nomenclature. Because the B48G cannot be stacked, the `stacknum` will always be `0`.

The UTP ports are “slot 1”, and the 10 GbE ports (if present) are “slot 2.”

For example, to configure the 2nd UTP port on the 2nd switch in the stack, we can use **`interface ethernet 2/1/2`** or **`e2/1/2`**. To configure the 1st 10 GbE port on the 4th switch in the stack, we can use **`interface ethernet 4/2/1`** or **`interface e4/2/1`**.

2.4.4 Interface ranges

Many commands are run on multiple interfaces, such as adding ports to a VLAN, or configuring many ports at once. For these commands, we can use a range. These take slightly different formats, depending on how you abbreviate the port name. To perform an operation on Ethernet slot 1, ports 1 through 4, use either **`ethernet 1/1 to 1/4`** or **`e1/1 to 1/4`**. Note that the second port number in the range does not have the “e” at the beginning; it is implied.

We can see the different methods in use in Example 2-3.

Example 2-3 Using interface ranges

```
telnet@m_Series(config)#vlan 300
telnet@m_Series(config-vlan-300)#tagged ethernet 5/1
telnet@m_Series(config-vlan-300)#tagged ethernet 5/2 to 5/4
telnet@m_Series(config-vlan-300)#tagged e1/5 to 1/7
telnet@m_Series(config-vlan-300)#sh vlan 300

PORT-VLAN 300, Name [None], Priority Level0, ARP Insp OFF
DHCP Snoop OFF, IP Source Insp OFF, Priority Force 0
Topo HW idx    : 65535    Topo SW idx: 257    Topo next vlan: 0
L2 protocols   : NONE
Tagged Ports   : ethe 5/1 to 5/7
```

```
telnet@m_Series(config-vlan-300)#
telnet@m_Series(config-vlan-300)#interface e5/1 to 5/5
telnet@m_Series(config-mif-5/1-5/5)#link-error-disable 5 10 15
Link-Error-Disable configured for port 5/1
Link-Error-Disable configured for port 5/2
Link-Error-Disable configured for port 5/3
Link-Error-Disable configured for port 5/4
Link-Error-Disable configured for port 5/5
telnet@m_Series(config-mif-5/1-5/5)#
telnet@m_Series(config-mif-5/1-5/5)#show running-config
[truncated for brevity]
interface ethernet 5/1
 link-error-disable 5 10 15
!
interface ethernet 5/2
 link-error-disable 5 10 15
!
interface ethernet 5/3
 link-error-disable 5 10 15
!
interface ethernet 5/4
 link-error-disable 5 10 15
!
interface ethernet 5/5
 link-error-disable 5 10 15
!
telnet@m_Series(config-mif-5/1-5/5)#
```

2.4.5 Specifying interfaces in stacked devices

The format for port number in stacked devices takes the format `e2/1/3`, where 2 is the unit number, 1 is the “slot” number (1 for the 1Gb ports, 2 for the 10 Gb ports) 3 for the port number.

2.5 Specifying subnet masks in the CLI

There are two methods of specifying an IP subnet mask, when one is needed: **1.2.3.4 255.255.255.0** or **1.2.3.4/24**; both formats have an equal result. The CLI automatically converts the **255.255.255.0** into the **/24**, which is what you will see in the configuration. We can see the two methods in use in Example 2-4.

Example 2-4 Subnet mask specification

```
telnet@m_Series(config)#interface e5/1
telnet@m_Series(config-if-e1000-5/1)#ip address 1.2.3.4/24
telnet@m_Series(config-if-e1000-5/1)#ip address 3.4.5.6 255.255.255.0
telnet@m_Series(config-if-e1000-5/1)#show running-config
[truncated]
interface ethernet 5/1
 ip address 1.2.3.4/24
 ip address 3.4.5.6/24
end
telnet@m_Series(config-if-e1000-5/1)#
```

2.6 The Element Manager (web GUI)

The Element Manager (web GUI) is not supported on the x-series.

Some of the products have a web GUI available, called the Element Manager. Many, if not all, configuration options can be executed using this interface.

Not all of the products have the web interface enabled by default. To enable it, use the **web-management** command in the configuration.

To configure the product using the web GUI, enter the username **set**. For the password, use the name of an SNMP community that you have configured for read-write access. (For details, see the section 13.1, “SNMP” on page 362).

As you might expect, this is not very secure. User access can be restricted by configuring specific Web-management users, as opposed to utilizing the SNMP community string as the password.

If you do choose to use the web GUI, you can use SSL for slightly better security. web access, such as telnet or SSH, can be restricted by Access Control Lists.

In this book we concentrate on using the CLI for configuration, because that is the method used by most network administrators.

3

Initial product setup

In this chapter, we detail the basic steps to follow to set up your new router or switch.

We are assuming that the product has already been physically installed according to the instructions in the *Hardware Installation Guide*.

3.1 Setting basic system identification

Here we describe the basic steps to take. Preferably, first read through Chapter 2, “Command Line Interface and Element Manager basics” on page 79 to familiarize yourself with the CLI.

1. Using the serial cable, configure the basic identification information for the product. Serial terminal connectivity can be accomplished using a straight-through DB9 serial cable with the following Terminal settings:
   - Type: VT-100
   - Speed: 9600 baud
   - Data bits: 8
   - Parity: None
   - Stop bits: 1
   - Flow Control: None
2. Enter the Global CONFIG level by typing the following statements:

   ```
   NetIron MLX-16>enable
   No password has been assigned yet...
   NetIron MLX-16#
   NetIron MLX-16#configure terminal
   NetIron MLX-16(config)#
   ```
3. From the Global CONFIG Level, enter the **`hostname`** ***`yourhostname`*** command to give the product a name that you will easily recognize.
4. Set the administrative contact and location for the product. This information is used by most SNMP management applications. Use the following commands:

   ```
   snmp-server contact Peter Mescher
   snmp-server location Los Gatos CA
   ```

3.2 Setting a management IP address

Most users will not use the serial port for all administration tasks. To enable configuration by a network connection, an IP address must be configured for the system.

Switches can either be managed through the out-of-band management port if available or in-band through a regular interface.

3.2.1 IBM g-series and x-series

The IBM g-series and x-series do not have a dedicated out-of-band management port. These switches can be managed in-band through any of the ports on the switch.

One way to manage the switch is to dedicate one physical port to be the management port connected to the management network. Access Control Lists can be used to restrict remote management through this one port for additional security.

For switches running Layer 2 images, a global IP address can be assigned to the switch using the `ip addr 1.2.3.4/24` command. The "/24" in the command indicates the length of the subnet mask; this can also be addressed in the traditional form of 255.255.255.0, for example, `ip addr 1.2.3.4 255.255.255.0`. To set up a default gateway for the switch, use the `ip default-gateway 1.2.3.1` command.

For switches running Layer 3 images, the switch will listen to any of the IP addresses configured on the interfaces. It is best to configure an IP address for a loopback interface to use as the management IP address because these interfaces never go down. To do so, enter the Interface CONFIG level for a loopback port using the `int loopback 1` command and assign an IP address using `ip addr 1.2.3.4/24`.

3.2.2 IBM m-series, r-series, s-series, and c-series

The IBM m-series, r-series, s-series, and c-series devices have a dedicated out-of-band management port available. On the chassis-based models with multiple management modules (each of which has a management port), the active module "owns" the IP address.

To assign a management IP address to these models, configure the IP address on the management 1 interface. To do this, use the `interface management 1` configuration command followed by `ip address 1.2.3.4/24`.

The management port and the ports used for network traffic on the interface modules have different data paths. However, the management port leverages the routing table from the network ports. If you want to configure a default gateway for the management port, create a static IP route in the routing table using the `ip route 1.2.3.4/24` command from the Global CONFIG level.

3.3 Enabling/disabling management methods

Before connecting the product to the network, you need to enable or disable the management methods as necessary. For information about securing the management methods in order to restrict access, see Chapter 13, "Manageability and monitoring" on page 361.

It is a good practice to explicitly enable or disable each management method, even if it is enabled or disabled by default on the particular product you are using.

3.3.1 Telnet

To enable/disable the telnet server, use the `[no] telnet server` command. (It is enabled by default on g-series and s-series products).

Attention: Avoid using telnet on any of the products because it is insecure; use SSH instead if possible (see the following sections).

3.3.2 SSH/SCP

SSH/SCP is automatically enabled when the product has a public/private key pair available. To generate a key pair, use the `crypto key generate` command.

3.3.3 Web server

The Web Management feature is not available on the x-series.

There are two web servers available, one that uses http, and one that uses https/SSL. To enable or disable the web server, use the `[no] web-management` command. The command for the https/SSL server is `web-management https`. (It is not enabled by default on any platform).

For the https/SSL server to function, you must generate or load an SSL certificate. To generate one (which will show up in your browser as unsigned), use the command `crypto-ssl certificate generate`.

An official certificate from a certificate authority can be loaded. For instructions on doing so, see "Configuring SSL Security for the Web Management Interface" in the *Configuration Guide*.

3.3.4 SNMP

To enable/disable the SNMP server, use the **`[no] snmp-server`** command.

3.4 Basic password configuration

The passwords discussed next are only the most basic ones and do not require a user name. Most users will want to set up individual user accounts or more sophisticated authentication methods, which we cover in Chapter 13, “Manageability and monitoring” on page 361.

3.4.1 Basic access

For accessing the product by the serial port, no password is required. While it is possible to set passwords for a serial user to elevate their access to the Privileged EXEC and CONFIG levels, any user with serial access can login to the User-mode prompt.

3.4.2 Telnet password

To set a password for telnet access, use the **`enable telnet password`** ***`yourpasswordhere`*** command. This password, like all of the global passwords, has no username.

This password will get the user to the basic User mode only. Further passwords are necessary to elevate access.

3.4.3 Enable and Config passwords

There are three types of passwords for Enable mode (also called Privileged EXEC mode), Read Only, Port Configuration, and Super User.

Read Only password

As the name implies, the Read Only password cannot be used to modify the configuration of the product. However, this password can perform any operation in EXEC level (either User or Privileged), so assign this level with care.

Set this password with **`enable read-only-password`** ***`yourpasswordhere`***.

Port Configuration password

The Port Configuration password can perform any read operations, but can only configure interface parameters. This level of access cannot change any other configuration options.

Set this password with **`enable port-config-password`** ***`yourpasswordhere`***.

Super User password

The Super User password can enable the user to perform any command on the unit.

Set this password with **`enable super-user-password`** ***`yourpasswordhere`***.

3.4.4 Password recovery

If you lose the super-user password, it can be easily reset with access to the serial port. To do so, initiate a reload. (This can be done by powering off the product if you have no way to get into enable mode).

Press the “b” key as many times as needed until the `Monitor>` prompt appears. Next, enter the **`no password`** command. This will erase the password. To then boot the system, enter the **`boot system flash primary`** command. The system will boot with no Super User password. You can then assign a new one normally after the product boots.

> **Important:** Because the Super User password can be reset from the serial port and there is no way to block the process, it is vital to ensure that physical access to the serial port is tightly controlled.

3.5 Hardware verification

To ensure the basic function of the product, enter the **`show module`** command. If the status for any module indicates an error, contact IBM support.

3.6 Initial code load

As part of the setup process, verify that the product is running the level and version of code you want to run long-term. Preferably, run the latest level of code in your products if this is a new network. If this product will be integrated into an existing network, run a level similar to that of your other products.

If you determine that a code change is necessary, see the procedures in 14.8, “Code upgrade process” on page 381.

3.6.1 Version discovery

The version of code running on the device can be found by using the **`show version`** command. To determine the version of code stored in flash memory, use the **`show flash`** command. The output of these commands varies by model.

IronWare: s-series, g-series, x-series

The version of code can be found under the Primary and Secondary code listings. In the case of stacked units, the available code will be listed separately for each stack unit. Example output is listed in Example 3-1.

Example 3-1   g Series show flash output

```
gSeries_5(config)#show flash
Stack unit 1:
  Compressed Pri Code size = 3167396, Version 05.0.01aT7e1
(fgs05001a.bin)
  Compressed Sec Code size = 3462298, Version 05.0.01aT7e1
(fgl05001a.bin)
  Compressed BootROM Code size = 416213, Version 05.0.00T7e5
  Code Flash Free Space = 1048576
Stack unit 2:
  Compressed Pri Code size = 3167396, Version 05.0.01aT7e1
(fgs05001a.bin)
gSeries_5(config)#
  Compressed Sec Code size = 3462298, Version 05.0.01aT7e1
(fgl05001a.bin)
  Compressed BootROM Code size = 416213, Version 05.0.00T7e5
  Code Flash Free Space = 1048576
```

There are two different code images available for the g-series, Layer 2 and Base Layer 3. (The stackable B50G model only supports the Layer 2 image). The file name for the Layer 2 image will have an “s” as the third character; the Layer 3 images will have an “l” (a lower-case “L”) as the third character.

The **show flash** output is similar to that for the g-series. Both the Active and Standby management modules (if applicable) are listed as shown in Example 3-2. The BootROM might not have a version number even close to that of the Primary and Secondary code images; *this is not usually an error.*

Example 3-2 s-series show flash output

```
telnet@FastIron SX 1600 Router#show flash
Active Management Module (Slot 10):
Compressed Pri Code size = 3319603, Version 05.0.00aT3e3
(sxr05000a.bin)
Compressed Sec Code size = 2502563, Version 05.0.00aT3e1
(sxs05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9699328
Standby Management Module (Slot 11):
Compressed Pri Code size = 3319603, Version 05.0.00aT3e3
(sxr05000a.bin)
Compressed Sec Code size = 2502563, Version 05.0.00aT3e1
(sxs05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9699328
telnet@FastIron SX 1600 Router#
```

There are three different loads available for the s-series, Layer 2, Base Layer 3, and Full Layer 3. The filename for the Layer 2 image will have an “s” as the third character, Base Layer 3 images will have an “r”, and Full Layer 3 will have an “l” (lower-case “L”).

For the x-series, only a Layer 2 image is available.

Multi-Service IronWare: c-series

The **show flash** output for a c-series shows multiple running code files. The version numbers that we are interested in are the primary and secondary Multi-Service IronWare images. The Monitor image and Boot image must be identical with each other.

An upgrade to the code might require an upgrade to the Monitor and Boot images. If an upgrade to the Monitor or Boot image is required, it will be noted in the release notes in a section titled something similar to “*Software Image files for Multi-Service IronWare R<code version>*”.

In Example 3-3, the version number we are concerned with is 3.8.0c. The T183 or T185 is irrelevant.

Example 3-3 c-series show flash output

```
telnet@NetIron CES 2048C>show flash
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Code Flash - Type MT28F128J3, Size 32 MB
  o IronWare Image (Primary)
    Version 3.8.0cT183, Size 9129481 bytes, Check Sum ee86
    Compiled on Mar 25 2009 at 21:19:22 labeled as ce03800c
  o IronWare Image (Secondary)
    Version 3.8.0cT183, Size 9129481 bytes, Check Sum ee86
    Compiled on Mar 25 2009 at 21:19:22 labeled as ce03800c
  o Monitor Image
    Version 3.8.0cT185, Size 342809 bytes, Check Sum 1cf4
    Compiled on Mar 25 2009 at 20:01:26 labeled as ceb03800c
  o Startup Configuration
    Size 379 bytes, Check Sum 3870
    Modified on 02:11:47 GMT+00 Fri Jan 12 1900

Boot Flash - Type AM29LV040B, Size 512 KB
  o Boot Image
    Version 3.8.0cT185, Size 342809 bytes, Check Sum 1cf4
    Compiled on Mar 25 2009 at 20:01:26 labeled as ceb03800c
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

telnet@NetIron CES 2048C>
```

Multi-Service IronWare: m-series

The output of **show flash** for the m-series is a little longer than for the other models, and determining if you are running the latest code is a little more complex. This is shown in Example 3-4.

Example 3-4 m Series show flash output (truncated)

```
telnet@m_Series>show flash
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Active Managment Module (Bottom Slot)
Code Flash - Type MT28F128J3, Size 32 MB
  o IronWare Image (Primary)
    Version 4.0.0dT163, Size 6625478 bytes, Check Sum 28f7
    Compiled on May 22 2009 at 23:03:58 labeled as xmr04000d
  o IronWare Image (Secondary)
    Version 4.0.0dT163, Size 6625478 bytes, Check Sum 28f7
    Compiled on May 22 2009 at 23:03:58 labeled as xmr04000d
```

```
  o LP Kernel Image (Monitor for LP Image Type 0)
    Version 3.5.0fT175, Size 387707 bytes, Check Sum 87ad
    Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
  o LP IronWare Image (Primary for LP Image Type 0)
    Version 4.0.0dT177, Size 4116213 bytes, Check Sum 4b29
    Compiled on May 22 2009 at 23:13:14 labeled as xmlp04000d
  o LP IronWare Image (Secondary for LP Image Type 0)
    Version 4.0.0dT177, Size 4116213 bytes, Check Sum 4b29
    Compiled on May 22 2009 at 23:13:14 labeled as xmlp04000d
  o Monitor Image
    Version 3.5.0fT165, Size 422466 bytes, Check Sum ab32
    Compiled on Oct 31 2008 at 12:04:30 labeled as xmb03500f
  o Startup Configuration
    Size 1636 bytes, Check Sum 6c40
    Modified on 15:33:54 GMT+00 Wed Jun 17 1903

Boot Flash - Type AM29LV040B, Size 512 KB
  o Boot Image
    Version 3.5.0T165, Size 424484 bytes, Check Sum b751
    Compiled on Jul 10 2007 at 19:13:56 labeled as xmprm03500
[truncated for brevity]
```

There is a lot of detail in Example 3-4 on page 99; all we need to be concerned with during the initial setup of a factory-fresh device is the IronWare image. If the IronWare version number (4.0.0d in our case) is not what you intend to run, you need to change it.

IronWare: r-series

The IronWare image for the management module is loaded with the IronWare image for the interface module in the Primary location on the code flash. The switch IronWare image for the management module is loaded with the IronWare image for the interface module in the Secondary location on the code flash. The monitor image for the management module is loaded with the monitor image for the interface module in the Monitor file on the code flash. The monitor image for the management module is also on the boot flash.

To display information concerning the contents of an IBM r-series switch, use the **show flash** command.

The other version numbers throughout the **show flash** will matter when we go over the instructions to upgrade code in 14.8, “Code upgrade process” on page 381.

3.6.2 Obtaining software

At the time of writing, the location to obtain code is:

http://www.ibm.com/systems/support/networking/

Select your product from the pull-down, “Downloads”.

That will take you to a special Brocade site that will contain only IBM-approved software releases for your product. The download site also contains manuals for the software.

3.7 DNS domain name and server

Some operations require communication with a DNS server. To configure this ability and specify one to four DNS resolvers, enter the **`ip dns server-address`** ***`1.1.1.1 [2.2.2.2 3.3.3.3 4.4.4.4]`*** configuration command.

To set the domain name the unit serves, use the **`ip dns domain-name`** ***`yourdomain.com`*** command.

3.8 NTP server

Most users will want to make sure that their log timestamps match up with the other devices in their environment. To do this, enter the **`sntp server`** ***`1.2.3.4`*** or **`sntp server`** ***`hostname`*** configuration command.

3.9 Remaining configuration

After doing the steps described so far, your remaining tasks will vary depending on your particular installation. Most users will perform tasks such as:

- Setting the initial SNMP configuration
- Configuring the physical attributes of the ports
- Configuring Users and Authentication
- Configuring VLANs
- Configuring Routing
- Choosing other features as desired, such as Power over Ethernet, IPv6, Stacking, and so on

4

Protocols and features

In this chapter we provide an extended listing of the software features and protocols supported on the IBM Data Center Networking b-type family of products. We cover the following products:

- IBM m-series Ethernet/IP Routers
- IBM r-series Ethernet Switches
- IBM s-series Ethernet Switches
- IBM c-series Ethernet Switches
- IBM x-series Ethernet Switches
- IBM g-series Ethernet Switches

We also supply a features matrix for these products.

4.1 IBM m-series Ethernet/IP Routers

The IBM m-series Ethernet/IP Routers are high-performing, multi-service IP/MPLS routers designed to address the most demanding performance, scalability, and security needs at the data center core, Internet/WAN border, and Enterprise campus backbone.

The m-series supports high-density, non-oversubscribed 4x10 GbE and 48x1 GbE Interface Modules, as well as Packet over SONET/SDH on available 8-port OC-48/12, and 2-port OC-192 Interface Modules.

The m-series is available in the following model configurations:

- IBM Ethernet Router B04M (4003-M04): A 4 RU chassis with 4 slots for interface modules, one dual-fan fan tray, and up to two management modules, three switch fabric modules, and three 1200 W power supplies.
- IBM Ethernet Router B08M (4003-M08): A 7 RU chassis with 8 slots for interface modules, one quad-fan fan tray, and up to two management modules, three switch fabric modules, and four 1200 W power supplies.
- IBM Ethernet Router B16M (4003-M16): A 14 RU chassis with 16 slots for interface modules, one front fan tray, two rear fan assemblies, and up to two management modules, four switch fabric modules, and eight 1200 W power supplies.
- IBM Ethernet router B32M (4003-M32): A 33 RU chassis with 32 slots for interface modules, two rear power fan assemblies, eight rear fan assemblies, and up two management modules, eight switch fabric modules, and eight 2400 W power supplies.

All m-series models run a common Operating System, Multi-Service IronWare. All software functionality is enabled in the base software and available across the product family.

4.1.1 Feature highlights for m-series

In this section we provide feature highlights of the m-series running Multi-Service IronWare R04.0.00f and higher:

- Wire-speed IPv4, IPv6, and MPLS routing featuring full Forwarding Information Base (FIB) programming in hardware
- Industry leading 320 Gbps link aggregation capability for aggregating up to 32 10GbE/OC-192 links
- Support for Packet over SONET/SDH

- Carrier-grade Quality of Service (QoS) for enabling converged multi-play networks
- Embedded, hardware-based sFlow traffic monitoring enabling network-wide accounting, utilization, reporting, capacity planning, intrusion detection, and more
- Classic Layer 2 support:
  - IEEE 802.1d (STP), 802.1w (RSTP), 802.1s (MSTP)
  - Single Spanning Tree Protocol (SSTP), PVST/PVST+, and Topology Groups, which enable either:
    - A single Spanning Tree instance for all VLANs
    - A separate Spanning Tree instance per VLAN
    - Sharing of a Spanning Tree instance across multiple VLANs per user configuration
  - IEEE 802.1q (VLAN tagging)
  - IEEE 802.1p (Class of Service Prioritization)
  - IEEE 802.1ad (Link Aggregation Control Protocol)
  - IEEE 802.1x (Port Security)
  - IGMPv1/v2/v3 & PIM-SM snooping
  - Metro Ring Protocol (MRP):
    - Alternative to Spanning Tree Protocol. Prevents Layer 2 loops
    - Provides fast, sub-second reconvergence in a Layer 2 ring topology
  - Virtual Switch Redundancy Protocol (VSRP):
    - Provides redundancy and sub-second failover in a Layer 2 mesh topology
    - Based on VRRPE, provides one or more backup switches for Layer 2 redundancy; the m-series also supports providing concurrent redundancy for Layer 3
- Advanced Layer 2 support: IEEE 802.1ag (Connectivity Fault Management)
- Layer 3:
  - Virtual Router Redundancy Protocol (VRRP) and VRRPE (Extended)
  - IPv4 Unicast support:
    - RIPv1/v2
    - OSPFv2
    - BGP-4
    - IS-IS

  - IPv4 Multicast support:
    - IGMPv1/v2/v3
    - PIM-DM
    - PIM-SM
    - PIM-SSM
    - MSDP
    - Anycast RP
  - IPv6 Unicast support:
    - RIPng
    - OSPFv3
    - BGP-4+
    - IS-IS for IPv6
  - IPv6 Multicast support:
    - PIM-SM
    - PIM-SSM
    - Anycast RP
    - MLDv1/v2
- Multi-VRF (Virtual Routing and Forwarding):

  VRF enables the creation of multiple routing instances on a single router, enhancing security
- Multiprotocol Label Switching (MPLS):
  - Label Distribution Protocol (LDP)
  - Fast Re-Route (Detour/Bypass)
  - Traffic Engineering: RSVP-TE, OSPF-TE, ISIS-TE
- Virtual Private Networks (VPN):
  - Virtual Private LAN Services (VPLS)
  - Virtual Leased Lines (VLL)
  - BGP/MPLS VPNs

4.1.2 Technical specifications for m-series

The m-series running Multi-Service IronWare R04.0.00f and higher conform to the summarized list shown in Table 4-1.

Table 4-1 m-series

IEEE compliance
► 802.3-2005 CSMA/CD Access Method and Physical Layer Specifications ► 802.3ab 1000BASE-T ► 802.3ae 10 Gigabit Ethernet ► 802.3x Flow Control ► 802.3ad Link Aggregation ► 802.1Q Virtual Bridged LANs ► 802.1D MAC Bridges ► 802.1w Rapid STP ► 802.1s Multiple Spanning Trees ► 802.1ad Provider Bridges; partial support: port-based and S-tagged service interface ► 802.1ag Connectivity Fault Management (CFM)

RFC compliance	
BGPv4	► RFC 4271 BGPv4 ► RFC 1745 OSPF Interactions ► RFC 1997 Communities & Attributes ► RFC 2439 Route Flap Dampening ► RFC 2796 Route Reflection ► RFC 1965 BGP4 Confederations ► RFC 2842 Capability Advertisement ► RFC 2918 Route Refresh Capability ► RFC 1269 Managed Objects for BGP ► RFC 2385 BGP Session Protection by TCP MD5 ► RFC 3682 Generalized TTL Security Mechanism, for eBGP Session Protection ► RFC 4273 BGP-4 MIB ► Draft-ieft-idr-restart Graceful Restart Mechanism for BGP
OSPF	► RFC 2328 OSPF v2 ► RFC 3101 OSPF NSSA ► RFC 1745 OSPF Interactions ► RFC 1765 OSPF Database Overflow ► RFC 1850 OSPF v2 MIB ► RFC 2370 OSPF Opaque LSA Option ► RFC 3630 TE Extensions to OSPF v2 ► RFC 3623 Graceful OSPF Restart
IS-IS	► RFC 1195 Routing in TCP/IP and Dual Environments ► RFC 1142 OSI IS-IS Intra-domain Routing Protocol ► RFC 2763 Dynamic Host Name Exchange ► RFC 2966 Domain-wide Prefix Distribution
RIP	► RFC 1058 RIP v1 ► RFC 1723 RIP v2 ► RFC 1812 RIP Requirements

RFC compliance	
IPv4 multicast	► RFC 1122 Host Extensions ► RFC 1112 IGMP ► RFC 2236 IGMP v2 ► RFC 3376 IGMP v3 ► RFC 3973 PIM-DM ► RFC 2362 PIM-SM ► RFC 2858 BGP-MP ► RFC 3618 MSDP ► RFC 3446 Anycast RP
General protocols	► RFC 791 IP ► RFC 792 ICMP ► RFC 793 TCP ► RFC 783 TFTP ► RFC 826 ARP ► RFC 768 UDP ► RFC 894 IP over Ethernet ► RFC 903 RARP ► RFC 906 TFTP Bootstrap ► RFC 1027 Proxy ARP ► RFC 951 BootP ► RFC 1122 Host Extensions for IP Multicasting ► RFC 1256 IRDP ► RFC 1519 CIDR ► RFC 1542 BootP Extensions ► RFC 1812 Requirements for IPv4 Routers ► RFC 1541 and 1542 DHCP ► RFC 2131 BootP/DHCP Helper ► RFC 3768 VRRP ► RFC 854 TELNET ► RFC 1591 DNS (client)
QoS	► RFC 2475 An Architecture for Differentiated Services ► RFC 3246 An Expedited Forwarding PHB ► RFC 2597 Assured Forwarding PHB Group ► RFC 2698 A Two Rate Three Color Marker
Management	► RFC 1354 IP Forwarding MIB ► RFC 2665 Ethernet Interface MIB ► RFC 1757 RMON Groups 1,2,3,9 ► RFC 2068 HTTP ► RFC 2030 SNTP ► RFC 2865 RADIUS ► RFC 3176 sFlow ► RFC 2863 Interfaces Group MIB

RFC compliance	
Other	► Draft-ieft-tcpm-tcpsecure TCP Security ► RFC 3704 Ingress Filtering for Multihomed Networks (uRPF) ► RFC 2784 Generic Routing Encapsulation (GRE) ► Draft-ieft-bfd-base Bidirectional Forwarding Detection (BFD) ► Draft-ieft-bfd-v4v6-1hop BFD for IPv4 and IPv6 (Single Hop); for OSPFv2, OSPFv3, IS-IS
IPv6 core	► RFC 2460 IPv6 Specification ► RFC 2461 IPv6 Neighbor Discovery ► RFC 2462 IPv6 Stateless Address Auto-Configuration ► RFC 4443 ICMPv6 ► RFC 4291 IPv6 Addressing Architecture ► RFC 3587 IPv6 Global Unicast Address Format ► RFC 2375 IPv6 Multicast Address Assignments ► RFC 2464 Transmission of IPv6 over Ethernet Networks ► RFC 2711 IPv6 Router Alert Option ► RFC 3596 DNS support
IPv6 routing	► RFC 2080 RIPng for IPv6 ► RFC 2740 OSPFv3 for IPv6 ► Draft-ieft-isis-ipv6 Routing IPv6 with IS-IS ► RFC 2545 Use of BGP-MP for IPv6
IPv6 multicast	► RFC 2710 Multicast Listener Discovery (MLD) for IPv6 ► RFC 3810 Multicast Listener Discovery Version 2 for IPv6 ► RFC 4604 IGMPv3 & MLDv2 for SSM ► Draft-ieft-ssm-arch SSM for IP ► RFC 2362 PIM-SM ► Draft-ieft-pim-sm-v2-new; partial support: SSM mode of operation
IPv6 transitioning	► RFC 2893 Transition Mechanisms for IPv6 Hosts and Routers ► RFC 3056 Connection of IPv6 Domains by IPv4 Clouds
MPLS	► RFC 3031 MPLS Architecture ► RFC 3032 MPLS Label Stack Encoding ► RFC 3036 LDP Specification ► RFC 2205 RSVP v1 Functional Specification ► RFC 2209 RSVP v1 Message Processing Rules ► RFC 3209 RSVP-TE ► RFC 3270 MPLS Support of Differentiated Services ► RFC 4090 Fast Reroute Extensions to RSVP-TE for LSP Tunnels; partial support: detour style ► RFC 3812 MPLS TE MIB ► Draft-ieft-bfd-mpls BFD for MPLS LSPs (RSVP-TE)

RFC compliance	
L3VPN	► RFC 2858 Multiprotocol Extensions for BGP-4 ► RFC 3107 Carrying Label Information in BGP-4 ► RFC 4364 BGP/MPLS IP VPNs ► Draft-ieft-idr-bgp-ext-communities BGP Extended Communities Attribute ► RFC 4576 Using LSA Options Bit to Prevent Looping in BGP/MPLS IP VPNs (DN Bit) ► RFC 4577 OSPF as the PE/CE Protocol in BGP/MPLS IP VPNs ► Draft-ieft-idr-route-filter Cooperative Route Filtering Capability for BGP-4 ► RFC 4382 MPLS/BGP Layer 3 VPN MIB
L2VPN and PWE3	► Draft-ieft-l2vpn-l2-framework Framework for Layer 2 Virtual Private Networks ► Draft-ieft-l2vpn-requirements Service Requirements for Layer 2 Provider Provisioned Virtual Private Networks ► RFC 4762 VPLS Using LDP Signaling ► Draft-ieft-pwe3-arch PWE3 Architecture ► RFC 4447 Pseudowire Setup and Maintenance using LDP ► RFC 4448 Encapsulation Methods for Transport of Ethernet over MPLS Networks ► Draft-ieft-pwe3-pw-tc-mib Definitions for Textual Conventions and OBJECT-IDENTITIES for Pseudo-Wires Management ► Draft-ieft-pwe3-pw-mib Pseudo Wire (PW) Management Information Base

Packet over SONET/SDH
► RFC 1661 The Point-to-Point Protocol (PPP) ► RFC 1662 PPP in HDLC-like Framing ► RFC 2615 PPP over SONET/SDH ► RFC 1332 Internet Protocol Control Protocol (IPCP) ► RFC 1377 The PPP OSI Network Layer Control Protocol (OSINLCP) ► RFC 2472 IPv6 over PPP ► RFC 3592 SONET/SDH Objects ► GR-253-CORE SONET Transport Systems: Common Generic Criteria ► G.707/Y.1322 Network Node Interface for SDH

MEF certification
► MEF 9 Certified—Abstract Test Suite for Ethernet Services at the UNI ► MEF 14 Certified—Abstract Test Suite for Traffic Management Phase 1

Element security options
▸ AAA ▸ RADIUS ▸ Secure Shell (SSH v2) ▸ Secure Copy (SCP v2) ▸ HTTPs ▸ TACACS/TACACS+ ▸ Username/Password (Challenge and Response) ▸ Bi-level Access Mode (Standard and EXEC Level) ▸ Protection against Denial of Service attacks, such as TCP SYN or Smurf Attacks

Network management and monitoring
▸ IronView Network Manager (INM) Web-based graphical user interface ▸ IBM Tivoli NetCOOL ▸ IBM Systems Director ▸ Integrated industry standard Command Line Interface (CLI) ▸ sFlow (RFC 3176) ▸ Telnet ▸ SNMP v1, v2c, v3 ▸ SNMP MIB II ▸ RMON

For the complete list, see the *NetIron Configuration Guide* for the software release in which you are interested. For a summarized list similar to the preceding table, see the *IBM m-series Ethernet routers Data Sheet*.

4.2 IBM r-series Ethernet Switches

The IBM r-series Ethernet Switches are high-performing, high-capacity Layer 2 switches with Full Layer 3 capabilities. Featuring low latency and jitter, the r-series are ideal for end-of-row and aggregation deployment in data centers, high performance computing environments, and Enterprise campus distribution.

The r-series supports high-density, non-oversubscribed 4x10 GbE and 48x1 GbE Interface Modules and the high-capacity, non-blocking 16x10 GbE Interface Module.

The r-series is available in the following model configurations:

- IBM Ethernet Switch B04R (4003-R04): A 4 RU chassis with 4 slots for interface modules, one dual-fan fan tray, and up to two management modules, three switch fabric modules, and three 1200 W power supplies.

- IBM Ethernet Switch B08R (4003-R08): A 7 RU chassis with 8 slots for interface modules, one quad-fan fan tray, and up to two management modules, three switch fabric modules, and four 1200 W power supplies.
- IBM Ethernet Switch B16R (4003-R16): A 14 RU chassis with 16 slots for interface modules, one front fan tray, two rear fan assemblies, and up to two management modules, four switch fabric modules, and eight 1200 W power supplies.

All r-series models run a common Operating System, Multi-Service IronWare for IBM r-series. All software functionality is enabled in the base software and available across the product family.

4.2.1 Feature highlights for r-series

In this section we provide feature highlights of the r-series running Multi-Service IronWare for IBM r-series R02.7.2a and higher:

- Extremely high densities scaling up to 512 10 GbE ports and 1,536 1 GbE ports in a single 32-slot chassis
- Extremely low latency and jitter allowing for network predictability
- End-to-end Quality of Service supported with hardware based honoring and marking, and congestion management
- Embedded, hardware-based sFlow traffic monitoring enabling network-wide accounting, utilization, reporting, capacity planning, intrusion detection, and more
- Classic Layer 2 support:
  - IEEE 802.1d (STP), 802.1w (RSTP), 802.1s (MSTP)
  - Single Spanning Tree Protocol (SSTP), PVST/PVST+, and Topology Groups, which enables either:
    - A single Spanning Tree instance for all VLANs
    - A separate Spanning Tree instance per VLAN
    - Sharing of a Spanning Tree instance across multiple VLANs per user configuration
  - IEEE 802.1q (VLAN tagging)
  - IEEE 802.1p (Class of Service Prioritization)
  - IEEE 802.1ad (Link Aggregation Control Protocol)
  - IEEE 802.1x (Port Security)
  - IGMPv1/v2/v3 & PIM-SM snooping

  - Metro Ring Protocol (MRP):
    - Alternative to Spanning Tree Protocol.
    - Prevents Layer 2 loops and provides fast, sub-second reconvergence in a Layer 2 ring topology.
  - Virtual Switch Redundancy Protocol (VSRP):
    - Provides redundancy and sub-second failover in a Layer 2 mesh topology
    - Based on VRRPE, provides one or more backup switches for Layer 2 redundancy
    - m-series also supports providing concurrent redundancy for Layer 3

- Layer 3:
  - Virtual Router Redundancy Protocol (VRRP) and VRRPE (Extended)
  - IPv4 Unicast support:
    - RIPv1/v2
    - OSPFv2
    - BGP-4
    - IS-IS
  - IPv4 Multicast support:
    - IGMPv1/v2/v3
    - PIM-DM
    - PIM-SM
    - PIM-SSM
    - MSDP
    - Anycast RP
  - IPv6 Unicast support:
    - RIPng
    - OSPFv3
    - BGP-4+
    - IS-IS for IPv6
  - IPv6 Multicast support:
    - PIM-SM
    - PIM-SSM
    - Anycast RP
    - MLDv1/v2

4.2.2 Technical specifications for r-series

The r-series running Multi-Service IronWare for IBM r-series R02.7.02a and higher conform to the summarized list shown in Table 4-2.

Table 4-2 r-series

IEEE compliance
► 802.3ae 10-Gigabit Ethernet ► 802.3x Flow Control ► 802.3ad Link Aggregation ► 802.1Q VLAN Tagging ► 802.1D Bridging ► 802.1w Rapid STP ► 802.1s Multiple Spanning Tree Protocol ► 802.1X User authentication ► 802.3 Ethernet Like MIB

RFC compliance	
BGPv4	► RFC 4271 BGPv4 ► RFC 1745 OSPF interactions ► RFC 1997 Communities & Attributes ► RFC 2439 route flap dampening ► RFC 2796 route reflection ► RFC 3065 BGP4 confederations ► RFC 3392 Capability Advertisement ► RFC 2918 Route Refresh Capability ► RFC 1269 Managed Objects for BGP ► RFC 1657 Managed Objects for BGP-4 using SMIv2 ► RFC 3682 Generalized TTTTL Security Mechanism for eBGP Session Protection ► RFC 2385 BGP Session Protection by TCTCP MD5 ► Draft-ieft-idr-restart Graceful Restart for BGP ► Draft-ieft-idr-route-filter
OSPF	► RFC 2178 OSPF ► RFC 1583 OSPF v2 ► RFC 3101 OSPF NSSA ► RFC 1745 OSPF Interactions ► RFC 1765 OSPF Database Overflow ► RFC 1850 OSPF v2 MIB and Traps ► RFC 2154 OSPF w/Digital Signatures (Password, MD-5) ► RFC 2328 OSPF v2 ► RFC 2370 OSPF Opaque LSA Option ► RFC 3623 Graceful OSPF Restart

RFC compliance	
IS-IS	► RFC 1195 Routing in TCTCP/IP and Dual Environments ► RFC 2763 Dynamic Host Name Exchange ► RFC 2966 Domain-wide Prefix Distribution ► RFC 3567 IS-IS Cryptographic Authentication (MDS)
RIP	► RFC 1058 RIP v1 ► RFC 1723 RIP v2 ► RFC 1812 RIP Requirements
IP Multicast	► RFC 1122 Host Extensions ► RFC 1256 ICMP Router Discovery Protocol ► RFC 1112 IGMP ► RFC 2236 IGMP v2 ► RFC 2362 PIM-SM ► RFC 3973 PIM-DM ► PIM-DM v1 ► D DVMRP v3-07 ► RFC 1075 DVMRP v2 ► RFC 2336 IGMP v2 ► RFC 3618 MSDP ► RFC 2283 MBGP ► RFC 2858 BGP-MP ► RFC 3376 IGMP v3 ► RFC 3446 Anycast RP ► RFC 4541 Considerations for IGMP and MLD Snooping

RFC compliance	
General protocols	► RFC 791 IP ► RFC 792 ICMP ► RFC 793 TCTCP ► RFC 783 TFTP ► RFC 826 ARP ► RFC 768 UDP ► RFC 894 IP over Ethernet ► RFC 903 RARP ► RFC 906 TFTP Bootstrap ► RFC 1027 Proxy ARP ► RFC 950 Subnets ► RFC 951 BootP ► RFC 1122 Host Requirements ► RFC 1256 IRDP ► RFC 1519 CIDR ► RFC 1542 BootP Extensions ► RFC 1812 General Routing ► RFC 1541 and 1542 DHCP ► RFC 2131 BootP/DHCP Helper ► RFC 3768 VRRP ► RFC 854 TELNET ► RFC 1591 DNS (client) ► RFC 2784 GRE ► RFC 1191 Path MTU Discovery ► RFC 896 Congestion Control ► RFC 3635 Pause Control ► RFC 1858 IP Fragment Filtering ► RFC 1340 Assigned Numbers
Management	► RFC 2578 SMIv2 ► RFC 2579 Textual Conventions for SMIv2 ► RFC 2665 Ethernet Interface MIB ► RFC 1354 IP Forwarding MIB ► RFC 1757 RMON Groups Partial 1, full for 2, 3, 9 ► RFC 2068 HTTTP ► RFC 2030 SNTP ► RFC 2138 RADIUS ► RFC 3176 sFlow ► Draft-ieft-tcpm-tcpsecure-00

RFC compliance	
IPv6 core	► RFC 2373 IPv6 Addressing architecture ► RFC 1886 DNS Extensions to support IPv6 ► RFC 1887 IPV6 Unicast address allocation architecture ► RFC 2374 IPv6 aggregatable global Unicast address format ► RFC 2450 Proposed TLA and NLA Assignment Rules ► RFC 2471 IPv6 testing address allocation ► RFC 2526 Reserved IPv6 subnet anycast address ► RFC 2928 Initial IPv6 sub TLA ID assignments ► RFC 2460 IPv6 Specification ► RFC 2461 IPv6 Neighbor Discovery ► RFC 2462 IPv6 Stateless Address Auto-configuration ► RFC 4443 ICMPv6 ► RFC 3513 IPv6 Addressing Architecture ► RFC 1981 IPv6 Path MTU Discovery ► RFC 3587 IPv6 Global Unicast Address Format ► RFC 2375 IPv6 Multicast Address Assignments ► RFC 2464 Transmission of IPv6 over Ethernet Networks ► RFC 2711 IPv6 Router Alert Option ► RFC 3363 DNS support
IPv6 routing	► RFC 2080 RIPng for IPv6 ► RFC 2740 OSPFv3 for IPv6 ► IETF Draft_ietf_isis_IPv6 IS-IS for IPv6 ► RFC 2545 Use of MP-BGP-4 for IPv6
IPv6 multicast	► RFC 2362 PIM-SM ► RFC 2710 Multicast Listener Discovery (MLD) for IPv6 ► RFC 3306 Unicast-Prefix-based IPv6 Multicast Addresses ► RFC 3810 MLDv2 ► RFC 4602 PIM-SM (Partial Address) ► draft-holbrook-idmr-igmpv3-ssm—IGMPv3 & MDLV2 for SSM ► Draft-ieft-ssm-arch SSM for IP
IPv6 transitioning	► RFC 2893 Transition Mechanisms for IPv6 Hosts and Routers ► RFC 3056 Connection of IPv6 Domains by IPv4 Clouds

Element security options
► AAA ► RADIUS ► Secure Shell (SSH v2) ► Secure Copy (SCP v2) ► TACACS/TACACS+ ► Username/Password (Challenge and Response) ► Bi-level Access Mode (Standard and EXEC Level) ► Protection against Denial of Service attacks, such as TCP SYN or Smurf Attacks

Network management and monitoring
► IronView Network Manager (INM) Web-based graphical user interface ► IBM Tivoli NetCOOL ► IBM Systems Director ► Integrated industry standard Command Line Interface (CLI) ► sFlow (RFC 3176) ► Telnet ► SNMP v1, v2c, v3 ► SNMP MIB II ► RMON

For a complete list, see the *BigIron RX Series Configuration Guide* for the software release in which you are interested. For a summarized list similar to the preceding table, see the *IBM r-series Ethernet Switches Data Sheet*.

4.3 IBM s-series Ethernet Switches

The IBM s-series Ethernet Switches are extremely flexible, high-port count chassis featuring full Layer 2 capabilities and options for Power over Ethernet (PoE) and Full Layer 3 (IPv4 or IPv6/IPv4), making this device ideal for aggregating large amounts of PoE end devices in an Enterprise Campus or as a distribution switch carrying converged video, voice, and data traffic.

The s-series supports non-oversubscribed 2x10 GbE and 24x1 GbE Interface Modules with Class 3 PoE capabilities. In addition, the management modules have options for 2x10 GbE ports on them.

The s-series is available in the following model configurations:

- IBM Ethernet Switch B08S (4003-S08): A 6 RU chassis with 8 slots for interface modules, one six-fan fan tray, and up to two management modules, two switch fabric modules, two 1200 W System (SYS) power supplies, and two 1250 W or 2500 W Power over Ethernet (PoE) power supplies.
- IBM Ethernet Switch B16S (4003-S16): A 14 RU chassis with 16 slots for interface modules, two rear fan assemblies, and up to two management modules, two switch fabric modules, four 1200 W System (SYS) power supplies, and four 1250 W or 2500 W Power over Ethernet (PoE) power supplies.

All s-series models run a common Operating System, IronWare. Most software functionality is enabled in the base software and includes Full Layer 2, Base Layer 3 (static IP routes), multicast snooping, Quality of Service, and comprehensive hardware-based security and policies including a wide range of Access Control Lists.

Two types of management and interface modules are available, those that support IPv4 and those that support IPv6/IPv4. A chassis must contain all of the same types of modules to operate, either all IPv4 or all IPv6/IPv4.

Upgradeable software activation available include:

- IPv4 Full Layer 3 Premium Activation (L3_PREM) for IPv4 Management Modules
- IPv4 Full Layer 3 Premium Activation (L3_PREM) for IPv6/IPv4 Management Modules
- IPv6/IPv4 Full Layer 3 Premium Activation (L3_PREM) for IPv6/IPv4 Management Modules

Each upgrade activation adds Full Layer 3 unicast and multicast routing capabilities, either IPv4 protocols support and/or IPv6 protocols support. See the next section for details on what the license activations add.

4.3.1 Feature highlights for s-series

In this section we provide feature highlights of the s-series running IronWare R05.1.00 and higher:

- Exceptional Class 3 (15.4 W/port) PoE capacity scaling to 384 10/100/1000 MbE PoE ports per chassis providing a convergence-ready (video/voice/data) infrastructure scalable to support future growth
- Combined SP/WRR queuing and cell-based switch fabric ensures low latency and jitter for voice and video traffic
- Intelligent PoE and configuration management with LLDP, LLDP-MED, and PoE Prioritization for IP phones
- Embedded, hardware-based sFlow traffic monitoring enabling network-wide accounting, utilization, reporting, capacity planning, intrusion detection, and more
- Classic Layer 2 support:
  - IEEE 802.1d (STP), 802.1w (RSTP), 802.1s (MSTP)
  - Single Spanning Tree Protocol (SSTP), PVST/PVST+, and Topology Groups, which enables either:
    - A single Spanning Tree instance for all VLANs
    - A separate Spanning Tree instance per VLAN
    - Sharing of a Spanning Tree instance across multiple VLANs per user configuration

- IEEE 802.1q (VLAN tagging)
- IEEE 802.1p (Class of Service Prioritization)
- IEEE 802.1ad (Link Aggregation Control Protocol)
- IEEE 802.1x (Port Security)
- IGMPv1/v2/v3, PIM-SM & MLDv1/v2 snooping
- Metro Ring Protocol (MRP):
  - Alternative to Spanning Tree Protocol. Prevents Layer 2 loops
  - Provides fast, sub-second reconvergence in a Layer 2 ring topology.
- Virtual Switch Redundancy Protocol (VSRP):
  - Provides redundancy and sub-second failover in a Layer 2 mesh topology
  - Based on VRRPE, provides one or more backup switches for Layer 2 redundancy
  - s-series also supports providing concurrent redundancy for Layer 3

- Layer 3 IPv4 (requires either IPv4 Full Layer 3 Activation or IPv6/IPv4 Full Layer 3 Activation):
  - Virtual Router Redundancy Protocol (VRRP) and VRRPE (Extended)
  - IPv4 Unicast support:
    - RIPv1/v2
    - OSPFv2
    - BGP-4
    - IS-IS
  - IPv4 Multicast support:
    - IGMPv1/v2/v3
    - PIM-DM
    - PIM-SM
    - PIM-SSM
    - MSDP
    - Anycast RP
- Layer 3 IPv6 (requires IPv6/IPv4 Full Layer 3 Activation for IPv6/IPv4 management and interface modules only):
  - IPv6 Unicast support:
    - RIPng
    - OSPFv3
    - BGP-4+
    - IS-IS for IPv6

- IPv6 Multicast support:
  - PIM-SM
  - PIM-SSM
  - Anycast RP
  - MLDv1/v2

For the most up-to-date list, see the software *Release Notes* and the *FastIron Configuration Guide* for the software version in which you are interested.

4.3.2 Technical specifications for s-series

The s-series running IronWare R05.1.00 and higher conform to the summarized list shown in Table 4-3.

Table 4-3 s-series

IEEE compliance
► 802.3 10Base-T
► 802.3u 100Base-TX
► 802.3u 100Base-FX
► 802.3u 100Base-LX
► 802.3z 1000Base-SX/LX
► 802.3ab 1000Base-T
► 802.3ae 10-Gigabit Ethernet
► 802.3af Power over Ethernet
► 802.3x Flow Control
► 802.3ad Link Aggregation
► 802.1d Ethernet Bridging
► 802.1D MAC Bridges
► 802.1p/q VLAN Tagging
► 802.1w Rapid Spanning Tree
► 802.1s Multiple Spanning Tree
► 802.1X Port-based Network Access Control
► 802.1Q Generic VLAN Registration Protocol (GVRP)
► 802.3 MAU MIB (RFC 2239)
► 802.3AB LLDP

RFC compliance	
Protocol support	► DNS Client ► RFC 1812 IP Requirements ► RFC 2338 VRRP ► VRRPE ► PVST/PVST+/PVRST
BGPv4	► RFC 1269 BGP-3 MIB ► RFC 1657 BGP-4 MIB ► RFC 1745 OSPF Interactions ► RFC 1771 BGP-4 ► RFC 1965 BGP-4 Confederations ► RFC 1997 Communities Attribute ► RFC 2385 TCP MD5 ► Authentication of BGP Session ► RFC 2439 Route Flap Dampening ► RFC 2796 Route Reflection ► RFC 2842 BGP4 Capabilities Advertisement ► RFC 2918 Route Refresh Capability
OSPF	► RFC 1583 and 2328 OSPF v2 ► RFC 1587 OSPF NSSA Option ► RFC 1745 OSPF Interactions ► RFC 1765 OSPF Database Overflow ► RFC 1850 OSPF Traps ► RFC 1850 OSPF v2 MIB ► RFC 2154 OSPF w/Digital Signatures (Password, MD-5) ► RFC 2178 OSPF v2 ► RFC 2370 OSPF Opaque LSA Option
RIP	► RFC 1058 RIP v1 ► RFC 1723 RIP v2
IP Multicast	► RFC 1112 IGMP ► RFC 2236 IGMP v2 ► RFC 3376 IGMP v3 ► IGMP Proxy ► DVMRP v3-07 ► RFC 1075 DVMRP ► RFC 1122 Host Extensions ► RFC 1256 ICMP Router Discovery Protocol ► PIM-DM v1 ► RFC 2362 PIM-SM ► PIM-SSM

RFC compliance	
General protocols	► RFC 768 UDP ► RFC 783 TFTP ► RFC 791 IP ► RFC 792 ICMP ► RFC 793 TCP ► RFC 826 ARP ► RFC 854 TELNET ► RFC 894 IP over Ethernet ► RFC 903 RARP ► RFC 906 TFTP Bootstrap ► RFC 1027 Proxy ARP ► RFC 1519 CIDR ► RFC 1541 and 2131 DHCP ► RFC 1591 DNS (client) ► RFC 1812 General Routing ► RFC 2338 VRRP

RFC compliance	
Management	► RFC 1157 SNMPv1 ► RFC 1191 Path MTU Discovery ► RFC 951 BootP ► RFC 1542 BootP Extensions ► RFC 1493 Bridge MIB ► RFC 1215 SNMP Generic Traps ► RFC 1354 IP Forwarding MIB ► RFC 1573 SNMP MIB II ► RFC 1757 RMON Groups 1,2,3,9 ► RFC 1905, 1906 SNMPv2c ► RFC 2030 SNTP ► RFC 2068 HTTP ► RFC 2818 HTTPS ► RFC 2138 RADIUS ► RFC 2571 Architecture Describing SNMP Framework ► RFC 3176 sFlow ► RFC 3411 SNMPv3 Framework ► RFC 2570 SNMPv3 Intro to Framework ► RFC 3412 SNMPv3 Processing ► RFC 3414 SNMPv3 USM ► RFC 2574 SNMPv3 User-based Security Model (USM) ► RFC 2573 SNMPv3 Applications ► RFC 2575 SNMP View-based Access Control Model SNMP (VACM) ► RFC 3415 SNMPv3VACM ► RFC 1643 Ethernet-like Interface MIB ► RFC 1354 IP Forwarding Table MIB ► RFC 1213 MIB-II ► RFC 1516 Repeater MIB ► RFC 1724 RIPv2 MIB ► RFC 2572 SNMP Message Processing and Dispatching

Other
► ANSI TIA 1057 LLDP-MED ► TACACS+ v1.78 ► MRP (Metro Ring Protocol) ► UDLD (Uni-directional Link Detection) ► IGMP Snooping ► Dynamic Filters and VLAN assignment ► CDP and FDP ► Configuration Logging

Quality of Service (QoS)

- MAC Address Mapping to Priority Queue
- ACL Mapping to Priority Queue
- ACL Mapping to ToS/DSCP
- ACL Mapping and Marking of ToS/DSCP
- DiffServ Support
- QoS Queue Management Using Weighted Round Robin (WRR), Strict Priority (SP), and a combination of WRR and SP

Element security options

- AAA
- RADIUS
- Secure Shell (SSH v2)
- Secure Copy (SCP v2)
- TACACS/TACACS+
- Username/Password (Challenge and Response)
- Bi-level Access Mode (Standard and EXEC Level)
- Protection against Denial of Service attacks, such as TCP SYN or Smurf Attacks

Network management and monitoring

- IronView Network Manager (INM) Web-based graphical user interface
- IBM Tivoli NetCOOL
- IBM Systems Director
- Integrated industry standard Command Line Interface (CLI)
- sFlow (RFC 3176)
- Telnet
- SNMP v1, v2c, v3
- SNMP MIB II
- RMON

For a complete list, see the *FastIron Configuration Guide* for the software release in which you are interested. For a summarized list similar to the preceding table, see the *IBM s-series Ethernet Switches Data Sheet*.

4.4 IBM c-series Ethernet Switches

The IBM c-series Ethernet Switches are extremely robust, richly featured switches that can support Advanced Layer 3 services such as MPLS, VPLS, and Multi-VRF routing. These devices are well suited for demanding Top-of-Rack (ToR) within the Data Center where the advanced security provided by Layer 3 partitioning with VRF might be required, or as a remote office border router where advanced Metro services are needed.

The compact, 1 RU switches are available in several different models, including:

- IBM Ethernet Switch B24C (C) (4002-A2C / 4002AC2): 24x 10/100/1000 MbE ports (RJ45) including 4x 100/1000 MbE combination ports (SFP) w/optional slot for 2-port 10 GbE module (CX4 or XFP)
- IBM Ethernet Switch B24C (F) (4002-B2C / 4002BC2): 24x 100/1000 MbE ports (SFP) including 4x 10/100/1000 MbE combination ports (RJ45) w/optional slot for 2-port 10 GbE module (CX4 or XFP)
- IBM Ethernet Switch B48C (C) (4002-A4C / 4002AC4): 48x 10/100/1000 MbE ports (RJ45) including 4x 100/1000 MbE combination ports (SFP)
- IBM Ethernet Switch B48C (F) (4002-B4C / 4002BC4)- 48x 100/1000 MbE ports (SFP)
- IBM Ethernet Switch B50C (C) (4002-A5C / 4002AC5): 48x 10/100/1000 MbE ports (RJ45) plus 2x 10 GbE ports (XFP)
- IBM Ethernet Switch B50C (F) (4002-B5C / 4002BC5): 48x 100/1000 MbE ports (SFP) plus 2x 10 GbE ports (XFP)

The c-series comes with all ports active, a 1+1 redundant 500 W power supply system, and resilient six-fan fan tray with Data Center designed, front-to-back airflow. All c-series models run a common Operating System, Multi-Service IronWare, based off the same OS the m-series Ethernet routers use. The base software functionality includes Full Layer 2, some Layer 3 support (RIPv1/v2 and static IP routes), multicast snooping, Quality of Service, and comprehensive hardware-based security and policies including a wide range of Access Control Lists.

Upgradeable software activation available include:

- Layer 3 Premium Activation (L3_PREM):

  Adds Layer 3 support
- Metro Edge Premium Activation (ME_PREM):
  - Adds Advance Layer 2 support
  - Adds Layer 3 (except BGP-4) support
  - Adds MPLS/VPLS/VRF support

See the next section for details on what the license activations add.

4.4.1 Feature highlights for c-series

In this section we provide feature highlights of the c-series running Multi-Service IronWare R03.9.00 and higher:

- Compact 1 RU IP/MPLS/VRF-enabled switch purpose-built for advanced Carrier Ethernet and large data center applications
- MEF 9 and MEF 14 certification with comprehensive Operations, Administration, and Maintenance (OAM) capabilities based on the IEEE 802.1ag-2007 and MEF Service OAM Framework
- Wire-speed, non-blocking performance in all configurations
- Deep egress buffering for transient bursts in traffic featuring 64 MB of buffering per 24x 1 GbE or 2x 10 GbE port groups
- Embedded, hardware-based sFlow traffic monitoring enabling network-wide accounting, utilization, reporting, capacity planning, intrusion detection, and more
- Classic Layer 2 support:
  - IEEE 802.1d (STP), 802.1w (RSTP), 802.1s (MSTP)
  - Single Spanning Tree Protocol (SSTP), PVST/PVST+, and Topology Groups, which enables either:
    - A single Spanning Tree instance for all VLANs
    - A separate Spanning Tree instance per VLAN
    - Sharing of a Spanning Tree instance across multiple VLANs per user configuration
  - IEEE 802.1q (VLAN tagging)
  - IEEE 802.1p (Class of Service Prioritization)
  - IEEE 802.1ad (Link Aggregation Control Protocol)
  - IEEE 802.1x (Port Security)
  - IGMPv1/v2/v3 and PIM-SM snooping
  - Metro Ring Protocol (MRP):
    - Alternative to Spanning Tree Protocol. Prevents Layer 2 loops
    - Provides fast, sub-second reconvergence in a Layer 2 ring topology.
  - Virtual Switch Redundancy Protocol (VSRP):
    - Provides redundancy and sub-second failover in a Layer 2 mesh topology
    - Based on VRRPE, provides one or more backup switches for Layer 2 redundancy
    - m-series also supports providing concurrent redundancy for Layer 3

- Advanced Layer 2 support (requires Metro Edge Premium Activation):
  - Ethernet Service Instance (ESI)
  - IEEE 802.1ad (Provider Bridges)
  - IEEE 802.1ag (Connectivity Fault Management)
  - IEEE 802.1ah (Provider Backbone Bridges)
- Layer 3:
  - Virtual Router Redundancy Protocol (VRRP) and VRRPE (Extended)
  - IPv4 Unicast support:

    RIPv1/v2
- Layer 3 (requires Layer 3 Premium Activation):
  - IPv4 Unicast support:

    BGP-4
- Layer 3 (requires Layer 3 or Metro Premium Activation):
  - Multi-VRF for IPv4 Unicast (OSPF, BGP-4, Static Routes)
  - IPv4 Unicast support:
    - OSPFv2
    - IS-IS
  - IPv4 Multicast support:
    - IGMPv1/v2/v3
    - PIM-DM
    - PIM-SM
    - PIM-SSM
    - MSDP
    - Anycast RP
- Multiprotocol Label Switching (MPLS):
  - Label Distribution Protocol (LDP)
  - Traffic Engineering: RSVP-TE, OSPF-TE, ISIS-TE
- Virtual Private Networks (VPN):
  - Virtual Private LAN Services (VPLS)
  - Virtual Leased Lines (VLL)

For the most up-to-date list, see the software *Release Notes* and the *NetIron Configuration Guide* for the software version in which you are interested.

4.4.2 Technical specifications for c-series

The c-series running Multi-Service IronWare R03.9.00 and higher conform to the summarized list shown in Table 4-4.

