

IBM Spectrum Accelerate Reference Architecture

Eyal Abraham
Patrick Pollard
Greg Treantos



Storage



Overview

IBM® Spectrum™ Accelerate, a member of the IBM Spectrum Storage™ family is an agile, software-defined storage solution for enterprise and cloud that builds on the customer-proven and mature IBM XIV® storage software. The key characteristic of Spectrum Accelerate is that it can be easily deployed and run on purpose-built or existing hardware chosen by the customer.

IBM Spectrum Accelerate enables rapid deployment of high-performance and scalable block data storage infrastructure over commodity hardware, either on-premises, or off-premises. Details of IBM Spectrum Accelerate software implementation and installation can be found in *IBM Spectrum Accelerate Deployment, Usage, and Maintenance*, SG24-8267 and associated product guide books.

This IBM Redbooks® publication introduces Spectrum Accelerate and describes the planning and preparation that are essential for a successful deployment of the solution, including a step-by-step approach. This approach can use either a graphical user interface-based method or a simple command line interface-based procedure.

This technical publication describes a Spectrum Accelerate implementation with Lenovo 3650M5 server hardware. Spectrum Accelerate can be deployed and run on purpose-built hardware chosen by the customer. This document is provided as a reference-architecture and as an example for such a general-purpose hardware configuration.

The IBM Spectrum Accelerate solution that is described in this document provides a total of 184 TB of usable capacity. The tested solution achieved peak performance of 440,000 IOPs and 3.4 GBps throughput.

Solution components

The solution consists of a 9-node Spectrum Accelerate cluster built from Lenovo 3650M5 servers, 10 GbE BNT, and general-purpose 1 GbE network switches. The solution also includes three UPS systems for power conditioning and protection, as depicted in Figure 1.

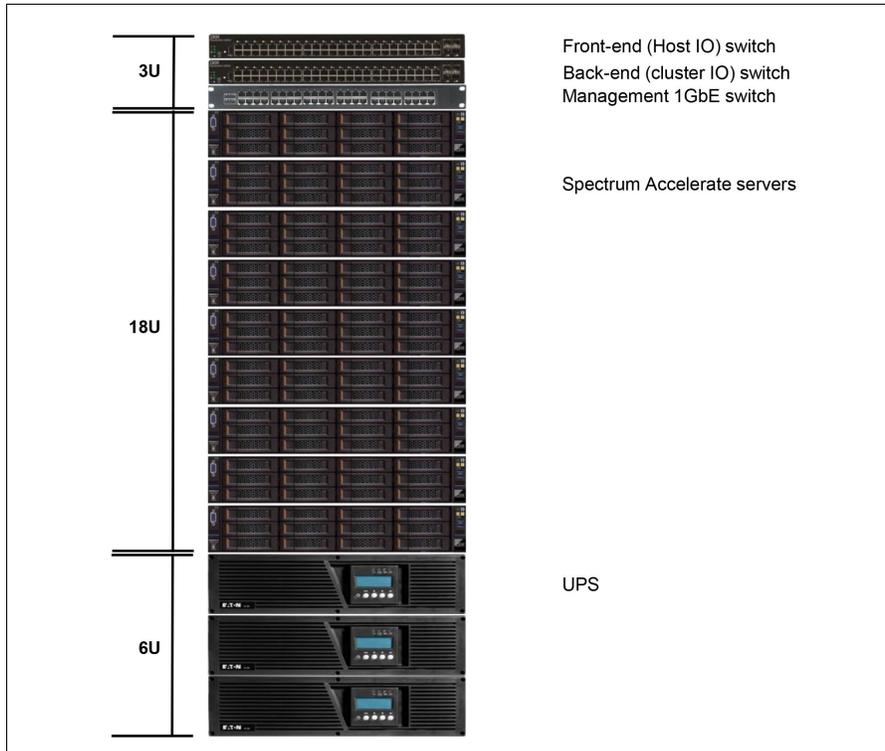


Figure 1 Physical rack installation

The solution is housed in a standard 19"/42U rack. The rack front and back views are shown in Figure 2.

