Redbooks
ibm.com/redbooks

SAN and Fabric Resiliency Best Practices for IBM b-type Products

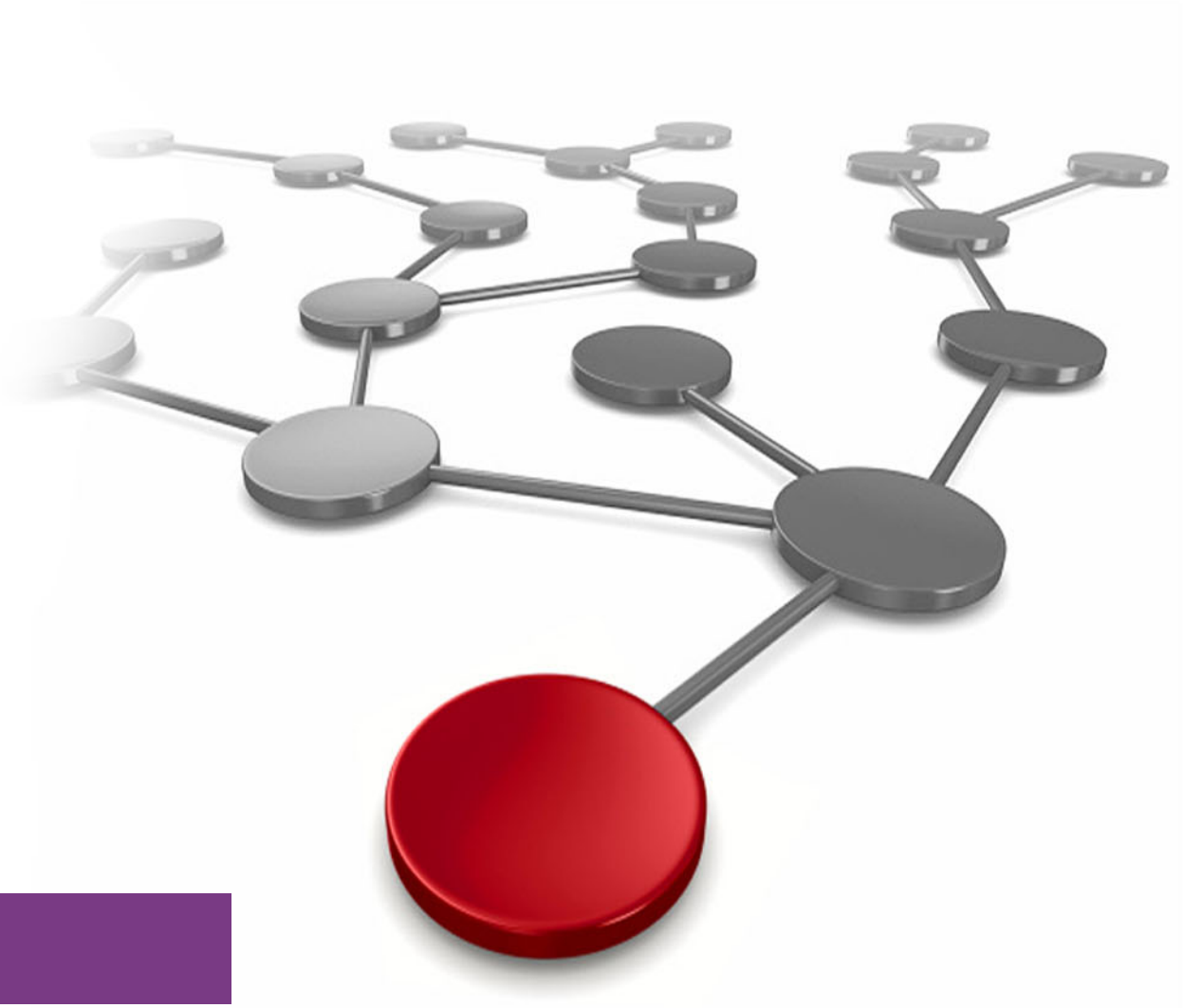
Ian Mac Quarrie

Jim Blue

David Green

David Lutz

Gavin O'Reilly



Storage

IBM

Redpaper

Fabric best practices using IBM SAN-b switches

This IBM® Redpaper® publication describes best practices for deploying and using advanced Broadcom Fabric Operating System (FOS) features to identify, monitor, and protect Fibre Channel (FC) SANs from problematic devices and media behavior.

FOS: This paper primarily focuses on the FOS command options and features that are available since version 8.2 with some coverage of new features that were introduced in 9.0.

This paper covers the following recent changes:

- SANnav

  Broadcom introduced SANnav in late 2019 as the replacement for BNA as the management interface. IBM support for SANnav started with SANnav 2.1.0a release in August 2020.

  IBM will continue support for BNA until April 30, 2023. For the next several years, we will see an increasing number of SANnav deployments as the number of BNA deployments declines. Because BNA will be around for the foreseeable future, this paper provides feature implementation details and examples for both management interfaces.

  Note: SANnav supports any switch that is running FOS 7.4.x and above.

- Fabric Performance Impact Notification

  Broadcom introduced a feature starting with FOS9.0 called *Fabric Performance Impact Notification* (FPIN). FPIN allows the fabric to notify end devices of congestion and physical layer issues that exist within the fabric.

  End devices that support FPIN can use fabric notifications to alter their behavior in ways that mitigate the effect fabric conditions can have on host applications, such as throttling down the amount of I/O that is sent or stopping the use of a path altogether.

  The behavior of end devices that support FPIN is vendor-specific.

  Note: FPIN is supported on Gen6 and Gen7 switches.

Introduction

Faulty or incorrectly configured devices, misbehaving hosts, and faulty or substandard FC media can significantly affect the performance of FC fabrics and the applications that they support. In most real-world scenarios, these issues cannot be corrected or mitigated within the fabric. Rather, the behavior must be addressed directly.

However, with the correct knowledge and capabilities, the fabric often can identify and in some cases, mitigate or protect against the effects of these misbehaving components to provide better fabric resiliency. This document concentrates specifically on Brocade Fabric Vision features (and related capabilities) that help provide optimum fabric resiliency.

For more information about the features that are described in this publication, see the product documents that are suitable for your FOS release.

The following documents are available at th Documentation tab of this Broadcom web page:

- *SAN Design and Best Practices*
- *Fabric OS Administrator's Guide*
- *Fabric OS Command Reference Manual*
- *Fabric OS Monitoring and Alerting Policy Suites Configuration Guide*
- *Fabric OS Flow Vision Configuration Guide*

For more information, see the *SANnav Management Portal Users Guide*.

Factors that affect fabric resiliency

The following common types of abnormal behavior originate from fabric components or attached devices:

- Faulty media (fiber-optic cables and Small Form-factor Pluggables (SFPs)/optics)

  Faulty media can cause frame loss because of excessive cyclic redundancy check (CRC) errors, Forward Error Correction (FEC) errors, invalid transmission words, and other conditions, which can result in I/O failure and application performance degradation.

- Misbehaving devices, links, or switches

  Occasionally, a condition arises where a device (server or storage array) or link (inter-switch link [ISL]) behaves erratically and causes disruptions in the fabric. If not immediately addressed, this situation might result in severe stress on the fabric.

- Congestion

  Congestion is caused by latencies or insufficient link bandwidth. End devices that do not respond as quickly as expected can cause the fabric to hold frames for excessive periods, which can result in application performance degradation or, in extreme cases, I/O failure.

- Credit loss

  Credit loss occurs when the receiving end of a link fails to acknowledge a request to transmit a frame because no buffers are available to receive the frame.

Faulty media

In addition to high-latency devices causing disruptions to data centers, fabric problems are often the result of faulty media. Faulty media can include bad cables, SFPs, extension equipment, receptacles, patch panels, improper connections, and so on.

Media can fault on any SAN port type and fail, often unpredictably and intermittently, which makes it even harder to diagnose. Faulty media that involves server/host and storage device ports (`F_Ports`) results in an impact to the end device that is attached to the `F_Port`, and to devices that are communicating with this device.

Failures on ISLs or `E_Ports` can result in an even greater impact. Many flows (host and target pairs) can simultaneously traverse a single `E_Port`. In large fabrics, this impact can be hundreds or thousands of flows.

If a media failure involves one of these links, it is possible to disrupt some or all of the flows that use the path. Severe cases of faulty media, such as a disconnected cable, can result in a complete failure of the media, which effectively brings a port offline.

This situation typically is easy to detect and identify. When it occurs on an `F_Port`, the effect is specific to flows that involve the `F_Port`. `E_Ports` typically are redundant; therefore, severe failures on `E_Ports` typically result only in a minor drop in bandwidth because the fabric automatically uses redundant paths.

Also, error reporting that is built into FOS readily identifies the failed link and port, which allows for simple corrective action and repair. With moderate cases of faulty media, failures occur, but the port can remain online or transition between online and offline.

This situation can cause repeated errors, which can occur indefinitely or until the media fails. When these types of failures occur on `E_Ports`, the result can be devastating because repeated errors can affect many flows, which can result in significant effects on applications that last for prolonged durations.

These failures include the following types of signatures:

- CRC errors on frames
- Invalid Transmission Words (includes encoder out errors)
- State Changes (ports going offline or online repeatedly)
- Credit loss; that is, complete loss of credit on a virtual channel (VC) on an `E_Port` prevents traffic from flowing on that VC, which results in frame loss and I/O failures for devices that use the VC

Misbehaving devices

Another common class of abnormal behavior originates from high-latency end devices (host or storage). A high-latency end device is one that does not respond as quickly as expected; therefore, it causes the fabric to hold frames for excessive periods. This situation can result in application performance degradation or in extreme cases, I/O failure. Common examples of moderate device latency include disk arrays that are overloaded and hosts that cannot process data as fast as requested.

For example, misbehaving hosts become more common as hardware ages. Bad host behavior often is caused by defective host bus adapter (HBA) hardware, bugs in the HBA firmware, and problems with HBA drivers.

Storage ports can produce the same symptoms because of defective interface hardware or firmware issues. Some arrays deliberately reset their fabric ports if they are not receiving host responses within their specified timeout periods.

Severe latencies are caused by badly misbehaving devices that stop receiving, accepting, or acknowledging frames for excessive periods. However, with the correct knowledge and capabilities, the fabric often can identify (and in some cases, mitigate or protect against) the effects of these misbehaving components to provide better fabric resiliency.

Congestion

Congestion occurs when the traffic being carried on a link exceeds its capacity. Sources of congestion can be links, hosts, or storage that is responding more slowly than expected. Congestion typically is because of fabric latencies or insufficient link bandwidth capacity.

As FC link bandwidth increased from 1 to 16 Gbps, instances of insufficient link bandwidth capacities radically decreased. Latencies, especially device latencies, are the major source of congestion in today's fabrics because of their inability to promptly return buffer credits to the switch.

Device-based latencies

A device that is experiencing latency responds more slowly than expected. The device does not return buffer credits (through `R_RDY` primitives) to the transmitting switch fast enough to support the offered load, even though the offered load is less than the maximum physical capacity of the link that is connected to the device.

Figure 1 shows the condition in which a buffer backup on ingress port 6 on B1 causes congestion upstream on S1, port 3. When all available credits are exhausted, the switch port that is connected to the device must hold more outbound frames until a buffer credit is returned by the device.

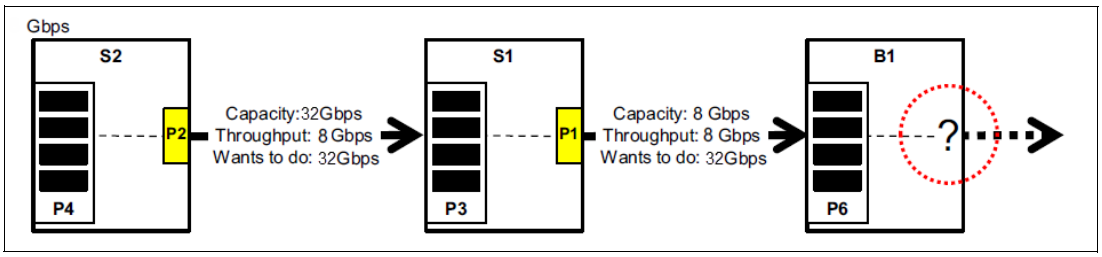


Figure 1 Device latency example

When a device does not respond in a timely fashion, the transmitting switch is forced to hold frames for longer periods, which results in high buffer occupancy, which results in the switch lowering the rate at which it returns buffer credits to other transmitting switches. This effect propagates through switches (and potentially multiple switches, when devices attempt to send frames to devices that are attached to the switch with the high-latency device), and ultimately affects the fabric.

Figure 2 shows how latency on a switch can propagate through the fabric.

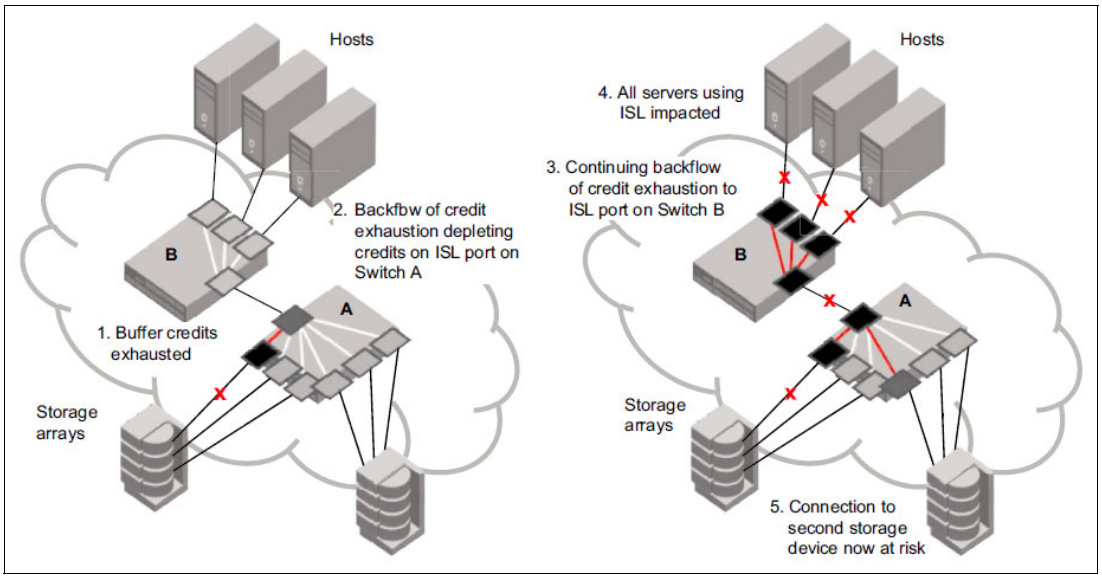


Figure 2 Latency on a switch can propagate through the fabric

Note: The effect on the fabric (and other traffic flows) varies based on the severity of the latency that is exhibited by the device. The longer the delay that is caused by the device in returning credits to the switch, the more severe the problem.

Moderate device latencies

Moderate device latencies from the fabric perspective are defined as those latencies that are not severe enough to cause frame loss. If the time between successive credit returns by the device is between a few hundred microseconds to tens of milliseconds, the device exhibits mild to moderate latencies because this delay typically is not enough to cause frame loss. This situation does cause a drop in application performance, but typically does not cause frame drops or I/O failures.

The effect of moderate device latencies on host applications might still be profound, based on the average disk service times that are expected by the application. Mission-critical applications that expect average disk service times of, for example, 10 ms, are severely affected by storage latencies in excess of the expected service times. Moderate device latencies traditionally were difficult to detect in the fabric.

Advanced monitoring capabilities that are implemented in Broadcom ASICs and FOS made these moderate device latencies much easier to detect by providing the following information and alerts:

- Switches in the fabric generate Fabric Performance Impact (FPI) alerts if FPI is enabled on the affected ports
- Elevated `tim_txcrd_z` counts on the affected `F_Port`, so the `F_Port` where the affected device is connected
- Potentially elevated `tim_txcrd_z` counts on all `E_Ports` that are carrying the flows to and from the affected `F_Port`/device

Note: Consider the following points:

- The `tim_txcrd_z` is defined as the number of times that the port was polled and that the port was unable to transmit frames because the transmit buffer-to-buffer credit (BBC) was zero. The purpose of this statistic is to detect congestion or a device that is affected by latency. This parameter is sampled at intervals of 2.5 microseconds, and the counter is incremented if the condition is true.

  Each sample represents 2.5 microseconds of time with zero Tx BBC. `tim_txcrd_z` counts are not an absolute indication of significant congestion or latencies and are just one of the factors in determining whether real latencies or fabric congestion are present. Some level of congestion is to be expected in a large production fabric, and is reflected in `tx_crd_z` counts. The Broadcom FPI feature was introduced to remove uncertainty around identifying congestion in a fabric.
- The `tim_latency_vc` is a Broadcom Gen5 Condor3 ASIC counter that measures the latency time that a frame incurs in the transmit queue of its corresponding VC. The purpose of this statistic is to directly measure the frame transmit latency of a switch port. Each unit of the counter value represents 250 nanoseconds of latency. The Broadcom FPI feature uses this counter to enhance the detection of devices introducing latency into
  the fabric.

Severe device latencies

Severe device latencies result in frame loss, which triggers the host Small Computer System Interface (SCSI) stack to detect failures and to retry I/Os. This process can take tens of seconds (possibly as long as 30 - 60 seconds), which can cause a noticeable application delay and potentially results in application errors. If the time between successive credit returns by the device is in excess of 100 ms, the device is exhibiting severe latency.

When a device exhibits severe latency, the switch is forced to hold frames for excessively long periods (on the order of hundreds of milliseconds). When this time becomes greater than the established timeout threshold, the switch drops the frame (per FC standards). Frame loss in switches is also known as *Fibre Channel Class 3* (*C3*) *discards* or *timeouts*.

Because the effect of device latencies often spreads through the fabric, frames can be dropped because of timeouts. This issue is not limited to only on the `F_Port` to which the misbehaving device is connected, but also on `E_Ports` carrying traffic to the `F_Port`. Dropped frames typically cause I/O errors that result in a host retry, which can result in significant decreases in application performance.

The implications of this behavior are compounded and exacerbated by the fact that frame drops on the affected `F_Port` (device) result not only in I/O failures to the misbehaving device (which are expected), but also on `E_Ports`, which might cause I/O failures for unrelated traffic flows involving other hosts (and typically are not expected).

Latencies on ISLs

Latencies on ISLs often are the result of back pressure from latencies elsewhere in the fabric. The cumulative effect of many individual device latencies can result in slowing the link. The link might be producing latencies if it is a long-distance link with distance delays or too many flows exist that use the same ISL.

Although each device might not appear to be a problem, the presence of too many flows with some level of latency across a single ISL or trunked ISL can become a problem. Latency on an ISL can ripple through other switches in the fabric and affect unrelated flows.

FOS can provide alerts and information indicating possible ISL latencies in the fabric through one or more of the following items:

- Switches in the fabric generate FPI Alerts if FPI is enabled on the affected ports
- C3 transmit discards (`er_tx_c3_timeout`) on the device `E_Port` or `EX_Port` that carries the flows to and from the affected `F_Port` or device
- `Broadcom` Monitoring Alerting Policy Suite (MAPS) alerts, if they are configured for C3 timeouts
- Elevated `tim_txcrd_z` counts on the affected `E_Port`, which also might indicate congestion
- C3 receives discards (`er_rx_c3_timeout`) on `E_Ports` in the fabric containing flows of a high-latency `F_Port`

Credit loss

Buffer credits are a part of the FC flow control and the mechanism that Fibre Channel connections use to track the number of frames that are sent to the receiving port. Whenever a frame is sent, the credit count is reduced by one. When the sending port runs out of credits, it cannot send more frames to the receiving port. When the receiving port successfully receives a frame, it tells the sending port that it has the frame by returning an `R_RDY` primitive.

When the sending port receives an `R_RDY`, it increments the credit count. Credit loss occurs when the receiving port does not recognize a frame (usually because of bit errors); therefore, it does not return an `R_RDY` or the sending port does not recognize the `r_rdy` (usually because of link synchronization issues).

Because FC links are never perfect, the occasional credit loss can occur. However, it becomes an issue only when all available credits are lost. Credit loss can occur on external and internal FC links. When credit loss occurs on external links, often it is caused by faulty media. Credit lost on internal ports often is associated with jitter, which in most cases is adjusted for by the internal adapter firmware.

The switch automatically tries to recover from a complete loss of credit on external links after 2 seconds by issuing a link reset. For the switch to perform automatic recovery from internal link credit loss, the Credit Loss Detection and Recovery feature must be enabled.

High-performance networks

With the use of low-latency Solid State Drives (SSD) and Flash controllers, the performance of the SAN becomes critical to achieving the full performance potential from those technologies. Eliminating latency from the SAN requires a level of planning and consideration that is often above what is necessary for traditional enterprise class storage, given the nominal operating ranges of those devices.

Poorly constructed and maintained SANs can add latency to the SCSI exchange completion times to varying degrees. This extra latency often can go undetected, or be considered insignificant for “spinning disk” subsystems, because it is often a small percentage of the response time those devices are capable of achieving. This response time can be in the 10s to 100s of milliseconds. This issue is not true of SSD/Flash storage, where the latency contribution from suboptimal SAN conditions can easily equal or exceed the capable response time for those technologies.

The Fabric Resiliency Best Practices that are described in this paper are especially critical because they pertain to maintaining a high-performance SAN. However, in addition to those practices, SAN design considerations also must be made about the use of mixed-speed devices and ISLs.

Mixed-speed SANs

It is generally required for multiple device speeds to exist in the SAN to enable the technology to be refreshed from one generation to the next. For that reason, the existence of mixed-speed devices cannot be avoided. However, mixed-speed devices that span more than one generation of technology should be avoided.

For example, mixing 4 Gb and 8 Gb devices generally is acceptable; however, mixing 4 Gb and 16 Gb is not acceptable. The speed matching that is required to accommodate these large speed differentials introduces latency and potential congestion points that can significantly degrade the performance and stability of SAN.

ISLs and multi-hop ISLs

Many flows between servers and storage, or storage-to-storage devices, must flow across the ISLs. Because of this requirement, ISLs are notorious for introducing latency into the transmission flows. The size and more importantly, the number of ISLs that is required between switches now must consider response time requirements and bandwidth requirements.

With storage devices getting into the submillisecond response times, ISLs must ensure that credits always are available so that frames are not delayed. This need might require multiple ISL trunks instead of fewer larger bandwidth trunks.

If frames must traverse multiple switches to reach their destination, delays can be introduced with each hop that is required between the source and destination switches. Multi-hop ISLs should never be used, except for being used for migration purposes on a temporary basis.

Where strict performance requirements exist, the use of ISLs for access to SSD/Flash should be avoided altogether.

Best practices: Consider the following best practices:

- Avoid introducing devices to the SAN that span more than one generation of technology.
- Avoid traversing ISLs when accessing SSD or Flash for high-performance use cases.

Designing resiliency into the fabric

This document is not intended to cover the general set of design considerations that are required for designing a Storage Area Network (SAN). However, a set of technologies must be considered to ensure that the fabric is resilient by design. This section includes preferred practices for each of the following areas:

- Hardware and FOS
- Virtual fabrics
- Flow management
- Routing policies
- Credit recovery tools
- Inter-switch link trunking
- Peer zoning
- Using a meaningful naming convention
- Dynamic port naming

For more information about architecture, topology, and capacity planning for a SAN, see *SAN Design and Best Practices*.

Hardware and FOS

To ensure that fabrics can maintain the highest availability, they must be running on supported hardware and current firmware levels. As hardware ages, components can become marginal and require the best in recovery retry logic that is provided by the latest firmware levels. When hardware reaches its end of life, several functions are disabled in the firmware, which also directly affects maintaining a high available fabric.

A balancing act always exists between running the latest firmware levels and running a stable, proven firmware level and managing resources to perform firmware upgrades. IBM recommends running the Broadcom target path firmware, which we recommend checking at least twice a year and planning for one upgrade a year.

For more information, see the following resources:

- *Brocade Software: Software Release Support and Posting Matrices*
- Brocade Product End-of-Life web page

Virtual fabrics

Virtual fabrics are not a new concept and have been around for a while. They also were the factory default setting on switches for many years. A *virtual fabric* is one or more logical switches that are connected with Inter Switch Links (ISLs) to create a fabric. Each logical switch that is connected must use the same FID (Fabric ID).

Although the use of logical switches does not create a failure boundary between logical switches because they use the same physical hardware, it does allow the different switches to provide logical isolation between different workloads. Each fabric has its own set of tables, such as zoning and name server lookup tables. Change notifications are limited to the logical switch.

It is common for distributed (open) workloads to use logical switches to separating things, such as tape and disk workloads, or production and test workloads FICON® workloads typically run on a physical switch. With FOS 9 FICON, workload now must be in a logical switch.

Best practice: Use virtual switches to provide logical isolation between different types of workloads.

Flow management

What was needed to allow fabrics to move frames quickly and efficiently changed over the years. In the early years, link speed was the main barrier, but faster and faster affordable links solved that issue. Later, it was the amount of bandwidth between devices that drove designs to use multiple links between switches and their devices, and between the switches in the fabric.

These days, managing the flows is the path a group of frames for a sequence use to get to the target end device. You now have high speed links and multiple paths; therefore, when it comes to flow management, many factors that can effect it that management, such as the following examples:

- Over subscription, which occurs when too many hosts request being sent to a single storage device
- Congestion because of the speed that request data is returned is faster than the source can absorb the data
- Congestion that occurs because too many sources request data that the internal structure can supply
- Congestion because an end device takes a pause or stops returning frame acknowledgments (buffer credits)
- Link quality, where a link in the path becomes marginal and drops frames or slows down

The following tools are available to help monitor and manage flows:

- FPI monitors frame flows at each egress port.
- Slow Drain Device Quarantine (SDDQ) moves a flow into a lower class of service so that it affects overall fabric performance less.
- Port Toggling disables or enables a port to force link resets to clear congestion.
- Port Fencing can disable a port to remove a link from the fabric.
- FPIN allows the fabric to notify a host that an issue exists with a path so the host can alert which paths it chooses to send flows across.

When designing large or high speed fabrics, special attention must be paid to the number and speed of the links between switches, the speed of the host devices versus the storage devices, and the number of host devices sending and requesting data to a set of storage ports.

These tools are an effective way to monitor the design and alert users when adjustments to that design are required.