Table 4-4 c-series

IEEE compliance
► 802.3 10Base-T ► 802.3u 100Base-TX, 100Base-FX, 100Base-LX ► 802.3z 1000Base-SX/LX ► 802.3ab 1000Base-T ► 802.3 CSMA/CD Access Method and Physical Layer Specifications ► 802.3ae 10 Gigabit Ethernet ► 802.3x Flow Control ► 802.3ad Link Aggregation ► 802.1Q Virtual Bridged LANs ► 802.1D MAC Bridges ► 802.1w Rapid STP ► 802.1s Multiple Spanning Trees ► 802.1x Port-based Network Access Control ► 802.1ad Provider Bridges ► 802.1ah Provider Backbone Bridges ► 802.1ag Connectivity Fault Management (CFM)

RFC compliance	
BGPv4	► RFC 4271 BGPv4 ► RFC 1745 OSPF Interactions ► RFC 1997 Communities and Attributes ► RFC 2439 Route Flap Dampening ► RFC 2796 Route Reflection ► RFC 1965 BGP4 Confederations ► RFC 2842 Capability Advertisement ► RFC 2918 Route Refresh Capability ► RFC 1269 Managed Objects for BGP ► RFC 2385 BGP Session Protection by TCP MD5 ► RFC 3682 Generalized TTL Security Mechanism, for eBGP Session Protection ► RFC 4273 BGP-4 MIB
OSPF	► RFC 2328 OSPF v2 ► RFC 3101 OSPF NSSA ► RFC 1745 OSPF Interactions ► RFC 1765 OSPF Database Overflow ► RFC 1850 OSPF v2 MIB ► RFC 2370 OSPF Opaque LSA Option

RFC compliance	
RIP	► RFC 1058 RIP v1 ► RFC 1723 RIP v2 ► RFC 1812 RIP Requirements
IS-IS	► RFC 1195 Routing in TCP/IP and Dual Environments ► RFC 1142 OSI IS-IS Intra-domain Routing Protocol ► RFC 2763 Dynamic Host Name Exchange ► RFC 2966 Domain-wide Prefix Distribution
IP Multicast	► RFC 1122 Host Extensions ► RFC 1112 IGMP ► RFC 2236 IGMP v2 ► RFC 3376 IGMP v3 ► RFC 3973 PIM-DM ► RFC 2362 PIM-SM
MPLS	► RFC 3031 MPLS Architecture ► RFC 3032 MPLS Label Stack Encoding ► RFC 3036 LDP Specification ► RFC 2205 RSVP v1 Functional Specification ► RFC 2209 RSVP v1 Message Processing Rules ► RFC 3209 RSVP-TE ► RFC 3270 MPLS Support of Differentiated Services ► RFC 3812 MPLS MIB ► Draft-ieft-bfd-mpls BFD for MPLS LSPs (RSVP-TE)
L2VPN and PWE3	► Draft-ieft-l2vpn-framework Framework for Layer 2 Virtual Private Networks ► Draft-ieft-l2vpn-requirements Service Requirements for Layer 2 Provider Provisioned Virtual Private Networks ► RFC 4762 VPLS using LDP Signaling ► Draft-ieft-pwe3-arch PWE3 Architecture ► RFC 4447 Pseudowire Setup and Maintenance using LDP ► RFC 4448 Encapsulation Methods for Transport of Ethernet over MPLS Networks ► Draft-ieft-pwe3-pw-tc-mib Definitions for Textual Conventions and OBJECT IDENTITIES for Pseudo-Wires Management ► Draft-ieft-pwe3-pw-mib Pseudo Wire (PW) Management Information Base

RFC compliance	
General protocols	► RFC 791 IP ► RFC 792 ICMP ► RFC 793 TCP ► RFC 783 TFTP ► RFC 826 ARP ► RFC 768 UDP ► RFC 894 IP over Ethernet ► RFC 903 RARP ► RFC 906 TFTP Bootstrap ► RFC 1027 Proxy ARP ► RFC 951 BootP ► RFC 1122 Host Extensions for IP Multicasting ► RFC 1256 IRDP ► RFC 1519 CIDR ► RFC 1542 BootP Extensions ► RFC 1812 Requirements for IPv4 Routers ► RFC 1541 and 1542 DHCP ► RFC 2131 BootP/DHCP Helper ► RFC 3768 VRRP ► RFC 854 TELNET ► RFC 1591 DNS (client)
Quality of Service	► RFC 2475 An Architecture for Differentiated Services ► RFC 3246 An Expedited Forwarding PHB ► RFC 2597 Assured Forwarding PHB Group ► RFC 2698 A Two Rate Three Color Marker
Other	► RFC 2665 Ethernet Interface MIB ► RFC 1757 RMON Groups 1,2,3,9 ► RFC 2068 HTTP ► RFC 2030 SNTP ► RFC 2865 RADIUS ► RFC 3176 sFlow ► RFC 2863 Interfaces Group MIB ► Draft-ieft-tcpm-tcpsecure TCP Security

MEF specifications
► MEF 2 Requirements and Framework for Ethernet Service Protection ► MEF 4 Metro Ethernet Network Architecture Framework Part 1: Generic Framework ► MEF 6.1 Metro Ethernet Services Definitions Phase 2 ► MEF 9 Abstract Test Suite for Ethernet Services at the UNI ► MEF 10.1 Ethernet Services Attributes Phase 2 ► MEF 11 User Network Interface (UNI) Requirements and Framework ► MEF 12 Metro Ethernet Network Architecture Framework Part 2: Ethernet Services Layer ► MEF 13 User Network Interface (UNI) Type 1 Implementation Agreement ► MEF 14 Abstract Test Suite for Traffic Management Phase 1 ► MEF 15 Requirements for Management of Metro Ethernet Phase 1 Network Elements ► MEF 17 Service OAM Framework and Requirements (partial) ► MEF 19 Abstract Test Suite for UNI Type 1
Element security options
► AAA ► RADIUS ► Secure Shell (SSH v2) ► Secure Copy (SCP v2) ► TACACS/TACACS+ ► Username/Password (Challenge and Response) ► Bi-level Access Mode (Standard and EXEC Level) ► Protection against Denial of Service attacks, such as TCP SYN or Smurf Attacks
Network management and monitoring
► IronView Network Manager (INM) Web-based graphical user interface ► IBM Tivoli NetCOOL ► IBM Systems Director ► Integrated industry standard Command Line Interface (CLI) ► sFlow (RFC 3176) ► Telnet ► SNMP v1, v2c, v3 ► SNMP MIB II ► RMON

For a complete list, see the *NetIron Configuration Guide* for the software release in which you are interested. For a summarized list similar to the preceding table, see the *IBM c-series Ethernet Switches Data Sheet*.

4.5 IBM x-series Ethernet Switches

The IBM x-series Ethernet Switch (4002-X2A / 4002AX2) is a high-performing, low-latency (1.5 µs), 1 RU Top-of-Rack (ToR), low-cost Aggregation switch featuring cut-through switching on 24 dual-speed 10/1 GbE ports plus four 1 GbE ports. The x-series is ideal for environments where performance and latency are critical such as High Performance Computing (HPC) and iSCSI storage environments as well as being used as a lightweight aggregation/distribution switch for 10/100/1000 MbE Edge/Access switches within the Data Center or Enterprise Campus.

The x-series comes with all ports active, a 1+1 redundant 300 W power supply system, and resilient triple-fan fan module featuring Data Center designed, front-to-back airflow. The device runs the IronWare Operating System. All supported software functionality is enabled in the base software and includes Full Layer 2, multicast snooping, Quality of Service, and a wide range of Access Control Lists.

4.5.1 Feature highlights for x-series

In this section we provide feature highlights of the x-series running IronWare R04.1.00c and higher:

- Flexibility to mix 10 GbE and 1 GbE servers on dual-speed 10/1 GbE ports, protecting investments and streamlining migration to 10 GbE-capable server farms
- Wire-speed performance with ultra-low-latency, cut-through, non-blocking architecture
- End-to-end QoS with hardware-based marking, queuing, and congestion management
- Embedded, hardware-based sFlow traffic monitoring enabling network-wide accounting, utilization, reporting, capacity planning, intrusion detection, and more
- Classic Layer 2 support:
  - IEEE 802.1d (STP), 802.1w (RSTP), 802.1s (MSTP)
  - Single Spanning Tree Protocol (SSTP) and PVST/PVST+, which enables either:
    - A single Spanning Tree instance for all VLANs
    - A separate Spanning Tree instance per VLAN
  - IEEE 802.1q (VLAN tagging)

- IEEE 802.1p (Class of Service Prioritization)
- IEEE 802.1ad (Link Aggregation Control Protocol)
- IGMPv1/v2/v3, PIM-SM snooping

For the most up-to-date list, see the software *Release Notes* and the *FastIron and TurboIron Configuration Guide* for the software version in which you are interested.

4.5.2 Technical specifications for x-series

The x-series running IronWare R04.1.00c and higher conform to the summarized list shown in Table 4-5.

Table 4-5 x-series

IEEE compliance
► 802.3x Flow Control ► 802.3ad Link Aggregation ► 802.1D MAC Bridging/STP ► 802.1Q VLAN Tagging ► 802.1w Rapid Spanning Tree Protocol (RSTP) ► 802.1s Multiple Spanning Tree Protocol (MSTP) ► 802.3AB LLDP ► 802.3 MAU MIB (RFC 2239)

RFC compliance	
Protocol support	► RFC 791 IP ► RFC 768 UDP ► RFC 783 TFTP ► RFC 792 ICMP ► RFC 793 TCP ► RFC 826 ARP ► RFC 894 IP over Ethernet ► RFC 903 RARP ► RFC 906 TFTP Bootstrap ► RFC 1027 Proxy ARP ► RFC 1519 CIDR ► RFC 1541 and 2131 DHCP ► RFC 1591 DNS (client)
IP Multicast	► RFC 1112 IGMP ► RFC 2236 IGMPv2 ► RFC 3376 IGMPv3 ► IGMP Proxy ► RFC 1122 Host Extensions

RFC compliance	
Management	► RFC 2571 Architecture for Describing SNMP Framework ► RFC 951 BootP ► RFC 1542 BootP Extensions ► RFC 2131 DHCP ► RFC 854 TELNET ► RFC 2865 RADIUS ► RFC 1493 Bridge MIB ► RFC 1643 Ethernet-like Interface MIB ► RFC 3176 sFlow ► RFC 1213 MIB-II ► RFC 1516 Repeater MIB ► RFC 1354 IP Forwarding Table MIB ► RFC 1757 RMON MIB ► RFC 2572 SNMP Message Processing and Dispatching ► RFC 1573 SNMP MIB II ► RFC 1157 SNMPv1/v2c ► RFC 3411 SNMPv3 Framework ► RFC 2570 SNMPv3 Intro to Framework ► RFC 3412 SNMPv3 Processing ► RFC 3414 SNMPv3 USM ► RFC 2574 SNMPv3 User-Based Security Model (USM) ► RFC 2573 SNMPv3 Applications ► RFC 2575 SNMP View-Based Access Control Model (VACM) ► RFC 3415 SNMPv3 VACM

Quality of Service (QoS)
► Rate Limiting ► Traffic Shaping ► MAC Address Mapping to Priority Queue ► ACL Mapping to Priority Queue ► ACL Mapping to ToS/DSCP ► ACL Mapping and Marking of ToS/DSCP ► QoS Queue Management using Weighted Round Robin (WRR), Strict Priority (SP), and a combination of WRR and SP

Element security options
► AAA ► RADIUS ► Secure Shell (SSH v2) ► Secure Copy (SCP v2) ► TACACS/TACACS+ ► Username/Password (Challenge and Response) ► Bi-level Access Mode (Standard and EXEC Level)

Network management and monitoring
► IronView Network Manager (INM) Web-based graphical user interface ► Integrated industry standard Command Line Interface (CLI) ► sFlow (RFC 3176) ► Telnet ► SNMP v1, v2c, v3 ► SNMP MIB II ► RMON

For a complete list, see the *TurboIron and FastIron Configuration Guide* for the software release in which you are interested. For a summarized list similar to the preceding table, see the *IBM x-series Ethernet Switches Data Sheet*.

4.6 IBM g-series Ethernet Switches

The IBM g-series Ethernet Switches are cost-effective, scalable, Power over Ethernet (PoE) capable switches with pay-as-you grow stackable models that enables reduced management overhead while scaling out infrastructure. The g-series is ideal to be deployed in Enterprise Campus wiring closets connecting to PoE devices such as Access Points, VoIP phones, and surveillance cameras, as well as end-user devices such as laptop and desktop computers. Alternatively, the g-series can be used as a lightweight Top of Rack switch in a Data Center.

The g-series is available in the following models:

- IBM Ethernet Switch B48G (4002-A4G / 4002AG4): 48x 10/100/1000 MbE ports (RJ45, PoE capable, non-stackable) including 4x 100/1000 MbE combination ports (SFP) w/optional slot for 2-port 10 GbE module (CX4 or XFP)
- IBM Ethernet Switch B50C (4002-A5G / 4002AG5): 48x 10/100/1000 MbE ports (RJ45, PoE capable, stackable) including 4x 100/1000 MbE combination ports (SFP) plus 2-port 10 GbE module (CX4)

The g-series comes with all ports active and 1+1 redundant 600 W power supplies with fans housed within the power supply units. The devices runs the IronWare Operating System based off the same OS the s-series Ethernet Switches run. Most supported software functionality is enabled in the base software and includes Full Layer 2, Base Layer 3 (static IP routes), multicast snooping, Quality of Service, and a wide range of Access Control Lists.

Upgradeable software activation available include:

- Edge Layer 3 Premium Activation (L3_EDGE) for IBM Ethernet Switch B48G only

  Adds Edge Layer 3 support (RIP, OSPF)

See the next section for details on what the license activations add.

4.6.1 Feature highlights for g-series

The IBM Ethernet Switch B48G model runs IronWare R04.3.01b and higher. The IBM Ethernet Switch B50G model runs IronWare R05.0.01a and higher. In this section we provide feature highlights of the g-series running those versions of IronWare:

- Field upgradeable to support Power over Ethernet (PoE) Class 3 (15.4 W/port) on all 48 ports
- IBM Ethernet Switch B50G: IronStack technology to support scaling up to eight B50G switches in a logical chassis and 384 PoE ports, with automatic healing in case of link or switch failures
- Plug-and-play support of IP endpoints such as VoIP phones utilizing IEEE 802.1ab LLDP and ANSI TIA 1057 LLDP-MED standards
- Comprehensive Enterprise-class security including protection against TCP SYN and ICMP Denial of Service attacks, STP Root Guard, and BPDU Guard
- Embedded, hardware-based sFlow traffic monitoring enabling network-wide accounting, utilization, reporting, capacity planning, intrusion detection, and more
- Classic Layer 2 support:
  - IEEE 802.1d (STP), 802.1w (RSTP), 802.1s (MSTP)
  - Single Spanning Tree Protocol (SSTP), PVST/PVST+, and Topology Groups, which enables either:
    - A single Spanning Tree instance for all VLANs
    - A separate Spanning Tree instance per VLAN
    - Sharing of a Spanning Tree instance across multiple VLANs per user configuration
  - IEEE 802.1q (VLAN tagging)
  - IEEE 802.1p (Class of Service Prioritization)
  - IEEE 802.1ad (Link Aggregation Control Protocol)
  - IEEE 802.1x (Port Security)

    - IGMPv1/v2/v3, PIM-SM & MLDv1/v2 snooping
    - Metro Ring Protocol (MRP):
        - Alternative to Spanning Tree Protocol. Prevents Layer 2 loops
        - Provides fast, sub-second reconvergence in a Layer 2 ring topology.
    - Virtual Switch Redundancy Protocol (VSRP):
        - Provides redundancy and sub-second failover in a Layer 2 mesh topology
        - Based on VRRPE, provides one or more backup switches for Layer 2 redundancy
        - m-series also supports providing concurrent redundancy for Layer 3
- Edge Layer 3 (IBM Ethernet Switch B48G only: requires Edge Layer 3 Activation):
    - IPv4 Unicast support:
        - RIPv1/v2
        - OSPFv2
- IronStack (IBM Ethernet Switch B50G only):

  Ability to stack up to 8 switches into a single logical chassis of 384, PoE capable ports

For the most up-to-date list, see the software *Release Notes* and the *FastIron Configuration Guide* for the software version in which you are interested.

4.6.2 Technical specifications for g-series

IBM Ethernet Switch B48G running IronWare R04.3.01b and higher and IBM Ethernet Switch B50G running IronWare R05.0.01a and higher conform to the summarized list shown in Table 4-6.

Table 4-6 g-series

IEEE compliance
► 802.1D-2004 MAC Bridging ► 802.1w Rapid Spanning Tree ► 802.1s Multiple Spanning Tree ► 802.1X Port-based Network Access Control ► 802.3 10Base-T ► 802.3ak CX4 ► 802.3ad Link Aggregation (Dynamic and Static) ► 802.3af Power over Ethernet ► 802.3u 100Base-TX ► 802.3x Flow Control ► 802.3z 1000Base-SX/LX ► 802.3ab 1000Base-T ► 802.3ae 10 Gigabit Ethernet ► 802.3 MAU MIB (RFC 2239) ► 802.3AB LLDP/LLDP-MED ► 802.1p Mapping to Priority Queue

RFC compliance	
Management	► RFC 2571 Architecture for Describing SNMP Framework ► RFC 2131 DHCP Relay ► RFC 1493 Bridge MIB ► RFC 1643 Ethernet Interface MIB ► RFC 1643 Ethernet MIB ► RFC 2068 Embedded HTTP ► RFC 2818 Embedded HTTPS ► RFC 3176 sFlow ► RFC 1213 MIB-II ► RFC 1516 Repeater MIB ► RFC 1724 RIP v1/v2 MIB ► RFC 1757 RMON MIB ► RFC 2572 SNMP Message Processing and Dispatching ► RFC 1573 SNMP MIB II ► RFC 2575 SNMP View-based Access Control Model SNMP ► RFC 1157 SNMPv1/v2c ► RFC 2573 SNMPv3 Applications ► RFC 2570 SNMPv3 Intro to Framework ► RFC 2574 SNMPv3 User-based Security Model ► RFC 854 TELNET Client and Server ► RFC 783 TFTP

Layer 2 features

- 4,096 VLANs
- 16,000 MAC Addresses
- 802.1s Multiple Spanning Tree
- Per VLAN spanning tree (PVST/PVST+/PVRST)
- Private VLAN
- Protocol VLAN (802.1v), Subnet VLAN
- Policy controlled MAC-based VLANs
- MAC Learning Disable
- Port Security
- MAC Address Locking
- Port-based Access Control Lists
- Dual Mode VLANs
- Fast Port Span
- BPDU Guard, Root Guard
- GARP VLAN Registration Protocol
- MAC-Layer Filtering
- Port-based, ACL-based, MAC filter-based, and VLAN-based Mirroring
- Single-instance Spanning Tree
- Trunk groups
- Trunk threshold
- Single link LACP
- Uni-Directional Link Detection (UDLD)
- Auto MDI/MDIX
- Port speed downshift and selective auto-negotiation
- Dynamic Voice VLAN Assignment
- Jumbo Frames up to 10,240 bytes for 10/100/1000 and 10GbE ports
- IGMP Snooping (v1/v2/v3)
- MLD Snooping (v1/v2)
- PIM-SM Snooping
- Private VLANs and uplink-switch
- Protected Link Groups
- Port Loop Detection
- VLAN based Static MAC Denial
- Flexible static multicast MAC address configuration

Layer 2 Metro features

- VLAN stacking (Q-in-Q)
- Metro Ring Protocol (MRP I and II)
- Virtual Switch Redundancy Protocol
- Topology Groups
- Super Aggregated VLANs (SAV)

Base Layer 3 features

- Virtual Interfaces (VE)
- Routed Interfaces
- IPv4 Static Routes
- Routing between directly connected subnets
- RIP v1/v2 announce
- Virtual Route Redundancy Protocol (VRRP)
- ECMP (B50G only)

Edge Layer 3 features (B48G upgrade only)

- Host routes
- OSPF
- RIP V1 , V2
- Route-only support
- Routes in hardware maximum: 1000

Quality of Service (QoS)

- MAC Address Mapping to Priority Queue
- ACL Mapping to Priority Queue
- ACL Mapping to ToS/DSCP
- Honoring DSCP and 802.1p
- ACL Mapping and Marking of ToS/DSCP
- DiffServ Support
- Classifying and Limiting Flows based on TCP flags
- DHCP Relay
- QoS Queue Management using Weighted Round Robin (WRR), Strict Priority (SP), and a combination of WRR and SP

Traffic management

- Inbound rate limiting per port
- ACL-based inbound rate limiting and traffic policies
- Outbound rate limiting per port and per queue
- Broadcast, multicast and unknown unicast rate limiting

Element security options

- AAA
- RADIUS
- Secure Shell (SSH v2)
- Secure Copy (SCP v2)
- TACACS/TACACS+
- Username/Password (Challenge and Response)
- Bi-level Access Mode (Standard and EXEC Level)

Network management and monitoring
► IronView Network Manager (INM) Web-based graphical user interface ► IBM Tivoli NetCOOL ► IBM Systems Director ► Integrated industry standard Command Line Interface (CLI) ► sFlow (RFC 3176) ► Telnet ► SNMP v1, v2c, v3 ► SNMP MIB II ► RMON

For a complete list, see the *FastIron Configuration Guide* for the software release in which you are interested. For a summarized list similar to the preceding table, see the *IBM g-series Ethernet Switches Data Sheet*.

4.7 Features matrix

In a 3-tier network design, the following terms can be used somewhat interchangeably:

- Access/Edge: Switches to which end devices connect are listed here:
  - Data Center (Access): Typically 1 RU / 2 RU switches, sometimes called “Top of Rack.” Denser, modular chassis can be used at the Access Layer, sometimes called “End of Row” or “Middle of Row.” End devices are typically servers or iSCSI storage devices. You might be looking for port-level security features, the ability to easily shape your broadcast domains, and traffic management features to ensure your high priority servers receive adequate bandwidth.
  - Enterprise Campus (Edge): Typically 1 RU / 2 RU switches, but sometimes denser, modular chassis can be used. These switches can be housed in wiring closets distributed on different floors within a building. A few examples of end devices are desktops, laptops, wireless access points, VoIP phones, and surveillance cameras. You might be looking for Power over Ethernet, auto-configuration of end devices with LLDP-MED or CDP, and user-level security AAA options such as 802.1x, RADIUS/TACACS+, or built-in Web Authorization servers.

- Aggregation/Distribution: Provides connectivity between Access/Edge switches. Typically higher bandwidth ports are required in order to provide acceptable bandwidth:
  - Data Center (Aggregation) and Enterprise Campus (Distribution): Switches here have essentially the same functionality, providing connectivity to the various edge switches in the network. Denser, modular chassis are typically used, with higher bandwidth ports to support “uplinks” from the Access/Edge and uplinks to the Core.
  - IP functionality such as Layer 3 routing might be needed at this layer. You might be looking to support a wide variety of IPv4 and IPv6 unicast and multicast protocols, the ability to virtualize your Layer 3 routing tables with Virtual Routing and Forwarding, and network-wide security and traffic management features.
- Core/Backbone: Provides connectivity over distance and/or outside the internal network:
  - Data Center (Core): Provides connectivity out to a WAN link, either an ISP or over leased lines to other Data Centers or to the Enterprise’s Campus networks. Advanced routing protocols such as BGP and services such as MPLS might be required at this layer. You might also be looking to implement secure Virtual Private Networks, tunneling, and advanced Traffic Engineering and monitoring features.
  - Enterprise Campus (Backbone): Backbone typically denotes a lighter-weight “core” that might not need to hold as many routes when peering to an ISP or back to the Data Center.

In a 2-tier design, the Core/Aggregation can be collapsed into a higher port density chassis, or Access/Aggregation collapsed. These are all possibilities with the high-performance characteristics of the IBM b-type DCN product line.

4.7.1 Modular routers and switches

Table 4-7 is a feature matrix based on the following configurations:

- IBM m-series running Multi-Service IronWare R04.0.00f+:

  All features are included in base software and denoted with a “X”.

- IBM r-series running Multi-Service IronWare for r-series R02.7.02a+:

  All features are included in base software and denoted with a “X”.

- IBM s-series running IronWare R05.1.00+:
  - “L2” denotes features supported in Layer 2 Image.
  - “BL3” denotes features supported in Base Layer 3 Image. Contains all features found in Layer 2 Image.
  - “FL3-IPv4” denotes features supported in IPv4 Full Layer 3 Image. Contains all features found in Base Layer 3 Image.
  - “FL3-IPv6” denotes features supported in IPv6/IPv4 Full Layer 3 Image. Contains all features found in IPv4 Full Layer 3 Image.

For the most up-to-date list and additional information, see the *Release Notes* and *Software Configuration Guides* for the appropriate software release.

Table 4-7 Modular routers and switches matrix

Category: Management features	**Feature description**	**IBM m-series**	**IBM r-series**	**IBM s-series**
Industry-standard Command Line Interface (CLI)	► Access by Serial (console) port, Telnet, and SSHv2 ► Two-tier authentication (Enable and Privileged EXEC)	X	X	L2
Web-based Graphical User Interface (GUI)	► Access by HTTP and HTTPS ► Device-level configuration	X	X	L2
Brocade Ironview Network Manager	► Network-wide GUI management software offered by Brocade	X	X	L2
Configuration file management	► Copy configuration files by TFTP and SCP	X	X	L2
User Accounts	► Setup individual user accounts	X	X	L2
Simple Network Management Protocol (SNMP) v1, v2 and v3	► Supports wide variety of Standard (Industry) and Enterprise (Proprietary) MIBs	X	X	L2

Category: Management features	Feature description	IBM m-series	IBM r-series	IBM s-series
Authentication, Authorization, and Accounting (AAA)	► RADIUS, TACACS, TACACS+ ► Login, Privileged EXEC, and CONFIG level Authentication by RADIUS, TACACS, TACACS+ ► Command Authorization by RADIUS, TACACS+ ► Accounting by RADIUS, TACACS+	X	X	L2
sFlow version 5	► Hardware-based, port-level, inbound packet sampling ► IPv6 support	X	X	L2
Remote Network MONitoring (RMON)	► Statistics (RMON Group 1) ► History (RMON Group 2) ► Alarms (RMON Group 3) ► Events (RMON Group 9)	X	X	L2

Category: Security features	Feature description	IBM m-series	IBM r-series	IBM s-series
Local passwords	► Supports local passwords ► Supports local passwords tied to user accounts	X	X	L2
802.1x Port Security	► Authenticate devices connected to a port ► Multiple Extensible Authentication Protocol (EAP) supported: MD5, TLS, TTLS, and PEAP ► Supports RADIUS Authentication ► Dynamic ACL, MAC filter, and VLAN assignment	X	X	L2
Multi-Device Port Authentication	► Authenticate devices connected to a port based on MAC address with RADIUS ► Block or forward traffic to restricted VLAN ► Dynamic VLAN assignments	X	X	L2
Multi-Device Port Authentication	► Dynamic ACL assignments			L2

Category: Security features	Feature description	IBM m-series	IBM r-series	IBM s-series
Media Access Control (MAC) Port Security	► Define a number of secure MAC addresses for an interface ► Interface permits only traffic with a source MAC address from the secure MAC address list	X	X	L2
Address Locking	► Limit number of devices that have access to a specific port		X	L2
MAC Filtering [Description pertains to selected platforms]	► Deny/permit traffic based on source and destination MAC address at an interface level ► Note: MAC Filtering for other devices might have different implications.			L2
MAC Filtering bypass of 802.1X [Description pertains to selected platforms]	► Define MAC filters that allow these devices to bypass 802.1X security	X	X	L2
Access Control Lists (ACLs): Layer 2	► MAC Address (Source & Destination) ► Ethertype ► VLAN ► ACLs can be applied as rules for permitting/denying traffic, traffic shaping, management access, and other features	X	X	
ACLs: Layer 3	► Standard: source IP address ► Extended: source & destination IP address, IP protocol information ► Logging of Denied packets ► Inbound packets only ► ACLs can be applied as rules for permitting/denying traffic, traffic shaping, management access, and other features	X	X	L2
ACLs for Outbound Packets	► Selected Layer 2 ACLs ► Standard and Extended Layer 3 ACLs	X		
Bridge Protocol Data Unit (BPDU) Guard	► Layer 2 Protection ► Prevents end devices from participating in STP	X	X	L2

Category: Security features	Feature description	IBM m-series	IBM r-series	IBM s-series
Root Guard	► Layer 2 Protection ► Enforces root bridge placement within the network	X	X	L2
Denial of Service (DoS) Protection	► Layer 3 Protection ► Protection from TCP SYN attacks ► Protection from ICMP (Smurf) attacks	X	X	L2
Dynamic Host Configuration Protocol (DHCP) Snooping	► Layer 3 Protection ► Intercept and filter untrusted DHCP packets and ward off Man-In-the-Middle type attacks.	X	X	L2
Dynamic Address Resolution Protocol (ARP) Inspection	► Layer 3: Intercept and filter packets with invalid IP to MAC address bindings and wards off Man-In-the-Middle type attacks such as ARP cache poisoning.	X	X	L2
IP Source Guard	► Layer 3 Protection ► Permit only DHCP packets and traffic from valid IP addresses	X	X	L2
CLI Logging	► Provides logging of all valid CLI commands from each user session into the system log	X	X	
Syslog	► Support multiple Syslog server logging	X	X	L2
Advanced Encryption Standard (AES) Support	► SNMPv3, SSHv2	X	X	L2
AES Support	► HTTPS, SCP	X	X	
Secure Management Protocols	► Secure Shell (SSH) version 2 ► Secure Copy (SCP) ► HTTPS	X	X	L2
Default Management Settings	► Telnet, SSH, HTTP, HTTPS, and SNMP servers are disabled by default	X	X	

Category: Security features	Feature description	IBM m-series	IBM r-series	IBM s-series
Web Authentication	► Port-based authentication that re-directs user web browser over HTTP or HTTPS to authentication page ► User authenticates using username/password, passcode, or none needed which is checked against local database or RADIUS server ► Device MAC address is then either permited until a trigger event or denied access to the switch			L2

Category: System features	Feature description	IBM m-series	IBM r-series	IBM s-series
Power over Ethernet (PoE)	► Class 3, 15.4 Watts per port			L2
Jumbo Frames	► Supported: 9,216 bytes	X	X	L2
Cisco Discovery Protocol (CDP) and Foundry Discovery Protocol (FDP)	► Allows configuration of a device to intercept and display CDP and FDP packets. ► This feature is useful for learning device and interface information in the network.	X	X	L2
High Availability	► Hitless Software Upgrade ► Hitless Layer 2 Switchover	X	X	L2
High Availability	► Hitless Layer 3 Failover (BGP and OSPF)	X		
802.1ag	► Connectivity Fault Management (CFM) enables ability to monitor health of service being delivered over the network	X		
Port Mirroring & Monitoring ► r-series: Port Mirror not supported on 16x10 GbE Interface Module (Port Monitor is supported)	► Basic: Mirrors traffic from a monitor port; supports inbound and outbound mirroring ► ACL-based: Mirrors traffic from a monitor port and filters by applying an ACL	X	X	L2

Category: System features	Feature description	IBM m-series	IBM r-series	IBM s-series
Link Layer Discovery Protocol (LLDP)	► IEEE 802.1ab: Layer 2 network discovery protocol enabling a device to advertise its capabilities and discover other LLDP-enabled devices		X	L2
Link Layer Discovery Protocol Media Endpoint Devices (LLDP-MED)	► Layer 2 network discovery protocol based on ANSI/TIA-1057 standard enabling devices to configure Media Endpoint devices such as IP telephones and security cameras			L2
Voice over IP Autoconfiguration	► Support autoconfiguration of VoIP using LLDP-MED and CDP			L2
Multi-port Static MAC Addresses	► Enable configuration of the same static MAC address to multiple ports ► Helps support some load balancing devices that announce an identical MAC address on multiple ports			L2
Port Flap Dampening	► Increase network stability by preventing a port from enabling immediately after it goes down	X	X	L2
Simple Network Time Protocol (SNTP)	► Supports specifying a SNTP server	X	X	L2
Port Configuration	► Auto-sensing: Automatically detects duplex settings ► Auto-negotiation: Automatically detects speed settings ► Auto-MDI/MDIX: Automatically detects cable connection type (straight-through vs crossover)	X	X	L2
Hardware Monitoring and Control	► Optical Monitoring of supported transceivers ► Temperature polling and automatic shutdown of device past unsafe limits	X	X	L2

Category: System features	Feature description	IBM m-series	IBM r-series	IBM s-series
Hardware Monitoring and Control	► Link Fault Signaling (LFS) for 10 GbE ports: Enables diagnostics on 10 GbE links ► Remote Fault Notification (RFN) on 1 GbE connections: Local port notifies remote port to disable itself when local port is disabled ► Virtual Cable Test (VCT): Enables diagnostics of copper wires			L2
Uni-Directional Link Detection (UDLD)	► Monitors a link between two devices and brings the ports on both ends of the link down if the link goes down at any point between the two devices	X	X	L2
Central Processing Unit (CPU) Protection	► Enahnces the efficiency of the CPU on an interface module and protects it from excessive amounts of network traffic.	X	X	L2
Protected Link Groups	► Minimized network disruption by specifying Standby links for an Active link carrying traffic ► Supported on 1 GbE ports only ► Superseded by Link Aggregation			L2

Category: Traffic Management features	Feature description	IBM m-series	IBM r-series	IBM s-series
802.1p	► Layer 2 Class of Service (CoS) for traffic prioritzation on 802.1q tagged frames utilizing 3-bit field (8 priority queues)	X	X	L2
Differentiated Service Codepoint (DSCP) / DiffServ	► Layer 3 class-based traffic prioritization utilizing 6-bit DSCP field (64 values) that can be mapped to a Per-Hop Behavior	X	X	L2

Category: Traffic Management features	Feature description	IBM m-series	IBM r-series	IBM s-series
Rate Limiting: s-series: Fixed rate limiting not supported on 10 GbE ports or tagged ports running Full L3 image	► Drops traffic on a port exceeding a configured threshold and criteria ► Applies to inbound ports ► Port-based: Defines a fixed threshold on a port ► Port-and-ACL-based: Defines a threshold for a port based on traffic matching an ACL ► Multicast, Broadcast, and Unknown-Unicast traffic for a port: Defines a threshold fir a oirt based on these traffic types	X	X	L2
Rate Limiting: m-series supports Inbound and Outbound port Rate Limiting for all policies	► Drops traffic on a port exceeding a configured threshold and criteria ► Outbound-Port-based: Defines a fixed threshold for an outbound port ► Port-and-priority based: Defines a threshold for a port and a particular priority queue ► Port-and VLAN-based- Defines a threshold a port and within a VLAN	X	X	
Traffic Shaping	► Outbound Rate Shaping at the port level allows packets to be stored in buffers if traffic exceeds a configured threshold			L2
Weighted Random Early Discard (WRED)	► Proactively monitors congestion and drops packets in a weighted manner taking into account priority if buffers start filling up	X	X	

Category: Traffic Management features	Feature description	IBM m-series	IBM r-series	IBM s-series
Traffic Scheduling	► The following scheduling schemes are supported: – Weighted Round Robin (WRR): All queues are serviced in a cycle but number of packets serviced for a particular queue based on queue weight – Strict Priority (SP): Assigns maximum weights to each queue, biasing higher queues over lower queues – Hybrid WRR and SP: Allow SP applied to specific queues and WRR to rest of queues			L2
Traffic Scheduling	► The following scheduling schemes are supported: – Weighted Fair Queuing (WFQ): Some weight-based bandwidth is allocated to all queues – Strict Priority (SP) scheduling: Ensures higher-priority traffic always serviced first – Mixed WFQ and SP: Provides SP for three highest priority queues and WFQ for remaining priority queues	X	X	

Category: Layer 2 features	Feature description	IBM m-series	IBM r-series	IBM s-series
802.1D	► Spanning Tree Protocol (STP): By default a separate instance of STP is run per VLAN if STP is enabled ► Single Spanning Tree Protocol (SSTP): Runs a single instance of STP across all VLANs and ports	X	X	L2
802.1w	► Rapid Spanning Tree Protocol (RSTP): By default, a separate instance of RSTP is run per VLAN if RSTP is enabled ► Single Spanning Tree Protocol (SSTP): Runs a single instance of RSTP across all VLANs and ports	X	X	L2

Category: Layer 2 features	Feature description	IBM m-series	IBM r-series	IBM s-series
802.1s	► Multiple Spanning Tree Protocol (MSTP): Allows mapping of several VLANs to a single spanning tree instance	X	X	L2
SuperSpan	► A STP enhancement that allows Service Providers (SPs) to use STP in both SP networks and customer networks	X	X	
Virtual Switch Redundancy Protocol (VSRP)	► Proprietary protocol and alternative to STP and Virtual Router Redundancy Protocol Extended (VRRPE) that provides redundancy and sub-second failover in Layer 2 and Layer 3 mesh topologies ► Based on VRRPE	X	X	L2
Metro Ring Protocol (MRP) Phase 1 / Phase 2	► Proprietary protocol and alternative to STP providing fast reconvergence in Layer 2 ring topologies	X	X	L2
Per-VLAN Spanning Tree (PVST) / PVST+	► Supports interoperability with Cisco proprietary protocol	X	X	L2
Topology Groups	► A named set of VLANs that share a Layer 2 topology ► Compatible with the following Layer 2 protocols: – 802.1D Spanning Tree Protocol – 802.1w Rapid Spanning Tree Protocol – Metro Ring Protocol – Virtual Switch Redundancy Protocol	X	X	L2
802.1q	► VLAN Tagging ► Allows traffic from multiple VLANs to traverse over the same network link ► VLAN Tag-type Translation	X	X	L2
802.1q-in-q	► Supports attaching a second 802.1q tag to an already tagged frame	X	X	L2

Category: **Layer 2 features**	**Feature description**	**IBM m-series**	**IBM r-series**	**IBM s-series**
VLANs	► Layer 2 Port-based VLANs ► Layer 3 Protocol-based (IPv4, dynamic IPv6, AppleTalk, IPX) VLANs ► Dual-mode VLAN ports: Allow untagged and tagged traffic through the same port ► Super Aggregated VLANs (SAV): Allow multiple VLANs within a VLAN to construct Layer 2 Paths and Channels ► Uplink Ports within a Port-based VLAN: Send all broadcast and unknown-unicast traffic to uplink port	X	X	L2
802.3ad	► Link Aggregation Control Protocol (LACP): Enables dynamic creation of Link Aggregation Groups (LAGs) ► Allows multiple load-sharing links between two devices	X	X	L2
Static Link Aggregation Groups (Trunk Groups)	► Manually configured LAGs / trunk groups ► Allows multiple load-sharing links between two devices ► Compatible with Cisco EtherChannel	X	X	L2
Multicast (Layer 2 traffic reduction)	► Internet Group Management Protocol (IGMP) v1, v2, v3 snooping ► IGMP snooping per VLAN ► IGMP Fast Leave ► Protocol Independent Multicast-Sparse Mode (PIM-SM) v2 snooping	X	X	L2
Multicast (Layer 2 traffic reduction)	► MLD v1/v2 Snooping ► MLDv1 Fast Leave			L2
Dynamic Host Configuration Protocol (DHCP) Assist	► Ensures that a DHCP server that manages multiple IP subnets can recognize the requester's IP subnet			L2

Category: Layer 3 features	Feature description	IBM m-series	IBM r-series	IBM s-series
Static IPv4 Routes	► Manually configured IPv4 routes ► Adress Resolution Protocol (ARP) entries	X	X	BL3
Routing Information Protocol (RIP) v1/v2 Advertising	► Advertises directly connected routes but does not learn RIP routes	X	X	BL3
RIPv1/v2	► IPv4 dynamic, distance-vector routing protoco ► Interior Gateway Protocol (IGP)l	X	X	FL3-IPv4
Open Shortest Path First (OSPF) v2	► IPv4 dynamic, link-state routing protocol ► Interior Gateway Protocol (IGP) ► MD5 authentication	X	X	FL3-IPv4
OSPFv2	► Graceful restart: Increases network resiliency by minimizing disruptions in forwarding during a router restart ► BiDirectional Forwarding Detection (BFD): Enables rapid detection of the failure on a forwarding path	X	X	
Intermediate System to Intermediate System (IS-IS)	► IPv4, dynamic link-state routing protocol ► Interior Gateway Protocol (IGP) ► MD5 authentication ► BiDirectional Forwarding Detection (BFD): Enables rapid detection of the failure on a forwarding path	X	X	
Border Gateway Protocol (BGP)-4	► IPv4 dynamic, path-vector protocol ► Exterior Gateway Protocol (EGP) ► MD5 authentication ► Equal-Cost Multiple-Path load sharing ► Confederations: sub-dividing an Autonomous System (AS) ► Aggregate route advertisement ► IP Address filtering ► Route reflectors ► Route flap dampening ► Fast external failover	X	X	FL3-IPv4
BGP-4	► Graceful restart ► Multi-hop eBGP	X	X	

Category: Layer 3 features	Feature description	IBM m-series	IBM r-series	IBM s-series
BGP-4	► 4-byte Autonomous System Numbers (ASN)	X		
Route Redistribution (IPv4)	► Distribute IPv4 routes between directly connected, static, and dynamically learned routes	X	X	FL3-IPv4
Static IPv6 Routes	► Manually configured IPv6 routes	X	X	FL3-IPv6
RIP next generation (RIPng)	► IPv6 dynamic, distance-vector routing protocol ► Interior Gateway Protocol (IGP)	X	X	FL3-IPv6
OSPFv3	► IPv6 dynamic, link state routing protocol ► Interior Gateway Protocol (IGP)	X	X	FL3-IPv6
IS-IS for IPv6	► IPv6, dynamic link-state routing protocol ► Interior Gateway Protocol (IGP)	X		
BGP4+	► IPv6 dynamic, path-vector protocol ► Exterior Gateway Protocol (EGP)	X	X	
Route Redistribution (IPv6)	► Distribute IPv4 routes between directly connected, static, and dynamically learned routes	X	X	FL3-IPv6
Multicast Routing (IPv4)	► IGMP v1, v2, v3 ► PIM-SM v1/v2 ► PIM-Dense Mode (PIM-DM) v2 ► Distance Vector Multicast Routing Protocol (DVMRP) ► PIM Anycast RP ► Multicast Source Discovery Protocol (MSDP)	X	X	FL3-IPv4
Multicast Routing (IPv4)	► PIM-Source Specific Mode (PIM-SSM) v4 ► Passive Multicast Route Insertion (PMRI)	X	X	
Multicast Routing (IPv6)	► Multicast Listener Discovery (MLD) v2 ► IPv6 PIM-SM	X	X	

Category: Layer 3 features	Feature description	IBM m-series	IBM r-series	IBM s-series
Multicast Routing (IPv6)	► IPv6 PIM-SSM ► IPv6 PIM Anycast RP ► IPv6 PMRI	X		
Virtual Router Redundancy Protocol (VRRP) / VRRP Extended (VRRPE)	► Provides redundant paths to routers within a LAN ► VRRPE utilizes a Master and Backup routers, from which the Master listens to a virtual IP address	X	X	FL3-IPv4
Multi-Virtual Routing and Forwarding (Multi-VRF)	► Sometimes known as “Multi-VRF CE” or “VRF-Lite” ► Enables multiple virtual routing/forwarding tables to be maintained on a single device	X		
Virtual Routing Interfaces	► Integrated Switch Routing (ISR): Enables a logical routing interface enabling Layer 3 routing of traffic between VLANs eliminating the need for an external router	X	X	BL3
Generic Routing Encapsulation (GRE) Tunnels	► IPv4 Point-to-Point ► GRE Keepalive	X	X	BL3
DHCP and BootP Relay	► Enable Layer 3 forwarding and return of DHCP and BootP requests from devices	X	X	BL3
ICMP Redirect Messages	► If traffic is misdirected notifies source and forwards traffic to appropriate router	X	X	FL3-IPv4
Policy Based Routing (PBR)	► Use ACLs and route maps to selectively modify and route IP packets in hardware	X	X	FL3-IPv4
Route-only Support	► Disable Layer 2 switching at a global or interface level	X	X	FL3-IPv4
IPv6-over-IPv4 Tunnels	► Create point-to-point IPv6 over IPv4 tunnels to move traffic through IPv6 domains	X	X	FL3-IPv6
Brocade Direct Routing (BDR)	► Hardware-based routing with configurable Content Addressable Memory (CAM) resource settings	X	X	

Category: Virtual Private Networks (VPN) features	**Feature description**	**IBM m-series**	**IBM r-series**	**IBM s-series**
Layer 2 VPN	► Virtual Private LAN Service (VPLS) – Multicast Snooping for VPLS – IGMP and PIM Proxy for VPLS – Disabling of VPLS Local switching – Local VPLS (without MPLS) ► BGP auto-discovery ► Virtual Leased Line (VLL) ► Local VLL ► Single tag ► Double tag	X		
Layer 3 VPN	► BGP/MPLS VPNs ► Multi-Virtual Routing and Forwarding ► Per-VRF VRRPE ► OSPF Sham link support ► Static Routes Across VRFs	X		

Category: Multiprotocol Label Switching (MPLS) features	**Feature description**	**IBM m-series**	**IBM r-series**	**IBM s-series**
MPLS	► Label Distribution Protocol (LDP) ► Resource Reservation Protocol-Traffic Engineering (RSVP-TE) ► CSPF ► OSPF-TE ► ISIS-TE ► Fast Re-route (Detour/Bypass) ► LSP Accounting ► Adaptive Labeled Switched Paths (LSPs) ► BiDirectional Forwarding Detection for RSVP-TE	X		

For the most up-to-date list, see the software *Release Notes* and the *NetIron Configuration Guide*, *BigIron RX Series Configuration Guide*, and the *FastIron Configuration Guide* for the software version in which you are interested.

4.7.2 Fixed port switches

Table 4-8 is a feature matrix based on the following configurations:

- IBM c-series running Multi-Service IronWare R03.9.00+
  - “Base” denotes features supported with Base Licensing.
  - “FL3” denotes features supported with Full Layer 3 Licensing.
  - “Metro” denotes features supported with Metro Licensing.
- IBM x-series running IronWare R04.1.00c+
  - “L2” denotes features supported in Layer 2 Image.
- IBM g-series: IBM Ethernet Switch B48G running IronWare R04.3.02a+ and IBM Ethernet Switch B50G running IronWare R05.0.02a+
  - “L2” denotes features supported in Layer 2 Image.
  - “BL3” denotes features supported in Base Layer 3 Image. Contains all features found in Layer 2 Image.
  - “EL3” denotes features supported in Edge Layer 3 Image; IBM Ethernet Switch B48G only. Contains all features found in Base Layer 3 Image.

For the most up-to-date list and additional information, see the *Release Notes* and *Software Configuration Guides* for the appropriate software release.

Table 4-8 Fixed port switches matrix

Category: Management features	**Feature description**	**IBM c-series**	**IBM x-series**	**IBM B48G**	**IBM B50G**
Industry-standard Command Line Interface (CLI)	► Access by Serial (console) port, Telnet, and SSHv2 ► Two-tier authentication (Enable and Privileged EXEC)	Base	L2	L2	L2
Web-based Graphical User Interface (GUI)	► Access by HTTP and HTTPS ► Device-level configuration			L2	L2
Brocade Ironview Network Manager	► Network-wide GUI management software offered by Brocade	Base		L2	
Configuration file management	► Copy configuration files by TFTP and SCP	Base	L2	L2	L2
User Accounts	► Setup individual user accounts	Base	L2	L2	L2

Category: Management features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Simple Network Management Protocol (SNMP) v1, v2 and v3	► Supports wide variety of Standard (Industry) and Enterprise (Proprietary) MIBs	Base	L2	L2	L2
Authentication, Authorization, and Accounting (AAA)	► RADIUS, TACACS, TACACS+ ► Login, Privileged EXEC, and CONFIG level Authentication by RADIUS, TACACS, TACACS+ ► Command Authorization by RADIUS, TACACS+ ► Accounting by RADIUS, TACACS+	Base	L2	L2	L2
sFlow version 5	► Hardware-based, port-level, inbound packet sampling	Base	L2	L2	L2
Remote Network MONitoring (RMON)	► Statistics (RMON Group 1) ► History (RMON Group 2) ► Alarms (RMON Group 3) ► Events (RMON Group 9)	Base	L2	L2	L2

Category: Security features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Local passwords	► Supports local passwords ► Supports local passwords tied to user accounts	Base	L2	L2	L2
802.1x Port Security	► Authenticate devices connected to a port ► Multiple Extensible Authentication Protocol (EAP) supported: MD5, TLS, TTLS, and PEAP ► Supports RADIUS Authentication ► Dynamic ACL, MAC filter, and VLAN assignment	Base		L2	L2
Multi-Device Port Authentication	► Authenticate devices connected to a port based on MAC address with RADIUS ► Block or forward traffic to restricted VLAN ► Dynamic VLAN assignments	Base		L2	L2

Category: Security features	**Feature description**	**IBM c-series**	**IBM x-series**	**IBM B48G**	**IBM B50G**
Media Access Control (MAC) Port Security	► Define a number of secure MAC addresses for an interface ► Interface permits only traffic with a source MAC address from the secure MAC address list	Base		L2	L2
Address Locking	► Limit number of devices that have access to a specific port			L2	L2
MAC Filtering [Description pertains to selected platforms]	► Deny/permit traffic based on source and destination MAC address at an interface level ► Note: MAC Filtering for other devices might have different implications.		L2	L2	L2
MAC Filtering bypass of 802.1X [Description pertains to selected platforms]	► Define MAC filters that allow these devices to bypass 802.1X security	Base		L2	L2
Access Control Lists (ACLs): Layer 2	► MAC Address (Source & Destination) ► Ethertype ► VLAN ► ACLs can be applied as rules for permitting/denying traffic, traffic shaping, management access, and other features	Base			
ACLs: Layer 3	► Standard: source IP address ► Extended: source & destination IP address, IP protocol information ► Inbound packets only ► ACLs can be applied as rules for permitting/denying traffic, traffic shaping, management access, and other features	Base	L2	L2	L2
ACLs: Layer 3	► Logging of Denied packets	Base	L2		
Bridge Protocol Data Unit (BPDU) Guard	► Layer 2 Protection ► Prevents end devices from participating in STP	Base	L2	L2	L2

Category: Security features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Root Guard	► Layer 2 Protection ► Enforces root bridge placement within the network	Base	L2	L2	L2
Denial of Service (DoS) Protection	► Layer 3 Protection ► Protection from TCP SYN attacks ► Protection from ICMP (Smurf) attacks	Base		L2	L2
Dynamic Host Configuration Protocol (DHCP) Snooping	► Layer 3 Protection ► Intercept and filter untrusted DHCP packets and ward off Man-In-the-Middle type attacks.	Base		L2	
Dynamic Address Resolution Protocol (ARP) Inspection	► Layer 3: Intercept and filter packets with invalid IP to MAC address bindings and wards off Man-In-the-Middle type attacks such as ARP cache poisoning.	Base		L2	
IP Source Guard	► Layer 3 Protection ► Permit only DHCP packets and traffic from valid IP addresses	Base		L2	L2
CLI Logging	► Provides logging of all valid CLI commands from each user session into the system log	Base			
Syslog	► Support multiple Syslog server logging	Base	L2	L2	L2
Advanced Encryption Standard (AES) Support	► SSHv2		L2	L2	L2
AES Support	► SNMPv3	Base		L2	
Secure Management Protocols	► Secure Shell (SSH) version 2 ► Secure Copy (SCP)	Base	L2	L2	L2
Secure Management Protocols	► HTTPS			L2	L2
Default Management Settings	► Telnet, SSH, HTTP, HTTPS, and SNMP servers are disabled by default	Base			

Category: System features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Power over Ethernet (PoE)	► Class 3, 15.4 Watts per port			L2	L2
Jumbo Frames	► Supported: 9,216 bytes	Base	L2	L2	L2
Cisco Discovery Protocol (CDP) and Foundry Discovery Protocol (FDP)	► Allows configuration of a device to intercept and display CDP and FDP packets. ► This feature is useful for learning device and interface information in the network.	Base	L2	L2	L2
802.1ag	► Connectivity Fault Management (CFM) enables ability to monitor health of service being delivered over the network	Metro			
Port Mirroring & Monitoring	► Basic: Mirrors traffic from a monitor port: supports inbound and outbound mirroring ► ACL-based: Mirrors traffic from a monitor port and filters by applying an ACL	Base	L2	L2	L2
Port Mirroring & Monitoring	► MAC filter-based: Mirrors traffic from a monitor port and filters by applying a MAC filter		L2	L2	L2
Link Layer Discovery Protocol (LLDP)	► IEEE 802.1ab: Layer 2 network discovery protocol enabling a device to advertise its capabilities and discover other LLDP-enabled devices		L2	L2	L2
Link Layer Discovery Protocol Media Endpoint Devices (LLDP-MED)	► Layer 2 network discovery protocol based on ANSI/TIA-1057 standard enabling devices to configure Media Endpoint devices such as IP telephones and security cameras			L2	L2
Voice over IP Autoconfiguration	► Support autoconfiguration of VoIP using LLDP-MED and CDP			L2	L2

Category: System features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Rate Limiting: x-series: Fixed rate limiting not supported on 10 GbE ports or tagged ports running Full L3 image	► Drops traffic on a port exceeding a configured threshold and criteria ► Fixed: Defines a fixed threshold on a port ► ACL-based: Defines a threshold based on traffic matching an ACL ► Multicast, Broadcast, and Unknown-Unicast traffic: Defines a threshold based on these traffic types	Base	L2	L2	L2
Rate Limiting	► VLAN-based: Drops traffic in a VLAN exceeding a configured threshold on a port	Base			
Multi-port Static MAC Addresses	► Enable configuration of the same static MAC address to multiple ports ► Helps support some load balancing devices that announce an identical MAC address on multiple ports		L2	L2	L2
Port Flap Dampening	► Increase network stability by preventing a port from enabling immediately after it goes down	Base	L2	L2	L2
Simple Network Time Protocol (SNTP)	► Supports specifying a SNTP server	Base	L2	L2	L2
Port Configuration	► Auto-sensing: Automatically detects duplex settings ► Auto-negotiation: Automatically detects speed settings ► Auto-MDI/MDIX: Automatically detects cable connection type (straight-through vs crossover)	Base	L2	L2	L2
Hardware Monitoring and Control	► Optical Monitoring of supported transceivers ► Temperature polling and automatic shutdown of device past unsafe limits	Base	L2	L2	L2

Category: System features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Hardware Monitoring and Control	► Link Fault Signaling (LFS) for 10 GbE ports: Enables diagnostics on 10 GbE links ► Remote Fault Notification (RFN) on 1 GbE connections: Local port notifies remote port to disable itself when local port is disabled ► Virtual Cable Test (VCT): Enables diagnostics of copper wires			L2	L2
Uni-Directional Link Detection (UDLD)	► Monitors a link between two devices and brings the ports on both ends of the link down if the link goes down at any point between the two devices	Base	L2	L2	L2
Central Processing Unit (CPU) Protection	► Enahnces the efficiency of the CPU on an interface module and protects it from excessive amounts of network traffic.	Base			
Protected Link Groups	► Minimized network disruption by specifying Standby links for an Active link carrying traffic ► Supported on 1 GbE ports only ► Superseded by Link Aggregation			L2	L2

Category: Traffic Management features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
802.1p	► Layer 2 Class of Service (CoS) for traffic prioritzation on 802.1q tagged frames utilizing 3-bit field (8 priority queues)	Base	L2	L2	L2
Differentiated Service Codepoint (DSCP) / DiffServ	► Layer 3 class-based traffic prioritization utilizing 6-bit DSCP field (64 values) that can be mapped to a Per-Hop Behavior	Base	L2	L2	L2

Category: Traffic Management features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Rate Limiting: Fixed rate limiting not supported on 10 GbE ports or tagged ports running L3 images	► Drops traffic on a port exceeding a configured threshold and criteria ► Applies to inbound ports ► Port-based: Defines a fixed threshold on a port ► Port-and-ACL-based: Defines a threshold for a port based on traffic matching an ACL ► Multicast, Broadcast, and Unknown-Unicast traffic for a port: Defines a threshold fir a oirt based on these traffic types	Base	L2	L2	L2
Rate Limiting	► Drops traffic on a port exceeding a configured threshold and criteria ► Port-and-priority based: Defines a threshold for a port and a particular priority queue			L2	L2
Traffic Shaping	► Outbound Rate Shaping at the port level allows packets to be stored in buffers if traffic exceeds a configured threshold		L2		
Traffic Scheduling	► The following scheduling schemes are supported: – Weighted Round Robin (WRR): All queues are serviced in a cycle but number of packets serviced for a particular queue based on queue weight – Strict Priority (SP): Assigns maximum weights to each queue, biasing higher queues over lower queues – Hybrid WRR and SP: Allow SP applied to specific queues and WRR to rest of queues		L2	L2	L2

Category: Traffic Management features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Traffic Scheduling	► The following scheduling schemes are supported: – Weighted Fair Queuing (WFQ): Some weight-based bandwidth is allocated to all queues – Strict Priority (SP) scheduling: Ensures higher-priority traffic always serviced first – Mixed WFQ and SP: Provides SP for three highest priority queues and WFQ for remaining priority queues	Base			

Category: Layer 2 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
802.1D	► Spanning Tree Protocol (STP): By default a separate instance of STP is run per VLAN if STP is enabled ► Single Spanning Tree Protocol (SSTP): Runs a single instance of STP across all VLANs and ports	Base	L2	L2	L2
802.1w	► Rapid Spanning Tree Protocol (RSTP): By default, a separate instance of RSTP is run per VLAN if RSTP is enabled ► Single Spanning Tree Protocol (SSTP): Runs a single instance of RSTP across all VLANs and ports	Base	L2	L2	L2
802.1s	► Multiple Spanning Tree Protocol (MSTP): Allows mapping of several VLANs to a single spanning tree instance	Base	L2	L2	L2

Category: Layer 2 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Virtual Switch Redundancy Protocol (VSRP)	► Proprietary protocol and alternative to STP and Virtual Router Redundancy Protocol Extended (VRRPE) that provides redundancy and sub-second failover in Layer 2 and Layer 3 mesh topologies ► Based on VRRPE	Base		L2	L2
Metro Ring Protocol (MRP) Phase 1 / Phase 2	► Proprietary protocol and alternative to STP providing fast reconvergence in Layer 2 ring topologies	Base		L2	L2
Per-VLAN Spanning Tree (PVST) / PVST+	► Supports interoperability with Cisco proprietary protocol	Base	L2	L2	L2
Topology Groups	► A named set of VLANs that share a Layer 2 topology ► Compatible with the following Layer 2 protocols: – 802.1D Spanning Tree Protocol – 802.1w Rapid Spanning Tree Protocol – Metro Ring Protocol – Virtual Switch Redundancy Protocol	Base		L2	L2
802.1q	► VLAN Tagging ► Allows traffic from multiple VLANs to traverse over the same network link ► VLAN Tag-type Translation	Base	L2	L2	L2
802.1q-in-q	► Supports attaching a second 802.1q tag to an already tagged frame	Base		L2	L2
VLANs	► Layer 2 Port-based VLANs ► Dual-mode VLAN ports: Allow untagged and tagged traffic through the same port	Base	L2	L2	L2

Category: Layer 2 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
VLANs	► Layer 3 Protocol-based (IPv4, dynamic IPv6, AppleTalk, IPX) VLANs ► Uplink Ports within a Port-based VLAN: Send all broadcast and unknown-unicast traffic to uplink port	Base		L2	L2
VLANs	► Super Aggregated VLANs (SAV): Allow multiple VLANs within a VLAN to construct Layer 2 Paths and Channels			L2	L2
802.3ad	► Link Aggregation Control Protocol (LACP): Enables dynamic creation of Link Aggregation Groups (LAGs) ► Allows multiple load-sharing links between two devices	Base	L2	L2	L2
Static Link Aggregation Groups (Trunk Groups)	► Manually configured LAGs / trunk groups ► Allows multiple load-sharing links between two devices ► Compatible with Cisco EtherChannel	Base	L2	L2	L2
Multicast (Layer 2 traffic reduction)	► Internet Group Management Protocol (IGMP) v1, v2, v3 snooping ► IGMP snooping per VLAN ► IGMP Fast Leave ► Protocol Independent Multicast-Sparse Mode (PIM-SM) v2 snooping	Base	L2	L2	L2
Multicast (Layer 2 traffic reduction)	► MLD v1/v2 Snooping ► MLDv1 Fast Leave			L2	L2
Dynamic Host Configuration Protocol (DHCP) Assist	► Ensures that a DHCP server that manages multiple IP subnets can recognize the requester's IP subnet		L2	L2	L2

Category: Advanced Layer 2 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Ethernet Service Instance (ESI)	► STP and RSTP for Customer-VLANs (C-VLANs), Service-VLANs (S-VLANs), and Backbone-VLANs (B-BLANs) when using the ESI framework ► Topology Groups within an ESI ► VLAN Groupsp within an ESI ► Static LAG (trunk groups) within an ESI	Metro			
802.1ad	► Provider Bridges: Enables a C-VLAN to be mapped to a S-VLAN	Metro			
802.3ah	► Provider Backbone Bridges (PBB): Expands Provider Bridge capabilities by adding PBB header including Backbone MAC Addresses	Metro			
Layer 2 Protocol Forwarding	► Allows/Blocks forwarding of Layer 2 protocol packets under a user-configured ESI	Metro			
VLAN Translation	► Assist with translating between S-VLANs across a provider boundary	Metro			

Category: Layer 3 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Static IPv4 Routes	► Manually configured IPv4 routes ► Adress Resolution Protocol (ARP) entries	Base		BL3	BL3
Routing Information Protocol (RIP) v1/v2 Advertising	► Advertises directly connected routes but does not learn RIP routes	Base		BL3	BL3
RIPv1/v2	► IPv4 dynamic, distance-vector routing protoco ► Interior Gateway Protocol (IGP)l	Base		EL3	

Category: Layer 3 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Open Shortest Path First (OSPF) v2	► IPv4 dynamic, link-state routing protocol ► Interior Gateway Protocol (IGP) ► MD5 authentication	Fl3 Metro		EL3	
OSPFv2	► Graceful restart: Increases network resiliency by minimizing disruptions in forwarding during a router restart ► BiDirectional Forwarding Detection (BFD): Enables rapid detection of the failure on a forwarding path	Fl3 Metro			
Intermediate System to Intermediate System (IS-IS)	► IPv4, dynamic link-state routing protocol ► Interior Gateway Protocol (IGP) ► MD5 authentication ► BiDirectional Forwarding Detection (BFD): Enables rapid detection of the failure on a forwarding path	Fl3 Metro			
Border Gateway Protocol (BGP)-4	► IPv4 dynamic, path-vector protocol ► Exterior Gateway Protocol (EGP) ► Graceful restart ► MD5 authentication ► Equal-Cost Multiple-Path load sharing ► Confederations: sub-dividing an Autonomous System (AS) ► Aggregate route advertisement ► IP Address filtering ► Route reflectors ► Route flap dampening ► Fast external failover	FL3			
Route Redistribution (IPv4)	► Distribute IPv4 routes between directly connected, static, and dynamically learned routes	Base		EL3	

Category: Layer 3 features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Multicast Routing (IPv4)	► IGMP v1, v2, v3 ► PIM-SM v1/v2 ► PIM-Dense Mode (PIM-DM) v2 ► PIM-Source Specific Mode (PIM-SSM) v4 ► Distance Vector Multicast Routing Protocol (DVMRP) ► PIM Anycast RP ► Multicast Source Discovery Protocol (MSDP) ► Passive Multicast Route Insertion (PMRI)	FL3 Metro			
Virtual Router Redundancy Protocol (VRRP) / VRRP Extended (VRRPE)	► Provides redundant paths to routers within a LAN ► VRRPE utilizes a Master and Backup routers, from which the Master listens to a virtual IP address	Base			
Multi-Virtual Routing and Forwarding (Multi-VRF)	► Sometimes known as “Multi-VRF CE” or “VRF-Lite” ► Enables multiple virtual routing/forwarding tables to be maintained on a single device	FL3 Metro			
Virtual Routing Interfaces	► Integrated Switch Routing (ISR): Enables a logical routing interface enabling Layer 3 routing of traffic between VLANs eliminating the need for an external router	Base		BL3	BL3
DHCP and BootP Relay	► Enable Layer 3 forwarding and return of DHCP and BootP requests from devices			BL3	BL3
ICMP Redirect Messages	► If traffic is misdirected notifies source and forwards traffic to appropriate router	Base			
Route-only Support	► Disable Layer 2 switching at a global or interface level	Base		EL3	

Category: Multiprotocol Label Switching (MPLS) features	Feature description	IBM c-series	IBM x-series	IBM g-series
MPLS	► Label Distribution Protocol (LDP) ► Resource Reservation Protocol-Traffic Engineering (RSVP-TE) ► CSPF ► OSPF-TE ► ISIS-TE ► Adaptive Labeled Switched Paths (LSPs)	Metro		

Category: Virtual Private Networks (VPN) features	Feature description	IBM c-series	IBM x-series	IBM g-series
Layer 2 VPN	► Virtual Private LAN Service (VPLS) – Multicast Snooping for VPLS – IGMP and PIM Proxy for VPLS – Disabling of VPLS Local switching – Local VPLS (without MPLS) – VLAN Translation ► Virtual Leased Line (VLL) ► Local VLL ► Single tag	Metro		

For the most up-to-date list, see the software *Release Notes* and the *NetIron Configuration Guide*, *FastIron and TurboIron Configuration Guide*, and the *FastIron Configuration Guide* for the software version in which you are interested.

