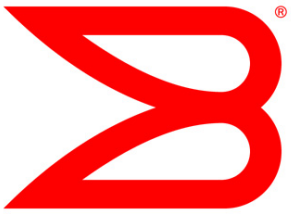


DATA CENTER

SAN Fabric Administration Best Practices Guide

Support Perspective

A high-level guide focusing on the tools needed to proactively configure, monitor, and manage the Brocade Fibre Channel Storage Area Network infrastructure.

BROCADE

CONTENTS

INTRODUCTION

For over 15 years Brocade has been developing, installing, and training customers on Fibre Channel (FC) Storage Area Networks (SANs) and, over time, has developed deep technical knowledge in administering SANs. This document is intended to be a high-level document based on Brocade experience, products, and features focusing on SAN fabric administration best practices guidelines for addressing configuration, monitoring, managing, and diagnosing the Brocade-based SAN infrastructure.

The guidelines in this document will not apply to every environment, but they will help guide you through the tools you need for successful administration of SAN fabrics. Please consult your Brocade sales representative or Brocade SE for details about the hardware and software products and features described in this document.

Note: *This is a "living" document that is continuously being expanded, so be sure to frequently check MyBrocade (my.brocade.com) for the latest update of this and other best practice documents. Future release of this document will cover additional topics such as best practices for routed fabrics, Dense Wavelength-Division Multiplexing (DWDM) connections, and access gateways. Refer to documents in the reference section for further details on the features and tools discussed in this guide. Refer to the SAN Design and Best Practices Guide for optimal design principles.*

AUDIENCE AND SCOPE

This document is intended for Storage Area Network-Fabric Administrators (including storage and network), Brocade certified Systems Engineers, IT architects, and System Integrators that provide value-added management solutions based on the latest product releases from Brocade.

The scope of this document is to address common issues faced by administrators in managing their SANs. The goal is to reduce the time needed for troubleshooting and dealing with application anomalies by using available tools to minimize fabricwide disruptions. The details outlined in this document are for 8 Gbps and 16 Gbps devices only.

Note: The features and functions covered in this document apply only to Brocade® Fabric OS®-based products. This document is not a replacement for product-specific manuals or detailed training on Brocade Fabric OS (FOS) or Brocade Network Advisor.

BROCADE TOOL SET

Brocade has built in an extensive set of SAN administration, usability, and RAS (Reliability, Accessibility, and Serviceability) features into the product line, including ASICs, Brocade FOS, Brocade Network Advisor, Brocade SAN Health®, and the Brocade SAN Health Professional management tool.

Brocade FOS: Brocade FOS has evolved through six generations of Fibre Channel speed transitions to provide a highly resilient platform for building next-generation Storage Area Network products. The operating system has evolved to provide two options for deployment. For very risk adverse customers running mission-critical applications, where stability and uptime are critical, upgrading within the minor release train with RAS improvements is the best option. Customers who want to take advantage of Brocade innovations in new products and features can continue to leverage the latest Brocade FOS release.

These are the key RAS features for the Target Path release:

- Credit recovery
- Bottleneck detection
- Port fencing

These are the key SAN resiliency features on the latest major Brocade FOS 7.0.x release:

- Credit loss detection and automatic recovery (Inter-Switch Link [ISL] and backend ports), including stuck Virtual Channels (VC)

- C3 discard frame logging and viewing
- Forward Error Correction on all 16 gigabit (Gbit) and 10 Gbit links
- In-flight encryption and compression
- D_Port support
- Advanced SFP monitoring (thresholds based on SFP type)
- Using E_Port top-talkers on 16 Gbps ISLs
- Access Gateway N_Port monitoring
- Duplicate World Wide Name (WWN) detection and resolution

Refer to the Brocade Fabric OS v7.0.x Release Notes, Brocade Fabric OS Administrator's Guide, and Brocade Fabric OS Command Reference Guide supporting Brocade Fabric OS v7.0.x for details on these new features.

Brocade Network Advisor and DCFM offer comprehensive monitoring and management support across multiple Brocade SAN, IP, and converged network fabrics. These applications equip administrators with configuration, zoning, visualization, analysis, and troubleshooting tools. Only Brocade Network Advisor is supported for management of switches operating with Brocade FOS v7.0 and later firmware versions.

Brocade SAN Health provides an accurate view of the SAN environment with fabric topologies and detailed performance metrics. Brocade SAN Health audit reports provide detailed color-coded hierarchical SAN insights from Brocade FOS, Brocade M-EOS, the Brocade Mi10K Director, and Cisco MDS switches. The tool supports discovery and reporting of both open systems and FICON fabrics.

EVOLUTION OF THE ENTERPRISE DATA CENTER

Fibre Channel-based SANs have evolved over the past 10 years, from SAN islands to a highly consolidated and complex infrastructure driven by server virtualization and high capacity storage arrays. Diverse workloads and traffic profiles going through the core network present a challenge in addressing intermittent anomalies in the fabric.