Best Practice: Consider the following best practices:

- Avoid the use of devices with different connection speeds.
- Manage the number of subscribing to device ports**.**

Routing policies

The routing policy determines the route or path frames take when traversing the fabric. The following routing policies are available:

- Default exchange-based routing (EBR)
- Port-based routing (PBR)
- Device-based routing (DBR)

Open systems FCP fabrics

Exchange Based Routing is always the preferred routing policy for FCP fabrics.

FICON fabrics

Before 2013, cascaded IBM FICON configurations supported only static PBR across ISLs. In this case, the ISL (route) for a specific port was assigned statically based on a round-robin algorithm at fabric login (FLOGI) time. PBR can result in some ISLs being overloaded.

In mid-2013, IBM z systems added support for DBR, which spread the routes across ISLs based on a device ID hash value. With the IBM z13® release in mid-2015, IBM added FICON Dynamic Routing (FIDR), which supports Brocade EBR to improve load balancing for cascaded FICON across ISLs.

For more information about the prerequisite z13 driver levels, adapter features, storage, and FOS levels to support FIDR, see the white paper *FICON Dynamic Routing (FIDR): Technology and Performance Implications*, WP102651.

FICON cascaded configurations with z13 and all other appropriate prerequisites should use EBR. All other FICON cascaded configurations should use DBR.

For more information about the FICON Dynamic Routing feature, see *Get More Out of Your IT Infrastructure with IBM z13 I/O Enhancements*, REDP-5134.

Note: Consider the following points:

- FICON should be EBR (if IBM z/OS® and z System supports FIDR) regardless of whether it is a FICON/FCP intermix. z System must be DBR if it does not support FIDR.
- As a best practice, use default exchange-based routing.

Enabling the routing policy

Enable the suitable routing policy that is based on the environment that the fabric supports.

Enable EBR by using the Advanced Performance Tuning Policy (**`aptpolicy`**) command.

Note: EBR is the default routing policy.

Example 1 shows the **`aptpolicy`** command that is used to set the EBR policy.

Example 1 The aptpolicy command for exchange-based routing

```
DCX1_Default:FID128:dlutz> aptpolicy 3
Policy updated successfully.

DCX1_Default:FID128:dlutz> aptpolicy
 Current Policy: 3

 3 : Default Policy
 1: Port Based Routing Policy
 2: Device Based Routing Policy (FICON support only)
 3: Exchange Based Routing Policy
```

Enable EBR policy for switches that support only FICON and the z/OS systems do not support FIDR.

Example 2 shows the **aptpolicy** command that is used to set the DBR policy.

Example 2 The aptpolicy command for device-based routing

```
DCX1_Default:FID128:dlutz> aptpolicy 2
Policy updated successfully.

DCX1_Default:FID128:dlutz> aptpolicy
 Current Policy: 2

 3 : Default Policy
 1: Port Based Routing Policy
 2: Device Based Routing Policy (FICON support only)
 3: Exchange Based Routing Policy
```

Credit recovery tools

FC traffic uses credit-based flow control in which each side of a connection provides a number of buffers. These buffers are advertised to the partner side of a connection as credits, and indicates how many frames can be outstanding.

Credits are replenished when transmissions are successful and acknowledged. In rare error scenario situations, credit-based flow control acknowledgments are not sent or not received, which leads to a credit loss condition. If the problem that causes this failure is persistent, credit loss can result in a stall of traffic.

FC credit-based recovery applies to external switch ports and back-end ports (ports that are connected to the core blade or core blade back-end ports) that are used for traffic within a switch. Traffic stalls on these internal back-end ports can have a wide effect, especially when they affect virtual circuits of an ISL. Broadcom (Brocade) introduced enhanced credit recovery tools to mitigate this type of problem. These tools can be enabled to automatically reset back-end ports when a loss of credits is detected on internal ports.

The following main choices are available for how the recovery can proceed when enabled:

- An escalating recovery based on the results of a single link reset only (**onLrOnly**).
- A threshold-based approach that uses multiple link resets (**onLrThresh**).

When used with the **onLrOnly** option, the recovery mechanism takes the following escalating actions:

1. When it detects credit loss, it performs a link reset and logs a RASlog message (RAS Cx-1014).
2. If the link reset fails to recover the port, the port reinitializes. A RASlog message is generated (RAS Cx-1015). The port reinitialization does not fault the blade.
3. If the port fails to reinitialize, the port is faulted. A RASlog message (RAS Cx-1016) is generated.
4. If a port is faulted and there are no more online back-end ports in the trunk, the core blade is faulted. (The port blade is always faulted.) A RASlog message is generated (RAS Cx-1017).

When used with the **onLrThresh** option, recovery is attempted through repeated link resets and a count of the link resets is kept. If the threshold of more than the configured threshold value (by using the **-lrthreshold** option) per hour is reached, the blade is faulted (RAS Cx-1018). Regardless of whether the link reset occurs on the port blade or on the core blade, the port blade always is faulted.

Best practice: Enable the credit tools with the `onLrOnly` recovery option.

Enabling the Credit Recovery Tool

Enable the Credit Recovery Tool with the link reset only (LROnly) option. Options changed with the different FOS releases, and some options are now hard-coded into the FOS. Other options do not apply to new hardware platforms.

Example 3 shows the **creditrecovmode** command to enable credit tools and display the credit tools settings for FOS V7.3.

Example 3 The creditrecovmode command for FOS V7.3

```
DCX1_SANA:FID16:dlutz> creditrecovmode --cfg on onLrOnly
DCX1_SANA:FID16:dlutz> creditrecovmode --fe_crdloss on

DCX1_SANA:FID16:dlutz> creditrecovmode --show
Internal port credit recovery is Enabled with LrOnly
LR threshold (not currently activated): 2
Fault Option (not currently activated): EDGEBLADE
C2 FE Complete Credit Loss Detection is Enabled
```

Example 4 shows the **creditrecovmode** command to enable credit tools and display the credit tools settings for FOS V7.4 and FOS V8.0.

Example 4 The creditrecovmode command for FOS V7.4 and V8.0

```
DCX1_SANA:FID16:dlutz> creditrecovmode --cfg on onLrOnly
DCX1_SANA:FID16:dlutz> creditrecovmode --fe_crdloss on
DCX1_SANA:FID16:dlutz> creditrecovmode --be_crdloss on
DCX1_SANA:FID16:dlutz> creditrecovmode --be_losync on

DCX1_SANA:FID16:dlutz> creditrecovmode --show
Internal port credit recovery is Enabled with LrOnly
LR threshold (not currently activated): 2
Fault Option (not currently activated): EDGEBLADE
C2 FE Complete Credit Loss Detection is Enabled
```

Example 5 shows the **creditrecovmode** command to enable credit tools and display the credit tools for FOS V8.1.

Example 5 The creditrecovmode command for FOS V8.1

```
F48a_Default:FID128:dlutz> creditrecovmode --cfg on onLrOnly
F48a_Default:FID128:dlutz>
F48a_Default:FID128:dlutz> creditrecovmode --be_crdloss on
F48a_Default:FID128:dlutz>
F48a_Default:FID128:dlutz> creditrecovmode --be_losync on
F48a_Default:FID128:dlutz>
F48a_Default:FID128:dlutz> creditrecovmode --show
Internal port credit recovery is Enabled with LrOnly
Back end port Loss of Sync's Link Reset is Enabled with LrOnly
```

Inter-switch link trunking

Trunking optimizes the use of bandwidth by allowing a group of ISLs to merge into a single logical link, which is called a *trunk group*. Traffic is distributed evenly and in order over this trunk group, which achieves greater performance with fewer links. Within the trunk group, multiple physical ports appear as a single port, which simplifies management.

Trunking improves system reliability by maintaining in-order delivery of data and avoiding I/O retries if one link within the trunk group fails.

Trunking provides excellent protection from credit lost on ISLs. If credit loss occurs on an ISL, frames continue to flow by using the other link until the switch can detect the credit loss (typically 2 seconds) and perform a link reset to recover the credits.

More IT environments are relying on server virtualization technologies that can share host adapter connections. Specifically, N_Port ID Virtualization (NPIV) allows many clients (servers, guest, or hosts) to use a single physical port on the SAN.

Each of these communications paths from server (virtual or otherwise) is a data flow that must be considered when planning for how many interswitch links are needed. These virtualized environments often lead to a situation where many data flows from the edge switches exist, which can lead to frame-based congestion if there are not enough ISL or trunk resources.

To avoid frame-based congestion in environments where many data flows exist between switches, it is better to create several two-link trunks than one large trunk with multiple links. For example, it is better to have two 2-link trunk groups than one 4-link trunk group.

Best practice: Use multiple trunked ISLs between switches in a fabric.

Peer zoning

Zoning is the method that is used to control which devices in a fabric are allowed to communicate (pass frames) with each other.

It is important to have a zoning methodology that allows only the intended devices to communicate with each other. Some zoning methods allow servers to talk to other servers, or storage devices to talk to other storage devices. Although these devices should ignore this type of traffic, resources are used during probing. In some cases, issues can exist because two devices that are not intended to communicate attempt to establish a communication path.

As of this writing, the most common zoning method is the initiator target zoning method in which a zone contains one initiator (host HBA) and one target (one storage port). With today's devices where servers have multiple HBAs and storage devices have multiple ports that are connected to provide more data paths for performance and redundancy, the number of zones that is required to establish this level of connectivity can be massive. It also introduces greater potential for errors in host to storage connections.

To reduce the number of zones, it is common to see a server HBA and multiple storage ports in a single zone, which allows for storage-to-storage communication. However, most storage devices recognize this issue and ignore the connection.

A newer zoning method is available that is called *peer zoning*. In this method, a zone can contain one or more HBA ports that are tagged as initiators with several storage ports that are tagged as target ports. When this type of zone is activated, the switch does not allow initiator-to-initiator communication or storage-to-storage communication. This method allows a single zone; for example, to contain a storage device and all of the host HBAs that must communicate with that storage.

Peer zoning is implemented in Broadcom switches with a new zone type called a *peer zone*.

When a zone is created, you add the **`--peerzone`** parameter to the **`zoneadd`** command. When you add members to the zone, you identify target ports by using the `-principal` parameter and initiator ports by using the `-member` parameter.

Figure 3 shows a sample setup with IBM SAN Volume Controller storage that communicates with ESX servers by using peer zoning.

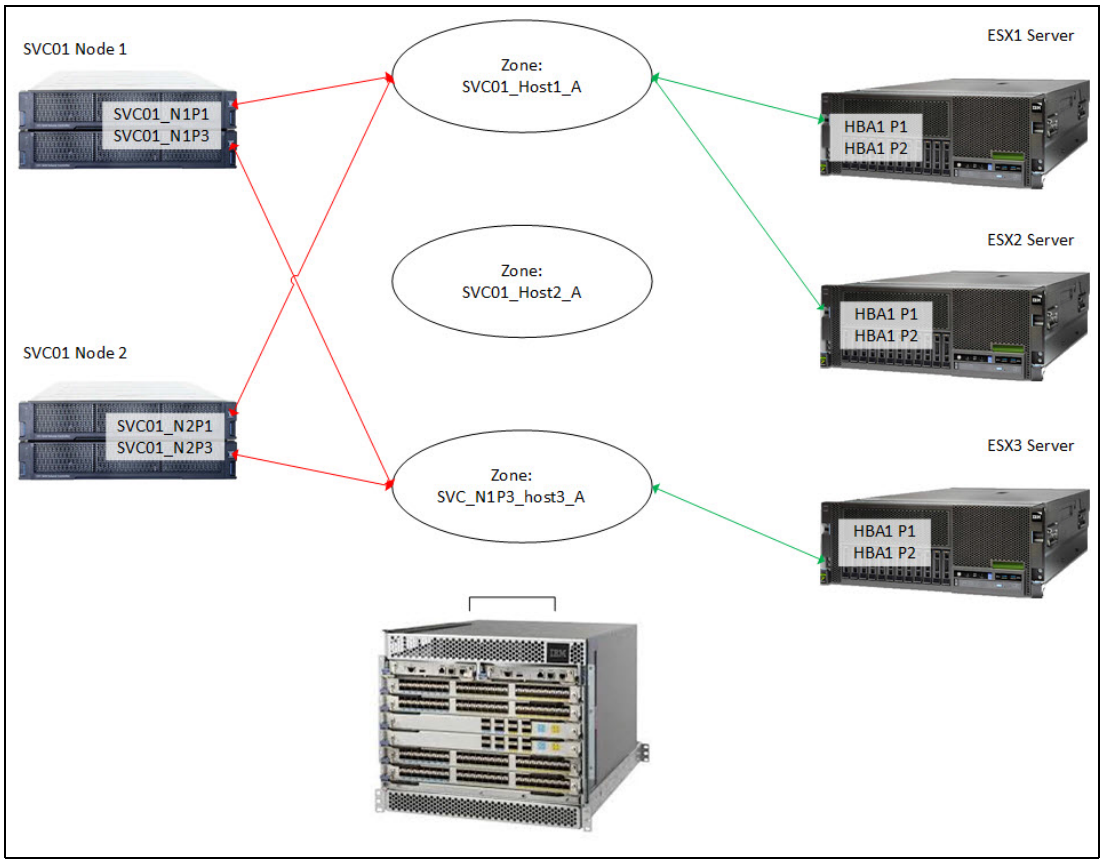


Figure 3 Using peer zoning

Example 6 shows the creation and addition of zones.

Example 6 Creating and adding zones

```
** Create storage peer zones for Host1 and Host3 connection groups.
zonecreate --peerzone SVC01_Host1_A -principal "SVC01_N1P1;SVC01_N2P1"
zonecreate --peerzone SVC01_Host3_A -principal "SVC01_N1P3;SVC01_N2P3"

** Servers ESX1, ESX3 to the HOST1 connection group.
zoneadd --peerzone SVC01_Host1_A -member "ESX1_HBA1_P1;ESX2_HBA1_P1"

** Servers ESX3 to the HOST3 connection group.
zoneadd --peerzone SVC01_Host3_A -member "ESX3_HBA1_P1"
```

Best practice: Use peer zoning for large fabrics.

Using a meaningful naming convention

The use of a strict and well-thought-out naming convention is critical to the reliable operation and serviceability of the environment.

> **Note:** Implementing and maintaining a meaningful naming convention is an important disaster prevention option that is available and requires no software to control. It provides administrators the ability to visually determine the site and detail of the device.

The naming convention can vary depending on the needs and architecture. The naming convention should be designed in a user-friendly fashion, and be consistent and documented.

User-friendly alias names ensure that zone members can be understood at a glance, and configuration errors can be minimized.

User-friendly Switch names ensure that no possible confusion exits when connecting and configuring the switch.

The naming convention can contain the following information:

`Location_FabricName_DeviceGroup_Devicename_Port/Domain`

Where:

- `Location` can contain DR to qualify a Disaster Recovery site, the room name, and the rack name.
- `FabricName` should contain the name of that fabric.
- `DeviceGroup` can mention the type of application or service, and identify the device as test or production.
- `Devicename` is the name of the device.
- `Port/Domain` is the description of the port or the Domain ID of the switch.

  Example 7 shows an example alias for different devices.

Example 7 Alias for a storage device

```
Alias Example for SVC device:
  NY_Green_SAPProd_DS5020-1_C1P2

    NY: New York
    Green: the name of the fabric
    SAPProd: This is an SAP production instance
    DS5020-1: This is a DS5020 number 1
    C1P2: The card and the port

Alias Example for SVC device:
  NY_Green_SAPProd_SVC1_IOGRP0_1
  DR_Blue_SAP2_SVC_N1S3_2

Alias for a server:
  NY_Green_backup_TSM02_fc3
  NY_blue_SAPProd _ ESX11_hba1

Naming convention for a Switch:
  NY_Green_Prod_DCX11_030
  NY_Green_Test_DCX12_031
```

- Alias for a zone

  This information differs from the other scheme because it can contain the name of multiple devices, or groups of devices, that are included in the same zone, as shown in the following example:

  `NY_Green_SAPProd _ESX11_FC1-TAPE33_Zone`

- Alias for a config

  A good practice is to name the config with the name of the fabric, and the version and date it was activated, as shown in the following example:

  `NY_Green_v3.7 _20180801_config`

Zone config versioning can help for fallback scenario and troubleshooting purposes.

Adding the word zone or alias also is useful to avoid confusion, especially when configuring zones by using the CLI.

Best practice: Use aliases and make all of the names meaningful.

Dynamic port naming

Every port on a Broadcom FC switch has a port name that by default is a slot number port number. The port name is displayed in many of the switch event messages, MAPS alert, and Network Advisor dashboards.

The use of a more meaningful port name makes these messages and dashboards more meaningful. It also makes identifying external devices that are causing fabric problems easier and quicker to identify.

The problem is that manually setting port names to more meaningful names is labor-intensive, and typically done only with scripts to set the port name to the alias name of the attached device.

With FOS V7.4, Broadcom introduced dynamic port names that dynamically set the port name to `<switch name>.<port type>.<port index>.<alias name>`. Dynamic port name is enabled by using the **`configure`** command and setting the dynamic port name to `on`.

In FOS V8 and higher, enhancements were made to allow configuring the dynamic port name by using any of the following fields:

- Switch Name
- Port Type
- Port Index
- F_Port Alias
- FDMI Host name
- Remote Switch Name
- Slot/Port Number

Example 8 shows examples of dynamically and manually set port names.

Example 8 Examples of dynamically and manually set port names

```
ANC_DCX1 > switchshow -portname
 103    2   39   20:67:00:05:1e:d0:b5:05    SANC_DCX1.E_PORT.103
 133    7    5   20:85:00:05:1e:d0:b5:05    CLSS14_HBA3
 134    7    6   20:86:00:05:1e:d0:b5:05    SANC_DCX1.F_PORT.134.(null)
```

```
 135    7    7   20:87:00:05:1e:d0:b5:05    DS5300_B2
 134    7    6   20:86:00:05:1e:d0:b5:05    SANC_DCX1.F_PORT.134.(null)
 145    7   17   20:91:00:05:1e:d0:b5:05    SANC_DCX1.F_PORT.145.XIV3_M5P
 147    7   19   20:93:00:05:1e:d0:b5:05    SANC_DCX1.F_PORT.147.DS4800_B2
 192    8    0   20:c0:00:05:1e:d0:b5:05    SANC_DCX1.(none).192
 208    8   16   20:d0:00:05:1e:d0:b5:05    SANC_DCX1.F_PORT.208.(null)
```

Example 9 shows setting and displaying the dynamic port name.

Example 9 Setting and displaying the dynamic port name

```
Brocade_def:FID128:admin > portname -d "S.T.I.F.A.R"
Brocade_def:FID128:admin > switchshow -portname

Index Port    PortWWN                    Name
==================================================
   8   8   20:08:00:05:33:a5:cf:20    ISL_F48_01_20_F48_02_20_16G
   9   9   20:09:00:05:33:a5:cf:20    ISL_F48_01_20_F48_02_20_8G
  10  10   20:0a:00:05:33:a5:cf:20    MarleneHBA2p0_F
  11  11   20:0b:00:05:33:a5:cf:20    MariaF_default
  12  12   20:0c:00:05:33:a5:cf:20    -
  13  13   20:0d:00:05:33:a5:cf:20    ISL_F48_1_20_F48svc_20_16G
  14  14   20:0e:00:05:33:a5:cf:20    -
  15  15   20:0f:00:05:33:a5:cf:20    -
  16  16   20:10:00:05:33:a5:cf:20    ISL_PFE_F48_01_20_R06low
  17  17   20:11:00:05:33:a5:cf:20    ISL_F48_1_20_F64_lo_20_1
  18  18   20:12:00:05:33:a5:cf:20    ISL_F48_1_20_F64_lo_20_2
  19  19   20:13:00:05:33:a5:cf:20    B-server_default
  20  20   20:14:00:05:33:a5:cf:20    -
  21  21   20:15:00:05:33:a5:cf:20    -
  22  22   20:16:00:05:33:a5:cf:20    ISL_PFE_1_20_DCX_20_1
  23  23   20:17:00:05:33:a5:cf:20    ISL_PFE_1_20_DCX_20_2
```

Another advantage of the use of consistent manual or automatic port naming is to create MAPS Port group based on port name.

The following command shows an example:

```
logicalgroup --create group_name -type type -feature feature_type -pattern pattern
```

For *feature_type*, port names or WWNs can be used, but not both. Quotation marks around the pattern value are required. If “!” is specified in the pattern, it must be within single quotation marks (`'!'`). You can specify only one feature as part of a group definition.

Example 10 shows creating a group that is named `GroupWith_ISL_F48`.

Example 10 Creating the group

```
switch:admin> logicalgroup --create GroupWith_ISL_F48 -type port -feature portname -pattern
"ISL_F48*"

PFE_F48_01def:FID128:admin> logicalgroup --show
-------------------------------------------------------------------------------------------------
Group Name                     |Predefined |Type             |Member Count |Members
-------------------------------------------------------------------------------------------------
GroupWith_ISL_F48              |No         |Port             |0            |
```

The following operators are available:

*****	Match any set of characters in the position that is indicated by the asterisk.
?	Match any single character in the position that is indicated by the question mark.
[expression]	Match any character that is defined by the expression inside the square brackets.
!	Match the string following and exclude any ports that match.

Best practice: Enable dynamic port naming.

Enabling the dynamic port name

On switches running FOS V7.4 and higher, enable the dynamic port name.

Note: To enable the dynamic port name on switches with virtual fabrics, run the **`configure`** command from all logical switches.

Example 11 shows the **`configure`** commands that are used to enable the dynamic port name.

Example 11 The configure command to enable the dynamic port name

```
F48a_Default:FID128:dlutz> configure

Not all options will be available on an enabled switch.
To disable the switch, use the "switchDisable" command.

Configure...
Fabric parameters (yes, y, no, n): [no] y

    WWN Based persistent PID (yes, y, no, n): [no]
    Allow XISL Use (yes, y, no, n): [yes]
    Dynamic Portname (on, off): [on] on
    Edge Hold Time(Low(80ms), Medium(220ms), High(500ms), UserDefined(80-500ms):
(80..500) [220]
```

Example 12 shows the **`portname -d`** command to configure the parameters to be displayed, where the following parameters are available: **`S- Switch Name`**, **`T- Port Type`**, **`I - Port Index`**, **`A - Alias name`**, **`F - FDMI Host name`**, **`R- Remote Switch Name`**, **`C- Slot / Port Number`** (only for SAN Director).

Example 12 The portname -d command to set the parameters and display the setting

```
PFE_F48_01_20:FID20:admin> portname -d "S.T.I.F.A.R"
PFE_F48_01_20:FID20:admin> portname -d
S.T.A.F.R
```

Maintaining an optimal FC SAN environment

In each subsequent release of FOS, Broadcom added and enhanced features to assist with monitoring, protecting, and troubleshooting fabrics. Most of the features were available since FOS V7.2.

Starting in FOS V7.2, this set of features is referred to as *Fabric Vision*. When implemented and administered correctly, these features can dramatically improve the reliability and resiliency of the fabric.

This section focuses on the Fabric Vision features that are available in FOS versions 7.2 - 7.4, and specifically on the following subset of features that apply to monitoring and alerting:

- Fabric Performance Impact
- Fabric Performance Impact Notifications method
- Port Toggling and Slow-Drain Device Quarantine
- Port Fencing

Fabric Performance Impact

FPI monitors congestion-related issues on all physical `E_Ports` and `F_Ports` at all times. FPI added automatic mitigation capabilities through Slow Drain Device Quarantine (SDDQ) and Port Toggle actions.

FPI detects different severity levels of latency and reports three latency states:

- The `IO_FRAME_LOSS` state is a severe level of latency. In this state, frame timeouts occurred or are likely to occur. Administrators must take immediate action to prevent application interruption.
- The `IO_PERF_IMPACT` state is a moderate level of latency. In this state, device-based latencies can negatively affect the overall network performance.
- The `IO_LATENCY_CLEAR` alert occurs when the latency conditions clear.

Administrators must act to mitigate the effect of latency devices. The separate states enable administrators to apply different MAPS actions for different severity levels.

On switches with FOS V8.1 and higher, set up MAPS to quarantine the port for `IO_FRAME_LOSS` events by using the SDDQ option (see “Enabling Monitoring Alerting Policy Suite” on page 50).

> **Note:** To use SDDQ, quality of service (QoS) must be enabled on all switches, which is the factory default.

Fabric Performance Impact Notifications method

FPIN, which is not to be confused with the FPI facility, is a method the fabric switches now use to notify host systems that one or more links in a path they are using is sub-optimal.

This method is useful when links are marginal and cause frame drops or slow transfers but not bad enough for the link to go offline. Historically, servers continue to send I/Os down the marginal path, which resulted in poor performance and more often I/O timeouts that caused recovery issues.

With FPIN, the switch can now advertise to the host that a marginal link exists in the path and for hosts that use FPIN that can suspend I/Os down that path. The result is a much better performance and recovery in these “sick but not dead” situations.

Enabling FPI

Enable FPI on switches that are running FOS V7.3 or V7.4. Switches that are running FOS V8.0 and higher FPI always have enabled by default. Example 13 shows how to enable FPI.

Example 13 Enabling FPI

```
Brocade_def:FID128:admin > mapsconfig -enableFPImon
```

Example 14 shows how to verify whether FPI Monitoring is enabled.

Example 14 Verifying that FPI Monitoring is enabled

```
Brocade_def:FID128:admin > mapsconfig --show
Configured Notifications:       RASLOG,SNMP,EMAIL,FENCE,SW_CRITICAL,SW_MARGINAL
Mail Recipient:                 Not Configured
FPI Monitoring:                 Enabled
Paused members :
===============
PORT :
CIRCUIT :
SFP :
```

Example 15 shows that FPI enabled or disabled is no longer displayed on the **mapsconfig** command because FPI is always enabled in FOS V8.x.

Example 15 FPI no longer displayed in FOS 8.x

```
F48a_Default:FID128:dlutz> mapsconfig --show
Configured Notifications:
RASLOG,SNMP,EMAIL,SW_CRITICAL,SW_MARGINAL,SFP_MARGINAL
Mail Recipient:                 dlutz@ca.ibm.com
Paused members :
===============
PORT :
CIRCUIT :
SFP :
```

Port Toggling and Slow-Drain Device Quarantine

In a fabric, many flows share a link or virtual channel (VC). However, the credits that are used to send traffic or packets across the link are common to all of the flows that use the same link. Therefore, a slow-draining device can slow down the return of credits and have a negative effect on the healthy flows through the link.