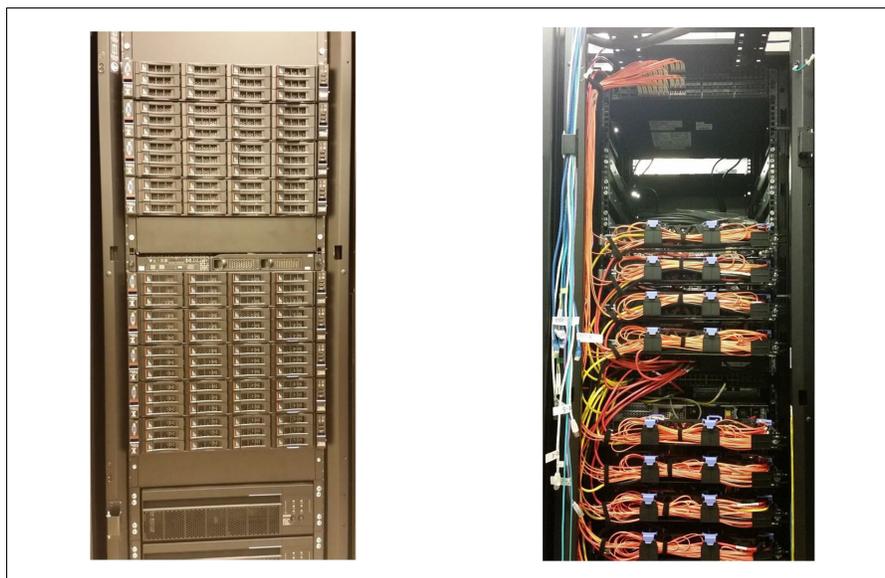


Figure 2 Rack front and back views

Figure 3 shows an illustration of the Spectrum Accelerate GUI.

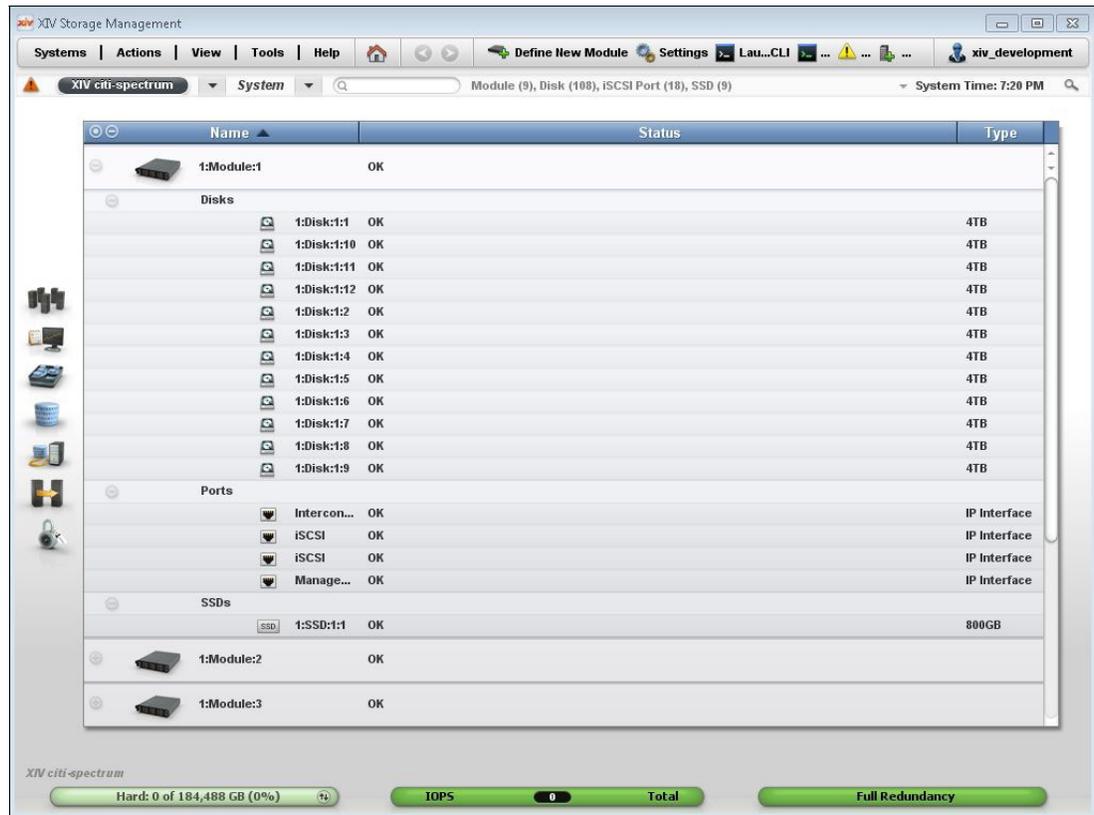


Figure 3 System GUI

Servers

The solution uses nine Lenovo 3650M5 servers that are configured with 14 drive bays, each carrying a 3.5" drive carriers, as illustrated in Figure 4 on page 6. Each server contains 12 x 4 TB NL-SAS HDD, a single 800 GB SSD, and a single 300 GB SAS HDD for hypervisor boot.

The network connectivity of the server includes two dual-port 10 GbE NICs. These four ports are used for front-end and back-end connectivity in a high availability configuration (see "Network" on page 6 for details).