5

Layer 2 switching

In this chapter we review the setup and configuration of all of the basic Layer 2 protocols, including VLANs, Spanning Tree, and Link Aggregation.

5.1 VLANs

In this section we assume that you are familiar with the basics of how VLANs work.

5.1.1 Basic VLAN configuration

We discuss the basics of VLAN configuration in the topics that follow.

Adding ports to a VLAN

Ports are added to VLANs on these products differently than on other networking products that you might be familiar with. Instead of adding a VLAN to the configuration for a particular port, you add the port to the configuration for the VLAN. This is done by changing into the VLAN config level (that is, **`vlan 13`**), and then adding ports to the VLAN.

Overall process

We add several ports to VLAN 13 on an s-series as shown in Example 5-1.

Example 5-1 Adding ports to a VLAN

```
BR-telnet@s_Series_R(config)#
BR-telnet@s_Series_R(config)#vlan 13
BR-telnet@s_Series_R(config-vlan-13)#untagged e10/1 eth 11/1 to 11/4
Added untagged port(s) ethe 10/1 ethe 11/1 to 11/4 to port-vlan 13.
BR-telnet@s_Series_R(config-vlan-13)#
BR-telnet@s_Series_R(config-vlan-13)#tagged eth 11/5
Added tagged port(s) ethe 11/5 to port-vlan 13.
BR-telnet@s_Series_R(config-vlan-13)#
```

Be aware of the following considerations:

- A “tagged” port is one that carries traffic with 802.1q-tagged Ethernet frames; basically, frames that already have VLAN information added to them.
- An untagged port is one that cannot carry tagged traffic.
- The default is that all ports start untagged in VLAN 1.
- A particular port can belong to multiple VLANs.

5.1.2 Dual-mode ports

There are some situations where a port will need to pass both tagged and untagged traffic, such as an IP phone that has a second port on it to plug in a PC. The traffic from the IP phone can be tagged (by the phone itself), while the PC traffic is untagged.

s-series, g-series, and x-series

To configure a dual-mode port, insert the port as tagged in the appropriate VLANs and then use the interface configuration to flag the port as a dual-mode port for a specific VLAN, for example **`dual-mode`** ***`10`***.

The VLAN specified **`dual-mode`** command is the VLAN that untagged traffic will be transmitted over. If no VLAN is specified with the command the default will be VLAN 1.

The interface can be tagged in multiple VLANs but can only be placed in dual-mode into one VLAN. We can see this process in Example 5-2.

Example 5-2 Configuring a port for dual-mode operation

```
BR-telnet@s_Series_R(config)#
BR-telnet@s_Series_R(config)#vlan 10
BR-telnet@s_Series_R(config-vlan-10)#tagged ethernet 11/1
Added tagged port(s) ethe 11/1 to port-vlan 10.
BR-telnet@s_Series_R(config-vlan-10)#
BR-telnet@s_Series_R(config-vlan-10)#vlan 13
BR-telnet@s_Series_R(config-vlan-13)#
BR-telnet@s_Series_R(config-vlan-13)#tagged ethernet 11/1
Added tagged port(s) ethe 11/1 to port-vlan 13.
BR-telnet@s_Series_R(config-vlan-13)#
BR-telnet@s_Series_R(config-vlan-13)#interface ethernet 11/1
BR-telnet@s_Series_R(config-if-e1000-11/1)#
BR-telnet@s_Series_R(config-if-e1000-11/1)#dual-mode 10
BR-telnet@s_Series_R(config-if-e1000-11/1)#
```

m-series, c-series, and r-series

On these devices, when you `tag` a port, the interface still stays untagged in the default VLAN of 1. This is in essence a dual-mode port. To have any untagged traffic pass through another VLAN on the interface, all you need to do is **`untag`** the interface into the VLAN you want.

5.1.3 Virtual interfaces on VLAN (products running router code only)

To create a virtual interface on a VLAN, change to the VLAN config level and run the command **`router-interface ve`** ***`1`***. You can then switch to **`interface`** ***`ve1`***, and assign an IP address to it, just as you might with any other interface. This IP address can be used as the default gateway for hosts attached to that VLAN.

Using virtual routing interfaces allows you to route between VLANs without needing an external router. We show this in Example 5-3.

Example 5-3 Virtual Interface Creation

```
telnet@m_Series(config)#
telnet@m_Series(config)#vlan 53
telnet@m_Series(config-vlan-53)#router-interface ve 56
telnet@m_Series(config-vlan-53)#
telnet@m_Series(config-vlan-53)#interface ve56
telnet@m_Series(config-vif-56)#
telnet@m_Series(config-vif-56)#ip address 4.5.6.7/22
telnet@m_Series(config-vif-56)#
```

5.1.4 VLAN groups

Just as you can configure multiple interfaces at once, you can also configure multiple VLANs at once. However, the mechanisms for doing so are different than those for interfaces.

VLAN groups also enable you to configure a single virtual interface for an entire group of VLANs. You can create up to 32 VLAN groups, and put up to 1000 VLANs in each group. The VLANs in the group must be in a contiguous range and not currently be configured on the product. After a VLAN is assigned to a VLAN group, it cannot be configured individually.

Ports assigned to the entire VLAN group (as opposed to an individual VLAN) must be added as tagged ports.

Overall process

To create a group, simply use the config command **`vlan-group`** ***`1`*** **`vlan`** ***`2 to 20`***. The whole process is shown in Example 5-4.

Example 5-4 VLAN groups

```
telnet@m_Series(config)#
telnet@m_Series(config)#vlan-group 1 vlan 67 to 70
```

```
telnet@m_Series(config-vlan-group-1)#
telnet@m_Series(config-vlan-group-1)#spanning-tree
telnet@m_Series(config-vlan-group-1)#exit
telnet#m_Series(config)#
telnet@m_Series(config)#vlan 67
error - vlan 67 is a member of vlan-group 1
telnet#m_Series(config)#
```

Creating a group router interface (s-series, g-series, and x-series only)

After you have added ports to your VLAN group, you can create a single virtual interface for the entire VLAN group. To do this run the vlan-group config command **group-router-interface**, and configure the interface with the command **interface group-ve *1***. The interface number assigned to your new group-router-interface is always equal to the VLAN group number. You cannot create more than one, nor pick your own number.

The creation of an interface is shown in Example 5-5.

Example 5-5 VLAN Group Virtual Interface

```
BR-FastIron SX 1600 Router(config)#vlan-group 2 vlan 45 to 49
BR-FastIron SX 1600 Router(config-vlan-group-2)#group-router-interface
error - failed to create ve because ports were not configured in this
vlan
BR-FastIron SX 1600 Router(config-vlan-group-2)#tagged ethernet 11/9
Added tagged port(s) ethe 11/9 to vlan-group 2.
BR-FastIron SX 1600 Router(config-vlan-group-2)#group-router-interface
BR-FastIron SX 1600 Router(config-vlan-group-2)#
BR-FastIron SX 1600 Router(config-vlan-group-2)#interface group-ve 2
BR-FastIron SX 1600 Router(config-vif-group-2)#
BR-FastIron SX 1600 Router(config-vif-group-2)#ip address 1.2.3.4/28
BR-FastIron SX 1600 Router(config-vif-group-2)#
```

5.1.5 Creating room for more VLANs

By default the number of VLANs on a particular product is either 64 (s-series, g-series, x-series) or 512 (m-series, r-series, c-series). These limits can, however, be increased.

To view the current limits on a product, run the command **show default values**. To change the limit for a particular value, run the config command **system-max *vlan***.

In Example 5-6 we illustrate the **show default values** command as well as modifying the VLAN max on a g-series.

Example 5-6 Increasing the maximum number of VLANs

```
FGS648P-STK Switch(config)#show default values
sys log buffers:50      mac age time:300 sec    telnet sessions:5
System Parameters    Default    Maximum    Current
igmp-max-group-addr  8192       32768      8192
ip-filter-sys        1021       1021       1021
l3-vlan              32         1024       32
mac                  16384      16384      16384
vlan                 64         4095       64
spanning-tree        32         255        32
mac-filter-port      32         256        32
mac-filter-sys       64         512        64
view                 10         65535      10
rmon-entries         6144       32768      6144
mld-max-group-addr   8192       32768      8192
igmp-snoop-mcache    4096       8192       4096
mld-snoop-mcache     4096       8192       4096
FGS648P-STK Switch(config)#
FGS648P-STK Switch(config)#system-max vlan 80
Reload required.  Please write memory and then reload or power cycle.
FGS648P-STK Switch(config)#
```

Note that this operation requires a reload of the product to take effect. (Changing any of these values requires a reload). Even on products with redundant management modules, a complete (disruptive) reload is still required.

5.1.6 MAC-based VLANs

These products have the ability to assign a particular VLAN based on the MAC address. This is usually done by retrieving the VLAN information from a RADIUS server. Full configuration details on this feature are beyond the intended scope of this book. For further information about this feature, see the *Configuration Guide*.

5.2 Spanning Tree Protocol

Spanning Tree Protocol (STP) is necessary for any switched network. It prevents broadcast traffic from travelling in an endless loop around the network, quickly overwhelming the network.

5.2.1 Enabling and disabling

STP can be enabled (or disabled) on either a global, interface, or VLAN level. In each case, the command is the same; simply use the command **`spanning-tree`** run from the appropriate configuration level. (Use **`spanning-tree single`** if run from the global level).

Rules for enabling and disabling

Configuration commands applied to a higher level will override any active configuration on a lower level on a one-time basis. For instance, if you have enabled spanning-tree on a VLAN, and then disable it on a global level, it will also be disabled on the VLAN. However, if you later re-enable it on the VLAN, it will still be disabled globally (except in places where you have explicitly enabled it).

Special notes for s-series, g-series, and x-series

On products running Layer 2 only images, spanning tree is enabled by default. On products running Layer 3 images, it is not. It is a good practice to explicitly enable or disable spanning tree on a global level to remove any sources of confusion.

Also, new VLANs will have the spanning tree state (enabled or disabled) according to the device defaults, *not* the global setting for spanning tree.

5.2.2 Spanning tree parameters

These are the spanning tree parameters.

Bridge priority

To set the bridge priority, use the **`spanning-tree priority`** command. This value can range from 0 to 65535. (If configuring on a global level, use **`spanning-tree single priority`**).

Port priority

To set port priority, use the config command (from the VLAN level) **`spanning-tree ethernet 5/1 priority 16`** (you cannot use the “e5/1” shorthand with this command). This value must be from 0 to 240, but only in increments of 16.
If configuring from the global level, use **`spanning-tree single ethernet 5/1 priority 16`**).

5.2.3 802.1w

802.1w is a form of Spanning Tree that offers more rapid convergence of the tree after a topology change. It can be enabled on a per VLAN basis using **`spanning-tree 802-1w`** from the VLAN config level. (On the m-series, r-series, and c-series, you can also use the VLAN-level command **`rstp`**.)

The priority is set in a similar fashion to STP: **`spanning-tree 802-1w priority`** ***`1`***. Each port in a VLAN used for a link between switches must be set to Point-to-Point mode using **`spanning-tree 802-1w ethernet`** ***`5/1`*** **`admin-pt2pt-mac`**.

To prevent unintended network re-convergence, ports that will be used to connect to edge devices must be set to edge-port mode. This is done (from the VLAN config level) with **`spanning-tree 802-1w ethernet`** ***`5/1`*** **`admin-edge-port`**. The command does not work on a range of ports; it must be entered individually for each port.

5.2.4 Spanning tree configuration example

In Example 5-7 we present a sample spanning tree configuration, using many of the options presented previously.

Example 5-7 Spanning tree

```
BR-FastIron SX 1600 Router(config)#
BR-FastIron SX 1600 Router(config)#spanning-tree single
BR-FastIron SX 1600 Router(config)#spanning-tree single priority 5
BR-FastIron SX 1600 Router(config)#spanning-tree single 802-1w
BR-FastIron SX 1600 Router(config)#no spanning-tree single
BR-FastIron SX 1600 Router(config)#vlan 43
BR-FastIron SX 1600 Router(config-vlan-43)#spanning-tree
BR-FastIron SX 1600 Router(config-vlan-43)#unit e 11/14 to 11/20
Added untagged port(s) ethe 11/14 to 11/20 to port-vlan 43.
BR-FastIron SX 1600 Router(config-vlan-43)#
BR-FastIron SX 1600 Router(config-vlan-43)#spanning-tree ethernet 11/14
priority 16
BR-FastIron SX 1600 Router(config-vlan-43)#spanning-tree 802-1w
BR-FastIron SX 1600 Router(config-vlan-43)#spanning-tree 802-1w
ethernet 11/15 admin-edge-port
BR-FastIron SX 1600 Router(config-vlan-43)#
```

5.3 Trunks and LAG groups

What these products see as “trunking” or “link aggregation” goes by other names by other vendors. In the case of these products “trunking” is their name for Link Aggregation; a group of physical ports that act as one logical port. All of the links in the trunk share a single MAC address and IP address.

The maximum number of links per group varies by product and, in the case of the s-series, the type of interface module, either IPv4 or IPv6/IPv4 modules:

- g-series: 8
- s-series:
  - IPv6/IPv4 modules: 8
  - IPv4 modules: 4
- m-series: 32
- c-series: 12

The ports must meet these requirements:

- All belong to the same VLANs
- Be the same speed, for example 10 GbE ports cannot be trunked with 10/100/1000 MbE ports
- Have the same port attributes
- Have the same ACLs
- Have no Layer 3 configuration (that is, IP address or OSPF configuration) prior to being placed in the trunk group.

Link aggregation/trunk groups can span multiple interface modules in a chassis-based device and multiple stacked switches on the B50G stackable switches. For additional rules, see the *Configuration Guides*.

There are two types of link aggregation you can configure: static and dynamic. Static trunk groups are compatible with Cisco EtherChannel. Dynamic Link Aggregation Groups (LAGs) are based on IEEE 802.3ad.

5.3.1 Creating a static group

Static groups are proprietary implementation of link aggregation. If your groups will not need to interoperate with other vendors’ products, static groups are the most straightforward to set up.

s-series, g-series, and x-series

Before configuring a trunk group, it is a best practice to unplug the ports that will be trunked together.

To create a trunk group, simply run the config-level command **trunk ethernet** ***1/1 to 1/4***. However, this command will only configure the trunk group. To actually activate it, use the command **trunk deploy**. We show this process in Example 5-8.

Example 5-8 Trunk creation on a g-series

```
FGS648P-STK Switch(config)#
FGS648P-STK Switch(config)#trunk switch ethernet 1/1/1 to 1/1/4
Trunk will be created in next trunk deploy.
FGS648P-STK Switch(config)#trunk deploy
FGS648P-STK Switch(config)#show trunk

Configured trunks:

Trunk ID: 1
Hw Trunk ID: 1
Ports_Configured: 4
Primary Port Monitored: Jointly
Ports        1/1/1   1/1/2   1/1/3   1/1/4
[truncated for brevity]
```

To add ports from multiple slots (s-series), stack units (g-series), or simply multiple ports, use the command: **trunk** ***ethernet 1/1 to 1/4 ethernet 2/3 ethernet 4/5 to 4/6***.

m-series, r-series, and c-series

On these devices, each link aggregation group has a unique name. After you name the link aggregation, you can then assign ports to it and set its attributes.

The basic steps for creating a link aggregation are as follows:

1. Create the link aggregation with your desired name with the config level command **lag** ***<group_name>*** **static**.
2. Add ports to the LAG with the command **ports ethernet** ***5/1 to 5/2***.
3. Designate one of the ports as the “primary” port with the command **primary-port** ***5/1***. (Note that it is just “5/1”, not “ethernet 5/1”.)
4. Activate the trunk with **deploy**.

We demonstrate the process of creating a link aggregation in Example 5-9.

Example 5-9 m-, r-, and c-series link aggregation

```
telnet@NetIron CES 2048C(config)#
telnet@NetIron CES 2048C(config)#lag group_1 static
telnet@NetIron CES 2048C(config-lag-group_1)#ports ethernet 1/6 to 1/7
telnet@NetIron CES 2048C(config-lag-group_1)#primary 1/7
telnet@NetIron CES 2048C(config-lag-group_1)#deploy
telnet@NetIron CES 2048C(config-lag-group_1)#
telnet@NetIron CES 2048C(config-lag-group_1)#sh lag
Total number of LAGs:          1
Total number of deployed LAGs: 1
Total number of trunks created:1 (254 available)
LACP System Priority / ID:     1 / 001b.ed38.f941
LACP Long timeout:             90, default: 90
LACP Short timeout:            3, default: 3

=== LAG "group_1" (static Deployed) ===
LAG Configuration:
   Ports:         ethe 1/6 to 1/7
   Port Count:    2
   Primary Port:  1/7
   Trunk Type:    hash-based

Deployment:  Trunk ID 1

Port  Link L2 State  Dupl Speed Trunk Tag Priori MAC            Name
Type
1/6   DisabNone     None None  1    No  level0 001b.ed38.f947
default-port
```

5.3.2 VLANs and groups

To add a trunk/link aggregation to a VLAN, simply add at least one of the ports; the remaining ports in the trunk will be added to the VLAN automatically. The reverse is true for removing a trunk from a VLAN.

5.3.3 Dynamic groups

Dynamic Link Aggregation Groups (LAG) utilize the 802.3ad standard for link aggregation. If your switch groups will need to interoperate with the products of other vendors, you will need to use this feature.

Dynamic groups are created by assigning a “key” to a set of ports; ports that have the same key as the port on the other end of the link.

s-series, g-series, and x-series

Dynamic groups are created by enabling link aggregation on specific ports, and assigning them a “key”. The key must match on both ends for the group to come online.

To do this, configure each port with the following commands:

1. **link-aggregate off**
2. **link-aggregate configure key *10000***
3. **link-aggregate active**

We show the process in Example 5-10.

Example 5-10 Creating a g-series dynamic group

```
FGS648P-STK Switch(config)#interface ethernet 1/1/1 ethernet 1/1/3
FGS648P-STK Switch(config-mif-1/1/1,1/1/3)#link-aggregate off
FGS648P-STK Switch(config-mif-1/1/1,1/1/3)#link-aggregate configure key
10000
FGS648P-STK Switch(config-mif-1/1/1,1/1/3)#link-aggregate active
FGS648P-STK Switch(config-mif-1/1/1,1/1/3)#
```

Note that we configured multiple ports at once. You can specify ranges, or configure ports one at a time.

m-series, r-series, and c-series

Dynamic LAGs on these devices are created in a similar fashion to static LAGs but with the “dynamic” attribute in the following command: **lag *<group_name> dynamic***. A key is created automatically on either end.

We show the process of creating dynamic LAGs in Example 5-11.

Example 5-11   Creating a m-series dynamic LAG

```
NetIron MLX-8(config)#lag "To_B48C" dynamic
NetIron MLX-8(config-lag-To_B48C)#ports ethernet 5/1 to 5/4
NetIron MLX-8(config-lag-To_B48C)#primary-port 5/1
NetIron MLX-8(config-lag-To_B48C)#deploy
NetIron MLX-8(config-lag-To_B48C)#exit
NetIron MLX-8(config)#show lag
Total number of LAGs:          1
Total number of deployed LAGs: 1
Total number of trunks created:1 (127 available)
LACP System Priority / ID:     1 / 001b.ed16.a800
LACP Long timeout:             90, default: 90
LACP Short timeout:            3, default: 3

=== LAG "To_B48C" (dynamic Deployed) ===
LAG Configuration:
   Ports:         ethe 5/1 to 5/4
   Port Count:    4
   Primary Port:  5/1
   Trunk Type:    hash-based
   LACP Key:      107

Deployment:  Trunk ID 2, Active Primary 5/1

Port  Link L2 State  Dupl Speed Trunk Tag Priori MAC            Name
5/1   Up   Forward   Full 1G    2     Yes level0 001b.ed16.a800
5/2   Up   Forward   Full 1G    2     Yes level0 001b.ed16.a800
5/3   Up   Forward   Full 1G    2     Yes level0 001b.ed16.a800
5/4   Up   Forward   Full 1G    2     Yes level0 001b.ed16.a800

Port  [Sys P] [Port P] [ Key ]
[Act][Tio][Agg][Syn][Col][Dis][Def][Exp][Ope]
5/1        1        1     107   Yes   L   Agg  Syn  Col  Dis  No   No
Ope
5/2        1        1     107   Yes   L   Agg  Syn  Col  Dis  No   No
Ope
5/3        1        1     107   Yes   L   Agg  Syn  Col  Dis  No   No
Ope
5/4        1        1     107   Yes   L   Agg  Syn  Col  Dis  No   No
Ope
NetIron MLX-8(config)#
```

6

Layer 3 routing

In this chapter we review the setup and configuration of all of the basic Layer 3 protocols, including RIP, OSPF, and BGP.

6.1 RIP and static routes

This section assumes that you are familiar with the basics of static routing and the RIP routing protocol. Examples in this section are based on the network setup shown in Figure 6-1.

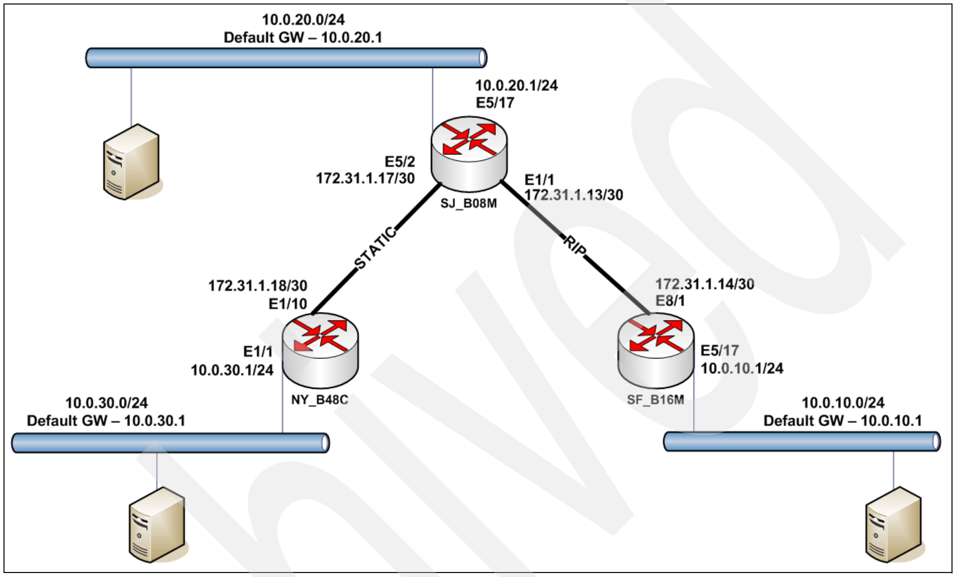


Figure 6-1 RIP lab network setup

6.1.1 Enabling and disabling RIP

In this section we describe how to enable or disable a RIP routing setup. The setup for this is shown in Figure 6-1, and in our examples we use SJ_B08M and SF_B16M routers.

There are three basic steps to enable RIP IP routing:

1. Enable RIP globally
2. Assign IP address to the routed interfaces
3. Enable RIP on the interfaces

We provide examples of those steps in the following sections. (In some of the examples we have outlined particular lines of code that we want to emphasize.)

Enabling RIP globally

With IBM b-type routers, RIP is enabled at an interface level. Before enabling RIP on particular interface, RIP has to be enabled globally.

With the **router rip** command in the config level context, you will come to the config-rip-router level as shown in Example 6-1.

Example 6-1 Enabling RIP globally

```
SJ_B08M#config t
SJ_B08M(config)#router rip
SJ_B08M(config-rip-router)#
```

In this level you can also configure RIP global parameters.

With this, RIP will be enabled globally. RIP can be globally disabled by using **no router rip** in the config level context as shown in Example 6-2.

Example 6-2 Disabling RIP globally

```
SJ_B08M#config t
SJ_B08M(config)#no router rip

Router rip now disabled. All rip config data will be lost when writing
to flash!!
```

Assigning IP address/mask to the routed interface

Each routed interface has to have an IP address/mask defined. In our example we are using interface E1/1 on the SJ_B08M router and interface E8/1 on the SF_B16M router for inter-router communication. Interfaces E5/17 on SJ_B08M router and E5/17 on SJ_B16M router are used for local network on each side. This can be achieved in the interface config level context as shown in Example 6-3 for our setup.

Example 6-3 Assigning IP addresses

```
SJ_B08M#config t
SJ_B08M(config)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip addr 172.31.1.13 255.255.255.252
SJ_B08M(config)#int e 5/17
SJ_B08M(config-if-e1000-5/17)#ip addr 10.0.20.1/24

SF_B16M#config t
SF_B16M(config)#interface e 8/1
SF_B16M(config-if-e10000-8/1)#ip addr 172.31.1.14 255.255.255.252
SF_B16M(config)#int e 5/17
SF_B16M(config-if-e1000-5/17)#ip addr 10.0.10.1/24
```

After assigning the IP addresses, we can verify connectivity between the two routers, SJ_B08M and SF_B16M, as shown in Example 6-4.

Example 6-4   Verifying inter router connectivity

```
SJ_B08M#ping 172.31.1.14
Sending 1, 16-byte ICMP Echo to 172.31.1.14, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 172.31.1.14     : bytes=16 time=603ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=603/603/603 ms.
SJ_B08M#
SF_B16M#ping 172.31.1.13
Sending 1, 16-byte ICMP Echo to 172.31.1.13, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 172.31.1.13     : bytes=16 time<1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=0/0/0 ms.
```

Routing information before enabling the RIP on interfaces is displayed in Example 6-5.

Example 6-5   Routing info before RIP is enabled

```
SJ_B08M(config)#show ip route
Total number of IP routes: 4
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type
Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    4d19h
2       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    0m24s
3       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d19h
4       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d1h
SJ_B08M(config)#
SF_B16M(config)#show ip route
Total number of IP routes: 4
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type
Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    1d17h
2       10.0.10.0/24       DIRECT          eth 5/17    0/0           D    0m50s
3       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    1d17h
4       172.31.1.12/30     DIRECT          eth 8/1     0/0           D    1d1h
SF_B16M(config)#
```

As you can see, we have no RIP routes present yet.

Enabling RIP on the interface

To enable RIP on the interface, the **`ip rip v2-only`** command can be used on the interface config level context as shown in Example 6-6.

Example 6-6 Enabling RIP on interfaces

```
SJ_B08M#config t
SJ_B08M(config)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip rip v2-only
SF_B16M#config t
SF_B16M(config)#interface e 8/1
SF_B16M(config-if-e10000-8/1)#ip rip v2-only
```

After enabling RIP, the routes need to be redistributed with the **`redistribute connected`** command in the rip-router config context level as shown in Example 6-7.

Example 6-7 Route redistribution

```
SJ_B08M#config t
SJ_B08M(config)#router rip
SJ_B08M(config-rip-router)#redistribute connected
SJ_B08M(config-rip-router)#
SF_B16M#config t
SF_B16M(config)#router rip
SF_B16M(config-rip-router)#redistribute connected
```

This will distribute all connected routes by the RIP protocol. Thus both routers will get information about routes from another router and with this information enable access from local networks.

In Example 6-8 you can see that RIP distributed routes are now in the route table of both routes, and networks from both sides can be reached.

Example 6-8 After RIP redistribution

```
SJ_B08M#show ip route
Total number of IP routes: 5
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    4d20h
2       10.0.10.0/24       172.31.1.14     eth 1/1     120/2         R    17m9s
3       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    38m44s
4       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d20h
```

```
5       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d2h
SJ_B08M#
SJ_B08M#ping 10.0.10.1
Sending 1, 16-byte ICMP Echo to 10.0.10.1, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 10.0.10.1       : bytes=16 time<1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=0/0/0 ms.
SJ_B08M#
SF_B16M#show ip route
Total number of IP routes: 5
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    1d17h
2       10.0.10.0/24       DIRECT          eth 5/17    0/0           D    31m10s
3       10.0.20.0/24       172.31.1.13     eth 8/1     120/2         R    5m47s
4       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    1d17h
5       172.31.1.12/30     DIRECT          eth 8/1     0/0           D    1d2h
SF_B16M#
SF_B16M#ping 10.0.20.1
Sending 1, 16-byte ICMP Echo to 10.0.20.1, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 10.0.20.1       : bytes=16 time=3ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=3/3/3 ms.
```

6.1.2 Configuring a static route

A static route is configured with the **ip route** command in the top level config context as shown in Example 6-10. In this example, before adding the static routes, we setup the environment as shown in Figure 6-1 on page 190. In this example the SJ_B08M and NY_B48C routers are used.

The following steps for the initial environment setup are shown in Example 6-9:

1. Assigning of the IP address on the interface E5/2 on the SJ_B08M router
2. Assigning of the IP address on the interface E1/10 on the NY_B48C router
3. Testing of the connectivity between routers with the **ping** command
4. Assigning of the IP address on interface E1/1 on the NYB48C router
5. Displaying the route information about both routers

Example 6-9   Adding static route - initial setup

```
SJ_B08M#config t
SJ_B08M(config)#int e 5/2
SJ_B08M(config-if-e1000-5/2)#ip addr 172.31.1.17/30
SJ_B08M(config-if-e1000-5/2)#exit
SJ_B08M(config)#
*************************************************************************************
****************
NY_B48C#config t
NY_B48C(config)#int e 1/10
NY_B48C(config-if-e1000-1/10)#ip addr 172.31.1.18/30
NY_B48C(config-if-e1000-1/10)#exit
NY_B48C(config)#exit
*************************************************************************************
****************
NY_B48C#ping 172.31.1.17
Sending 1, 16-byte ICMP Echo to 172.31.1.17, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 172.31.1.17     : bytes=16 time=376ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=376/376/376 ms.
SJ_B08M#ping 172.31.1.18
Sending 1, 16-byte ICMP Echo to 172.31.1.18, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 172.31.1.18     : bytes=16 time<1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=0/0/0 ms.
*************************************************************************************
****************
NY_B48C#config t
NY_B48C(config)#int e 1/1
Y_B48C(config-if-e1000-1/1)#ip addr 10.0.30.1/24
NY_B48C(config-if-e1000-1/1)#exit
NY_B48C(config)#exit
*************************************************************************************
****************
NY_B48C#show ip route
Total number of IP routes: 2
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       10.0.30.0/24       DIRECT          eth 1/1     0/0           D    0m8s
2       172.31.1.16/30     DIRECT          eth 1/10    0/0           D    8m34s
NY_B48C#
SJ_B08M#show ip route
```

```
Total number of IP routes: 6
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    4d21h
2       10.0.10.0/24       172.31.1.14     eth 1/1     120/2         R    45m41s
3       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    1h7m
4       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d21h
5       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d2h
6       172.31.1.16/30     DIRECT          eth 5/2     0/0           D    6m11s
SJ_B08M#
```

Now we are ready to add the static route to enable routing between the 10.0.20.0/24 and 10.0.30.0/24 networks as shown in Example 6-10. A static route needs to be added on both routers.

Example 6-10 Adding static route

```
SJ_B08M#config t
SJ_B08M(config)#ip route 10.0.30.0/24 172.31.1.18
SJ_B08M(config)#
SJ_B08M(config)#show ip route
Total number of IP routes: 7
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    4d21h
2       10.0.10.0/24       172.31.1.14     eth 1/1     120/2         R    54m49s
3       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    1h16m
4       10.0.30.0/24       172.31.1.18     eth 5/2     1/1           S    1m47s
5       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d21h
6       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d2h
7       172.31.1.16/30     DIRECT          eth 5/2     0/0           D    15m18s
SJ_B08M(config)#
NY_B48C#config t
NY_B48C(config)#ip route 10.0.20.0/24 172.31.1.17
NY_B48C(config)#
NY_B48C(config)#show ip route
Total number of IP routes: 3
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
```

```
1       10.0.20.0/24      172.31.1.17     eth 1/10    1/1            S    1m30s
2       10.0.30.0/24      DIRECT          eth 1/1     0/0            D    6m33s
3       172.31.1.16/30    DIRECT          eth 1/10    0/0            D    14m59s
NY_B48C(config)#
```

After adding the routes, we can check the connectivity using the **ping** command as shown in Example 6-11.

Example 6-11 Connectivity check

```
SJ_B08M#ping 10.0.30.1
Sending 1, 16-byte ICMP Echo to 10.0.30.1, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 10.0.30.1       : bytes=16 time<1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=0/0/0 ms.
SJ_B08M#
NY_B48C#ping 10.0.20.1
Sending 1, 16-byte ICMP Echo to 10.0.20.1, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 10.0.20.1       : bytes=16 time=1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=1/1/1 ms.
NY_B48C#
```

Null route

It is possible to define a so called “null” route for a specific network or a host address. Such a route means that all IP packets are dropped. It is configured like a static route using **ip route** but by using a “null” interface (sometimes called “null 0”) as shown in Example 6-12.

Example 6-12 Defining the null route

```
SJ_B08M#config t
Warning: 1 user(s) already in config mode.
SJ_B08M(config)#ip route 10.99.99.0 255.255.255.0 null0
SJ_B08M(config)#show ip route
Total number of IP routes: 8
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    4d21h
2       10.0.10.0/24       172.31.1.14     eth 1/1     120/2         R    1h2m
3       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    1h24m
4       10.0.30.0/24       172.31.1.18     eth 5/2     1/1           S    9m55s
5       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d21h
```

```
6       10.99.99.0/24      DIRECT          drop        1/1           S    0m6s
7       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d3h
8       172.31.1.16/30     DIRECT          eth 5/2     0/0           D    23m26s
```

6.1.3 Configuring a default route

The default route is configured with the **ip route** command in the top level config context as shown in Example 6-13.

Example 6-13 Defining the default route

```
SJ_B08M(config)#ip route 0.0.0.0 0.0.0.0 172.31.1.14
SJ_B08M(config)#show ip route
Total number of IP routes: 8
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    0m4s
        0.0.0.0/0          172.31.1.14     eth 1/1     1/1           S    0m4s
2       10.0.10.0/24       172.31.1.14     eth 1/1     120/2         R    1h5m
3       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    1h27m
4       10.0.30.0/24       172.31.1.18     eth 5/2     1/1           S    12m43s
5       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d21h
6       10.99.99.0/24      DIRECT          drop        1/1           S    2m55s
7       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d3h
8       172.31.1.16/30     DIRECT          eth 5/2     0/0           D    26m14s
SJ_B08M(config)#
```

If the default route is not present or defined, then the default network can be defined using the **ip default-network networkaddress** command as shown in Example 6-14.

Example 6-14 Adding the default network

```
SJ_B08M(config)#ip default-network 10.0.20.0/24
```

6.1.4 Displaying route information

To display routing information, the **show ip route** command can be used as shown in Example 6-15.

Example 6-15 Displaying the route information

```
SJ_B08M(config)#show ip route
Total number of IP routes: 8
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       0.0.0.0/0          10.64.208.1     mgmt 1      1/1           S    0m3s
2       10.0.10.0/24       172.31.1.14     eth 1/1     120/2         R    1h8m
3       10.0.20.0/24       DIRECT          eth 5/17    0/0           D    1h29m
4       10.0.30.0/24       172.31.1.18     eth 5/2     1/1           S    15m11s
5       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d21h
6       10.99.99.0/24      DIRECT          drop        1/1           S    5m23s
7       172.31.1.12/30     DIRECT          eth 1/1     0/0           D    1d3h
8       172.31.1.16/30     DIRECT          eth 5/2     0/0           D    28m42s
```

Table 6-1 shows the possible route types.

Table 6-1 Route types

Route type	Description
B	The route was learned from BGP.
D	The destination is directly connected to the router where the command was executed.
R	The route was learned from RIP.
S	Static route.
*	Candidate for the default route.
O	The route is an OSPF route.

If the **`ospf`** option is used to display routes, the types shown in Table 6-2 are possible.

Table 6-2 OSPF route types

Route type	Description
O	OSPF intra area route. This is a route in the same area.
IA	Inter area OSPF route. This is a route between two areas.
E1	External OSPF route type 1.
E2	External OSPF route type 2.

It is also possible to display just RIP routing information with the **`show ip route rip`** command.

6.2 OSPF

This section assumes that you are familiar with the basics of the OSPF routing protocol. Examples in this section are based on the network setup shown in Figure 6-2.

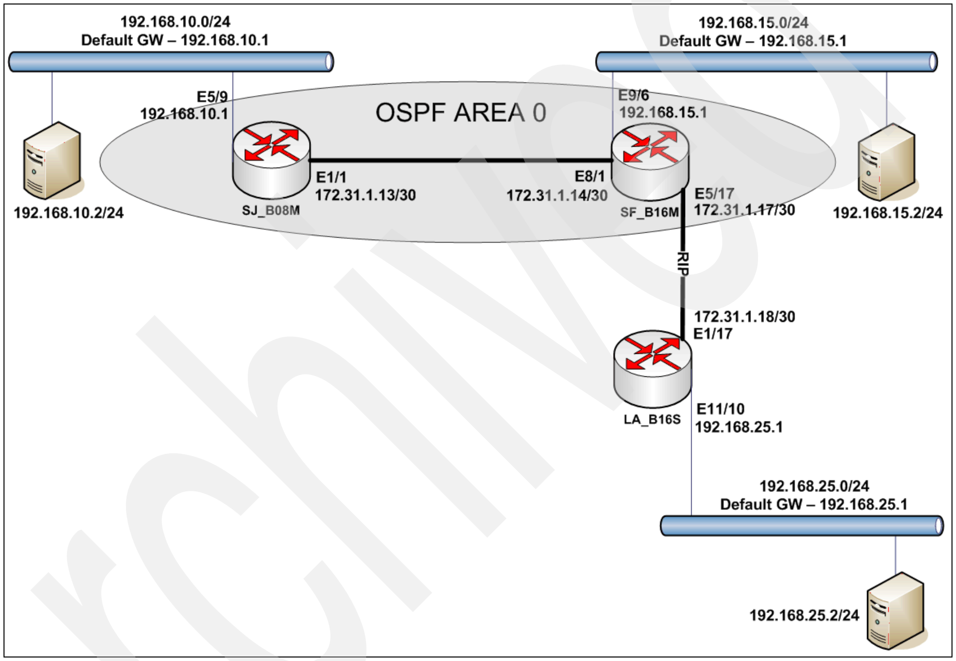


Figure 6-2 OSPF lab network setup

As we can see in Figure 6-2, we have two routers in the RIP domain: SF_B16M and LA_B16S, and two routers in the OSPF domain: SFB16M and SJ_B08M. Before performing the OSPF setup we need to set up the RIP domain. The following steps outline this setup:

- LA_B16S router:

  a. Enable RIP globally.
  b. Assign the IP address to the E1/17 routing interface.
  c. Enable RIP v2 routing protocol on E1/17 interface.
  d. Enable redistribution for the RIP domain
  e. Assign the IP address to the local network interface E11/10.

f. Test the connectivity to the device on the local network with the **ping** command.

These steps are shown in Example 6-16.

Example 6-16 LA_B16S RIP setup

```
LA_B16S#config t
LA_B16S(config)#router rip
***********************************************************************
LA_B16S(config-rip-router)#interface e 1/17
LA_B16S(config-if-e1000-1/17)#ip address 172.31.1.18 255.255.255.252
***********************************************************************
LA_B16S(config-if-e1000-1/17)#ip rip v2-only
***********************************************************************
LA_B16S(config)#router rip
LA_B16S(config-rip-router)#redistribution
***********************************************************************
LA_B16S(config-rip-router)#interface e 11/10
LA_B16S(config-if-e1000-11/10)#ip address 192.168.25.1 255.255.255.0
***********************************************************************
LA_B16S(config-if-e1000-11/10)#exit
LA_B16S(config)#exit
LA_B16S#ping 192.168.25.2
Sending 1, 16-byte ICMP Echo to 192.168.25.2, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 192.168.25.2    : bytes=16 time=1ms TTL=128
Success rate is 100 percent (1/1), round-trip min/avg/max=1/1/1 ms.
LA_B16S#
```

- SF_B16M router:

  a. Enable RIP globally.
  b. Assign the IP address to the E5/17 routing interface.
  c. Enable RIP v2 routing protocol on E5/17 interface.
  d. Enable redistribution for the RIP domain.
  e. Assign the IP address to the local network interface E9/6.
  f. Test the connectivity to the device on the local network with the ping command.

These steps are shown in Example 6-17.

Example 6-17 SF_B16M RIP setup

```
SF_B16M#config t
SF_B16M(config)#router rip
***********************************************************************
```

```
SF_B16M(config-rip-router)#interface e 5/17
SF_B16M(config-if-e1000-5/17)#ip address 172.31.1.17/30
***********************************************************************
SF_B16M(config-if-e1000-5/17)#ip rip v2-only
***********************************************************************
SF_B16M(config-if-e1000-5/17)#router rip
SF_B16M(config-rip-router)#redistribute connected
***********************************************************************
SF_B16M(config-rip-router)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#ip address 192.168.15.1/30
***********************************************************************
SF_B16M(config-if-e1000-9/6)#exit
SF_B16M(config)#exit
SF_B16M#ping 192.168.15.2
Sending 1, 16-byte ICMP Echo to 192.168.15.2, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 192.168.15.2    : bytes=16 time=100ms TTL=128
Success rate is 100 percent (1/1), round-trip min/avg/max=100/100/100
ms.
```

After the routers are setup we can verify the routing table and try to **ping** the local network from the other router. The results are shown in Example 6-18.

Example 6-18 RIP setup verification

```
LA_B16S#show ip route
Total number of IP routes: 5, avail: 262138 (out of max 262144)
B:BGP D:Connected  R:RIP  S:Static  O:OSPF *:Candidate default
        Destination     NetMask         Gateway         Port       Cost   Type
1       10.64.208.0     255.255.240.0   0.0.0.0         mgmt1      1      D
2       172.31.1.12     255.255.255.252 172.31.1.17     1/17       2      R
3       172.31.1.16     255.255.255.252 0.0.0.0         1/17       1      D
4       192.168.15.0    255.255.255.252 172.31.1.17     1/17       2      R
5       192.168.25.0    255.255.255.0   0.0.0.0         11/10      1      D
*******************************************************************************
LA_B16S#ping 192.168.15.1
Sending 1, 16-byte ICMP Echo to 192.168.15.1, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 192.168.15.1    : bytes=16 time<1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=0/0/0 ms.
*******************************************************************************
SF_B16M#show ip route
Total number of IP routes: 5
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
```

```
        Destination        Gateway         Port        Cost          Type
Uptime
1       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d16h
2       172.31.1.12/30     DIRECT          eth 8/1     0/0           D    4d0h
3       172.31.1.16/30     DIRECT          eth 5/17    0/0           D    1d2h
4       192.168.15.0/30    DIRECT          eth 9/6     0/0           D    2d18h
5       192.168.25.0/24    172.31.1.18     eth 5/17    120/2         R    6m22s
******************************************************************************
SF_B16M#ping 192.168.25.1
Sending 1, 16-byte ICMP Echo to 192.168.25.1, timeout 5000 msec, TTL 64
Type Control-c to abort
Reply from 192.168.25.1    : bytes=16 time<1ms TTL=64
Success rate is 100 percent (1/1), round-trip min/avg/max=0/0/0 ms.
```

Now we are ready to configure OSPF between SJ_B08M and SJ_B16M routers.

6.2.1 Basic configuration steps

Following are the initial basic configuration steps for OSPF setup. In our example, we perform these steps on both routers (SJB08M and SJB16M). Also, interface E1/1 on the SJB08M router and interface E8/1 on the SJB16M router participate in the OSPF area. Before continuing, we need to perform the initial setup of interfaces, which includes the setup of IP addresses on both routing interfaces and local network interface E5/9 on the router SJB08M. These steps can be accomplished with the commands shown in Example 6-16 on page 201.

1. Enable OSPF globally in the global config level with the **router ospf** as shown in Example 6-19.

Example 6-19 Enabling OSPF globally

```
SJ_B08M#config t
SJ_B08M(config)#router ospf
SF_B16M#config t
SF_B16M(config)#router ospf
```

2. Globally configure area 0 with the **area 0** command in the ospf config level (config-ospf-router) as shown in Example 6-20.

Example 6-20 Global area configuration

```
SJ_B08M(config-ospf-router)#area 0
SF_B16M(config-ospf-router)#area 0
```

3. Assign the specific interface to the area in the interface config level (that is, config-if-e10000-8/1) with the **ip ospf area 0** command as shown in Example 6-21

Example 6-21 Assigning interfaces to area 0

```
SJ_B08M(config-ospf-router)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip ospf area 0
SF_B16M(config-ospf-router)#interface e 8/1
SF_B16M(config-if-e10000-8/1)#ip ospf area 0
```

When a specific port is assigned to an area, the subnets which are assigned to that port are included in the IP route table.

6.2.2 Redistribution

Redistribution is used to redistribute route information in the OSPF setup. All redistribution commands are executed in the ospf config level (config-ospf-router). The following options are available:

- Redistribution connected

  The **`redistribution connected`** command as shown in Example 6-22, is used to redistribute connected routes.

Example 6-22 Redistribution connected

```
SJ_B08M(config-if-e10000-1/1)#router ospf
SJ_B08M(config-ospf-router)#redistribution connected
*****************************************************************************
SF_B16M(config-if-e10000-8/1)#show ip route
Total number of IP routes: 6
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type
Uptime
1       10.64.208.0/20     DIRECT          mgmt 1      0/0           D    4d17h
2       172.31.1.12/30     DIRECT          eth 8/1     0/0           D    4d1h
3       172.31.1.16/30     DIRECT          eth 5/17    0/0           D    1d3h
4       192.168.10.0/30    172.31.1.13     eth 8/1     110/10        O2   0m5s
5       192.168.15.0/30    DIRECT          eth 9/6     0/0           D    2d19h
6       192.168.25.0/24    172.31.1.18     eth 5/17    120/2         R    1h13m
*****************************************************************************
SF_B16M(config-if-e10000-8/1)#router ospf
SF_B16M(config-ospf-router)#redistribution connected
*****************************************************************************
SJ_B08M(config)#show ip route
Total number of IP routes: 5
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
```

```
       Destination       Gateway        Port       Cost          Type
Uptime
1      10.64.208.0/20    DIRECT         mgmt 1     0/0           D    7d20h
2      172.31.1.12/30    DIRECT         eth 1/1    0/0           D    4d2h
3      172.31.1.16/30    172.31.1.14    eth 1/1    110/10        O2   0m28s
4      192.168.10.0/30   DIRECT         eth 5/9    0/0           D    2d20h
5      192.168.15.0/30   172.31.1.14    eth 1/1    110/10        O2   0m28s
```

As you can see, after enabling redistribution on one router, the routes from that router appear on the other router.

- Redistribution static:

  The **redistribution static** command is shown in Example 6-23.

Example 6-23 Redistribution static

```
SJ_B08M(config-ospf-router)#show ip route static
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       10.10.10.0/24      192.168.10.2    eth 5/9     1/1           S    0m12s
***********************************************************************
SJ_B08M(config)#router ospf
SJ_B08M(config-ospf-router)#redistribution static
***********************************************************************
SF_B16M(config)#show ip route ospf
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       10.10.10.0/24      172.31.1.13     eth 8/1     110/10        O2   22m2s
2       192.168.10.0/30    172.31.1.13     eth 8/1     110/10        O2   40m29s
```

As you can see, the static route from the SJ_B08M router is distributed by OSPF to the SF_B16M router.

- Redistribution rip

  The **redistribution rip** command is used to redistribute RIP routes from attached RIP domain. This means that RIP routes are sent out as OSPF LSAs (Link State Advertisement). The example is shown in Example 6-24.

Example 6-24 RIP redistribution

```
SF_B16M(config)#show ip route rip
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
```

```
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       192.168.25.0/24    172.31.1.18     eth 5/17    120/2         R    1h59m
***********************************************************************
SF_B16M(config)#router ospf
SF_B16M(config-ospf-router)#redistribution rip
***********************************************************************
SJ_B08M(config-ospf-router)#show ip route ospf
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type Uptime
1       172.31.1.16/30     172.31.1.14     eth 1/1     110/10        O2   43m35s
2       192.168.15.0/30    172.31.1.14     eth 1/1     110/10        O2   43m35s
3       192.168.25.0/24    172.31.1.14     eth 1/1     110/10        O2   0m10s
```

We can see that the RIP routes from SF_B16M are redistributed to the SJ_B08M router.

When the RIP domain is participating in the OSPF setup, the RIP enabled router can also send out OSPF routes as RIP updates. This is achieved with the **redistribution** command in the rip config level as shown in Example 6-25.

Example 6-25 OSPF routes distributed into RIP domain

```
SF_B16M(config)#router rip
SF_B16M(config-rip-router)#redistribute ospf
LA_B16S#show ip route rip
Start index: 1  B:BGP D:Connected  R:RIP  S:Static  O:OSPF *:Candidate default
        Destination     NetMask         Gateway         Port       Cost   Type
        10.10.10.0      255.255.255.0   172.31.1.17     1/17       2
        172.31.1.12     255.255.255.252 172.31.1.17     1/17       2
        192.168.10.0    255.255.255.252 172.31.1.17     1/17       2
        192.168.15.0    255.255.255.252 172.31.1.17     1/17       2
```

With RIP participating in the OSPF setup it is important that RIP routers are able to learn the default routes redistributed by OSPF. This is achieved with the **ip rip learn-default** command in interface config level as shown in Example 6-26.

Example 6-26 RIP learn-default

```
LA_B16S#config t
LA_B16S(config)#interface e 1/17
LA_B16S(config-if-e1000-1/17)#ip rip learn-default
```

This command can also be applied at the global level.

6.2.3 Displaying various OSPF information

In this section we describe commands that can be used to display various OSPF routing information.

OSPF route information

The OSPF route information can be displayed with the **show ip ospf route** command as shown in Example 6-27.

Example 6-27 OSPF route information

```
SJ_B08M(config)#show ip route ospf
Type Codes - B:BGP D:Connected I:ISIS O:OSPF R:RIP S:Static; Cost - Dist/Metric
ISIS Codes - L1:Level-1 L2:Level-2
OSPF Codes - i:Inter Area 1:External Type 1 2:External Type 2 s:Sham Link
        Destination        Gateway         Port        Cost          Type
Uptime
1       172.31.1.16/30     172.31.1.14     eth 1/1     110/10        O2   1h1m
2       192.168.15.0/30    172.31.1.14     eth 1/1     110/10        O2   1h1m
3       192.168.25.0/24    172.31.1.14     eth 1/1     110/10        O2   18m0s
```

This command is very useful as displaying all routes with the **show ip route** command can be very long, and because of this it is not s easy to read.

OSPF link state

The OSPF link state database can be displayed with the **show ip ospf database** command as shown in Example 6-28.

Example 6-28 OSPF link state

```
SJ_B08M(config)#show ip ospf database

Link States

Index Area ID         Type LS ID            Adv Rtr         Seq(Hex) Age  Cksum
1     0               Rtr  172.31.1.13      172.31.1.13     80000005 1917 0x8560
2     0               Rtr  172.31.1.14      172.31.1.14     80000005 1891 0x835f
3     0               Net  172.31.1.13      172.31.1.13     80000003 1917 0x33bf

Type-5 AS External Link States

Index Age  LS ID           Router           Netmask  Metric   Flag Fwd Address
1     1891 192.168.15.0    172.31.1.14      fffffffc 0000000a 0000 0.0.0.0
2     1917 192.168.10.0    172.31.1.13      fffffffc 0000000a 0000 0.0.0.0
```

```
3     1126 192.168.25.0    172.31.1.14     ffffff00 0000000a 0000 0.0.0.0
4     1891 172.31.1.16     172.31.1.14     fffffffc 0000000a 0000 0.0.0.0
5     717  10.10.10.0      172.31.1.13     ffffff00 0000000a 0000 0.0.0.0
```

OSPF area type information

The OSPF area type information can be displayed with the **`show ip ospf area`** command as shown in Example 6-29.

Example 6-29 OSPF area type information

```
SJ_B08M(config)#show ip ospf area
Number of Areas is 1

Indx Area            Type   Cost       SPFR       ABR   ASBR  LSA   Chksum(Hex)
1    0               normal 0          6          0     1     3     00013681
```

OSPF neighbor information

The OSPF neighbor information can be displayed with the **`show ip ospf neighbor`** command as shown in Example 6-30.

Example 6-30 OSPF neighbor info

```
SF_B16M(config)#show ip ospf neighbor
Number of Neighbors is 1, in FULL state 1

Port  Address         Pri State      Neigh Address   Neigh ID        Ev Opt Cnt
8/1   172.31.1.14     1   FULL/DR    172.31.1.13     172.31.1.13     6  66  0
SJ_B08M(config)#show ip ospf neighbor
Number of Neighbors is 1, in FULL state 1

Port  Address         Pri State      Neigh Address   Neigh ID        Ev Opt Cnt
1/1   172.31.1.13     1   FULL/BDR   172.31.1.14     172.31.1.14     5  66  0
```

In this example we see that **SF_B16M** is the designated router (DR) and **SJ_B08M** is the backup designated router (BDR).

OSPF interface information

The OSPF interface information, which includes area ID and adjacency info, can be displayed with the **`show ip ospf interface`** command as shown in Example 6-31.

Example 6-31 OSPF interface information

```
SJ_B08M(config)#show ip ospf interface

eth 1/1, OSPF enabled
```

```
     IP Address 172.31.1.13, Area 0
     Database Filter: Not Configured
     OSPF state DR, Pri 1, Cost 1, Options 2, Type broadcast Events 4
     Timers(sec): Transmit 1, Retrans 5, Hello 10, Dead 40
     DR:  Router ID 172.31.1.13       Interface Address 172.31.1.13
     BDR: Router ID 172.31.1.14       Interface Address 172.31.1.14
     Neighbor Count = 1, Adjacent Neighbor Count= 1
     Neighbor:         172.31.1.14 (BDR)
     Authentication-Key: None
     MD5 Authentication: Key None, Key-Id None, Auth-change-wait-time 300
```

OSPF virtual link information

The OSPF virtual link information can be displayed with the **`show ip ospf virtual-link`** command as shown in Example 6-32.

Example 6-32 OSPF virtual link information

```
SJ_B08M(config)#show ip ospf virtual-link
No ospf virtual-link entries available
```

In our example we did not define any OSPF virtual links.

OSPF trap information

The information about SNMP traps which were triggered by OSPF can be displayed with the **`show ip ospf trap`** command as shown in Example 6-33.

Example 6-33 OSPF trap information

```
SJ_B08M(config)#show ip ospf trap
Interface State Change Trap:                      Enabled
Virtual Interface State Change Trap:              Enabled
Neighbor State Change Trap:                       Enabled
Virtual Neighbor State Change Trap:               Enabled
Interface Configuration Error Trap:               Enabled
Virtual Interface Configuration Error Trap:       Enabled
Interface Authentication Failure Trap:            Enabled
Virtual Interface Authentication Failure Trap:    Enabled
Interface Receive Bad Packet Trap:                Enabled
Virtual Interface Receive Bad Packet Trap:        Enabled
Interface Retransmit Packet Trap:                 Disabled
Virtual Interface Retransmit Packet Trap:         Disabled
Originate LSA Trap:                               Disabled
Originate MaxAge LSA Trap:                        Disabled
Link State Database Overflow Trap:                Disabled
Link State Database Approaching Overflow Trap:    Disabled
```

OSPF border information

To display information about Area Border Routers (ABRs) or Autonomous System Boundary Routers (ASBRs), the **`show ip ospf border`** command can be used as shown in Example 6-34.

Example 6-34 OSPF border information

```
SJ_B08M(config)#show ip ospf border
      router ID       router type next hop router outgoing interface Area
1     172.31.1.14     ASBR        172.31.1.14     1/1                0
SF_B16M(config)#show ip ospf border
      router ID       router type next hop router outgoing interface Area
1     172.31.1.13     ASBR        172.31.1.13     8/1                0
```

OSPF configuration information

To display information about the OSPF general configuration, the **`show ip ospf config`** command can be used as shown in Example 6-35.

Example 6-35 OSPF configuration information

```
SJ_B08M(config)#show ip ospf config
Router OSPF: Enabled

Graceful Restart: Disabled
Graceful Restart Helper: Enabled
Graceful Restart Time: 120
Graceful Restart Notify Time: 0

Redistribution: Enabled
Default OSPF Metric: 10
OSPF Auto-cost Reference Bandwidth: Disabled
OSPF Redistribution Metric: Type2

OSPF External LSA Limit: 14447047

OSPF Database Overflow Interval: 0

RFC 1583 Compatibility: Enabled

Router id: 172.31.1.13
Interface State Change Trap:                      Enabled
Virtual Interface State Change Trap:              Enabled
Neighbor State Change Trap:                       Enabled
```

```
Virtual Neighbor State Change Trap:               Enabled
Interface Configuration Error Trap:               Enabled
Virtual Interface Configuration Error Trap:       Enabled
Interface Authentication Failure Trap:            Enabled
Virtual Interface Authentication Failure Trap:    Enabled
Interface Receive Bad Packet Trap:                Enabled
Virtual Interface Receive Bad Packet Trap:        Enabled
Interface Retransmit Packet Trap:                 Disabled
Virtual Interface Retransmit Packet Trap:         Disabled
Originate LSA Trap:                               Disabled
Originate MaxAge LSA Trap:                        Disabled
Link State Database Overflow Trap:                Disabled
Link State Database Approaching Overflow Trap:    Disabled

OSPF Area currently defined:
Area-ID          Area-Type Cost
0                normal    0
```

6.2.4 OSPF timers

There are several OSPF timers which can be adjusted if required. The following examples show which commands can be used to adjust those timers.

Hello interval

The interval between hello packets can be set from 1-65535 seconds (default value is 10 seconds), at the interface level in the ospf config level with the **`ip ospf hello-interval`** ***`value`*** command as shown in Example 6-36.

Example 6-36 OSPF hello interval

```
SJ_B08M(config)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip ospf hello-interval 20
```

Hello dead interval

The time that a neighbor router waits for a hello packet before it declares that the current router is down can be set from 1-65535 seconds (default value is 40 seconds), at the interface level in the ospf config level with the **`ip ospf dead-interval`** ***`value`*** command as shown in Example 6-37.

Example 6-37 OSPF hello dead interval

```
SJ_B08M(config)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip ospf dead-interval 50
```

Retransmit interval

The interval between retransmission of LSAs to adjacent routers for the particular interface can be set from 1-3600 seconds (default value is 5 seconds), on the interface level in the ospf config level with the **ip ospf retransmit-interval *value*** command as shown in Example 6-38.

Example 6-38 OSPF retransmit interval

```
SJ_B08M(config)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip ospf retransmit-interval 8
```

Transmit delay

The time for transmitting link state update packets on a particular interface can be set from 1-3600 seconds (default value is 1 second), on the interface level in the ospf config level with the **ip ospf transmit-delay *value*** command as shown in Example 6-39.

Example 6-39 OSPF transmit delay

```
SJ_B08M(config)#interface e 1/1
SJ_B08M(config-if-e10000-1/1)#ip ospf transmit-delay 3
```

Delay and hold timers

The SPF delay time defines how quickly after a topology change that the router will wait to start SPF calculation. The default time for this time is 5 seconds. The SPF hold time defines how long the router will wait before two SPF calculations. The default value for this time is 10 seconds. Both timers are defined in ospf config level (config-ospf-router) with the **timers spf *DelayValue HoldValue*** command as shown in Example 6-40.

Example 6-40 OSPF delay and hold timers

```
SJ_B08M(config)#router ospf
SJ_B08M(config-ospf-router)#timers spf 10 20
```

All these configuration changes can be removed or undone by using **no** in front of the respective commands.

All the timer information can be displayed using the **show ip ospf** command as show in Example 6-41.

Example 6-41 OSPF timers information

```
SF_B16M(config)#show ip ospf
OSPF Version                  Version 2
```

```
Router Id                     172.31.1.14
ASBR Status                   Yes
ABR Status                    No         (0)
Redistribute Ext Routes from  Connected RIP
Initial SPF schedule delay    0          (msecs)
Minimum hold time for SPFs    0          (msecs)
Maximum hold time for SPFs    0          (msecs)
External LSA Counter          5
External LSA Checksum Sum     000294a4
Originate New LSA Counter     19
Rx New LSA Counter            12
External LSA Limit            14447047
Database Overflow Interval    0
Database Overflow State :     NOT OVERFLOWED
RFC 1583 Compatibility :      Enabled
NSSA Translator:              Enabled
Graceful Restart:             Disabled,   timer 120
Graceful Restart Helper:      Enabled
```

6.2.5 Default route propagation

To propagate the default route into OSPF the **default-information-originate** command must be used in the ospf config level as OSPF does not propagate default routes by default. The command is shown in Example 6-42.

Example 6-42 OSPF default route propagation

```
SJ_B08M(config)#router ospf
SJ_B08M(config-ospf-router)#default-information-originate
```

6.2.6 OSPF load sharing

Up to eight equal paths can be configured for load sharing. The default is to use four paths. To change the number of paths the **ip load-sharing *value*** command in the config level can be used as shown in Example 6-43.

Example 6-43 OSPF load sharing

```
SJ_B08M(config)#router ospf
SJ_B08M(config-ospf-router)#ip load-sharing 5
```

6.2.7 OSPF stub area

OSPF routers within a stub area cannot send or receive external Link State Advertisements (LSAs) and must use a default route to the area's Are a Border Router (ABR) or Autonomous System Boundary Router (ASBR) to send traffic out of the area.

To define the stub area with the particular cost of entering and leaving that area, the **area** ***AreaValue*** **stub** ***StubValue*** command in ospf config level (config-ospf-router) can be used as shown in Example 6-44.

Example 6-44 OSPF stub area

```
SJ_B08M(config)#router ospf
SJ_B08M(config-ospf-router)#area 1 stub 10
```

In this example we defined cost 10 for area 1.

6.2.8 OSPF totally stubby area

By default, summary Link State Advertisements (LSA type 3) are sent into stub areas. You can further reduce the number of LSAs sent into a stub area by configuring a totally stubby area.

To define a totally stubby area, the **area** ***AreaValue*** **stub no-summary** command in ospf config level (config-ospf-router) can be used as shown in Example 6-45.

Example 6-45 OSPF totally stubby

```
SJ_B08M(config)#router ospf
SJ_B08M(config-ospf-router)#area 1 stub 10 no-summary
```

6.2.9 Virtual links

Virtual links can be defined to reach the backbone area (area 0) through a non-backbone area. A virtual link is not associated with the physical link, but tunnels packets through a non-backbone area to reach area 0. The following rules apply:

- Virtual links must be configured between two ABRs.
- The transit area through which virtual link is configured must have full routing information.
- The transit area cannot be a stub area.

Note: The use of virtual links can only be a temporary fix because it brings the additional complexity and troubleshooting to any internetwork communication.

6.3 BGP

In this section we assume that you are familiar with the basics of the BGP routing protocol. Examples are based on the network setup shown in Figure 6-3.