To remedy this situation, the following actions were introduced:

- Port Toggling

  Port Toggling takes a port offline for a specified period of time and then brings it back online. The intent is that cycling the port resets the attached device and enables the switch and device to start from a clean initialization point.

- Slow-Drain Device Quarantine (SDDQ)

  SDDQ use the QoS facility. SDDQ enables MAPS to identify a slow-draining device with FPI events and quarantine it by automatically moving all traffic destined to the `F_Port` that is connected to the slow-draining device to a low-priority VC so that the traffic in the original VC does not experience back-pressure.

  When a device is marked as being slow-draining, only the flows that are destined to it are shifted to the low-priority Virtual Circuit (VC). Flows in the reverse direction are *not* affected.

  SDDQ action is blocked when at least one of the following conditions exist:

  – The total number of ports in a zone is greater than 32.
  – Defzone is labeled as `all access`, and no user-defined zoning configuration is used.

FOS V8.1.0 introduces an unquarantine action (`UNQUAR`) for Gen6 switches. The unquarantine action moves a previously quarantined slow drain device out of the quarantine automatically, if the slow drain behavior cleared for a defined timeout period. An unquarantine action can be applied with the `IO_LATENCY_CLEAR` state monitoring. Users can configure an unquarantine timeout value along with the unquarantine action.

FOS V8.1.0 includes the `UNQUAR` action of the default FPI rules in the predefined conservative policy and moderate policy. A use case for the unquarantine action is when a path is needed for different traffic during the day and night, where night traffic must be quarantined and day traffic must use the high priority virtual channel.

The maximum number of devices that can be isolated per unit (chassis or fixed-port switch) is 32 (the default value is 10).

SDDQ and FICON

In most cases, do not enable the SDDQ feature in a FICON environment. The way FICON operates does not benefit from the use of SDDQ.

When SDDQ is enabled, all traffic to a slow-draining device is moved to the lowest-priority virtual circuit. In a single-zone environment, all traffic in the zone is affected.

SDDQ on long-distance links

SDDQ is supported on long-distance links. In long-distance mode, the QoS VC priority is still maintained. However, you must specifically enable QoS again because by default, enabling a long-distance mode disables QoS mode.

After you specifically enable QoS mode again, you must disable and enable the port to bring up the link in the long-distance and QoS modes. The use of MAPS rule enables delaying quarantining the port until after a specified number of rule violations, and is automatically unquarantined one hour after device latency clear.

The port toggling (PT) action and the SDDQ action are mutually exclusive. When the **`mapsconfig`** command is used, you cannot enable the SDDQ and PT actions at the same time.

Enabling SDDQ

On switches with FOS V8.1 and higher, enable SDDQ for the `MAPS IO_FRAME_LOSS` events with the unquarantine option.

Note: To use the SDDQ switch, QoS must be enabled on all switches.

Complete the following steps:

1. To verify whether QoS is active on the ISL switch ports, run the **islshow** command on the switch and look for `QoS` next to each ISL link. You can view the port QoS setting by running the **portcfgshow** command. If QoS is not enabled, run the **portcfgqos** command. Example 16 shows the **islshow** command on an ISL that has QoS active.

   Example 16 The islshow command showing the QoS setting

   ```
   SANA_DCX1:FID16:dlutz> islshow
     1:199-> 28 10:00:00:05:33:99:12:02  20 SANA_DCX2       sp:  4.000G bw:
   4.000G TRUNK QoS
   ```

   Example 17 shows the **portcfgshow** command on ports with the default QoS AutoEnable setting.

   Example 17 The portcfgshow command showing the QoS setting

   ```
   SANA_DCX1:FID16:dlutz> portcfgshow
   Ports of Slot 2        16  17  18  19    20  21  22  23  29  30  31
   ----------------------+---+---+---+---+-----+---+---+---+---+---+---
   Speed                  AN  AN  AN  AN    AN  AN  AN  AN  AN  AN  AN
   Fill Word(On Active)    0   0   0   0     0   0   0   0   0   0   0
   Fill Word(Current)      0   0   0   0     0   0   0   0   0   0   0
   AL_PA Offset 13        ..  ..  ..  ..    ..  ..  ..  ..  ..  ..  ..

   QoS Port               AE  AE  AE  AE    AE  AE  AE  AE  AE  AE  AE
   EX Port                ..  ..  ..  ..    ..  ..  ..  ..  ..  ..  ..
   ```

2. To enable SDDQ, update the suitable `IO_FRAME_LOSS MAPS` rules to use the SDDQ action.
3. Start the MAPS configuration dialog box by clicking **Monitor** → **Fabric Vision** → **MAPS** → **Configure**. In the MAPS Configure dialog box, select the suitable MAPS policy and click **Edit**. Edit the `IO_FRAME_LOSS` rule on the FPI tab and select the **SDDQ** option.

Figure 4 on page 25 shows the FPI rule `IO_FRAME_LOSS` with the SDDQ and unquarantine action enabled.

Add/Edit Rule

Rule Type: Base Rule / Rule On Rule (RoR)

Rule Name: Auto / Custom

FPI_IO_FRAME_LOSS_sddq

Auto-Append (_number) if rule name already exists

Severity: Default

Measure: Device Latency Impact (DEV_LATENCY_IMPACT)

Threshold Value: == IO_FRAME_LOSS

Time Base: None

Actions:
- RAS Log Event
- SNMP Trap
- E-mail
- FMS
- SDDQ
- Un-Quarantine
  - 0 Days
  - 0 Hours 30 Minutes 0 Seconds
- Toggle Duration 2 (2-3600)

Quiet Time
- 0 Days
- 0 Hours 1 Minutes 0 Seconds

(Applicable only for RAS Log, SNMP Trap and E-mail)

Figure 4 Network Advisor update MAPS FPI IO_FRAME_LOSS rule

Note: Run several weeks with this rule enabled but with the SDDQ facility disabled to ensure that the rule works as expected.

4. Enable the SDDQ facility on the MAPS configuration dialog box by selecting the fabric and clicking **Actions**.

   Figure 5 shows the MAPS Policy actions dialog box with the FPI SDDQ action enabled.

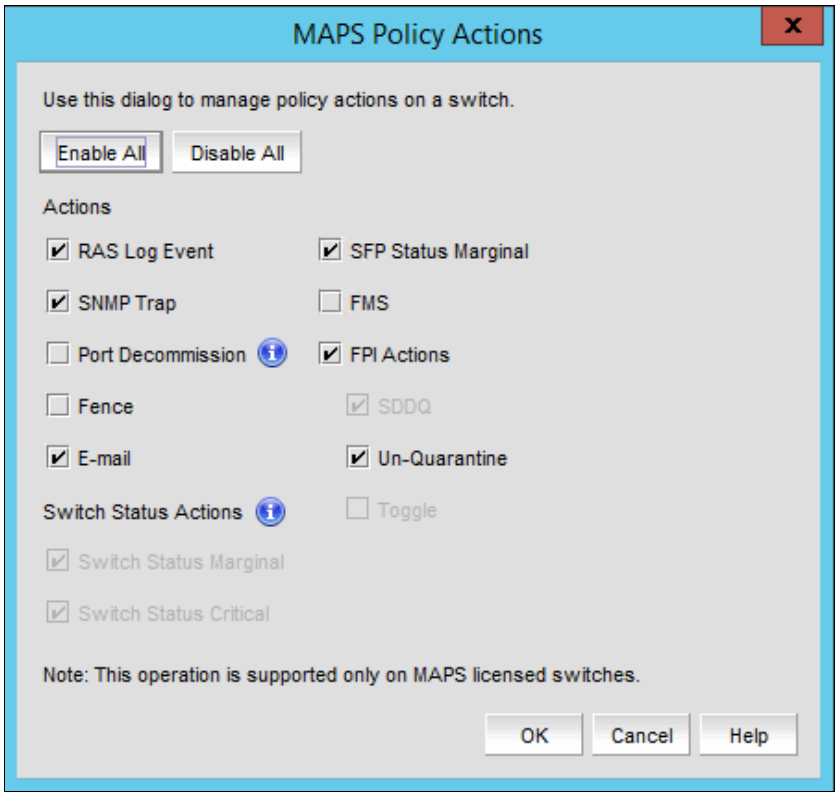


Figure 5 MAPS Policy Actions window with SDDQ enabled

Port Fencing

The Broadcom MAPS Port Fencing action protects against faulty components and conditions that affect links by automatically blocking ports when predefined thresholds are reached.

Enabling MAPS rules with the Port Fencing option must be used with care so that fencing ports occur only on ports that feature severe issues. As a preferred practice, only MAPS rules for host ports for the CRC and Link Reset thresholds are enabled for port fencing.

Before enabling MAPS policy, it is advised to disable fencing for the MAPS policy action. This enablement globally disables fencing, even if a specific rule of a policy is configured with a fencing action. Disabling fencing allows you to monitor the behavior of the policy without any risk to unnecessarily fence a port and create an unexpected impact during normal operations.

When the behavior of the MAPS policy is considered good, the fence MAPS policy action can be enabled.

Configuring Port Fencing

Complete the following steps to configure Port Fencing:

1. From the MAPS configuration, click **Monitor** → **Fabric Vision** → **MAPS** → **Configure**, select the MAPS policy to modify, and then, click **Edit**.
2. Select the rule to modify and click the left arrow, or create a rule and select the **Fence** checkbox to enable the port fencing action.
3. Click the right arrow to transfer the rule to the selected policy.

   Figure 6 shows the editing of a MAPS rule.

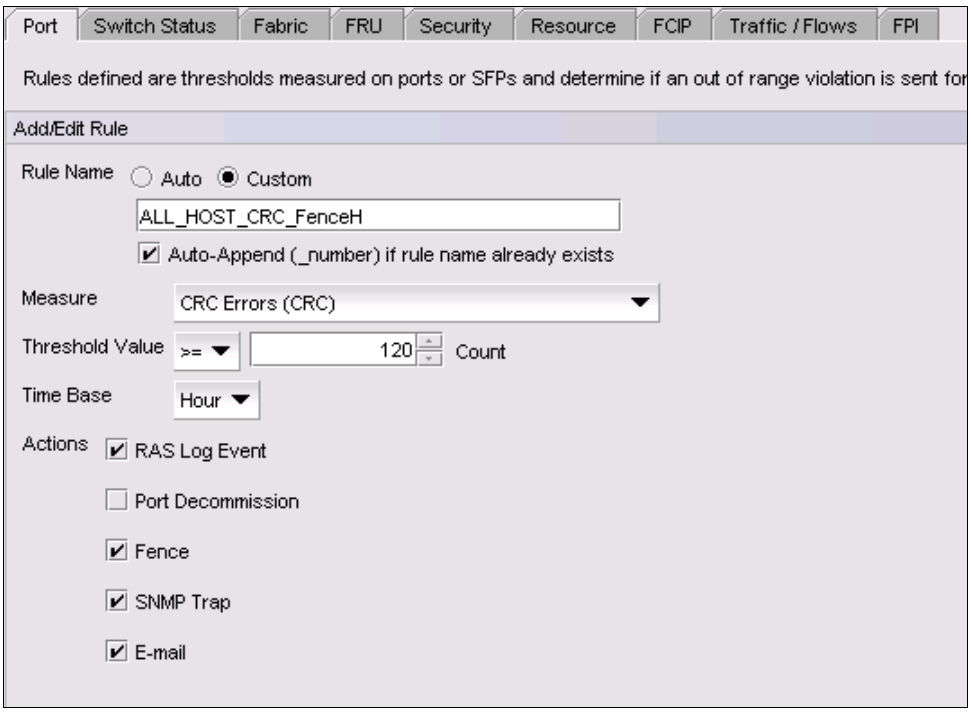


Figure 6 MAPS edit rule

4. After the rules are modified or created, activate the MAPS policy and monitor to ensure that the rules are operating correctly. Then, enable the port fencing facility. From the MAPS configuration, click **Actions** and select the **Fence** checkbox.

Figure 7 shows the activate MAPS policy actions.

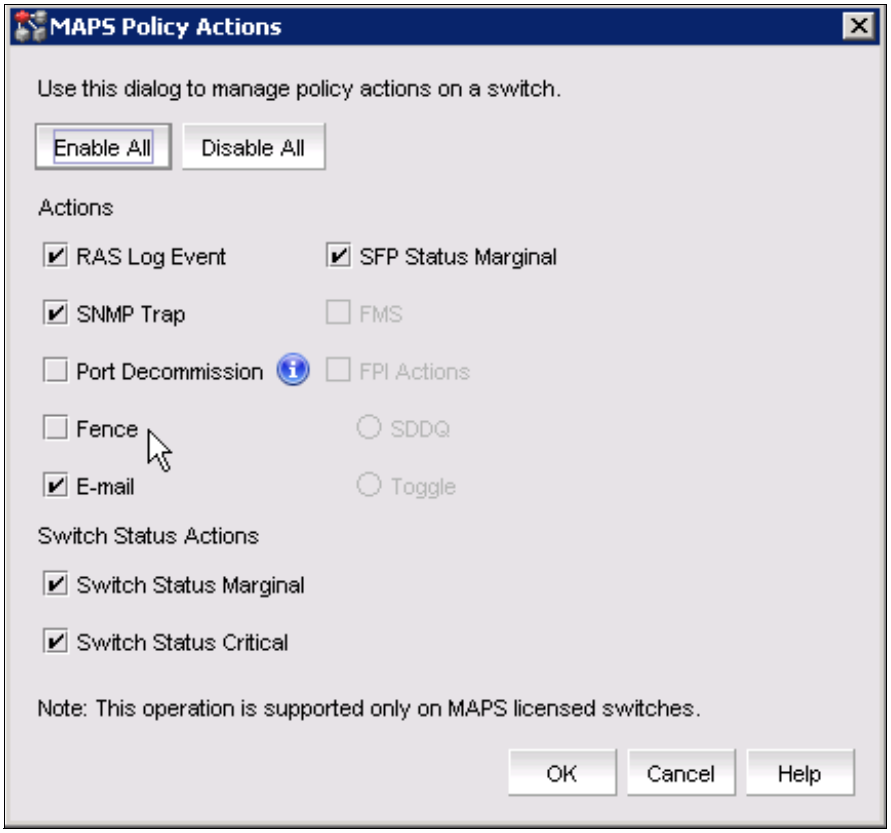


Figure 7 MAPS activate actions

Preferred settings

Enable port fencing on well-managed fabrics with high availability requirements, and then only on host ports for the link reset and CRC metrics.

Maintaining an optimal FCIP SAN environment

Starting with FOS V7.2, Broadcom added features to assist with monitoring, protecting, and troubleshooting FCIP connectivity. These features are part of Fabric Vision MAPS. When implemented and administered with care, these features can positively affect the reliability and resiliency of FCIP solutions and overall SAN fabric operations.

This section focuses on methods, tools, and features that are available in FOS 8.2 and 9.0 and especially as they apply to the following topics:

- Bandwidth Validation
- Using Multiple Circuits
- Effective Monitoring by using Monitoring and Alerting Policy Suite (MAPS)
- Troubleshooting

Ensuring suitable FCIP bandwidth

The task of determining the necessary FCIP bandwidth can be broken down into two general objectives. The first involves defining the bandwidth that is required to meet the needs of the business; the second involves validating that the bandwidth that is delivered matches the projected and expected bandwidth of the solution as implemented.

This publication is not intended to be a broad architectural review, but rather is more specific to the type of questions that must be considered in addition to tests and methodologies that can be performed to arrive at the answers to the objective of maintaining an optimal fabric environment. Many implementation and design considerations for an FC storage area network exist. For more information, see the Broadcom white paper *SAN Design and Best Practices*.

The tools discussion in this section focuses on what is available on Broadcom extension switches and director blades.

Determining bandwidth needs

Several basic questions must be answered when determining how much bandwidth is necessary for the needs of a business. First, the amount of data that is to be transferred or replicated must be calculated. Next, the amount of time for how quickly this data must be transferred from the local to remote site needs to be determined. When these values are known, the simple math of dividing the total data amount by the time determines a starting point for the minimum amount of bandwidth necessary.

To determine the current amount of data to be transferred or replicated, the following questions must be considered:

- How much data will need to be initially transferred for synchronized copies at the local and remote data sites?
- How frequently is the initial data being changed?
- What is the profile of the change rate over a period?
- What is the maximum amount of the data change rate?

The starting point for data sizing is the first question. By using a data replication solution as an example, the initial amount of data to be synchronized to a remote data site must be determined. This answer requires determining which local data sets, volumes, and consistency groups must be replicated to the remote data site, and then, calculating the total size of all of the data to be transferred.

Tip: IBM Storage Insights or IBM Spectrum Control can be used to monitor performance and provide data to determine bandwidth requirements.

If multiple storage systems are replicating data between the local and remote data sites, all of the individual replication streams must be known to obtain a reliable answer for the amount of data to be initially transferred. Therefore, the characteristics of data transfer and replication applications sharing a specific FCIP tunnel must be understood.

Because data sets, volumes, and consistency groups are rarely consistent in size, most storage systems apply a fairness algorithm so that each item to be transferred has equal portions of the bandwidth. With a mixture of large and small volumes and consistency groups, the balanced transfer rates result in smaller volumes or consistency groups being synchronized before the larger ones.

After a smaller volume or consistency group is synchronized, most systems begin to transfer data that was changed in the source volume or consistency group while the larger volumes and consistency groups are still being synchronized. This sizing scope is determined with questions about change frequency, workload profile, maximum change, and change rate over a period.

When the amount of initial data to be transferred or replicated is understood, the next step involves determining the rate of change of the data to be mirrored. This value is not a percentage of how much of each volume or consistency group is changing; rather, it is the amount of data that changes in terms of size, such as bytes. Depending on the type of data and applications that use the data, the change rate might be consistent over time, or it might vary greatly over a specific period.

The best answer for the change rate of each data set, volume, or consistency group is the maximum change rate for a set period, such as 24 hours. The change rates for the various data units to be transferred or replicated are determined based on a common period.

The combination of the initial amount of data to be synchronized with the amount of change data is the total scope of the data to be transferred or replicated. The next step is to determine what the business needs or requires in terms of how quickly the data can be initially transferred and then, what the recovery point objective (RPO) must be.

When the time factor is known, the bandwidth that is needed for the FCIP tunnel is a math exercise to determine the bandwidth value in bits per second. This bandwidth setting is what is needed for the current amount of data to be replicated.

At this point, one other factor must be considered: growth over time. Most businesses experience growth of their replication needs over time. The exercises and calculations were for the current needs, and no consideration was made for data growth.

As a business grows and expands, the amount of data to be synchronized and the change rate likely increases over time. Therefore, meeting the bandwidth needs for the moment is likely to be insufficient for the future operations of the replication solution.

The current bandwidth requirement must be adjusted for future needs based on trend metrics. If a business is experiencing data growth of approximately 25% per year, the replication needs in a year likely are to experience similar growth. Although no rigid rules exist for “future proofing” bandwidth needs, this point must be considered and evaluated and result in an adjustment to the current bandwidth needs to account for future data growth.

Broadcom extension switches and blades feature an optional compression feature that can be considered when determining the bandwidth needs for mirror and replication solutions. Broadcom hardware and the Fabric Operating System (FOS) provide several compression modes. However, the compression modes depend on a number of factors, such as the hardware platform, the protocol (FCIP or IP extension) and the available tunnel bandwidth.

For more information, see *Brocade Fabric OS Extension Configuration Guide*.

Note: Throughput for any compression mode depends on the compressibility of the data to be replicated. Do not expect compression to provide any reduction factor of bandwidth needs.

Actual versus allocated bandwidth

Continuing with the replication solution scenario, in this section, we discuss a simple example of bandwidth needs.

A company has 17 TB of data that must be replicated, and the maximum change rate is 3 TB per day for a total of 20 TB. The business needs of the company state that initial synchronization and starting with the change, data during the synchronization period is to be completed within 24 hours.

Trending data shows that the growth rate of the data is under 10% per year. Therefore, the following calculations are used:

```
((17 TB + 3 TB) * 8 bits/ B) / (24 hours * 3600 second/hour) = 1.852 Gbps
```

To account for the growth over one year, the bandwidth that is needed is adjusted:

```
1.852 Gbps * 1.10 = 2.037 Gbps
```

By rounding down, the company must plan for approximately 2 Gbps total bandwidth between the local and remote data sites to meet the current replication requirements and remain viable for almost a year into the future. With redundant fabrics, one implementation design for this replication solution example can be composed of single 1 Gbps links per fabric across their two redundant fabrics.

Therefore, the bandwidth needs for replication were determined and implemented as 1 Gbps FCIP tunnel per fabric between the local and remote data sites. The next suggested step is to verify that the bandwidth meets the design target before the FCIP tunnels are put into production.

Several WAN analysis tools are designed for testing connections, tracing routes, and estimating the end-to-end IP path performance characteristics between the local and remote data sites. These tools are available by using the **portCmd** command with the following options:

portCmd --tperf This option generates traffic over a circuit to test the network link for issues, such as maximum throughput, congestion, loss percentage, out of order delivery, and other network metrics.

portCmd --ping This option tests the connection between a local Ethernet port and a destination IP address. If testing a VLAN connection, a VLAN tag table entry must be manually added on the local and remote sides of the route.

portCmd --traceroute This option traces routes from the local Ethernet port to a destination IP address. When tracing a route across a VLAN, a VLAN tag table entry must be manually added on the local and remote sides of the route.

portShow fciptunnel This command can be used to view the configuration, status, operational and performance statistics for a designated FCIP Tunnel.

The reason for the use of some of these tools before the FCIP tunnel going into production is that the tunnel cannot be passing any other traffic while the WAN Tool option is running. The WAN tool can be run on multiple circuits consecutively, but only one wtool session per circuit is allowed at a time.

One notable feature of the WAN tool is the ability to test the end-to-end connectivity with jumbo frames to verify that the end-to-end pathway fully supports jumbo frames.

Note: For more information about setting up dashboards and configuring the dashboard widgets, see *Brocade Fabric OS Extension Configuration Guide* for your specific release.

After an FCIP tunnel and its associated circuits are in production mode, the ability to use the WAN analysis tools is disruptive and therefore not likely to be used. A method is available to check the throughput by using TCP/IP metrics that are displayed from the extension switch data.

TCP/IP data can be used to determine what a circuit can deliver. A good validation is if the TCP/IP viewpoint shows the potential for throughput is equal to or greater than the configured bandwidth.

For example, consider an FCIP tunnel that has two circuits on two different GE ports, and each circuit is configured for a maximum bandwidth of 100 Mbps. By using the TCP/IP metric data from the TCP session data of the extension switch, the TCP/IP point shows a potential bandwidth of more than 400 Mbps for each circuit.

In this example, the potential is greater than the configured solution, which is a good check. Although TCP/IP shows that the circuit is capable of much greater throughput, the 100 Mbps threshold limits the traffic level that the extension switch allows.

Example 18 shows two methods to check how much bandwidth was configured for circuits by using the **portshow fciptunnel** command. The first version that uses the command shows the basic configuration of the FCIP tunnel and its associated circuits; the second version of the command shows the breakout bandwidth allocations that are based on priority.

Example 18 The portshow fciptunnel command options

```
portshow fciptunnel -c

 Tunnel Circuit  OpStatus  Flags    Uptime  TxMBps  RxMBps ConnCnt CommRt Met/G
-------------------------------------------------------------------------------
 24    -         Up      -------d-    24d5h   23.21    0.11   6      -       -
 24    0 ge1     Up      ----a---4    24d5h   11.54    0.05   6    20/100   0/-
 24    1 ge3     Up      ----a---4    24d5h   11.67    0.05   6    20/100   0/-
-------------------------------------------------------------------------------

portshow fciptunnel -c --ha --qos

 Tunnel Circuit  OpStatus  Flags    Uptime  TxMBps  RxMBps ConnCnt CommRt Met/G
-------------------------------------------------------------------------------
 24    -         Up      cM-----d-    24d5h    0.00    0.00   6      -       -
 24    0 ge1     Up      ----a---4    24d5h    0.00    0.00   6     0/100   0/-
 24    1 ge3     Up      ----a---4    24d5h    0.00    0.00   6     0/100   0/-
 24    -         Up      hM-----d-    24d5h    0.00    0.00   6      -       -
 24    0 ge1     Up      ----a---4    24d5h    0.00    0.00   6    10/100   0/-
 24    1 ge3     Up      ----a---4    24d5h    0.00    0.00   5    10/100   0/-
 24    -         Up      mM-----d-    24d5h   23.21    0.11   6      -       -
 24    0 ge1     Up      ----a---4    24d5h   11.54    0.05   6     6/100   0/-
 24    1 ge3     Up      ----a---4    24d5h   11.67    0.05   5     6/100   0/-
 24    -         Up      lM-----d-    24d5h    0.00    0.00   6      -       -
 24    0 ge1     Up      ----a---4    24d5h    0.00    0.00   6     4/100   0/-
 24    1 ge3     Up      ----a---4    24d5h    0.00    0.00   6     4/100   0/-
-------------------------------------------------------------------------------
```

Each circuit is configured for a maximum bandwidth of 100 Mbps. Although each circuit is connected to a GE Ethernet port with a capacity of 1000 Mbps, the extension switch limits the maximum throughput to only 100 Mbps. The method for manually checking the throughput is a similar type of situation. If the TCP/IP driver shows more bandwidth than configured, the circuit is in a good condition.

Manually checking what the TCP/IP driver considers the available bandwidth to be requires performing some calculations by using the size of the receive window divided by the Round-Trip Time (RTT), as shown in the following simple version of the equation:

<size of receive window in bits> / <RTT in seconds> = <detected bandwidth in Mbps>

By using the **portshow fciptunnel** command and specific options, the necessary information to perform the calculation is displayed. Most FCIP tunnel implementations do not use QoS zoning, which results in the replication traffic being passed with medium QoS.

Therefore, only the medium TCP flows must be measured, as shown in Example 19.