The CPU and memory configuration were selected to support IBM Spectrum Accelerate release 11.5.0.d and accommodate future solutions such as a converged infrastructure, or ability to use more DRAM. The following is a high-level bill of material (see "Bill of Material" on page 12 for a detailed bill of material):

- ▶ 9 x Lenovo 3650M5
 - 2 x Intel Xeon Processor E5-2620 v3 6C 2.4GHz
 - 256 GB TruDDR4 DRAM

Each server (IBM Spectrum Accelerate node) is equipped with:

- 1 x 300 GB SAS HDD
- 12 x 4 TB SAS HDD
- 1 x 800 GB SSD
- Emulex VFA5 2x10 GbE SFP+ PCIe
 - 2 x Dual-port 10 GbE NICs

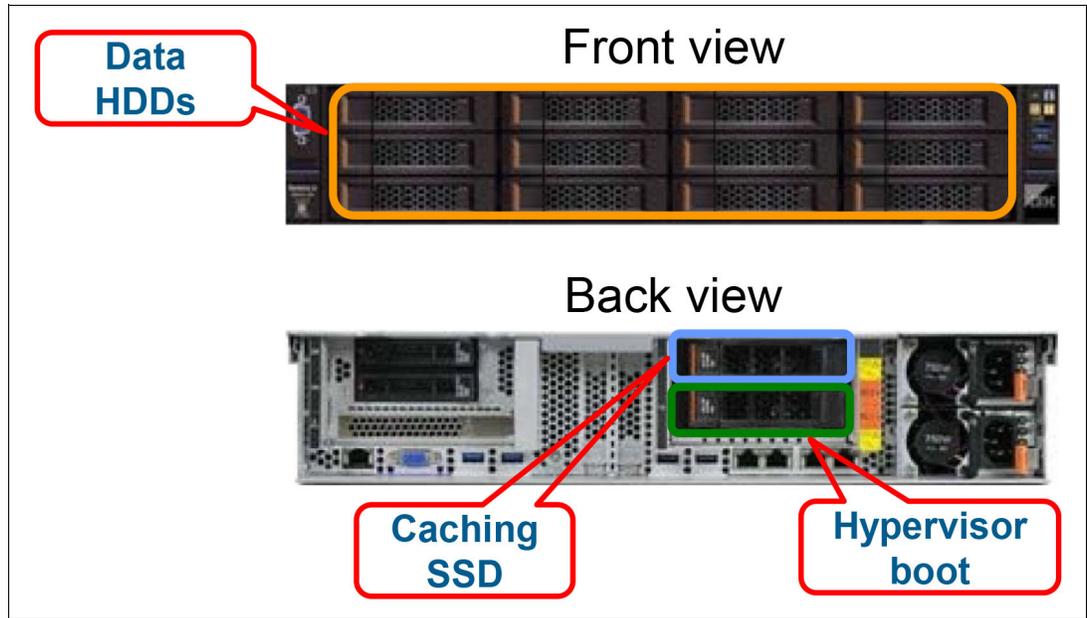


Figure 4 Lenovo 3650M5 server view

Network

The network configuration consists of three networks: One 1-GbE IP network for management of the solution components, and two 10-GbE networks for the data path. Each one of these networks uses a separate switch:

- ▶ 2 x 10 GbE Switch G8264:
 - One switch for back-end traffic and one for front-end host connectivity
 - High availability configuration between NIC ports
- ▶ 1 GbE Switch for management access

The 10 GbE networks can be converged in a single switch. However, to build a high-availability setup, two switches are required and the data-path must be split between the two (see “Data-Path networks: iSCSI host attach and Interconnect” on page 8 for details).

The reference architecture used a single VLAN for all networks with four subnets: Management, back-end Interconnect, and two for iSCSI host connectivity.

Network configuration and NIC/port assignments are done by using the ESX Hypervisor vSwitch. For vSwitch configuration, see the Spectrum Accelerate Installation Guide and ESX literature.

Figure 5 illustrates the solution vSwitch configuration for the Management network.