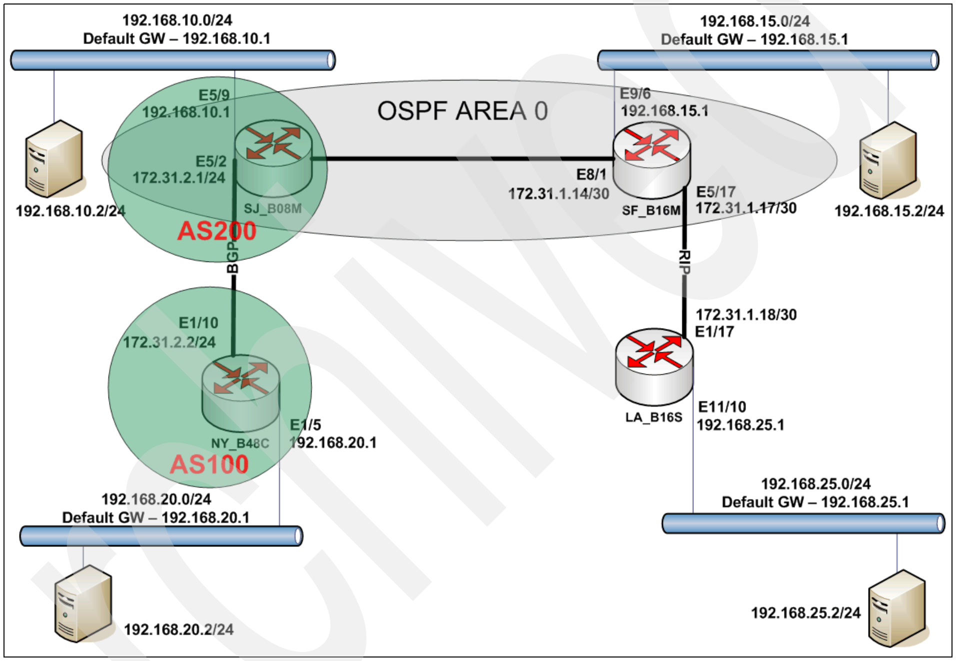


Figure 6-3 BGP lab network setup

In our example we have two autonomous systems (AS): AS100 with the router **NY_B48C** and AS 200 with the router **SJ_B08M**.

6.3.1 Configuring peers

To configure BGP peers, the following steps have to be performed. In our setup the commands must be executed on both routers participating in the BGP setup as shown in Figure 6-3. Before continuing our setup, we need to set up IP addresses as defined in Figure 6-3.

Activating BGP mode

With the **router bgp** command from the config context, BGP mode is activated as shown in Example 6-46.

Example 6-46 Activate BGP

```
NY_B48C(config)#router bgp
BGP: Please configure 'local-as' parameter in order to run BGP4.
SJ_B08M(config)#router bgp
BGP: Please configure 'local-as' parameter in order to run BGP4.
```

Defining local Autonomous System (AS)

The local AS can be defined with the **local-as LocalValue** command as shown in Example 6-47.

Example 6-47 Defining local AS

```
NY_B48C(config-bgp)#local-as 100
SJ_B08M(config-bgp)#local-as 200
```

The **value** defines if peering will be EBGP (External BGP) or IBGP (Internal BGP). If the value is the same on both sides, BGP will initiate an internal session, and if the values are different it will initiate external session.

Defining remote Autonomous System (AS)

The **neighbor IP remote-as RemoteValue** command can be used to define remote AS as shown in Example 6-48.

Example 6-48 Defining remote AS

```
NY_B48C(config)#router bgp
NY_B48C(config-bgp)#neighbor 172.31.2.1 remote-as 200
SJ_B08M(config)#router bgp
SJ_B08M(config-bgp)#neighbor 172.31.2.2 remote-as 100
```

As already mentioned if the **LocalValue** and **RemoteValue** are the same then BGP will initiate an internal session, otherwise BGP will initiate an external session.

6.3.2 Verifying peering

To verify the status of BGP peering, the **show ip bgp neighbors** command can be used as shown in Example 6-49.

Example 6-49 Verifying BGP peering

```
NY_B48C(config)#show ip bgp neighbors
    Total number of BGP Neighbors: 1
1   IP Address: 172.31.2.1, AS: 200 (EBGP), RouterID: 172.31.1.13, VRF: default-vrf
    State: ESTABLISHED, Time: 0h1m5s, KeepAliveTime: 60, HoldTime: 180
       KeepAliveTimer Expire in 46 seconds, HoldTimer Expire in 169 seconds
    Minimal Route Advertisement Interval: 0 seconds
       RefreshCapability: Received
    Messages:    Open    Update  KeepAlive Notification Refresh-Req
       Sent    : 1       0       2         0            0
       Received: 1       0       2         0            0
    Last Update Time: NLRI       Withdraw          NLRI       Withdraw
                  Tx: ---        ---          Rx: ---        ---
    Last Connection Reset Reason:Unknown
    Notification Sent:     Unspecified
    Notification Received: Unspecified
    Neighbor NLRI Negotiation:
      Peer Negotiated IPV4  unicast  capability
      Peer configured for IPV4 unicast  Routes
    As-path attribute count: 0
    Outbound Policy Group:
       ID: 1, Use Count: 1
    TCP Connection state: ESTABLISHED, flags:00000044 (0,0)
     Maximum segment size: 1460
    TTL check: 0, value: 0, rcvd: 64
       Byte Sent:   83, Received: 83
       Local host:  172.31.2.2, Local  Port: 179
       Remote host: 172.31.2.1, Remote Port: 8249
       ISentSeq: 4252753534  SendNext: 4252753618  TotUnAck:          0
       TotSent:           84  ReTrans:           0  UnAckSeq: 4252753618
       IRcvSeq:  1208541986  RcvNext:  1208542070  SendWnd:       64981
       TotalRcv:          84  DupliRcv:          0  RcvWnd:        65000
       SendQue:            0  RcvQue:            0  CngstWnd:       3102
***********************************************************************
SJ_B08M(config-bgp)#show ip bgp neighbors
    Total number of BGP Neighbors: 1
1   IP Address: 172.31.2.2, AS: 100 (EBGP), RouterID: 192.168.20.1, VRF: default-vrf
    State: ESTABLISHED, Time: 0h2m21s, KeepAliveTime: 60, HoldTime: 180
       KeepAliveTimer Expire in 31 seconds, HoldTimer Expire in 150 seconds
```

```
Minimal Route Advertisement Interval: 0 seconds
   RefreshCapability: Received
Messages:    Open    Update  KeepAlive Notification Refresh-Req
   Sent    : 1       0       3         0            0
   Received: 1       0       3         0            0
Last Update Time: NLRI       Withdraw         NLRI       Withdraw
              Tx: ---        ---           Rx: ---       ---
Last Connection Reset Reason:Unknown
Notification Sent:     Unspecified
Notification Received: Unspecified
Neighbor NLRI Negotiation:
  Peer Negotiated IPV4  unicast  capability
  Peer configured for IPV4 unicast  Routes
Neighbor AS4 Capability Negotiation:
As-path attribute count: 0
Outbound Policy Group:
   ID: 1, Use Count: 1
TCP Connection state: ESTABLISHED, flags:00000033 (0,0)
 Maximum segment size: 1460
TTL check: 0, value: 0, rcvd: 64
   Byte Sent:   102, Received: 102
   Local host:  172.31.2.1, Local  Port: 8249
   Remote host: 172.31.2.2, Remote Port: 179
   ISentSeq: 1208541986  SendNext: 1208542089  TotUnAck:          0
   TotSent:         103  ReTrans:           0  UnAckSeq: 1208542089
   IRcvSeq:  4252753534  RcvNext:  4252753637  SendWnd:       64981
   TotalRcv:        103  DupliRcv:          0  RcvWnd:        65000
   SendQue:           0  RcvQue:            0  CngstWnd:       3102
```

As you can see, peering was successfully established.

The State field has the possible values shown in Table 6-3.

Table 6-3 State files values

State	Explanation
Idle	Waiting for Start event (neighbor command to be typed in).
Connect	Waiting for the TCP connection to complete: 1. If TCP connect succeeds: Sends BGP open message to peer. Changes state to OpenSent. 2. If TCP connect fails: Continues to listen for a connection from remote peer. Changes state to Active.
Active	BGP is trying to acquire a peer by initiating a TCP connection.
OpenSent	Waiting for OPEN message from the peer. If OPEN message is received, KEEPALIVE is sent to the peer.
OpenConfirm	Waiting for a KEEPALIVE message from the peer. When KEEPALIVE message is received, the state changes to Established.
Established	In this state peers can exchange routing information.

6.3.3 Injecting routes

The BGP routing table is empty until routes are injected into it from the IGP (Interior Gateway Protocol). The injection can be achieved with two possible commands:

- The **redistribution** command
- The **network** command

If the **redistribution** command is used from IGP to BGP, this can result in publishing private network information to destinations outside of the Autonomous System (AS). This can also mean that you publish unregistered addresses which are used internally. For this reason, use full redistribution with caution.

To inject the route with the **network** command, this has to be executed in the BGP config context (config-bgp) as shown in Example 6-50.

Example 6-50 Injecting route

```
SJ_B08M(config)#router bgp
SJ_B08M(config-bgp)#network 10.10.10.0/24
SJ_B08M(config-bgp)#show ip bgp
Total number of BGP Routes: 1
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i internal, S stale
Origin codes: i - IGP, e - EGP, ? - incomplete
    Network            Next Hop        Metric LocPrf Weight Path
*>  10.10.10.0/24      192.168.10.2    1      100    32768  i
NY_B48C(config)#show ip bgp
Total number of BGP Routes: 1
Status codes: s suppressed, d damped, h history, * valid, > best, i internal, S stale
Origin codes: i - IGP, e - EGP, ? - incomplete
    Network            Next Hop        Metric LocPrf Weight Path
*>  10.10.10.0/24      172.31.2.1      1      100    0      200 i
```

6.3.4 Verifying BGP routes

The BGP routes can be displayed using the **show ip bgp** command as shown in Example 6-51.

Example 6-51 Displaying BGP routes

```
NY_B48C(config)#show ip bgp
Total number of BGP Routes: 1
Status codes: s suppressed, d damped, h history, * valid, > best, i internal, S stale
Origin codes: i - IGP, e - EGP, ? - incomplete
    Network            Next Hop        Metric LocPrf Weight Path
*>  10.10.10.0/24      172.31.2.1      1      100    0      200 i
```

In the output we can see the parameters shown in Table 6-4.

Table 6-4 Parameters

Parameter	Explanation
Status codes	The list of characters that display the route status will appear on the left column for each route: ► s - Suppressed entry ► * - Valid entry ► > - The prefix has multiple entries, this indicates the best route ► i - The prefix was learned by BGP
Origin codes	This character will indicate the route origin and will appear right to the Path field: ► i - The routes have been installed using the “network” command ► e - The routes have learned by EGP ► ? - The routes have been learned by redistribution
Network	The network address and prefix.

Parameter	Explanation
Next Hop	The next hop advertised by a BGP neighbor. Next hop means next BGP peer.
Path	The route's AS path.

6.3.5 Internal BGP

In this section we discuss update source loopback and route distribution.

Update source loopback

When comparing IBGP to EBGP, IBGP peers use loopbacks. The use of loopback decouples IBGP from relying on the availability of the physical interface for establishing TCP connections.

The use of loopbacks is also a requirement of IBGP as it requires a full mesh. Full mesh means there is a need for another path for a TCP connection to the peer.

The **`update-source loopback N`** command defines that loopback interface **`N`** will be used as a source IP address when BGP packets will be sent to a neighbor. This command is executed as part of the **`neighbor`** command in the BGP config context (config-bgp) as shown in Example 6-52. For this the loopback interface has to be defined.

Example 6-52 Update source loopback

```
NY_B48C(config)#interface loopback 1
NY_B48C(config-lbif-1)#ip address 10.1.1.1 255.255.255.255
NY_B48C(config-lbif-1)#router bgp
NY_B48C(config-bgp)#neighbor 172.31.2.1 update-source loopback 1
Please clear the neighbor session for the parameter change to take effect!
NY_B48C(config-bgp)#clear ip bgp neighbor all
```

Route distribution

By default the synchronization between IGP and BGP is disabled. Synchronization can be enabled with the **`synchronization`** command as shown in Example 6-53.

Example 6-53 BGP synchronization

```
NY_B48C(config)#router bgp
NY_B48C(config-bgp)#synchronization
```

If synchronization is enabled, the IGP route table must contain a route before BGP can advertise it.

The **no synchronization** command decouples BGP from IGP performance and ti also prevents any possible route-flapping.

The **auto-summary** command summarizes routes which are redistributed from IGP to BGP grouping individual subnets into a single prefix. The command is executed in the BGP config context as shown in Example 6-54.

Example 6-54 BGP auto-summary

```
NY_B48C(config)#router bgp
NY_B48C(config-bgp)#auto-summary
```

> **Important:** Using **auto-summary** can cause that block of addresses to be advertised even that some subnets are not existing in the network.

As the update to the outbound routes using new or changed outbound policy or filter is not automatically cleared, the **clear ip bgp neighbor** command must be used. The **clear ip bgp neighbor** command can be used in two different ways:

- **clear ip bgp neighbor all** – This will clear the whole neighbor information.
- **clear ip bgp neighbor NeighborAddress** — This will clear information about that particular network address only.

> **Attention:** Use **clear ip bgp neighbor all** with caution, because it can cause unwanted tearing down of critical routes with ISP peers.

Next hop

Each BGP route has a next hop. When the next hop cannot be reached, no routes will be installed. In an example where the IBGP router cannot reach the next hop to a given destination, it will not install any routes that have that next hop.

This situation can be fixed by using the **next-hop-self** command on the IBGP neighbor. In this case, the router will set itself as the next hop for any routes it advertises to that neighbor. The **next-hop-self** command is used as part of the **neighbor** command in the BGP config context as shown in Example 6-55.

Example 6-55 BGP next hop

```
NY_B48C(config)#router bgp
NY_B48C(config-bgp)#neighbor 172.31.2.1 next-hop-self
```

6.3.6 EBGP multihop

When EBGP routers cannot be connected directly, the **`ebgp-multihop Value`** command can be used to specify how many hops away the BGP peer is. **`Value`** represents the number of hops. The **`Ebgp-multihop`** command is specified as part of the **`neighbor`** in BGP config context as shown in Example 6-56.

Example 6-56 EBGP multihop

```
NY_B48C(config)#router bgp
NY_B48C(config-bgp)#neighbor 172.31.2.1 ebgp-multihop 2
Please clear the neighbor session for the parameter change to take
effect!
```

7

Stacking (B50G model only)

Stacking is an easy, cost-efficient way to provide some measure of switch scalability for customers that do not require the size or speed of a chassis-based switch.

Stacking is only available on the IBM Ethernet Switch B50G (4002-G5A, 4002AG5) model. No other g-series model, and no other series of the products, have that feature.

7.1 Basic stack building steps

Building a basic stack from scratch is straightforward:

1. Cable the units.
2. Attach a console cable to the unit you intend to be the “primary” stack unit.
3. Add **`stack enable`** to the configuration
4. Run the command **`stack secure-setup`**
5. Follow the prompts and watch the units reload
6. Do a **`write mem`** to commit the configuration.

While there are manual configuration methods available, they are not necessary for any but the most esoteric configurations, and only serve to add additional complexity.

7.2 Stack cabling

There are two topologies available for a stack: ring, and linear.

- The only difference between the two topologies is that in the linear topology, the units on the “ends” of the line each have one stacking port free. (This port can be used as a 10 Gb Ethernet uplink, if you desire).
- However, for reliability reasons, always use the ring topology, unless you have some compelling reason to use linear. For more details, see *IBM b-type Data Center Networking - Design and Best Practices Introduction*, SG24-7786.

7.3 Stack setup

In the following topics we show how to set up a stack.

7.3.1 Preparation

Cable the units using your desired topology.

Ensure that you do not have routing code loaded on the products. If you have Edge Layer 3 code loaded, the stack will not work properly. To do this, run the command **`show version`**. The third character of name of the current running code must be an “S”, not an “L”. We show sample output from this command in Example 7-1.

Example 7-1 show version output

```
gSeries_7(config)#
gSeries_7(config)#show version
  SW: Version 05.0.01aT7e1 Copyright (c) 1996-2008 Foundry Networks,
Inc.
      Compiled on Mar 09 2009 at 13:29:35 labeled as FGS05001a
      (3167396 bytes) from Primary fgs05001a.bin
      BootROM: Version 05.0.00T7e5 (FEv2)
  HW: Stackable FGS648P-POE (PROM-TYPE FGS648-STK-U)
=======================================================================
STACKID 1: SL 1: FGS-48G 48-port Management Module + PoE
      Serial  #: CH09092423
      P-ASIC  0: type D804, rev 01
      P-ASIC  1: type D804, rev 01
=======================================================================
STACKID 1: SL 2: FGS-2XGC 2-port 10G Module (2-CX4)
=======================================================================
  400 MHz Power PC processor 8245 (version 129/1014) 66 MHz bus
  512 KB boot flash memory
 8192 KB code flash memory
  256 MB DRAM
The system uptime is 58 seconds
The system : started=warm start   reloaded=by "reload"

gSeries_7(config)#
```

In our example, we are currently running FG**S**05001a. Because the third character is an “S”, this unit is ready to stack.

If your switches have been used for anything else, it is a best practice to erase the configuration entirely in all switches except the one you want to use as the stack controller. This is done from the enable-level command **`erase configuration.`**

7.3.2 Enable stacking

To enable stacking, go into config mode, and enter the command **`stack enable`**.

7.3.3 Secure setup

Leave config mode, and run the command **stack secure-setup**. (**stack secure-setup** is *not* a config-level command.)

During the secure-setup process, it will verify the units you want to include in the stack, ask you to choose a backup unit (used to configure the stack if the primary fails), and ask you to verify the stack IDs of each unit. After the process is complete, every unit in the stack will reload.

After the setup is over, execute a **write mem**. See Example 7-2.

Example 7-2 standard stack secure-setup

```
FGS648P-STK Switch#
FGS648P-STK Switch#show running
Current configuration:
!
ver 05.0.01aT7e1
!
stack unit 1
  module 1 fgs-48-port-management-module
  module 2 fgs-cx4-2-port-10g-module
  priority 128
!
end

FGS648P-STK Switch#
FGS648P-STK Switch#config t
FGS648P-STK Switch(config)#stack enable
Enable stacking. This unit actively participates in stacking
FGS648P-STK Switch(config)#exit
FGS648P-STK Switch#
FGS648P-STK Switch#stack secure-setup
FGS648P-STK Switch#Discovering the stack topology...

Current Discovered Topology - RING

Available UPSTREAM units
Hop(s)  Id   Type    Mac Address
1       new  FGS648  001b.ed87.9b40
2       new  FGS648  001b.ed87.9940

Available DOWNSTREAM units
Hop(s)  Id   Type    Mac Address
1       new  FGS648  001b.ed87.9940
```

```
2       new  FGS648  001b.ed87.9b40

Do you accept the topology (RING) (y/n)?: y

Selected Topology:
Active  Id   Type    Mac Address
        1    FGS648  001b.ed86.fa40

Selected UPSTREAM units
Hop(s)  Id   Type    Mac Address
1       3    FGS648  001b.ed87.9b40
2       2    FGS648  001b.ed87.9940

Selected DOWNSTREAM units
Hop(s)  Id   Type    Mac Address
1       2    FGS648  001b.ed87.9940
2       3    FGS648  001b.ed87.9b40

Do you accept the unit id's (y/n)?: y

FGS648P-STK Switch#

Election, was active, no role change, assigned-ID=1
Election, was active, no role change, assigned-ID=1
reset unit 2: diff bootup id=1
reset unit 3: diff bootup id=1

Config changed due to add/del units. Do write mem if you want to keep
it
```

Author note: Despite the switch telling you to do a **write mem** ***now, wait until the stack units come back as "Ready". If you do a*** **write mem** ***now, the written configuration will not have the new stack units in it.***

```
FGS648P-STK Switch#
FGS648P-STK Switch#

Election, was alone --> active, assigned-ID=1
Detect stack member 2 POE capable
Detect stack member 3 POE capable
Done hot swap: Set stack unit 3 to Ready
Election, was active, no role change, assigned-ID=1
Done hot swap: Set stack unit 2 to Ready
```

```
Config changed due to add/del units. Do write mem if you want to keep
it

FGS648P-STK Switch#
FGS648P-STK Switch#show running

Current configuration:
!
ver 05.0.01aT7e1
!
stack unit 1
  module 1 fgs-48-port-management-module
  module 2 fgs-cx4-2-port-10g-module
  priority 128
  stack-port 1/2/1 1/2/2
stack unit 2
  module 1 fgs-48-port-management-module
  module 2 fgs-cx4-2-port-10g-module
  stack-port 2/2/1 2/2/2
stack unit 3
  module 1 fgs-48-port-management-module
  module 2 fgs-cx4-2-port-10g-module
  stack-port 3/2/1 3/2/2
stack enable
!
end

FGS648P-STK Switch#
FGS648P-STK Switch#wr mem
Write startup-config done.
FGS648P-STK Switch#Flash Memory Write (8192 bytes per dot) .Flash to
Flash Done.
FGS648P-STK Switch#
FGS648P-STK Switch#
```

Note that we perform a **write mem** after the reboot of the other stack members completes. This commits the new stack configuration to flash.

To view the new stack, run **show stack**. See Example 7-3.

*Example 7-3 **example** show stack **output***

```
FGS648P-STK Switch#
FGS648P-STK Switch#
FGS648P-STK Switch#sh stack
alone: standalone, D: dynamic config, S: static config
ID   Type   Role    Mac Address    Pri State    Comment
1  S FGS648 active  001b.ed86.fa40 128 local    Ready
2  S FGS648 standby 001b.ed87.9940   0 remote   Ready
3  S FGS648 member  001b.ed87.9b40   0 remote   Ready

    active      standby
     +---+       +---+       +---+
 -2/1| 1 |2/2--2/1| 2 |2/2--2/1| 3 |2/2-
  |   +---+       +---+       +---+   |
  |                                  |
  |-------------------------------------|

Current stack management MAC is 001b.ed86.fa40
FGS648P-STK Switch#
```

Now that the stack has been built, take note of a few considerations:

- The different members of the stack are referred to by both their ID and their MAC address. There are LEDs on the front panel of the switches that will let you determine which stack member is which at a glance.
- All management will be done through the active stack controller. Even if you were to plug a console cable directly into a member that was not the active controller, all operations will be performed as if you are plugged into the active controller. If you want to perform an operation on a particular member, such as removing it from the stack, you must un-cable it first.
- You will have no control over the stacking interfaces. If you attempt to enter the interface configuration level on the ports used for stacking, the CLI will return an error.

7.4 Adding a member to a stack

To add a member to the stack, simply cable the new member into the topology, and re-run **stack secure-setup**. As with building a new stack, you must clear out the configuration of the new unit prior to adding it to the stack. If you do not, **secure-setup** might not discover the new device.

7.5 Removing a member from a stack

To remove a member from the stack, un-cable the stack member, login into the console of the removed device, and run the enable command **stack unconfigure me**.

7.6 Managing a stack

The topics show how to manage the stack.

7.6.1 Interface management

Interfaces are numbered on all g-series units by *module/slot/port*. As mentioned earlier, the 10 Gb interfaces (slot 2) that are used for stacking are inaccessible for configuration purposes. However, if you have cabled your stack using a linear topology, you can configure the 10Gb interfaces that are not cabled for stacking. The ports used for stacking are labeled in the **show stack** output.

7.6.2 General network management

From the perspective of all the protocols that run on the switch (STP, IP, and so on), a stack operates as a single unit;, it is only from a physical perspective that they are separate. Each unit in the stack shares the same configuration, the same management IP, the same set of VLANs and even the same switch MAC address.

7.6.3 Stack MAC address

By default, the MAC address of the stack is the built-in MAC of the active stack controller. In the event of a failover to the standby stack controller, the MAC address will change to that of the standby; this is undesired behavior in many management situations.

To set the MAC address to a fixed value, run the configuration command **stack mac *1234.5678.9012***. The best value to use is the MAC address of the active controller, but you can use any valid MAC that does not conflict with another address on the network.

8

Quality of Service

In this chapter we review the setup and configuration of Quality of Service (QoS) implementation on the IBM b-type switches and routers.

8.1 FastIron

In this section we show examples of QoS implementation on the FastIron software platform, which is used in s-series, g-series, and x-series switches. We assume that you are familiar with the basics of QoS.

8.1.1 Assigning QoS priorities to the incoming traffic

By default, all untagged traffic coming into switches is placed into the best effort queue *qosp0*. This queue has the lowest priority.

It is possible to assign QoS priority to untagged incoming traffic based on the incoming switch port, whether ingress port or based on the MAC address. The necessary commands are shown in Example 8-1.

Example 8-1 Assigning QoS priorities

```
LA_B16S#config t
LA_B16S(config)#interface e 11/10
LA_B16S(config-if-e1000-11/10)#priority 5

LA_B16S#config t
LA_B16S(config)#
LA_B16S(config)#static-mac-address 1122.3344.EEFF ethernet 11/10 priority 4
```

Both commands can be disabled with the use of **no** before the respective command.

To see the QoS priority queue profiles, use the **show qos-profiles** command as seen in Example 8-2.

Example 8-2 Viewing the QoS profile settings

```
telnet@LA_B16S(config)#show qos-profiles all
bandwidth scheduling mechanism: weighted priority
Profile qosp7     : Priority7(Highest) bandwidth requested  75% calculated  75%
Profile qosp6     : Priority6          bandwidth requested   7% calculated   7%
Profile qosp5     : Priority5          bandwidth requested   3% calculated   3%
Profile qosp4     : Priority4          bandwidth requested   3% calculated   3%
Profile qosp3     : Priority3          bandwidth requested   3% calculated   3%
Profile qosp2     : Priority2          bandwidth requested   3% calculated   3%
Profile qosp1     : Priority1          bandwidth requested   3% calculated   3%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   3% calculated   3%
telnet@LA_B16S(config)#
```

As you can see, *qosp0* has the lower resources allocated and is considered “best effort,” while the *qosp7* queue profile has the most resources allocated.

> **Note:** When there are no port based VLANs defined on the FastIron switch, then the **static-mac-address** command is available in the global config level. In a case when port based VLANs are defined, this command is available in VLAN config level.
>
> The **Priority** command is always executed on the interface level.

8.1.2 Marking

Marking is the process when the Layer 2 CoS and/or Layer 3 DSCP values of the packet are changed before leaving the switch. For example, marking can also be used when the packet arrives from the device that does not support CoS and/or DSCP tagging.

Even if marking is not used, the switch still performs mappings to HW queues. Marking is performed using ACLs.

8.1.3 Configuring DSCP based QoS

By default, DSCP honoring is not enabled in FastIron switches for switched and routed traffic.

To honor DSCP on the g-series switches, use the **trust dscp** command on the interface config level as shown in Example 8-3.

Example 8-3 DSCP on g-series

```
B48G(config-if-e1000-11)trust dscp
```

To honor DSCP on s-series switches, it can be enabled using ACLs as shown in Example 8-4.

Example 8-4 DSCP on s-series

```
LA_B16S(config)#access-list 101 permit ip any any dscp-cos-mapping
```

With such an ACL, all IP traffic will honor DSCP information in the IP packet when mapping to HW queues.

Both commands can be disabled with the use **no** before the respective command.

8.1.4 Configuring the QoS mappings

There are two QoS mappings used in the FastIron switches:

- DSCP → Internal Forwarding Priority
- Internal Forwarding Priority → Hardware Forwarding Queue

The DSCP and Internal Forwarding Priority mappings can be seen using the **show qos-tos** command. The default mappings are shown in Example 8-5.

Example 8-5 Viewing the DSCP and Internal Forwarding Priority default mappings

```
telnet@LA_B16S(config)#show qos-tos
DSCP-->Traffic-Class map: (DSCP = d1d2: 00, 01...63)

     d2|  0   1   2   3   4   5   6   7   8   9
  d1   |
  -----+-----------------------------------------
   0   |  0   0   0   0   0   0   0   0   1   1
   1   |  1   1   1   1   1   1   2   2   2   2
   2   |  2   2   2   2   3   3   3   3   3   3
   3   |  3   3   4   4   4   4   4   4   4   4
   4   |  5   5   5   5   5   5   5   5   6   6
   5   |  6   6   6   6   6   6   7   7   7   7
   6   |  7   7   7   7

Traffic-Class-->802.1p-Priority map (use to derive
DSCP--802.1p-Priority):

Traffic | 802.1p
Class   | Priority
--------+---------
   0    |    0
   1    |    1
   2    |    2
   3    |    3
   4    |    4
   5    |    5
   6    |    6
   7    |    7
--------+---------
```

From the foregoing example, you can see that DSCP markings of 00 - 07 are placed in an Internal Forwarding Priority (IFP) of 0, markings of 08 - 15 into an IFP of 1, markings of 16 - 13 into an IFP of 2, and so forth.

From there, the IFPs are mapped to the Hardware Forwarding Queues that we saw earlier in this chapter. In the default settings, IFP 0 is mapped to *qosp0*, IFP 1 is mapped to *qosp1*, IFP 2 is mapped to *qosp2*, and so forth.

It is possible to change the default mappings for DSCP to Internal Forwarding Priority (0 - 7) using the command shown in Example 8-6.

Example 8-6 DSCP to Internal Forwarding Priority

```
LA_B16S(config)#qos-tos map dscp-priority 0 2 3 4 to 1
```

It is possible to map up to 8 DSCP priority values to one internal forwarding priority.

It is also possible to change default mappings from Internal Forwarding Priority to Hardware Forwarding Queue (qosp0 - qosp7) using the command shown in Example 8-7.

Example 8-7 Internal Forwarding Priority to Hardware Forwarding Queue

```
LA_B16S(config)#qos tagged-priority 3 qosp0
802.1p priority 3 mapped to qos profile qosp0
```

In this example, we changed the mapping for 802.1p priority 3 to the qosp0 HW queue. To check our settings, we use the **show qos-tos** command as seen in Example 8-8.

Example 8-8 Viewing the QoS DSCP settings after our configuration changes

```
telnet@LA_B16S(config)#show qos-tos
DSCP-->Traffic-Class map: (DSCP = d1d2: 00, 01...63)

     d2|  0   1   2   3   4   5   6   7   8   9
  d1   |
  -----+-----------------------------------------
   0   |  1   0   1   1   1   0   0   0   1   1
   1   |  1   1   1   1   1   1   2   2   2   2
   2   |  2   2   2   2   3   3   3   3   3   3
   3   |  3   3   4   4   4   4   4   4   4   4
   4   |  5   5   5   5   5   5   5   5   6   6
   5   |  6   6   6   6   6   6   7   7   7   7
   6   |  7   7   7   7

Traffic-Class-->802.1p-Priority map (use to derive
DSCP--802.1p-Priority):
```

```
Traffic | 802.1p
Class   | Priority
--------+---------
   0    |    0
   1    |    1
   2    |    2
   3    |    0
   4    |    4
   5    |    5
   6    |    6
   7    |    7
--------+---------
```

Both commands can be disabled with the use **`no`** before the respective command.

8.1.5 Scheduling

Scheduling is used to define how internal forwarding queues are handled. Several queueing methods can be selected:

- Weighted Round Robin (WRR): With this method, all queues are serviced in each cycle. In each rotation, service forwards a specific number of packets from each queue based on the wight assigned to the particular queue.
  To select the WRR scheduling method, use the command shown in Example 8-9. The command is executed on the global config level.

Example 8-9 Selecting WRR

```
LA_B16S(config)#qos mechanism weighted
bandwidth scheduling mechanism: weighted priority
Profile qosp7     : Priority7(Highest) bandwidth requested  75% calculated  75%
Profile qosp6     : Priority6          bandwidth requested   7% calculated   7%
Profile qosp5     : Priority5          bandwidth requested   3% calculated   3%
Profile qosp4     : Priority4          bandwidth requested   3% calculated   3%
Profile qosp3     : Priority3          bandwidth requested   3% calculated   3%
Profile qosp2     : Priority2          bandwidth requested   3% calculated   3%
Profile qosp1     : Priority1          bandwidth requested   3% calculated   3%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   3% calculated   3%
LA_B16S(config)#
```

- Strict Priority (SP): This method will strictly serve queues based on the priority. This can mean that it serves as many packets as possible before moving to the next queue. To select the SP scheduling method, use the command shown in Example 8-10. The command is executed on the global config level.

Example 8-10 Selecting SP

```
LA_B16S(config)#qos mechanism strict
bandwidth scheduling mechanism: strict priority
Qos profile bandwidth percentages are ignored
LA_B16S(config)#
LA_B16S(config)#no qos mechanism strict
bandwidth scheduling mechanism: weighted priority
Profile qosp7     : Priority7(Highest) bandwidth requested  75% calculated  75%
Profile qosp6     : Priority6          bandwidth requested   7% calculated   7%
Profile qosp5     : Priority5          bandwidth requested   3% calculated   3%
Profile qosp4     : Priority4          bandwidth requested   3% calculated   3%
Profile qosp3     : Priority3          bandwidth requested   3% calculated   3%
Profile qosp2     : Priority2          bandwidth requested   3% calculated   3%
Profile qosp1     : Priority1          bandwidth requested   3% calculated   3%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   3% calculated   3%
LA_B16S(config)#
```

- Hybrid WRR and SP - this method combines WRR and SP scheduling. In this combination, by default the SP method is assigned to queues qosp7 and qosp6 and WRR method for queues qosp5 to qosp0. To select the WRR+SP scheduling method, use the command shown in Example 8-11. The command is executed on the global config level.

Example 8-11 Selecting WRR+SP

```
LA_B16S(config)#qos mechanism mixed-sp-wrr
bandwidth scheduling mechanism: mixed weighted priority with strict priority
Profile qosp7     : Priority7(Highest) Set as strict priority
Profile qosp6     : Priority6          Set as strict priority
Profile qosp5     : Priority5          bandwidth requested  25% calculated  25%
Profile qosp4     : Priority4          bandwidth requested  15% calculated  15%
Profile qosp3     : Priority3          bandwidth requested  15% calculated  15%
Profile qosp2     : Priority2          bandwidth requested  15% calculated  15%
Profile qosp1     : Priority1          bandwidth requested  15% calculated  15%
Profile qosp0     : Priority0(Lowest)  bandwidth requested  15% calculated  15%
LA_B16S(config)#
LA_B16S(config)#no qos mechanism mixed-sp-wrr
bandwidth scheduling mechanism: weighted priority
Profile qosp7     : Priority7(Highest) bandwidth requested  75% calculated  75%
Profile qosp6     : Priority6          bandwidth requested   7% calculated   7%
Profile qosp5     : Priority5          bandwidth requested   3% calculated   3%
Profile qosp4     : Priority4          bandwidth requested   3% calculated   3%
Profile qosp3     : Priority3          bandwidth requested   3% calculated   3%
Profile qosp2     : Priority2          bandwidth requested   3% calculated   3%
Profile qosp1     : Priority1          bandwidth requested   3% calculated   3%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   3% calculated   3%
LA_B16S(config)#
```

All commands can be disabled with the use **no** before the respective command.

Note: The default scheduling method is WRR. When **no** is used to disable other scheduling methods, the WRR method will be restored as shown in Example 8-10 on page 239 and Example 8-11 on page 239.

8.1.6 Configuring the QoS queues

Certain parameters can be changed for QoS queues:

- Queue name:

  The default names of the queues are qosp7, qosp6, qosp5, qosp4, qosp3, qosp2, qosp1, and qosp0. To change the queue name, use the command shown in Example 8-12. This command is executed on the global level. The new name can be up to 32 characters long.

Example 8-12 Queue name

```
LA_B16S(config)#qos name qosp7 top-priority
LA_B16S(config)#
LA_B16S(config)#show qos-profiles all
bandwidth scheduling mechanism: weighted priority
Profile top-priori: Priority7(Highest) bandwidth requested  75% calculated  75%
Profile qosp6     : Priority6          bandwidth requested   7% calculated   7%
Profile qosp5     : Priority5          bandwidth requested   3% calculated   3%
Profile qosp4     : Priority4          bandwidth requested   3% calculated   3%
Profile qosp3     : Priority3          bandwidth requested   3% calculated   3%
Profile qosp2     : Priority2          bandwidth requested   3% calculated   3%
Profile qosp1     : Priority1          bandwidth requested   3% calculated   3%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   3% calculated   3%
LA_B16S(config)#
LA_B16S(config)#qos name qosp7 top-pri
Error - qosp7 is not a current qos profile name
Current qos profile names are qosp0    qosp1    qosp2    qosp3    qosp4
qosp5    qosp6    top-priority
LA_B16S(config)#
```

After the queue is renamed, the new name has to be used in configuration commands. If you try to use old name, you will see an error as shown in Example 8-12. To return to the original name, you must execute another **qos name** command with the current and original name as parameters.

Note: Preferably use a name length of up to 10 characters so that full names can be displayed in **show** commands.

- Minimum bandwidth percentages of the WRR queues:

  It is possible to adjust default minimum bandwidth percentage of WRR queues. This can be achieved with the command shown in Example 8-13 in the global config level.

Example 8-13 Minimum bandwidth % for WRR queues

```
LA_B16S(config)#qos profile qosp7 30 qosp6 20 qosp5 10 qosp4 10 qosp3 8 qosp2 8 qosp1 8 qosp0 6
bandwidth scheduling mechanism: weighted priority
Profile qosp7     : Priority7(Highest) bandwidth requested  30% calculated  30%
Profile qosp6     : Priority6          bandwidth requested  20% calculated  20%
Profile qosp5     : Priority5          bandwidth requested  10% calculated  10%
Profile qosp4     : Priority4          bandwidth requested  10% calculated  10%
Profile qosp3     : Priority3          bandwidth requested   8% calculated   8%
Profile qosp2     : Priority2          bandwidth requested   8% calculated   8%
Profile qosp1     : Priority1          bandwidth requested   8% calculated   8%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   6% calculated   6%
LA_B16S(config)#
LA_B16S(config)#no qos profile qosp7 30 qosp6 20 qosp5 10 qosp4 10 qosp3 8 qosp2 8 qosp1 8 qosp0 6
bandwidth scheduling mechanism: weighted priority
Profile qosp7     : Priority7(Highest) bandwidth requested  75% calculated  75%
Profile qosp6     : Priority6          bandwidth requested   7% calculated   7%
Profile qosp5     : Priority5          bandwidth requested   3% calculated   3%
Profile qosp4     : Priority4          bandwidth requested   3% calculated   3%
Profile qosp3     : Priority3          bandwidth requested   3% calculated   3%
Profile qosp2     : Priority2          bandwidth requested   3% calculated   3%
Profile qosp1     : Priority1          bandwidth requested   3% calculated   3%
Profile qosp0     : Priority0(Lowest)  bandwidth requested   3% calculated   3%
LA_B16S(config)#
```

All queues have to be included in the command and values for all the queues has to total 100%. You can use **no** in front of a command to return to the original values, as shown in Example 8-13.

- Bandwidth Allocations of the Hybrid WRR and SP Queues:

  It is also possible to change the bandwidth percentage in the WRR and SP scheduling method. This can be achieved with the command shown in Example 8-14 in the global config level.

Example 8-14 Queue bandwidth % for WRR+SP queues

```
LA_B16S(config)#qos profile qosp7 sp qosp6 sp qosp5 35 qosp4 25 qosp3 10 qosp2 10 qosp1 10 qosp0 10
bandwidth scheduling mechanism: mixed weighted priority with strict priority
Profile qosp7     : Priority7(Highest) Set as strict priority
Profile qosp6     : Priority6          Set as strict priority
Profile qosp5     : Priority5          bandwidth requested  35% calculated  35%
Profile qosp4     : Priority4          bandwidth requested  25% calculated  25%
Profile qosp3     : Priority3          bandwidth requested  10% calculated  10%
Profile qosp2     : Priority2          bandwidth requested  10% calculated  10%
Profile qosp1     : Priority1          bandwidth requested  10% calculated  10%
Profile qosp0     : Priority0(Lowest)  bandwidth requested  10% calculated  10%
LA_B16S(config)#
LA_B16S(config)#no qos profile qosp7 sp qosp6 sp qosp5 35 qosp4 25 qosp3 10 qosp2 10 qosp1 10 qosp0 10
bandwidth scheduling mechanism: mixed weighted priority with strict priority
Profile qosp7     : Priority7(Highest) Set as strict priority
Profile qosp6     : Priority6          Set as strict priority
Profile qosp5     : Priority5          bandwidth requested  25% calculated  25%
Profile qosp4     : Priority4          bandwidth requested  15% calculated  15%
Profile qosp3     : Priority3          bandwidth requested  15% calculated  15%
Profile qosp2     : Priority2          bandwidth requested  15% calculated  15%
Profile qosp1     : Priority1          bandwidth requested  15% calculated  15%
Profile qosp0     : Priority0(Lowest)  bandwidth requested  15% calculated  15%
LA_B16S(config)#LA_B16S(config)#
```

All queues have to be included in the command and values for all queues except qosp7 and qosp6 have to total 100%. You can use **no** in front of the command to return to the original values as shown in Example 8-14.

> **Note:** This command can only be executed if the WRR and SP scheduling method is enabled. Only queues qosp7 and qosp6 support SP mode, all other queues are part of the WRR model and have to have bandwidth defined.

8.1.7 Viewing QoS settings

To view the QoS settings, use the command shown in Example 8-15.

Example 8-15 QoS settings

```
LA_B16S(config)#show qos-profiles all
bandwidth scheduling mechanism: mixed weighted priority with strict priority
Profile qosp7     : Priority7(Highest) Set as strict priority
Profile qosp6     : Priority6          Set as strict priority
Profile qosp5     : Priority5          bandwidth requested  25% calculated  25%
Profile qosp4     : Priority4          bandwidth requested  15% calculated  15%
Profile qosp3     : Priority3          bandwidth requested  15% calculated  15%
Profile qosp2     : Priority2          bandwidth requested  15% calculated  15%
Profile qosp1     : Priority1          bandwidth requested  15% calculated  15%
Profile qosp0     : Priority0(Lowest)  bandwidth requested  15% calculated  15%
LA_B16S(config)#
```

The QoS DSCP based settings can be displayed with the command shown in Example 8-16. The command will display current mappings for DSCP to forwarding priority, and then forwarding priority to 802.1p priority.

Example 8-16 QoS DSCP settings

```
LA_B16S(config)#show qos-tos
DSCP-->Traffic-Class map: (DSCP = d1d2: 00, 01...63)

     d2|  0   1   2   3   4   5   6   7   8   9
  d1   |
  -----+----------------------------------------
   0   |  1   0   1   1   1   0   0   0   1   1
   1   |  1   1   1   1   1   1   2   2   2   2
   2   |  2   2   2   2   3   3   3   3   3   3
   3   |  3   3   4   4   4   4   4   4   4   4
   4   |  5   5   5   5   5   5   5   5   6   6
   5   |  6   6   6   6   6   6   7   7   7   7
   6   |  7   7   7   7

Traffic-Class-->802.1p-Priority map (use to derive
DSCP--802.1p-Priority):

Traffic | 802.1p
Class   | Priority
--------+---------
   0    |    0
   1    |    1
   2    |    2
   3    |    3
   4    |    4
   5    |    5
   6    |    6
   7    |    7
--------+---------

LA_B16S(config)#
```

The d1 and d2 value together represent the DSCP value. For example, if we want to know the DSCP mapping for value 63, we need to look under d1=6 and d2=3 and we see that we get the value 7.

8.1.8 Rate limiting: s-series and x-series

Rate limiting is used to limit the maximum number of bytes the port can receive. Everything above this limit will be dropped. Rate limiting is performed in the hardware.

Rate limiting is specified in bits per second (bps) and it applies to a one second interval. Unused bandwidth is not carried over to the next interval.

> **Attention:** Avoid using rate limiting on ports that are used for Spanning Tree Protocol (STP) control traffic because this can disrupt proper operation of STP.

There are several configuration rules for rate limiting setup:

- Rate limiting is available only on inbound ports.
- Fixed rate limiting is not supported on 10-Gigabit Ethernet ports.
- Fixed rate limiting is not supported on tagged ports in the base Layer 3 and full Layer 3 images.
- The rate limit on IPv6 hardware takes several seconds to take effect at higher configured rate limit values. For example, if the configured rate limit is 750 Mbps, line-rate limiting can take up to 43 seconds to take effect.

Configuring a port-based rate limiting policy

To configure port based rate limiting policy, use the command shown in Example 8-17. This command is executed on the interface config level.

Example 8-17 Port based rate limit

```
LA_B16S(config)#interface e 11/10
LA_B16S(config-if-e1000-11/10)#
LA_B16S(config-if-e1000-11/10)#rate input fixed 400000
Rate Limiting on Port 11/10 - Config: 400000 bps, Actual: 400 Kbps
LA_B16S(config-if-e1000-11/10)#
LA_B16S(config-if-e1000-11/10)#no rate input fixed 400000
LA_B16S(config-if-e1000-11/10)#
```

In our example, we set 400000 bps (50000 bytes) limit. To disable rate limiting on the port level, use the **no** command as shown in Example 8-17.

> **Note:** The minimum rate limit level is 64 Kbps.

Configuring an ACL-based rate limiting policy

ACL based rate limiting applies to the policy where rate limiting is applied to the matching IP traffic. ACL based rate limiting is configured with the use of traffic policies which are then referenced in the ACL. The same traffic policies can be referenced in many ACLs.

The traffic policies become active on the ports to which the ACLs are bound.

Displaying the fixed rate limiting configuration

To display the fixed rate limiting configuration, use the command shown in Example 8-18.

Example 8-18 Displaying rate limiting configuration

```
LA_B16S(config)#show rate-limit fixed
Total rate-limited interface count: 3.
 Port    Configured Input Rate       Actual Input Rate
11/10                  400000                  400000
11/11                   70000                   70000
11/12                  800000                  800000
LA_B16S(config)#
```

8.1.9 Rate shaping: s-series and x-series

Rate shaping can be performed on outbound traffic and can be used to control the bandwidth of outbound traffic. With this approach, traffic can be smoother for the neighboring devices.

Before sending out, packets are stored in the buffers and then forwarded out at the defined rate. Rate shaping is performed at a 1 byte token level.

There are several configuration rules for rate shaping setup:

- Outbound rate shapers can be configured only on physical ports, not on virtual or loopback ports.
- For trunk ports, the rate shaper must be configured on individual ports of a trunk using the **`config-trunk-ind`** command (trunk configuration level); you cannot configure a rate shaper for a trunk itself, only on the individual trunk ports.
- When outbound rate shaping is enabled on a port on an IPv4 device, the port’s QoS queuing method (qos mechanism) will be strict mode. This applies to IPv4 devices only. On IPv6 devices, the QoS mechanism is whatever method is configured on the port, even when outbound rate shaping is enabled.
- You can configure a rate shaper for a port and for the individual priority queues of that port. However, if a port rate shaper is configured, that value overrides the rate shaper value of a priority queue if the priority queue’s rate shaper is greater than the rate shaper for the port.
- On s-series devices, configured rate shaper values are rounded up to the nearest multiple of 651 Kbps. The maximum configurable limit is 2665845 Kbps.

- On x-series devices, configured rate shaper values are rounded to the nearest values programmable by the hardware. You can specify a value up to the port’s line rate, which is 1000000000 for 10 GbE ports.

Configuring outbound rate shaping for a port

To configure the outbound rate shaping on a particular port, use the command shown in Example 8-19. This command is executed on the interface config level.

Example 8-19 Port rate shaping

```
LA_B16S(config)#interface e 11/10
LA_B16S(config-if-e1000-11/10)#
LA_B16S(config-if-e1000-11/10)#rate-limit output shaping 1300
Outbound Rate Shaping on Port 11/10 Config: 1300 Kbps, Actual: 1302
Kbps
LA_B16S(config-if-e1000-11/10)#
LA_B16S(config-if-e1000-11/10)#no rate-limit output shaping 1300
LA_B16S(config-if-e1000-11/10)#
```

The behavior of the above example is different on s-series and x-series devices:

- On s-series, the configured outbound rate shaping of 1300 Kbps is automatically rounded to 1302Kbps, the nearest multiple of 651 Kbps.
- On x-series, the configured outbound rate shaping of 1300 Kbps is automatically rounded to 1344 Kbps, which is the nearest value programmable by hardware.

To disable rate shaping on the port level, use the **no** command as shown in Example 8-19.

Configuring outbound rate shaping for a specific priority

To configure outbound rate shaping for a specific priority level, use the command shown in Example 8-20. This command is executed on the interface config level.

Example 8-20 Priority rate shaping

```
LA_B16S(config)#interface e 11/12
LA_B16S(config-if-e1000-11/12)#
LA_B16S(config-if-e1000-11/12)#rate-limit output shaping 600 priority 5
Outbound Rate Shaping on Port 11/12 for Priority 5
         Config: 600 Kbps, Actual: 651 Kbps
LA_B16S(config-if-e1000-11/12)#
LA_B16S(config-if-e1000-11/12)#no rate-limit output shaping 600
priority 5
LA_B16S(config-if-e1000-11/12)#
```

In our example, we defined outbound rate shaping of 600 Kbps for priority 5, which was automatically rounded to 651 Kbps for s-series or 504 Kbps for x-series. To disable rate shaping on the priority level, use the **no** command as shown in Example 8-20.

Configuring outbound rate shaping for a trunk port

It is possible to define rate shaping for every port in the trunk, which is supported on static and LACP trunks. The example commands of such a configuration are shown in Example 8-21. The command for rate shaping of trunk ports is executed on the trunk config level.

Example 8-21 Trunk rate shaping

```
LA_B16S(config)#trunk e 11/14 to 11/17
Trunk will be created in next trunk deploy.
LA_B16S(config)#
LA_B16S(config)#trunk deploy
LA_B16S(config)#
LA_B16S(config)#trunk e 11/14 to 11/17
LA_B16S(config-trunk-11/14-11/17)#
LA_B16S(config-trunk-11/14-11/17)#config-trunk-ind
Trunk port monitoring, if any, has been been removed
LA_B16S(config-trunk-11/14-11/17)#
LA_B16S(config-trunk-11/14-11/17)#rate-limit output shaping ethe 11/15 600
Outbound Rate Shaping on Port 11/15 Config: 600 Kbps, Actual: 651 Kbps
LA_B16S(config-trunk-11/14-11/17)#
LA_B16S(config-trunk-11/14-11/17)#rate-limit output shaping ethe 11/17 1300
Outbound Rate Shaping on Port 11/17 Config: 1300 Kbps, Actual: 1302 Kbps
LA_B16S(config-trunk-11/14-11/17)#
LA_B16S(config-trunk-11/14-11/17)#no rate-limit output shaping ethe 11/17 1300
LA_B16S(config-trunk-11/14-11/17)#
```

In this example, we defined trunk on ports *E11/14* to *E11/17* and then defined rate shaping for the ports *E11/15* and *E11/17*. To remove rate shaping on a particular trunk port, you can use the **no** command in front of the respective rate shaping command as shown in Example 8-21.

Displaying rate shaping configurations

To display the current rate shaping configuration, use the command shown in Example 8-22.

Example 8-22 Rate shaping configuration

```
LA_B16S(config)#show rate-limit output-shaping
Outbound Rate Shaping Limits in Kbps:
 Port   PortMax   Prio0   Prio1   Prio2   Prio3   Prio4   Prio5   Prio6   Prio7
11/12         -       -       -       -       -       -     651       -       -
```

```
11/15      651      -       -       -       -       -       -       -       -
LA_B16S(config)#
```

8.1.10 Rate limiting: g-series

Rate limiting is used to limit the maximum number of bytes that the port can receive. Everything above this limit will be dropped. Rate limiting is performed in the hardware.

Rate limiting is specified in bits per second (bps) and it applies to one second interval. Unused bandwidth is not carried over to the next interval.

> **Attention:** Avoid using rate limiting on ports that are used for Spanning Tree Protocol (STP) control traffic because this can disrupt proper operation of STP.

Configuring fixed rate limiting on inbound ports

Inbound rate limiting is supported on:

- GbE ports
- 10-GbE ports
- Trunk ports

Inbound rate limiting is not supported on:

- Ports on which LACP is enabled
- Virtual interfaces
- Loopback interfaces

The minimum rate limit is 65 Kbps and the maximum is defined by the port type (1G, 10G). An example of the rate limiting command is shown in Example 8-23. The command is executed on the interface config level.

Example 8-23 Inbound port rate limiting

```
LG_B50G(config)#interface e 1/1/20
LG_B50G(config-if-e1000-1/1/20)#rate-limit input fixed 500000
Rate Limiting on Port 1/1/20 - Config: 500000 Kbps, Actual: 500000 Kbps
LG_B50G(config-if-e1000-1/1/20)#
```

Rate limiting can be disabled with the use of `no` in front of the respective enable command.

Configuring fixed rate limiting on outbound ports

Rate limiting on outbound ports is supported at a port level and on a port priority level. There are several configuration rules for rate limiting setup on the outbound ports:

- Outbound rate limiting is supported on:
  - GbE ports
  - 10-GbE ports
  - Trunk ports
- Outbound rate limiting is not supported on:
  - Ports on which LACP is enabled
  - Virtual interfaces
  - Loopback interfaces
- Because of the hardware architecture of the g-series, the effect of outbound rate limiting differs on GbE ports compared to 10-GbE ports. For example, applying the same rate limiting value on GbE and 10-GbE ports will produce different results.
- It is possible to configure both outbound port-based rate limiting and outbound port- and priority-based rate limiting on a single physical port or trunk port. However, if a priority-based limit for a given port is greater than the port based rate limit, then the port-based rate limit will override the priority-based rate limit. Similarly, if the port based rate limit is greater than the priority-based limit, then the priority-based rate limit will override the port based rate limit.

Port-based rate limiting

To enable port based outbound rate limiting use the command shown in Example 8-24. The command is executed on the interface level.

Example 8-24 Port based outbound rate limiting

```
LG_B50G(config)#interface e 1/1/20
LG_B50G(config-if-e1000-1/1/20)#rate-limit output fixed 50
Outbound Rate Limiting on Port 1/1/20 Config: 50 Kbps, Actual: 65 Kbps
LG_B50G(config-if-e1000-1/1/20)#
```

In this example, the configured value of 50 Kbps was rounded up to the minimum possible rate limit for 1G ports which is 65 Kbps.

Rate limiting can be disabled with the use of **no** in front of the respective enable command.

Port-based and priority-based rate limiting

To enable port priority based outbound rate limiting, use the command shown in Example 8-25. The command is executed on the interface level.

Example 8-25 Port priority based outbound rate limiting

```
LG_B50G(config)#interface e 1/1/20
LG_B50G(config-if-e1000-1/1/20)#rate-limit output fixed 2000 priority 5
Outbound Rate Limiting on Port 1/1/20 for Priority 5
         Config: 2000 Kbps, Actual: 1950 Kbps
LG_B50G(config-if-e1000-1/1/20)#
```

In this example, the configured value of 2000 Kbps was rounded to the nearest 65 Kbps increment that is used on 1G ports.

Rate limiting can be disabled with **no** in front of the respective enable command.

Configuring an ACL based rate limiting policy

ACL based rate limiting applies to the policy where rate limiting is applied to the matching IP traffic. ACL based rate limiting is configured with the use of traffic policies which are then referenced in the ACL. The same traffic policies can be referenced in many ACLs.

The traffic policies become active on the ports to which ACLs are bound.

Displaying the fixed rate limiting configuration

It is possible to display rate limiting configuration for inbound and outbound setups.

Inbound ports

To display the inbound rate setup, use the command shown in Example 8-26.

Example 8-26 Inbound ports rate info

```
LG_B50G#show rate-limit input
Total rate-limited interface count: 1.
   Port    Configured Input Rate        Actual Input Rate
 1/1/20                   500000                   500000
LG_B50G#
```

Outbound ports

To display the outbound rate setup, use the command shown in Example 8-27.

Example 8-27 Outbound ports rate info

```
LG_B50G#show rate-limit output
Outbound Rate Shaping Limits in Kbps:
  Port  PortMax   Prio0   Prio1   Prio2   Prio3   Prio4   Prio5   Prio6   Prio7
1/1/20       65       -       -       -       -       -    1950       -       -
LG_B50G#
```

8.2 NetIron m-series and c-series

In this section, we show examples of QoS implementation on the NetIron software platform which is used in m-series routers. We assume that you are familiar with the basics of QoS.

On m-series routers, QoS can be configured for ingress and egress traffic. In the following sections, we outline procedures for both options. In addition, certain procedures also affect ingress and egress traffic.

8.2.1 Configuring ingress QoS procedures

In the following sections we show how to configure ingress QoS.

Configuring ingress decode policy maps

If you do not want to use the default decode policy maps, then you must first define custom ingress policy maps. These maps are defined globally and they can be applied globally on all ports or on a locally specified port.

Policy maps are defined in the QoS configuration level, which can be reached using the command shown in Example 8-28.

Example 8-28 QoS configuration level

```
SF_B16M#config t
SF_B16M(config)#qos-mapping
SF_B16M(config-qos-mapping)#
```

It is possible to define three types of custom ingress policy maps, as described in the following topics.

- Configuring ingress decode DSCP policy maps:

  First, you define the name of the DSCP decode policy map with the command shown in Example 8-29.

Example 8-29 DSCP decode policy map name

```
SF_B16M(config-qos-mapping)#dscp decode-map Map1
SF_B16M(config-qos-mapping-dscp-decode)#
SF_B16M(config-qos-mapping-dscp-decode)#exit
SF_B16M(config-qos-mapping)#no dscp decode-map Map1
SF_B16M(config-qos-mapping)#
```

The name can be up to 64 character long and the same name can be used for different types of policy maps.

> **Important:** The name **`default-map`** cannot be used, because it is reserved for the default policy map name.

The defined map can be deleted by using the **`no`** command in front of the definition command as shown in Example 8-29.

Next, you configure the map for the particular DSCP bits of incoming packets. With this configuration it is defined as to which router internal priority (0-7) those DSCP bits are mapped. Optionally it is also possible to define the drop precedence (0-3).

This configuration is performed in the policy map level (Example 8-30).

Example 8-30 Configuring DSCP maps

```
SF_B16M(config-qos-mapping)#dscp decode-map Map1
SF_B16M(config-qos-mapping-dscp-decode)#dscp-value 30 to priority 4 drop-precedence 1
SF_B16M(config-qos-mapping-dscp-decode)#
```

It is possible to specify more DSCP values if you want to map them to the same priority and drop precedence. For the DSCP values, if you do not specify the mappings, the default mappings will be used.

With the **`no`** command, you can remove the defined maps and return them to the default values. There are two scenarios for this case:

– Removing the priority and drop-precedence at the same time is shown by using the **`no`** command with the **`to priority`** parameter (Example 8-31).

Example 8-31 Removing the full DSCP mapping

```
SF_B16M(config-qos-mapping-dscp-decode)#no dscp-value 30 to priority 4
SF_B16M(config-qos-mapping-dscp-decode)#
```

– Removing of only the drop-precedence is achieved with the **no** command in front of the whole original command (Example 8-32).

Example 8-32 Removing only the drop-precedence DSCP mapping

```
SF_B16M(config-qos-mapping-dscp-decode)#no dscp-value 30 to priority 4 drop-precedence 1
SF_B16M(config-qos-mapping-dscp-decode)#
```

In this case, the priority mapping will stay as previously defined.

- Configuring ingress decode PCP policy maps:

  First, you define the name of the PCP decode policy map with the command shown in Example 8-29.

Example 8-33 PCP decode policy map name

```
SF_B16M(config-qos-mapping)#pcp decode-map Map2
SF_B16M(config-qos-mapping)#
SF_B16M(config-qos-mapping-pcp-decode)#exit
SF_B16M(config-qos-mapping)#no pcp decode-map Map2
SF_B16M(config-qos-mapping)#
```

The name can be up to 64 character long and the same name can be used for different types of policy maps.

> **Important:** The name **default-map** cannot be used, because it is reserved for the default policy map name.

The defined map can be deleted by using the **no** command in front of the definition command as shown in Example 8-33.

Next, you configure the map for a particular PCP bits of incoming packets. With this configuration it is defined to which router the internal priority (0-7) that those PCP bits are mapped. Optionally it is also possible to define drop precedence (0-3).

This configuration is performed in the policy map level (Example 8-34).

Example 8-34 Configuring PCP maps

```
SF_B16M(config-qos-mapping)#pcp decode-map Map2
SF_B16M(config-qos-mapping-pcp-decode)#pcp-value 6 to priority 4 drop-precedence 1
SF_B16M(config-qos-mapping-pcp-decode)#
```

It is possible to specify more PCP values if you want to map them to the same priority and drop precedence.

For the PCP values, if you do not specify the mappings, the default mappings will be used.

With the **no** command you can remove the defined maps and return them to the default values. For this case there are two scenarios:

– Removing of priority and drop-precedence at the same time is shown by using the **no** command with the **to priority** parameter (Example 8-35).

Example 8-35 Removing the full PCP mapping

```
SF_B16M(config-qos-mapping-pcp-decode)#no pcp-value 6 to priority 4
SF_B16M(config-qos-mapping-pcp-decode)#
```

– Removing of only the drop-precedence is achieved with the **no** command in front of the entire original command as shown in Example 8-36.

Example 8-36 Removing only the drop-precedence PCP mapping

```
SF_B16M(config-qos-mapping-pcp-decode)#no pcp-value 6 to priority 4 drop-precedence 1
SF_B16M(config-qos-mapping-pcp-decode)#
```

In this case, priority mapping will stay as previously defined.

- Configuring ingress decode EXP policy maps:

  First, you define the name of the EXP decode policy map with the command shown in Example 8-37.

Example 8-37 EXP decode policy map name

```
SF_B16M(config-qos-mapping)#exp decode-map Map3
SF_B16M(config-qos-mapping-exp-decode)#
SF_B16M(config-qos-mapping-exp-decode)#exit
SF_B16M(config-qos-mapping)#no exp decode-map Map3
SF_B16M(config-qos-mapping)#
```

The name can be up to 64 character long and the same name can be used for different types of policy maps.

Important: The name **default-map** cannot be used because it is reserved for the default policy map name.

The defined map can be deleted by using the **no** command in front of the definition command as shown in Example 8-37.

Next we configure the map for a particular EXP bits of incoming packets, defining to which router the internal priority (0-7) of those EXP bits are mapped. Optionally it is also possible to define the drop precedence (0-3).

This configuration is performed in the policy map level (Example 8-38).

Example 8-38 Configuring EXP maps

```
SF_B16M(config-qos-mapping)#exp decode-map Map3
SF_B16M(config-qos-mapping-exp-decode)#exp-value 5 to priority 4 drop-precedence 1
SF_B16M(config-qos-mapping-exp-decode)#
```

It is possible to specify more EXP values if you want to map them to the same priority and drop precedence.

For the EXP values, if you do not specify the mappings, the default mappings will be used.

With the **`no`** command, you can remove the defined maps and return them to default values. For this case there are two scenarios for this case:

– Removing of priority and drop-precedence at the same time is shown by using the **`no`** command **`to priority`** parameter (Example 8-39).

Example 8-39 Removing the full EXP mapping

```
SF_B16M(config-qos-mapping-exp-decode)#no exp-value 5 to priority 4
SF_B16M(config-qos-mapping-exp-decode)#
```

– Removing of only the drop-precedence is achieved with the **`no`** command in front of the entire original command as shown in Example 8-40.

Example 8-40 Removing only the drop-precedence EXP mapping

```
SF_B16M(config-qos-mapping-exp-decode)#no exp-value 5 to priority 4 drop-precedence 1
SF_B16M(config-qos-mapping-exp-decode)#
```

In this case, the priority mapping will stay as previously defined.