Example 19 The portshow fciptunnel command

```
portshow fciptunnel -cd --ha --qos --tcp

  TCP Connection 24.0 HA-Type:Main Pri:Medium Conn:0x02f77549
  ===================================================
   Local / Remote Port      : 3226 / 49671
   Duration                 : 24d5h
   MSS                      : 1460 bytes
   ARL Min / Cur / Max      : 3000 / 33032 / 50000
   ARL Reset Algo           : StepDown
   Send Window
    Size / Scale            : 1240064 / 9
    Slow Start Threshhold   : 16777216
    Congestion Window       : 16854320
    Pkts InFlight           : 0
   Recv Window
    Size / Scale            : 1249792 (Max:1249792) / 9
   SendQ Nxt / Min / Max    : 0xe12631b8 / 0xe122e5a8 / 0xe12631b8
   RecvQ Nxt / Min / Max    : 0x8e4328c7 / 0x8e4328c7 / 0x8e563ac7
   RecvQ Pkts               : 745107039
   Sender Stats
    Sent Bytes / Pkts       : 1372336980073 / 1126716089
    Unacked Data            : 216080
    Retransmits Slow / Fast : 342 / 96497 (High:0)
    SlowStart               : 0
   Reciever Stats
    Recv Bytes / Pkts       : 28859842579 / 578869489
    Out-of-Order            : 0 (High:45)
    Duplicate ACKs          : 368298
   RTT / Variance (High)    : 46 ms (72 ms) / 0 ms (29 ms)
```

The receive window size in the example output is in bytes, not bits; therefore, it must be converted by multiplying this value by 8 to get bits. The Round-trip Time value is shown in milliseconds, which must be converted to seconds by dividing the RTT value by 1000. In the our example, the equation becomes:

```
(1249792 bytes * 8 bits/byte) / (46 ms / 1000 msps) = 217.355 Mbps
```

The TCP/IP driver that is used in the Broadcom extension switch and blade is programmed to create enough TCP sessions to drive the circuit to its maximum with a measure of failover capability. In Example 20, two TCP sessions are identified by the unique TCP connection identifier. In the example, the second TCP session includes similar values for the receive window size and Round-trip Time.

Example 20 TCP sessions

```
TCP Connection 24.1 HA-Type:Main Pri:Medium Conn:0x02f7755d
  ===================================================
   Local / Remote Port      : 3225 / 55818
   Duration                 : 24d5h
   MSS                      : 1460 bytes
   ARL Min / Cur / Max      : 3000 / 43000 / 50000
```

```
ARL Reset Algo           : StepDown
Send Window
 Size / Scale            : 1240064 / 9
 Slow Start Threshhold   : 16777216
 Congestion Window       : 16907848
 Pkts InFlight           : 0
Recv Window
 Size / Scale            : 1249792 (Max:1249792) / 9
SendQ Nxt / Min / Max    : 0x5b5158b5 / 0x5b4d22d1 / 0x5b5158b5
RecvQ Nxt / Min / Max    : 0x8e2f1c40 / 0x8e2f1c40 / 0x8e422e40
RecvQ Pkts               : 1933390132
Sender Stats
 Sent Bytes / Pkts       : 1374055066258 / 1127990519
 Unacked Data            : 275940
 Retransmits Slow / Fast : 347 / 81260 (High:0)
 SlowStart               : 0
Reciever Stats
 Recv Bytes / Pkts       : 28858431967 / 581754054
 Out-of-Order            : 0 (High:69)
 Duplicate ACKs          : 361130
RTT / Variance (High)    : 46 ms (71 ms) / 0 ms (28 ms)
```

The two TCP sessions are capable of driving traffic at the rate of 434.71 Mbps, which is greater than the configured 100 Mbps for the circuit. Therefore, the configured bandwidth can be fully used. This verification was performed while the FCIP tunnel and its circuits were in production.

Link quality

The characteristics of the FC and IP protocols are not an easy match. FC is based on lossless connections between device ports with a high emphasis on in-order delivery of frames, but IP is based on the assumption that some degree of packet loss occurs.

The FC-SCSI protocol is sensitive to response times and higher response times (measured as the round-trip times) result in lower throughput. FC-SCSI also is sensitive to fluctuating latencies, and tends to have issues where consistent latency does not exist. As a result, the FCIP links have a lower tolerance to out of order, slow start, and retransmits than typical IP links.

Jitter

Jitter is the variation in the delay of received packets. At the sending side, packets are sent in a continuous stream with the packets spaced evenly apart. Because of network congestion, incorrect queuing, or configuration errors, this steady stream can become bursty, or the delay between each packet can vary instead of remaining constant.

Jitter is measured as a percentage of variance of how much the round trip (RTT) changes based on the average time. If the RTT average time is 50 ms with Jitter of 5%, the RTT times range 47.5 ms - 52.5 ms. Although some degree of jitter is expected and allowed by IP protocols, extreme jitter results in degraded performance.

Best Practice: Jitter should not vary by more than 10 to 15%.

Retransmits

Packet loss is more common in IP networks than frame discards in FC networks. The IP protocol accommodates this situation by having the receiving end request a retransmission of one or more packets whenever a packet loss event is detected. The need to retransmit packets increases the latency to the end FC device ports. If the latency is sufficiently large, SCSI timeouts or replication suspensions might occur because of increased latency.

The acceptable retransmit levels in an FCIP network typically is much lower where retransmit levels should be under 0.05% and preferably 0.01% or less.

Retransmits are reported by extension switches as packet loss (pktloss) and can be seen by using the CLI command, as shown in the following example:

```
portshow fciptunnel -cd --ha --qos --tcp command
```

Best Practice: Retransmits should by 0.05% or less and 0.01% is even better.

Out-of-order packets

When an IP-FCIP packet is received and is more than three positions in the receiving packet flow, it is discarded and a retransmit situation results. If a packet is received within three or less positions when compared to how it was sent, it is a an out-of-order situation. For example, IP-FCIP packets are sent in the order of 1 - 2 - 3 - 4 - 5. A retransmit situation occurs if the receiving order is 2 - 3 - 4 - 5 - 1; an out-of-order situation example is 2 - 3 - 1 - 4 - 5.

The receiving extension switch handles out-of-order packets by moving each out of position packet into its correct position in the receive stream of packets before the extraction process of the encapsulated FC frames begins. Therefore, out-of-order packets introduce some degree of latency in the receiving extension switch, and the degree of latency is not a severe as a retransmit situation. However, if the number of out-of-order packets is significant over time, the effect of the resulting latency can cause time-out situations or other errors.

Best Practice: Out of order should be 0.05% or less.

Using multiple circuits

An FCIP tunnel is a single Inter-Switch Link (ISL) that contains at least one, or more, circuits. When multiple circuits are used to create the FCIP tunnel, it is also known as an *extension trunk*. A circuit is a physical connection between a pair of IP addresses that are associated with the local and remote endpoints of an FCIP tunnel. Circuits provide the pathways for traffic flows between the local and remote interfaces at each end of the tunnel.

Multiple circuits can be configured per Ethernet port by assigning them unique IP interfaces. When configuring a circuit, the IP addresses for the local and remote interfaces are provided, and each circuit must be composed of unique pairs of local and remote IP interfaces, as shown in Figure 8.

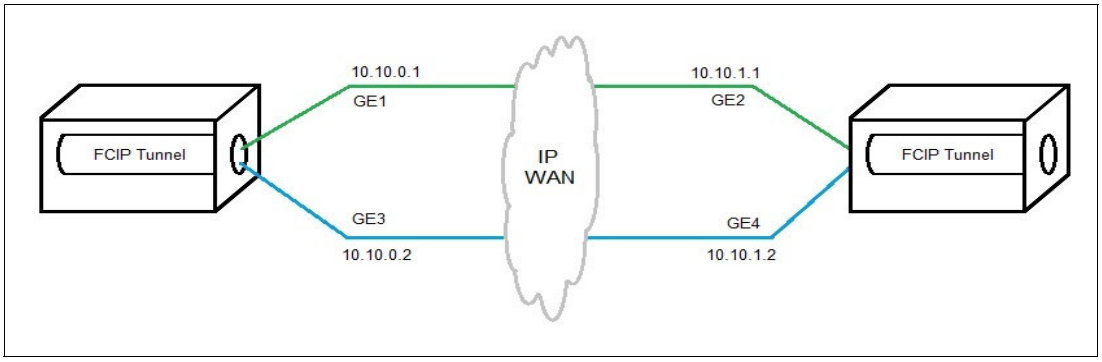


Figure 8 FCIP Tunnel with multiple circuits

Extension Trunking

Extension Trunking is a method for the aggregation and management of the use of IP WAN bandwidth. Extension Trunking provides a method of using redundant paths over the WAN for lossless failover and increased resiliency because of WAN failure. Extension Trunking also improves load balancing on a weighted round-robin basis. Trunking is enabled by creating multiple circuits within a FCIP tunnel so that the tunnel can use the circuits that are passing traffic between multiple local and remote addresses.

Extension Trunking provides lossless link loss (LLL), which ensures that all data lost in-flight is retransmitted and reassembled back in order before being delivered to upper layer protocols. This feature is essential for FICON environments to prevent interface control checks (IFCCs), and open systems replication environments to prevent SCSI time-outs.

While multiple FCIP tunnels can be defined between pairs of extension switches or extension blades, this design defeats the benefits of a multiple-circuit FCIP tunnel. Defining two tunnels between a pair of extension switches or blades is not as redundant or fault-tolerant as having multiple circuits in one FCIP tunnel. The advantage of multiple circuits is where some circuits within an Extension Trunk can be configured as failover circuits or as spillover circuits.

The *failover circuit* is essentially a standby circuit to be used when an active circuit fails. A *spillover circuit* is a secondary circuit that is used only during periods of high traffic usage. When configuring an FCIP tunnel with multiple circuits, failover circuits and spillover circuits cannot be used at the same time.

For solutions that have high bandwidth requirements between the local and remote data sites, multiple FCIP tunnels might be necessary. In this situation, each FCIP tunnel is composed of multiple circuits. When multiple parallel FCIP tunnels are created between the local and remote data sites, lossless dynamic load sharing (DLS) must be enabled. This function is enabled to correctly handle routing updates that occur when FCIP tunnels come up or go down. Each routing update can cause dropped or unroutable frames if the destination is by way of a peer tunnel.

Adaptive Rate Limiting

Adaptive Rate Limiting (ARL) is a licensed feature and can be implemented on individual circuits. ARL enables the extension switch to change the circuit's throughput rate by working with the IP network. ARL uses information from the TCP sessions to dynamically adjust the throughput rate for the circuit.

ARL is implemented by configuring minimum and maximum bandwidth rates for a circuit. The minimum configured bandwidth is maintained, and the maximum configured rate is not exceeded.

When replication traffic starts passing over a circuit, ARL starts at the minimum rate and attempts to increase the data throughput until it reaches the maximum configured rate (or TCP information indicates that no more bandwidth is available). If the throughput data rate is not at the maximum configured rate, ARL routinely checks TCP conditions for more bandwidth.

If the indicators show that more bandwidth is available, ARL continues to increase up to the maximum configured bandwidth rate. If problems are detected on the circuit, ARL reduces the rate.

ARL is recommended if the long distance WAN path is shared with multiple types of mirror and replication traffic such that the bandwidth cannot be guaranteed. The following common scenarios are examples of where ARL can be effective:

- Storage and non-storage traffic are sharing a single WAN link
- More than one extension interface uses a single WAN link that was dedicated to a specific storage system
- Any combination of the first two scenarios

A key design element of a SAN solution is the use of redundant fabrics. This design element also should be applied to WAN connectivity between the local and remote site. If redundant fabrics with separate extension switches use a single WAN link between sites, any interruption of the WAN link is disruptive to the redundant fabrics. The distribution of the connections can be based at the following levels:

- Circuit, where one or more circuits per fabric use separate WAN links
- Fabric, where all circuits for a specific fabric include dedicated WAN links

Preferred settings

Create FCIP tunnels with multiple circuits between a pair of extension switches or blades with ARL configured using multiple WAN links if possible. If multiple FCIP tunnels with protocol optimization features are used, such as FICON Acceleration or Broadcom Open Systems Tape Pipelining (OSTP), each tunnel must be allocated to a logical switch/logical fabric (LS/LF), depending on the specific set of feature configurations that must be used.

Best practice: When multiple circuits are used for a FCIP Tunnel, be sure the RTT of each circuit is within 5% or less of each other.

Ensure effective monitoring

On switches running FOS V8.2 or higher, use the MAPS to monitor the FCIP tunnels and circuits. To ensure that consistent settings and policies are enabled on all switches in the fabric, use Network Advisor or SANnav.

MAPS is a ready-made solution for policy-based threshold monitoring and alerting. By using pre-built, policy-based rules templates, MAPS simplifies threshold configuration, monitoring, and alerting. The following sections focus on the FCIP Health category for monitoring.

Fabrics with long distance connections for replication and data migration that have high availability requirements must have strict monitoring of those connections. As a preferred practice, use a custom set of MAPS rules that combines a number of thresholds from the default aggressive policy, in addition to customized rules for the solution's round-trip time (RTT) and circuit utilization.

Table 1 lists the monitoring items and the MAPS thresholds for the three default policies.

Table 1 Default MAPS FCIP monitoring thresholds

Monitoring statistic	Unit	Default FCIP monitoring threshold		
		Aggressive	Moderate	Conservative
Circuit packet loss percentage (`CIR_PKTLOSS`)	Percentage per minute	0.01	0.05	0.1
Circuit state change (`CIR_STATE`)	Changes per hour	0	3	5
Circuit utilization percentage (`CIR_UTIL`)	Percentage per hour	60	75	90
Circuit jitter (`JITTER`)	Percentage of delay change per minute	5	15	20
Circuit round trip time (`RTT`)	Milliseconds	250	250	250
Tunnel (`STATE_CHG`)	Changes per minute	0	1	3

Typical IP traffic can tolerate higher rates of packet loss and the resulting retransmission of FCIP packets than the FC layer of FCIP can tolerate. Many cases exist in which storage devices experience replication failures because of packet loss on the FCIP tunnel, yet the IP infrastructure did not record or report any issues. For FCIP metrics, we suggest logging any packet loss of 0.01%, and to alert and take action for packet loss of 0.05% or higher.

Preferred FCIP MAPS settings summary

Enable MAPS with a custom policy that is based on the default Aggressive policy on all switches that are running FOS V8.2 or higher. The monitoring options that must be customized are Round-trip Time and Jitter. Round-trip Time varies greatly with every installation and the distance involved; therefore, after you determine what the normal RTT value is, establish a logging threshold when RTT exceeds 10% and an alerting threshold when it exceeds 20%. For Jitter metrics, we suggest logging when Jitter reaches 10% and to alert when Jitter becomes 25% or higher.

Be aware that excessive packet loss, variable RTT times, or inconsistent Jitter are indications of issues within the LAN or WAN pathway between the local and remote sites. Although the extension switch is recording alerts with these parameters, the source of the issue is not with or within the extension switch.

Accelerators

Response times can play a major role in the amount of data that can be sent over an extended link. To help with this segment of a replication solution, accelerators are available where the local FCIP switch or a device in the IP network can send responses on behalf of a remote switch or device. The idea is that an accelerator is the remote switch or device sends a suitable response; however, the “proxy” action of the accelerator allows data flows to begin quicker, which increases throughput.

However, the early responses can create issues when the early “proxy” response differs from the response that is sent by the remote switch or device. For many storage systems, accelerators in the extended link are not supported.

Best practice: Do not use FCIP write acceleration with IBM disk storage devices. The use of FCIP Tape Acceleration for tape devices can be deployed if the path with FCIP Tape Acceleration does not have any disk traffic.

Troubleshooting

Two primary data resources are available to use when troubleshooting issues in the FCIP SAN environment. The error log and MAPS alerts are the key reference points and both data resources must be consulted at each end of a problematic FCIP tunnel. MAPS alerts for elevated pktloss, excessive jitter, or RTT threshold alerts are indicators of issues in the LAN/WAN pathway between a pair of extension switches.

Consult the error log to determine if any reasons for problems are listed. Error log messages state why a circuit or tunnel went down or indicate some internal issue within the extension switch. The messages of interest are network or remote close, keep-alive timeout, and retransmits exceeded. Many of these messages are indicators if the root source of a problem is inside the extension switch, outside the extension switch with the LAN/WAN pathway, or with the remote extension switch.

The information from these two data resources can provide helpful clues towards further investigations and resolution.

Monitoring a FC SAN environment

This section describes how to maintain and optimal FC SAN environment.

Monitoring Alerting Policy Suite

MAPS provides an easy-to-use solution for policy-based threshold monitoring and alerting. When configured correctly, MAPS provides real-time monitoring and logging of many switch metrics, which eliminates the need to clear stats and monitor the stats over time. With MAPS, events are logged that enable you to see exactly when these threshold events occurred. This information then enables you to correlate events that occur during a performance or host impacting event.

MAPS alerting through email or SNMP notifies you of conditions that require attention, which enables marginal links or components to be resolved before they affect operations.

MAPS also can be used to automatically fence, toggle, or quarantine ports, and provide levels of protection for the fabric without operations intervention.

MAPS was introduced in FOS V7.2 and replaces Fabric Watch as the preferred monitoring tool. With FOS V7.4, Fabric Watch is no longer available and MAPS is enabled by default.

FOS V7.4 introduced a basic set of monitoring rules to monitor overall switch status, FRU health, and base FPI, and to monitor events without the need for a Fabric Vision or Fabric Watch License. To use the full suite of MAPS monitoring, reporting, and actions capabilities, a Fabric Vision (or Fabric Watch and Advanced Performance Monitoring) license is required.

FOS V7.4 also added several new monitoring metrics, including several metrics for FCIP circuits and tunnels. Two new actions for FPI events were also added where MAPS can be set up to behave in the following ways:

- Toggle (bounce the port offline then back online) when FPI detects congestion on the port
- Quarantine the flow (place the flow into a lower quality of service) to minimize the effect the congested flow has on the rest of the fabric

FOS V8.0 added FCIP monitors to include monitoring Ethernet ports and tunnel metrics.

FOS V8.1 added Rule On Rule, which allows a rule to monitor how many times another rule was triggered, which enables a threshold to be exceeded a number of times and not generate alerts or actions. However, if the threshold is exceeded several times, it can be alerted or actioned.

FOS V8.1 added the FPI SDDQ unquarantine action that allows quarantined ports to be automatically unquarantined based on a time value or the `IO_Latency_Clear` event.

Configuring MAPS

Enable a default MAPS policy (typically the `dlft_conservative` policy) unless you have a high availability or high performance fabric. If so, see the advice for a custom policy. To enable a MAPS policy with Network Advisor, complete the following steps:

1. Select **Monitor** → **Fabric Vision** → **MAPS** → **Configure**.

   Figure 9 shows the Network Advisor Fabric Vision MAPS Configure menu.

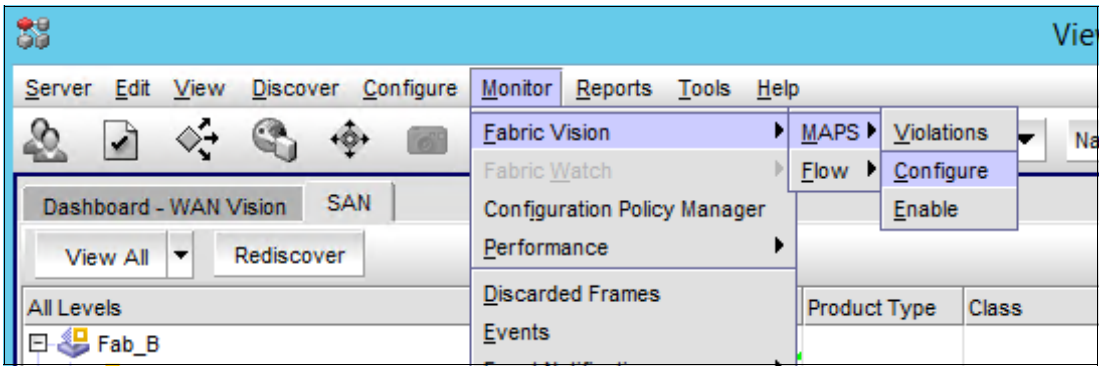


Figure 9 Network Advisor Fabric Vision MAPS Configure menu

2. Set up the suitable MAPS actions and, at a minimum, the RAS Log Event should be selected. Do not select the Fencing action unless a custom policy is being used with suitable port fencing metrics and thresholds.

Figure 10 shows the Network Advisor MAPS Policy Actions dialog.

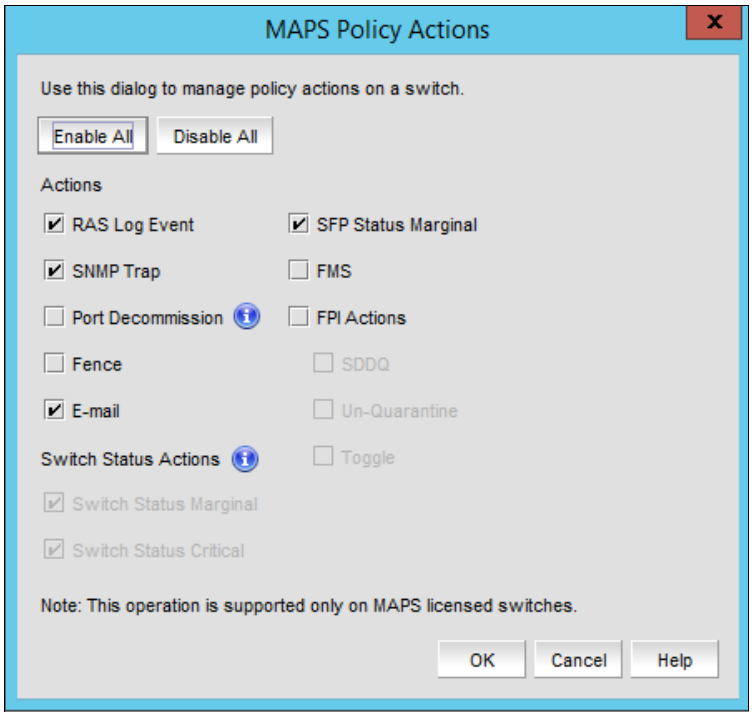


Figure 10 MAPS Policy Actions dialog

Note: IBM suggests enabling MAPS with the `dflt_conservative` policy, with the RAS Log, SNMP, and E-mail actions selected. IBM does not recommend selecting the Fencing action.

Smaller installations that do not have Network Advisor running FOS V7.2 or V7.3 can enable MAPS by running **`mapsconfig --enablemaps`**. In FOS V7.4 and later, MAPS is enabled by default. However, unless you have a license, you can use only the limited base monitoring policy.

The `dflt_conservative` policy can be enabled by using the **`mapspolicy --enable`** command. The MAPS actions can be set up by using the **`mapconfig --actions`** command (see Example 21).

Note: For switches with Virtual Fabrics enabled, all logical switches must have MAPS enabled and configured. The use of the **`FOSEXEC --fid all`** command issues the commands to all logical switches.

Example 21 CLI commands enabling MAPS and MAPS actions

```
SANA_DCX2:FID16:dlutz> fosexec --fid all -cmd "mapspolicy --enable
dflt_conservative_policy"
---------------------------------------------------
"mapspolicy" on FID 128:
---------------------------------------------------
"mapspolicy" on FID 4:
```

```
SANA_DCX2:FID16:dlutz> fosexec --fid all -cmd "mapsconfig --actions
RASLOG,SNMP,EMAIL,SW_CRITICAL,SW_MARGINAL,SFP_MARGINAL"
---------------------------------------------------
"mapsconfig" on FID 128:
---------------------------------------------------
"mapsconfig" on FID 4:
```

Custom policy for high availability fabrics

Fabrics that have high availability and performance requirements need a more stringent approach to monitor and alter attributes than the default policies provide.

As a preferred practice, we suggest the use of a custom set of MAPS rules that combine the thresholds from the moderate policy and aggressive policy into a single custom policy. The custom policy has aggressive rules that log minor threshold violations so that you can detect marginal conditions and act on them in a proactive way.

The custom policy uses moderate rules to alert you through SNMP or email for conditions that need more immediate attention. It also includes a set of safety net rules to use automatic actions, such as for port fencing, or port quarantine for severe issues. The overall strategy is to copy (clone) the moderate default policy because many default rules (such as power supply or SFP thresholds) exist that we want to use without the need to define every rule.

Port metrics

For the port metrics that you want to customize, delete the default policy equivalent rules and replace them with your thresholds, as listed in Table 2. Port thresholds must be customized for the `Non_E_F_PORTS`, `ALL_E_PORTS`, `ALL_OTHER_F_PORTS`, `ALL_HOSTS_PORTS`, and `ALL_TARGET_PORTS` groups.

Table 2 Custom MAPS port metrics threshold

Type	Time	OP	RASLOG	Alert	Fence (host only)
C3TXTO	min	ge	3	20	
CRC	min	ge	10	20	40
CRC	hour	ge			240
ITW	min	ge	20	40	
LF	min	ge	3	5	
LOSS_SIGNAL	min	ge	3		
LOSS_SYNC	min	ge	3		
LR	min	ge	5	10	20
LR	hour	ge			60
PE	min	ge	3	7	
RX/TX/UTIL	hour	ge	75	90	
STATE_CHG	min	ge	5	10	

Note: Thresholds in the Fence column must be created only in the `ALL_HOST_PORTS` group.

Switch fabric metrics

For switch metrics, set the `EPORT_DOWN` and `FAB_SEG` thresholds to greater than or equal to 1 to create alerts for every `E_Port` link that is down, or switch segmentation, by using the thresholds that are listed in Table 3.

Table 3 Custom MAPS switch metrics thresholds

Type	Timebase	OP	RASLOG alert
EPORT_DOWN	min	ge	1
FAB_SEG	min	ge	1
FLOGI	min	ge	5

FPI metrics

For FPI metrics at FOS V7.x and FOS V8.0, MAPS must be set up to log FPI events to the RASLOG. Most fabrics experience congestion because of the shared workloads in the fabric. Because the number of FPI events can be large, alerting on FPI events is not recommended, as shown in Example 22.

Example 22 The mapsrule CLI command to enable logging for FPI events

```
mapsRule --create FPI_IO_PERF_IMPACT_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_PERF_IMPACT -action RASLOG
-policy <policy_name>
mapsRule --create FPI_IO_FRAME_LOSS_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_FRAME_LOSS -action RASLOG
-policy <policy_name>
mapsRule --create FPI_IO_LATENCY_CLEAR_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_LATENCY_CLEAR -action RASLOG
-policy <policy_name>
```

Although SDDQ was introduced in FOS V7.4, MAPS can be configured only to quarantine a port. To remove the port from quarantine requires manual operator intervention, which can easily be missed.