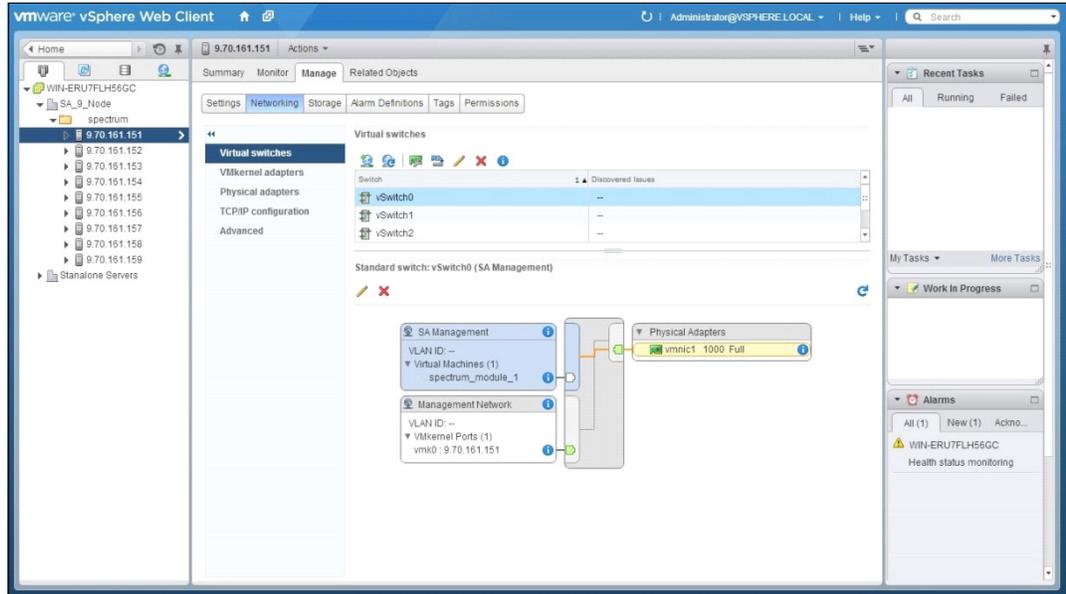


Figure 5 vSwitch-0 configuration: Management network

Figure 6 illustrates the solution vSwitch configuration for the Interconnect.

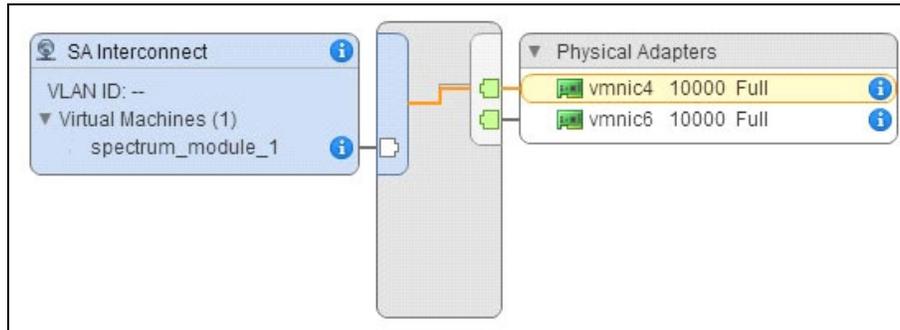


Figure 6 vSwitch-1 configuration: Interconnect

Figure 7 illustrates the solution vSwitch configuration for the host attach.

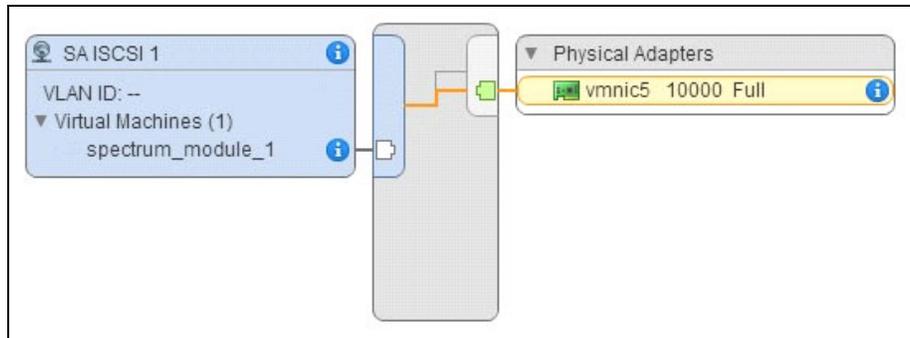


Figure 7 vSwitch-2 configuration: iSCSI host attach

Management network

The management network provides management connectivity to all systems components. This network provides connectivity to an external management workstation for installation, management, and maintenance.

Because the management network is not in the system's data path, it can be built with a single switch. However, a failure of this single switch disables management access to the solution.

This solution takes advantage of the server integrated management module (IMM) interface. This interface is useful during buildup on the servers in the cluster and for various server-level maintenance activities.

The management and the IMM networks can use separate switches or be converged in a single switch. To save on network IP addresses, the IMM network can be designated in a private 10.x.x.x network.