Binding ingress decode policy maps

After custom maps are defined they need to be bound to be used. The ingress decode policy map can be mapped globally or per port based. It is possible to map a default policy map, an all zero policy map, or a custom policy map. For PCP, it is additionally possible to map the following predefined maps: 7P1D, 6P2D, 5P3D.

`Default-map` will assign a default map and an all-zero-map which maps all DSCP bits to zero priority and zero drop precedence.

The following sections show how to bind the various types of maps:

1. Binding ingress decode DSCP policy maps:

   To bind the ingress decode DSCP policy map globally, use the command shown in Example 8-41.

Example 8-41 Binding DSCP policy map globally

```
SF_B16M(config)#qos dscp decode-policy Map1
SF_B16M(config)#
SF_B16M(config)#no qos dscp decode-policy Map1
SF_B16M(config)#
```

   Instead of the *Map1* it is possible to use *default-map* or *all-zero-map*.

   You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-41.

   To bind the DSCP policy map at the port level, use the command shown in Example 8-42.

Example 8-42 Binding DSCP policy map on the port

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos dscp decode-policy Map1
SF_B16M(config-if-e1000-9/6)#
SF_B16M(config-if-e1000-9/6)#no qos dscp decode-policy Map1
SF_B16M(config-if-e1000-9/6)#
```

   Instead of *Map1* it is possible to use *default-map* or *all-zero-map*.

   You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-42.

2. Binding ingress decode PCP policy maps:

   To bind the ingress decode PCP policy map globally, use the command shown in Example 8-43.

Example 8-43 Binding PCP policy map globally

```
SF_B16M(config)#qos pcp decode-policy Map2
SF_B16M(config)#
SF_B16M(config)#no qos pcp decode-policy Map2
SF_B16M(config)#
```

   Instead of *Map2* it is possible to use *default-map* or *all-zero-map or 7P1D, 6P2D, 5P3D*.

   You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-43.

To bind the PCP policy map on the port level, use the command shown in Example 8-44.

Example 8-44 Binding PCP policy map on the port

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos pcp decode-policy Map2
SF_B16M(config-if-e1000-9/6)#
SF_B16M(config-if-e1000-9/6)#no qos pcp decode-policy Map2
SF_B16M(config-if-e1000-9/6)#
```

Instead of *Map2* it is possible to use *default-map* or *all-zero-map or 7P1D, 6P2D, 5P3D*.

You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-44.

3. Binding ingress decode EXP policy maps:

   To bind the ingress decode EXP policy map globally, use the command shown in Example 8-45.

Example 8-45 Binding EXP policy map globally

```
SF_B16M(config)#qos exp decode-policy Map3
SF_B16M(config)#
SF_B16M(config)#no qos exp decode-policy Map3
SF_B16M(config)#
```

Instead of *Map3* it is possible to use *default-map* or *all-zero-map.*

You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-45.

To bind the EXP policy map on the port level, use the command shown in Example 8-46

Example 8-46 Binding EXP policy map on the port

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos exp decode-policy Map3
SF_B16M(config-if-e1000-9/6)#
SF_B16M(config-if-e1000-9/6)#no qos exp decode-policy Map3
SF_B16M(config-if-e1000-9/6)#
```

Instead of *Map3* it is possible to use *default-map* or *all-zero-map*.

You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-46.

Configuring a force priority

Forcing the priority can be used when multiple possible priority values are available for a particular ingress port. The following options are available:

- Force priority for a port:

  After the priority is defined on a particular port, it is possible to force that priority has precedence above other priority values for incoming packets. The commands to achieve this are shown in Example 8-47. The commands are executed on the interface config level.

Example 8-47 Force priority for a port

```
SF_B16M(config)#interface e 9/10
SF_B16M(config-if-e1000-9/10)#priority 5
SF_B16M(config-if-e1000-9/10)#priority force
SF_B16M(config-if-e1000-9/10)#
```

  By using a **`no priority force`** command, the force priority can be disabled.

- Force drop precedence for a port:

  After a drop precedence is defined on a particular port, it is possible to force that drop precedence has precedence above other priority values for incoming packets. The commands to achieve this are shown in Example 8-48. The commands are executed on the interface config level.

Example 8-48 Force drop precedence for a port

```
SF_B16M(config)#interface e 9/11
SF_B16M(config-if-e1000-9/11)#drop-precedence 2
SF_B16M(config-if-e1000-9/11)#drop-precedence force
SF_B16M(config-if-e1000-9/11)#
```

  With the use of the **`no drop-precedence force`** command the force drop precedence can be disabled.

- Configuring a force priority for a VLAN:

  By default, VLANs have the priority 3. It is possible to define different priority (3-7) and force that priority to have precedence above other priority values for incoming packets. The commands to achieve this are shown in Example 8-49. The commands are executed on the VLAN config level.

Example 8-49 Force priority for VLAN

```
SF_B16M(config)# vlan 21
SF_B16M(config-vlan-21)#priority 5
SF_B16M(config-vlan-21)#priority force
SF_B16M(config-vlan-21)#
```

With the use of the **no priority force** command the force priority can be disabled.

- Force priority to the DSCP value:

  To force the DSCP value on the incoming packets on the ingress port over other priority values, use the command shown in Example 8-50. The command is executed on the interface config level.

Example 8-50 Force priority for DSCP

```
SF_B16M(config)#interface e 9/12
SF_B16M(config-if-e1000-9/12)#qos dscp force
SF_B16M(config-if-e1000-9/12)#
```

With the use of the **no qos dscp force** command the dscp force priority can be disabled.

- Force priority to the EXP value:

  To force the EXP value on the incoming packet on the ingress port over other priority values, use the command shown in Example 8-51. The command is executed on the interface config level.

Example 8-51 Force priority for EXP

```
SF_B16M(config)#interface e 9/12
SF_B16M(config-if-e1000-9/12)#qos exp force
SF_B16M(config-if-e1000-9/12)#
```

With the use of the **no qos exp force** command the EXP force priority can be disabled.

- Force priority to the PCP value:

  To force the PCP value on the incoming packets on the ingress port over other priority values, use the command shown in Example 8-52. The command is executed on the interface config level.

Example 8-52 Force priority for PCP

```
SF_B16M(config)#interface e 9/12
SF_B16M(config-if-e1000-9/12)#qos pcp force
SF_B16M(config-if-e1000-9/12)#
```

With the use of the **no qos pcp force** command the pcp force priority can be disabled.

- Force priority to a value specified by an ACL:

  It is possible to use the **priority-force** parameter within an ACL to apply the priority to specific traffic defined in the ACL.

8.2.2 Configuring egress QoS procedures

Configuration of egress QoS is similar to the ingress QoS configuration. The only difference is that instead of decode maps we now have encode maps, as we are marking outgoing packets on egress ports, meaning that we are encoding QoS information in the packets.

Because of this similarity. we just show the commands. For details about the mappings and other configuration settings, see “Configuring ingress QoS procedures” on page 251.

Configuring egress encode policy maps

Policy maps are defined in QoS configuration level, which can be reached with the command shown in Example 8-53.

Example 8-53 QoS configuration level

```
SF_B16M#config t
SF_B16M(config)#qos-mapping
SF_B16M(config-qos-mapping)#
```

It is possible to define three types of custom ingress policy maps, as described in the following topics.

- Configuring egress encode DSCP policy maps:

  First, you define the name of the DSCP encode policy map with the command shown in Example 8-54.

Example 8-54 DSCP encode policy map name

```
SF_B16M(config)#qos-mapping
SF_B16M(config-qos-mapping)#dscp encode-map Map10
SF_B16M(config-qos-mapping-dscp-encode)#
SF_B16M(config-qos-mapping-dscp-encode)#exit
SF_B16M(config-qos-mapping)#no dscp encode-map Map10
SF_B16M(config-qos-mapping)#
```

  The defined map can be deleted by using the **`no`** command in front of the definition command as shown in Example 8-54.

  Next, you configure the map for particular DSCP bits for outgoing packets. With this configuration it is defined as to which DSCP bits internal priority (0-7) and optionally drop precedence (0-3) are mapped.

  This configuration is performed in the policy map level (Example 8-55).

Example 8-55 Configuring DSCP encode maps

```
SF_B16M(config-qos-mapping)#dscp encode-map Map10
SF_B16M(config-qos-mapping-dscp-encode)#priority 4 drop-precedence 1 to dscp-value 30
SF_B16M(config-qos-mapping-dscp-encode)#
```

For the DSCP values, you do not specify the mappings as the default mappings will be used.

If the *drop-precedence* value is not specified any matched value (0-3) is allowed.

With the **no** command you can remove the defined maps and return them to the default values.

- Configuring egress encode PCP policy maps:

First, you define the name of the PCP encode policy map with the command shown in Example 8-56.

Example 8-56 PCP encode policy map name

```
SF_B16M(config)#qos-mapping
SF_B16M(config-qos-mapping)#pcp encode-map Map11
SF_B16M(config-qos-mapping-pcp-encode)#
SF_B16M(config-qos-mapping-pcp-encode)#exit
SF_B16M(config-qos-mapping)#no pcp encode-map Map11
SF_B16M(config-qos-mapping)#
```

The defined map can be deleted by using the **no** command in front of the definition command as shown in Example 8-56.

Next, you configure the map for a particular PCP value for outgoing packets. With this configuration we define which PCP value internal priority (0-7) and optionally drop precedence (0-3) are mapped.

This configuration is performed in the policy map level (Example 8-57).

Example 8-57 Configuring PCP encode maps

```
SF_B16M(config-qos-mapping)#pcp encode-map Map11
SF_B16M(config-qos-mapping-pcp-encode)#priority 4 drop-precedence 1 to pcp-value 6
SF_B16M(config-qos-mapping-pcp-encode)#
```

For the PCP values, you do not specify the mappings as the default mappings will be used.

If the drop-precedence value is not specified, any matched value (0-3) is allowed.

With the **no** command, you can remove the defined maps and return them to the default values.

- Configuring egress encode EXP policy maps:

  First, you define the name of the EXP encode policy map with the command shown in Example 8-58.

Example 8-58 EXP encode policy map name

```
SF_B16M(config)#qos-mapping
SF_B16M(config-qos-mapping)#exp encode-map Map12
SF_B16M(config-qos-mapping-exp-encode)#
SF_B16M(config-qos-mapping-exp-encode)#exit
SF_B16M(config-qos-mapping)#no exp encode-map Map12
SF_B16M(config-qos-mapping)#
```

  The defined map can be deleted by using the **no** command in front of the definition command as shown in Example 8-58.

  Next, you configure the map for a particular EXP value for outgoing packets. With this configuration we define which EXP value internal priority (0-7) and optionally drop precedence (0-3) are mapped.

  This configuration is performed in the policy map level (Example 8-59).

Example 8-59 Configuring EXP encode maps

```
SF_B16M(config-qos-mapping)#exp encode-map Map12
SF_B16M(config-qos-mapping-exp-encode)#priority 4 drop-precedence 1 to
exp-value 5
SF_B16M(config-qos-mapping-exp-encode)#
```

  For the EXP values, you do not specify the mappings the default mappings will be used.

  If the *drop-precedence* value is not specified, then any matched value (0-3) is allowed.

  With the **no** command, you can remove the defined maps and return then to the default values.

Binding an egress encode policy map

The rules for binding egress encode maps are the same as for ingress decode maps as described in “Binding ingress decode policy maps” on page 255.

The following sections show how to bind various types of maps.

1. Binding an egress encode DSCP policy map:

   To bind the egress encode DSCP policy map globally, use the command shown in Example 8-41.

Example 8-60 Binding DSCP encode policy map globally

```
SF_B16M(config)#qos dscp encode-policy Map10
SF_B16M(config)#
SF_B16M(config)#no qos dscp encode-policy Map10
SF_B16M(config)#
```

Instead of *Map10* it is possible to use *default-map* or *all-zero-map*.

You can use the **`no`** command in front of the bind command to remove the binding as shown in Example 8-60.

To bind the DSCP policy map at the port level, use the command shown in Example 8-61.

Example 8-61 Binding DSCP encode policy map on the port

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos dscp encode-policy Map10
SF_B16M(config-if-e1000-9/6)#
SF_B16M(config-if-e1000-9/6)#no qos dscp encode-policy Map10
SF_B16M(config-if-e1000-9/6)#
```

Instead of the *Map10* it is possible to use default-map or *all-zero-map.*

You can use the **`no`** command in front of the bind command to remove the binding as show in Example 8-61.

2. Binding an egress encode PCP policy map:

   To bind the egress encode PCP policy map globally use the command shown in Example 8-62.

Example 8-62 Binding PCP encode policy map globally

```
SF_B16M(config)#qos pcp encode-policy Map11
SF_B16M(config)#
SF_B16M(config)#no qos pcp encode-policy Map11
SF_B16M(config)#
```

Instead of *Map11* it is possible to use *default-map* or *all-zero-map*.

You can use the **`no`** command in front of the bind command to remove the binding as shown in Example 8-62.

To bind the PCP policy map at the port level, use the command shown in Example 8-63.

Example 8-63 Binding PCP encode policy map on the port

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos pcp encode-policy Map11
```

```
SF_B16M(config-if-e1000-9/6)#
SF_B16M(config-if-e1000-9/6)#no qos pcp encode-policy Map11
SF_B16M(config-if-e1000-9/6)#
```

Instead of *Map11* it is possible to use *default-map* or *all-zero-map*.

You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-63.

3. Binding an egress encode EXP policy map:

   To bind the egress encode EXP policy map globally, use the command shown in Example 8-64.

Example 8-64 Binding EXP encode policy map globally

```
SF_B16M(config)#qos exp encode-policy Map12
SF_B16M(config)#
SF_B16M(config)#no qos exp encode-policy Map12
SF_B16M(config)#
```

Instead of *Map12* it is possible to use *default-map* or *all-zero-map*.

You can use the **no** command in front of the bind command to remove the binding as shown in Example 8-64.

To bind the EXP policy map on the port level, use the command shown in Example 8-65.

Example 8-65 Binding EXP encode policy map on the port

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos exp encode-policy Map12
SF_B16M(config-if-e1000-9/6)#
SF_B16M(config-if-e1000-9/6)#no qos exp encode-policy Map12
SF_B16M(config-if-e1000-9/6)#
```

Instead of *Map12* it is possible to use *default-map* or *all-zero-map.*

You can use the no command in front of the bind command to remove the binding as shown in Example 8-65.

8.2.3 Configuring QoS procedures applicable to ingress and egress

There are a couple of procedures that are applicable for ingress and egress QoS traffic processing. We discuss those procedures in the following sections.

Enabling a port to use the DEI bit

The Drop Eligible Indicator (DEI) which is part of the Service VLAN tag (S-TAG - as defined in 802.1ad) can be used to define drop precedence as the DEI bit indicates if packet is drop eligible.

With the **qos use-dei** command on the interface config level, the router will use the DEI bit to identify a drop precedence of the incoming packet on the ingress port.

For the outgoing packet on the egress port the drop precedence will be encoded in the DEI bit as follows:

- Drop precedence 2 and 3 => DEI = 1
- Drop precedence 0 and 1 => DEI = 0

The command for enabling the use of the DEI bit is shown in Example 8-66. This command can only be used on the interface level.

Example 8-66 Use of DEI bit

```
SF_B16M(config)#interface e 9/8
SF_B16M(config-if-e1000-9/8)#qos use-dei
SF_B16M(config-if-e1000-9/8)#
SF_B16M(config-if-e1000-9/8)#no qos use-dei
SF_B16M(config-if-e1000-9/8)#
```

With the use of the **no qos use-dei** command, the DEI bit is disabled as shown in Example 8-66.

Specifying the trust level

The trust level is used to specify from whereabouts in the packet received on the interface device will extract the QoS value.

There are three possible values for the trust level:

- **cos**: The device uses the IEEE 802.1p (CoS) priority value in the packet's Ethernet frame header. Use this trust option when you plan to mark the packet's DSCP value based on the incoming IEEE 802.1p value.
- **ip-prec**: The device uses the three most-significant bits in the packet's ToS field and interprets them as an IP precedence value. Use this trust option when the incoming packet is from a device that does not support DSCP and you need to mark the packet for QoS on DSCP devices.
- **dscp**: The device uses the six most-significant bits in the packet's ToS field and interprets them as a DSCP value.

Typical commands setting different trust options are shown in Example 8-67.

Example 8-67 Setting the trust option

```
SF_B16M(config)#interface e 9/8
SF_B16M(config-if-e1000-9/8)#qos-tos trust cos
SF_B16M(config-if-e1000-9/8)#
SF_B16M(config-if-e1000-9/8)#qos-tos trust ip-prec
SF_B16M(config-if-e1000-9/8)#
SF_B16M(config-if-e1000-9/8)#qos-tos trust dscp
SF_B16M(config-if-e1000-9/8)#
```

Enable marking

Marking is used to modify an outbound packet's IEEE 802.1p priority field, DSCP field, or both, to match the results of the QoS mappings performed by the device. When marking is enabled on the interface it applies to the arriving packets through that interface.

There are two possible marking values:

- **`cos`**: The device changes the outbound packet's IEEE 802.1p priority value to match the results of the device's QoS mapping from the specified trust level.
- **`dscp`**: The device changes the outbound packet's IEEE 802.1p priority value to match the results of the device's QoS mapping from the specified trust level.

In Example 8-68, we enable marking on the interface *E9/6* and the enabling **`pcp`** encode policy on egress interface *E9/15*.

Example 8-68 Marking example

```
SF_B16M(config)#interface e 9/6
SF_B16M(config-if-e1000-9/6)#qos-tos mark cos
SF_B16M(config-if-e1000-9/6)#interface e 9/15
SF_B16M(config-if-e1000-9/15)#qos pcp encode-policy on
SF_B16M(config-if-e1000-9/15)#
```

Attention:

- The **`qos pcp encode-policy on`** command must be configured when the **`qos-tos mark cos`** command is configured.
- The **`qos pcp encode-policy`** command is on by default and does not require explicit configuration unless it has been configured to be off.
- It is not possible to apply an ACL to an interface in the outbound direction to change the priority of certain types of traffic.

Packet mapping commands

Packet mapping commands are used to specify what CoS, IP-Precedence, or DSCP value received on an ingress port (based on the trust level) will be used to determine the QoS value marked on an outgoing packet on an egress port.

These mappings are used when packet marking is enabled.

The following mappings can be changed:

- CoS → DSCP mappings:

  The command shown in Example 8-69 shows a typical mapping. This command is executed on global config level.

Example 8-69 CoS to DSCP mappings

```
SF_B16M(config)#qos-tos map cos-dscp 0 31 23 47 15 5 53 39
SF_B16M(config)#ip rebind-acl all
SF_B16M(config)#
```

  You need to specify eight DSCP values to match all CoS values (0-7).

- IP Precedence → DSCP mappings:

  The command shown in Example 8-70 shows a typical mapping. This command is executed on global config level.

Example 8-70 IP Precedence to DSCP mappings

```
SF_B16M(config)#qos-tos map ip-prec-dscp 0 31 23 47 15 5 53 39
SF_B16M(config)#ip rebind-acl all
SF_B16M(config)#
```

  You need to specify eight DSCP values to match all IP precedence values (0-7).

- DSCP → DSCP mappings:

  The command shown in Example 8-71 shows a typical mapping. This command is executed on global config level.

Example 8-71 DSCP to DSCP mappings

```
SF_B16M(config)#qos-tos map dscp-dscp 0 to 10
SF_B16M(config)#ip rebind-acl all
SF_B16M(config)#
```

  You need to specify up to eight DSCP values that you want to change.

> **Attention:** To place a qos-tos mapping change into effect, you must enter the **ip rebind-acl all** command at the global config level after making the mapping change. This applies to all mappings that are configured using the **qos-tos map** command.

All mapping commands can be removed by using the **no** command. This will return default values of all mappings.

Support for super aggregate VLANs

When a super-aggregate VLAN is used, its untagged interface can be configured to copy the QoS bits from the tag value set by the edge device to the tag value set by the core device. This is only supported if the incoming packet has ETYPE 0x8100. It can be configured using the command shown in Example 8-72.

Example 8-72 SA VLAN QoS

```
SF_B16M(config)#interface e 9/5
SF_B16M(config-if-e1000-9/5)#qos decode-cvlan-pcp
```

Support for QoS configurations on LAG ports

There are two types of rules that apply to QoS configuration on LAG ports:

- LAG configuration rules where QoS values are applied directly to the port:
  a. Each port that is configured into the LAG, must have the same priority, priority force, drop-precedence, and drop-precedence force configuration.
  b. If you have already formed a LAG with the same configuration, you can change the configuration by making changes to the LAG's primary port.
  c. If the LAG configuration is deleted, each of the ports in the LAG (primary and secondary) will inherit the QoS configuration of the primary port.
- LAG configuration rules for QoS configurations using commands that begin with the **qos** keyword:
  a. The secondary ports configured in the LAG must not have any QoS values configured on them.
  b. The **qos** commands that are configured on the primary port are applied to all ports in the LAG.
  c. After the LAG is formed, you can change the QoS configuration for all ports in the LAG by making changes to the LAG's primary port, but you cannot change the QoS configurations directly on any secondary ports.
  d. If the LAG is deleted, the QoS configuration will only be retained on the primary port.

8.2.4 Displaying QoS information

It is possible to display information about the QoS configuration.

QoS decode policy map configurations

To display the decode policy map, use the commands shown in Example 8-73. It is possible to display the configuration for DSCP, PCP and EXP decode maps.

Example 8-73 Displaying decode map configurations

```
SF_B16M(config)#show qos-map dscp decode-map Map1
DSCP decode map Map1
  DSCP  0 to priority  0 drop-precedence  0
  DSCP  1 to priority  0 drop-precedence  0
  DSCP  2 to priority  0 drop-precedence  1
  DSCP  3 to priority  0 drop-precedence  1
...
DSCP 60 to priority  7 drop-precedence  2
  DSCP 61 to priority  7 drop-precedence  2
  DSCP 62 to priority  7 drop-precedence  3
  DSCP 63 to priority  7 drop-precedence  3
SF_B16M(config)#
SF_B16M(config)#show qos-map pcp decode-map Map2
PCP decode map Map2
  PCP  0 to priority  0 drop-precedence  0
  PCP  1 to priority  1 drop-precedence  0
  PCP  2 to priority  2 drop-precedence  0
  PCP  3 to priority  3 drop-precedence  0
  PCP  4 to priority  4 drop-precedence  0
  PCP  5 to priority  5 drop-precedence  0
  PCP  6 to priority  4 drop-precedence  1
  PCP  7 to priority  7 drop-precedence  0
SF_B16M(config)#
SF_B16M(config)#show qos-map exp decode-map Map3
EXP decode map Map3
  EXP  0 to priority  0 drop-precedence  0
  EXP  1 to priority  1 drop-precedence  0
  EXP  2 to priority  2 drop-precedence  0
  EXP  3 to priority  3 drop-precedence  0
  EXP  4 to priority  4 drop-precedence  0
  EXP  5 to priority  4 drop-precedence  1
  EXP  6 to priority  6 drop-precedence  0
  EXP  7 to priority  7 drop-precedence  0
SF_B16M(config)#
```

QoS encode policy map configurations

To display the encode policy map, use the commands shown in Example 8-74. It is possible to display configuration for DSCP, PCP and EXP encode maps.

Example 8-74 Displaying encode map configurations

```
SF_B16M(config)#show qos-map dscp encode-map Map10
DSCP encode map Map10
  Priority  0 drop-precedence  0 to DSCP  0
  Priority  0 drop-precedence  1 to DSCP  2
  Priority  0 drop-precedence  2 to DSCP  4
  Priority  0 drop-precedence  3 to DSCP  6
  Priority  1 drop-precedence  0 to DSCP  8
...
Priority  6 drop-precedence  3 to DSCP 54
  Priority  7 drop-precedence  0 to DSCP 56
  Priority  7 drop-precedence  1 to DSCP 58
  Priority  7 drop-precedence  2 to DSCP 60
  Priority  7 drop-precedence  3 to DSCP 62
SF_B16M(config)#
SF_B16M(config)#show qos-map pcp encode-map Map11
PCP encode map Map11
  Priority  0 drop-precedence  0 to PCP  0
  Priority  0 drop-precedence  1 to PCP  0
  Priority  0 drop-precedence  2 to PCP  0
  Priority  0 drop-precedence  3 to PCP  0
  Priority  1 drop-precedence  0 to PCP  1
...
Priority  6 drop-precedence  2 to PCP  6
  Priority  6 drop-precedence  3 to PCP  6
  Priority  7 drop-precedence  0 to PCP  7
  Priority  7 drop-precedence  1 to PCP  7
  Priority  7 drop-precedence  2 to PCP  7
  Priority  7 drop-precedence  3 to PCP  7
SF_B16M(config)#
SF_B16M(config)#show qos-map exp encode-map Map12
EXP encode map Map12
  Priority  0 drop-precedence  0 to EXP  0
  Priority  0 drop-precedence  1 to EXP  0
  Priority  0 drop-precedence  2 to EXP  0
  Priority  0 drop-precedence  3 to EXP  0
  Priority  1 drop-precedence  0 to EXP  1
...
Priority  6 drop-precedence  3 to EXP  6
  Priority  7 drop-precedence  0 to EXP  7
  Priority  7 drop-precedence  1 to EXP  7
  Priority  7 drop-precedence  2 to EXP  7
  Priority  7 drop-precedence  3 to EXP  7
SF_B16M(config)#
```

It is also possible to display decode and encode maps for *all-zero-map* and *default-map*.

QoS policy map binding configurations

To display the decode policy map binding, use the commands shown in Example 8-75. It is possible to display bindings globally or on the port level.

Example 8-75 Displaying map binding configurations

```
SF_B16M(config)#show qos-map binding global
 qos pcp decode-policy default-map
 qos pcp encode-policy default-map
 qos exp decode-policy default-map
 qos exp encode-policy default-map
 qos dscp decode-policy default-map
 qos dscp encode-policy default-map
SF_B16M(config)#
SF_B16M(config)#show qos-map binding 9/10
 priority force
 priority 5
SF_B16M(config)#show qos-map binding 9/11
 drop-precedence force
 drop-precedence 2
SF_B16M(config)#show qos-map binding 9/12
 qos pcp force
 qos exp force
 qos dscp force
SF_B16M(config)#
```

Displaying QoS packet and byte counters

It is possible to enable and display the QoS packet and byte counters as shown in the following sections.

Enabling QoS packet and byte counters

Packet and byte counters can be enabled with the command shown in Example 8-76. This command is run in the global config level.

Example 8-76 Enabling QoS counters

```
SF_B16M(config)#enable-qos-statistics
SF_B16M(config)#
```

Counters can be disabled with the use of the **no** option in front of the command.

Displaying QoS packet and byte counters

An example of the packet and byte counters is shown in Example 8-77.

Example 8-77 Packet and byte counters

```
SF_B16M(config)#show np qos statistics eth 9/10
Port 9/10
  Ingress counters:
    COS 0: packets 0                    bytes 0
    COS 1: packets 0                    bytes 0
    COS 2: packets 0                    bytes 0
    COS 3: packets 0                    bytes 0
    COS 4: packets 0                    bytes 0
    COS 5: packets 0                    bytes 0
    COS 6: packets 0                    bytes 0
    COS 7: packets 0                    bytes 0
  Egress counters:
    COS 0: packets 0                    bytes 0
    COS 1: packets 0                    bytes 0
    COS 2: packets 0                    bytes 0
    COS 3: packets 0                    bytes 0
    COS 4: packets 0                    bytes 0
    COS 5: packets 0                    bytes 0
    COS 6: packets 0                    bytes 0
    COS 7: packets 0                    bytes 0
SF_B16M(config)#
```

As you can see, statistics are displayed for each priority level for ingress and egress counters. Counters can also be displayed for other type of interfaces, for example, packet-over-sonet.

Clearing QoS packet and byte counters

If required, it is possible to clear the counters with the command shown in Example 8-78.

Example 8-78 Clearing the packet and byte counters

```
SF_B16M(config)#show np qos statistics eth 9/6
Port 9/6
```

```
  Ingress counters:
    COS 0: packets 1                     bytes 277
    COS 1: packets 0                     bytes 0
    COS 2: packets 0                     bytes 0
    COS 3: packets 0                     bytes 0
    COS 4: packets 0                     bytes 0
    COS 5: packets 0                     bytes 0
    COS 6: packets 0                     bytes 0
    COS 7: packets 0                     bytes 0
  Egress counters:
    COS 0: packets 0                     bytes 0
    COS 1: packets 0                     bytes 0
    COS 2: packets 0                     bytes 0
    COS 3: packets 0                     bytes 0
    COS 4: packets 0                     bytes 0
    COS 5: packets 0                     bytes 0
    COS 6: packets 0                     bytes 0
    COS 7: packets 8                     bytes 6224
SF_B16M(config)#clear np qos statistics ethernet 9/6
SF_B16M(config)#show np qos statistics eth 9/6
Port 9/6
  Ingress counters:
    COS 0: packets 0                     bytes 0
    COS 1: packets 0                     bytes 0
    COS 2: packets 0                     bytes 0
    COS 3: packets 0                     bytes 0
    COS 4: packets 0                     bytes 0
    COS 5: packets 0                     bytes 0
    COS 6: packets 0                     bytes 0
    COS 7: packets 0                     bytes 0
  Egress counters:
    COS 0: packets 0                     bytes 0
    COS 1: packets 0                     bytes 0
    COS 2: packets 0                     bytes 0
    COS 3: packets 0                     bytes 0
    COS 4: packets 0                     bytes 0
    COS 5: packets 0                     bytes 0
    COS 6: packets 0                     bytes 0
    COS 7: packets 0                     bytes 0
SF_B16M(config)#
```

Counters can also be cleared for other type of interfaces, that is, packet-over-sonet.

8.2.5 Configuring packet drop priority using WRED

NetIron m-series routes support packet drop priority using WRED. There are several steps required to enable WRED, and some of them optional.

Enabling WRED

WRED can be enabled with the command shown in Example 8-79. The command is executed on the global config level.

Example 8-79 Enabling WRED

```
SF_B16M(config)#qos queue-type 4 wred enable
SF_B16M(config)#
```

In this example, we enabled WRED for queue type 4. WRED needs to be enabled for any forwarding queue where you want its operation. There are eight forwarding queues available numbered 0-7. By default, when you enable WRED on a particular queue the default values will be used.

The WRED on a particular queue can be disabled with the **no** command in front of the command for enabling it.

It is possible to define the following custom parameters.

Setting the averaging-weight (Wq) parameter

To set the Wq parameter, use the command in Example 8-80. The command is executed on the global config level.

Example 8-80 Wq parameter

```
SF_B16M(config)#qos queue-type 2 wred averaging-weight 2
Error - Please enable WRED for queue 2 before attempting to set
parameters
SF_B16M(config)#qos queue-type 2 wred enable
SF_B16M(config)#qos queue-type 2 wred averaging-weight 2
SF_B16M(config)#
```

The Wq parameter can be set for a particular queue type, in our example the value of 2 was set for the queue type 2.

> **Note:** The queue has to be enabled for WRED before the Wq parameter can be set. In Example 8-80 you can see the error message issued if the queue is not WRED enabled.

There are 13 possible values for the Wq parameter as shown in Table 8-1.

Table 8-1 Wq parameter values

Averaging weight setting	Wq value as a percentage
1	50%
2	25%
3	12.5%
4	6.2%
5	3.12%
6	1.56%
7	0.78%
8	0.4%
9	0.2%
10	0.09%
11	0.05%
12	0.02%
13	0.01%

The command can be disabled with the **`no`** command in front of the respective command.

Configuring the maximum instantaneous queue size

It is possible to specify the maximum size to which the queue is allowed to grow. An example of the command to achieve this is shown in Example 8-81. The command is executed on the global config level. The maximum size is specified in Kilobytes.

Example 8-81 Maximum queue size

```
SF_B16M(config)#qos queue-type 2 max-queue-size 32
SF_B16M(config)#
```

In this example, we set 32Kbytes as the maximum queue size for the queue type 2.

The command can be disabled with the **no** command in front of the respective command.

Configuring the drop precedence parameters

The DSCP or TOS bits in packets are used to prioritize packet delivery for specified queue types. The values can be from 0-3. The packet with value 0 will be less likely to be dropped than the packet with the value of 3.

> **Note:** In addition to bits in the DSCP, the DP option can use other fields (in the PCP header or the EXP bit header) to control WRED in the priority queues.
>
> Packets that do not have the DSCP or TOS value set are assigned a drop precedence equal to the DSCP or TOS level of 0.

It is possible to set the following to affect the behavior of drop precedence parameters:

- Setting the maximum drop probability
- Setting the minimum and maximum average queue size
- Setting the maximum packet size

Restoring default WRED parameters

To restore the default WRED parameters use the command shown in Example 8-82. The command is executed on the global config level.

Example 8-82 Restore default WRED parameters

```
SF_B16M(config)#show qos wred
QType Enable AverWeight MaxQSz DropPrec MinAvgQSz MaxAvgQSz MaxDropProb MaxPktSz
  0     No
  1     No
  2     Yes   2(25. 0%)     32      0           448      1024          2%    16384
                                   1           384      1024          4%    16384
                                   2           320      1024          9%    16384
                                   3           256      1024          9%    16384
  3     No
  4     Yes    4(6.25%)   1024      0           448      1024          2%    16384
                                   1           384      1024          4%    16384
                                   2           320      1024          9%    16384
                                   3           256      1024          9%    16384
  5     No
  6     No
  7     No
SF_B16M(config)#qos queue-type 2 wred default-params
SF_B16M(config)#show qos wred
```

```
QType Enable AverWeight MaxQSz DropPrec MinAvgQSz MaxAvgQSz MaxDropProb MaxPktSz
  0     No
  1     No
  2     Yes    4(6.25%)     32     0          448      1024          2%    16384
                                   1          384      1024          4%    16384
                                   2          320      1024          9%    16384
                                   3          256      1024          9%    16384
  3     No
  4     Yes    4(6.25%)   1024     0          448      1024          2%    16384
                                   1          384      1024          4%    16384
                                   2          320      1024          9%    16384
                                   3          256      1024          9%    16384
  5     No
  6     No
  7     No
SF_B16M(config)#
```

Displaying the WRED configuration

To display the WRED parameters, use the command shown in Example 8-83. The command is executed on the global config level.

Example 8-83 Displaying WRED parameters

```
SF_B16M(config)#show qos wred
QType Enable AverWeight MaxQSz DropPrec MinAvgQSz MaxAvgQSz MaxDropProb MaxPktSz
  0     No
  1     No
  2     Yes   2(25. 0%)     32     0          448      1024          2%    16384
                                   1          384      1024          4%    16384
                                   2          320      1024          9%    16384
                                   3          256      1024          9%    16384
  3     No
  4     Yes    4(6.25%)   1024     0          448      1024          2%    16384
                                   1          384      1024          4%    16384
                                   2          320      1024          9%    16384
                                   3          256      1024          9%    16384
  5     No
  6     No
  7     No
SF_B16M(config)#
```

8.2.6 Egress port and priority based rate shaping

NetIron m-series routers support egress rate shaping. It can be applied on the port level or for each priority queue on the specified port.

Rate shaping is used to smoothen the traffic, especially when the traffic is bursty. When rate shaping is used the traffic that exceeds the threshold is buffered not dropped as it might be if rate limiting was used.

> **Note:** Because excess traffic is buffered, rate shaping must be used with caution. In general, it is not advisable to rate shape delay-sensitive traffic.

Configuring port-based rate shaping

When configuring port based rate shaping, use the following minimum values and increments as shown in Table 8-2.

Table 8-2 Port-based rate shaping interval table

Range	Increment supported within the range (bytes)
0 - 10M	8,333
10M - < 100M	20,833
100 M - < 1G	208,333
1G - 10G	2,083,333

> **Note:** The egress rate shaping burst size for a port-based shaper is 10000 bytes.

The example of port based rate shaping configuration is shown in Example 8-84. The command is executed on the interface config level.

Example 8-84 Port based rate shaping

```
SF_B16M(config)#interface e 5/17
SF_B16M(config-if-e1000-5/17)#qos shaper 200000
SF_B16M(config-if-e1000-5/17)#
```

The command can be disabled by using the **`no`** command in front of the respective command.

Configuring port and priority-based rate shaping

The example of port and priority based rate shaping configuration is shown in Example 8-85. The command is executed on the interface config level.

Example 8-85 Port and priority based rate shaping

```
SF_B16M(config)#interface e 5/17
SF_B16M(config-if-e1000-5/17)#qos shaper priority 3 400000
SF_B16M(config-if-e1000-5/17)#
```

In this example, we set rate shaping for priority Level 3.

The command can be disabled by using the **`no`** command in front of the respective command.

8.2.7 Traffic manager statistics display

Counters are available to track the packets and bytes that enter the ingress traffic manager and exit the egress traffic manager. Data from these counters can be displayed as described in the following sections.

Displaying all traffic manager statistics for a router

To display the traffic manager statistics of the router, use the command shown in Example 8-86.

Example 8-86 Router traffic manager statistic

```
SF_B16M#show tm statistics
--------- Ports 4/1 - 4/4 ---------
Ingress Counters:
   Total Ingress Pkt Count:                 3690
   EnQue Pkt Count:                         3680
   EnQue Byte Count:                        294400
   DeQue Pkt Count:                         3680
   DeQue Byte Count:                        294400
   TotalQue Discard Pkt Count:              10
   TotalQue Discard Byte Count:             640
   Oldest Discard Pkt Count:                0
   Oldest Discard Byte Count:               0
```

```
Egress Counters:
   EnQue Pkt Count:                          3580
   EnQue Byte Count:                         229760
   Discard Pkt Count:                        0
   Discard Byte Count:                       0

--------- Ports 4/5 - 4/8 ---------
...
```

Displaying traffic manager statistics for a port group

To display the traffic manager statistic of the port group, use the command shown in Example 8-87.

Example 8-87 Port group traffic manager statistic

```
SF_B16M#show tm statistics ethernet 5/5
--------- Ports 5/1 - 5/20 ---------
Ingress Counters:
   Total Ingress Pkt Count:                  104939
   EnQue Pkt Count:                          104929
   EnQue Byte Count:                         55777520
   DeQue Pkt Count:                          104929
   DeQue Byte Count:                         55777520
   TotalQue Discard Pkt Count:               10
   TotalQue Discard Byte Count:              640
   Oldest Discard Pkt Count:                 0
   Oldest Discard Byte Count:                0
Egress Counters:
   EnQue Pkt Count:                          152593
   EnQue Byte Count:                         83525568
   Discard Pkt Count:                        0
   Discard Byte Count:                       0

SF_B16M#
```

You can select any port in the given port group to get the display for the whole port group.

Displaying traffic manager statistics for an interface module

To display the traffic manager statistic of the interface module, use the command shown in Example 8-88.

Example 8-88 Interface module traffic manager statistic

```
SF_B16M#show tm statistics slot 5
--------- Ports 5/1 - 5/20 ---------
Ingress Counters:
   Total Ingress Pkt Count:                    104988
   EnQue Pkt Count:                            104978
   EnQue Byte Count:                           55796896
   DeQue Pkt Count:                            104978
   DeQue Byte Count:                           55796896
   TotalQue Discard Pkt Count:                 10
   TotalQue Discard Byte Count:                640
   Oldest Discard Pkt Count:                   0
   Oldest Discard Byte Count:                  0
Egress Counters:
   EnQue Pkt Count:                            152667
   EnQue Byte Count:                           83555136
   Discard Pkt Count:                          0
   Discard Byte Count:                         0

SF_B16M#
```

Clearing traffic manager statistics

It is possible to clear the traffic manager statistics. Statistics can be cleared on the port group or slot as shown in Example 8-89.

Example 8-89 Clearing traffic manager statistics

```
SF_B16M#show tm statistics slot 5
--------- Ports 5/1 - 5/20 ---------
Ingress Counters:
   Total Ingress Pkt Count:                    104988
   EnQue Pkt Count:                            104978
   EnQue Byte Count:                           55796896
   DeQue Pkt Count:                            104978
   DeQue Byte Count:                           55796896
   TotalQue Discard Pkt Count:                 10
   TotalQue Discard Byte Count:                640
   Oldest Discard Pkt Count:                   0
   Oldest Discard Byte Count:                  0
```

```
Egress Counters:
   EnQue Pkt Count:                          152667
   EnQue Byte Count:                         83555136
   Discard Pkt Count:                        0
   Discard Byte Count:                       0

SF_B16M#clear tm statistics slot 5
TM Statistics of all ports on slot 5 have been cleared.
SF_B16M#show tm statistics slot 5
--------- Ports 5/1 - 5/20 ---------
Ingress Counters:
   Total Ingress Pkt Count:                  0
   EnQue Pkt Count:                          0
   EnQue Byte Count:                         0
   DeQue Pkt Count:                          0
   DeQue Byte Count:                         0
   TotalQue Discard Pkt Count:               0
   TotalQue Discard Byte Count:              0
   Oldest Discard Pkt Count:                 0
   Oldest Discard Byte Count:                0
Egress Counters:
   EnQue Pkt Count:                          0
   EnQue Byte Count:                         0
   Discard Pkt Count:                        0
   Discard Byte Count:                       0

SF_B16M#
```

9

Network traffic security

In this chapter we cover the following network security topics:

- Access Control Lists
- MAC Port Security
- DHCP Snooping and IP Source Guard

9.1 Basic IP Access Control Lists

Access Control Lists permit or deny access to IP devices based on the address or protocol within a network packet. ACLs are programmed into the Content Addressable Memory (CAM) and handled in hardware.

There are two types of Access Control Lists: Standard and Extended. Standard access lists can filter traffic based only on source IP address. Extended ACLs can filter based on source and/or destination IP, host name, protocol (such as ICMP, TCP, UDP, OSPF, and so on), TCP/UDP port number, and other criteria.

On the m-series and c-series, both inbound and outbound ACLs are supported. On the r-series, s-series, g-series, and x-series, only inbound ACLs are supported.

ACLs can be applied to physical interfaces, trunk groups/LAGs, or virtual interfaces.

Access lists can also be either numbered or named. If the access list is numbered, ACLs 1-99 are reserved for standard ACLs, while 100-199 are for Extended ACLs. One difference between numbered and names ACLs is that numbered ACLs require you to provide each rule for the ACL in one line from the Global CONFIG level, while a named ACL will take you to a different command prompt where you enter a rule line by line.

9.1.1 ACL rules

Each ACL entry is evaluated in the order you entered it. As soon as a packet matches one of the entries, it is sent or discarded, depending on the rule.

The default action when no ACLs are configured on an interface is to permit all traffic. However, after an ACL is applied to an interface an implicit **`deny all`** is appended to the ACL entry which results in the discarding of *any* traffic that has not matched a **`permit`** entry somewhere in the ACL. If you want to transmit any traffic that has not been explicitly dropped, you *must* insert a rule permitting all remaining traffic to pass.

> **Note:** Which rules must be applied, and where, is beyond the intended scope of this book.

9.1.2 Defining an ACL

The best way to define an ACL is with named ACLs using the following Global CONFIG level command: **`ip access-list standard | extended`** ***`ACL_name`***. Use **`standard`** to configure a standard named ACL and **`extended`** to configure an extended named ACL.

Following this command, you will be in the ACL CONFIG level. At this level, you simply enter in your ACL rules, one after another.

If you prefer to use numbered ACLs, you can use the **`access-list`** command from the Global CONFIG level.

Named Standard ACL rules

Each entry in a named standard ACL contains:

- An action; either **`permit`** or **`deny`**
- An IP address range *or* a specific IP address/host
- An indication if denial (or permitting) has to be logged (optional)

The full syntax is:

```
deny | permit <source-ip> | <hostname> <wildcard> [log] or
deny | permit <source-ip>/<mask-bits> | <hostname> [log] or
deny | permit host <source-ip> | <hostname> [log] or
deny | permit any [log]
```

Extended ACL rules

Each entry in an extended ACL generally contains the following elements:

- An action, which can be either **`permit`** or **`deny`**.
- A protocol, which can be simply **`ip`**, a named IP protocol (such as TCP, UDP, OSPF, IGMP, ICMP, and so on) or an IP protocol number.
- The IP source address range, a specific IP/host, or **`any`**.
- (optional) If this is a UDP or TCP rule, an operator such as **`eq`** (equals), **`gt`** (greater than) or **`lt`** (less than), followed by a port number or protocol name.
- A destination IP range or specific IP/host or **`any`**.
- (optional, m-series and c-series only) **`fragment`** or **`non-fragment`**, to match fragmented (or not) IP packets.
- (optional) If this is an ICMP rule, the ICMP message number.
- (optional) If this is a TCP rule, the keyword **`established`**, which will only match packets with the ACK or RST bits set.

- (optional) A destination UDP or TCP operator and port number (if applicable).
- (optional) An indication if denial (or permitting) has to be logged.

There are other parameters (such as **precedence**, **tos**, and so on) available, which are explained in greater detail in the reference manuals.

To give you an idea of all that an ACL rule can contain, the full syntax of a named extended ACL rules on the s-series, g-series, and x-series is as follows:

```
deny | permit
<ip-protocol>
<source-ip> | <hostname> <wildcard>
[<operator> <source-tcp/udp-port>]
<destination-ip> | <hostname> <wildcard>
[<icmp-num> | <icmp-type>]
[<tcp/udp comparison operator> <destination-tcp/udp-port>]
[dscp-cos-mapping]
[dscp-marking <0-63> [802.1p-priority-marking <0 –7>... |
dscp-cos-mapping]]
[dscp-matching <0-63>]
[log]
[precedence <name> | <0 – 7>]
[tos <0 – 63> | <name>]
[traffic policy <name>]
```

For the m-series and c-series, the syntax is as follows:

```
deny | permit
<ip-protocol>
<source-ip> | <hostname> <wildcard>
[<operator> <source-tcp/udp-port>]
<destination-ip> | <hostname> <wildcard>
[<operator> <destination-tcp/udp-port>]
[<icmp-type>]
[established]
[precedence <name> | <num>]
[tos <number>]
[dscp-mapping <number>]
[dscp-marking <number>]
[fragment | non-fragment]
[option value | name | keyword]
[ priority <priority-value> | priority-force <priority-value> |
priority-mapping <priority-value> ]
[mirror]
```

For the r-series, the syntax is as follows:

```
deny | permit <ip-protocol>
<source-ip> | <hostname> <wildcard>
[<operator> <source-tcp/udp-port>]
<destination-ip> | <hostname> <wildcard>
[<operator> <destination-tcp/udp-port>]
[match-all <tcp-flags>] [match-any <tcp-flags>]
[<icmp-type>] [established] [precedence <name> | <num>]
[tos <number>] [dscp-matching <number>]
[802.1p-priority-matching <number>]
[dscp-marking <number> 802.1p-priority-marking <number>
internal-priority-marking <number>] | [dscp-marking
<number> dscp-cos-mapping] | [dscp-cos-mapping]
[fragment] [non-fragment] [first-fragment]
[fragment-offset <number>]
[spi <00000000 - ffffffff>] [log]
```

Later, we show some example rules.

ACL wildcards

As part of most ACL rules, a range of IP addresses is supplied, which can appear in one of two forms: CIDR and bitmask. The Classless Inter-Domain Routing (CIDR) format is the standard notation for an IP network, where the network address is followed by a slash and the number of significant bits, that is,/24. The bitmask format is slightly different from what you might be accustomed to for a subnet mask. Instead of a 24-bit mask taking the form 255.255.255.0, it is instead expressed as 0.0.0.255. Either format will work, although the bitmask format is what will appear in your running configuration.

9.1.3 Example ACLs and ACL rules

In the following topics, we supply some sample named ACLs to give you a basic overview and idea as to their proper syntax and application.

Standard ACL

Our ACL in Example 9-1 is explicitly allowing traffic from two networks, denying it (with logging) from a third, and denying (without logging) all the rest.

Example 9-1 A sample standard ACL

```
telnet@m_Series(config)#ip access-list standard sample_1
telnet@m_Series(config-std-nacl)#
telnet@m_Series(config-std-nacl)#permit 1.2.3.4/24
```

```
telnet@m_Series(config-std-nacl)#permit 2.3.4.5 0.0.0.255
telnet@m_Series(config-std-nacl)#deny 3.4.5.6/24 log
telnet@m_Series(config-std-nacl)#exit
telnet@m_Series(config)#
telnet@m_Series(config)#show access-list name sample_1

Standard IP access list  sample_1
permit 1.2.3.0 0.0.0.255
permit 2.3.4.0 0.0.0.255
deny 3.4.5.0 0.0.0.255 log
[Even though it is not shown, the ACL will drop any traffic that has
not matched one of the three rules. This means there is an implied
“deny any” at the end.]
telnet@m_Series(config)#
```

Extended ACL

Our sample extended ACL in Example 9-2 accomplishes the following tasks:

- Drops all ping traffic.
- Lets all remaining ICMP traffic through.
- Only allows a web host to communicate with destinations through port 80, and ensures that the web host does not try to initiate any communications of its own.
- Allows TFTP traffic only from a single workstation.
- Allows Telnet traffic only from a single workstation to only a single destination.
- Drops all other telnet traffic on that workstation’s network and logs it.
- Drops all remaining traffic (implied.)

Example 9-2   A sample extended ACL

```
telnet@m_Series(config)#
telnet@m_Series(config)#ip access-list extended sample_2
telnet@m_Series(config-ext-nacl)#
telnet@m_Series(config-ext-nacl)#deny icmp any any echo
telnet@m_Series(config-ext-nacl)#deny icmp any any echo-reply
telnet@m_Series(config-ext-nacl)#permit icmp any any
telnet@m_Series(config-ext-nacl)#
telnet@m_Series(config-ext-nacl)#permit tcp any host 1.2.3.4 eq 80
telnet@m_Series(config-ext-nacl)#permit tcp host 1.2.3.4 eq 80 any
established
telnet@m_Series(config-ext-nacl)#
telnet@m_Series(config-ext-nacl)#permit udp host 2.3.4.5 eq tftp any
telnet@m_Series(config-ext-nacl)#
telnet@m_Series(config-ext-nacl)#permit tcp host 3.4.5.6 eq telnet host
4.5.6.7
```

```
telnet@m_Series(config-ext-nacl)#
telnet@m_Series(config-ext-nacl)#deny tcp 3.4.5.0 255   0.0.0.255 eq
telnet host 4.5.6.7 log
[Even though it is not shown, the ACL will drop any traffic that has
not matched one of the three rules.]
telnet@m_Series(config-ext-nacl)#exit
telnet@m_Series(config)#
telnet@m_Series(config)#show access-list named sample_2

Extended IP access list  sample_2
    0: deny icmp any any echo
    1: deny icmp any any echo-reply
    2: permit icmp any any
    3: permit tcp any host 1.2.3.4 eq http
    4: permit tcp host 1.2.3.4 eq http any established
    5: permit udp host 2.3.4.5 eq tftp any
    6: permit tcp host 3.4.5.6 eq telnet host 4.5.6.7
    7: deny tcp 3.4.5.0 0.0.0.255 eq telnet host 4.5.6.7 log

telnet@m_Series(config)#

[Even though it is not shown, the ACL will drop any traffic that has
not matched one of the three rules. This means there is an implied
“deny any” at the end.]
```

9.1.4 Applying an ACL

To apply an access-list to an IP interface, simply enter the interface configuration level, and run the command **ip access-group ACL_name in**. The **in** refers to whether this is going to be used as an inbound or outbound access list. Most access lists are used for inbound traffic. The outbound ACL feature (using the out option) is only available on c-series and m-series products.

ACL re-application

If you have changed an ACL, you can apply the new ACL to interfaces using the config level command **ip rebind-acl *ACL_name***. In place of a specific ACL name or number, you can also rebind all ACLs.

Per port - per VLAN application

You can apply ACLs only to traffic on a certain VLAN on a particular port. Doing so can be useful for filtering traffic in one of your more central switches, instead of having to configure ACLs on each edge switch in your network.

The first step is to enable this function (it is not enabled by default.) To do this, run the config level command **enable acl-per-port-per-vlan**. Save it to the startup configuration and reload the product. (The reload is required.)

To apply an ACL to a particular VLAN on a particular port, go into the interface level config and enter the command **per-vlan *12***. From there, apply the ACLs to the VLAN with the **ip access-list *sample_1 in*** command. We show this process in Example 9-3.

Example 9-3 Applying a ACL to a VLAN on a particular port

```
FastIron(config)#int e 1/23
FastIron(config-if-e1000-1/23))#per-vlan 12
FastIron(config-if-e1000-1/23-vlan-12))#ip access-group 10 in
```

9.2 MAC Port Security

MAC Port Security is available to limit the number of devices that can communicate over a given network port. This prevents users from inserting unauthorized switches on the network.

To enable this feature on a particular port, enter the interface config level and enter the command **port security**. This puts you into the port security config level. Next, enter the command **enable**.

Maximum MACs per port

The default maximum number of MAC addresses on a port when port security is enabled is 1. If you want to increase this maximum (for instance, if a port will be used for an IP-phone chained to a PC), enter the command **maximum *2*** from the port security config level.

Age-out of learned MAC addresses

Oddly, the default is never to age-out learned MAC addresses. This means manual intervention is required if a different device was ever attached to the network port. (You have to disable and then re-enable port-security on the affected port.) To change this, enter the command **age *5*** at the port security config level. (The unit for this command is minutes).

Action to take during a violation

You can choose one of two actions to take in the case of a violation: either drop the frames from the violating MAC, or shut the port down entirely for a period of time.

On the s-series, g-series, and x-series, the following considerations apply:

- To shut down traffic from the violating MAC for a period of time, enter the command **violation restrict *5***. After the time (in minutes; 5 in this case) has elapsed, it will perform another check on the next connection attempt; if the original address has aged-out, then the new MAC will not be restricted any more.
- To shut the port down entirely for a period of time, enter the port security level command **violation shutdown *5***. The “5” indicates the number of minutes to shut the port down. If you have swapped out the device attached to the port, you simply have to wait for the MAC to age out. (assuming you have set the ageing timer.)

On the m-series, c-series, and r-series, the following considerations apply:

- To shut down a port after a certain number of violations, enter the port security level command **violation restrict** to enable the feature. You can also specify for the port to shut down after a burst of security violations, for example **violation restrict 100** will shutdown the port after 100 violations occurring within a second. The **restrict-max-deny 50** command is used to define how many security violations can occur on the port before the port is shut down, where “50” indicates the number of violations that can occur before the port is shut down.
- To shut down a port for a certain amount of time after any security violation occurs, use the **violation shutdown *5*** where “5” indicates the number of minutes to shut the port down,

Example

We show the overall enabling, along with some of the configuration options in Example 9-4.

Example 9-4 Port Security

```
BR-FastIron SX 1600 Router(config)#interface ethernet 11/1
BR-FastIron SX 1600 Router(config-if-e1000-11/1)#port security
BR-FastIron SX 1600 Router(config-port-security-e1000-11/1)#enable
BR-FastIron SX 1600 Router(config-port-security-e1000-11/1)#maximum 2
BR-FastIron SX 1600 Router(config-port-security-e1000-11/1)#violation
restrict 5
BR-FastIron SX 1600 Router(config-port-security-e1000-11/1)#
BR-FastIron SX 1600 Router(config-port-security-e1000-11/1)#interface
ethernet 11/2
BR-FastIron SX 1600 Router(config-if-e1000-11/2)#port security
BR-FastIron SX 1600 Router(config-port-security-e1000-11/2)#enable
BR-FastIron SX 1600 Router(config-port-security-e1000-11/2)#maximum 16
```

```
BR-FastIron SX 1600 Router(config-port-security-e1000-11/2)#violation
shutdown 5
BR-FastIron SX 1600 Router(config-port-security-e1000-11/2)#
```

9.3 DHCP Snooping

Note: Be aware that the DHCP Snooping feature is not currently available on the IBM Ethernet Switch B50G or the IBM x-series.

To prevent an unauthorized DHCP server from disrupting your network, you can enable the DHCP Snooping feature. If you enable this feature, only DHCP servers on ports you specifically designate will work. Any other DHCP servers will not function.

A prerequisite for this feature is the acl-per-port-per-vlan option. To enable this option (which requires a reboot) run the config level command **enable acl-per-port-per-vlan**.

Next, enable DHCP Snooping on the VLANs where you want to use this feature. This is done with the config level command **ip dhcp snooping vlan *2*** (s-series) or **ip dhcp-snooping vlan *2*** (m-series and c-series).

Finally, you need to set the ports that have legitimate DHCP servers to “trusted”. Until you do so, DHCP will not work on the VLAN. This is done by entering the interface config level and entering the command **dhcp snooping trust** (s-series and G4A) or **dhcp-snooping-trust** (c-series and m-series).

We go through the entire process in Example 9-5. In our example, we enable snooping on VLAN 2, and we set up port 11/1 as our trusted DHCP server. (We have already enabled acl-per-port-per-vlan.)

Example 9-5 DHCP Snooping

```
FastIron(config)#ip dhcp snooping vlan 2
FastIron(config)#interface ethernet 11/1
FastIron(config-if-e1000-11/1)#dhcp-snooping-trust
FastIron(config-if-e1000-11/1)#exit
```

9.4 IP Source Guard

Note: Be aware that the IP Source Guard feature is not currently available on the IBM Ethernet Switch B50G or the IBM x-series.

To prevent IP address spoofing by DHCP-connected hosts, you can use the IP Source Guard feature. This feature monitors the content of DHCP packets and use that information to drop any packets that originate from an IP address different from the one assigned by DHCP. Until the host obtains an address from DHCP, all other IP traffic will be blocked.

This feature is meant to be used on DHCP client ports, such as those used to connect to PCs. While it is possible to run it on ports connected to switches (or on tagged ports), there are multiple restrictions on doing so. (See the *Configuration Guide* for details.)

A prerequisite for this feature is enabling the DHCP Snooping feature in the previous section.

After you have DHCP-snooping enabled, enabling IP Source Guard is straightforward: enter the interface config level, and enter the command **`source-guard enable`** (s-series and G4A) or **`source-guard`** (m-series and c-series).

We demonstrate the process on an m-series in Example 9-6.

Example 9-6 IP Source Guard

```
telnet@m_Series(config)#interface ethernet 5/1
telnet@m_Series(config-if-e1000-5/1)#source-guard
telnet@m_Series(config-if-e1000-5/1)#
```

9.5 MAC Address Authentication (also known as Multi-Device Port Authentication)

> **Note:** Be aware that the MAC Address Authentication feature is not currently available on the IBM x-series.

It is becoming very common to restrict access to a network to certain MAC addresses only. For instance, this can help prevent access to a private network by a visitor who simply plugs their laptop into a port on an unused desk. This feature is known as “Multi-Device Port Authentication.”

> **Note:** On the IBM Ethernet Switch B50G, MAC Address Authentication cannot be used on the same port as MAC Port Security.

At a product level, this feature uses a RADIUS server to verify the MAC address. (The MAC address is used as both the username and password submitted to the RADIUS server.)

You can even use this feature to dynamically assign a particular MAC to a specific VLAN and/or assign a particular ACL to it. (See the *Configuration Guide* for details on these advanced features.)

As a prerequisite to this feature, you must set up a RADIUS server, and define it on the product. The most basic way to do this is with the config level command **`radius-server host`** ***`1.2.3.4`***. Additional syntax can be defined for specifying different ports or servers for RADIUS accounting, authentication, or authorization and the syntax is:

```
radius-server host <ip-addr> | <server-name> [auth-port <number>
acct-port <number>]
```

Where:

- **`host <ip-addr> | <server-name>`** parameter is either an IP address or an ASCII text string.
- **`<auth-port>`** parameter is the Authentication port number; it is an optional parameter. The default is 1812.
- **`<acct-port>`** parameter is the Accounting port number; it is an optional parameter. The default is 1813.

MAC Authentication itself is enabled using the **`mac-authentication enable`** command at the config level.

You then apply it to interfaces by running the command, **`mac-authentication enable ethernet`** ***`5/1`***. You can also specify a port range, or `all` to enable it on all ports. The feature will not work on trunk ports.

We show a configuration example for this feature in Example 9-7.

Example 9-7 MAC Address Authentication

```
telnet@m_Series(config)#radius-server host 1.2.3.4
telnet@m_Series(config)#mac-authentication enable
telnet@m_Series(config)#mac-authentication enable ethernet 5/1
Error - port 5/1 is a trunk port. Mac-Authentication cannot be enabled.
telnet@m_Series(config)#mac-authentication enable ethernet 5/5
telnet@m_Series(config)#mac-authentication enable ethernet 5/6 to 5/8
telnet@m_Series(config)#
```

10

PoE and LLDP

In this chapter, we describe how to configure IBM b-type network devices to support Power over Ethernet (PoE) devices and Link Layer Discovery Protocol (LLDP).

10.1 Power over Ethernet (PoE)

In this section we assume that you are familiar with the principles of Power over Ethernet (PoE).

10.1.1 Enabling and disabling PoE on the port

PoE can be enabled or disabled as described in the following topics.

Enabling support for legacy devices

PoE can be enabled on the interface level with the **`inline power`** command as shown in Example 10-1.

Example 10-1 Enabling inline power on the interface 1/1/46

```
Edge_L2_B50G>show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 960000 mWatts.

Power Allocations:      Requests Honored 1 times

 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
--------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/6  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/7  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/8  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/9  Off     Off            0          0  n/a      n/a         3  n/a
1/1/10  Off     Off            0          0  n/a      n/a         3  n/a
.....
1/1/46  Off     Off            0          0  n/a      n/a         3  n/a
1/1/47  Off     Off            0          0  n/a      n/a         3  n/a
1/1/48  Off     Off            0          0  n/a      n/a         3  n/a
--------------------------------------------------------------------------
 Total                         0          0
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/46
Edge_L2_B50G(config-if-e1000-1/1/46)#
```

```
Edge_L2_B50G(config-if-e1000-1/1/46)#inline power
Edge_L2_B50G(config-if-e1000-1/1/46)#
Edge_L2_B50G(config-if-e1000-1/1/46)#exit
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#exit
Edge_L2_B50G#
Edge_L2_B50G#show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 944600 mWatts.

Power Allocations:      Requests Honored 2 times

 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
--------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/6  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/7  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/8  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/9  Off     Off            0          0  n/a      n/a         3  n/a
1/1/10  Off     Off            0          0  n/a      n/a         3  n/a
.....
1/1/44  Off     Off            0          0  n/a      n/a         3  n/a
1/1/45  Off     Off            0          0  n/a      n/a         3  n/a
1/1/46  On      Off            0          0  n/a      n/a         3  n/a
1/1/47  Off     Off            0          0  n/a      n/a         3  n/a
1/1/48  Off     Off            0          0  n/a      n/a         3  n/a
--------------------------------------------------------------------------
 Total                         0          0
```

PoE can be disabled at the port level with the **no inline power** command as shown in Example 10-2.

Example 10-2 Disabling inline power on the interface 1/1/46

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/46
Edge_L2_B50G(config-if-e1000-1/1/46)#
Edge_L2_B50G(config-if-e1000-1/1/46)#no inline power
Edge_L2_B50G(config-if-e1000-1/1/46)#
```

```
Edge_L2_B50G(config-if-e1000-1/1/46)#PoE: Power disabled on port 1/1/46
because of admin off.