With FOS V8.1 and the **`unquarantine`** action, you can update the `IO_FRAME_LOSS` event to automatically quarantine a port for a specific time interval, at which time it is unquarantined, as shown in Example 23.

Example 23 The mapsrule CLI command to enable sddq

```
mapsRule --create FPI_IO_FRAME_LOSS_sddq -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_FRAME_LOSS -action
RASLOG,SDDQ,UNQUAR -uqrt=30 -uqrt_unit min -policy <policy_name>
```

With FOS V8.2.1, a rule can be set up for FPI `IO_FRAME_LOSS` with an action of alerting so that you can be alerted if ports are being repeatedly put in and out of quarantine state, as shown in Example 24.

Example 24 The mapsRule CLI command to enable altering for repeat sddq events

```
mapsRule --createRoR FPI_IO_FRAME_LOSS_ror -group ALL_PORTS -monitor
FPI_IO_FRAME_LOSS_sddq -op ge -timebase DAY -value 5 -action RASLOG,EMAIL,SNMP
-policy <policy_name>
```

FCIP metrics

Typical IP traffic can tolerate higher rates of packet loss (driving retransmits) than FC over IP (FCIP) can tolerate. Many cases exist in which storage devices experience replication failures because of packet loss on the FCIP tunnel where the IP infrastructure does not record or report any issues. For FCIP metrics, we suggest that you log any packet loss over 5%, and alert and take action for packet loss over 10%, as listed in Table 4.

Table 4 Custom MAPS FCIP metrics thresholds

Type	Time	Op	RASLOG	Alert
PKTLOSS	min	ge	0.05	0.1
JITTER	min	ge	10	20
RTT	none	ge	see note	see note

Note: Round Trip Time varies greatly with every installation. Therefore, after you establish what the normal RTT value is, establish a logging threshold when RTT exceeds 10%, and an alerting threshold when it exceeds 25%.

Preferred settings

Enable MAPS with the default conservative policy on all switches running FOS V7.2 or higher. For fabrics that have high availability requirements, create a custom policy to provide more monitoring and alerting for marginal issues.

Monitoring by using Network Advisor

Network Advisor 12 introduced dashboards and were enhanced in subsequent releases. Dashboards are a visual way to view key fabric metrics to help quickly identify issues.

The dashboard displays different widgets that contain switch and port status, port thresholds, performance monitors. and other items. Network Advisor comes with some standard dashboards, such as Product Status and Traffic and SAN Port Health, and you can create more custom dashboards.

A dashboard provides a high-level overview of the network and the current states of managed devices. You can easily check the status of the devices in the network. The dashboard also provides several features to help you quickly access reports, device configurations, and system event logs.

The custom Fabric Health dashboard that is shown in Figure 11 has several widgets defined that can quickly show the current state and health of the fabric. As shown at the top of the figure, the Scope field defines which switches and what time frame is used to populate the widgets.

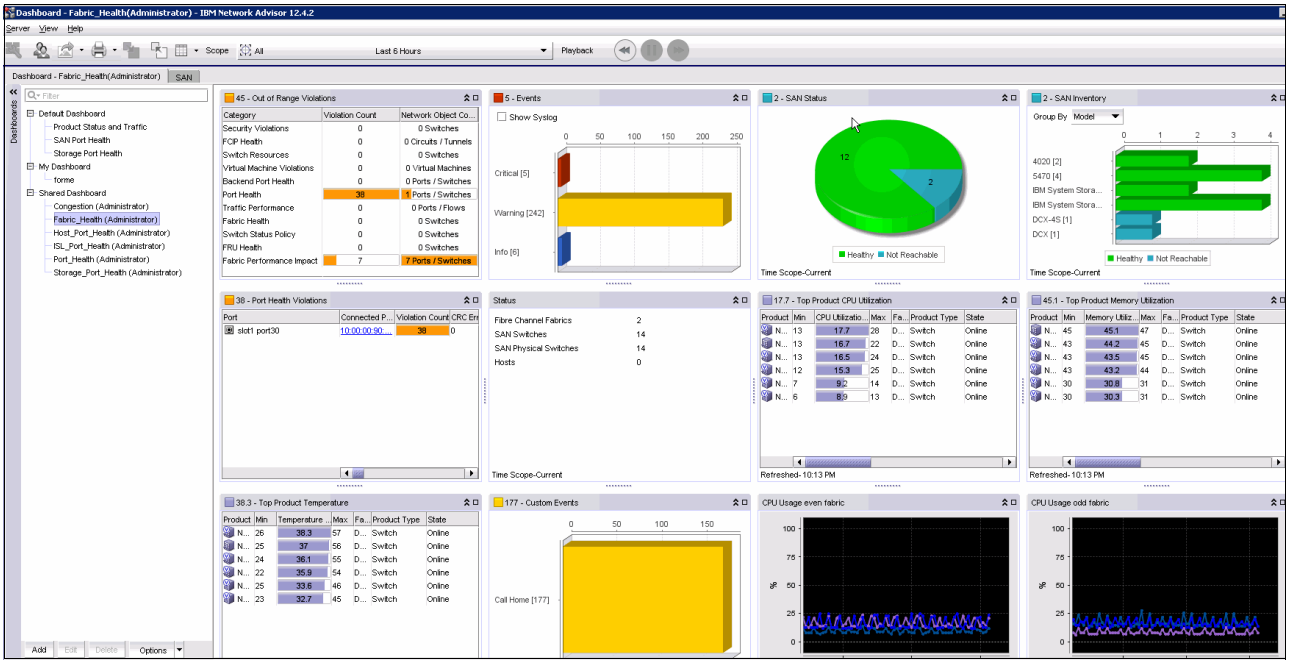


Figure 11 Network Advisor Fabric Health dashboard

Widgets, such as the Out of Range widget that shows the number of ports that had violations for each category for the selected time range, enable by double-clicking the category. Dialog boxes open in which you can drill down to the specific details for the violations. Similarly, you can use the Events widgets to click the event severity to display the individual event messages.

Note: For more information about setting up dashboards and configuring the dashboard widgets, see the *Brocade Network Advisor SAN User Manual* for your release by searching in the Brocade Document Library website.

One of the more powerful features of the dashboards is the ability to select the time frame or which fabric is used to populate the widgets. You can set the time frame for the last 24 hours to see what issues occurred in the past day to monitor for marginal issues that might be occurring, or dial down the scope to 30 minutes to focus on the current metrics for real time problem investigation.

Figure 12 shows the dashboard time and fabric scope selection window.

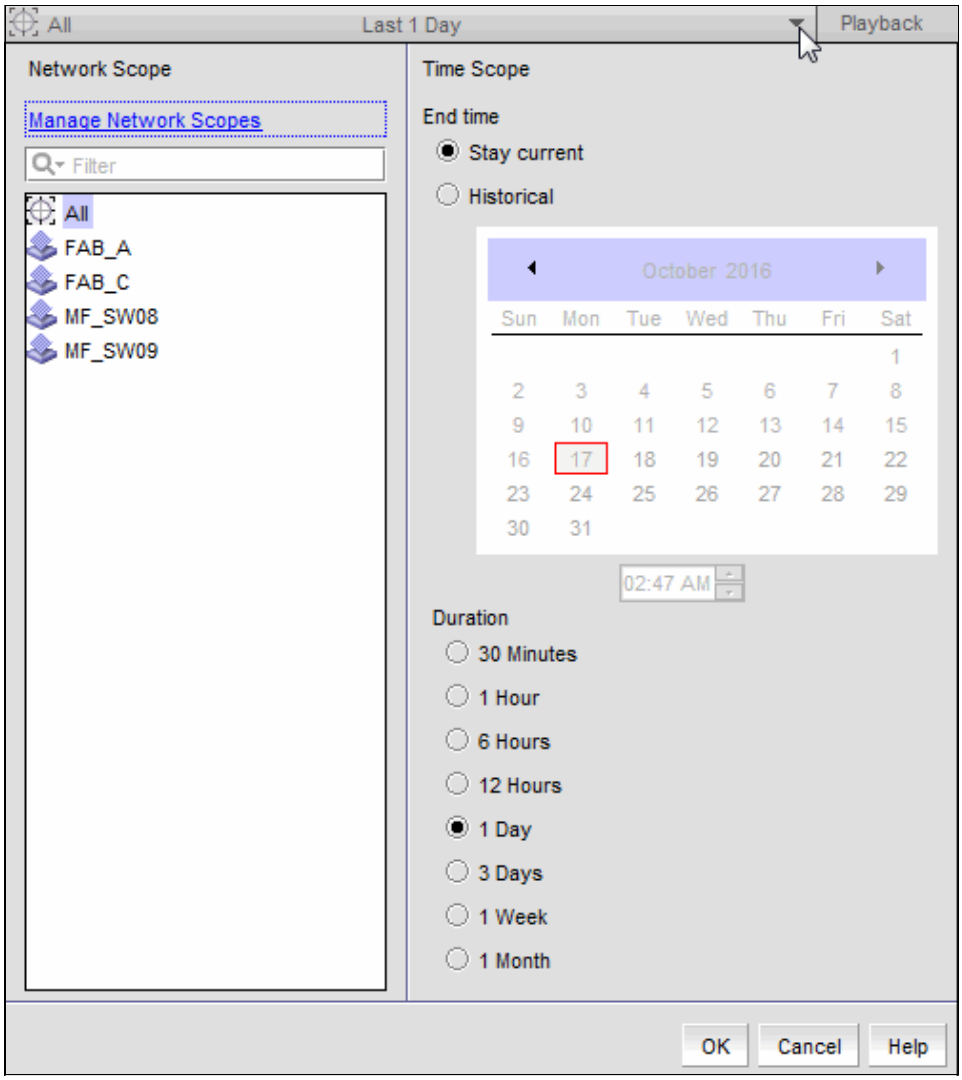


Figure 12 Network Advisor dashboard scope selection

Another useful feature of the dashboard widgets is the ability to double-click most of the widget metrics to see more information or a graph of the metric.

Figure 13 shows the ITW port widget. By double-clicking the port name, a chart that shows the ITW occurrences opens. Figure 13 also shows the Host port ITW widget that shows 427 ITWs.

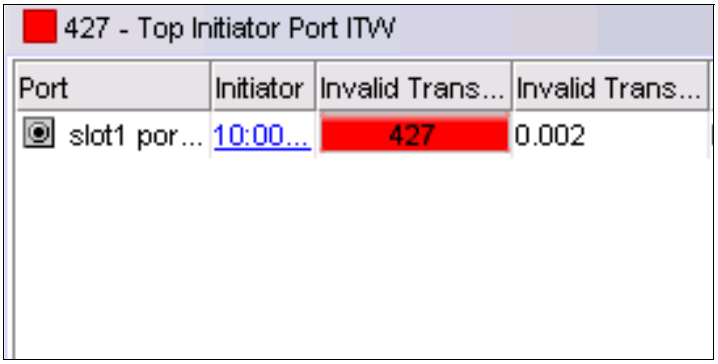


Figure 13 Network Advisor dashboard ITW widget

Figure 14 shows the ITW graph after double-clicking the port on the ITW widget.

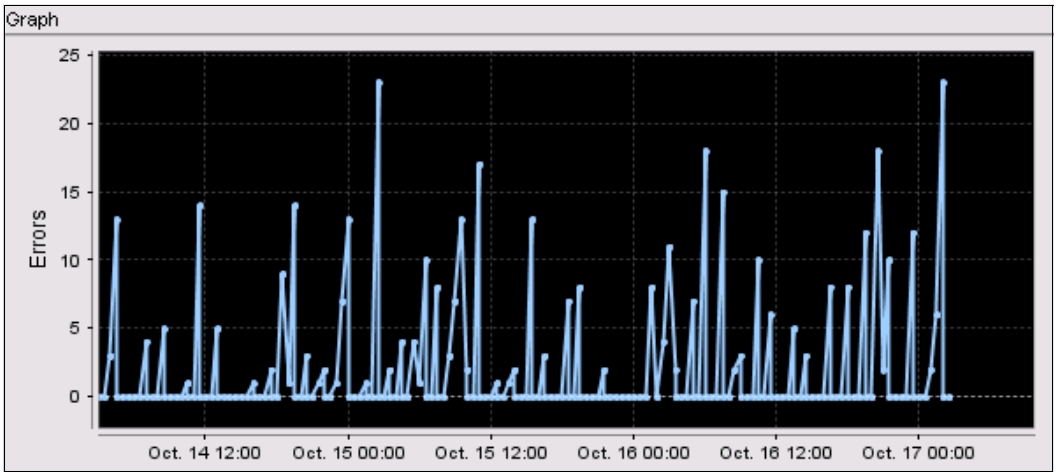


Figure 14 Network Advisor Dashboard ITW graph

As a preferred practice, create a customized dashboard to monitor the overall fabric health, and a dashboard to monitor port metrics. Optionally, create specialized custom dashboards to show port metrics for storage devices, server ports, and ISLs. These dashboards are used during major incidents, and can help identify whether a storage port, server port, or ISL is causing a problem.

Preferred settings

Create Fabric Health and Ports dashboards with the widgets that are listed in Table 5.

Table 5 Suggested Product Status and Ports dashboards widgets

Dashboard name	Status widget	Performance widget
Fabric Health	SAN Inventory	Top Product CPU
	Events	Top Product Memory
	Status	
	Out of Range Violations	
	Port Health Violations	
	Custom Events	
	Bottlenecked Ports	
All Ports	Port Health Violations	Top Port C3 Discards
		Top Port C3 Discards RX TO
		Top Port C3 Discards TX TO
		Top Port CRC
		Top Port ITW
		Top Port Link Failures
		Top Port Link Resets
		Top Port PCS Block Errors
All Ports (continued)	Port Health Violations	Top Port Sync Losses
		Top Port Utilization Percent

Optionally, create Host, Storage, and ISL port dashboards (see Table 6), which can be useful when determining problems.

Table 6 Optional Host, Storage, and ISL dashboard widgets

Dashboard	Status widget	Performance widget
Host Ports	Initiator Port Health Violations	Top Initiator Ports ITWs
	Initiator Bottleneck Ports	Top Initiator Port Link Failures
		Top Initiator Ports C3 Discards RX TO
		Top Initiator Ports C3 Discards TX TO
		Top Initiator Port CRC Errors
		Top Initiator Port Link Resets
		Top Initiator PCS Block Errors
		Top Initiator Port Sync Losses
		Top Initiator Port Utilization
Storage Ports	Target Port Health Violations	Top Target Ports C3 Discards RX TO

Dashboard	Status widget	Performance widget
	Target Bottleneck Ports	Top Target Ports C3 Discards TX TO
		Top Target Ports CRC Errors
		Top Target Ports Encode Error
		Top Target Port Link Failures
		Top Target Port Sync Loss
		Top Target Port Link Resets
		Top Target Port ITWs
		Top Target Ports PCS Block Errors
ISL Ports	ISL Port Health Violations	Top ISL Port Utilization
	ISL Bottleneck Ports	Top ISL Ports CRC Errors
		Top ISL Ports Encode Error
		Top ISL Port Link Failures
		Top ISL Port Sync Loss
		Top ISL Port Link Resets
		Top ISL Ports C3 Discards RX TO
		Top ISL Ports C3 Discards TX TO

Creating FCIP Dashboard

You can use the built-in WAN Vision dashboard or create a custom FCIP Dashboard by using the widgets that are listed in Table 7 to monitor fabric FCIP metrics.

Table 7 Optional FCIP dashboard widgets

Dashboard	Status widgets	Performance widgets
FCIP		Top Circuit FC Extension Utilization
		Top Circuit IP Extension Utilization
		Top Circuit Jitter
		Top Circuit Utilization
		Top Duplicate Ack
		Top Slow Start
		Top Tunnel Dropped Packets
		Top Tunnel Utilization

Enabling Monitoring Alerting Policy Suite

On switches running FOS V7.2 or later, use the MAPS to monitor the switches over Fabric Watch. To ensure that consistent settings and policies are enabled on all switches in the fabric, use Network Advisor.

Best practice: Enable MAPS monitoring and alerting. For more information about how to get started with MAPS by using SANnav, see this IBM Support video.

Enabling Monitoring Alerting Policy Suite

To enable MAPS, complete the following steps on switches that are running FOS V7.x. (Switches running FOS V8.x have MAPS always enabled):

1. Log in to Network Advisor and click **Monitor** → **Fabric Vision** → **MAPS** → **Enable**.

   Figure 15 shows the Network Advisor menu options to enable MAPS.

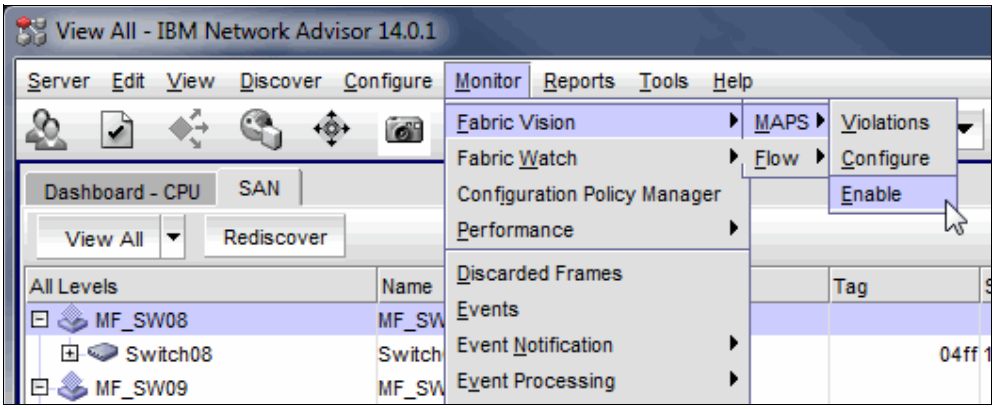


Figure 15 Network Advisor MAPS Enable menu

2. Select the switches that you want to enable MAPS on by selecting them in the Available Switches pane and click the right arrow to move them to the Selected Switches pane. After all of the switches that you want to enable MAPS on are selected, click **OK** to enable MAPS on those switches.

Figure 16 shows the Network Advisor MAPS enable switch selection window.

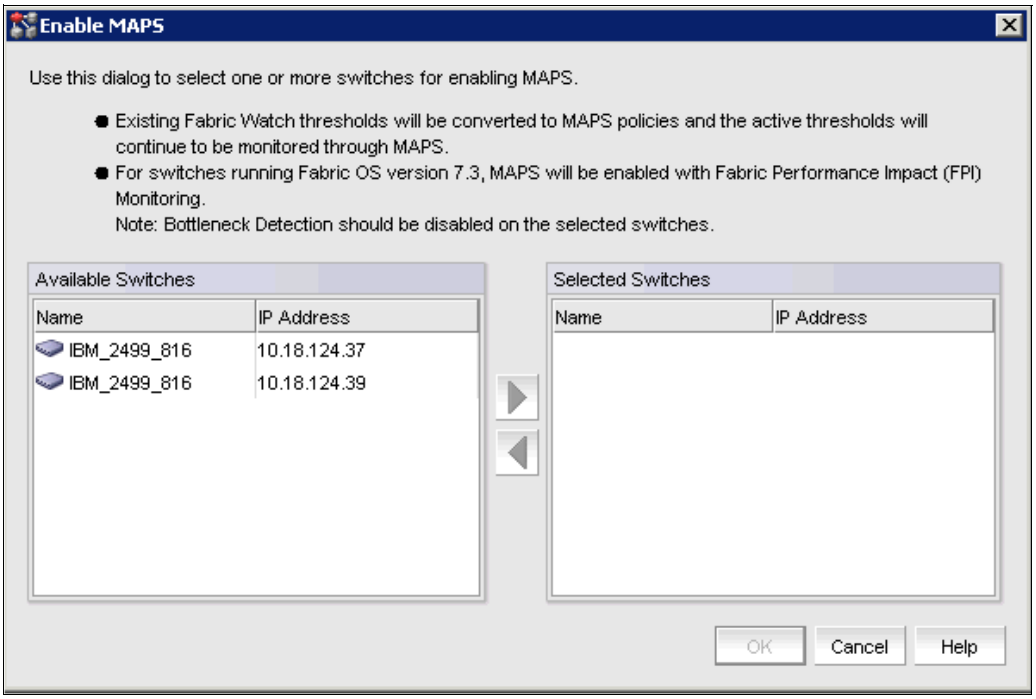


Figure 16 Network Advisor MAPS enable switch selection

Configuring MAPS Actions

To configure MAPS Actions, complete the following steps:

1. After MAPS is enabled, click **Monitor** → **Fabric Vision** → **MAPS** → **Configure**.

   Figure 17 shows the Network Advisor menu options to open the MAPS configuration window.

Figure 17 Network Advisor MAPS open Configure menu

2. In the MAPS Configuration window, select the switches on which to enable policy actions. To select multiple switches, hold the Ctrl key while selecting the switches. After the switches are selected, click **Actions**.

Figure 18 shows the MAPS Configuration window.

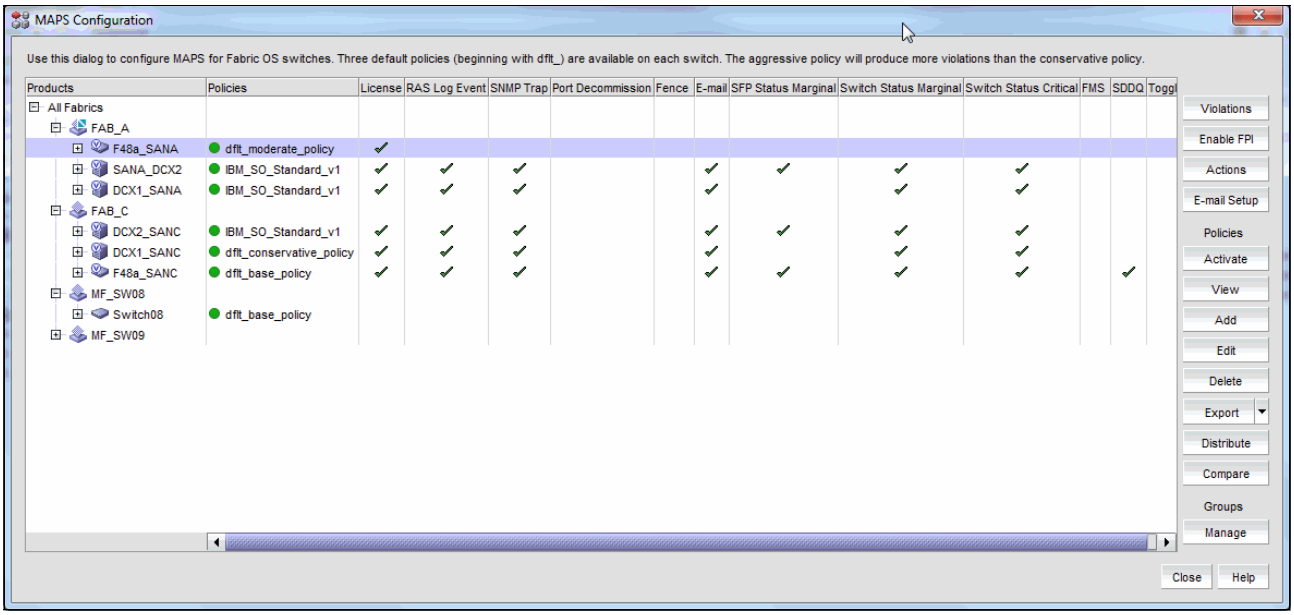


Figure 18 Network Advisor MAPS Configuration

3. In the MAPS Policy Actions dialog box, select the **RAS Log Event**, **SNMP Trap**, **E-mail**, **Switch Status Marginal**, **Switch Status Critical**, and **SFP Status Marginal** checkboxes. For switches with FOS V7.4 and later, select **FPI Actions** and **SDDQ**. Click **OK**.

> **Note:** Do not enable the Fence action.

Figure 19 shows the MAPS Policy Actions dialog box.

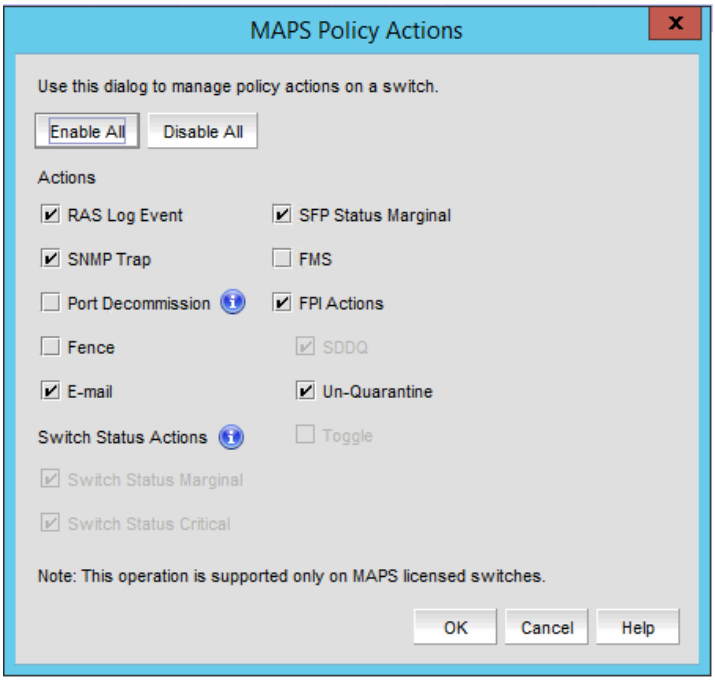


Figure 19 Network Advisor MAPS Policy Actions

Enabling MAPS default policy

In the MAPS Configuration dialog box, expand the list of available policies for each of the switches. Select the `dft_conservative_policy` for each switch. To select a policy for each switch, hold the Ctrl key while selecting the policies. After policies for each switch are selected, click **Activate**. Figure 20 shows the MAPS Configuration dialog box with `dft_conservative_policy` selected.