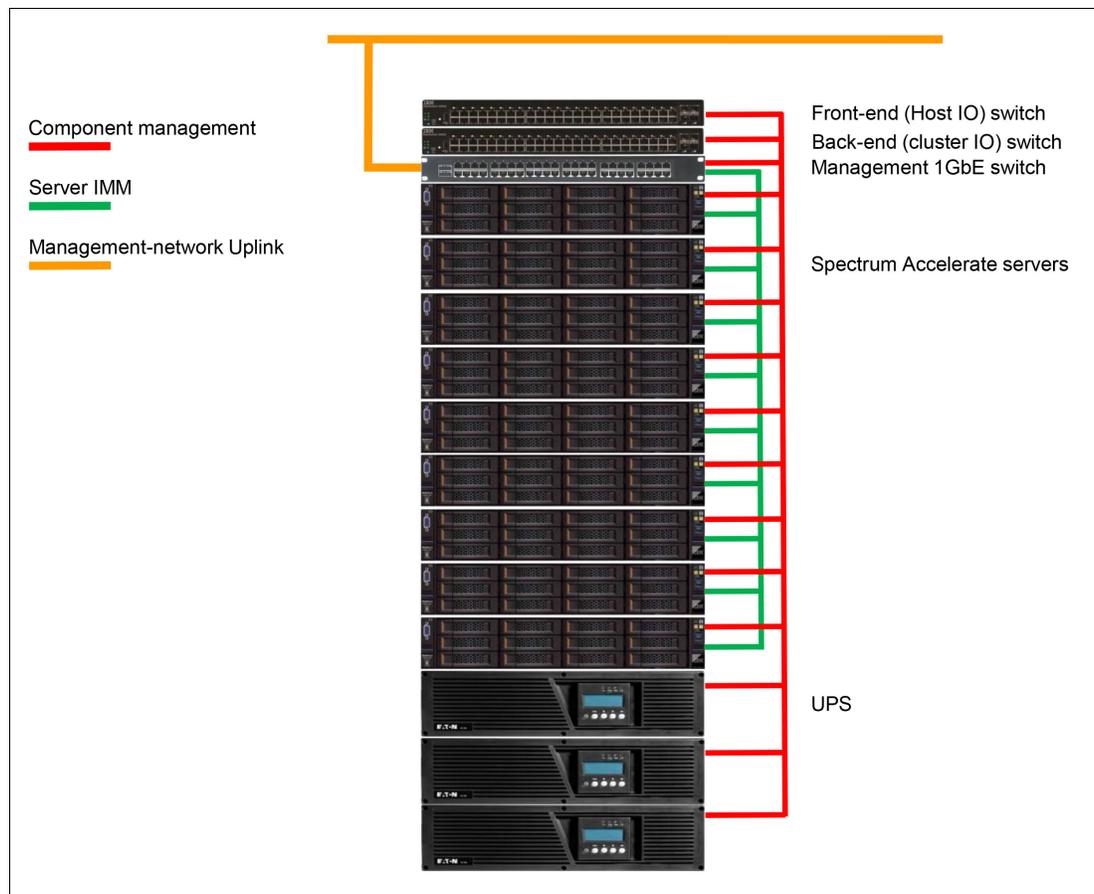


Figure 8 1 GbE Management network infrastructure

Data-Path networks: iSCSI host attach and Interconnect

The Data-Path networks are configured across two 10-GbE G8264 switches. The switches are configured with SFP optics, but DAC cables can be used for the internal rack if you are interested in reducing costs.

Spectrum Accelerate data-path has two types of network traffic: Interconnect IO (“Back-End”) and iSCSI Host IO (“Front-End”). Interconnect IO traffic includes data and control passed

between the nodes, and iSCSI Host IO is the iSCSI traffic that carries client-host and application data.

This solution is configured to provide high availability through network redundancy. The redundancy is achieved by separating the front-end and back-end networks, and using two switches and four NIC ports on each node, as shown in Figure 9:

- ▶ iSCSI host attach (Front-end): Configured to use port-1 on each of the two NICs and split across two switches.
- ▶ Interconnect (Back-end): Configured to use port-2 on each NIC and split across two switches.