Edge_L2_B50G(config-if-e1000-1/1/46)#
```

IBM b-type switches can automatically detect if a connected PoE device is a legacy device or an 802.3af compliant device. Legacy devices are supported without any additional configuration.

Disabling support for legacy devices

To disable support for legacy devices, the global command **`no legacy-inline-power`** can be used as shown in Example 10-3.

Example 10-3 Disabling support for legacy devices

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no legacy-inline-power
Edge_L2_B50G(config)#
```

10.1.2 CDP PoE requirement detection

Some PoE capable devices advertise their power requirement using CDP (Cisco Discovery Protocol). IBM b-type network switches support the CDP protocol for power requirement.

To enable an IBM b-type switch to detect the CDP power requirement, use the **`cdp run`** command at the global config level as shown in Example 10-4.
To disable the feature, use **`no cdp run`** as also shown in Example 10-4.

Example 10-4 Enabling and disabling CDP

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#cdp run
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#exit
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no cdp run
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#exit
Edge_L2_B50G#
```

The following requirements must be met:

- If a port is configured with a maximum power level or a power class for a power consuming device, the power level or power class takes precedence over the CDP power requirement. Therefore, if it is required that the device adheres to the CDP power requirement, then power level or power class on the port must not be configured.
- The IBM b-type switch PoE will adjust a port's power only if there are available power resources on the device.

10.1.3 Power level

By default, IBM b-type switches deliver 15.4W (Class 3) of power when no maximum value is specified.

Maximum PoE power can be set in two different ways:

- Specifying PoE power in milliwatts (mW)

  To specify the maximum power in milliwatts, the **`inline power power-limit Value`** command can be used at the interface level, where **`Value`** represents the power value in milliwatts, as shown in Example 10-5.

Example 10-5 Power limit setup with mW

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/46
Edge_L2_B50G(config-if-e1000-1/1/46)#
Edge_L2_B50G(config-if-e1000-1/1/46)#inline power power-limit 9000
Warning: Inline power configuration on port 1/1/46 has been modified.
Edge_L2_B50G(config-if-e1000-1/1/46)#
Edge_L2_B50G(config-if-e1000-1/1/46)#show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 951000 mWatts.

Power Allocations:      Requests Honored 4 times


 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
--------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a         3  n/a
```

```
....
1/1/46  On      Off            0       9000  n/a      n/a         3  n/a
1/1/47  Off     Off            0          0  n/a      n/a         3  n/a
1/1/48  Off     Off            0          0  n/a      n/a         3  n/a
--------------------------------------------------------------------------
 Total                         0       9000
```

- Specifying the PoE power based on the 802.3af power class

  To specify maximum power using the 802.3af power class the **`inline power power-by-class ClassValue`** command can be used, where **`ClassValue`** represents 802.3af power class. An example of the command is shown in Example 10-6.

Example 10-6 Power limit setup with power class

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/45
Edge_L2_B50G(config-if-e1000-1/1/45)#
Edge_L2_B50G(config-if-e1000-1/1/45)#inline power power-by-class 2
Warning: Inline power configuration on port 1/1/45 has been modified.
Edge_L2_B50G(config-if-e1000-1/1/45)#show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 953000 mWatts.

Power Allocations:      Requests Honored 3 times


 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
--------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a         3  n/a
....
1/1/44  Off     Off            0          0  n/a      n/a         3  n/a
1/1/45  On      Off            0       7000  n/a      n/a         3  n/a
1/1/46  On      Off            0          0  n/a      n/a         3  n/a
1/1/47  Off     Off            0          0  n/a      n/a         3  n/a
1/1/48  Off     Off            0          0  n/a      n/a         3  n/a
--------------------------------------------------------------------------
 Total                         0       7000
```

Only one maximum type can be defined on the port.

Table 10-1 shows the power consumption of the 802.3af PoE power classes.

Table 10-1 Power consumption

Class	Usage	Power (Watts)
0	default	15.4
1	optional	4
2	optional	7
3	optional	15.4

10.1.4 Power priority

It is possible to configure the power priority which will take precedence in case of power supply failures or reduced PoE power supply. The command, **`inline power priority Priority,`** at the interface level, as shown in Example 10-7, can be used to set up the PoE priority.

Example 10-7 PoE priority setting

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/46
Edge_L2_B50G(config-if-e1000-1/1/46)#
Edge_L2_B50G(config-if-e1000-1/1/46)#inline power priority 1
Warning: Inline power configuration on port 1/1/46 has been modified.
Edge_L2_B50G(config-if-e1000-1/1/46)#show inline power

Power Capacity:        Total is 960000 mWatts. Current Free is 960000 mWatts.

Power Allocations:     Requests Honored 6 times


 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
----------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a         3  n/a
....
```

```
1/1/46  On      Off            0          0  n/a      n/a         1  n/a
1/1/47  Off     Off            0          0  n/a      n/a         3  n/a
1/1/48  Off     Off            0          0  n/a      n/a         3  n/a
--------------------------------------------------------------------------
 Total                         0          0
```

The PoE priority (**Priority** parameter) can have three values:

- 1 - Critical
- 2 - High
- 3 - Low

In the case of a power failure firstly ports with priority 1 (Critical), then 2 (High) and then 3 (Low) will be supplied with the required power.

10.1.5 Resetting PoE parameters

To override or reset POE port parameters including the power priority, power class, and maximum power level, each POE parameter must be specified in the CLI command line, as shown in the following examples.

Example 10-8 is for the **inline power priority 1 power-limit 3000** command.

Example 10-8 Inline power setting of priority and power limit in mW together

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/34
Edge_L2_B50G(config-if-e1000-1/1/34)#inline power priority 1 power-limit 3000
Edge_L2_B50G(config-if-e1000-1/1/34)#show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 957000 mWatts.

Power Allocations:      Requests Honored 8 times


 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
--------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a         3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a         3  n/a
....
1/1/33  Off     Off            0          0  n/a      n/a         3  n/a
```

```
1/1/34  On      Off            0       3000  n/a      n/a          1  n/a
1/1/35  Off     Off            0          0  n/a      n/a          3  n/a
```

Example 10-9 is for the **inline power priority 2 power-by-class 1** command.

Example 10-9 Inline power setting of priority and power limit with power class together

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#interface ethernet 1/1/34
Edge_L2_B50G(config-if-e1000-1/1/34)#inline power priority 2 power-by-class 1
Edge_L2_B50G(config-if-e1000-1/1/34)#show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 957000 mWatts.

Power Allocations:      Requests Honored 8 times


 Port   Admin   Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
        State   State   Consumed  Allocated                          Error
--------------------------------------------------------------------------
1/1/1-  Off     Off            0          0  n/a      n/a          3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a          3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a          3  n/a
1/1/1-  Off     Off            0          0  n/a      n/a          3  n/a
 1/1/5  Off     Off            0          0  n/a      n/a          3  n/a
....
1/1/33  Off     Off            0          0  n/a      n/a          3  n/a
1/1/34  On      Off            0       4000  n/a      n/a          2  n/a
1/1/35  Off     Off            0          0  n/a      n/a          3  n/a
....
```

10.1.6 Displaying PoE information

The **show inline power** command can be used to display the operational status of PoE. An example of the command is shown in Example 10-10.

Example 10-10 Displaying inline power info

```
Edge_L2_B50G#show inline power

Power Capacity:         Total is 960000 mWatts. Current Free is 960000 mWatts.

Power Allocations:      Requests Honored 9 times
```

```
 Port    Admin    Oper    ---Power(mWatts)---  PD Type  PD Class  Pri  Fault/
         State    State   Consumed  Allocated                          Error
---------------------------------------------------------------------------
1/1/1-   Off      Off            0          0  n/a      n/a         3  n/a
1/1/1-   Off      Off            0          0  n/a      n/a         3  n/a
1/1/1-   Off      Off            0          0  n/a      n/a         3  n/a
1/1/1-   Off      Off            0          0  n/a      n/a         3  n/a
 1/1/5   Off      Off            0          0  n/a      n/a         3  n/a
 1/1/6   Off      Off            0          0  n/a      n/a         3  n/a
 1/1/7   Off      Off            0          0  n/a      n/a         3  n/a
 1/1/8   Off      Off            0          0  n/a      n/a         3  n/a
 1/1/9   Off      Off            0          0  n/a      n/a         3  n/a
1/1/10   Off      Off            0          0  n/a      n/a         3  n/a
....
```

To display details about PoE power supplies the **`show inline power detail`** command can be used as shown in Example 10-11.

Example 10-11   Displaying detailed inline power info

```
Edge_L2_B50G#show inline power detail


Power Supply Data On stack 1:
++++++++++++++++++

Power Supply #1:
        Max Curr:       10.0 Amps
        Voltage:        48.0 Volts
        Capacity:       480 Watts
Power Supply #2:
        Max Curr:       10.0 Amps
        Voltage:        48.0 Volts
        Capacity:       480 Watts


POE Details Info. On Stack 1 :


General PoE Data:
+++++++++++++++++
Firmware
Version
--------
04.0.004.0.0


Cumulative Port State Data:
++++++++++++++++++++++++++

#Ports    #Ports     #Ports   #Ports    #Ports       #Ports     #Ports
Admin-On  Admin-Off  Oper-On  Oper-Off  Off-Denied   Off-No-PD  Off-Fault
-------------------------------------------------------------------------
0         24         0        24        0            0          0
1         23         0        24        0            1          0



Cumulative Port Power Data:
++++++++++++++++++++++++++

#Ports  #Ports  #Ports        Power       Power
Pri: 1  Pri: 2  Pri: 3  Consumption  Allocation
-----------------------------------------------
0       0       0           0.0   W     0.0   W
0       1       0           0.0   W     0.0   W


Edge_L2_B50G#
```

10.2 Link Layer Discovery Protocol (LLDP)

This section assumes that you are familiar with the principles of the Link Layer Discovery Protocol (LLDP).

10.2.1 Configuration considerations

Here are various configuration considerations when LLDP is used:

- LLDP is supported on Ethernet interfaces only.
- If a port is 802.1X-enabled, the transmission and reception of LLDP packets will only take place while the port is authorized.

- Cisco Discovery Protocol (CDP) and Foundry Discovery Protocol (FDP) run independently of LLDP. Therefore, these discovery protocols can run simultaneously on the same device.
- By default, the IBM b-type networking device limits the number of neighbors per port to four, and staggers the transmission of LLDP packets on different ports, in order to minimize any high-usage spikes to the CPU.
- By default, the IBM b-type device forwards.
- Ports that are in blocking mode (spanning tree) can still receive LLDP packets from a forwarding port.
- Auto-negotiation status indicates what is being advertised by the port for 802.3 auto-negotiation.

10.2.2 Enabling and disabling LLDP

By default LLDP is disabled. It is enabled by default on individual ports, but to run it, it has to be enabled globally.

Enabling LLDP

To enable LLDP, the **`lldp run`** command at the global config level can be used as shown in Example 10-12.

Example 10-12 Enabling LLDP

```
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp run
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show running
Current configuration:
!
ver 05.0.02aT7e1
!
....
lldp run
!
!
!
!
end
```

Disabling LLDP

To disable LLDP, the **no lldp run** command at the global config level can be used as shown in Example 10-13.

Example 10-13 Disabling LLDP

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp run
Edge_L2_B50G(config)#
```

10.2.3 Enabling support for tagged LLDP packets

By default IBM b-type devices do not accept tagged LLDP packets from other vendors devices. To enable that support use the **lldp tagged-packets process** command at the global config level as shown in Example 10-14.

Example 10-14 Enabling LLDP packet acceptance

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp tagged-packets process
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show running
Current configuration:
!
ver 05.0.02aT7e1
!
....
lldp tagged-packets process
lldp run
!
!
!
!
end
```

When this feature is enabled, the device will accept incoming LLDP tagged packets if the VLAN tag matches any of the following elements:

- A configured VLAN on the port
- The default VLAN for a tagged port
- The configured untagged VLAN for a dual-mode port

The support for tagged LLDP packets can be disabled with the **no lldp tagged-packets process** command as shown in Example 10-15.

Example 10-15 Disabling LLDP packet acceptance

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp tagged-packets process
Edge_L2_B50G(config)#
```

10.2.4 Changing LLDP operating mode

After LLDP is enabled at the global level, LLDP packets can be transmitted and received on the device ports. It is possible to disable the ability of a particular port to transmit or receive LLDP packed, or even changing the operational mode to only transmit or receive LLDP information.

Different operating modes can be configured for each port on the IBM b-type device.

Enabling receive and transmit mode

By default, receive and transmit mode is enabled on all ports. After this has been disabled on a particular port, it can be enabled with the **lldp enable ports PortList** command at the global config level, where **PortList** is the list of ports, as shown in Example 10-16.

Example 10-16 Enabling receive and transmit mode

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp enable ports e 1/1/30 e 1/1/40
Edge_L2_B50G(config)#
```

Disabling receive and transmit mode

With the **no lldp enable ports PortList** command, at the global config level, it is possible to disable receive and transmit mode on the ports listed in the **PortList** as shown in Example 10-17.

Example 10-17 Disabling receive and transmit mode

```
Edge_L2_B50G#
Edge_L2_B50G#config t
```

```
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable ports e 1/1/30 e 1/1/40
Edge_L2_B50G(config)#show running
Current configuration:
!
ver 05.0.02aT7e1
!
....
no lldp enable ports ethe 1/1/30 ethe 1/1/40
lldp run
!
!
!
!
end
```

In both cases ports can be listed individually, in a range using the **`to`** keyword to specify the range, and by using the **`all keyword`** all ports can be specified.

Enabling receive only mode

To enable receive only mode on the ports, with default configuration, the **`no lldp enable transmit ports PortList`** command can be used to simply disable the transmit mode as shown in Example 10-18.

Example 10-18 Enabling receive only mode

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable transmit ports e 1/1/30 e 1/1/40
Edge_L2_B50G(config)#show running
Current configuration:
!
ver 05.0.02aT7e1
!
....
no lldp enable transmit ports ethe 1/1/30 ethe 1/1/40
lldp run
!
!
!
!
end
```

If the ports are already in transmit only mode, they can be put into receive only mode by the following steps, as shown in Example 10-19.

1. Disabling transmit only mode with the **`no lldp enable transmit ports PortList`** command
2. Enabling receive only mode with the **`lldp enable receive ports PortList`** command

Example 10-19 Enabling receive only mode II

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable transmit ports e 1/1/20 to 1/1/25
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp enable receive ports e 1/1/20 to 1/1/25
```

Disabling receive only mode

To disable receive only mode on specific ports **`no lldp enable receive ports PortList`** command can be used, as shown in Example 10-20.

Example 10-20 Disabling receive only mode

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable receive ports e 1/1/16
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show running
Current configuration:
!
ver 05.0.02aT7e1
!
....
no lldp enable receive ports ethe 1/1/16
lldp run
!
!
!
!
end
```

In both cases ports can be listed individually, or in a range using the **`to`** keyword to specify the range, and by using the **`all`** keyword all ports can be specified.

Enabling transmit only mode

To enable transmit only mode on the ports, with default configuration, the **`no lldp enable receive ports PortList`** command can be used to simply disable receive mode as shown in Example 10-21.

Example 10-21 Enabling transmit only mode

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable receive ports e 1/1/18
Edge_L2_B50G(config)#
```

If the ports are already in receive only mode, they can be put into transmit only mode by the following steps, as shown in Example 10-22:

1. Disabling receive only mode with the **`no lldp enable receive ports PortList`** command
2. Enabling transmit only mode with the **`lldp enable transmit ports PortList`** command

Example 10-22 Enabling transmit only mode II

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable receive ports e 1/1/18
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp enable transmit ports e 1/1/18
Edge_L2_B50G(config)#
```

Disabling transmit only mode

To disable transmit only mode on specific ports, the **`no lldp enable transmit ports PortList`** command can be used, as shown in Example 10-23.

Example 10-23 Disabling transmit only mode

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable transmit ports e 1/1/18
Edge_L2_B50G(config)#
```

In both cases ports can be listed individually, in a range using **to** keyword to specify the range, and by using the **all** keyword all ports can be specified.

10.2.5 LLDP neighbors

It is possible to specify the maximum number of the LLDP neighbors for which LLDP data will be retained. This maximum can be specified per device or per port.

Per device

To specify the maximum number of LLDP neighbors data per device, the **lldp max-total-neighbors Value** command can be used as shown in Example 10-24.

Example 10-24 Maximum LLDP neighbors per device

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp max-total-neighbors 200
Edge_L2_B50G(config)#
```

The **Value** can range between 16 and 65536, and the default **Value** is 392.

The **show lldp** command can be used to display the currently configured value, as shown in Example 10-25.

Example 10-25 LLDP info

```
Edge_L2_B50G#show lldp
LLDP transmit interval           : 30 seconds
LLDP transmit hold multiplier    : 4  (transmit TTL: 120 seconds)
LLDP transmit delay              : 2 seconds
LLDP SNMP notification interval  : 5 seconds
LLDP reinitialize delay          : 2 seconds
LLDP-MED fast start repeat count : 3

LLDP maximum neighbors           : 200
LLDP maximum neighbors per port  : 4
Edge_L2_B50G#
```

Per port

The maximum number of LLDP neighbors data per port can be defined by using the command **lldp max-neighbors-per-port Value** as shown in Example 10-26.

Example 10-26 Maximum LLDP neighbors per port

```
Edge_L2_B50G#
Edge_L2_B50G#config t
```

```
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp max-neighbors-per-port 6
Edge_L2_B50G(config)#
```

The **Value** can range between 1 and 64, the default **Value** is 4.

The **show lldp** command can be used to verify the currently configured value.

10.2.6 LLDP SNMP notifications and syslog messages

SNMP notification and syslog messages provides information for management application.

When LLDP is enabled, LLDP SNMP notification and syslog messages are disabled by default.

Enabling

When SNMP notifications are enabled, corresponding syslog messages are also enabled. To enable them, the **lldp enable snmp notifications ports PortsList** command can be used as shown in Example 10-27.

Example 10-27 Enabling SNMP notifications

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp enable snmp notifications ports all
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show running
Current configuration:
!
ver 05.0.02aT7e1
!
....
lldp enable snmp notifications ports ethe 1/1/1 ethe 1/1/5 to 1/1/48 ethe 1/2/1 to
1/2/2
lldp run
!
!
!
!
end
```

Disabling

To disable SNMP notifications and syslog messages, the **`no lldp enable snmp notifications ports PortsList`** command can be used as shown in Example 10-28.

Example 10-28 Disabling SNMP notifications

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp enable snmp notifications ports all
Edge_L2_B50G(config)#
```

In both cases ports can be listed individually, in a range using the **`to`** keyword to specify the range, and by using the **`all`** keyword all ports can be specified.

Time between SNMP notifications and syslog messages

When notifications and messages are enabled the device will send up to one SNMP notification and a corresponding syslog message every five seconds.

It is possible to specify this interval to range between 5 to 3600 (1 hour) seconds with the **`lldp snmp-notification-interval Interval`** command, as shown in Example 10-29.

Example 10-29 Time between notifications

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp snmp-notification-interval 10
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show lldp
LLDP transmit interval         : 30 seconds
LLDP transmit hold multiplier  : 4  (transmit TTL: 120 seconds)
LLDP transmit delay            : 2 seconds
LLDP SNMP notification interval  : 10 seconds
LLDP reinitialize delay        : 2 seconds
LLDP-MED fast start repeat count : 3

LLDP maximum neighbors         : 392
LLDP maximum neighbors per port  : 4
```

Note: As LLDP syslog messages are rate limited, some LLDP information supplied by the system will not match the current LLDP statistic as shown in the **`show lldp statistic`** command output.

10.2.7 Minimum time between LLDP transmissions

By default, LLDP frames are transmitted every two seconds. If required, the delay can be set to range between 1 and 8192 seconds with the **`lldp transmit-delay Interval`** as shown in Example 10-30.

Example 10-30 Minimum time between LLDP transmissions

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp transmit-delay 5
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show lldp
LLDP transmit interval          : 30 seconds
LLDP transmit hold multiplier   : 4  (transmit TTL: 120 seconds)
LLDP transmit delay             : 5 seconds
LLDP SNMP notification interval : 10 seconds
LLDP reinitialize delay         : 2 seconds
LLDP-MED fast start repeat count : 3

LLDP maximum neighbors          : 392
LLDP maximum neighbors per port  : 4
```

The LLDP transmit delay prevents the LLDP agent from transmitting a lot of successive LLDP frames in a short time period when rapid changes occur in LLDP.

Note: The **`Interval`** value must not be greater than one quarter of the LLDP transmission interval, as explained in the next section.

10.2.8 Interval between regular LLDP transmissions

The default interval between regular LLDP transmissions is 30 seconds. It is possible to change that interval to range between 5 to 32768 seconds with the **`lldp transmit-interval Interval`** command, as shown in Example 10-31.

Example 10-31   Interval between regular LLDP transmissions

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp transmit-interval 40
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show lldp
LLDP transmit interval          : 40 seconds
LLDP transmit hold multiplier   : 4  (transmit TTL: 160 seconds)
LLDP transmit delay             : 5 seconds
LLDP SNMP notification interval  : 10 seconds
LLDP reinitialize delay         : 2 seconds
LLDP-MED fast start repeat count : 3

LLDP maximum neighbors          : 392
LLDP maximum neighbors per port  : 4
```

10.2.9 Holdtime multiplier for transmit TTL

The holdtime multiplier for transmit TTL is used to compute the actual TTL (time to live) value used in the LLDP frame. The TTL value is the time for how long the receiving device has to maintain the information in its Management Information Base (MIB). The default holdtime multiplier is four, and if required, this value can be changed between 2 and 10 with the **`lldp transmit-hold Value`** command as shown in Example 10-32.

Example 10-32   Holdtime multiplier

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp transmit-hold 2
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show lldp
LLDP transmit interval          : 40 seconds
LLDP transmit hold multiplier   : 2  (transmit TTL: 80 seconds)
LLDP transmit delay             : 5 seconds
LLDP SNMP notification interval  : 10 seconds
LLDP reinitialize delay         : 2 seconds
LLDP-MED fast start repeat count : 3

LLDP maximum neighbors          : 392
LLDP maximum neighbors per port  : 4
```

TTL is then calculated by multiplying the LLDP transmit interval by the holdtime multiplier. For example, if the LLDP transmit interval is 40 and the holdtime multiplier is 2, then the value 80 will be encoded as the TTL field in the LLDP header.

> **Note:** If transmit interval and/or transmit holdtime are set to inappropriate values, this can cause the LLDP agent to transmit information with TTL values that are excessively high, thus affecting how long a receiving device will retain the LLDP information if it is not refreshed.

10.2.10 Minimum time between port reinitialization

The minimum time between port reinitialization defines how many seconds the device will wait when LLDP is disabled on the port before it allows LLPD reenablement on the port. The default value is two seconds. This time can be changed with the command **`lldp reinit-delay Value`**, as shown in Example 10-33.

Example 10-33 Minimum time between port reinitialization

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp reinit-delay 8
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#show lldp
LLDP transmit interval          : 40 seconds
LLDP transmit hold multiplier   : 2  (transmit TTL: 80 seconds)
LLDP transmit delay             : 5 seconds
LLDP SNMP notification interval  : 10 seconds
LLDP reinitialize delay         : 8 seconds
LLDP-MED fast start repeat count : 3

LLDP maximum neighbors          : 392
LLDP maximum neighbors per port  : 4
```

The **`Value`** can range between 1 and 10 seconds.

10.2.11 Advertised LLDP TLVs

When LLDP is enabled, some of the TLVs (type-length-value) are advertised by default and some are not. You can change these with the following commands:

- To enable not advertised TLVs, use **`lldp advertise TLVName Parameters`**.
- To disable advertised TLVs, use **`no lldp advertise TLVName Parameters`**.

Table 10-2 lists the initial advertised status and the parameters that can be used in the **`lldp advertise`** command or the **`no lldp advertise`** command.

Table 10-2 Advertised status

TLV	Default advertise mode	TLVName in advertise command	Parameters in advertise command	Comments
Management address	advertised	N/A	N/A	IPv4 address which can be used to manage the device
Port description	advertised	port-description	ports **`PortList`**	N/A
System capabilities	advertised	system-capabilities	ports **`PortList`**	Possible values are: ▶ Repeater ▶ Bridge ▶ WLAN access point ▶ Router ▶ Telephone ▶ DOCSIS cable device ▶ Station only ▶ Other
System description	not advertised	system-description	ports **`PortList`**	Same as sysDescr MIB object in MIB.II.
System name	advertised	system-name	ports **`PortList`**	N/A

Examples of both commands are shown in Example 10-34.

Example 10-34 LLDP advertise examples

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp advertise port-description ports e 1/1/17
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp advertise port-description ports all
Edge_L2_B50G(config)#
```

802.1 capabilities

In Table 10-3 we show the 802.1 capabilities and the parameters that can be used.

Table 10-3 802.1 capabilities

TLV	Default advertise mode	TLVName in advertise command	Parameters in advertise command	Comments
VLAN name	not advertised	vlan-name	vlan **`VLANID`** ports **`PortList`**	N/A
Untagged VLAN ID	advertised	port-vlan-id	ports **`PortList`**	N/A

Examples of the commands for 802.1 capabilities are shown in Example 10-35.

Example 10-35 LLDP advertise examples - 802.1 capabilities

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp advertise vlan-name vlan 77 ports e 1/1/10 to 1/1/15
LLDP warning: port 1/1/10 is not a member of VLAN 77
LLDP warning: port 1/1/11 is not a member of VLAN 77
LLDP warning: port 1/1/12 is not a member of VLAN 77
LLDP warning: port 1/1/13 is not a member of VLAN 77
LLDP warning: port 1/1/14 is not a member of VLAN 77
LLDP warning: port 1/1/15 is not a member of VLAN 77
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp advertise vlan-name vlan 77 ports e 1/1/10 to 1/1/15
Edge_L2_B50G(config)#
```

802.3 capabilities

In Table 10-4 we show the 802.3 capabilities and the parameters that can be used in the command.

Table 10-4 802.3 capabilities

TLV	Default advertise mode	TLVName in advertise command	Parameters in advertise command	Comments
Link aggregation information	advertised	link-aggregation	ports **`PortList`**	N/A
MAC/PHY configuration and status	advertised	mac-phy-config-status	ports **`PortList`**	► The following information is included: ► Auto-negotiation capability and status ► Speed and duplex mode ► Flow control capabilities for auto-negotiation ► Port speed down-shift and maximum port speed advertisement
Maximum frame size	advertised	max-frame-size	ports **`PortList`**	► Advertise value depends if **`aggregated-vlan`** or **`jumbo`** commands are in effect
Power-via-MDI information	not advertised	power-via-mdi	ports **`PortList`**	The following information is included: ► POE capability ► POE status ► PSE power pair ► Power class

Examples of the commands are shown in Example 10-36.

Example 10-36 LLDP advertise examples - 802.3 capabilities

```
Edge_L2_B50G#
Edge_L2_B50G#config t
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#lldp advertise link-aggregation ports e 1/1/1
Edge_L2_B50G(config)#
Edge_L2_B50G(config)#no lldp advertise link-aggregation ports e 1/1/1
Edge_L2_B50G(config)#
```

In all cases, **PortList** ports can be listed individually, in a range using the **to** keyword to specify the range. By using the **all** keyword, all ports can be specified.

All advertised parameters can be verified locally with the command **show lldp local-info** as shown in Example 10-37.

Example 10-37 Local advertisement information

```
Edge_L2_B50G# show lldp local-info
Local port: 1/1/1
  + Chassis ID (MAC address): 001b.ed87.9b40
  + Port ID (MAC address): 001b.ed87.9b40
  + Time to live: 80 seconds
  + System name         : "Edge_L2_B50G"
  + Port description    : "GigabitEthernet1/1/1"
  + System capabilities : bridge
    Enabled capabilities: bridge
  + 802.3 MAC/PHY          : auto-negotiation enabled
    Advertised capabilities: 10BaseT-HD, 10BaseT-FD, 100BaseTX-HD,
                             100BaseTX-FD, 1000BaseT-HD, 1000BaseT-FD
    Operational MAU type   : 1000BaseSX-FD
  + Link aggregation: aggregated (aggregated port ifIndex: 1)
  + Maximum frame size: 1522 octets
  + Port VLAN ID: none
  + Management address (IPv4): 192.168.0.100

Local port: 1/1/2
  + Chassis ID (MAC address): 001b.ed87.9b40
....
Local port: 1/1/48
  + Chassis ID (MAC address): 001b.ed87.9b40
  + Port ID (MAC address): 001b.ed87.9b6f
  + Time to live: 80 seconds
  + System name         : "Edge_L2_B50G"
  + Port description    : "GigabitEthernet1/1/48"
  + System capabilities : bridge
    Enabled capabilities: bridge
  + 802.3 MAC/PHY          : auto-negotiation enabled
    Advertised capabilities: 10BaseT-HD, 10BaseT-FD, 100BaseTX-HD,
                             100BaseTX-FD, fdxSPause, fdxBPause,
1000BaseT-HD,
                             1000BaseT-FD
    Operational MAU type   : 1000BaseT-FD
  + Link aggregation: not capable
  + Maximum frame size: 1522 octets
  + Port VLAN ID: 20
  + Management address (IPv4): 192.168.0.100

Edge_L2_B50G#
```

11

Multicast

In this chapter, we describe the various multicast protocols supported on the IBM Data Center Networking b-type family of products and show you how to set up a multicast capable network.

We explore both Layer 2 multicast snooping as well as Layer 3 IPv4 multicast routing.

11.1 IP Multicast introduction

There are several ways (Figure 11-1) to establish communication and exchange traffic within a packet switching network:

- Unicast: One-to-one communication; a single sender sends information destined for a single receiver in the network. A receiver is a host interested in establishing communication with the sender.
- Anycast: One-to-many association, one-to-one communication; a single sender is aware of multiple receivers but selects only one of them to communicate with.
- Multicast: One-to-many communication; a single sender sends information to multiple receivers in the network.
- Broadcast: One-to-all communication; a single sender sends information destined for all hosts in the network. Some hosts might be receivers, while others might not. In Ethernet networks, this communication is limited to the broadcast domain.

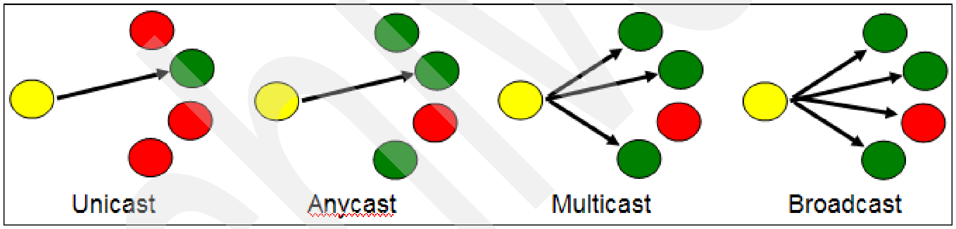


Figure 11-1 Illustration of different forwarding methods

Many applications can take advantage of efficient one-to-many communication especially if information is being delivered in a single direction. Some examples include video transmission (IP Television, surveillance cameras), distribution of real-time information (sports scores, stock quotes), online collaboration (video conferencing, remote learning), and file distribution.

Broadcasting such data is not efficient because there might be a relatively small number of hosts within a network that desires the information being sent. Broadcast traffic is also bounded within the local network or broadcast domain.

Using unicast transmission protocols to deliver these types of communications requires the source to open individual transmission connections with each receiver. The source also needs to be aware of each receiver wanting to subscribe to the information being sent. Figure 11-2 shows such a transmission on a small scale; you can see that the same traffic is duplicated in all network segments before reaching each receiver.

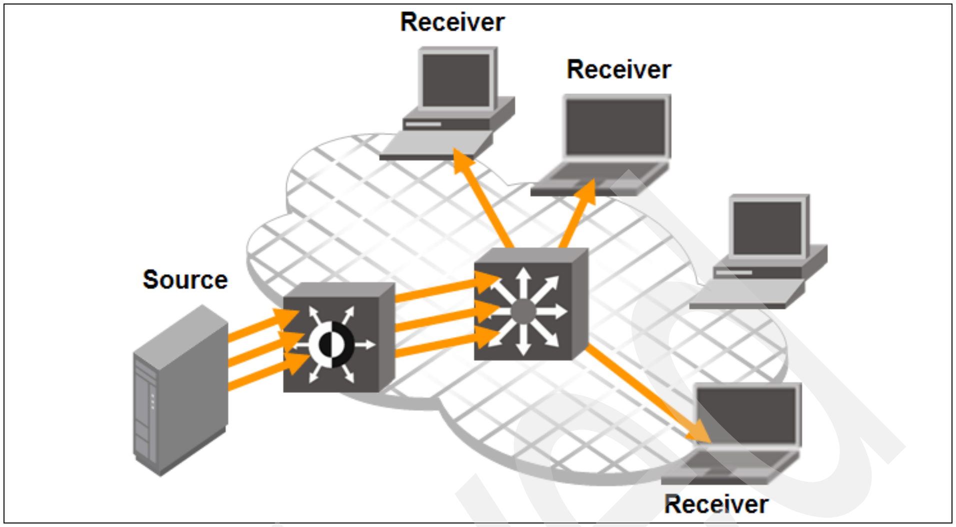


Figure 11-2 Unicast transmission - Duplicate data streams

With a network that supports IP multicast, only a single stream of data needs to be sent from the source into the network. The network will then forward only a single copy of the data stream towards the receiver. When the stream reaches the network device that one or more receivers are connected to, the device will make copies of the stream to each receiver (Figure 11-3).

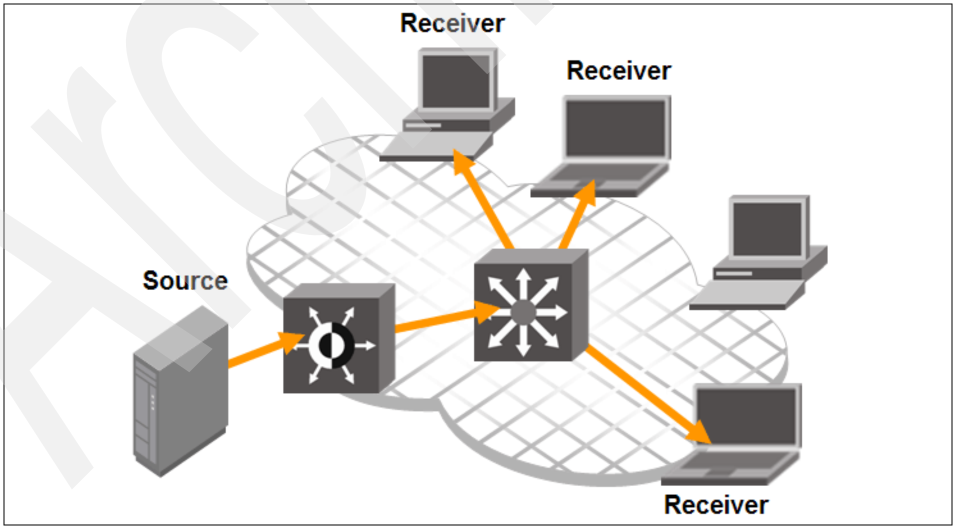


Figure 11-3 Multicast transmission - Efficient use of bandwidth

11.2 IP Multicast protocol

Here we describe key concepts of the IP Multicast:

- IP multicast group: The source and all hosts interested in the data stream being sent from that source (receivers) form an IP multicast group.
- IP multicast group address: This is the IP address that identifies an IP multicast group. Sources use this IP address as the destination IP of their IP packets.
- IP multicast group membership: For IPv4, the Internet Group Management Protocol (IGMP) is used to manage group membership. For IPv6, the Multicast Listener Discovery (MLD) protocol is used. These protocols are used by hosts to tell the router that they want to subscribe to an IP multicast group, thereby becoming a receiver. The router will maintain a membership group table. A router will also periodically query the receivers to make sure they still want to be a part of a multicast group.
- IP multicast routing: The Protocol Independent Multicast (PIM) protocol is used to route IP multicast packets. PIM does not discover routes but uses the unicast routing tables already on the router. PIM builds a multicast tree for each IP multicast group which is used to forward the multicast packets. Distance Vector Multicast Routing Protocol (DVMRP) is also supported, but not as widely used.
- Snooping: In a Layer 2 network, the default behavior for a switch is to broadcast unknown unicast and multicast packets. If IGMP or MLD snooping is enabled on a Layer 2 switch, the switch will monitor multicast group membership reports and maintains a forwarding table. It will then only forward multicast packets towards ports that have receivers in a multicast group.

11.2.1 Multicast IP Address Range (IPv4)

The IP address range of 224.0.0.0 - 239.255.255.255 is designated as the multicast address range by the Internet Assigned Numbers Authority (IANA), the group which controls the numbers for protocols and IP Address allotments globally. It is also defined in RFC 3171. Within the IP multicast address range:

- 224.0.0.0 - 224.0.0.255: Reserved for well-known multicast addresses. Some of these addresses are used for network protocols to communicate with each other and are not routed.
- 224.0.1.0 - 238.255.255.255: Globally scoped addresses for transmission between different Autonomous Systems (AS) and the Internet.
  - 232.0.0.0 - 232.255.255.255: Reserved for the Protocol Independent Multicast-Source Specific Multicast (PIM-SSM) protocol.

- 239.0.0.0 - 239.255.255.255: Administratively scoped or local multicast addresses fall into this range. These multicast addresses are for use within an AS and can be subdivided further using multicast boundaries.

11.2.2 Multicast MAC Address Range (Layer 2)

Multicast IP addresses are mapped directly to MAC addresses. A MAC address is comprised of 6 octets (48-bits) while an IP address is 4 octets (32-bits). The first three octets in a multicast MAC address is comprised an Organizationally Unique Identifier (OUI) assigned by the IANA, “01:00:5E”. The next bit in the multicast MAC address is always ‘0’. Following that, the lower 23 bits of the IP multicast address are used (Figure 11-4).

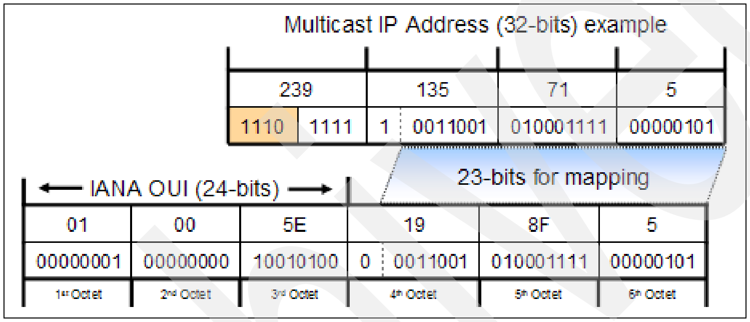


Figure 11-4 Multicast IP Address to MAC Mapping

Because the IP multicast address range is 224.0.0.0/4, the first four bits are always “1110”. Therefore a total of (32 - 23 - 4) = 5 bits are not used for mapping, so there is a possibility for up to 32 IP multicast groups to use the same MAC address. If there are multiple receivers in the same subnet that are subscribing to different multicast groups that have those characteristics, they will each receive multiple streams even though they might not have intended to join the other multicast groups. Use care in planning which IP multicast addresses to use.

11.2.3 Internet Group Management Protocol (IGMP)

There are three versions of IGMP:

- IGMPv1 (RFC 1112): Defines “join” message that hosts use to subscribe to a multicast group. Routers use a timer-based method to remove hosts from a multicast group.
- IGMPv2 (RFC 2236): Defines “leave group” message that hosts can use to un-subscribe to a multicast group. This allows for more timely updates of multicast group membership.

IGMPv1 and IGMPv2 are used to support PIM-DM (Dense Mode) and PIM-SM (Sparse Mode). IGMPv2 is backwards compatible with IGMPv1.

- IGMPv3 (RFC 3376): Provides the ability for hosts to define an “include” list of source addresses that they want to receive multicast traffic from and an “exclude” list that they do not want to receive multicast traffic from. This additional information allows for leaner and more secure multicast trees to be built.

IGMPv3 is used to support PIM-SSM (Source-Specific Multicast). IGMPv3 is backwards compatible with IGMPv2, however if you are using SSM then all devices must be configured to support IGMPv3.

11.2.4 Protocol Independent Multicast (PIM)

Here are commonly used terms when discussing multicast-capable routers:

- Multicast tree: A multicast tree is comprised of a Root Node and one or more nodes that are leaf or intermediate nodes.
- Node: Refers to a router in the multicast tree.
- Root Node: The node that initiates the tree building process and is also the router that sends the multicast packets down the multicast delivery tree. This is the Designated Router (DR) of the source subnet in a (S, G) tree, and the Rendezvous Point (RP) when using a shared (*, G) tree.
- Designated Router (DR): The router responsible in a subnet for managing the multicast routing protocols and multicast group tables. If multiple routers are connected to the source subnet an election process occurs to select a DR, typically the router with the highest IP address.
- Source-Group pair (S, G): “S” is the IP multicast source address. “G” is the IP multicast group address. A multicast tree is built for each (S, G).
- (*, G) pair: The “*“ represents all sources. The “G” represents the IP multicast group address. A shared multicast tree is built for all (*, G).
- Upstream: Represents the direction from which a router receives multicast packets. An upstream router is a node that sends multicast packets.
- Downstream: Represents the direction to which a router forwards multicast packets. A downstream router is a node that receives multicast packets from upstream transmissions.
- Leaf nodes: Routers that do not have any downstream routers.
- Intermediate nodes: Routers that are in the path between source routers and leaf routers.

- Group presence: Means that a multicast group has been learned from one of the directly connected interfaces. Members of the multicast group are present on the router.

Figure 11-5 shows the network topology that we use in our examples:

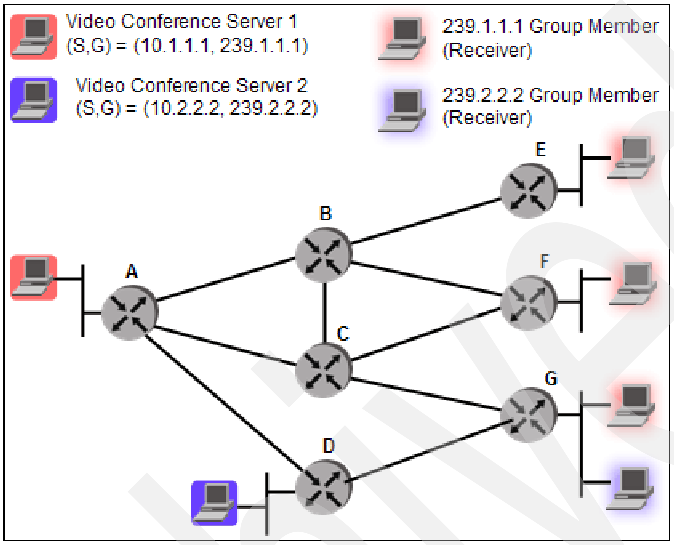

Video Conference Server 1
(S,G) = (10.1.1.1, 239.1.1.1)
Video Conference Server 2
(S,G) = (10.2.2.2, 239.2.2.2)
239.1.1.1 Group Member
(Receiver)
239.2.2.2 Group Member
(Receiver)
A
B
C
D
E
F
G

Figure 11-5 An example of a network topology with multicast sources and receivers

Within the host-receiver subnet, IGMP is used to send a message to the router to join a multicast group. A multicast group membership table is built and maintained by the multicast-capable router.

On the host-source side, the source just sends IP multicast packets addressed to the IP multicast group address. The multicast routing protocol will then build the multicast tree from the source towards the receivers in the network. How this tree is built depends on the type of PIM routing protocol being used.

The following configurations are supported:

- PIM-DM (Dense Mode):
  - Simplest of the PIM protocols to deploy at the cost of additional overhead.
  - Periodically floods all branches in the network and then prunes back branches that have no multicast group receivers.
  - Most effective in environments with receivers in many subnets.
  - Creates a separate multicast tree for each (Source, Group) pair, with the Designated Router (DR) of the source being the Root Node.

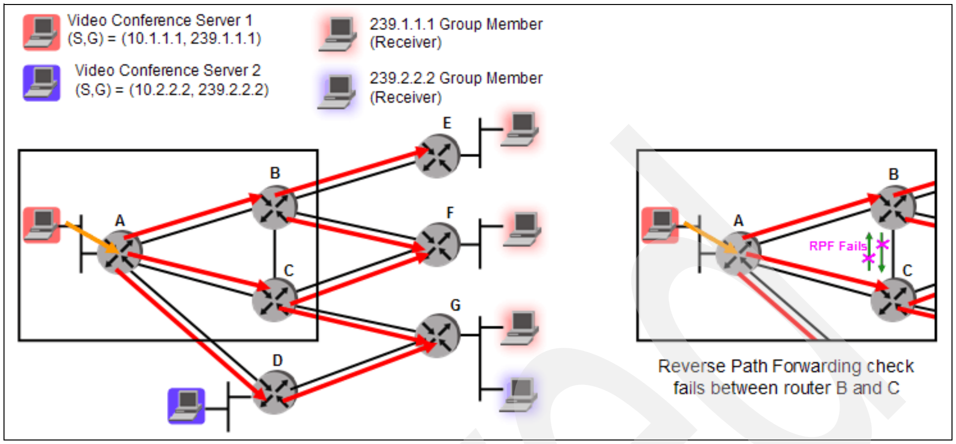
- PIM-SM (Sparse Mode):
  - Utilizes a Rendezvous Point (RP) which registers all sources and receivers in the network.
  - A shared multicast tree (*,G) is used with the RP as the Root Node. The shared tree is built with only explicit “join” requests from a receiver.
  - The Designated Router of the source will initially unicast all multicast traffic to RP. The RP will use its shared tree to distribute traffic.
    - Eventually a shortest path (S, G) tree is created and used for an established multicast group.
  - Makes efficient use of bandwidth because does not flood multicast traffic to uninterested receivers.
  - Much more scalable and becoming the most popular implementation but requires more design considerations.
- PIM-SSM (Source-Specific Mode):
  - Similar to PIM-SM, but instead of simply joining an IP multicast group the receiver uses IGMPv3 to also control from which sources they are interested in by building either an “include” list or an “exclude” list.
  - If multiple sources are in the same multicast group, helps increase security and reduce traffic by only subscribing to streams from specified sources.
  - Requires additional administrative overhead.

PIM-Dense Mode (PIM-DM)

While unicast forwarding uses the destination IP address to decide the next-hop, multicast forwarding uses the source IP address to decide the next-hop with a method called Reverse Path Forwarding (RPF).

When an IP multicast packet arrives at a router’s ingress interface it will check to see if there is a routing entry to the source IP address through that interface. If there is such an entry, this means that the packet came upstream, or from the direction of the Root Node. If there is no routing entry for the packet that the multicast packet did not come from the direction of the Root Node, and therefore was misdirected. The router will drop the packet and notify the neighboring router of the error.

Figure 11-6 shows the PIM-DM Multicast Tree using Reverse Path Forwarding.



Video Conference Server 1
(S,G) = (10.1.1.1, 239.1.1.1)
Video Conference Server 2
(S,G) = (10.2.2.2, 239.2.2.2)
239.1.1.1 Group Member
(Receiver)
239.2.2.2 Group Member
(Receiver)
A
B
C
D
E
F
G
RPF Fails
Reverse Path Forwarding check
fails between router B and C

Figure 11-6 PIM-DM Multicast Tree using Reverse Path Forwarding

In Figure 11-6, the multicast forwarding tree for the source at 10.1.1.1 is shown. Here is an example of how it was built utilizing RPF:

1. To initiate the multicast stream, the source will begin sending packets with a destination IP address of 239.1.1.1, and a source IP address of 10.1.1.1.
   – Router A is the Designated Router of the source subnet, so becomes the Root Node for this (S, G) group.
2. Router A begins sending the multicast packets all neighboring routers, Routers B, C, and D.
3. Routers B, C, and D will do a RPF check.
   – The routers will see if a routing entry exists to 10.1.1.0 (the network that 10.1.1.1 is in) through the port that Router A is connected to.
   – If we assume the Layer 3 network is configured correctly with a routing protocol such as OSPF or RIP and all paths are equal, the RPF check must pass. The router will make a note of this so they don't have to do a RPF check on all subsequent packets.
   – Because the RPF check passes, the packets are forwarded towards each router's neighbor, besides Router A where it came from
4. Taking a closer look at Router B, it will forward packets to Routers C and E.
   – On Router C, the entry in the routing table back to 10.1.1.0 must be through Router A (1 hop) instead of through Router B (2 hops), with link costs being equal. Therefore RPF fails and the packet is dropped and Router B notified.
   – The RPF check must pass on Router E.

5. Some routers such as Router F have two RPF neighbors, for example, Router G has both Router C and D upstream that passes RPF. The PIM protocol will optimize and drop one of the paths by doing a rate-limited prune.

Figure 11-7 shows what an optimized tree might look like at this stage.

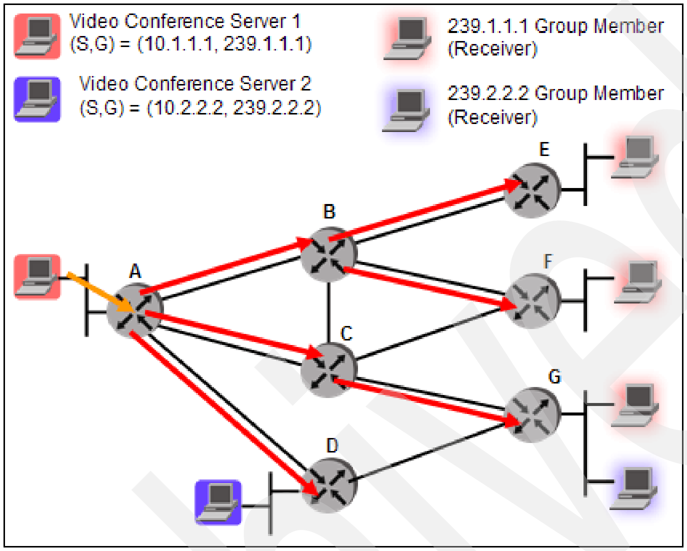

Video Conference Server 1
(S,G) = (10.1.1.1, 239.1.1.1)
Video Conference Server 2
(S,G) = (10.2.2.2, 239.2.2.2)
239.1.1.1 Group Member
(Receiver)
239.2.2.2 Group Member
(Receiver)
E
B
F
A
C
G
D

Figure 11-7 (S, G) tree with only one path to each downstream node

You can also see that Router D does not have any multicast group presence (receivers) for (10.1.1.1, 239.1.1.1) even though it is part of the tree and receiving traffic. With PIM-DM, this is called “flooding,” where all branches are sent multicast traffic periodically even if it is not requested. The branches that do not have any multicast group presence will send a “prune” message to its upstream router, telling the router it does not need multicast packets for that (S, G) group. If Router D ever receives a IGMP join message for that (S, G) group later, it can send a PIM “graft” message to rejoin that (S, G) tree.

Keep in mind that there is a separate multicast tree for each (S, G) pair.

Figure 11-8 is an illustration of what (10.2.2.2, 239.2.2.2)'s RPF tree might look like.

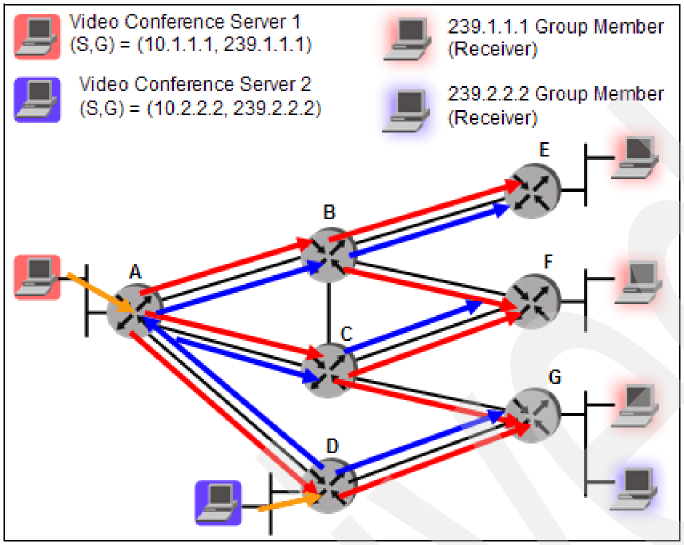

Video Conference Server 1
(S,G) = (10.1.1.1, 239.1.1.1)
Video Conference Server 2
(S,G) = (10.2.2.2, 239.2.2.2)
239.1.1.1 Group Member
(Receiver)
239.2.2.2 Group Member
(Receiver)
A
B
C
D
E
F
G

Figure 11-8 RPF tree

For (10.2.2.2, 239.2.2.2), the Root Node is Router D.

PIM-Sparse Mode (PIM-SM)

Instead of building a separate multicast tree for each (Source, Group) pair, PIM-Sparse Mode (PIM-SM) initially maintains a single shared tree, (*, G). This shared tree originates at a defined router in the network called the Rendezvous Point (RP) which is the Root Node of the shared tree. There can be multiple RPs configured for redundancy but only one can be active at a time within a multicast group.

RP information is distributed through a PIM-SM domain by a BootStrap Router (BSR). A PIM-SM domain is a contiguous set of routers that all implement PIM and are configured to operate within a common boundary. There can be multiple BSRs configured for redundancy but only one can be active at a time within a PIM-SM domain. It is preferable that the RP and BSR be the same router.

Another main difference between PIM-DM and PIM-SM is that the shared multicast tree is built and maintained from explicit "join" and "prune" from Designated Routers within the domain instead of flooding the network. The following is an example for a tree for a (*, G) group that has interested receivers but no source yet for the group.

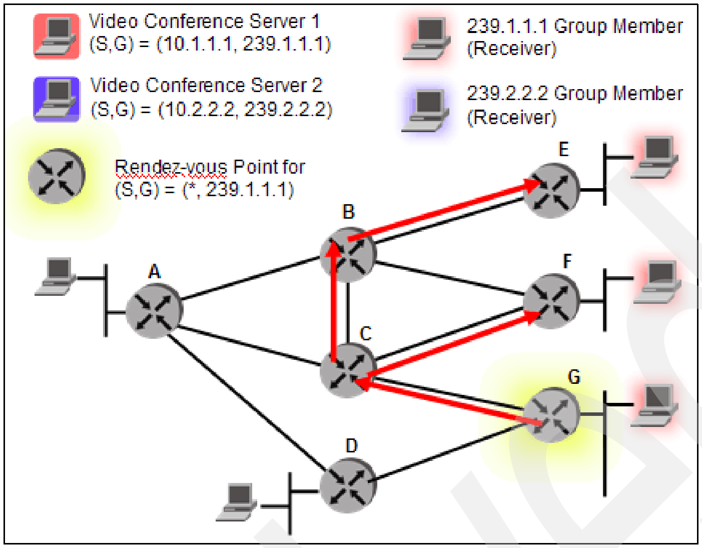


Figure 11-9 PIM-SM multicast tree with interested receivers but no source

Router G is the Rendezvous Point for (*, 239.1.1.1) and builds the tree based on explicit “join” and “prune” received from the various DRs in the network. Any source multicasting into the 239.1.1.1 uses this shared tree. There can be a different RPs for different multicast groups, but the simplest implementations uses one RP for all multicast groups, or (*, *).

When a source joins, it will send its multicast stream by unicast towards the RP. After it reaches the RP, data will then flow downstream the shared tree towards the receivers.

Note: On some platforms having a receiver on the path from the source to the Rendezvous Point might be restricted.

Figure 11-10 is an example of a source joining the network and its unicast path to the RP.

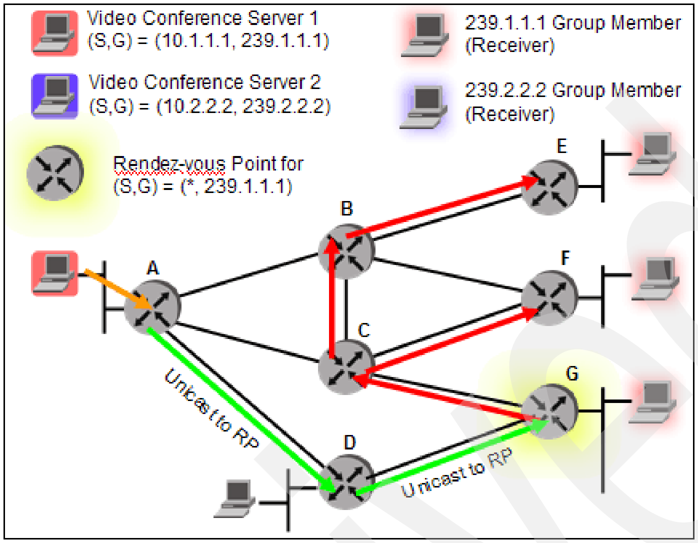


Figure 11-10 PIM-SM source multicast packets first goes to the RP, then downstream

As you can see, there are pros and cons with using a shared tree. Some of the pros are less operational overhead because only one tree needs to be maintained for a multicast group. However, this might lead to inefficient delivery of the multicast streams because the shared tree might not be the shortest path to the receivers.

Therefore a new (S, G), Shortest Path Tree (SPT) is built immediately after the first multicast packet reaches a receiver. The receiver will send a “join” message to the DR of the source, building a new tree. Branches will then start sending “Prune” messages to the RP after they are on the new (S, G) SPT, resulting in an optimized path. New receivers on the (S, G) SPT are able to have their DRs send “join” messages directly to the source DR instead of having to send to the RP to send back to the source DR.

Figure 11-11 shows the shortest path tree.

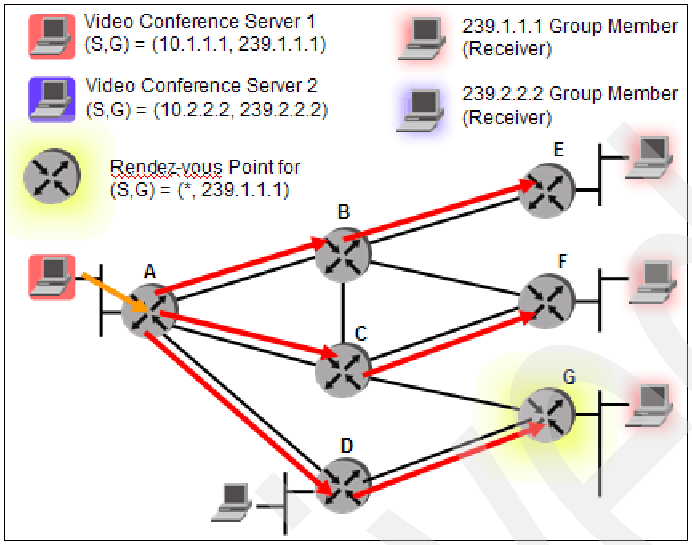


Figure 11-11 PIM-SM Shortest Path Tree

PIM-Source Specific Multicast (PIM-SSM)

PIM-Source Specific Multicast (PIM-SSM) is similar to PIM-Sparse Mode (PIM-SM) but also differs in the following methods:

- Allows a receiver to specify a specific list of sources to either “include” or “exclude” for the multicast group they are interested in
- Requires either IGMPv3 (IPv4) or MLDv2 (IPv6) which contains the source IP address information
- Builds a Shortest Path Tree (SPT) immediately without the need for a Rendezvous Point (RP) because the receivers know the source IP address and can unicast a “join” message directly back.

SSM also introduces the concept of channels, which consist of a single source and multiple receivers who specifically “subscribe” to receive broadcasts from that source or “unsubscribe” to stop. Therefore, there can be multiple multicast tree entries for a single (S, G).

PIM-SSM is given the reserved IP multicast address range of 232.0.0.0/8.

Related protocols

In the topics that follow we discuss some related protocols.

Multicast Source Discovery Protocol (MSDP)

This is used by PIM-SM routers to exchange source information across PIM-SM domains. Routers running MSDP can discover PIM-SM sources in other PIM-SM domains. MSDP typically use the Rendezvous Point routers in each PIM-SM domain as the MSDP peers to exchange information. Routers that run MSDP usually also run BGP, therefore it is preferable to use the same source address used by BGP as the MSDP source address.

Anycast RP (MSDP and PIM)

Anycast RP is a method for providing load balancing and redundancy within a PIM-SM domain.

With MSDP Anycast RP, all RPs within a domain are configured with the same Anycast RP address which is typically a loopback IP address. MSDP is used between all RPs in a mesh configuration to keep all RPs synchronized with the active sources.

PIM-SM routers are configured to register (statically or dynamically) with the RP using the Anycast RP address. Because multiple RPs have the same Anycast RP address, an Interior Gateway Protocol (IGP) such as OSPF routes the PIM-SM router to the RP with the best route. This helps keep the loads on the RPs distributed. If any RPs within the Anycast RP group goes down, the IGP will automatically re-calculate a new best route.

In PIM Anycast RP, the RP address of the Anycast RP is a shared IP address used among multiple PIM routers, known as PIM RP. The PIM RP routers create an Anycast RP set. Each of these routers are configured with two IP address; a shared RP address in their loopback address and a separate, unique IP address. The unique IP address is used for peering.

When a source is activated in a PIM Anycast RP domain, the PIM RP First Hop will register the packet and create the (S, G) state. If there are external peers in the Anycast RP set, the router will re-encapsulate and unicast it to all peers if needed to ensure consistent state synchronization.

Distance Vector Multicast Routing Protocol (DVMRP)

DVMRP is supported on the products, but is not as widely used as PIM. DVMRP uses IGMP to manage group membership and exchange information with other DVMRP routers. DVMRP uses flooding, pruning, and Reverse Path Forwarding (RPF) mechanisms to build a multicast group. We do not go into the details of DVMRP.

11.3 Multicast configuration

In Table 11-1 we show the features of multicast and the IBM series that they apply to.

Table 11-1 Multicast features

Multicast feature	Feature description	IBM m-series	IBM r-series	IBM s-series
Multicast (Layer 2 traffic reduction)	► Internet Group Management Protocol (IGMP) v1, v2, v3 snooping ► IGMP snooping per VLAN ► IGMP Fast Leave ► Protocol Independent Multicast-Sparse Mode (PIM-SM) v2 snooping	X	X	L2
Multicast (Layer 2 traffic reduction)	► MLD v1/v2 Snooping ► MLDv1 Fast Leave			L2
Multicast Routing (IPv4)	► PIM-Source Specific Mode (PIM-SSM) v4 ► Passive Multicast Route Insertion (PMRI)	X	X	
Multicast Routing (IPv4)	► IGMP v1, v2, v3 ► PIM-SM v1/v2 ► PIM-Dense Mode (PIM-DM) v2 ► Distance Vector Multicast Routing Protocol (DVMRP) ► PIM Anycast RP ► Multicast Source Discovery Protocol (MSDP)	X	X	FL3-IPv4
Multicast Routing (IPv6)	► Multicast Listener Discovery (MLD) v2 ► IPv6 PIM-SM	X	X	
Multicast Routing (IPv6)	► IPv6 PIM-SSM ► IPv6 PIM Anycast RP ► IPv6 PMRI	X		

Multicast features	Feature description	IBM c-series	IBM x-series	IBM B48G	IBM B50G
Multicast (Layer 2 traffic reduction)	► Internet Group Management Protocol (IGMP) v1, v2, v3 snooping ► IGMP snooping per VLAN ► IGMP Fast Leave ► Protocol Independent Multicast-Sparse Mode (PIM-SM) v2 snooping	Base	L2	L2	L2
Multicast (Layer 2 traffic reduction)	► MLD v1/v2 Snooping ► MLDv1 Fast Leave			L2	L2
Multicast Routing (IPv4)	► IGMP v1, v2, v3 ► PIM-SM v1/v2 ► PIM-Dense Mode (PIM-DM) v2 ► PIM-Source Specific Mode (PIM-SSM) v4 ► Distance Vector Multicast Routing Protocol (DVMRP) ► PIM Anycast RP ► Multicast Source Discovery Protocol (MSDP) ► Passive Multicast Route Insertion (PMRI)	FL3 Metro			

See Chapter 4, “Protocols and features” on page 103 for more information.