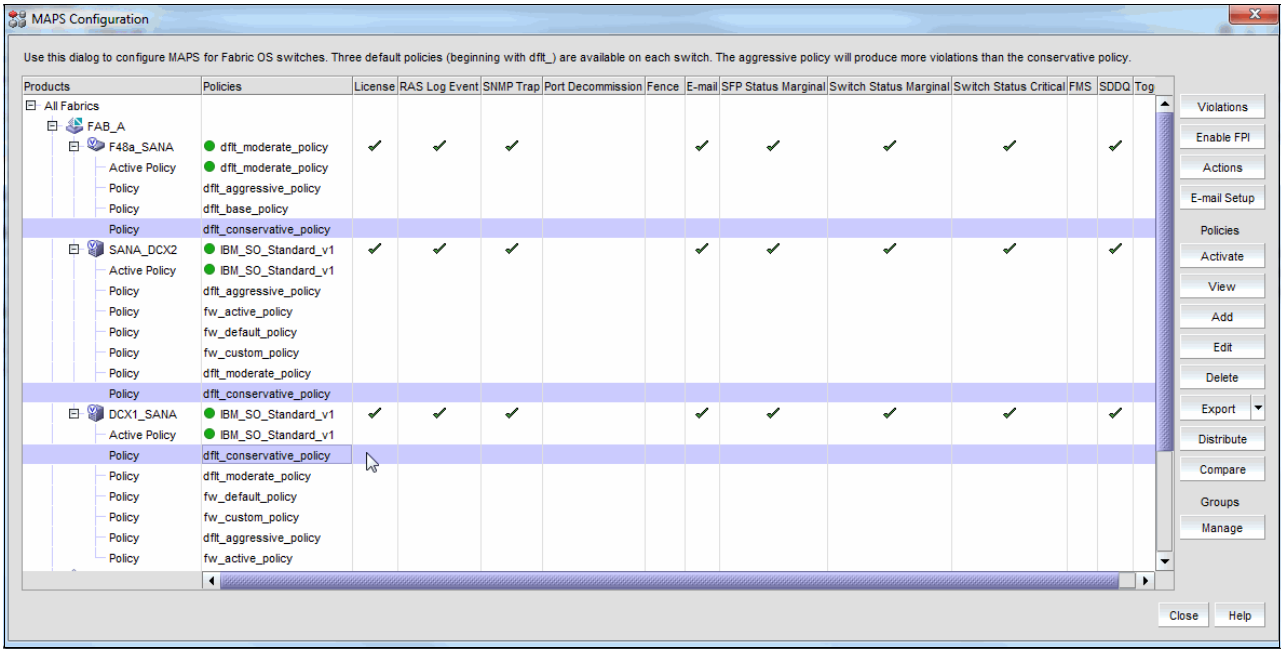


Figure 20 Network Advisor MAPS Configuration with conservative policy selected

MAPS is now enabled with the default conservative policy.

Creating a custom MAPS policy for high availability fabrics

For switches that require strict monitoring to provide highly available fabrics, implement custom MAPS rules for the port and switch thresholds that are listed in the tables in this section by completing the following steps:

1. Create a custom MAPS policy by making a copy of the default moderate policy by running the **`mapspolicy --clone`** command, as shown in Example 25.

   Example 25 The mapspolicy clone command

   ```
   mapsPolicy --clone dflt_moderate_policy -name IBM_Custom_policy
   ```

2. Create custom rules for the port metrics that are listed in Table 8.

Table 8 Custom MAPS port metrics threshold

Type	Time	OP	RASLOG	Alert	Fence (host only)
C3TXTO	min	ge	3	20	
CRC	min	ge	10	20	40
CRC	hour	ge			240
ITW	min	ge	20	40	

Type	Time	OP	RASLOG	Alert	Fence (host only)
LF	min	ge	3	5	
LOSS_SIGNAL	min	ge	3		
LOSS_SYNC	min	ge	3		
LR	min	ge	5	10	20
LR	hour	ge			60
PE	min	ge	3	7	
RX/TX/UTIL	hour	ge	75	90	
STATE_CHG	min	ge	5	10	

3. Display existing rules in the newly created customer policy by running the **mapspolicy --show <policyname>** command, as shown in Example 26.

Example 26 The mapspolicy show command for default CRC metrics

```
F48a_Default:FID128:dlutz> mapsPolicy --show IBM_Custom_policy | grep -i CRC

defALL_D_PORTSCRC_2 |ALL_D_PORTS(CRC/MIN>2) |RASLOG,SNMP,EMAIL |
defALL_D_PORTSCRC_D1000 |ALL_D_PORTS(CRC/DAY>1000) |RASLOG,SNMP,EMAIL |
defALL_D_PORTSCRC_H60 |ALL_D_PORTS(CRC/HOUR>60) |RASLOG,SNMP,EMAIL
defALL_E_PORTSCRC_10 |ALL_E_PORTS(CRC/MIN>10) |RASLOG,SNMP,EMAIL |
defALL_E_PORTSCRC_20 |ALL_E_PORTS(CRC/MIN>20) |FENCE,DECOM,SNMP,EMAIL     |
defALL_HOST_PORTSCRC_10 |ALL_HOST_PORTS(CRC/MIN>10) |RASLOG,SNMP,EMAIL |
defALL_HOST_PORTSCRC_20 |ALL_HOST_PORTS(CRC/MIN>20) |FENCE,DECOM,SNMP,EMAIL |
defALL_OTHER_F_PORTSCRC_10 |ALL_OTHER_F_PORTS(CRC/MIN>10) |RASLOG,SNMP,EMAIL |
defALL_OTHER_F_PORTSCRC_20|ALL_OTHER_F_PORTS(CRC/MIN>20)|FENCE,DECOM,SNMP,EMAIL
defALL_TARGET_PORTSCRC_10 |ALL_TARGET_PORTS(CRC/MIN>10) |FENCE,DECOM,SNMP,EMAIL
defALL_TARGET_PORTSCRC_5 |ALL_TARGET_PORTS(CRC/MIN>5) |RASLOG,SNMP,EMAIL |
defNON_E_F_PORTSCRC_10 |NON_E_F_PORTS(CRC/MIN>10) |RASLOG,SNMP,EMAIL |
defNON_E_F_PORTSCRC_20 |NON_E_F_PORTS(CRC/MIN>20) |FENCE,SNMP,EMAIL |
```

4. Remove any existing default maps rules for metrics that are shown in the tables in this section by running the **mapspolicy --delrule** commands, as shown in Example 27.

Example 27 mapspolicy delrule for default CRC metrics

```
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_D_PORTSCRC_2
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_D_PORTSCRC_D1000
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_D_PORTSCRC_H60
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_E_PORTSCRC_10
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_E_PORTSCRC_20
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_HOST_PORTSCRC_10
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_HOST_PORTSCRC_20
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_OTHER_F_PORTSCRC_10
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_OTHER_F_PORTSCRC_20
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_TARGET_PORTSCRC_10
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_TARGET_PORTSCRC_5
mapsPolicy --delrule IBM_Custom_policy -rulename defNON_E_F_PORTSCRC_10
```

5. Create custom rules by running the **mapsrule --create** command, and add these rules to the newly created custom policy, as shown in Example 28.

Example 28 mapsrule create for custom CRC metrics

```
mapsRule --create ALL_E_CRC_log -group ALL_E_PORTS -monitor CRC -op ge
-timebase MIN -value 10 -action RASLOG -policy IBM_Custom_policy
mapsRule --create ALL_E_CRC_alert -group ALL_E_PORTS -monitor CRC -op ge
-timebase MIN -value 20 -action RASLOG,EMAIL,SNMP -policy IBM_Custom_policy
mapsRule --create ALL_TARGET_CRC_log -group ALL_TARGET_PORTS -monitor CRC -op
ge -timebase MIN -value 10 -action RASLOG -policy IBM_Custom_policy
mapsRule --create ALL_TARGET_CRC_alert -group ALL_TARGET_PORTS -monitor CRC -op
ge -timebase MIN -value 20 -action RASLOG,EMAIL,SNMP -policy IBM_Custom_policy
mapsRule --create ALL_OTHER_F_CRC_log -group ALL_OTHER_F_PORTS -monitor CRC -op
ge -timebase MIN -value 10 -action RASLOG -policy IBM_Custom_policy
mapsRule --create ALL_OTHER_F_CRC_alert -group ALL_OTHER_F_PORTS -monitor CRC
-op ge -timebase MIN -value 20 -action RASLOG,EMAIL,SNMP -policy
IBM_Custom_policy
mapsRule --create NON_E_F_CRC_log -group NON_E_F_PORTS -monitor CRC -op ge
-timebase MIN -value 10 -action RASLOG -policy IBM_SO_Standard_v81_F48
mapsRule --create NON_E_F_CRC_alert -group NON_E_F_PORTS -monitor CRC -op ge
-timebase MIN -value 20 -action RASLOG,EMAIL,SNMP -policy IBM_Custom_policy
mapsRule --create ALL_HOST_CRC_log -group ALL_HOST_PORTS -monitor CRC -op ge
-timebase MIN -value 10 -action RASLOG -policy IBM_SO_Standard_v81_F48
mapsRule --create ALL_HOST_CRC_alert -group ALL_HOST_PORTS -monitor CRC -op ge
-timebase MIN -value 20 -action RASLOG,EMAIL,SNMP -policy IBM_Custom_policy
mapsRule --create ALL_HOST_CRC_fence -group ALL_HOST_PORTS -monitor CRC -op ge
-timebase MIN -value 40 -action RASLOG,EMAIL,SNMP,DECOM,FENCE -policy
IBM_Custom_policy
mapsRule --createRoR ALL_HOST_CRC_ror -group ALL_HOST_PORTS -monitor
ALL_HOST_CRC_alert -op ge -timebase HOUR -value 10 -action
RASLOG,EMAIL,SNMP,DECOM,FENCE -policy IBM_Custom_policy
mapsRule --create ALL_D_PORTSCRC_3_log -group ALL_D_PORTS -monitor CRC -op ge
-timebase MIN -value 3 -action RASLOG -policy IBM_Custom_policy
mapsRule --create ALL_D_PORTSCRC_D1500_log -group ALL_D_PORTS -monitor CRC -op
ge -timebase DAY -value 1500 -action RASLOG -policy IBM_Custom_policy
mapsRule --create ALL_D_PORTSCRC_H90_log -group ALL_D_PORTS -monitor CRC -op ge
-timebase HOUR -value 90 -action RASLOG -policy IBM_Custom_policy
```

Figure 21 shows the custom port metrics in Network Advisor policy editor.

Rules

Groups / Rules	Severity	Monitor Condition	Time Base	RAS Log Event	SNMP Trap	E-mail	FMS	Port Decommission	Fence	SFP Status Margin
ALL_E_PORTS										
ALL_E_C3TXTO_log	⚠	C3TXTO >= 3	Min	✓						
ALL_E_C3TXTO_alert	⚠	C3TXTO >= 20	Min	✓	✓	✓				
ALL_E_CRC_log	⚠	CRC >= 10	Min	✓						
ALL_E_CRC_alert	⚠	CRC >= 20	Min	✓	✓	✓				
ALL_E_ITW_log	⚠	ITW >= 21	Min	✓						
ALL_E_ITW_alert	⚠	ITW >= 40	Min	✓	✓	✓				
ALL_E_LOSS_SYNC_log	⚠	LOSS_SYNC >= 3	Min	✓						
ALL_E_LR_log	⚠	LR >= 5	Min	✓						
ALL_E_LR_alert	⚠	LR >= 10	Min	✓	✓	✓				
ALL_E_PE_log	⚠	PE >= 3	Min	✓						
ALL_E_PE_alert	⚠	PE >= 7	Min	✓	✓	✓				
ALL_E_STATE_CHG_log	⚠	STATE_CHG >= 5	Min	✓						
ALL_E_STATE_CHG_alert	⚠	STATE_CHG >= 10	Min	✓	✓	✓				
ALL_F_PORTS										
defALL_F_PORTSDEV_NPIV_LOGINS_PER_90	⚠	DEV_NPIV_LOGINS > 90 %	None	✓	✓	✓				
ALL_HOST_PORTS										
ALL_HOST_C3TXTO_log	⚠	C3TXTO >= 3	Min	✓						
ALL_HOST_C3TXTO_alert	⚠	C3TXTO >= 20	Min	✓	✓	✓				
ALL_HOST_CRC_log	⚠	CRC >= 10	Min	✓						
ALL_HOST_CRC_alert	⚠	CRC >= 20	Min	✓	✓	✓				
ALL_HOST_CRC_ror	⚠	ALL_HOST_CRC_alert >= 10	Hour	✓	✓	✓		✓	✓	
ALL_HOST_CRC_fence	⚠	CRC >= 40	Min	✓	✓	✓		✓	✓	
ALL_HOST_ITW_log	⚠	ITW >= 21	Min	✓						
ALL_HOST_ITW_alert	⚠	ITW >= 40	Min	✓	✓	✓				
ALL_HOST_LOSS_SYNC_log	⚠	LOSS_SYNC >= 3	Min	✓						
ALL_HOST_LR_log	⚠	LR >= 5	Min	✓						
ALL_HOST_LR_alert	⚠	LR >= 10	Min	✓	✓	✓				
ALL_HOST_LR_ror	⚠	ALL_HOST_LR_alert >= 10	Hour	✓	✓	✓		✓	✓	
ALL_HOST_LR_fence	⚠	LR >= 20	Min	✓	✓	✓		✓	✓	
ALL_HOST_PE_log	⚠	PE >= 3	Min	✓						

Figure 21 Network Advisor ports policy editor

6. Create custom rules for the switch fabric metrics that are listed in Table 9 to alert on any E_Port down or fabric segmentation event.

Table 9 Custom MAPS switch fabric metrics

Type	Timebase	OP	RASLOG alert
EPORT_DOWN	min	ge	1
FAB_SEG	min	ge	1

7. Remove any existing default maps rules for metrics by running the **`mapspolicy --delrule`** commands, as shown in Example 29.

Example 29 The mapsPolicy delrule command for default switch fabric metrics

```
mapsPolicy --delrule IBM_Custom_policy -rulename defSWITCHFAB_SEG_4
mapsPolicy --delrule IBM_Custom_policy -rulename defSWITCHEPORT_DOWN_4
```

8. Create custom rules by running the **mapsrule --create** command, and add these rules to the newly created custom policy, as shown in Example 30.

Example 30 The mapsrule create command for custom switch fabric metrics

```
mapsRule --create SWITCH_EPORT_DOWN_alert -group SWITCH -monitor EPORT_DOWN -op
ge -timebase MIN -value 1 -action RASLOG,EMAIL,SNMP -policy IBM_Custom_policy
mapsRule --create SWITCH_FAB_SEG_alert -group SWITCH -monitor FAB_SEG -op ge
-timebase MIN -value 1 -action RASLOG,EMAIL,SNMP -policy IBM_Custom_policy
```

Figure 22 shows the custom switch fabric metrics in Network Advisor policy editor.

Rules

Groups / Rules	Severity	Monitor Condition	Time Base	RAS Log Event	SNMP Trap	E-mail	FMS	Quiet Time
System Groups								
SWITCH								
defSWITCHL2_DEVCNT_PER_90	⚠	L2_DEVCNT_PER > 90 %	None	✓	✓	✓		
defSWITCHZONE_CHG_10	⚠	ZONE_CHG > 10	Day	✓	✓	✓		
defSWITCHLSAN_DEVCNT_PER_90	⚠	LSAN_DEVCNT_PER > 90 %	None	✓	✓	✓		
defSWITCHFAB_CFG_4	⚠	FAB_CFG > 4	Min	✓	✓	✓		
SWITCH_FAB_SEG_alert	⚠	FAB_SEG >= 1	Min	✓	✓	✓		
defSWITCHZONE_CFGSZ_PER_90	⚠	ZONE_CFGSZ_PER > 90 %	None	✓	✓	✓		
defSWITCHDID_CHG_1	⚠	DID_CHG > 1	Min	✓	✓	✓		
defSWITCHBB_FCR_CNT_MAX	⚠	BB_FCR_CNT > 16	None	✓	✓	✓		
SWITCH_EPORT_DOWN_alert	⚠	EPORT_DOWN >= 1	Min	✓	✓	✓		

Figure 22 Network Advisor switch fabric policy editor

Many fabric performance impact events can exist that you want to be aware of but that do not warrant real-time alerting. Set up a set of custom FPI MAPS rules to log only FPI events for periodic review to identify congesting workloads, or for analysis of a performance issue.

9. Remove any existing default maps FPI rules by running the **mapspolicy --delrule** commands, as shown in Example 31.

Example 31 The mapspolicy delrule command for default FPI metrics

```
mapsPolicy --delrule IBM_Custom_policy -rulename
defALL_PORTS_IO_FRAME_LOSS_UNQUAR
mapsPolicy --delrule IBM_Custom_policy -rulename defALL_PORTS_IO_LATENCY_CLEAR
mapsPolicy --delrule IBM_Custom_policy -rulename
defALL_PORTS_IO_PERF_IMPACT_UNQUAR
```

10. Create custom rules by running the **mapsrule --create** command, and add these rules to the newly created custom policy, as shown in Example 32.

Example 32 The mapsrule create command for custom FPI metrics for FOS V7.x

```
mapsRule --create FPI_IO_PERF_IMPACT_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_PERF_IMPACT -action RASLOG
-policy IBM_Custom_policy
mapsRule --create FPI_IO_FRAME_LOSS_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_FRAME_LOSS -action RASLOG
-policy IBM_Custom_policy
mapsRule --create FPI_IO_LATENCY_CLEAR_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_LATENCY_CLEAR -action RASLOG
-policy IBM_Custom_policy
```

Figure 23 shows the custom FPI metrics in Network Advisor policy editor.

Rules

Groups / Rules	Severity	Monitor Condition	Time Base	RAS Log Event	SNMP Trap	E-mail	FMS	Toggle (Duration)	SDDQ	Un-Quaran
ALL_PORTS										
FPI_IO_FRAME_LOSS_sddq		DEV_LATENCY_IMPACT == IO_FRAME_LOSS	None	✓					✓	30 Minutes
FPI_IO_FRAME_LOSS_ror		FPI_IO_FRAME_LOSS_sddq > 5	Day	✓	✓	✓				
FPI_IO_LATENCY_CLEAR_log		DEV_LATENCY_IMPACT == IO_LATENCY_CLEAR	None	✓						
FPI_IO_PERF_IMPACT_log		DEV_LATENCY_IMPACT == IO_PERF_IMPACT	None	✓						
ALL_E_PORTS										
ALL_E_UTIL_alert		UTIL >= 90 %	Hour	✓	✓	✓				
ALL_E_UTIL_log		UTIL >= 75 %	Hour	✓						
ALL_E_RX_log		RX >= 75 %	Hour	✓						
ALL_E_TX_alert		TX >= 90 %	Hour	✓	✓	✓				
ALL_E_RX_alert		RX >= 90 %	Hour	✓	✓	✓				
ALL_E_TX_log		TX >= 75 %	Hour	✓						

Figure 23 Network Advisor FPI policy editor

With FOS V8.x, the SDDQ unquarantine action was added. This action allows MAPS rules to be set up to automatically quarantine a congesting workload, and then, have the quarantine removed after a specified time, as shown in Example 33.

Example 33 The mapsrule create command for custom FPI metrics for FOS 8.x

```
mapsRule --create FPI_IO_PERF_IMPACT_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_PERF_IMPACT -action RASLOG
-policy IBM_Custom_policy
mapsRule --create FPI_IO_FRAME_LOSS_sddq -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_FRAME_LOSS -action
RASLOG,SDDQ,UNQUAR -uqrt=30 -uqrt_unit min -policy IBM_Custom_policy
mapsRule --create FPI_IO_LATENCY_CLEAR_log -group ALL_PORTS -monitor
DEV_LATENCY_IMPACT -op eq -timebase NONE -value IO_LATENCY_CLEAR -action RASLOG
-policy IBM_Custom_policy
```

Note: Monitoring FCIP extension metrics is critical for the proper operation of FCIP tunnels.

11. Create a set of custom rules to log marginal FCIP tunnel behavior and alert when tunnel performance approaches impactive levels based on Table 10.

Table 10 Custom maps FCIP metrics thresholds

Type	Time	Op	RASLOG	Alert
PKTLOSS	min	ge	0.05	0.1
JITTER	min	ge	10	20
RTT	none	ge	see note	see note

12. Create custom rules by running the **`mapsrule --create`** command, and add these rules to the newly created custom policy for each of the QoS Circuit and Tunnel groups, as shown in Example 34.

Example 34 mapsrule create for custom FCIP metrics

```
mapsRule --create ALL_CIRCUIT_MED_QOS_PKTLOSS_log -group ALL_CIRCUIT_MED_QOS
-monitor PKTLOSS -op ge -timebase MIN -value 0.05 -action RASLOG -policy
IBM_Custom_policy
```

```
mapsRule --create ALL_CIRCUIT_MED_QOS_PKTLOSS_alert -group ALL_CIRCUIT_MED_QOS
-monitor PKTLOSS -op ge -timebase MIN -value 0.1 -action RASLOG,EMAIL,SNMP
-policy IBM_Custom_policy

mapsRule --create ALL_TUNNEL_MED_QOS_PKTLOSS_log -group ALL_TUNNEL_MED_QOS
-monitor PKTLOSS -op ge -timebase MIN -value 0.05 -action RASLOG -policy
IBM_Custom_policy
mapsRule --create ALL_TUNNEL_MED_QOS_PKTLOSS_alert -group ALL_TUNNEL_MED_QOS
-monitor PKTLOSS -op ge -timebase MIN -value 0.1 -action RASLOG,EMAIL,SNMP
-policy IBM_Custom_policy

mapsRule --create ALL_CIRCUITS_JITTER_log -group ALL_CIRCUITS -monitor JITTER
-op ge -timebase NONE -value 10 -action RASLOG -policy IBM_Custom_policy
mapsRule --create ALL_CIRCUITS_JITTER_alert -group ALL_CIRCUITS -monitor JITTER
-op ge -timebase NONE -value 20 -action RASLOG,EMAIL,SNMP -policy
IBM_Custom_policy
```

Figure 24 shows the custom extensions metrics in Network Advisor policy editor.

Rules

Groups / Rules	Monitor Condition	Time Base	RAS Log Event	SNMP Trap	E-mail
ALL_CIRCUITS					
defALL_CIRCUITS_IP_UTIL_P_90	IP_UTIL > 90 %	Min	✓	✓	✓
defALL_CIRCUITSCIR_PKTLOSS_PER_5	CIR_PKTLOSS > 0.5 %	Min	✓	✓	✓
defALL_CIRCUITSCIR_STATE_5	CIR_STATE > 5	Min	✓	✓	✓
defALL_CIRCUITS_IP_PKTLOSS_P_5	IP_PKTLOSS > 0.5 %	Min	✓	✓	✓
ALL_CIRCUITS_JITTER_alert	Jitter >= 20 mSec	None	✓	✓	✓
defALL_CIRCUITS_IP_RTT_250	IP_RTT > 250 mSec	None	✓	✓	✓
defALL_CIRCUITSCIR_UTIL_90	CIR_UTIL > 90 %	Min	✓	✓	✓
ALL_CIRCUITS_IP_JITTER_log	IP_JITTER >= 10 %	None	✓		
ALL_CIRCUITS_IP_JITTER_alert	IP_JITTER >= 20 %	None	✓	✓	✓
defALL_CIRCUITS_RTT_250	RTT > 250 mSec	None	✓	✓	✓
ALL_CIRCUITS_JITTER_log	Jitter >= 10 mSec	None	✓		
ALL_CIRCUIT_F_QOS					
ALL_CIRCUIT_HIGH_QOS					
ALL_CIRCUIT_MED_QOS					
ALL_CIRCUIT_MED_QOS_PKTLOSS_alert	PKTLOSS >= 0.1 %	Min	✓	✓	✓
ALL_CIRCUIT_MED_QOS_PKTLOSS_log	PKTLOSS >= 0.05 %	Min	✓		
defALL_CIRCUIT_MED_QOS_UTIL_PER_90	UTIL > 90 %	Hour	✓	✓	✓
ALL_CIRCUIT_LOW_QOS					
ALL_CIRCUIT_IP_LOW_QOS					
ALL_CIRCUIT_IP_MED_QOS					
ALL_CIRCUIT_IP_HIGH_QOS					
ALL_TUNNELS					
defALL_TUNNELSSTATE_CHG_3	TUNNEL_STATE > 3	Min	✓	✓	✓
defALL_TUNNELS_IP_UTIL_P_90	IP_UTIL > 90 %	Hour	✓	✓	✓
defALL_TUNNELSUTIL_PER_90	TUNNEL_UTIL > 90 %	Hour	✓	✓	✓
ALL_TUNNEL_F_QOS					
ALL_TUNNEL_LOW_QOS					
ALL_TUNNEL_MED_QOS					
ALL_TUNNEL_MED_QOS_PKTLOSS_alert	PKTLOSS >= 0.1 %	Min	✓	✓	✓

Figure 24 Network Advisor extensions policy editor

Configuring Network Advisor dashboards

Complete the following steps to configure Network Advisor dashboards:

1. Create Fabric Health and Ports dashboards by using the widgets that are listed in Table 11.

Table 11 Suggested Product Status and Ports dashboards widgets

Dashboard name	Status widget	Performance widget
Fabric Health	SAN Status	Top Product CPU
	SAN Inventory	Top Product Memory
	Events	
	Status	
	Out of Range Violations	
	Port Health Violations	
Ports	Port Health Violations	Top Port C3 Discards
		Top Port C3 Discards RX TO
		Top Port C3 Discards TX TO
		Top Port CRC
		Top Port ITW
		Top Port Link Failures
		Top Port Link Resets
		Top Port Too Long Errors
		Top Port Utilization Percent

2. Optionally, create a Host, Storage, and ISL port dashboards (which can be useful when doing problem determination), as listed in Table 12.

Table 12 Optional Host, Storage, and ISL dashboard widgets

Dashboard	Status widget	Performance widget
Host Ports	Initiator Port Health Violations	Top Initiator Ports C3 Discards
	Initiator Bottleneck Ports	Top Initiator Ports CRC Errors
		Top Initiator Ports Encode Error
		Top Initiator Port Link Failures
		Top Initiator Port Sync Loss
		Top Initiator Port Link Resets
		Top Initiator Port ITWs
Storage Ports	Target Port Health Violations	Top Target Ports C3 Discards
	Target Bottleneck Ports	Top Target Ports CRC Errors

Dashboard	Status widget	Performance widget
		Top Target Ports Encode Error
		Top Target Port Link Failures
		Top Target Port Sync Loss
		Top Target Port Link Resets
		Top Target Port ITWs
ISL Ports	ISL Port Health Violations	Top ISL Ports C3 Discards
	ISL Bottleneck Ports	Top ISL Ports CRC Errors
		Top ISL Ports Encode Error
		Top ISL Port Link Failures
		Top ISL Port Sync Loss
		Top ISL Port Link Resets
		Top ISL Ports C3 Discards RX TO
		Top ISL Ports C3 Discards TX TO

3. Create a dashboard by selecting the **My Dashboard** group and clicking **Add**.

   Figure 25 shows the Network Advisor window that is used to add a dashboard.