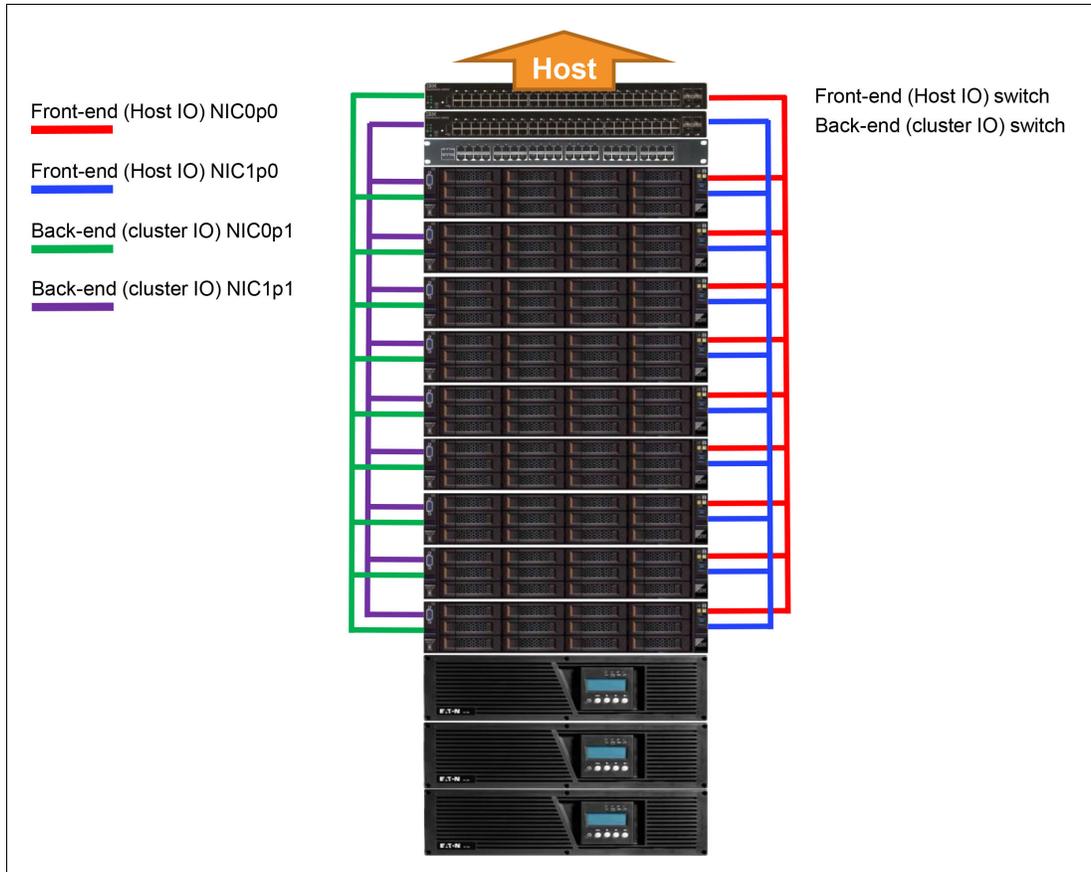


Figure 9 10 GbE Data-Path network infrastructure

Power

The Spectrum Accelerate solution power plant was modeled according to a modified XIV Storage System UPS power protection scheme. This solution uses three UPS units to power all the components in the rack. All servers have dual power supplies and are cabled in groups of three to the three UPS units as shown in Figure 10 on page 10.

10 GbE network switches are equipped with a dual power supply connection and are distributed between the UPS systems to reduce the probability of a dual network switch loss (also see “Data-Path networks: iSCSI host attach and Interconnect” on page 8).

Each UPS is connected to a single-phase 30A 220 V for data center power.

UPS devices, including the APC SURT6000RMLX device used in this reference architecture, include a status monitoring port. An external management workstation can be used to monitor the UPS state, but this configuration was beyond the scope of this solution and was not implemented.