Next, we explore how to implement a multicast-enabled network based on the Protocol Independent Multicast (PIM) protocols.

11.3.1 Design considerations

Multicast routing and multicast snooping be configured on the same device. However, routing and snooping cannot be configured on the same routing interface or VLAN. This is because if a packet gets snooped, it will not be handled by the routing layer, and vice-versa. Typically you will want multicast routing on the routing interfaces (physical or VE), and snooping within a VLAN.

IGMP is enabled automatically whenever an interface is considered for PIM because it is a necessary requirement. IGMPv1, v2, and v3 are supported by all platforms and the default is IGMPv2. You must explicitly enable IGMPv3 either on a global or per-interface level.

Different multicast groups, interfaces, and routers can run their own versions of IGMP. There might be minor incompatibilities between IGMPv2 and IGMPv3. For example, an interface running IGMPv2 can recognize IGMPv3 packets but cannot process them. Also, a router running IGMPv3 can recognize and process IGMPv2 packets, but when that router sends queries to that interface, it will be in the IGMPv3 format so might not be recognized. So it is not a best practice to mix IGMPv2 and IGMPv3 within a multicast domain.

IGMPv3 is required on all devices if you want to implement PIM-SSM.

PIM-DMv1 and PIM-DMv2 are supported by all devices. The primary difference between the two version are:

- PIM-DMv1 uses IGMP to send messages
- PIM-DMv2 sends messages to the well-known multicast address 224.0.013 (ALL-PIM-ROUTERS) with protocol number 103. This is the default version.

The CLI commands for both PIM-DM versions are the same except for the command to enable each version. You cannot mix PIM-DM versions on devices that are connected to each other.

11.3.2 Configuring PIM-SM

Figure 11-12 illustrates the network on which we are enabling multicast in our examples, with the goal of supporting video streaming.

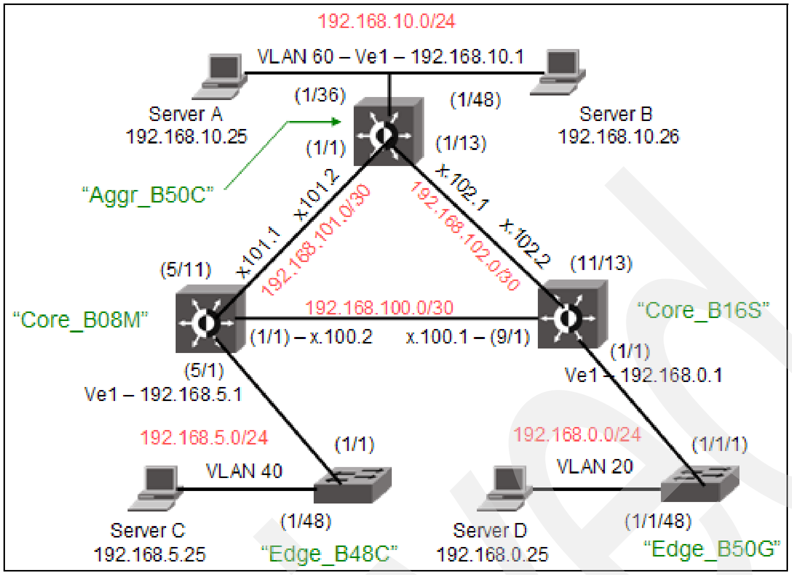


Figure 11-12 Network topology used in our examples

Our network is set up as illustrated in Figure 11-12:

- "Core_B08M" and "Core_B16S" are Layer 3 routing only.
- "Aggr_B50C" has routing interfaces as well as well as a VLAN.
- "Edge_B48C" and "Edge_B50G" are only operating on Layer 2 switching.
- OSPF is running between "Core_B08M", "Core_B16S", and "Aggr_B50C".

> **Note:** If you have a video streaming server on Server A and multicast it, all devices in VLAN 60 will be able to see it, but not past VLAN 60. This is because unknown unicast and multicast addresses are broadcasted. The packets are dropped at "Aggr_B50C" because there is no multicast routing protocols enabled yet.

We take the following steps:

1. Enable PIM on each router.
2. Enable PIM-SM on each routing interface.
3. Identify PIM-SM Bootstrap Router (BSR) candidate(s).
4. Identify PIM-SM Rendezvous Point (RP) candidate(s).
5. Enable IGMP snooping in each VLAN.
6. Enable PIM snooping for devices that support it.

Enabling PIM on each router

Use the **(config)#router pim command** as shown in Example 11-1 at the Global CONFIG level to enable PIM or DVMRP on each router. No reboot is required.

Example 11-1 Enabling multicast routing protocols on a router

```
telnet@Core_B08M(config)#router pim
  pim            Enable pim
telnet@Core_B08M(config)#router pim
telnet@Core_B08M(config-pim-router)#
```

We also used this command on “Core_B08S” and “Aggr_B50C”.

The **(config)#no router pim** command can be used to remove the PIM configuration.

> **Note:** On the s-series, you can use the **(config-pim-router)#disable-pim** command if you want keep the configuration but just want to disable PIM.

Enabling PIM-SM on each routing interface

Use the **(config-if-e10000-1/1)#ip pim-sparse** command as shown in Example 11-2 at the Interface CONFIG level to enable PIM-SM for that interface.

Example 11-2 Enabling PIM-SM for a routing interface

```
telnet@Core_B08M(config-if-e10000-1/1)#ip pim-sparse
  pim-sparse                Enable PIM Sparse mode on interface
telnet@Core_B08M(config-if-e10000-1/1)#ip pim-sparse
telnet@Core_B08M(config-if-e10000-1/1)#
```

We used this command on the following interfaces:

- “Core_B08M”: (1/1), (5/11), (ve1)
- “Core B16S”: (9/1), (11/13), (ve1)
- “Aggr_B50C”: (1/1), (1/13), (ve1)

Identifying PIM-SM bootstrap router (BSR) candidates

You must configure at least one interface on a router to be a BSR candidate. The BSR is responsible for distributing the Rendezvous Point (RP) information throughout the PIM-SM domain. The command syntax from the PIM-Router CONFIG level:

Syntax: `[no] bsr-candidate ethernet [<slotnum>/]<portnum> | loopback <num> | ve <num> <hash-mask-length> [<priority>]`

The BSR candidate can be configured either on a physical Ethernet interface, a loopback interface, or a virtual interface (ve).

The `<hash-mask-length>` parameter specifies the number of bits in a group address that are significant when calculating the group-to-RP mapping. You can specify a value from 1 – 32.

Preferred: Set the `<hash-mask-length>` value to ‘30’ for IPv4 networks.

The `<priority>` parameter specifies a BSR priority from 0 - 255. When the election process for BSR takes place, the candidate BSR with the highest priority becomes the BSR. The default is 0.

We configure a single BSR on the “Core_B08M” router on interface Ethernet 5/11 with the suggested `hash-mask-length` of 30 and a priority of 255, as shown in Example 11-3.

Example 11-3 Configuring a PIM-SM Bootstrap Router (BSR) candidate

```
telnet@Core_B08M(config-pim-router)#bsr-candidate
  bsr-candidate                 Set candidate bootstrap router
telnet@Core_B08M(config-pim-router)#bsr-candidate ethernet
  ethernet   Ethernet interface
telnet@Core_B08M(config-pim-router)#bsr-candidate ethernet 5/11
  SLOT/PORT
telnet@Core_B08M(config-pim-router)#bsr-candidate ethernet 5/11 30
  DECIMAL   Hash mask length (1 to 32)
telnet@Core_B08M(config-pim-router)#bsr-candidate ethernet 5/11 30 255
  DECIMAL   BSR priority (0 to 255)
  <cr>
telnet@Core_B08M(config-pim-router)#bsr-candidate ethernet 5/11 30 255
telnet@Core_B08M(config-pim-router)#
```

We configure another BSR on “Core_B16S” to add resiliency to the network, but with a lower priority, as shown in Example 11-4.

Example 11-4 Configuring another BSR candidate for resiliency

```
telnet@Core_B16S(config-pim-router)#bsr-candidate ethernet 11/13 30 200
BSR address: 192.168.102.2, hash mask length: 30, priority: 200
telnet@Core_B16S(config-pim-router)#
```

We can use the following **show** commands to see the current status of the BSR in the network as shown in Example 11-5 and Example 11-6.

Example 11-5 Examining BSR status from "Core_B08M"

```
telnet@Core_B08M(config-pim-router)#show ip pim bsr
PIMv2 Bootstrap information

This system is the elected Bootstrap Router (BSR)
  BSR address: 192.168.101.1
  Uptime: 10:26:54, BSR priority: 255, Hash mask length: 30
  Next bootstrap message in 00:00:20

Next Candidate-RP-advertisment in 00:00:40
  RP: 192.168.101.1
    group prefixes:
    224.0.0.0 / 4

  Candidate-RP-advertisement period: 60
telnet@Core_B08M(config-pim-router)#
```

Example 11-6 Examine BSR status from "Core_B16S"

```
telnet@Core_B16S(config-pim-router)#show ip pim bsr
PIMv2 Bootstrap information, state=CAND-BSR
  BSR address: 192.168.101.1
  BSR priority: 255, Hash mask length: 30

This system is a candidate BSR
  Candidate BSR address: 192.168.102.2
  Priority: 200, hash mask length: 30

Next Candidate-RP-advertisment in 00:00:30
  RP: 192.168.102.2
    group prefixes:
    224.0.0.0 / 4

  Candidate-RP-advertisement period: 60
telnet@Core_B16S(config-pim-router)#
```

Identifying PIM-SM rendezvous point (RP) candidates

You must configure at least one interface on a router to be a RP candidate. The RP maintains the IP multicast shared tree (*, G), receives all DR PIM-SM control messages, and is where multicast traffic from new sources first travel to.

Note: While it is not required that the BSR and RP be the same router, it is preferable to do so.

The command syntax from the PIM-Router CONFIG level:

Syntax: [no] rp-candidate ethernet [<slotnum>/]<portnum> | loopback <num> | ve <num>

The RP candidate can be configured either on a physical Ethernet interface, a loopback interface, or a virtual interface (ve).

Note: On the m-series, a GRE tunnel can also be configured to be a RP candidate.

We go ahead and configure our RP to be the same interface as our BSR as shown in Example 11-7.

Example 11-7 Configuring a PIM-SM Rendezvous Point (RP) candidate

```
telnet@Core_B08M(config-pim-router)#rp-candidate
  rp-candidate                  Configure candidate rendezvous point
(RP)
telnet@Core_B08M(config-pim-router)#rp-candidate
  add        Add group prefix to existing RP candidate
  delete     Delete group prefix from existing RP candidate
  ethernet   Specify physical interface for new RP candidate
  loopback   Specify loopback interface for new RP candidate
  pos        Specify physical interface for new RP candidate
  tunnel     Specify tunnel interface for new RP candidate
  ve         Specify virtual interface for new RP candidate
telnet@Core_B08M(config-pim-router)#rp-candidate ethernet
  SLOT/PORT
telnet@Core_B08M(config-pim-router)#rp-candidate ethernet 5/11
  <cr>
telnet@Core_B08M(config-pim-router)#rp-candidate ethernet 5/11
telnet@Core_B08M(config-pim-router)#
```

By default, this command configures the router as a candidate RP for all 224.0.0.0/4. As a result, the router is a candidate RP for all valid PIM-SM group numbers.

We can add another RP candidate in network for more resiliency, and as shown in Example 11-8, we configure another.

Example 11-8 Configuring another RP candidate for resiliency

```
telnet@Core_B16S(config-pim-router)#rp-candidate ethernet 11/13
telnet@Core_B16S(config-pim-router)#
```

You can change this by adding or deleting specific address ranges using the following commands:

Syntax: `[no] rp-candidate add <group-addr> <mask-bits>`

The `<group-addr> <mask-bits>` specifies the group address and the number of significant bits in the subnet mask.

The foregoing command replaces the default address range. You can also delete an address range from the range you added:

Syntax: `[no] rp-candidate delete <group-addr> <mask-bits>`

> **Note:** It is possible to set a static RP IP address, but this is not desirable. Multiple RP candidates can be configured to provide resiliency.

Up to now we have the following in the running-config for our router PIM settings as shown in Example 11-9 on page 348.

Example 11-9 running-config excerpt of router PIM settings

```
router pim
 bsr-candidate ethernet 5/11 30 255
 rp-candidate ethernet 5/11
 rp-candidate add 224.0.0.0 4
```

We also have PIM-SM enabled on all our routing interfaces.

Enabling IGMP snooping in each VLAN.

On all devices, IP Multicast Traffic Reduction using IGMP Snooping can be turned on at a Global-level. To enable IP Multicast Traffic Reduction with IGMP snooping, use the following command:

Syntax: `[no] ip multicast active | passive`

Active means that the device actively sends out IGMP queries to identify IP multicast groups on the network and makes entries in the IGMP table based on the Group Membership reports received from the network.

> **Note:** Routers in the network generally handle this operation. Use the active IGMP mode only when the device is in a stand-alone Layer 2 Switched network with no external IP multicast router attachments.

Passive means when passive IGMP mode is enabled, the device listens for IGMP Group Membership reports but does not send IGMP queries. The passive mode is sometimes called “IGMP snooping”.

The foregoing command will enable IGMP snooping on all VLANs.

> **Note:** IBM m-series and c-series only support the **`ip multicast`** command at the Global-level. All other devices that support IGMP snooping support configuration on a VLAN level as well.

We configure IGMP Snooping on “Aggr_B50C”, “Edge_B48C”, and “Edge_L2_B50G” as shown in Example 11-10.

Example 11-10 Turning on IGMP Snooping globally

```
Edge_L2_B50G(config)#ip multicast
  multicast                    Set IGMP snooping globally
Edge_L2_B50G(config)#ip multicast passive
  passive             IGMP snooping: device listens for IGMP packets
Edge_L2_B50G(config)#ip multicast passive
Edge_L2_B50G(config)#
```

Enabling PIM Snooping for devices that support it

IGMP Snooping allows Layer 2 devices to forward multicast traffic only towards receivers within that network that requested it. However, when multiple PIM-SM routers are connected through a Layer 2 device, the switch normally doesn’t look at the PIM information and always forwards it to the downstream routers even though they might not need it.

PIM-SM Snooping requires IGMP Snooping to be enabled and allows a Layer 2 device to listen to PIM-SM “join” and “prune” messages to make intelligent forwarding decisions.

Because there are no intermediate Layer 2 devices between our PIM-SM routers, we do not need to enable PIM Snooping. However, the commands to enable PIM Snooping are similar to those of IGMP Snooping:

`Syntax:` `[no] ip pimsm-snooping`

We can check the settings of IP multicast traffic reduction using the command as shown in Example 11-11.

Example 11-11   Reviewing the IP multicast traffic reduction settings

```
telnet@Edge_B48C(config)#show ip multicast
Global Multicast Traffic Reduction Configuration
   IGMP Snooping State:    Passive    Version             :         2
   Group Interval     :        260    Query Interval      :       125
   Max Response Time  :         10    Robustness Var      :         1
   Last Member Qry Int:          5    Last Member Qry Count:        3
   Querier Exp Tm     :        255
   IGMP Proxy         :   Disabled    Proxy Interval      : N/A
   Filter             :   Disabled    Tracking            :   Disabled

   PIM Snooping       :   Disabled
   PIM Prune Wait Time:          3
   PIM Proxy          :   Disabled    Proxy Interval      :        60
VLAN snooping configurations:

VLAN ID 1
 Inherits Global Configurations. Entries 0
VLAN ID 40
 Inherits Global Configurations. Entries 0

telnet@Edge_B48C(config)#
```

11.4 Conclusion

With multicast now enabled in the network, we can try streaming a video across the network. Using a streamer such as the free VideoLAN Layer (VLC), we try broadcasting from Server A to a multicast IP address, 239.1.2.3 and we try viewing it across the network from Server D. The results are shown in Figure 11-13.

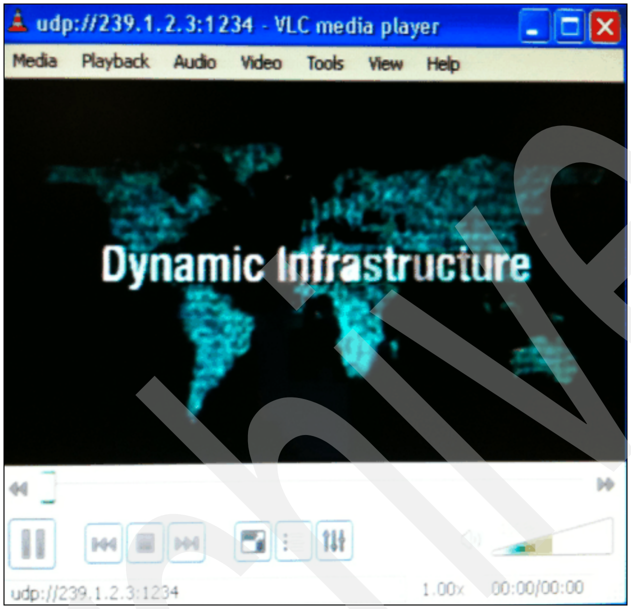


Figure 11-13 Video streaming from 239.1.2.3

With multicast streaming, we can significantly reduce the load on our networks while simultaneously enabling richer content.

12

High availability network functions

In this chapter, we review the functions provided by the products to enable high availability networks:

- Virtual Router Redundancy Protocol (VRRP) and VRRP Extended (VRRPE)
- Virtual Switch Redundancy Protocol (VSRP)
- Metro Ring Protocol (MRP)
- Protected Link Groups

12.1 Virtual Router Redundancy Protocol (VRRP)

VRRP is available on devices running IronWare Base or Full Layer 3 images and devices running Multi-Service IronWare.

VRRP is a way for one router to take over for another in the event of a failure. When one router fails, a backup takes over the first router's IP address.

Within each router running VRRP, you create one or more Virtual Routers, each of which you assign a Virtual Router ID (VRID). Within a particular network, multiple routers can have the same VRID, which forms a group. Each router in the VRID group must have a different IP address, but from these, a single IP address is selected to be the VRID IP address. A Master router is elected based on assigned priority while all other routers act as backup.

Within a router, you can also have VRRP monitor the interfaces through the router uplink (called a track port). Multiple interfaces on a router can be assigned to monitor the main uplink and take over ownership of the IP address if the uplink goes down.

Each router has a priority that you designate. This way, several routers can act as backup for the same owner. The owner has a priority of 255, and each backup some priority less than that.

12.1.1 Configuration steps

To configure VRRP, perform the following steps:

1. Run the config level command **router vrrp** to enable the protocol on the primary router.
2. Enter the Interface CONFIG level for the interface that will own the IP address.
3. Assign the IP address to the interface (example: **ip address** ***1.2.3.1/24***).
4. Enter the command **ip vrrp vrid** ***1*** (assigns this to VRID group 1).
5. Designate the router as the owner with the **owner** command.
6. Enter the IP address for the VRID with **ip-address** ***1.2.3.1*** (selects the VRID IP from one of the IPs in the VRID group).
7. If desired, designate a track port with **track-port ethernet** ***5/1***. Set a "track priority" using **owner track-priority** ***20***. (Use a priority less than the backup priority of your backup routers.)
8. Enter the command **activate**.
9. On the backup router, run **router vrrp**, and then enter the interface config level for the appropriate interface.

10. Assign an IP address to the interface. (ex. **ip address *1.2.3.2/24***)
    Important: *Do not* assign the VRID's IP address to the interface. This produces an address conflict.
11. Enter the command **ip vrrp vrid *1***.
12. Designate the router as a backup with the **backup priority *100*** command.
13. Enter the IP address for the VRID with **ip-address *1.2.3.1***.
14. If desired, designate a track port with **track-port ethernet *5/1***. Set a "track priority" using **backup priority 100 track-priority *19***. (Use a priority less than the backup priority of your other backup routers.)
15. Enter the command **activate**.

We step through the configuration of the owner router in Example 12-1 on page 355 and the backup router in Example 12-2 on page 355.

Example 12-1 VRRP owner configuration

```
telnet@m_Series(config)#router vrrp
telnet@m_Series(config)#interface ethernet 5/1
telnet@m_Series(config-if-e1000-5/1)#ip address 1.2.3.1/24
telnet@m_Series(config-if-e1000-5/1)#ip vrrp vrid 1
telnet@m_Series(config-if-e1000-5/1-vrid-1)#owner
telnet@m_Series(config-if-e1000-5/1-vrid-1)#ip-address 1.2.3.1
telnet@m_Series(config-if-e1000-5/1-vrid-1)#track-port ethernet 5/2
telnet@m_Series(config-if-e1000-5/1-vrid-1)#owner track-priority 20
telnet@m_Series(config-if-e1000-5/1-vrid-1)#activate
```

Example 12-2 VRRP backup configuration

```
BR-FGS648P-STK Router(config)#
BR-FGS648P-STK Router(config)#router vrrp
BR-FGS648P-STK Router(config)#interface ethernet 1/1/2
BR-FGS648P-STK Router(config-if-e1000-1/1/2)#ip address 1.2.3.2/24
BR-FGS648P-STK Router(config-if-e1000-1/1/2)#ip vrrp vrid 1
BR-FGS648P-STK Router(config-if-e1000-1/1/2-vrid-1)#backup priority 100
BR-FGS648P-STK Router(config-if-e1000-1/1/2-vrid-1)#ip-address 1.2.3.1
BR-FGS648P-STK Router(config-if-e1000-1/1/2-vrid-1)#track-port ethernet
1/1/3
BR-FGS648P-STK Router(config-if-e1000-1/1/2-vrid-1)#backup priority 100
track-priority 19
BR-FGS648P-STK Router(config-if-e1000-1/1/2-vrid-1)#activate
VRRP router 1 for this interface is activating
BR-FGS648P-STK Router(config-if-e1000-1/1/2)#
```

12.1.2 VRRP Extended (VRRPe)

> **No**te: VRRP Extended (VRRPe) is not available on the IBM g-series.

VRRPe is available on devices running IronWare Full Layer 3 images and devices running Multi-Service IronWare.

VRRPe is a proprietary protocol that makes some minor changes to base VRRP (it is not available on the g-series products):

- IP addresses are not assigned to the router interfaces running VRRPe. Instead, a VRID IP address is assigned for all routing interfaces in the VRID group.
- You can ping the VRID's IP.
- All routers are "backup" with the current Owner of the address determined by priority.
- The track-port priority *reduces* the router priority by the track-port priority value, instead of making it *equal to* the track-port priority.
- In the event of a priority "tie", the router with the highest IP address wins.

VRRPe has a configuration parameter called "slow start" that reduces "thrashing" in the event of an unstable interface. If you enable this option, a router will wait for a period of time before re-taking ownership of the VRID.

To configure VRRPe on each router, perform the following steps:

1. Run the command **router vrrp-extended**.
2. If desired, run the command **slow-start *30***, where the 30 is the desired delay, in seconds.
3. Enter the interface config level for the interface you want to configure VRRPe on.
4. Assign an IP address to the interface. This *cannot* be the same IP address you will use for the VRID.
5. Enter the command **ip vrrp-extended vrid *1***.
6. Enter the IP address for the VRID with the command **ip-address *1.2.3.4***.
7. Designate a track port with **track-port ethernet *5/1***.
8. Set the backup priority and the track priority reduction amount with **backup priority *160* track-priority *100***. All routers receive a "backup priority", even the router you intend to be the primary router.
9. Run the command **activate**.

We demonstrate the configuration process in Example 12-3.

Example 12-3 VRRPe example

```
telnet@m_Series(config)#router vrrp-extended
telnet@m_Series(config-vrrpe-router)#slow-start 30
telnet@m_Series(config-vrrpe-router)#interface ethernet 5/1
telnet@m_Series(config-if-e1000-5/1)#ip address 3.4.5.2/24
telnet@m_Series(config-if-e1000-5/1)#ip vrrp-extended vrid 1
telnet@m_Series(config-if-e1000-5/1-vrid-1)#backup priority 100
track-priority 20
telnet@m_Series(config-if-e1000-5/1-vrid-1)#ip-address 3.4.5.1
telnet@m_Series(config-if-e1000-5/1-vrid-1)#track-port ethernet 5/2
telnet@m_Series(config-if-e1000-5/1-vrid-1)#activate
telnet@m_Series(config-if-e1000-5/1)#
```

12.2 Virtual Switch Redundancy Protocol (VSRP)

Note: VSRP is not available on the IBM x-series.

VSRP is a proprietary protocol built on top of VRRPe, and provides similar function for Layer 2 forwarding, offering potentially faster failover than STP or RSTP. VSRP provides both Layer 2 and Layer 3 redundancy.

Unlike VRRPe, which will interoperate with any downstream switch, VSRP downstream devices must be “VSRP aware.” All IBM b-type DCN devices are VSRP aware.

VSRP is enabled on a per-VLAN basis.

Configuring VSRP is very easy. Simply follow these steps on each:

1. Enter the VLAN config level where you want to set up VSRP.
2. Enter the command **vsrp vrid *1***.
3. Set the priority and track priority with **backup priority *100*** (The backup priority is used for the election of a master, with the highest priority winning. In the case of a tie, the backup with the highest IP address wins)
4. Set the track port with **track-port ethernet *5/2***. (For VSRP, this is an interface on the router that is outside the VRID group but is tracked by VRID. This is typically the outbound port. If this port goes down, the track-port priority is subtracted from the VRID router priority, which might cause a re-negotiation of the Master)

5. Configure Layer 3 redundancy with **ip-address 1.2.3.4.** (VSRP Layer 3 redundancy works like VRRPe, so a single VRID IP address is used)
6. Start VSRP with **enable**.

We demonstrate the steps in Example 12-4.

Example 12-4 VSRP setup

```
BR-FGS648P-STK Router(config)#vlan 2
BR-FGS648P-STK Router(config-vlan-2)#vsrp vrid 1
BR-FGS648P-STK Router(config-vlan-2-vrid-1)#backup priority 100
track-priority 20
BR-FGS648P-STK Router(config-vlan-2-vrid-1)#track-port ethernet 1/1/5
BR-FGS648P-STK Router(config-vlan-2-vrid-1)#ip-address 1.2.3.4
BR-FGS648P-STK Router(config-vlan-2-vrid-1)#enable
VSRP vrid 1 for this VLAN is activating
BR-FGS648P-STK Router(config-vlan-2)#
```

12.3 Metro Ring Protocol (MRP)

Metro Ring Protocol (MRP) is a protocol that quickly restores connectivity in the event of a link outage. It does this faster than RSTP. As the name implies, it is used with ring-shaped topologies, as might be used in a network spanning a metropolitan area. It can be combined with VSRP within a data center to provide fast Layer 2 network resiliency across a large network.

Multiple MRP rings can share both nodes and interfaces. MRP rings are built on top of VLANs.

MRP has the concept of a *master node*. This is the node that controls the ring, and monitors it for link failures.

The details of how MRP works are beyond the scope of this book; if you want to know more about the protocol, see *IBM b-type Data Center Networking - Design and Best Practices Introduction*, SG24-7786.

Implementation: To implement MRP, carry out the following steps
(the first step is performed only on one node, the rest apply to all the nodes):

1. On one node, disable one of the ring interfaces; this prevents a loop from forming.
2. Enter the VLAN config level for the VLAN that the ring will be built on.
3. Enter the command **metro-ring *1***.

4. Designate two interfaces that provide connectivity to the two “legs” of the ring with the command **ring-interface ethernet *5/1* ethernet *5/2***.
5. If this is the master node, enter the command **master**.
6. Enter the command **enable** to start MRP.

We demonstrate this process in Example 12-5.

Example 12-5 MPR implementation

```
telnet@NetIron MLX-16 Router(config)#
telnet@NetIron MLX-16 Router(config)#interface ethernet 5/1
telnet@NetIron MLX-16 Router(config-if-e1000-5/1)#disable
telnet@NetIron MLX-16 Router(config-if-e1000-5/1)#exit
telnet@NetIron MLX-16 Router(config)#vlan 123
telnet@NetIron MLX-16 Router(config-vlan-123)#tagged ethernet 5/1 to
5/4
telnet@NetIron MLX-16 Router(config-vlan-123)#metro-ring 1
telnet@NetIron MLX-16 Router(config-vlan-123-mrp-1)#ring-interface
ethernet 5/1 ethernet 5/2
telnet@NetIron MLX-16 Router(config-vlan-123-mrp-1)#master
telnet@NetIron MLX-16 Router(config-vlan-123-mrp-1)#enable
telnet@NetIron MLX-16 Router(config-vlan-123-mrp-1)#
```

12.4 Uni-directional link detection (ULD)

It is possible to have an ethernet link between two switches/routers pass through another device such as a DWDM. It is possible for such devices to fail in such a way that traffic only flows in one direction. When this occurs, the link might still appear as “up” in each switch/router, but not actually work. Unidirectional Link Detection (ULD) prevents such a state by sending keep-alive packets. If the packets are not received, the link will be shut down. (Then other recovery, such as MRP or STP reconvergence, can take place.)

To configure ULD on a particular interface, enter the config-level command **link-keepalive ethernet *1/1/1***.

Tip: If using MRP, the *Configuration Guide* suggests setting the pre-forwarding time higher than the default of 300 ms (such as 400 ms or 500 ms.) You can do this with the MRP-level config command **preforwarding-time *400***.

If enabling ULD on a trunk or LAG, it must be done individually for each physical port.

12.5 Protected Link Groups

Protected Link Groups are available on the IBM s-series and IBM g-series.

Protected Link Groups are a way of designated backups for the failure of a given link or trunk. They are different from trunks in that only one link in a link group is active at a given time. Also, the various links in a Protected Link Group can be of different types. For instance, a group can contain a 10GbE link and a 6-link 1GbE trunk.

To create a Protected Link Group, enter the config-level command `protected-link-group 1 ethernet 5/1 ethernet 7/1`. You can put as many interfaces as you like in the group.

Next, designate an active port. This is the port that will be used by default for traffic within the group. This is set with the command `protected-link-group 1 active-port ethernet 5/1`.

13

Manageability and monitoring

In this chapter, we review the various methods that you can use to manage the products, monitor their status, observe their traffic, and manage users.

13.1 SNMP

SNMP is a common way to manage network devices. In addition, SNMP traps send out notifications of events that have occurred that might require attention.

13.1.1 SNMP trap receiver

To set a destination for SNMP traps, use the config level command ***snmp-server host 1.2.3.4 version <v1 | v2c | v3> <community string> port <UDP port #>***

Some trap receivers might be configured to only accept traps from certain IP addresses. To make sure that all traps originate from a particular address, use **snmp-server trap-source *management 1***. This will send all traps from the first IP address configured on **management 1**. You can use any interface (management, loopback, router, ethernet, SONET, and so on) on the product as a trap source, for example **snmp-server trap-source *ethernet 1/1.***

13.1.2 SNMP community string

SNMP v1 and v2 control access by the SNMP community string. On the s-series and g-series, there is a default read-only string **public**. It is generally a best practice to change eliminate this; do so with the config level command **no snmp-server community public ro**.

To add a SNMP community string, use the command **snmp-server community *community_string* ro** to create a read-only community string, or **snmp-server community *community_string* rw** to create and read-write string.

To apply an ACL (standard ACL-only) to requests from that community, append the ACL number to the end of the command.

13.2 Securing administrator access

The default user/security configuration is neither very flexible nor secure.
As stated before, the default security configurations are not suitable for most environments. By default, there is no concept of a user account; anyone with serial, telnet, or SSH access can at the least run any **show** command they desire, which might present a security risk.

13.2.1 Initial access method securing

In this section we discuss topics related to access security.

Telnet server

By default, telnet is enabled on the s-series, x-series, and g-series; on the m-series, c-series, and r-series Telnet is disabled by default. To disable Telnet access, use the config level command **`no telnet server`**. If you do choose to leave telnet access available, it is preferable to create a telnet password with the config level command **`enable telnet password`** ***`yourPasswordHere`***.

SSH

SSH is far more secure than telnet, and is the method we prefer. To enable the SSH server, generate an encryption key with the config level command **`crypto key generate`**. This action will also enable the SSH server.

Serial access

While you can secure access to **`enable`** and **`config`** mode, you cannot prevent a user from accessing the base user prompt by the serial port. The serial port can also be used to reset the super-user password. For this reason it is *vital* that you secure physical access to the serial port.

Web server

Very few administrators will use the web GUI to configure their device. To disable the web server, use the config level command **`no web-management`**.

13.2.2 Basic enable password configuration

If you will only have a single administrator, you do not necessarily have to configure any users at all; you can just assign single passwords to different access levels for the **`enable`** command.

Read-only password

To create an “enable” password for users to enter the Privileged EXEC level that will not be able to enter the Global CONFIG level, assign an “enable” read-only password using the Config level command **`enable read-only password`** ***`yourPassWordHere`***.

Port configuration

To create an **`enable`** password for users to only modify port level configurations but not modify any Global configuration parameters, use the Config level command **`enable port-config-password`** ***`yourPasswordHere`***.

Super-user password

As the name implies, the user with this password will have access to any command available on the product. This password is set with **`enable super-user-password`** ***`yourPasswordHere`***.

13.2.3 Creating local user accounts

Your user accounts might be stored as part of the configuration, or they might be obtained remotely from a RADIUS, TACACS, or TACACS+ server. In this section we discuss creating local accounts.

To create local user accounts, use the config level command **`username`** ***`yourUserNameHere`*** **`privilege`** ***`0`*** **`password`** ***`usersPassWordHere`***. The **`privilege`** parameter (0 in our example) can have one of the following values: 0=Super-User, 4=Port-Config-Only, 5=Read-Only.

Before you can create user accounts with lower access levels, you must set a global super-user password, or create a super-user account.

13.2.4 Applying user account security

To actually use the user accounts (either local or remote), you need to use the **`aaa`** commands.

If you are using user accounts, you probably do not need to set the super-user/port-config/read-only passwords.

Telnet/SSH protection

The basic **`aaa`** command you are likely to configure is the one for Telnet/SSH access. This is **`aaa authenticating login default`** ***`local`***. This command tells the product to use the local username/password list (that you defined in section 13.2.3, “Creating local user accounts”.

If you like, you can specify a list of access methods, in the order you want to try them. Your choices are as follows:

- `local`: Use the locally-defined usernames and passwords.
- `TACACS`, `TACACS+`, or `RADIUS`: Use TACACS, TACACS+, or a RADIUS server for authentication. (see 13.2.6, “Configuring remote access servers” on page 366).
- `line`: Use the `telnet` password defined with the `enable telnet password` ***yourPassWordHere*** command.
- `enable:` Use the passwords set with `enable super-user-password / port-config-password / read-only-password` ***yourPassWordHere*** commands.

If at *any* level, you are rejected, the login will fail. For example, if you enter in a username valid on RADIUS, but the wrong password, you will be immediately rejected. The system will not even try authentication methods further down in your list.

Automatic privilege level setting: If you want, you can automatically put users in the privilege level they have been authorized for, enabling them to bypass the `enable` command. Set this option with the `aaa authentication login privilege-mode` command. (Note: This option will not work with methods that do not allow you to specify a privilege level, such as plain TACACS).

Serial port authentication

If required you can also configure user authentication for the serial port; however, if you are using remote authentication, and the remote-authentication server goes down, you will be unable to access your product without a reboot and password bypass (see 13.2.5, “Bypassing password authentication” for details). The command to enable this is `enable aaa console`.

To “log out” of the serial port (which then requires the password on the next use), simply enter the command `exit` from the user level prompt. You will then see a username prompt. You can also set a console time-out with `console time-out` ***10***. The “10” is the number of minutes of inactivity before the logout takes place.

You can also force a logout of all console ports by issuing the `kill console all` command. The `kill` command can also be used to force logout of Telnet and SSH sessions as well using `kill <console | telnet | ssh> <session # | all>.`

13.2.5 Bypassing password authentication

If you have forgotten your passwords, or are unable to access your user accounts, you can bypass password authentication with the serial port. Because of this, it is best to secure physical access to the device.

To bypass device authentication, follow these steps:

1. Connect to the product with the serial cable and reboot the product.
2. Keep pressing the “b” key to start the Boot Monitor.
3. Enter the command **no password**. (The command cannot be abbreviated)
4. Enter the command **boot system flash** ***primary***. (or ***secondary***) Again, you cannot abbreviate.
5. From there, go to the **enable** and **config** prompts as usual and fix your authenticating problem. (Set a new password, fix remote authentication servers, and so on.)

13.2.6 Configuring remote access servers

Before you can actually use remote access servers, you must tell the product where they are.

TACACS/TACACS+

To define a TACACS or TACACS+ server, enter the config level command **tacacs-server host** ***1.2.3.4***. You can define up to eight, simply by running the command multiple times. (They will be used in the order you define them.)

If you are using TACACS+, and your server configuration requires a key, use the command **tacacs-server host** ***1.2.3.4*** **key** ***yourKeyHere***.

TACACS+ Authorization

To get a user’s privilege level from a TACACS+ server, use the command **aaa authorization exec default tacacs+**.

After you set this, you will need to set up an Attribute-Value pair on your TACACS+ server specifying the user’s privilege level. This A-V pair is the **foundry-privlvl**. This must be set to either 0 (super-user), 4 (port-config), or 5 (read-only).

The following sample coding is given in the manual for a user entry:

```
user=bob {
   default service = permit
   member admin
   #Global password
   global = cleartext "cat"
   service = exec {
      privlvl = 15
         }
}
```

You can also use the authorization functions to restrict certain users to certain commands. For more information about this function, see the *Configuration Guide*.

TACACS+ Accounting

To record logins/logouts on a TACACS+ server (not available with TACACS), enter the command **`aaa accounting exec default start-stop tacacs+`**.

To record all commands entered by users, enter the command **`aaa accounting commands 0 default start-stop tacacs+`**.

RADIUS

Technically, RADIUS stands for Remote Dial-In User Service, but it is now used for much more, such as MAC authentication, dynamic VLAN assignment, and so on. In this case, we are configuring RADIUS to authenticate your users, which is close to its original purpose.

To define a RADIUS server to the product, use the command **`radius-server host`** *`1.2.3.4`*. You can define up to eight, simply by running the command multiple times. (They will be used in the order you define them.)

RADIUS authorization: To get a user’s privilege level from a RADIUS server, use the command **`aaa authorization exec default RADIUS`**.

After you set this, you will need to set up some proprietary attributes on your RADIUS server if you want to use the authorization functions. Use Vendor-ID 1991, Vendor Type 1. The attributes can be found in Table 13-1. You *must* use all three attributes, even if you do not want to restrict command usage. To allow the user to use all commands, set the **`foundry-command-string`** attribute to **`*`**, and set the **`foundry-command-exception-flag`** attribute to **`0`**.

Table 13-1 Vendor Specific Attributes

Attribute name	Attribute ID	Data type	Description
foundry-privilege-level	1	integer	Specifies the privilege level for the user. This attribute can be set to one of the following values: • **0** - Super User level: Allows complete read-and-write access to the system. This is generally for system administrators and is the only management privilege level that allows you to configure passwords. • **4** - Port Configuration level: Allows read-and-write access for specific ports but not for global (system-wide) parameters. • 5 - Read Only level: Allows access to the Privileged EXEC mode and CONFIG mode of the CLI but only with read access.
foundry-command-string	2	string	Specifies a list of CLI commands that are permitted or denied to the user when RADIUS authorization is configured. The commands are delimited by semicolons (;). You can specify an asterisk (*) as a wildcard at the end of a command string. For example, the following command list specifies all show and debug ip commands, as well as the write terminal command: show *; debug ip *; write term*
foundry-command-exception-flag	3	integer	Specifies whether the commands indicated by the foundry-command-string attribute are permitted or denied to the user. This attribute can be set to one of the following values: • **0** - Permit execution of the commands indicated by foundry-command-string, deny all other commands. • **1** - Deny execution of the commands indicated by foundry-command-string, permit all other commands.

13.3 Port Mirroring

Port Mirroring is a commonly-used network troubleshooting feature. With this feature, you can copy traffic either to or from a particular port (called a “monitored port” to another port (called a “mirror” port.) Both inbound (ingress) and outbound (egress) traffic can be monitored.

13.3.1 Configuration restrictions

There are certain restrictions on port mirroring configurations, which vary by product model. Some restrictions, however, are universal:

- Trunk/LAG ports cannot be mirror ports, but they can be monitored.
- Only one mirror port can be used per mirror port for egress monitoring.
- Only one mirror port can be used per mirror port for ingress monitoring.

Considerations: To configure port mirroring, you first specify the mirror port, which is the port that the monitored port is copied to. Then you enable monitoring on the monitored port. Mirror and monitored ports can be configured for ingress and egress traffic:

- s-series: Only 1 ingress and 1 egress mirror port can be configured per port region. A port region is either a group of 12 1 GbE ports, for example ports 1-12 on a 24x1 GbE Interface Module, or a single 10 GbE port. Each 10 GbE port can have one ingress mirror port and one egress mirror port.
- x-series and G4A devices: Only 1 ingress and 1 egress mirror port can be configured per device
- c-series and G5A devices: Only 1 ingress and 1 egress mirror port can be configured per port region. A port region is a group of 24x 1 GbE ports, for example ports 1-24, ports 25-48, or a single 10 GbE port. Each 10 GbE port can have one ingress mirror port and one egress mirror port.

13.3.2 Port mirroring configuration

To designate a mirror port, enter the config level command **`mirror-port ethernet`** ***`5/1`***. On s-series, x-series, and g-series devices you can also append **`input`** or **`output`** to the end of the command if you want to restrict it to being an egress or ingress mirror.

Any frames that the mirror port is not able to handle will be discarded.

If you are mirroring a LAG or a trunk, configuring the mirror port in the first port in the LAG or trunk will automatically configure mirroring in the remaining ports.

We show a mirroring example in Example 13-1.

Example 13-1 Mirror port example

```
BR-FGS648P-STK Router(config)#mirror ethernet 1/1/1
BR-FGS648P-STK Router(config)#show mirror
Mirror port 1/1/1
  Input monitoring      : (Stk1/S1) 1
  Output monitoring     : (Stk1/S1) 1
BR-FGS648P-STK Router(config)#
```

After a mirror port is configured, you can specify a monitored port from which traffic will be copied to the mirror port specified. We show a configuration example in Example 13-2.

Example 13-2 Monitored port example

```
BR-FGS648P-STK Router(config)#interface ethernet 1/1/25
BR-FGS648P-STK Router(config-if-e1000-1/1/25)#monitor ethernet 1/1/1
input
BR-FGS648P-STK Router(config-if-e1000-1/1/25)#
BR-FGS648P-STK Router(config-if-e1000-1/1/25)#show monitor
Monitored Port 1/1/25
  Input mirrored by: (Stk1/S1) 1
  Output mirrored by: None
BR-FGS648P-STK
Router(config-if-e1000-1/1/25)#
```

13.4 sFlow

sFlow is a technology that can sample inbound traffic on a port and send header information from those samples to a central server (called the collector) for analysis.

13.4.1 Setting up sFlow

To set up sFlow, perform the following steps:

1. To set the sFlow collector, use the config level command **`sflow destination`** ***`1.2.3.4`***.
2. Start sFlow with **`sflow enable`**.
3. Go to the interface config level (either for a single interface or multiple interfaces)
4. Run the command **`sflow forwarding`**.
5. (optional) You can modify the sampling rate for this port using the **`sflow sample`** ***`4096`*** command from the interface config level.

> **Note:** The management port (if applicable) will not forward sFlow packets, so the sFlow collector must be accessible by IP from one of the regular traffic ports.

One parameter you might want to adjust is the sampling interval. By default, one out of every 2048 packets will be sampled. You can adjust this by using the config level command sflow sample 8192. Preferably, do not decrease this value below the default of 2048. On the g-series and s-series, the actual used value of this command is the next-highest odd power of 2 (if the value is not equal to an odd power of 2.) For your reference, available used values are: 2, 8, 32, 128, 512, 2048, 8192, 32768, 131072, 524288, 2097152, 8388608, 33554432, 134217728, 536870912, and 2147483648.

sFlow also sends port counter data periodically. To adjust the interval in which the product will do so, use the command **`sflow polling-interval`** ***`30`***.
(To disable this function, enter **`0`** for the value.

13.4.2 sFlow ACL rules (c-series and m-series only)

To configure a **`permit`** ACL rule that will forward all matched traffic to your sFlow collector, simply append **`copy-sflow`** to the end of the rule.

14

Maintenance

In this chapter, we review the basic software/firmware maintenance operations common on the covered products.

14.1 Trivial File Transfer Protocol server

A Trivial File Transfer Protocol (TFTP) server must be run somewhere on your management network and accessible to the product. A functioning TFTP server is considered a customer responsibility, and no TFTP software is either supplied or endorsed by IBM.

However, for the purposes of this book, we used tftpd32, and it appeared to work successfully; nevertheless, any use of this transfer is at your own discretion.

14.2 Secure Copy (SCP)

Secure Copy (SCP) is another method that can be used to transfer files to and from your product.

14.2.1 SCP client

Unlike TFTP, file transfers with SCP are initiated from your configuration workstation, not from the product itself. SCP is automatically enabled upon generation of an SSH key. To set up an SSH key, see 3.3.2, “SSH/SCP” on page 94.

If SSH is disabled, so is SCP.

14.2.2 The pscp product used in this book

Like TFTP, IBM neither supplies nor endorses any particular SCP client. However, for this Redbooks publication, we used `pscp` for all of our examples and it appeared to work. Be aware that any product you use is at your own discretion.

14.3 Software and storage components

For each device, there are several software components that might need to be updated. Each of these components are stored in different locations.

- Application Image: The Multi-Service IronWare (m-series, c-series), Multi-Service IronWare for IBM r-series, and IronWare (s-series, g-series, x-series) images that contains the base OS and software functionality.
- Boot Image: Initializes the device and loads the Application Image.

- Monitor Image (m-, r-, and c-series only): Provides routing image handling and memory initialization.
  - m-series: A separate boot and monitor image is available.
  - r-series: A combined boot and monitor image is available and copied to a single location.
  - c-series: A single boot and monitor image is available and copied to both the boot and monitor locations.
- FPGA Images (m-series, r-series): The m-series and r-series have FPGAs on the management and switch fabric modules that can be field-upgradeable. The m-series also has FPGAs on the interface modules that are field-upgradeable.

When updating the Application Image, it might not be necessary to update the boot, monitor, or FPGA images. See the Release Notes for the product and version for dependencies.

On the m-series and r-series, each Interface module runs their version of the foregoing software images in addition to the management modules. For the m-series and r-series, there are Unified software images that can help simplify the upgrade process. These include:

- m-series: A Unified software image is available which includes the Application image for the management module and all interface modules. FPGA images will still need to be upgraded separately, as do boot and monitor images but these are updated very infrequently.
- r-series: A Unified software image is available which includes the Application image for the management module and all interface modules, as well as the boot-monitor image.
- m-series - The Unified software image contains all Management.

The aforementioned software images are stored in different locations:

- Flash Memory: The Flash Memory is where the Application Image is stored.It is divided into a primary and secondary partition. This allows you to load two different software versions on the same device in case you want to do some testing.
- Boot Flash Memory: There is a memory location reserved for the Boot Image.
- Monitor Flash Memory (m-series, c-series): There is a memory location reserved for the Monitor Image.
- Monitor Flash Memory (r-series): On the r-series, the Boot and Monitor Flash Memory location is the same.

Other locations that files can be stored are to be copied from/to are as follows:

- PCMCIA Card (m-series and r-series): The m-series and r-series management modules have two slots to support a 128 MB PCMCIA ATA Flash Memory card
- TFTP Server
- Remote storage (copied by SCP)

Configuration files can also be copied to/from the devices. These include:

- **`running-config`**: The currently running configuration file on the device.
- **`startup-config`**: The configuration file that is loaded into the running-config upon device bootup.

14.4 File Management (m-series, r-series, c-series)

The file system on the router products can hold arbitrary files beyond the boot images and the startup configuration. The built-in flash memory can hold a total of 32 MB of data, although most of this is used with the primary and secondary boot images.

The filenames of primary, secondary, startup-config, running-config, monitor, and boot-parameter, are “reserved” for particular purposes, so do not use these filenames for other uses.

14.4.1 File management commands

The following commands are available to manage the files:

- **`delete`** ***`filename:`*** Deletes a file
- **`rename`** ***`oldFilename newFilename:`*** Renames a file
- **`more`** ***`filename:`*** Displays the contents of a file
- **`dir:`** Lists the contents of a directory, along with file system free space.
- **`append`** ***`filename1 filename2`***: Adds the contents of *filename1* to the end of *filename2*.
- **`delete`** ***`filename:`*** Deletes a file
- **`cp`** ***`filename1 filename2`***: Creates a copy of *filename1* with the name *filename2*. The **`cp`** command is *very* different from the **`copy`** command, which is primarily used to transfer files to/from TFTP servers.

14.4.2 PCMCIA Card (m-series)

The m-series has two PCMCIA slots in the management module(s) meant to house flash cards. You can use these for the storage code or configuration files.

The two PCMCIA slots are the directories /slot1 and /slot2. (The built-in flash memory is both the root directory and /flash). The file management commands discussed in 14.4, “File Management (m-series, r-series, c-series)” on page 376 can also be applied to files in the PCMCIA slots.

The following commands are available on the PCMCIA cards only:

- **`md`** ***`dirname`***: Creates a directory.
- **`cd`** ***`dirname:`*** Changes the current working directory.
- **`undelete`** ***`filename`***: Un-deletes an accidentally deleted file (this command is *not* available with the base flash).
- **`rmdir`** ***`dirname`***: Removes a directory.
- **`pwd`**: Displays the current directory.
- **`cp`** ***`sourcefile destfile`***
- You can specify the directory as part of filenames, for instance, to copy the file foo.cfg from slot 1 to slot 2, you can use the command **`cp /slot1/foo.cfg /slot2/foo.cfg`**.

14.5 Configuration maintenance

All maintenance operations that involve loading configuration files on and off of the product involve either Trivial File Transfer Protocol (TFTP) or Secure Copy Protocol (SCP).

The startup-config is the configuration file that loads when the device first boots up. The running-config is the currently running configuration on the device.

All TFTP or SCP commands are run from enable mode.

14.5.1 Configuration transfers

In this section, we discuss methods to transfer the configurations.

TFTP

After starting the TFTP server on a management station, you are ready to copy configuration files to/from the product. This is done by using the **`copy`** command. The general format for the copy command is:

```
copy <from where> <to where> <additional attributes>
```

- To copy the running configuration from the device to a TFTP server, use the command **`copy running-config tftp`** ***`1.2.3.4 filename`***.
- To copy the startup configuration from the device to a TFTP server, use the command **`copy startup-config tftp`** ***`1.2.3.4 filename`***.
- To copy a configuration file from a tftp server to the startup configuration, use **`copy tftp startup-config`** ***`1.2.3.4 filename`***.
- To load a configuration from a tftp server and make it the current running configuration, use **`copy tftp running-config`** ***`1.2.3.4 filename`***.

> **Note:** There are some configuration changes that require a reload; this is still the case if you load a new configuration by TFTP. Commands that require a reload to take effect are **`jumbo`** (s-, g-, x-series), some stacking configuration changes (g-series, G5A only) and global configuration commands that start with **`system-max`**. The system will not reload without your permission.

SCP

To transfer configurations on and off of the products, certain “special” filenames are used to direct files to/from the correct location on the products. The running configuration has the filename runConfig, while the startup configuration is startConfig.

With SCP, the default copy into the starting configuration actually appends the new configuration to the existing one. If you want to *replace* the current running configuration with a new one, copy into the file startConfig-overwrite.

To append a configuration file from the local hard drive to the device’s startup-config, use:

```
scp c:\<config_file.cfg> user@1.2.3.4:startConfig
```

To replace a configuration file from the local hard drive to the device’s running-config, use:

```
scp c:\<config_file.cfg> user@1.2.3.4:runConfig-overwrite
```

To copy the running-config from the device to a local location, use:

```
scp user@1.2.3.4:runConfig c:\<config_file.cfg>
```

14.5.2 Configuration inserting

If you want to simply apply new commands to an existing configuration, you can build the commands in the text editor of your choice and simply paste the text into your terminal program.

14.6 The copy command

The syntax for the copy command is complicated, and varies by product model. In general, the syntax is usually: **`copy`** ***`sourceDevice destinatonDevice sourceParameters destinationParameters`***.

> **Note:** If TFTP is a source *or* destination the TFTP parameters go first.

> **Important:** Many of these commands are used for upgrading software images or FPGA software. See 14.8, “Code upgrade process” on page 381 for details before using these commands.

Sources include the following possibilities:

- **`flash:`** Internal flash memory
- **`running-config:`** (PCMCIA or tftp destinations only)
- **`startup-config:`** (PCMCIA or tftp destination only)
- **`slot1`** or **`slot2`**: PCMCIA slots (m- and r-series only)
- **`tftp`**

Destinations include the following possibilities:

- **`flash:`** Internal flash memory
- **`image`**: For use with Unified images (m- and r-series only)
- **`running-config:`** (PCMCIA or tftp sources only)
- **`startup-config:`** (PCMCIA or tftp sources only)
- **`lp:`** Line card (m- and r-series only)
- **`slot1`** or **`slot2`**: PCMCIA slots (m-series only)
- **`snm:`** Switch fabric module (m-series only)
- **`tftp`**

Here are some examples of a **`copy`** command:

- **`copy tftp flash 1.2.3.4 rmb02702b.bin monitor copy-boot:`**
  Copies rmb02702b.bin from TFTP server at 1.2.3.4 to the monitor/boot on r-series

- **copy slot2 image xm04000f.bin:**
  Copies Unified software image xm04000f.bin from slot 2 in PCMCIA on m-series
- **copy flash tftp 1.2.3.4 ce03900.bin primary:**
  Copies the Application image stored in the primary flash partition to the TFTP server at 1.2.3.4 to the file name ce03900.bin

To delve into the parameters for a given source or destination:

- **flash:**
  - If **flash** is both the source and the destination:
    - Source parameters are: **primary** or **secondary**
    - Destination parameters are: <cr> (if the source was **primary**, the destination will be the secondary, and vice-versa) or **standby** (m-series only; refers to the backup management module)
    - **monitor** or **boot** (c-series and m-series): Append to the line to copy the image to the monitor or boot location
    - **monitor copy-boot** (r-series): Append to the line to copy the boot-monitor image to the appropriate location.
  - If the destination is **lp, system,** or **snm**: the source parameter is *filename*
  - **tftp:** remember that, as noted previously, if **tftp** is the source or destination, put the tftp parameters first:
    - If tftp is the source, the allowed **flash** parameters are***: filename,*** **boot, monitor, primary,** or **secondary**
    - If tftp is the destination, the allowed **flash** parameters are ***filename,*** **primary or secondary**.
- **image:** No additional parameters are required on the m-series. On the r-series, you will need to append **lp-boot** and **mp-boot**.
- **slot1** or **slot2*****: filename*** of the file to be copied or copied to
- **tftp:** As noted before, if **tftp** is the source or destination, the tftp parameters go first, even if TFTP is the destination. The parameters are ***1.2.3.4 filename*** where “1.2.3.4” is the IP address of the TFTP server and “filename” the file to be copied or copied to
- **running-config or startup-config:** no parameters
- **snm:** Use the parameter **sbridge all**

- **`lp:`** There are several allowed parameters when using **`lp`** as a destination
  - **`monitor`**
  - **`primary`**
  - **`secondary`**
  - **`fpga-all all`**: Using this parameter will upgrade all the FPGAs on all line cards at once. If you want to do something different, check the release notes for the version of software you are installing.

    We provide instructions or examples for this command later in the chapter.

14.7 Other SCP transfers

You can use SCP to transfer files other than configuration files.

14.7.1 s-series, g-series, and x-series

The only possible file copy operation possible, other than the configuration files discussed previously, are the primary and secondary boot images.

To copy an image into primary boot file, use the following command: **`scp`** ***`sourcefilename username@1.2.3.4:`*****`flash:primary.bin`**. For the secondary boot file, replace **`primary.bin`** with **`secondary.bin`**.

14.7.2 m-series, r-series, and c-series

To copy the management module main images: primary, secondary, and monitor, use the following scp command format: **`scp`** ***`sourcefilename username@1.2.3.4:flash:primary.`***

To copy other files (including files such as FPGA images, other configuration files, and so on) use the format **`scp`** ***`sourcefilename username@1.2.3.4:flash:otherfilename.`*** To copy files onto the PCMCIA cards, replace **`flash`** with **`slot1`** or **`slot2`**.

14.8 Code upgrade process

The code upgrade process varies dramatically by platform. For the latest information, consult the Release Notes issued with the code release you are installing.

Note: Read the Release Notes before upgrading to be informed about all dependencies, requirements, and other important notes for the particular software release. Also review the software upgrade instructions in the Release Notes.

Our instructions will use TFTP, but you might use SCP for certain upgrade operations. See the appropriate Configuration Guide for your product for details.

All of the products have support for a primary and secondary flash partition for holding Application images. Best practice is to load new code versions into the secondary flash location, and only if the boot is successful, copy it to the primary.

Important: See the Release Notes to determine which file is used for which purpose. Be very careful when loading images that you use the correct file in each step.

14.8.1 Obtaining code

The first step prior to upgrading your product is, of course, obtaining code. The code is available on the support website for the Systems Networking products. To access this website, navigate to:

http://www.ibm.com/systems/support/

Then click the link to **Systems Networking**. From here, select your product from the drop-down menu, click **Go** and then **Downloads**.

You will be presented with links to the various major releases available for your product. Note that you *must* pick the appropriate product from the download; even if two models are similar. For instance, the B50G runs a different code load (and has different options from) the B48G, which in turn runs different code than the B48C.

Clicking on the link to download code will lead you to an IBM-customized website provided by Brocade. From here, you can access both the code images and the manuals for the particular release you are installing.

After filling out the Export Control Form, you will be presented with the available images for your product. Only select the image for which you have purchased a license. (If you are unsure, contact your IBM representative or Business Partner for further information).

At this point, you will be presented with the actual files to download. Which files go where will be discussed in detail in the section on each product.

14.8.2 Software upgrade: g-series and x-series

Upgrading the software on the g-series and x-series requires only two steps: upgrading the Boot ROM, and upgrading the Application image (IronWare). (Note: The Boot ROM image might not change with every release).

The following IronWare images are available for the g-series and x-series:

- Layer 2
- Base Layer 3 (g-series only)
- Edge Layer 3 (B48G only, requires optional license)

While the main software might be loaded by SCP, the Boot ROM can only be loaded by TFTP.

For our example, we are loading 5.0.01a, into the secondary flash partition of a B50G. After we verify that it loads successfully, we copy it into the primary code slot.

Upgrading boot code

To upgrade the boot code, use the enable-level command **copy tftp flash** ***1.2.3.4 fgz05000.bin*** **bootrom**. We use this command in Example 14-1.

> **Note:** Upgrading the boot code *might* be disruptive to network traffic. Only upgrade boot code during a maintenance window. Perform the boot code upgrade from the serial port, because you might lose your Telnet/SSH session during the upgrade.

On a stacked devices, the copy command will copy the images to all stacked units.

Example 14-1 Upgrading g-series BootROM

```
FGS648P-STK Switch#copy tftp flash 10.64.210.47 g5.0.01a/fgz05000.bin
bootrom
FGS648P-STK Switch#Flash Memory Write (8192 bytes per dot)
...................................................(Boot Flash Update)
Erase........Write....................................................
TFTP to Flash Done.
FGS648P-STK Switch#
```

Because our Boot ROM upgrade was successful, we now load the Base Layer 3 image into the secondary code slot. We accomplish this using **copy tftp flash *1.2.3.4 fgs05001a.bin* secondary**. This is shown in Example 14-2.

Example 14-2 Loading new code into the secondary flash slot

```
telnet@FGS648P-STK Switch#copy tftp flash 10.64.210.47
g5.0.01a/FGS05001a.bin secondary
telnet@FGS648P-STK Switch#Flash Memory Write (8192 bytes per dot)
........................................................................
........................................................................
........................................................................
........................................................................
........................................................................
.................................
TFTP to Flash Done.
telnet@FGS648P-STK Switch#
```

On stacked units, the code will be copied to the designated flash partition on all stacked units (this is not shown on the console).

We now want to request that the g-series boot once from the secondary flash slot. This is accomplished with the command **boot system flash secondary**.

Note: Run this command from the enable prompt, *not* the config prompt. Running the command from the config prompt will make the change permanent. We are testing the new code; if it fails, we want the product to go back to using the primary code image.

Note: Reloading the product *is* disruptive to network traffic until the product completes the reload process.

Copying secondary code image

Assuming the system came up normally, copy the secondary code image into the primary slot with the command **copy flash flash primary**. This final process is shown in Example 14-3. (On stacked units, this will copy the secondary code slot from the stack controller to the primary slot on all stack units. It will ignore the content of the secondary bank on the non-controller units).

Example 14-3 Final s-series upgrade result

```
BR-telnet@s_series#show flash
Active Management Module (Slot 10):
```

```
Compressed Pri Code size = 3319603, Version 05.0.00aT3e3
(sxr05000a.bin)
Compressed Sec Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9175040
BR-telnet@s_series#copy flash flash primary
BR-telnet@s_series#Flash Memory Write (8192 bytes per dot)
.......................................................................
.......................................................................
.......................................................................
.......................................................................
..........................................................Flash to
Flash Done.
BR-telnet@s_series#sh flash
Active Management Module (Slot 10):
Compressed Pri Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed Sec Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9699328
BR-telnet@s_series#
```

14.8.3 Software upgrade: s-series

Upgrading the code on the s-series requires only two steps: upgrading the Boot ROM, and upgrading the main software. (Note: The Boot ROM image might not change with every release).

The following IronWare images are available for the s-series:

- Layer 2
- Base Layer 3
- Full Layer 3 (requires optional license)

While the main software can be loaded by SCP, the Boot ROM can only be loaded by TFTP.

All loaded code is automatically copied over to the backup management module (if present).

For our example, we are loading 5.0.00a, Base Layer 3 image into the secondary flash partition. After we verify that it loads successfully, we copy it into the primary code slot.

Upgrading boot code

To upgrade the boot code, use the enable-level command **`copy tftp flash`** ***`1.2.3.4 sxz05000.bin`*** **`bootrom`**. We use this command in Example 14-4.

Note: Upgrading the boot code *might* be disruptive to network traffic. Only upgrade boot code during a maintenance window. Perform he boot code upgrade from the serial port, because you might lose your Telnet/SSH session during the upgrade.

Example 14-4 Upgrading s-series BootROM

```
BR-telnet@s_series# copy tftp flash 10.64.210.47 sxz05000.bin bootrom
SW-telnet@s_series#Load to buffer (8192 bytes per dot)
................................................................Write
to boot flash.........
TFTP to Flash Done.
BR-telnet@s_series#
BR-telnet@s_series#show flash
Active Management Module (Slot 10):
Compressed Pri Code size = 3319603, Version 05.0.00aT3e3
(sxr05000a.bin)
Compressed Sec Code size = 2502563, Version 05.0.00aT3e1
(sxs05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9437184
```

(We don't have a "before" `show flash` because our test s-series was already at the latest level).