Fabric usage has also changed. There are more high-availability clusters, such as IBM HACMP, VMware, and Microsoft Windows. Workload has also become much more complex. Instead of simple host target port pairs, you now see hypervisors such as VMware VSphere, Windows Hyper-V, and IBM VIOS servicing large numbers of virtualized hosts. This makes it much more difficult to isolate application problems when application performance becomes a problem.

Storage virtualization has created its own special I/O requirements, adding a degree of complexity to the I/O complex previously unseen outside of very complex mainframe environments.

All this has a serious impact on storage—particularly fabric—problem determination. There are more entities to manage such as Logical Unit Numbers (LUNs), hosts, storage, and virtual machines (VMs), and more potential problems. Also, the operational environment is much more difficult to troubleshoot than it was even a few years ago. Rogue or badly behaving devices have much more impact on production environment than they did previously, and management tools have not kept up with the changes.

There is an increase in virtualized hosts running in hypervisor clusters accessing virtualized storage, which could potentially put a strain on the storage infrastructure, especially when there are a high number of virtual hosts per physical server and all accessing the same storage infrastructure.

Many of the new behaviors induced by innovations in workload and storage infrastructures have generated a corresponding difference in fabric traffic patterns and fabric manageability. For example, there is a significant increase in very short frames, such as those encapsulating SCSI reserves and in-band Fibre Channel control frames used by workload and storage virtualization products.

N_Port ID Virtualization (NPIV) hides flow information that was previously reported on individual ports.

The result of all this change is the appearance of increasing issues with application performance that seem to be associated with storage performance in some way but that cannot be sufficiently identified so that corrective action can be taken.

SAN ADMINISTRATOR DILEMMA

When application performance problems become obvious, the SAN complex is frequently blamed. SAN administrators usually have no metrics that might point to some other component in the infrastructure. Frequently, the result is very long delays before the culprit or culprits are identified and measures are taken to address the problem. The impact of such outages range from an inconvenience to a massive outage, where mission-critical application availability is compromised and the enterprise is seriously affected. The experience is never a positive one.

Brocade recognized the need for improved monitoring and problem determination aids and started a series of initiatives to address the problem of relevant performance and problem determination metrics in the fabric. Bottleneck detection is one of the first deliverables of this work. Bottleneck detection is designed to positively identify bottlenecks in the fabric.

Two types of bottlenecks are detected:

- Bandwidth-based bottlenecks are determined by high link utilization. These are called congestion bottlenecks. Congestion bottlenecks are relatively easy to detect and, in effect, can be detected by other Brocade products such as Brocade Fabric Watch. Bottleneck detection provides an alternative mechanism and more information about the congestion.
- Device latency-based bottlenecks, called latency bottlenecks are much more difficult to detect. This is the primary focus of bottleneck detection, and the focus of much of the remainder of this section.

Latency detection is frame-based and identifies buffer credit problems. One of the major strength of Fibre Channel is that it creates lossless connections by implementing a flow control scheme based on buffer credits. The disadvantage of such an approach is that the number of available buffers is limited and may eventually be totally consumed.

The temporary unavailability of buffer credits creates a temporary bottleneck. The longer the credits are unavailable, the more serious the bottleneck. Whereas temporary credit unavailability is expected in normal Fibre Channel operation, the longer durations are of most concern.

Long periods without buffer credits are typically manifested as performance problems and are usually the result of device latencies. Exceptional situations cause fabric back pressure that can extend all the way across the fabric and back. Excessive back pressure can create serious problems in an operational SAN.

Chronic back pressure can exacerbate the effect of hardware failures and misbehaving devices and can also contribute to serious operational issues, as the existence of existing bottlenecks increases the probability of a failure.

There are several common sources of high latencies:

- Storage ports (targets) often produce latencies that can slow down applications, because they do not deliver data at the rate expected by the host platform. Even well-architected storage array performance can deteriorate over time. For example, LUN provisioning policies such as allocating too many LUNs behind a given port can contribute to poor performance of the storage, if the control processor in the array cannot deliver data from all the LUNs quickly enough to satisfy read requests. The overhead of dealing with a very large number of LUNs may cause slow delivery.
- Hosts (initiators) may also produce significant latencies by requesting more data than they are capable of processing in a timely manner.

- Distance links can frequently consume all the buffer credits reserved for them and create a serious bottleneck in the middle of a fabric, which can have serious consequences for any applications sharing that link.
- Misbehaving devices such as defective Host Bus Adapters (HBAs) can create havoc in a well-constructed SAN and increase the threat to the fabric.

Eliminating bottlenecks contributes to the overall stability of a fabric by reducing the effects of other events in the SAN. Back pressure in the fabric, produced by latencies and congestion, exacerbates the effects of other events in a SAN and reduces the ability of the fabric to deal with problems such as misbehaving devices. At best, application performance is impacted. In extreme cases, SAN outages can occur.