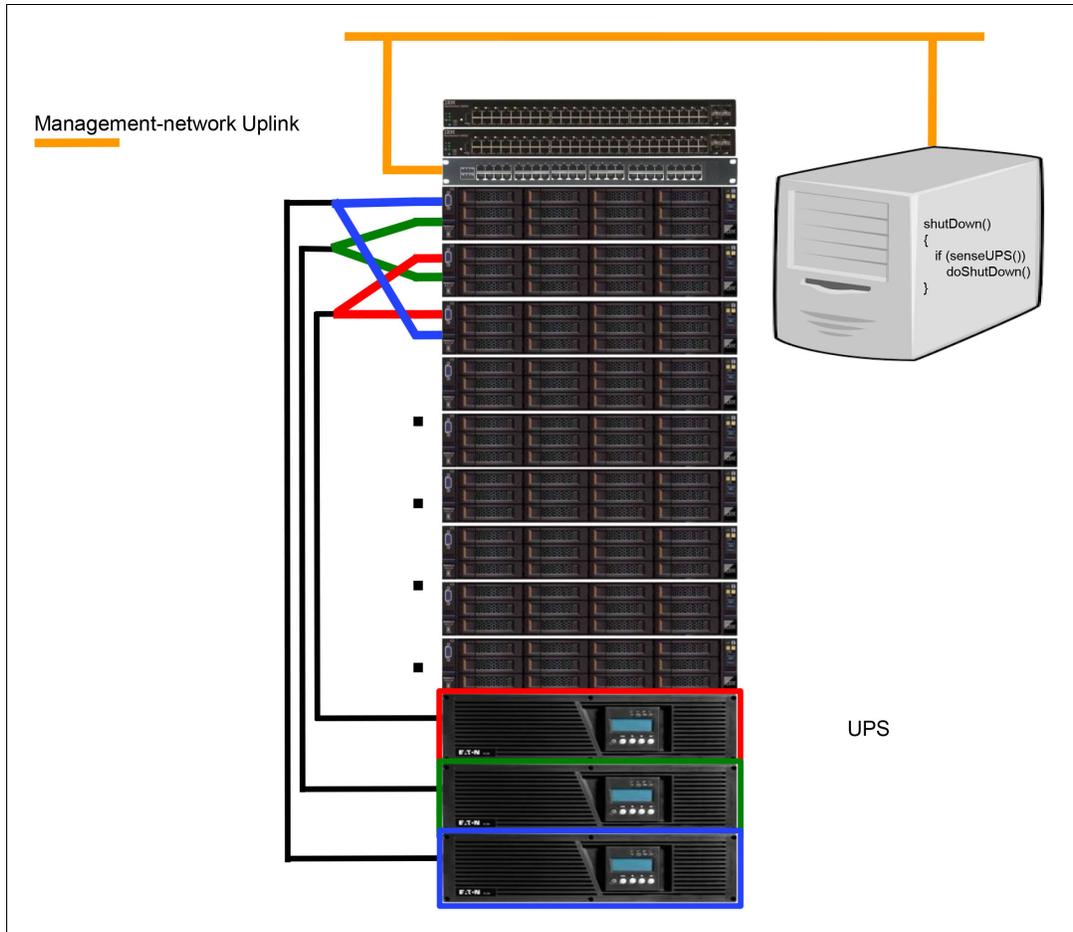


Figure 10 UPS power and shut-down management

Preferred practices

This section lists preferred practices for installing and configuring the solution

Server preparation

Before hypervisor installation, download and install the Lenovo Bootable Media Creator server tool for the x3650M5. Use this tool to create the fix packs to be used with the bootable media creator.

The tool will prompt you for machine type and fixes. Use the tool to collect all relevant fixes for the server, including but not limited to the following list:

- ▶ IMM fixes
- ▶ BIOS fixes
- ▶ HDD firmware

- ▶ SAS controller firmware
- ▶ iSCSI adapter firmware

Hypervisor installation

Install the ESXi 5.5 update 2 from the IBM Support Portal at:

<https://www-947.ibm.com/support/entry/portal/support>

The Lenovo hypervisor build includes all drives relevant for the Lenovo systems. After hypervisor installation, verify that all driver levels are at the latest levels. If not, search the Support Portal for the latest ESXi driver.

Network configuration

Verify that the switch firmware is at the latest levels.

Verify that the MTU setup of the Interconnect network matches the MTU defined by your network administrator. The MTU setup in this reference architecture was set to 9000. For more information about switch setup and configuration, see the *IBM BNT RackSwitch G8264 ISCLI Command Reference Guide* at:

<http://www-01.ibm.com/support/docview.wss?uid=isg3T7000297>

Deployment workstation

Prepare a separate Windows or Linux workstation that will act as a deployment and management host. This workstation will contain Spectrum Accelerate. In addition to the Spectrum Accelerate license and code obtained through IBM Passport Advantage®, you must also download and install the XIV GUI from IBM Fix Central at:

<https://www.ibm.com/support/fixcentral/>

The workstation needs to have access to the ESX management network and the VM management network

Documentation

For more information about the IBM Spectrum Accelerate solution and the other topics described in this document, see these resources:

- ▶ Spectrum Accelerate documentation is available in the IBM Knowledge Center at:
https://www.ibm.com/support/knowledgecenter/STZSWD_11.5.0/xiv_sds_kc_welcome.html
- ▶ IBM Spectrum Accelerate: Deployment, Usage, and Maintenance, SG24-8267:
<http://www.redbooks.ibm.com/abstracts/sg248267.html?open>

Bill of Material

This section provides the bill of material for the servers, Ethernet switches, and UPS devices:

- Configuration for the nine Lenovo 3650M5 servers is listed in Figure 11.