Because our Boot ROM upgrade was successful, we now load the Base Layer 3 image into the secondary code slot. We accomplish this using **`copy tftp flash`** ***`1.2.3.4 sxl05000a.bin`*** **`secondary`**. This is shown in Example 14-5.

Example 14-5 Loading new code into the secondary flash slot

```
BR-telnet@s_series#copy tftp flash 10.64.210.47
/sx5.0.00a/sxl05000a.bin secondary
BR-telnet@s_series#Flash Memory Write (8192 bytes per dot)
.......................................................................
.......................................................................
.......................................................................
```

```
.......................................................................
...........................................................
TFTP to Flash Done.
BR-telnet@s_series#show flash
Active Management Module (Slot 10):
Compressed Pri Code size = 3319603, Version 05.0.00aT3e3
(sxr05000a.bin)
Compressed Sec Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 7864320
BR-telnet@s_series#
```

If you look carefully at the `show flash` output, you can see that the secondary code image file changed from sxs0500a.bin in Example 14-4 on page 386 to sxl0500a.bin in Example 14-5 on page 386.

We now want to request that the s-series boot once from the secondary flash slot. This is accomplished with the command **`boot system flash secondary`**.

Note: Run this command from the enable prompt, *not* the config prompt. Running the command from the config prompt will make the change permanent. We are testing the new code; if it fails, we want the product to go back to using the primary code image.

Note: Reloading the product *is* disruptive to network traffic until the product completes the reload process.

Copying secondary code image

Assuming the system came up normally, copy the secondary code image into the primary slot with the command **`copy flash flash primary`**. This final process is shown in Example 14-6.

Example 14-6 Final s-series upgrade result

```
BR-telnet@s_series#show flash
Active Management Module (Slot 10):
Compressed Pri Code size = 3319603, Version 05.0.00aT3e3
(sxr05000a.bin)
Compressed Sec Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9175040
```

```
BR-telnet@s_series#copy flash flash primary
BR-telnet@s_series#Flash Memory Write (8192 bytes per dot)
.......................................................................
.......................................................................
.......................................................................
.......................................................................
..........................................................Flash to
Flash Done.
BR-telnet@s_series#sh flash
Active Management Module (Slot 10):
Compressed Pri Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed Sec Code size = 2798275, Version 05.0.00aT3e2
(/sx5.0.00a/sxl05000a.bin)
Compressed BootROM Code size = 524288, Version 05.0.00T3e5
Code Flash Free Space = 9699328
BR-telnet@s_series#
```

14.8.4 Hitless software upgrade: s-series

The s-series switches are capable of being upgraded without disruption to most Layer 2 production traffic.

> **Important:** At the time of writing, the hitless upgrade process on the s-series is supported only if *all* of the following conditions are met:
>
> - You are currently running 5.0.00 or higher. (this must be the case for all IBM-supported units).
> - The release notes state that upgrades from the release you are currently running to the release you want to run can be performed in a hitless fashion. (For instance, a hitless *downgrade* from 5.1 to 5.0 is *not* supported, while a hitless upgrade *from* 5.0 to 5.1 is).
> - You can tolerate the interruption of Layer 3 protocols, including routed IP traffic. (The interruption of basic routed traffic will not last long, but it will be interrupted).
> - You have two management modules installed (must be matched).
> - You have serial console access to *both* management modules.
> - You can tolerate the interruption of traffic over the interfaces built-in to the management modules, if applicable. (Traffic will not be interrupted for very long, but it will be interrupted).
> - You meet any other conditions detailed in the release notes.

The hitless upgrade process is quite straightforward:

> **Note:** Perform the upgrade commands from a serial console port. Also, make sure you have serial console access available to the standby console port, in case the upgrade fails.

First, load your new code into the code slot using **copy tftp flash *1.2.3.4 sxz05100.bin* primary**. (or secondary, if desired).

Next, run the enable-level command **hitless-reload**.

14.8.5 Software upgrade: c-series

Upgrading the code on the c-series requires only three steps: upgrading the Monitor image, the Boot flash and the main software. (Note: The Boot and Monitor Images might not change with every release).

There is a single Multi-Service IronWare image available for this device that includes all routing functionality, dependent on the license activation you have on your device. A single boot/monitor image is also available that is copied to both the boot and monitor image locations.

While the Application image can be loaded by SCP, the Boot ROM image and Monitor image can only be loaded by TFTP.

For our example, we are loading 3.8.00c into the secondary flash partition. After we verify that it loads successfully, we copy it into the primary flash partition.

Image backup

Preferably, back up your current running code image before beginning the upgrade process. Do this with the command **cp *primary backupfilename.img***. You will only be able to fit a single image backup in the memory of the c-series, so you will need to delete it before another upgrade.

Upgrading boot and monitor code

To upgrade the monitor code, use the enable-level command **copy tftp flash *1.2.3.4 ceb03800c.bin* monitor**. We use this command in Example 14-7.

The Boot flash upgrade is a very similar process. In the software release we are using, they are even the same file as the Monitor image. (Check the release notes for details on your particular release). Use the command **copy tftp flash *1.2.3.4 ceb03800c.bin* boot**.

Note: Upgrading boot code *might* be disruptive to network traffic. Only upgrade boot code during a maintenance window. Perform the boot code upgrade from the serial port because you might lose your Telnet/SSH session during the upgrade.

Example 14-7 Upgrading c-series Monitor image and Boot flash

```
telnet@NetIron CES 2048C#
telnet@NetIron CES 2048C#copy tftp flash 10.64.210.47 /c3.8.00c/
ce03800c.binb03800c.bin03800c.bin  monitor
................................TFTP: Download to MP mon flash done.
telnet@NetIron CES 2048C#copy tftp flash 10.64.210.47
/c3.8.00c/ceb03800c.bin boot
................................TFTP: Download to MP boot flash done.
Copy to boot flash:
Erasing.......Writing.................................................
..............Done
telnet@NetIron CES 2048C#
telnet@NetIron CES 2048C#show flash
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
Code Flash - Type MT28F128J3, Size 32 MB
  o IronWare Image (Primary)
    Version 3.8.0cT183, Size 9129481 bytes, Check Sum ee86
    Compiled on Mar 25 2009 at 21:19:22 labeled as ce03800c
  o IronWare Image (Secondary)
    Version 3.8.0cT183, Size 9129481 bytes, Check Sum ee86
    Compiled on Mar 25 2009 at 21:19:22 labeled as ce03800c
  o Monitor Image
    Version 3.8.0cT185, Size 342809 bytes, Check Sum 1cf4
    Compiled on Mar 25 2009 at 20:01:26 labeled as ceb03800c
  o Startup Configuration
    Size 379 bytes, Check Sum 3870
    Modified on 16:08:27 GMT+00 Wed Jan 03 1900

Boot Flash - Type AM29LV040B, Size 512 KB
  o Boot Image
    Version 3.8.0cT185, Size 342809 bytes, Check Sum 1cf4
    Compiled on Mar 25 2009 at 20:01:26 labeled as ceb03800c
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
telnet@NetIron CES 2048C#
```

Because our Boot and monitor upgrade was successful, we now load the Base Layer 3 image into the secondary code slot. We accomplish this using **copy tftp flash *1.2.3.4 ce03800c.bin* secondary**. This is shown in Example 14-8.

Example 14-8 Loading new code into the secondary flash slot

```
telnet@NetIron CES 2048C#
telnet@NetIron CES 2048C#copy tftp flash 10.64.210.47
/c3.8.00c/ce03800c.bin secondary
.......................................................................
.......................................................................
.............................................................TFTP:
Download to secondary flash done.
telnet@NetIron CES 2048C#
```

We now want to request that the c-series boot once from the secondary flash slot. This is accomplished with the command **`boot system flash secondary`**.

> **Note:** Run this command from the enable prompt, *not* the config prompt. Running the command from the config prompt will make the change permanent. We are testing the new code; if it fails, we want the product to go back to using the primary code image.

> **Note:** Reloading the product *is* disruptive to network traffic until the product completes the reload process.

Assuming the system came up normally, copy the secondary code image into the primary slot with the command **`copy flash flash primary`**.

14.8.6 Software upgrade: m-series

Upgrading the code in an m-series unit is a multi-step process. Each step must be carried out carefully and in the correct order for the upgrade to be successful.

1. Take an inventory of your current code levels
2. Cross-check the inventory with the requirements of the new code
3. Upgrade the main Multi-Service IronWare image with the Unified image which upgrades the Management and Interface modules together
4. Upgrade the management monitor and boot images (if required)
5. Upgrade the interface card monitor and boot images (if required)
6. Upgrade the MBRIDGE FPGA in the management module(s) (if required)
7. Upgrade the interface FPGA in the interface modules (if required)
8. Perform a reload coherence check

9. Reload the software

MP in this upgrade process refers to the Management Processor on the Management modules. LP refers to the Line Processor on the Interface modules.

Current code inventory

The code inventory is taken using the **`show version`** command. From this output, you will need to take note of the following items:

- Management Module(s):
  - Multi-Service IronWare Image name (for the slot you are currently running; primary or secondary)
  - LP Kernel (Monitor) Image name (a copy of the LP images are kept on the Management module)
  - LP IronWare Image name (a copy of the LP images are kept on the Management module)
  - Boot Image name
  - Monitor Image name
  - MBRIDGE revision (Management module FPGA)
- Interface Card(s):
  - Multi-Service IronWare Image name
  - Boot Image name
  - Monitor Image name

    The Boot, Monitor, and IronWare images must be the same among all your line cards.
  - FPGA versions (differ by line card type)

The output from the m-series in our setup is shown in Example 14-9.

Example 14-9 m-series show version

```
telnet@m_Series(config)#
telnet@m_Series(config)#show version
HW: NetIron MLX 16K Router
Backplane (Serial #: XXYYYYYYYY,  Part #: 95605-300C)
NI-X-SF Switch Fabric Module 1 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
NI-X-SF Switch Fabric Module 2 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
```

```
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
NI-X-SF Switch Fabric Module 3 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
NI-X-SF Switch Fabric Module 4 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
=======================================================================
===
SL M2: NI-MLX-MR Management Module Active (Serial #: XXYYYYYYYY, Part
#: 35524-103C):
Boot     : Version 3.5.0T165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:13:56 labeled as xmprm03500
 (424484 bytes) from boot flash
Monitor  : Version 3.5.0fT165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:04:30 labeled as xmb03500f
 (422466 bytes) from code flash
IronWare : Version 4.0.0dT163 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on May 22 2009 at 23:03:58 labeled as xmr04000d
 (6625478 bytes) from Primary
Board ID : 00 MBRIDGE Revision : 21
916 MHz Power PC processor  (version 8003/0101) 166 MHz bus
512 KB Boot Flash (AM29LV040B), 32 MB Code Flash (MT28F128J3)
1024 MB DRAM
Active Management uptime is 5 days 2 hours 46 minutes 5 seconds
=======================================================================
===
SL M1: NI-MLX-MR Management Module Standby (Serial #: XXYYYYYYYY, Part
#: 35524-103C):
Boot     : Version 3.5.0T165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:13:56 labeled as xmprm03500
 (424484 bytes) from boot flash
Monitor  : Version 3.5.0fT165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:04:30 labeled as xmb03500f
```

```
 (422466 bytes) from code flash
IronWare : Version 4.0.0dT163 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on May 22 2009 at 23:03:58 labeled as xmr04000d
 (6625478 bytes) from Primary
Board ID : 00 MBRIDGE Revision : 21
916 MHz Power PC processor  (version 8003/0101) 166 MHz bus
512 KB Boot Flash (AM29LV040B), 32 MB Code Flash (MT28F128J3)
1024 MB DRAM
Standby Management uptime is 5 days 2 hours 45 minutes 7 seconds
=======================================================================
===
SL 1: NI-MLX-10Gx4 4-port 10GbE Module (Serial #: XXYYYYYYYY, Part #:
35600-202D)
Boot     : Version 3.5.0T175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:14:32 labeled as xmlprm03500
 (387133 bytes) from boot flash
Monitor  : Version 3.5.0fT175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
 (387707 bytes) from code flash
IronWare : Version 4.0.0dT177 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on May 22 2009 at 23:13:14 labeled as xmlp04000d
 (4116213 bytes) from Primary
FPGA versions:
Valid PBIF Version = 3.14, Build Time = 12/17/2008 14:32:00

Valid XPP Version = 5.07, Build Time = 12/9/2008 16:39:00

Valid XGMAC Version = 0.12, Build Time = 11/10/2008 15:50:00

X10G2MAC 0
X10G2MAC 1
666 MHz MPC 8541 (version 8020/0020) 333 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 286331153 Bytes BRAM
PPCR0: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
PPCR1: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
LP Slot 1 uptime is 5 days 2 hours 45 minutes 1 seconds
=======================================================================
===
SL 5: NI-MLX-1Gx20-SFP 20-port 1GbE-100FX Module (Serial #: XXYYYYYYYY,
Part #: 35604-102C)
```

```
Boot     : Version 3.5.0T175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:14:32 labeled as xmlprm03500
 (387133 bytes) from boot flash
Monitor  : Version 3.5.0fT175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
 (387707 bytes) from code flash
IronWare : Version 4.0.0dT177 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on May 22 2009 at 23:13:14 labeled as xmlp04000d
 (4116213 bytes) from Primary
FPGA versions:
Valid PBIF Version = 3.14, Build Time = 12/17/2008 14:32:00

Valid XPP Version = 5.07, Build Time = 12/9/2008 16:39:00

BCM5695GMAC 0
BCM5695GMAC 1
666 MHz MPC 8541 (version 8020/0020) 333 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 286331153 Bytes BRAM
PPCR0: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
LP Slot 5 uptime is 5 days 2 hours 45 minutes 3 seconds
=======================================================================
===
SL 9: NI-MLX-1Gx20-GC 20-port 10/100/1000 Copper Module (Serial #:
XXYYYYYYYY, Part #: 35603-102C)
Boot     : Version 3.5.0T175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:14:32 labeled as xmlprm03500
 (387133 bytes) from boot flash
Monitor  : Version 3.5.0fT175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
 (387707 bytes) from code flash
IronWare : Version 4.0.0dT177 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on May 22 2009 at 23:13:14 labeled as xmlp04000d
 (4116213 bytes) from Primary
FPGA versions:
Valid PBIF Version = 3.14, Build Time = 12/17/2008 14:32:00

Valid XPP Version = 5.07, Build Time = 12/9/2008 16:39:00
```

```
BCM5695GMAC 0
BCM5695GMAC 1
666 MHz MPC 8541 (version 8020/0020) 333 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 286331153 Bytes BRAM
PPCR0: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
LP Slot 9 uptime is 5 days 2 hours 45 minutes 3 seconds
=======================================================================
===
All show version done
telnet@m_Series(config)#
```

In our example, we have the following code inventory:

- Management modules:
  - Multi-Service IronWare xmr04000d
  - Monitor xmb03500f
  - Boot xmprm03500
  - MBRIDGE 21
- Interface modules:
  - Multi-Service IronWare xmlp04000d
  - Monitor xmlb03500f
  - Boot xmlprm03500
  - FPGA: 10GbE module
    - PBIF 3.14
    - XPP 5.07
    - XGMAC 0.12
  - FPGA: 20-port GigE modules (both SFP and Copper)
    - PBIF 3.14
    - XPP 5.07

  If we had a POS module, we need to note the versions for the PBIF, XPP, and STATS FPGAs.

Version cross-check

We are going to be installing IronWare 4.0.00b on our device — this will be a downgrade from the currently installed code level.

The FPGA images are a little harder to identify; while **show version** might label a FPGA in each interface module with an identical name, such as “PBIF”, they actually refer to different FPGA images, depending on the type of the line card.

From the release notes for this version, in the section entitled: “Image Files for Multi-Service IronWare R04.0.00b” we see the following requirements:

- Management Module:
  - Multi-Service IronWare xmr04000b
  - Monitor xmb03500f
  - Boot xmprm03500
  - MBRIDGE 21
- Interface modules:
  - Multi-Service IronWare xmlp04000b
  - Monitor xmlb03500f
  - Boot xmlprm03500
  - FPGA: 10GbE
    - PBIF (PBIFSP2.bin) 3.14
    - XPP (XPPSP2.bin) 5.07
    - XGMAC (XGMACSP2.bin) 0.12
  - FPGA: 20-port GigE modules
    - PBIF (PBIFSP2.bin) 3.14
    - XPP (XPPSP2.bin) 5.07
  - FPGA: POS (not installed in our example)
    - XPP (XPPOC.bin) 5.07
    - PBIF (PBIFOC.bin) 3.04
    - STATS (STATSOC.bin) 2.6

If we compare this list with our current installed code level, we see that all the FPGA levels and Boot and Monitor images are at their correct levels. The only code that needs changing is the IronWare itself.

If we were actually performing a code upgrade (or downgrade, in our case), of course, we only load the code necessary for the upgrade process. However, for the sake of completeness, in this section we go through the upgrade process for all of the code loaded on the product, including FPGAs code and Boot and Monitor images.

Multi-Service IronWare code upgrade

For convenience, Multi-Service IronWare Images for both the management and interface modules are available as a single downloadable package, in our case xm04000b.

To start the upgrade, run the command **`copy tftp image`** ***`1.2.3.4 xm04000b.bin`***. We see the process take place in Example 14-10.

Example 14-10 r-series upgrade example

```
telnet@m_Series#
telnet@m_Series#copy tftp image 10.64.210.47 m4.0.00b/xm04000b.bin
.......................................................................
..............................................................
Download combined image from tftp is done - Start Copying Individual
Images.
1) Copy LP Application Image.
Copy to LP primary flash.
Copy file LP Application on MP to file primary on all LP slots
File Download: LP Application (MP) -> primary (LP 1) is done.
File Download: LP Application (MP) -> primary (LP 9) is done.
File Download: LP Application (MP) -> primary (LP 5) is done.
File download to interface module is done (3 successful)
Save a copy to MP's flash, please
wait...................................................................
.......................................................................
.......................................................................
.................Done
Copy LP PRIMARY IMAGE to standby MP, please wait.
Start code flash synchronization to standby MP.
Code flash synchronization to standby MP is done.
2) Copy MP Application Image.
Copy to MP primary
flash..................................................................
.........................................................done.
Copy MP PRIMARY IMAGE to standby MP, please wait.
Start code flash synchronization to standby MP.
Code flash synchronization to standby MP is done.
telnet@m_Series#
```

Management boot and monitor upgrade

Upgrading the interface card boot and monitor code will not be required on every code release and will be needed very infrequently.

To upgrade boot or monitor code on the management module, enter the command **copy tftp flash *1.2.3.4 xmprm03500.bin boot***. We go through this process in Example 14-11. (If we wanted to upgrade monitor code, we replace “boot” with “monitor” and use the appropriate image).

Example 14-11 Upgrading management module boot code.

```
m_Series(config)#
```

```
m_Series#copy tftp flash 10.64.210.47 m4.0.00b/xmprm03500.bin boot
.......................................................................
................................TFTP: Download to MP boot flash done.
Copy to boot flash:
Erasing.......Writing.................................................
..Done
Copy MP BOOT IMAGE to standby MP, please wait.
Not Needed. [This is not needed because we aren’t actually doing an
upgrade here, but if we were, we would see the code being loaded onto
the alternate management moudle.]
m_Series#
m_Series#
```

Interface card boot and monitor upgrade

Upgrading interface card boot and monitor code will not be required on every code release and will be needed very infrequently.

To upgrade boot or monitor code on the interface module, enter the command **copy tftp flash *1.2.3.4 xmlprm03500.bin boot* all**. We go through this process in Example 14-12.

If we wanted to upgrade monitor code, we replace “boot” with “monitor” and use the appropriate image. Our example command will upgrade all line cards at once.

Example 14-12 Upgrading interface module boot code

```
m_Series#
m_Series#copy tftp lp 10.64.210.47 m4.0.00b/xmlprm03500.bin boot all
.......................................................................
.......................................................TFTP: Download
to LP boot done.
Copy file m4.0.00b/xmlprm03500.bin from tftp on MP to file boot on all
LP slots
File Download: m4.0.00b/xmlprm03500.bin from t (MP) -> boot (LP 1) is
done.
File Download: m4.0.00b/xmlprm03500.bin from t (MP) -> boot (LP 9) is
done.
File Download: m4.0.00b/xmlprm03500.bin from t (MP) -> boot (LP 5) is
done.
File download to interface module is done (3 successful)
m_Series#
m_Series#
```

MBRIDGE FPGA upgrade

Upgrading the MBRIDGE FPGA on the management module might not be required on every code release.

To upgrade the MBRIDGE FPGA code, enter the command **copy tftp mbridge *1.2.3.4 mbridge_04000b.xsvf***. You can see this process in Example 14-13.

Example 14-13 MBRIDGE code upgrade

```
m_Series#
m_Series#copy tftp mbridge 10.64.210.47 m4.0.00b/mbridge_04000b.xsvf
.......................................................................
.......................................................................
...................TFTP: Download MBRIDGE to memory done.
Copy to MBRIDGE PROM...............................................Save
the new MBRIDGE to
flash..................................................................
.......................................................................
.........................Done
Copy MBRIDGE IMAGE to standby MP, please wait.
Not Needed.

Author note: This is not needed because we aren’t actually doing an
upgrade here, but if we were, we would see the code being loaded onto
the alternate management moudle.

m_Series#
m_Series#
```

Interface module FPGA upgrade

Upgrading the FPGA images on the interface modules might not be required on every code release. All of the FPGA code is available in one consolidated download package; this is the one to obtain.

To upgrade the FPGAs on the Interface modules with the Unified LP FPGA image, enter the command **copy tftp lp *1.2.3.4 lpfpga04000b.bin* fpga-all all**. You can see this process in Example 14-14.

Example 14-14 Upgrading interface module FPGA code

```
m_Series#
m_Series#copy tftp lp 10.64.210.47 m4.0.00b/lpfpga04000b.bin fpga-all
all
Bundle FPGA download begins: 9 images in the bundle.
```

```
.......................................................................
.......................................................................
...........................TFTP: Download to LP FPGA ALL done.
Copying FPGA images to the applicable slot(s), this might take several
minutes...
Copying 1st image (PBIF - Ethernet) to slot(s) 1, 5, 9 skipped, same
version exists. Use "force overwrite" if required.
Copying 4th image (XPP - Ethernet) to slot(s) 1, 5, 9 skipped, same
version exists. Use "force overwrite" if required.
Copying 7th image (XGMAC - Ethernet) to slot(s) 1 skipped, same version
exists. Use "force overwrite" if required.
No FPGA image to be copied. [This is not needed because we aren't
actually doing an upgrade here, but if we were, we would see the code
being loaded onto the interface modules.]
Bundle FPGA copy to interface module(s) completed.
m_Series#
m_Series#
```

Reload coherence check

Prior to actually booting into the new code, run a reload coherence check to verify that all the software that is loaded on the product is compatible.

To do this, run the command **`reload-check`**. If there are no problems, the product will simply respond with "`Checking for coherence... Done.`" If there are problems with the load, this command will produce output similar to Example 14-15.

Example 14-15 Reload coherence check

```
NetIron# reload-check
Checking for coherence... done.
Warning: The new MP application (3 6 0 13) will not be compatible with
the new LP #application (3 6 0 0) version
Warning: The new MP application (3 6 0 0) will not be compatible with
the new LP #application (3 5 0 0) version
Warning: The new LP application (3 6 0 13) on MP will not be compatible
with the new LP # application (3 6 0 0) version
Warning: The new LP monitor (3 6 0 0) on MP will not be compatible with
the new LP monitor (3 5 0 0) version
Warning: The new LP # application (3 6 0 0) will not be compatible with
the new LP monitor (3 5 0 0) version.
Warning: The new LP # application (3 6 0 0) will not be compatible with
the new LP monitor (3 5 0 0) version.
Warning: The new LP PBIF FPGA will not be compatible with the new LP #
application.
```

```
Warning: The new LP XPP FPGA will not be compatible with the new LP #
application.
Warning: The new LP XPP XGMAC will not be compatible with the new LP #
application.
Warning: The new LP PBIF-OC FPGA will not be compatible with the new LP
# application.
Warning: The new LP XPP-OC FPGA will not be compatible with the new LP
# application.
Warning: The new LP SPP STATS-OC will not be compatible with the new LP
# application.
Are you sure? (enter ‘y’ or ‘n’):
```

If this occurs, do not reload the device and compare the software image versions found on your device with those noted in the Release Notes. If everything looks compatible and the **reload-check** still fails, contact IBM Support.

Product reload

Assuming the coherence check passed, you can now reload the product with the **reload** command. If you want to perform a hitless upgrade, see “Hitless upgrades” on page 407.

Upgrade verification

Now that our upgrade process is complete, we can verify the upgrade with another run of **show version**. Our final output can be seen in Example 14-16.

Example 14-16 Final upgrade results

```
m_Series#
m_Series#show version

W: NetIron MLX 16K Router
Backplane (Serial #: XXYYYYYYYY,  Part #: 95605-300C)
NI-X-SF Switch Fabric Module 1 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
NI-X-SF Switch Fabric Module 2 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
NI-X-SF Switch Fabric Module 3 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
```

```
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
NI-X-SF Switch Fabric Module 4 (Serial #: XXYYYYYYYY,  Part #:
35523-302A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
=======================================================================
===
SL M2: NI-MLX-MR Management Module Active (Serial #: XXYYYYYYYY, Part
#: 35524-103C):
Boot     : Version 3.5.0T165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:13:56 labeled as xmprm03500
 (424484 bytes) from boot flash
Monitor  : Version 3.5.0fT165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:04:30 labeled as xmb03500f
 (422466 bytes) from code flash
IronWare : Version 4.0.0bT163 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Feb 27 2009 at 21:45:56 labeled as xmr04000b
 (6609952 bytes) from Primary
Board ID : 00 MBRIDGE Revision : 21
916 MHz Power PC processor  (version 8003/0101) 166 MHz bus
512 KB Boot Flash (AM29LV040B), 32 MB Code Flash (MT28F128J3)
1024 MB DRAM
Active Management uptime is 4 minutes 17 seconds
=======================================================================
===
SL M1: NI-MLX-MR Management Module Standby (Serial #: XXYYYYYYYY, Part
#: 35524-103C):
Boot     : Version 3.5.0T165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:13:56 labeled as xmprm03500
 (424484 bytes) from boot flash
Monitor  : Version 3.5.0fT165 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:04:30 labeled as xmb03500f
 (422466 bytes) from code flash
IronWare : Version 4.0.0bT163 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Feb 27 2009 at 21:45:56 labeled as xmr04000b
 (6609952 bytes) from Primary
```

```
Board ID : 00 MBRIDGE Revision : 21
916 MHz Power PC processor  (version 8003/0101) 166 MHz bus
512 KB Boot Flash (AM29LV040B), 32 MB Code Flash (MT28F128J3)
1024 MB DRAM
Standby Management uptime is 3 minutes 17 seconds
=======================================================================
===
SL 1: NI-MLX-10Gx4 4-port 10GbE Module (Serial #: XXYYYYYYYY, Part #:
35600-202D)
Boot     : Version 3.5.0T175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:14:32 labeled as xmlprm03500
 (387133 bytes) from boot flash
Monitor  : Version 3.5.0fT175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
 (387707 bytes) from code flash
IronWare : Version 4.0.0bT177 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Feb 27 2009 at 22:09:14 labeled as xmlp04000b
 (4110942 bytes) from Primary
FPGA versions:
Valid PBIF Version = 3.14, Build Time = 12/17/2008 14:32:00

Valid XPP Version = 5.07, Build Time = 12/9/2008 16:39:00

Valid XGMAC Version = 0.12, Build Time = 11/10/2008 15:50:00

X10G2MAC 0
X10G2MAC 1
666 MHz MPC 8541 (version 8020/0020) 333 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 286331153 Bytes BRAM
PPCR0: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
PPCR1: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
LP Slot 1 uptime is 3 minutes 16 seconds
=======================================================================
===
SL 5: NI-MLX-1Gx20-SFP 20-port 1GbE-100FX Module (Serial #: XXYYYYYYYY,
Part #: 35604-102C)
Boot     : Version 3.5.0T175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:14:32 labeled as xmlprm03500
 (387133 bytes) from boot flash
```

```
Monitor  : Version 3.5.0fT175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
 (387707 bytes) from code flash
IronWare : Version 4.0.0bT177 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Feb 27 2009 at 22:09:14 labeled as xmlp04000b
 (4110942 bytes) from Primary
FPGA versions:
Valid PBIF Version = 3.14, Build Time = 12/17/2008 14:32:00

Valid XPP Version = 5.07, Build Time = 12/9/2008 16:39:00

BCM5695GMAC 0
BCM5695GMAC 1
666 MHz MPC 8541 (version 8020/0020) 333 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 286331153 Bytes BRAM
PPCR0: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
LP Slot 5 uptime is 3 minutes 16 seconds
=======================================================================
===
SL 9: NI-MLX-1Gx20-GC 20-port 10/100/1000 Copper Module (Serial #:
XXYYYYYYYYY, Part #: 35603-102C)
Boot     : Version 3.5.0T175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Jul 10 2007 at 19:14:32 labeled as xmlprm03500
 (387133 bytes) from boot flash
Monitor  : Version 3.5.0fT175 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Oct 31 2008 at 12:06:22 labeled as xmlb03500f
 (387707 bytes) from code flash
IronWare : Version 4.0.0bT177 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Feb 27 2009 at 22:09:14 labeled as xmlp04000b
 (4110942 bytes) from Primary
FPGA versions:
Valid PBIF Version = 3.14, Build Time = 12/17/2008 14:32:00

Valid XPP Version = 5.07, Build Time = 12/9/2008 16:39:00

BCM5695GMAC 0
BCM5695GMAC 1
666 MHz MPC 8541 (version 8020/0020) 333 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
```

```
512 MB DRAM, 8 KB SRAM, 286331153 Bytes BRAM
PPCR0: 768K entries CAM, 8192K PRAM, 2048K AGE RAM
LP Slot 9 uptime is 3 minutes 16 seconds
=======================================================================
===
All show version done
m_Series#
m_Series#
```

If you look carefully through the example, you will see that our primary IronWare image in both the management modules and the interface modules has now changed to 4.0.00b (from 4.0.00d).

It is also a good practice to verify that all of your interface modules come up after the reload. This is done with a **show module** command, as seen in Example 14-17. If any modules show an error status, contact IBM Support.

Example 14-17 show module output.

```
m_Series#
m_Series#show module

       Module                                Status      Ports  Starting
MAC
M1 (upper): NI-MLX-MR Management Module     Standby (Ready State)
M2 (lower): NI-MLX-MR Management Module     Active
F1: NI-X-SF Switch Fabric Module         Active
F2: NI-X-SF Switch Fabric Module         Active
F3: NI-X-SF Switch Fabric Module         Active
F4: NI-X-SF Switch Fabric Module         Active
S1: NI-MLX-10Gx4 4-port 10GbE Module  CARD_STATE_UP    4
001b.ed22.8400
S2:
S3:
S4:
S5: NI-MLX-1Gx20-SFP 20-port 1GbE-100FX Module  CARD_STATE_UP    20
001b.ed22.84c0
S6:
S7:
S8:
S9: NI-MLX-1Gx20-GC 20-port 10/100/1000 Copper Module  CARD_STATE_UP
20   001b.ed22.8580
S10:
S11:
S12:
```

```
S13:
S14:
S15:
S16:
m_Series#
m_Series#
```

Hitless upgrades

The m-series products are capable of being upgraded with minimal disruption to production traffic if the product is equipped with a redundant management modules. Layer 2 switching and Layer 3 forwarding continues to occur, as well as Layer 3 routing with Graceful OSPF and BGP restart configured. Some functionality is not supported though, including FPGA image upgrades and MPLS features. See the *Hardware Installation Guide* for additional details.

Before you plan a hitless upgrade, there are conditions that *must* be met.

Important: At the time of writing, the hitless upgrade process on the m-series is supported only if *all* of the following conditions are met:

- You have two management modules installed
- You have serial console access to *both* management modules
- If you are using OSPF and/or BGP, you have enabled OSPF Graceful Restart and/or BGP Graceful Restart
- You do not need to change router configuration during the 1 to 10 minute process
- You have not made any changes requiring a software reload
- You do not need to upgrade any FPGA images
- You are not downgrading code
- You meet any other conditions detailed in the release notes

The process for carrying is quite simple. Using the instructions in the previous section, load new Boot and Monitor images (if necessary), then load the new Multi-Service IronWare image, but instead of then doing a **`reload`**, use the command **`hitless-reload mp`** ***`primary`*** **`lp`** ***`primary`***.

14.8.7 Software upgrade: r-series

Upgrading the code in an r-series unit is a multi-step process. Each step must be carried out carefully and in the correct order for the upgrade to be successful.

1. Take an inventory of your current code levels
2. Cross-check the inventory with the requirements of the new code
3. Upgrade the main Multi-Service IronWare for IBM r-series image with the Unified image which upgrades the Management and Interface modules together, along with the Boot and Monitor images
4. Upgrade the MBRIDGE FPGA in the management module(s) (if required)
5. Verify the images have copied successfully
6. Reload the software

MP in this upgrade process refers to the Management Processor on the Management modules. LP refers to the Line Processor on the Interface modules.

Current code inventory

The code inventory is taken using the **`show version`** command. From this output, you will need to take note of the following items:

- Management Module(s)
  - Multi-Service IronWare for IBM r-series Image name (for the slot you are currently running; primary or secondary)
  - Boot Image name
  - Monitor Image name
  - MBRIDGE revision (Management module FPGA)
- Interface Card(s)
  - Multi-Service IronWare for IBM r-series Image name
  - Boot Image name
  - Monitor Image name

The output from the r-series in our setup is shown in Example 14-18.

Example 14-18 r-series show version

```
telnet@RX-8(config)#show version
======================================================================
HW: BigIron RX Router
BI-RX-8-S Backplane (Serial #: XXYYYYYYYY,  Part #: XXYYYYYYYY)
RX-BI-SFM3 Switch Fabric Module 1 (Serial #: XXYYYYYYYY,  Part #:
31523-100A)
FE 1: Type fe200,  Version 2
FE 2: Type fe200,  Version 2
FE 3: Type fe200,  Version 2
=======================================================================
SL M2: RX-BI-MR2 Management Module (High Value) Active (Serial #:
XXYYYYYYYY, Pa
rt #: 31524-000A):
Boot     : Version 2.7.2cT145 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:03:02 labeled as rmb02702c
 (432020 bytes) from boot flash
Monitor  : Version 2.7.2cT145 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:03:02 labeled as rmb02702c
 (432020 bytes) from code flash
IronWare : Version 2.7.2cT143 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:16:42 labeled as rmpr02702c
 (4476358 bytes) from Primary
Board ID : 00 MBRIDGE Revision : 21
916 MHz Power PC processor  (version 8003/0101) 166 MHz bus
512 KB Boot Flash (AM29LV040B), 32 MB Code Flash (MT28F128J3)
2048 MB DRAM
Active Management uptime is 53 days 3 hours 38 minutes 50 seconds
=======================================================================
===
SL 1: RX-BI-2XG 2-port 10GbE Module (Serial #: XXYYYYYYYY, Part #:
TEST000000)
Boot     : Version 2.7.2cT155 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:02:40 labeled as rlb02702c
 (306154 bytes) from boot flash
Monitor  : Version 2.7.2cT155 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:02:40 labeled as rlb02702c
 (306154 bytes) from code flash
IronWare : Version 2.7.2cT157 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:26:26 labeled as rlp02702c
 (2316161 bytes) from Primary
FAP 1 version: 2
FAP 2 version: 2
FAP 3 version: 0
FAP 4 version: 0
660 MHz Power PC processor 440GP (version 8020/0020) 330 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
```

```
512 MB DRAM, 8 KB SRAM, 0 Bytes BRAM
LP Slot 1 uptime is 26 days 3 hours 28 minutes 14 seconds
==========================================================================
SL 3: RX-BI-16XG 16-port 10GbE Module (Serial #: XXYYYYYYYY, Part #:
60-1001337-
01)
Boot     : Version 2.7.2cT155 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:02:40 labeled as rlb02702c
 (306154 bytes) from boot flash
Monitor  : Version 2.7.2cT155 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:02:40 labeled as rlb02702c
 (306154 bytes) from code flash
IronWare : Version 2.7.2cT157 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:26:26 labeled as rlp02702c
 (2316161 bytes) from Primary
FAP 1 version: 3
FAP 2 version: 3
FAP 3 version: 3
FAP 4 version: 3
660 MHz Power PC processor 440GP (version 8020/0020) 330 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 0 Bytes BRAM
LP Slot 3 uptime is 53 days 3 hours 39 minutes 14 seconds
==========================================================================
SL 6: RX-BI-24C 24-port 1 GbE Copper Module (Serial #: XXYYYYYYYY, Part
#: 31521
-001D)
Boot     : Version 2.7.2cT155 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:02:40 labeled as rlb02702c
 (306154 bytes) from boot flash
Monitor  : Version 2.7.2cT155 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:02:40 labeled as rlb02702c
 (306154 bytes) from code flash
IronWare : Version 2.7.2cT157 Copyright (c) 1996-2007 Foundry Networks,
Inc.
Compiled on Nov  5 2009 at 18:26:26 labeled as rlp02702c
 (2316161 bytes) from Primary
FAP 1 version: 2
FAP 2 version: 2
FAP 3 version: 0
FAP 4 version: 0
660 MHz Power PC processor 440GP (version 8020/0020) 330 MHz bus
512 KB Boot Flash (AM29LV040B), 16 MB Code Flash (MT28F640J3)
512 MB DRAM, 8 KB SRAM, 0 Bytes BRAM
LP Slot 6 uptime is 53 days 3 hours 38 minutes 53 seconds
=======================================================================
All show version done
telnet@RX-8(config)#
```

In our example, we have the following code inventory:

- Management modules:
  - Multi-Service IronWare for IBM r-series: rmpr02702c
  - Monitor: rmb02702c
  - Boot: rmb02702c
  - MBRIDGE: 21
- Interface modules:
  - Multi-Service IronWare for IBM r-series: rlp02702c
  - Monitor: rlb02702c
  - Boot: rlb02702c

Version cross-check

We are going to be installing Multi-Service IronWare for IBM r-series R02.7.02a on our device — this will be a downgrade from the currently installed code level.

From the Release Notes for this version, in the section entitled: “Image Files for Multi-Service IronWare R02.7.00a” we see the following requirements:

- Management Module
  - Multi-Service IronWare for IBM r-series: rmb02702a
  - Monitor: rmb02702a
  - Boot: rmb02702a
  - MBRIDGE 21
- Interface modules:
  - Multi-Service IronWare for IBM r-series: rlp02702c
  - Monitor: rlb02702c
  - Boot: rlb02702c

If we compare this list with our current installed code level, we see that all the MBRIDGE FPGA level is at their correct level. We need to upgrade the Application, Boot, and Monitor code using the Unified image.

If we were actually performing a code upgrade (or downgrade, in our case), of course, we only load the code necessary for the upgrade process. However, for the sake of completeness, in this section we go through the upgrade process using the Unified image and MBRIDGE FPGA image.

Multi-Service IronWare for IBM r-series code upgrade

For convenience, Multi-Service IronWare for IBM r-series, Boot, and Monitor images for both the management and interface modules are available as a Unified downloadable package, in our case rx02702a.bin.

To start the upgrade, run the command **`copy tftp image 10.106.3.162 rx02702a.bin lp-boot mp-boot`**. We see the process take place in Example 14-19.

Example 14-19 r-series upgrade example

```
telnet@RX-8#copy tftp image 10.106.3.162 rx02702a.bin lp-boot mp-boot
........................................................................
........................................................................
.................................................................
Download combined image from tftp is done - Start Copying Individual
Images.
1) Copy LP Monitor Image.
Copy to LP monitor and boot flash.
Copy file LP Monitor on MP to file monitor on all LP slots
File Download: LP Monitor (MP) -> monitor (LP 1) is done.
File Download: LP Monitor (MP) -> monitor (LP 3) is done.
File Download: LP Monitor (MP) -> monitor (LP 6) is done.
File download to interface module is done (3 successful)
Save a copy to MP's flash, please
wait..........................................
........................................................................
.........Done
2) Copy LP Application Image.
Copy to LP primary flash.
Copy file LP Application on MP to file primary on all LP slots
File Download: LP Application (MP) -> primary (LP 6) is done.
File Download: LP Application (MP) -> primary (LP 3) is done.
File Download: LP Application (MP) -> primary (LP 1) is done.
File download to interface module is done (3 successful)
Save a copy to MP's flash, please
wait..........................................
........................................................................
...............................................Done
3) Copy MP Monitor Image.
```

```
Copy to MP code
flash...........................................................
...............................................done.
Copy to MP boot
flash...Erasing......Writing....................................
................Done
done.
4) Copy MP Application Image.
Copy to MP primary
flash........................................................
.......................................................................
.......................................................................
......done.
telnet@RX-8#
```

MBRIDGE FPGA upgrade

Upgrading the MBRIDGE FPGA on the management module might not be required on every code release.

To upgrade the MBRIDGE FPGA code, enter the command **copy tftp mbridge *1.2.3.4 mbridge_04000b.xsvf***.

Verify Image Copy

On the r-series, you can check to make sure that the images have been copied correctly by using the **show flash** command. As you can see in Example 14-20, the primary flash partition now holds R02.7.02a.

Example 14-20 r-series show flash after upgrade

```
telnet@RX-8#show flash
====================================================================
Active Management Module (Right Slot)
Code Flash - Type MT28F128J3, Size 32 MB
  o IronWare Image (Primary)
    Version 2.7.2aT143, Size 4476202 bytes, Check Sum d9c7
    Compiled on Sep 29 2009 at 17:18:26 labeled as rmpr02702a
  o IronWare Image (Secondary)
    Version 2.7.3T143, Size 4471787 bytes, Check Sum 9797
    Compiled on Aug 10 2009 at 01:24:06 labeled as rmpr02703b93
  o LP Kernel Image (Monitor for LP Image Type 0)
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
  o LP IronWare Image (Primary for LP Image Type 0)
    Version 2.7.2aT157, Size 2316212 bytes, Check Sum cf64
    Compiled on Sep 29 2009 at 17:28:32 labeled as rlp02702a
  o LP IronWare Image (Secondary for LP Image Type 0)
    Version 2.2.1T157, Size 2472983 bytes, Check Sum cbe2
    Compiled on Jan 25 2006 at 12:18:14 labeled as rlp02201b173
  o Boot-Monitor Image
    Version 2.7.2aT145, Size 432020 bytes, Check Sum 6362
    Compiled on Sep 29 2009 at 17:04:42 labeled as rmb02702a
  o Startup Configuration
    Size 8928 bytes, Check Sum 8bdf
    Modified on 21:21:07 GMT+00 Wed Dec 23 2009

Boot Flash - Type AM29LV040B, Size 512 KB
  o Boot-Monitor Image
    Version 2.7.2aT145, Size 432020 bytes, Check Sum 6362
    Compiled on Sep 29 2009 at 17:04:42 labeled as rmb02702a
====================================================================
Line Card Slot 1
Code Flash: Type MT28F640J3, Size 16 MB
  o IronWare Image (Primary)
    Version 2.7.2aT157, Size 2316212 bytes, Check Sum cf64
    Compiled on Sep 29 2009 at 17:28:32 labeled as rlp02702a
  o IronWare Image (Secondary)
    Version 2.2.1T157, Size 2472983 bytes, Check Sum cbe2
    Compiled on Jan 25 2006 at 12:18:14 labeled as rlp02201b173
  o Boot-Monitor Image
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
Boot Flash: Type AM29LV040B, Size 512 KB
  o Boot-Monitor Image
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
====================================================================
Line Card Slot 3
Code Flash: Type MT28F640J3, Size 16 MB
  o IronWare Image (Primary)
    Version 2.7.2aT157, Size 2316212 bytes, Check Sum cf64
    Compiled on Sep 29 2009 at 17:28:32 labeled as rlp02702a
  o IronWare Image (Secondary)
    Version 2.7.1bT157, Size 2353556 bytes, Check Sum a981
    Compiled on Jun  3 2009 at 18:56:38 labeled as rlp02701b
  o Boot-Monitor Image
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
Boot Flash: Type AM29LV040B, Size 512 KB
  o Boot-Monitor Image
```

```
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
====================================================================
Line Card Slot 6
Code Flash: Type MT28F640J3, Size 16 MB
  o IronWare Image (Primary)
    Version 2.7.2aT157, Size 2316212 bytes, Check Sum cf64
    Compiled on Sep 29 2009 at 17:28:32 labeled as rlp02702a
  o IronWare Image (Secondary)
    Version 2.2.1T157, Size 2472983 bytes, Check Sum cbe2
    Compiled on Jan 25 2006 at 12:18:14 labeled as rlp02201b173
  o Boot-Monitor Image
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
Boot Flash: Type AM29LV040B, Size 512 KB
  o Boot-Monitor Image
    Version 2.7.2aT155, Size 306154 bytes, Check Sum 60d5
    Compiled on Sep 29 2009 at 17:04:18 labeled as rlb02702a
=======================================================================
All show flash done
telnet@RX-8#
```

Product reload

Assuming that the coherence check passed, you can now reload the product with the **reload** command. If you want to perform a hitless upgrade, see “Hitless upgrades” on page 407.

Upgrade verification

Now that our upgrade process is complete, verify the upgrade with another run of **show version**.

Hitless upgrades

The m-series products are capable of being upgraded with minimal disruption to production traffic if the product is equipped with a redundant management modules. Layer 2 switching continues to occur, as well as Layer 3 routing with Graceful OSPF and BGP restart configured. Some functionality is not supported though, including FPGA image upgrades. See the Hardware Installation Guide for additional details.

Before you plan a hitless upgrade, there are conditions that *must* be met.

> **Important:** At the time of writing, the hitless upgrade process on the m-series is supported only if *all* of the following conditions are met:
>
> - You have two management modules installed.
> - You have serial console access to *both* management modules.
> - If you are using OSPF and/or BGP, you have enabled OSPF Graceful Restart and/or BGP Graceful Restart.
> - You do not need to change router configuration during the 1 to 10 minute process.
> - You have not made any changes requiring a software reload.
> - You do not need to upgrade any FPGA images.
> - You are not downgrading code.
> - You meet any other conditions detailed in the release notes.

The process for carrying out this upgrade is quite simple. Using the instructions in the previous section, load new Boot and Monitor images (if necessary), then load the new Multi-Service IronWare for IBM r-series image, but instead of then doing a **`reload`**, use the command **`hitless-reload.`**

15

Interface hardware configuration

In this chapter, we provide information regarding how to physically configure the interfaces on your product. We discuss both Ethernet and SONET interfaces.

Information about IP configuration, VLANs, switching and routing, and so on, is covered in other chapters.

15.1 Ethernet interfaces

Most users will need to perform relatively few configuration operations on Ethernet interfaces; the default suffice under most circumstances. Most, if not all, of the commands are performed at the Interface CONFIG level.

15.1.1 Port Name

You might want to assign a more meaningful name to a port other than “Ethernet 2/5”. This is done with the **port-name** command, as shown in Example 15-1.

Example 15-1 port-name command

```
telnet@m_Series(config-if-e1000-5/1)#show interface brief | include 5/1
5/1   DisabNone      None None  None  No  level0 001b.ed22.84c0
telnet@example(config)#interface ethernet 5/1
telnet@example(config-if-e1000-5/1)#port-name Peter
telnet@m_Series(config-if-e1000-5/1)#show interface brief | include 5/1
5/1   DisabNone      None None  None  No  level0 001b.ed22.84c0 Peter
telnet@m_Series(config-if-e1000-5/1)#
```

15.1.2 Enabling/disabling

On devices running IronWare (s-series, g-series, and x-series), all ports are enabled by default on switches, and disabled by default on routers.

On devices running Multi-Service IronWare (m-series, c-series) and Multi-Service IronWare for IBM r-series, the ports are disabled by default.

It is a good practice to explicitly set the state for a particular port, no matter what kind of device it is.

The commands to do this are **enable** and **disable**. Prefixing the **enable** or **disable** with a **no** has no effect. The command is accepted, but it has no effect.

We show an example of port enabling and disabling in Example 15-2.

Example 15-2 Enabling/disabling a port

```
telnet@m_Series(config-if-e1000-5/1)#show interface brief | include 5/1
5/1  DisabNone     None None  None  No  level0 001b.ed22.84c0 Peter
telnet@m_Series(config-if-e1000-5/1)#
telnet@m_Series(config-if-e1000-5/1)#enable
telnet@m_Series(config-if-e1000-5/1)#
```

```
telnet@m_Series(config-if-e1000-5/1)#show interface brief | include 5/1
5/1   Down None      None None  None  No  level0 001b.ed22.84c0 Peter
[the interface is now “Down” (because it isn’t actually connected) but
it is no longer disabled]
telnet@m_Series(config-if-e1000-5/1)#
telnet@m_Series(config-if-e1000-5/1)#disable
telnet@m_Series(config-if-e1000-5/1)#
telnet@m_Series(config-if-e1000-5/1)#show interface brief | include 5/1
5/1   DisabNone      None None  None  No  level0 001b.ed22.84c0 Peter
telnet@m_Series(config-if-e1000-5/1)#
telnet@m_Series(config-if-e1000-5/1)#no disable
telnet@m_Series(config-if-e1000-5/1)#show interface brief | include 5/1
5/1   DisabNone      None None  None  No  level0 001b.ed22.84c0 Peter
[no disable had no effect on the port’s status, which is still
disabled]
telnet@m_Series(config-if-e1000-5/1)#
```

15.1.3 Speed and duplex (Gigabit Ethernet only)

The switches support auto-sensing, auto-negotiation, and auto-MDI/MDIX detection. While **auto** might be a nice first choice for speed and duplex settings, unfortunately, speed and duplex auto-negotiation might have issues. This is usually the case when one end of the link is set to auto-negotiate, and the other one is hard-coded. Failure to properly autonegotiate can result in connecting at sub-optimal port speeds or even a failure to connect.

The speed and duplex setting for gigabit ports are set with the **speed-duplex** interface command. The following options exist across the product line:

- `10-full`
- `10-half`
- `100-full`
- `100-half`
- `1000-full (master/slave clock negotiation is implied)`
- `1000-half (master/slave clock negotiation is implied)`
- `1000-full-master (s-, g-, x-series only)`
- `1000-full-slave (s-, g-, x-series only)`
- `1000-master (s-, g-, x-series only)`
- `auto`

The 1000-master/1000-full-master/1000-slave-master options forces the interface to 1000-full, and hard-codes the device to own the clock pulse; this option removes all negotiation. (With **1000-full**, the two devices must still negotiate who is in charge of setting the clock pulse.)

At gigabit speeds, UTP ports on the products will attempt to auto-negotiate first, no matter what the individual port is set to, and then only pay attention to your manual speed-duplex settings if the initial auto-negotiation fails.

Speed and duplex restrictions

Depending on platform, various speed and duplex settings might not be available. Note that the <tab> or ? lists all available options on the entire platform, even if only some, or none, of the config options are available. These settings only apply to the 10/100/1000 MbE RJ45 (copper) ports except on the x-series. SFP and other port types are fixed to **`auto`**.

x-series

The 10/1 GbE SFP+ ports on the x-series support either 10 GbE or 1 GbE. The default is 10 GbE. To support a speed of 1 GbE using a 1 GbE transceiver, use the **`speed-duplex 1000`** command to configure the interface speed.

15.1.4 VLAN

With the products covered in this guide, you do not actually set the VLAN for a port from the interface configuration; instead, this is done from the VLAN CONFIG level itself. We cover this in detail in 5.1.1, “Basic VLAN configuration” on page 176.

15.1.5 Jumbo frames

The method for enabling jumbo frames varies by model.

m-series, r-series, and c-series

To enable jumbo frames, enter the global config command **`default-max-frame-size`** *`2345`*. Set this value to your needed IP MTU + 18. The allowed range for this value is 1298 to 9216. The command requires a reload.

IP MTU is set separately; see Chapter 6, “Layer 3 routing” on page 189.

s-series, g-series, and x-series

To enable jumbo frames on these products, enter the config command **`jumbo`**. This command requires a reload.

This enables jumbo frames up to 10240, which can allow an IP packet up to 10222. On s-series and B48G switches with an L3 license, the IP MTU must be configured separately. It can be done with the **`ip mtu`** command either from the Global CONFIG level to configure it globally or within an Interface CONFIG level.

15.2 Packet Over SONET (POS)

Packet Over SONET (POS) interfaces are available on the IBM m-series.

While Packet Over SONET links are becoming less common, many users will still need to configure them for long-haul WAN links.

The parameters for these links are very specific by provider, so the defaults will probably not work for you and it is beyond the scope of this book to discuss all the possible variations.

15.2.1 Enabling/disabling

The POS ports are enabled and disabled in the same way as the Ethernet ports.

15.2.2 Encapsulation type

The default encapsulation type is **`PPP`**. Other available options are **`HDLC`** or **`frame-relay`**. If choosing frame-relay, the default is Cisco-compatible, but you can add **`ietf`** at the end for RFC 1490 compatibility, if needed.

The interface config command to set this is **`encapsulation`** ***`HDLC`***.

15.2.3 Maximum frame size

To set the maximum frame size for all POS interfaces, enter the config command **`pos-default-max-frame-size`** ***`2345`***. The allowed range for this setting is 1284 to 9216. This command requires a reload to take effect. This value cannot be set on an individual interface basis.

The IP MTU is set separately; see Chapter 6, “Layer 3 routing” on page 189.

15.2.4 Other frame relay parameters

Here we discuss the other frame relay parameters.

DTE/DCE

The NetIron products are Data Terminal Equipment (DTE)-only, so the device on the other end must be set up as Data Carrier Equipment (DCE.) There is an explicit interface command for this, **`frame-relay intf-type dte`**, but it is really a placeholder for possible future function, because **`dce`** is not a valid option at this time.

DLCI

The Data Link Circuit Identifier is set with the interface command `frame-relay interface-dlci` *11.*

LMI

The default Local Management Interface (LMI) is `lmi` (Cisco compatible), but your other options are `ansi` or `ccitt`.

15.2.5 Clock source

By default, the NetIron uses an internal clock; use this if you are connecting two NetIron's back-to-back with no network in between. If you are connecting the SONET link to a network with its own clock source, run the `clock line` interface command. (The default is `internal`.)

15.2.6 MTU

The default MTU is 4470 bytes. To change it, use the `mtu 1234` interface command.

15.2.7 Port bandwidth

You can choose to run your POS ports at either 622Mbps or 155Mbps with the `bandwidth 155` interface command.

15.2.8 Frame type

The frame type can be set to either `SDH` or `SONET` using the `pos framing` *sdh* interface command.

15.2.9 Other settings

You can also modify the CRC Length, enable/disable keep-alives, enable/disable ATM scrambling, and configure Automatic Protection Switching (APS). For details on these lesser-used features, see the NetIron Configuration Guide.

15.3 Interface monitoring

The basic command to do a quick check of the status of your interfaces is **show interface brief**. This command will give a one-line summary of each port, along with listing the MAC address. Sample output (from a router) is shown in Example 15-3.

Example 15-3 show interface brief

```
telnet@m_Series>show interface brief

Port  Link L2 State  Dupl Speed Trunk Tag Priori MAC            Name
1/1   DisabNone     None None  None  No  level0 001b.ed22.8400
1/2   DisabNone     None None  None  No  level0 001b.ed22.8401
1/3   DisabNone     None None  None  No  level0 001b.ed22.8402
1/4   DisabNone     None None  None  No  level0 001b.ed22.8403
5/1   Down None     None None  None  No  level0 001b.ed22.84c0
5/2   Down None     None None  None  No  level0 001b.ed22.84c1
5/3   DisabNone     None None  None  No  level0 001b.ed22.84c2
5/4   DisabNone     None None  None  No  level0 001b.ed22.84c3
5/5   DisabNone     None None  None  No  level0 001b.ed22.84c4
5/6   DisabNone     None None  None  No  level0 001b.ed22.84c5
5/7   DisabNone     None None  None  No  level0 001b.ed22.84c6
5/8   DisabNone     None None  None  No  level0 001b.ed22.84c7
5/9   DisabNone     None None  None  No  level0 001b.ed22.84c8
5/10  DisabNone     None None  None  No  level0 001b.ed22.84c9
5/11  DisabNone     None None  None  No  level0 001b.ed22.84ca
5/12  DisabNone     None None  None  No  level0 001b.ed22.84cb
5/13  DisabNone     None None  None  No  level0 001b.ed22.84cc
5/14  DisabNone     None None  None  No  level0 001b.ed22.84cd
5/15  DisabNone     None None  None  No  level0 001b.ed22.84ce
5/16  DisabNone     None None  None  No  level0 001b.ed22.84cf
5/17  DisabNone     None None  None  No  level0 001b.ed22.84d0
5/18  DisabNone     None None  None  No  level0 001b.ed22.84d1
5/19  DisabNone     None None  None  No  level0 001b.ed22.84d2
5/20  DisabNone     None None  None  No  level0 001b.ed22.84d3
9/1   Up   Forward  Full 1G    None  No  level0 001b.ed22.8580
9/2   Down None     None None  None  No  level0 001b.ed22.8581
9/3   Down None     None None  None  No  level0 001b.ed22.8582
9/4   Down None     None None  None  No  level0 001b.ed22.8583
9/5   Down None     None None  None  No  level0 001b.ed22.8584
9/6   Down None     None None  None  No  level0 001b.ed22.8585
9/7   Down None     None None  None  No  level0 001b.ed22.8586
9/8   Down None     None None  None  No  level0 001b.ed22.8587
9/9   Down None     None None  None  No  level0 001b.ed22.8588
```

```
9/10  Down None     None None  None  No  level0 001b.ed22.8589
9/11  Down None     None None  None  No  level0 001b.ed22.858a
9/12  Down None     None None  None  No  level0 001b.ed22.858b
9/13  Down None     None None  None  No  level0 001b.ed22.858c
9/14  Down None     None None  None  No  level0 001b.ed22.858d
9/15  Down None     None None  None  No  level0 001b.ed22.858e
9/16  Down None     None None  None  No  level0 001b.ed22.858f
9/17  Down None     None None  None  No  level0 001b.ed22.8590
9/18  Down None     None None  None  No  level0 001b.ed22.8591
9/19  Down None     None None  None  No  level0 001b.ed22.8592
9/20  Down None     None None  None  No  level0 001b.ed22.8593
mgmt1 Up   Forward  Half 100M  None  Yes level0 001b.ed22.8400

Port  Link L2 State  Dupl Speed Trunk Tag Priori MAC         Name
lb1   Up   N/A        N/A  N/A   None  N/A N/A    N/A
```

show interface brief lists all interfaces: trunks, “real” interfaces, the management port, loopbacks, virtual interfaces, SONET ports, and so on.

show interface *ethernet 9/1* (or whatever interface you want information about) will give you a wealth of statistics and status information. Example 15-4 illustrates **show interface ethernet 9/1** output.

Example 15-4 show interface

```
telnet@m_Series#show interface ethernet 9/1
GigabitEthernet9/1 is up, line protocol is up
  STP Root Guard is disabled, STP BPDU Guard is disabled
  Hardware is GigabitEthernet, address is 001b.ed22.8580 (bia
001b.ed22.8580)
  Configured speed auto, actual 1Gbit, configured duplex fdx, actual
fdx
  Member of VLAN 1 (untagged), port is in untagged mode, port state is
Forwarding
  STP configured to ON, Priority is level0, flow control enabled
  Priority force disabled, Drop precedence level 0, Drop precedence
force disabled
  dhcp-snooping-trust configured to OFF
  mirror disabled, monitor disabled
  Not member of any active trunks
  Not member of any configured trunks
  No port name
  Internet address is 11.1.1.2/30, MTU 1548 bytes, encapsulation
ethernet
  300 second input rate: 26 bits/sec, 0 packets/sec, 0.00% utilization
  300 second output rate: 26 bits/sec, 0 packets/sec, 0.00% utilization
  3179 packets input, 228191 bytes, 0 no buffer
  Received 240 broadcasts, 0 multicasts, 2939 unicasts
  0 input errors, 0 CRC, 0 frame, 0 ignored
  0 runts, 0 giants
  NP received 3179 packets, Sent to TM 3179 packets
  NP Ingress dropped 0 packets
  3179 packets output, 228264 bytes, 0 underruns
  Transmitted 239 broadcasts, 0 multicasts, 2940 unicasts
  0 output errors, 0 collisions
  NP transmitted 3179 packets, Received from TM 3179 packets
telnet@m_Series#
```

15.4 Configuring multiple interfaces

To save on time, you can configure multiple interfaces with the same options at one time. To do this, enter the command **interface** ***ethernet 11/1 to 11/4***.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks publications

For information about ordering these publications, see “How to get Redbooks publications” on page 428. Note that some of the documents referenced here might be available in softcopy only.

- *Introduction to Storage Area Networks*, SG24-5470

Online resources

These websites are also relevant as further information sources:

- IBM System Storage hardware, software, and solutions:

  http://www.storage.ibm.com

- IBM System Networking:

  http://www-03.ibm.com/systems/networking/

- IBM b-type Ethernet switches and routers:

  http://www-03.ibm.com/systems/networking/hardware/ethernet/b-type/index.html

- Brocade Resource Center:

  http://www.brocade.com/data-center-best-practices/resource-center/index.page

- Brocade and IBM Ethernet Resources:

  http://www.brocade.com/microsites/ibm_ethernet/resources.html

- IBM System Storage, Storage Area Networks:

  http://www.storage.ibm.com/snetwork/index.html

How to get Redbooks publications

You can search for, view, or download Redbooks publications, Redpapers publications, Technotes, draft publications, and Additional materials, as well as order hardcopy Redbooks publications, at this website:

ibm.com/redbooks

Help from IBM

IBM Support and downloads:

ibm.com/support

IBM Global Services:

ibm.com/services

Index

IBM b-type Data Center Networking: Product Introduction and Initial Setup

(0.5” spine)
0.475”<->0.875”
250 <-> 459 pages

IBM b-type Data Center Networking

Product Introduction and Initial Setup

Learn about the products in the IBM b-type portfolio

Discover how to quickly perform an initial setup

Read about advanced features

As organizations drive to transform and virtualize their IT infrastructures to reduce costs, and manage risk, networking is pivotal to success. Optimizing network performance, availability, adaptability, security, and cost is essential to achieving the maximum benefit from your infrastructure.

In this IBM Redbooks publication, we address the requirements:

- Expertise to plan and design networks with holistic consideration of servers, storage, application performance, and manageability
- Networking solutions that enable investment protection with performance and cost options that match your environment
- Technology and expertise to design and implement and manage network security and resiliency
- Robust network management software for integrated, simplified management that lowers operating costs of complex networks

IBM and Brocade have entered into an agreement to provide expanded network technology choices with the new IBM b-type Ethernet Switches and Routers, to provide an integrated end-to-end resiliency and security framework.

Combined with the IBM vast data center design experience and the Brocade networking expertise, this portfolio represents the ideal convergence of stren gth and intelligence. For organizations striving to transform and virtualize their IT infrastructure, such a combination can help you reduce costs, manage risks, and prepare for the future. This book is meant to be used along with *IBM b-type Data Center Networking: Design and Best Practices Introduction*, SG24-7786.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7785-00 ISBN 0738434396