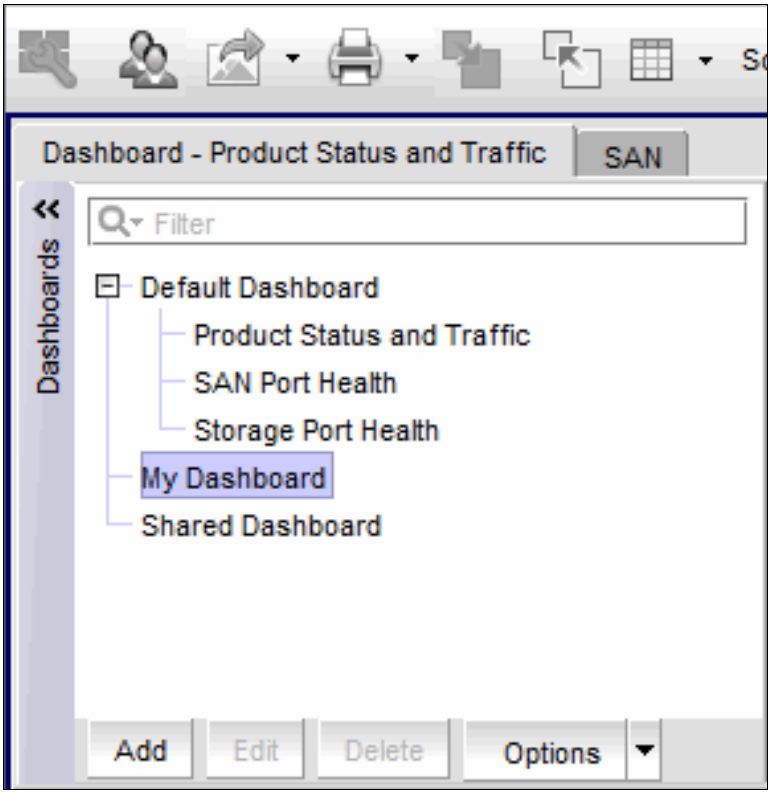


Figure 25 Network Advisor: Add dashboard

4. Enter the dashboard name in the Name entry field in the Add Dashboard dialog box and click **OK**.

   Figure 26 shows the Add Dashboard dialog box.

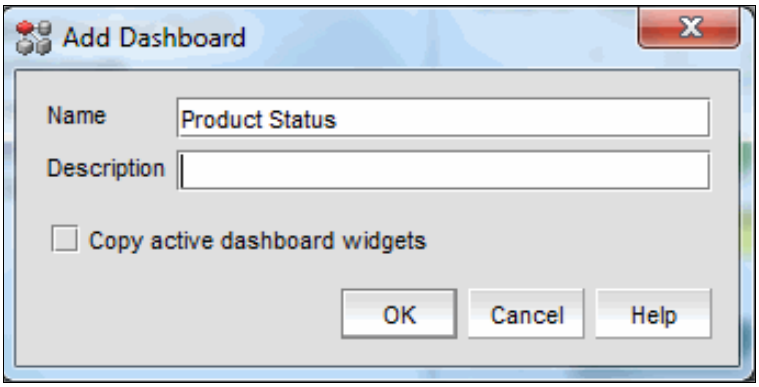


Figure 26 Network Advisor Add Dashboard dialog box

5. Use the Customize Dashboard tool to add widgets to the empty dashboard.

   Figure 27 shows the icon to start the Customize Dashboard tool.

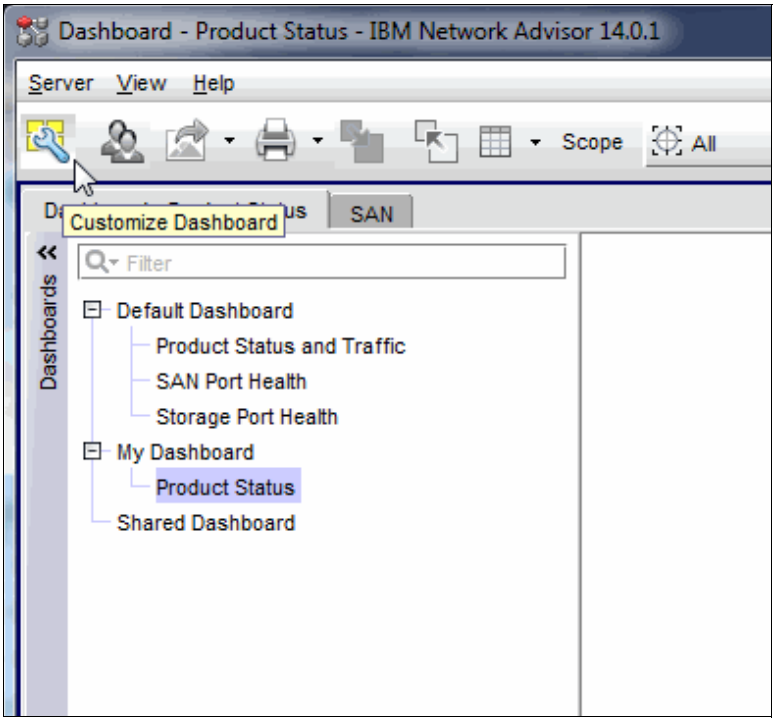
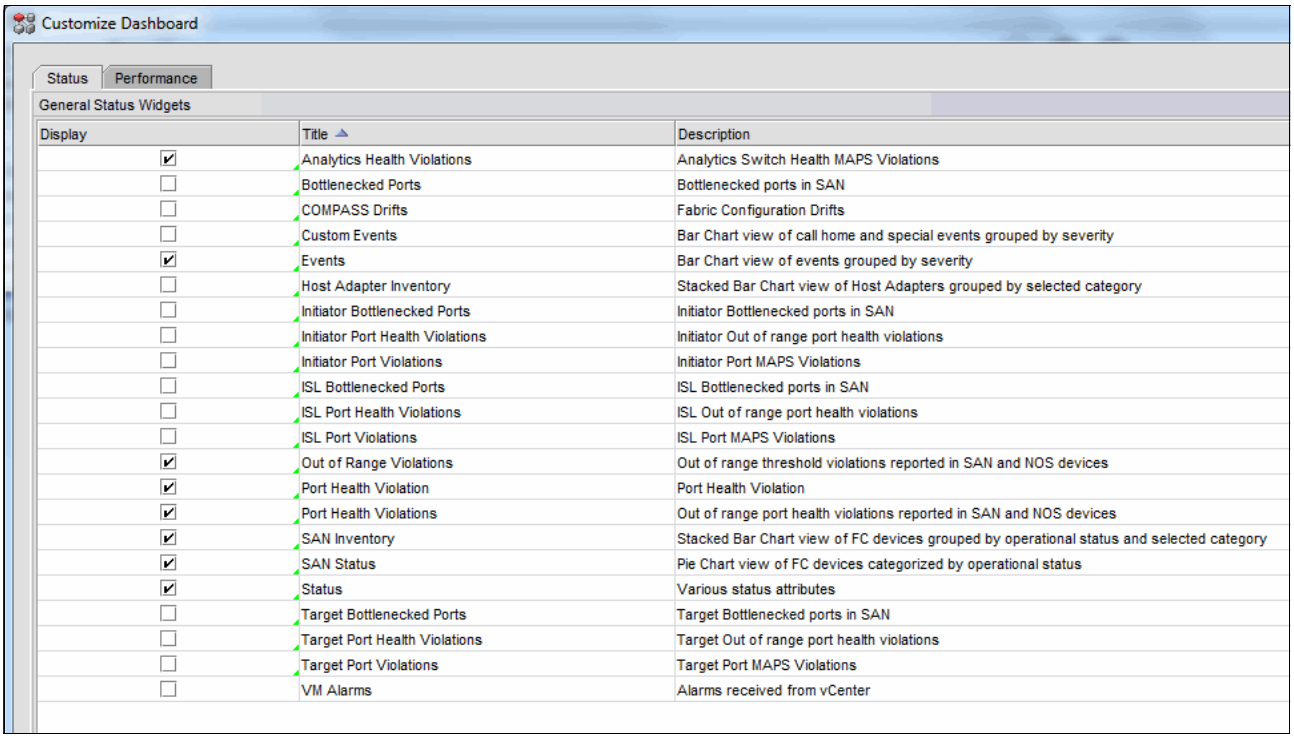


Figure 27 Network Advisor Customer Customize Dashboard icon

6. To add widgets to the dashboard, choose the required widgets by selecting the checkbox that is next to the widget titles.

Figure 28 shows the Customize Dashboard Status dialog box.

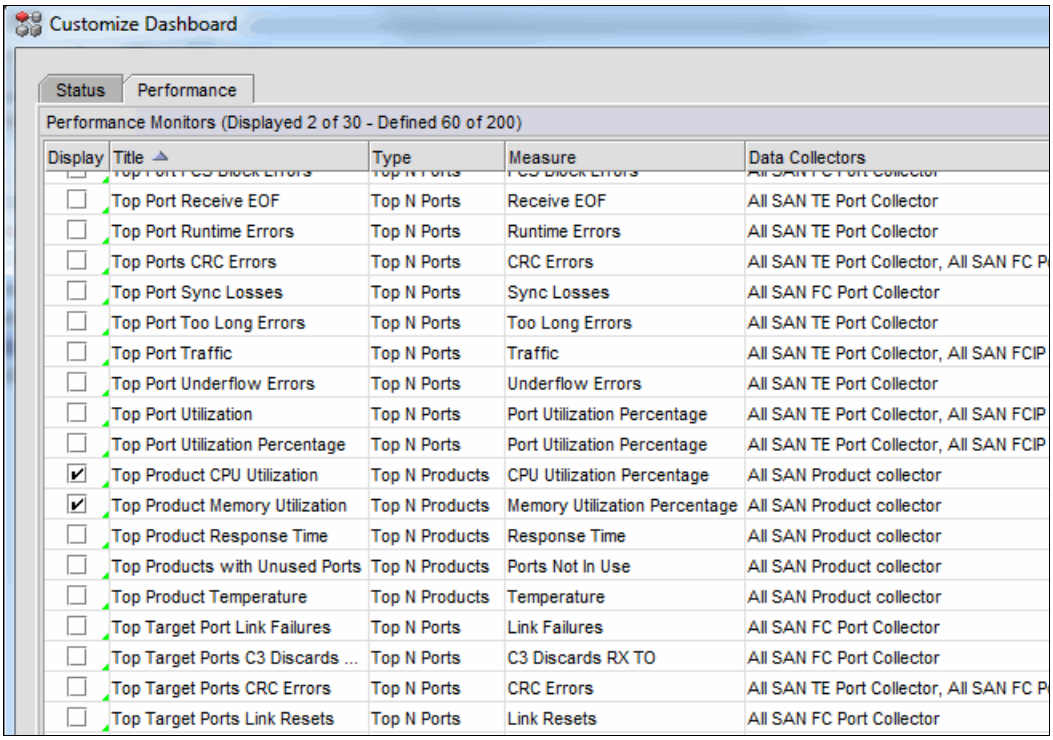


Display	Title	Description
☑	Analytics Health Violations	Analytics Switch Health MAPS Violations
☐	Bottlenecked Ports	Bottlenecked ports in SAN
☐	COMPASS Drifts	Fabric Configuration Drifts
☐	Custom Events	Bar Chart view of call home and special events grouped by severity
☑	Events	Bar Chart view of events grouped by severity
☐	Host Adapter Inventory	Stacked Bar Chart view of Host Adapters grouped by selected category
☐	Initiator Bottlenecked Ports	Initiator Bottlenecked ports in SAN
☐	Initiator Port Health Violations	Initiator Out of range port health violations
☐	Initiator Port Violations	Initiator Port MAPS Violations
☐	ISL Bottlenecked Ports	ISL Bottlenecked ports in SAN
☐	ISL Port Health Violations	ISL Out of range port health violations
☐	ISL Port Violations	ISL Port MAPS Violations
☑	Out of Range Violations	Out of range threshold violations reported in SAN and NOS devices
☑	Port Health Violation	Port Health Violation
☑	Port Health Violations	Out of range port health violations reported in SAN and NOS devices
☑	SAN Inventory	Stacked Bar Chart view of FC devices grouped by operational status and selected category
☑	SAN Status	Pie Chart view of FC devices categorized by operational status
☑	Status	Various status attributes
☐	Target Bottlenecked Ports	Target Bottlenecked ports in SAN
☐	Target Port Health Violations	Target Out of range port health violations
☐	Target Port Violations	Target Port MAPS Violations
☐	VM Alarms	Alarms received from vCenter

Figure 28 Network Advisor Customer Dashboard Status dialog box

7. Select the **Performance** tab to add performance widgets.

   Figure 29 shows the Customize Dashboard Performance dialog box.



Display	Title	Type	Measure	Data Collectors
☐	Top Port Receive EOF	Top N Ports	Receive EOF	All SAN TE Port Collector
☐	Top Port Runtime Errors	Top N Ports	Runtime Errors	All SAN TE Port Collector
☐	Top Ports CRC Errors	Top N Ports	CRC Errors	All SAN TE Port Collector, All SAN FC P
☐	Top Port Sync Losses	Top N Ports	Sync Losses	All SAN FC Port Collector
☐	Top Port Too Long Errors	Top N Ports	Too Long Errors	All SAN TE Port Collector
☐	Top Port Traffic	Top N Ports	Traffic	All SAN TE Port Collector, All SAN FCIP
☐	Top Port Underflow Errors	Top N Ports	Underflow Errors	All SAN TE Port Collector
☐	Top Port Utilization	Top N Ports	Port Utilization Percentage	All SAN TE Port Collector, All SAN FCIP
☐	Top Port Utilization Percentage	Top N Ports	Port Utilization Percentage	All SAN TE Port Collector, All SAN FCIP
☑	Top Product CPU Utilization	Top N Products	CPU Utilization Percentage	All SAN Product collector
☑	Top Product Memory Utilization	Top N Products	Memory Utilization Percentage	All SAN Product collector
☐	Top Product Response Time	Top N Products	Response Time	All SAN Product collector
☐	Top Products with Unused Ports	Top N Products	Ports Not In Use	All SAN Product collector
☐	Top Product Temperature	Top N Products	Temperature	All SAN Product collector
☐	Top Target Port Link Failures	Top N Ports	Link Failures	All SAN FC Port Collector
☐	Top Target Ports C3 Discards ...	Top N Ports	C3 Discards RX TO	All SAN FC Port Collector
☐	Top Target Ports CRC Errors	Top N Ports	CRC Errors	All SAN TE Port Collector, All SAN FC P
☐	Top Target Ports Link Resets	Top N Ports	Link Resets	All SAN FC Port Collector

Figure 29 Network Advisor Customize Dashboard Performance dialog box

Figure 30 shows the completed dashboard with widgets.

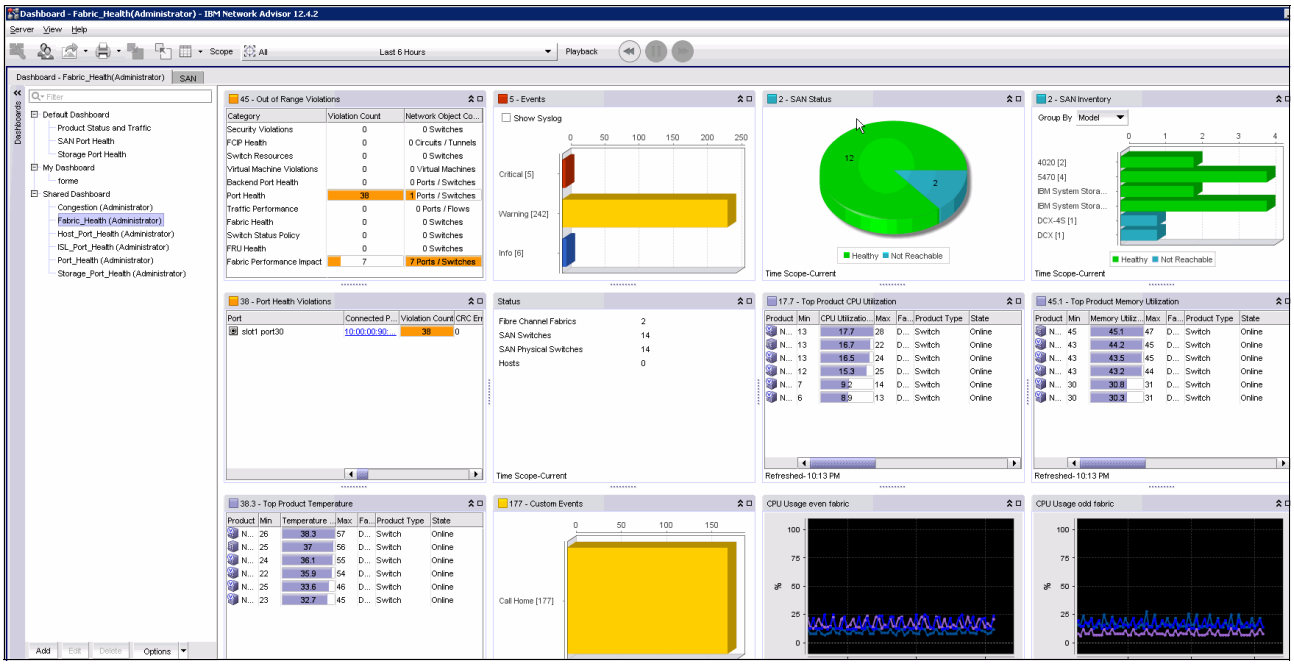


Figure 30 Network Advisor Fabric Health dashboard

Best practice: Configure and use Network Advisor dashboards.

SANnav dashboards

Dashboards are a powerful tool that are used to gain insights by eliminating the guess work and providing a seamless view of what occurring in the SAN fabric. The use of dashboards speeds up daily operations by reducing troubleshooting time and increasing the speed of resolution.

SANnav Management Portal offers a set of predefined monitoring dashboards that are aimed to satisfy the monitoring needs of most customers. Dashboards include the following key capabilities:

- Real-time monitoring of network health status and performance
- Predefined dashboards
- Customizable dashboard templates for static and dynamic views
- Immediately widgets to monitor switch and port status for error and performance statistics
- Customizable, more granular content that uses network scope and date range

The following predefined dashboard are available:

- Health Summary

  This dashboard monitors the health of all SAN entities (hosts, storage, switches, and fabrics) through established Broadcom best practices, MAPS violations, and hardware statuses. SANnav then provides insights and recommended actions to further assess the issue.

- Network Port Traffic Conditions

  Monitors all ports in the fabric for congestion while visually tracking its severity overtime. This dashboard also allows users to take a deeper look into the causes of such congestion through port and flow-level investigation.

- Extension

  This dashboard focuses on monitoring extension-related statistics, and provides insights into tunnel and circuit performance.

Figure 31 shows the default Health Summary dashboard.

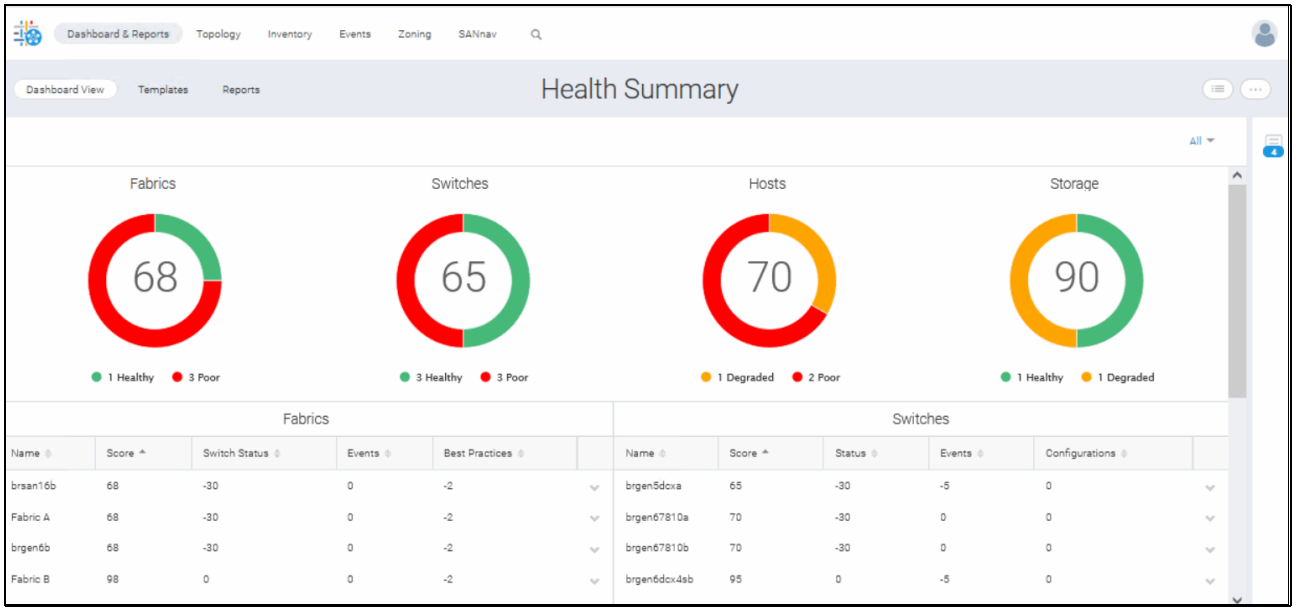


Figure 31 Default Health Summary dashboard

> **Note:** Real-time troubleshooting of congestion and oversubscription events with the Network Port Traffic Conditions dashboard requires Fabric OS version 8.2.1b or greater.
>
> For more information about how to investigate these events, see SANnav Management Portal and SANnav Global View.

As a best practice, creating custom dashboards is highly recommended. This task is performed by using system-defined product status and performance widgets. The default dashboard is the default view when you log in to SANnav.

The Health Summary dashboard is the default dashboard after installing SANnav. After custom dashboards are created, replace the default dashboard with the newly customized dashboard.

As a preferred practice, create a set of customized dashboards to monitor the overall fabric health, and a dashboard to monitor port health. Optionally, create specialized custom dashboards to show port metrics for storage devices, host server ports, and ISLs. These dashboards are used during major incident troubleshooting and investigation that can help identify whether a storage port, server port, or ISL is causing a problem.

> **Best practice:** Create custom SANnav dashboards.

More than 80 status and performance widgets are available.

Note: For more information about setting up dashboards and configuring dashboard widgets, see the SANnav Management Portal User Guide for your release by searching the Broadcom Document Library web page.

Create Fabric Health and Ports dashboards with the widgets that are listed in Table 13.

Table 13 Suggested Product Status and Ports dashboards widgets

Dashboard name	Status widgets	Performance widgets
Fabric Health	Events Summary	Top Product CPU Utilization
	Fabrics (Chart)	Top Product Memory Utilization
	FPI Violations	
	Out of Range Violations	
	Product Status	
All Ports	Port Health Violations	Top Port BB Credit Zero
		Top Port C3 Discard Rx Timeout
		Top Port C3 Discards
		Top Port C3 Discards Tx Timeouts
		Top Port CRC Errors
		Top Port CRC Errors with bad EOF
		Top Port Encode Error In
		Top Port Encode Erorr Out
		Top Port Frame Too Long Errors
		Top Port Invalid Transmissions
		Top Port Link Failures
		Top Port Link Resets
		Top Port Link Losses
		Top Port Traffic
		Top Port Utilization Percentage
		Top Port with Bad EOF
		Top Port PCS Block Errors

The predefined Extension dashboard can be used for FCIP monitoring. A custom dashboard does *not* need to be created.

Optionally, create Host, Storage, and ISL port dashboards, as listed in Table 6, which can be useful when performing problem determination.

Table 14 Optional Host, Storage and ISL dashboard widgets

Dashboard name	Status widgets	Performance widgets
Host Ports	Initiator Port Health Violations	Top Initiator Port Link Failures
	Initiator Port out of range violations	Top Initiator Port C3 Discards Rx Timeout
		Top Initiator Port CRC Errors
		Top Initiator Port Link Resets
		Top Initiator Port PCS Block Errors
		Top Initiator Port Sync Losses
		Top Initiator Port Utilization
Storage Ports	Target Port Health Violations	Top Target Port Link Failures
	Target Port Out Of Range Violations	Top Target Ports C3 Discard Rx Timeout
		Top Target Ports CRC Errors
		Top Target Ports Link Resets
		Top Target Ports PCS Block Errors
		Top Target Ports Sync Losses
		Top Target Port Utilization
ISL Ports	ISL Port Health Violations	Top ISL Port Link Failures
	ISL Port Out of Range Vilolations	Top ISL Ports C3 Discards Rx Timeout
		Top ISL Ports CRC Errors
		Top ISL Ports Link Resets
		Top ISL Ports PCS Block Errors
		Top ISL Port Sync Losses
		Top ISL Port Utilization

IBM Storage Insights

IBM Storage Insights is a cloud-based software tool that can collect performance data and metadata from IBM storage systems. The licensed version of the tool also can collect data from non-IBM storage systems.

IBM Storage Insights features extensive alerting and reporting capabilities. IBM Storage Insights also allows you to open tickets against monitored IBM storage directly from IBM Storage Insights if you need technical support. It also enables IBM Storage support to view performance data and collect support log data. It is recommend that you register for IBM Storage Insights and use it as part of your monitoring solution.

For more information about IBM Storage Insights, see this IBM Support web page.

IBM Storage Insights can monitor your storage and send you alerts for performance related problems and issues with specific devices. Some common alerts include response times for volumes, or the status of a storage system.

Figure 32 shows the IBM Storage Insights dashboard. In this dashboard, you can see that the monitored storage systems that encountered problems are listed at the top of the dashboard.