<u>Model</u>	<u>Part Num</u>	<u>Description</u>	<u>QTY</u>
5462AC1		System x3650 M5, 2x6c,256GB	OEM
	A5FQ	System x3650 M5 PCIe Riser 1 (2 x8 FH/FL + 1 x8 ML2 Slots)	1
	5977	Select Storage devices - no IBM-configured RAID required	1
	A5FX	System x Enterprise 2U Cable Management Arm (CMA)	1
	A5FF	System x3650 M5 12x 3.5" Base without Power Supply	1
	A5UT	Emulex VFA5 2x10 GbE SFP+ PCIe Adapter for IBM System x	1
	A5EU	System x 750W High Efficiency Platinum AC Power Supply	2
	A5EJ	Addl Intel Xeon Processor E5-2620 v3 6C 2.4GHz 15MB 1866MHz 85W	1
	9206	No Preload Specify	1
	A5B7	16GB TruDDR4 Memory (2Rx4, 1.2V) PC4-17000 CL15 2133MHz LP RDIMM	16
	A5EY	System Documentation and Software-US English	1
	A1ML	IBM Integrated Management Module Advanced Upgrade	1
	A5ED	Intel Xeon Processor E5-2620 v3 6C 2.4GHz 15MB Cache 1866MHz 85W	1
	A40Q	Emulex VFA5 ML2 Dual Port 10GbE SFP+ Adapter for IBM System x	1
	A3YY	N2215 SAS/SATA HBA for IBM System x	1
	A5GE	x3650 M5 12x 3.5" HS HDD Assembly Kit	1
	A5VD	IBM 2TB 7.2K 6Gbps NL SATA 3.5" G2HS 512e HDD	1
	A56J	S3700 800GB SATA 3.5" MLC HS Enterprise SSD for IBM System x	1
	A5VQ	IBM 4TB 7.2K 12Gbps NL SAS 3.5" G2HS 512e HDD	12
	6311	2.8m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
	A5FV	System x Enterprise Slides Kit	1
	A5FC	System x3650 M5 WW Packaging	1
	ASFE	Notice for Advanced Format 512e Hard Disk Drives	1
	A5G1	System x3650 M5 EIA Plate	1
	9297	2U Bracket for Emulex 10GbE Virtual Fabric Adapter for IBM System x	2
	A2HP	Configuration ID 01	1
	A5V5	System x3650 M5 Right EIA for Storage Dense Model	1
	A5FM	System x3650 M5 System Level Code	1
	A5FH	System x3650 M5 Agency Label GBM	1
	A5EA	System x3650 M5 Planar	1
	ASQB	System x3650 M5 Rear 2x 3.5" HDD Label (Cascaded)	1
	A5G5	System x3650 M5 Riser Bracket	1
	A5FT	System x3650 M5 Power Paddle Card	1
	A5GL	System x3650 M5 Rear 2x 3.5" HDD Kit (Cascaded)	1
	5374CM1	HIPO : Configuration Instruction	1
	A2HP	Configuration ID 01	1
	A2JX	Controller 01	1
	A46U	N2215 SAS/SATA HBA Placement	1

Figure 11 Lenovo 3650M5 servers configuration

- Configuration for the two 10 GbE switches is listed in Figure 12.

Part No.	Description	Quantity
Misc	G8264	
7309G64	IBM System Networking RackSwitch G8264 (rear to front)	1
00D6185	IBM System Networking Adjustable 19in 4 Post Rail Kit	1
39Y7937	1.5m, 10A/100-250V, C13 to IEC 320-C14 Rack Power Cable	2
46C3447	IBM SFP+ SR Transceiver	48 (**)

Figure 12 10 GbE switch configuration

- The three UPS devices are APC model SURT6000RMLX.

Authors

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

Eyal Abraham is the manager of the XIV Storage Solution Engineering team at IBM. He has been with IBM XIV for six years, joining the team as a storage specialist following the acquisition of XIV. Eyal and his team have worldwide responsibility for designing, testing, and proving IT solutions with XIV and Spectrum Accelerate for IBM clients. The team works with clients, IBM Business Partners, and IBM employees. Eyal has held various engineering, professional services and consulting positions at IBM as well as at EMC. Eyal holds an MBA from Northeastern University and an Electrical Engineering degree from Tel-Aviv University, Israel.

Patrick Pollard is an IBM XIV World Wide Storage Solution Architect based in Littleton, Massachusetts. Patrick has been with IBM and XIV since 2008, and was one of the original XIV team members since the acquisition. Patrick has been involved with the storage industry and various storage-based technology companies since 1980. As World Wide Technical Specialist and Subject Matter Expert, Patrick has led the introduction of many of XIV's key features such as Remote Mirroring, Hyper Scale Mobility, and most recently Real time Compression and Spectrum Accelerate. Patrick has expertise in SAN, Fibre Channel and iSCSI protocols, IBM Mainframe FICON® and ESCON as well as SCSI. Patrick is also involved in the delivery of IBM Flash-based solutions and has worked on many internal qualifications extending the support of IBM Flash in the marketplace.

Greg Treantos is a part of the XIV Solutions Engineering team at IBM. He has been with XIV for four years and IBM for 25 years. Currently, Greg is the XIV Solutions Engineering team's Subject Matter Expert for Spectrum Accelerate. Greg holds a BS in Education from Fitchburg State University and a B.S. in Information Technology from the University of Massachusetts Lowell.

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

Bertrand Dufrasne
IBM ITSO

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

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 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>

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 grant 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.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

© Copyright International Business Machines Corporation 2015. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

This document REDP-5260-00 was created or updated on August 31, 2015.

Send us your comments in one of the following ways:

- ▶ Use the online **Contact us** review Redbooks 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 U.S.A.

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:

FICON®	IBM®	Redbooks (logo)  ®
IBM Spectrum™	Passport Advantage®	XIV®
IBM Spectrum Storage™	Redbooks®	

The following terms are trademarks of other companies:

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

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

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



REDP-5260-00

ISBN 0738454419

Printed in U.S.A.

Get connected