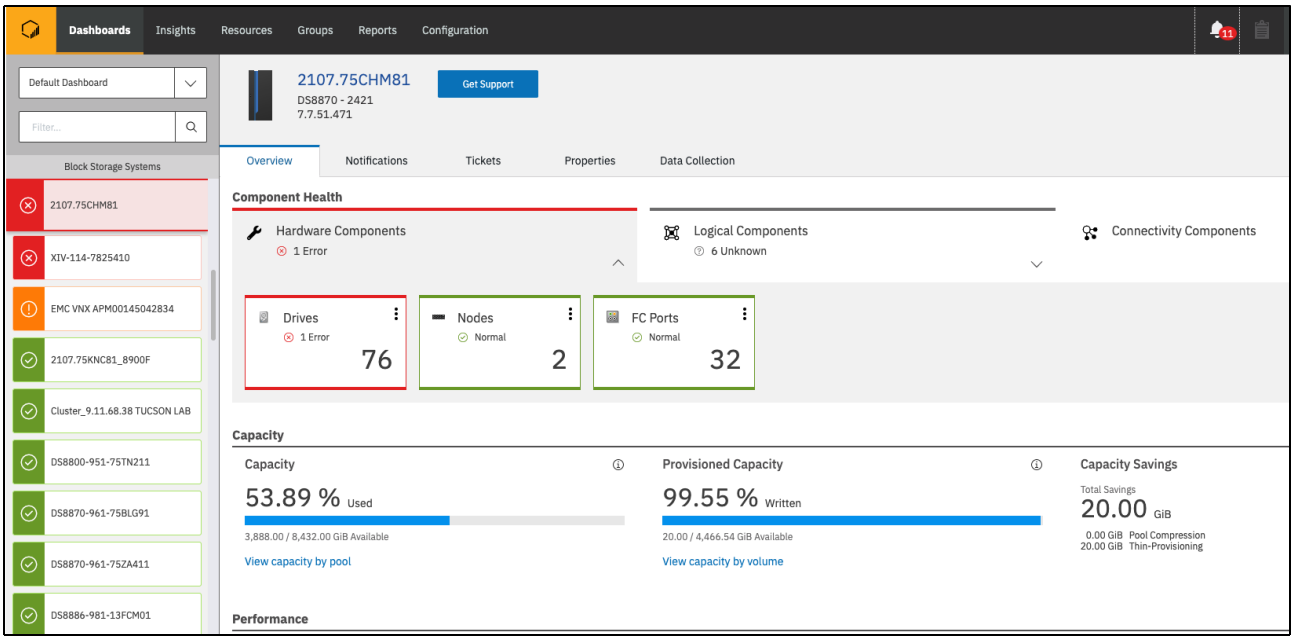


Figure 32 Storage Insights Dashboard

The dashboard also provides some information about which component of the storage system is a problem. You can use IBM Storage Insights to begin troubleshooting the problem and if necessary, open a ticket against the storage system. You also can configure alerts for that storage system so that you are notified of future problems with it. For more information, see this IBM Support web page.

IBM Storage Insights also supports monitoring Broadcom fabrics. Implementing the free version of IBM Storage Insights enhances the IBM support teams capabilities to provide quick problem resolutions by having quick access to switch information.

The Pro version of IBM Storage Insights more provides access to performance history and the ability to set up real-time alerting.

IBM Storage Insights and IBM Storage Insights Pro can collect metadata and performance data from your switches similar to storage systems provided performance data. You can configure alerts for your switches, fabrics, and switch ports to be alerted about error counters.

IBM Storage Insights should be a part of your monitoring strategy. Figure 33 shows a preview example of the list of physical switches in IBM Storage Insights.

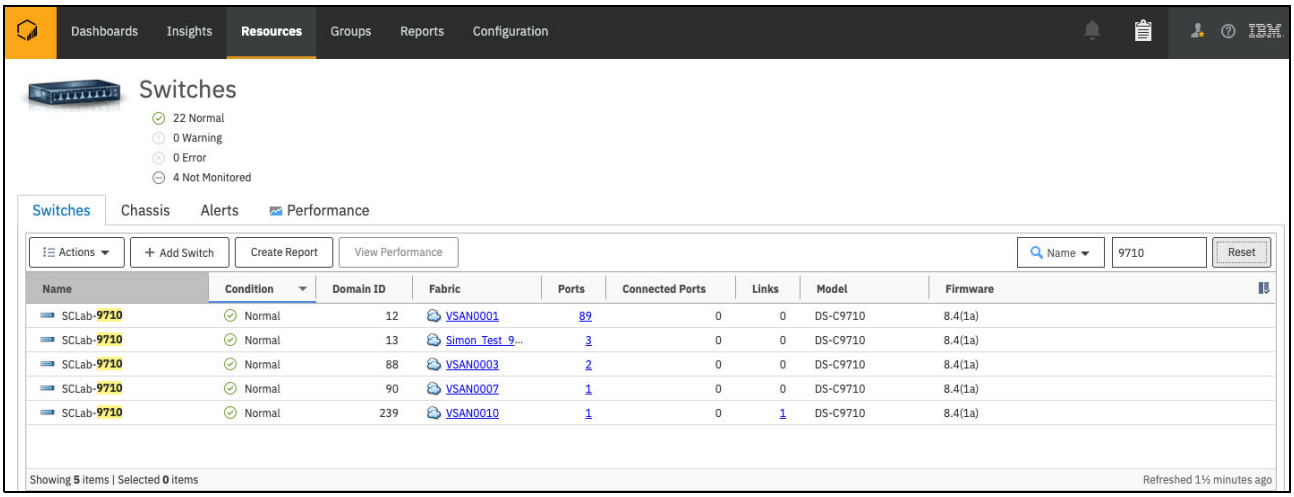


Figure 33 IBM Storage Insights physical switches view

Figure 34 shows an example of drilling down into a switch in IBM Storage Insights.

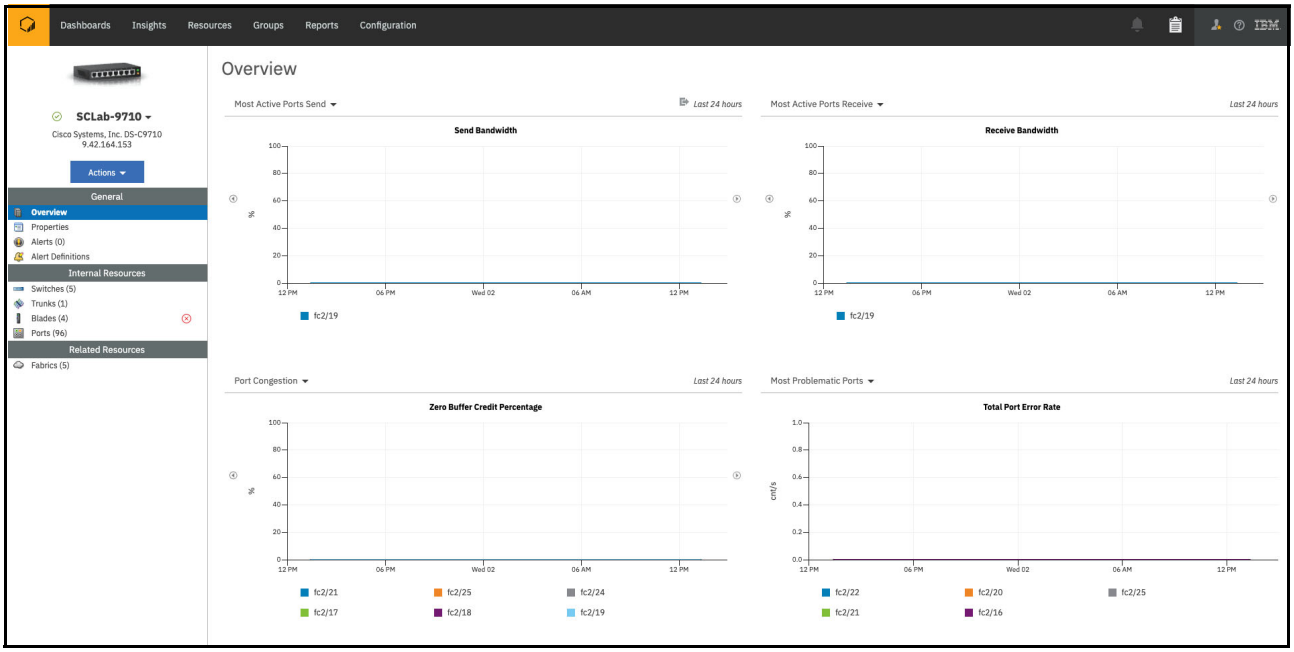


Figure 34 IBM Storage insights - detailed view

Figure 33 and Figure 34 show a preview of an expected feature. These examples that show fabric support might not be the same as in the final product.

Best practice: Use IBM Storage Insights to monitor your IBM storage, including switches.

FICON

With the release of FOS V8.1.0c, Broadcom and IBM qualified FICON multi-hop configurations. Initial multi-hop configurations are limited to three hops with specific configurations.

For more information, see the *FICON Multihop: Requirements and Configurations* white paper.

Enabling in-order delivery

FICON is sensitive to receiving commands and responses in-order and frames within a specific sequence. FICON requires In Order Delivery (IOD) be enabled to ensure that in-order delivery occurs during some recovery scenarios.

Example 35 shows the **`iodset`** and **`iodshow`** commands to enable frame IOD.

Example 35 The iodset and iodshow commands

```
DCX1_SANA:FID16:dlutz> iodset
IOD is set

DCX1_SANA:FID16:dlutz> iodshow
IOD is set
```

Best practice: Enable IOD on switches that support FICON.

Example 36 shows the **`iodreset`** and **`iodshow`** commands to disable frame IOD.

Example 36 The iodreset command

```
DCX1_SANA:FID16:dlutz> iodreset
IOD is not set

DCX1_SANA:FID16:dlutz> iodshow
IOD is not set
```

Access control and zoning

FICON switches can be set up in one of two ways:

- FICON switches can be set up to be like older ESCON directors, which did not use zoning, but used Allow/Prohibit matrixes that were created by the ESCON director console, or mainframe software, such as hardware configuration definition (HCD). To set up a switch in this manner, set the default access to `no access` and do not implement any zoning. You can then create Allow/Prohibit matrixes by using Network Advisor or mainframe software.
- FICON switches also can be set up by using zoning. Because FICON enforces host-to-device access through the Input Output Configuration DataSet (IOCDS), separate zones do *not* need to be created for each host to device connection. It is a common practice to place all FICON host and devices into a single zone.

As a preferred practice, Broadcom and IBM suggest one zone for all FICON connectivity.

Although domain index zoning is supported, WWN zoning for QoS is advised in environments where N_Port ID Virtualization (NPIV) is deployed.

Note: A mix of index zoning and WWN zoning is not recommended.

Disable the default zone, save it, and enable your new zoneset, as shown in Example 37.

Example 37 Enabling the new zoneset

```
Brocade_def:FID128:admin> defzone --no access
You are about to set the Default Zone access mode to No Access
Do you want to set the Default Zone access mode to No Access ? (yes, y, no, n):
[no] y
Brocade_def:FID128:admin> cfgsave
You are about to save the Defined zoning configuration. This
action will only save the changes on Defined configuration.
If the update includes changes to one or more traffic isolation
zones, you must issue the 'cfgenable' command for the changes
to take effect.
Do you want to save the Defined zoning configuration only?  (yes, y, no, n): [no]
y
Updating flash ...
Brocade_def:FID128:admin> cfgenable mynewzone
You are about to enable a new zoning configuration.
This action will replace the old zoning configuration with the
current configuration selected. If the update includes changes
to one or more traffic isolation zones, the update may result in
localized disruption to traffic on ports associated with
the traffic isolation zone changes.
Do you want to enable 'mynewzone ' configuration  (yes, y, no, n): [no] y
zone config " mynewzone " is in effect
Updating flash ...
Brocade_def:FID128:admin>
```

Configuring the FICON switch routing policy

Example 38 shows the suggested Device Based Routing (DBR) policy.

Example 38 Suggested DBR policy

```
Brocade_def:FID128:admin> aptpolicy
 Current Policy: 3
 3 : Default Policy
 1: Port Based Routing Policy
 2: Device Based Routing Policy (FICON support only)
 3: Exchange Based Routing Policy
Brocade_def:FID128:admin> switchdisable
Brocade_def:FID128:admin> aptpolicy 2
Policy updated successfully.
Brocade_def:FID128:admin> aptpolicy
 Current Policy: 2
 3 : Default Policy
 1: Port Based Routing Policy
 2: Device Based Routing Policy (FICON support only)
 3: Exchange Based Routing Policy
Brocade_def:FID128:admin> switchenable
```

Configuring SCC policies

To configure the SCC policies, complete the following steps:

1. Use **secPolicyCreate** to create and activate the policy, as shown in Example 39.

Example 39 Using secPolicy to create and activate policy

```
Brocade_def:FID128:admin > secpolicyshow
_____________________________________________________
                ACTIVE POLICY SET
_____________________________________________________
                DEFINED POLICY SET

Brocade_def:FID128:admin > secpolicycreate "SCC_POLICY",
"10:00:c4:f5:7c:96:0a:c0;10:00:c4:f5:7c:95:ab:e4"
SCC_POLICY has been created.

Brocade_def:FID128:admin > secpolicysave
secpolicysave command was completed successfully.

Brocade_def:FID128:admin > secpolicyshow
_____________________________
                ACTIVE POLICY SET
_____________________________________
                DEFINED POLICY SET
SCC_POLICY
   WWN                     DId swName
   ------------------------------------------------
   10:00:c4:f5:7c:95:ab:e4  57 PFE_up_def
   10:00:c4:f5:7c:96:0a:c0  49 PFE_lo_def
```

2. Activate the policy as shown in Example 40 (note that `swName` might show `Unknown` status).

Example 40 Activate the policy

```
Brocade_def:FID128:admin > secpolicyactivate
About to overwrite the current Active  Policy Set.
ARE YOU SURE  (yes, y, no, n): [no] y
secpolicyactivate command was completed successfully.
Brocade_def:FID128:admin > secpolicyshow

_____________________________________________________
                ACTIVE POLICY SET
SCC_POLICY
   WWW                     DId swName
   -------------------------------------------------
   10:00:c4:f5:7c:95:ab:e4  57 PFE_up_def
   10:00:c4:f5:7c:96:0a:c0  49 PFE_lo_def

_____________________________________________________
                DEFINED POLICY SET
SCC_POLICY
   WWW                     DId swName
   -------------------------------------------------
   10:00:c4:f5:7c:95:ab:e4  57 PFE_up_def
   10:00:c4:f5:7c:96:0a:c0  49 PFE_lo_def
```

3. Enter the **`fddCfg`** command to enable the ACL fabric-wide consistency policy and enforce a strict SCC policy, as shown in Example 41.

Example 41 The fddCfg command to enable ACL fabric wide consistency policy

```
Brocade_def:FID128:admin > fddcfg --fabwideset "SCC:S"
Brocade_def:FID128:admin >
```

4. To create a policy that includes all the switches in the fabric, enter the command that is shown in Example 42.

Example 42 Create a policy to include all the switches in the fabric

```
Brocade_def:FID128:admin > secPolicyCreate SCC_POLICY "*"
SCC_POLICY has been created.
```

5. Configure the switch parameters by using the **`configure`** command, as shown in Example 43. The configuration change is saved only if all parameters are acknowledged (only the values to change are displayed here). The switch must be disabled *before* the **`configure`** command is run.

Example 43 Switch configure command to setup SCC policies for FICON

```
Brocade_def:FID128:admin > switchdisable
```

Choose a unique domain ID (20 in this example):

```
Brocade_def:FID128:admin> configure
Configure...
  Fabric parameters (yes, y, no, n): [no] y
    Domain: (1..239) [4] 20
```

Disable device probing:

```
Disable Device Probing: (0..1) [0] 1
```

Leave E_D_TOV value at 2 seconds (2000) unless connected to extension equipment. In some cases, when connecting to extension equipment, it must be set to 5 seconds (5000):

```
E_D_TOV: (1000..5000) [2000]
```

The preferred practice is to set the domain ID to be insistent. Setting the insistent domain ID is required for two-byte addressing.

```
Insistent Domain ID Mode (yes, y, no, n): [no] y
```

Note: Before setting HIF mode, the following attributes must be configured:

- An insistent domain ID (IDID)
- A fabric-wide consistency policy → SCC:S (Strict mode)
- A valid SCC_Policy (configured and activated)

If one of these parameters is not configured, the following error message is produced before HIF is enabled:

```
Error: Unable to set HIF Mode. No valid SCC policy or Fabric wide(SCC:S)
configuration
```

Example 44 shows the configuration process.

Example 44 Setting the insistent domain ID

```
Brocade_def:FID128:admin> configure
Configure...
  Fabric parameters (yes, y, no, n): [no] y
    High Integrity Fabric Mode (yes, y, no, n): [no] y
```

Enable High-Integrity Fabric mode:

```
Brocade_def:FID128:admin > configure

Configure...
Fabric parameters (yes, y, no, n): [no] y
High Integrity Fabric Mode (yes, y, no, n): [no] y
HIF mode is enabled on the switch
Brocade_def:FID128:admin > switchenable
```

Port-based routing (PBR) and device-based routing (DBR) are qualified for IBM z Systems®. However, see your system qualification letter for current support information (log in required).

Access gateway

BladeCenter and chassis-style systems typically feature embedded switches. These switches can operate in native fabric mode or access gateway (AG) mode. AG mode uses NPIV to connect the devices in the chassis to the network instead of native fabric mode, which operates as a standard switch, and requires its own fabric domain and a copy of the name server and configuration databases. In AG mode, the embedded switch does not use any of these items.

Embedded switches can support trunking, which often requires an optional license. Trunking allows transparent failover and failback within the trunk group. Trunked links are more efficient and can distribute I/O more evenly across all the links in the trunk group.

Best practice: Consider the following best practices:

- Run Access Gateways in AG mode.
- For a chassis with high throughout or high availability goals, use F_port trucking.

For more information, see *Brocade Access Gateway Administrator's Guide*.

Frame Viewer

Frames that are discarded are sent to the CPU for processing. During subsequent CPU processing, information about the frame, such as SID, DID, and transmit port number, is retrieved and logged. This information is maintained for a specific fixed number of frames.

Frame Viewer captures FC frames that are dropped. Depending on the hardware platform and FOS version, captured frames can be for different types, such as timeouts, unroutable, or unreachable destinations that are received on an Edge ASIC (an ASIC with front-end (FE) ports). If the frame is dropped on a Core ASIC, the frame is not captured by Frame Viewer.

By default, only the timeout frame type is enabled. If you are seeing high frame discards for other reasons, you might want to enable Frame Viewer to log other discard types to capture the source and destination address to help isolate the issue.

Example 45 shows the **framelog** command to show the frame viewer status.

Example 45 The framelog --status command output

```
F48a_Default:FID128:dlutz> framelog --status
Service Status:                 Enabled
Enabled Disc Frame Types: timeout du type1miss unroute type2miss type6miss
```

Example 46 shows the framelog command to enable logging of different frame types on a GEN 5 switch running FOS V8.1.

Example 46 The framelog --enable command on a GEN5 switch

```
F48a_Default:FID128:dlutz> framelog --enable -type timeout
Nothing to do: service is already enabled.
F48a_Default:FID128:dlutz> framelog --enable -type du
F48a_Default:FID128:dlutz> framelog --enable -type unroute
F48a_Default:FID128:dlutz> framelog --enable -type type1miss
F48a_Default:FID128:dlutz> framelog --enable -type type2miss
F48a_Default:FID128:dlutz> framelog --enable -type type6miss
F48a_Default:FID128:dlutz> framelog --enable -type all
Nothing to do: service is already enabled.
```

Example 47 shows the framelog command to enable logging of different frame types on a GEN 4 switch running FOS V7.4.

Example 47 The framelog --enable command on a GEN 4 switch

```
SANA_DCX2:FID16:dlutz> framelog --enable -type du
Error: Feature requested is not supported
SANA_DCX2:FID16:dlutz> framelog --enable -type unroute
Error: Feature requested is not supported
```

Figure 48 shows the **framelog** command to show C3 frame Tx timeouts.

Example 48 The framelog command output

```
framelog --show -n 1200
==========================================================================
        Sun Oct 16 23:49:07 EDT 2016
==========================================================================
Log              TX    RX
timestamp        port  port SID       DID      SFID DFID Type Count
==========================================================================
Sep 18 05:56:05  2/22 -1/-1 0x14a040 0x0a4e00 128 128 timeout 20
Sep 18 05:56:04  9/46 -1/-1 0x14a040 0x0a4e00 128 128 timeout 20
Sep 18 05:39:08  -1/-1 3/16 0x14a040 0x0a4e00 128 128 timeout 2
Sep 18 05:39:08  9/46 -1/-1 0x14a040 0x0a4e00 128 128 timeout 20
Sep 04 05:21:56 9/46 -1/-1 0x14a040 0x0a4e00 128 128 timeout 20
Sep 04 05:21:55 -1/-1  1/1 0x140100 0x0ae940 128 128 timeout 8
Aug 03 12:00:44 2/22 -1/-1 0x14a940 0x0a1d03 128 128 timeout 9
Aug 03 12:00:44 2/22 -1/-1 0x149a40 0x0a3900 128 128 timeout 11
Aug 02 04:47:23 -1/-1 4/2 0x143200 0x0a7100 128 128 timeout 8
Aug 02 04:47:23 2/22 -1/-1 0x14acc0 0x0a2d00 128 128 timeout 2
Aug 02 04:47:23 2/22 -1/-1 0x14a940 0x0a1d03 128 128 timeout 17
Aug 02 04:47:23 2/22 -1/-1 0x142000 0x0ad1c0 128 128 timeout 1
```

Forward Error Correction

Brocade Gen5 (16 Gbps) platforms support Forward Error Correction (FEC) that automatically corrects bit errors. This function enhances the link reliability, and improves resiliency with the presence of marginal media. FEC is preferred between all supported devices. FEC is mandatory on Gen6 link (32 Gbps) speed.

FEC on Gen5 can correct up to 11-bit errors in every 2112-bit transmission in a 10 Gbps/16 Gbps data stream in frames and primitives. FEC is enabled by default on the back-end (BE) links of Condor 3 ASIC-based switches and blades, and minimizes the loss of credits on BE links.

FEC also is enabled by default on FE links when connected to another FEC-capable device. FEC on Gen6 uses a more robust coding algorithm that corrects up to seven 10-bit streams and detects up to 14 10-bit streams, without the requirement that the errors be in a burst.

FEC is mandatory on Gen6 platforms for 32 Gbps speed to ensure that the bit-error rate stays within the standard requirement. Condor 4 ASIC automatically turns on FEC when a port operates at 32 Gbps speed, and cannot be disabled.

Enable FEC on 10 Gbps/16 Gbps connections when both ends of the link support it.

Summary of best practices

Table 15 is a summary of the preferred features and capabilities to improve the overall resiliency of FOS-based FC fabric environments:

Table 15 Summary of best practices

Section	Practice	Reference
Designing	Avoid introducing devices to the SAN that span more than one generation of technology.	“ISLs and multi-hop ISLs” on page 8
Designing	Avoid traversing ISLs when accessing SSD/Flash for high performance use cases.	“ISLs and multi-hop ISLs” on page 8
Designing	Use virtual switches to provide logical isolation between different types of workloads.	“Virtual fabrics” on page 9
Designing	Avoid using devices with different connection speeds.	“Flow management” on page 9
Designing	Manage the number of subscribing to device ports.	“Flow management” on page 9
Designing	Use default exchange based routing.	“Routing policies” on page 10
Designing	Enable the credit tools with the onLrOnly recovery option.	“Credit recovery tools” on page 12
Designing	Use multiple trunked ISLs between switches in a fabric.	“Inter-switch link trunking” on page 14
Designing	Use Peer Zoning for large fabrics.	“Peer zoning” on page 15
Designing	Use aliases and make all of the names meaningful.	“Using a meaningful naming convention” on page 17
Designing	Enable Dynamic Port Naming.	“Dynamic port naming” on page 18
Maintaining	Jitter should not vary by more than 10 to 15%.	“Jitter” on page 34
Maintaining	Retransmits should by 0.05% or less and 0.01% is even better.	“Retransmits” on page 35
Maintaining	Out of order packets should be 0.05% or less.	“Out-of-order packets” on page 35
Designing	When multiple circuits for a FCIP Tunnel are used, be sure the RTT of each circuit is within 5% or less of each other.	“Using multiple circuits” on page 35
Designing	Do not use FCIP write acceleration with IBM disk storage devices. The use of FCIP Tape Acceleration for tape devices can be deployed as long as the path with FCIP Tape Acceleration does not have any disk traffic.	“Accelerators” on page 39
Monitoring	Use IBM Storage Insights to monitor your storage, including switches.	“IBM Storage Insights” on page 68
Monitoring	Enable FPI or Bottleneck monitoring.	
Monitoring	Enable MAPS monitoring and alerting.	“Enabling Monitoring Alerting Policy Suite” on page 50
Monitoring	Configure and use Network Advisor Dashboards.	“Configuring Network Advisor dashboards” on page 60
Monitoring	Create custom SANnav dashboards.	“SANnav dashboards” on page 64

Section	Practice	Reference
FICON	Enable IOD on switches supporting FICON.	“Enabling in-order delivery” on page 70
Access Gateway	Run Access Gateways in AG mode.	“Access gateway” on page 74
Access Gateway	For Chassis with high throughout or high availability goals use F_port trucking.	“Access gateway” on page 74

Authors

This paper was written by a team of specialists from around the world. The content is based on Broadcom documentation and is presented in a form that specifically identifies IBM preferred practices:

Jim Blue
IBM Systems

David Green
IBM Systems

David Lutz
IBM Systems

Ian Mac Quarrie
IBM Systems

Gavin O'Reilly
IBM GTS

Thanks to the following Bert Dufrasne, IBM Redbooks, IBM Systems, for his contributions to this project.

Special thanks to Broadcom® for their support of this paper, and the following people in particular:

Brian Larsen
Tim Werts

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 IBM Redbooks 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 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 US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

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, MD-NC119, Armonk, NY 10504-1785, USA

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 jurisdictions 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 provide in any way it believes appropriate without incurring any obligation to you.

The performance data and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions.

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.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

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 actual people or business enterprises 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. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation, registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at https://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks or registered trademarks of International Business Machines Corporation, and might also be trademarks or registered trademarks in other countries.

FICON®	IBM z13®	z Systems®
IBM®	Redbooks®	z/OS®
IBM z Systems®	Redbooks (logo) ®	z13®

The following terms are trademarks of other companies:

Broadcom®

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

REDP-4722-05

ISBN 0738460257

Printed in U.S.A.