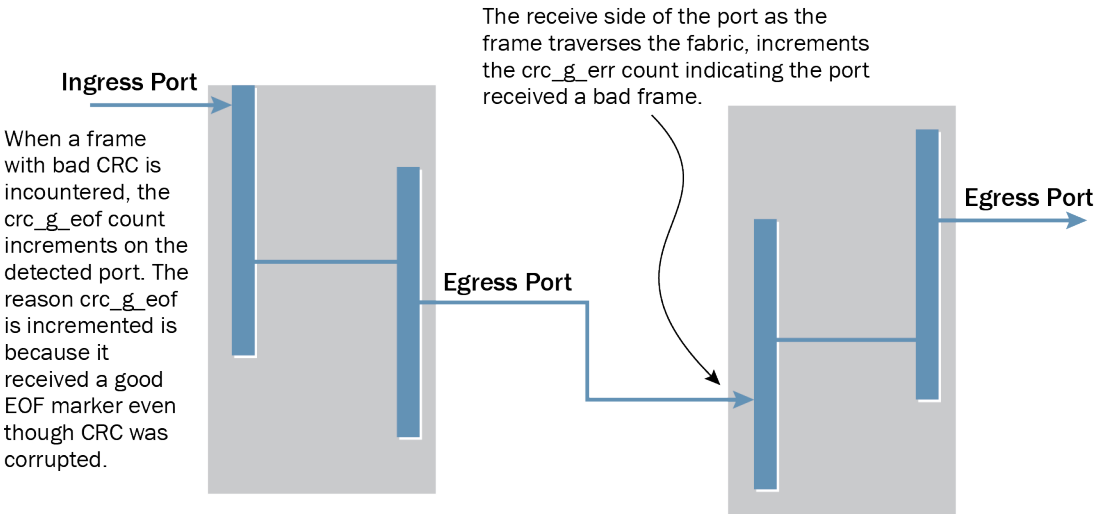
Another reason for connectivity issues could be a marginal link. A marginal link involves the connection between switches or between the switch and the device. Isolating the exact cause of a marginal link involves analyzing and testing many of the components that make up the link (including the switch port, switch SFP, cable, edge device, and edge device SFP). Brocade Fabric OS provides various port statistics and error counters (using the non-disruptive CLI command **portErrShow**) to help troubleshoot a marginal link. Brocade enhanced the toolkit with integrated diagnostic-port capability for 16 Gbps platforms running Brocade FOS v7.0.0 or later (requires 10 Gbps or 16 Gbps Brocade branded SFPs). The diagnostic port allows the administrator to diagnose link-level faults.

These are some of the most important counters:

er_enc_in: This counter indicates where the encoding was corrupted. Encoding errors can occur inside or outside of the frame. These errors impact ordered sets. If encoding errors occur inside the frame, this counter increments and is not logged as a "Class3" discard, since the frame is unreadable. When encoding errors occur outside of the frame, the enc_out counter will increment.

er_enc_out: This is the same as er_enc_in, except that encoding errors occur outside the frames.

er_crc: Cyclic Redundancy Check (CRC) errors indicate frame corruption in the associated frame. There are two types of CRC errors that can be logged on a B-Series switch, and together they can assist in determining where the error was introduced into the fabric. There is the CRC with good end of frame (EOF) (crc g_eof) and a plain CRC (with bad EOF). When a frame with a CRC error is first detected with a complete frame, a CRC with good EOF is logged. Once the CRC is detected, the good EOF (EOFn) is replaced with a bad EOF (EOFni). When a CRC with good EOF is detected at the port, it indicts the transmitter or path from the sending side as a possible culprit.



er_bad_eof: For the incoming frame, a found SOF (start of frame), but no known FC EOF is found. In other words, the EOF is damaged beyond recognition.

Class 3 Discards: There are multiple class 3 (C3) discards, but the primary ones of concern are those due to timeout conditions (er_c3_timeout). There are two types of C3 timeout discards that are logged, receive (RX) and transmit (TX). Both of these C3 timeout discards function in a similar manner. If an R_RDY (buffer-to-buffer credit) or VC_RDY has an encoding error, a credit is lost, which may impact performance on that link.

Loss of Sync: The number of times a synchronization error occurs on the port. This means that two devices failed to communicate at the same speed. Synchronization errors are always accompanied by a link failure. Loss of synchronization errors frequently occur due to a faulty SFP or cable.

Link Loss: The number of times a link failure occurs on a port or sends or receives Not Operational (NOS). Both physical and hardware problems can cause link failures. Link failures also frequently occur due to a loss of synchronization or a loss of signal.

Loss of Sig: The number of times that a signal loss occurs in a port. Signal loss indicates that no data is moving through the port. A loss of signal usually indicates a hardware problem.

FABRIC CONFIGURATION

Fabrics can be architected to mitigate some impacts of device latency. Isolating the device flows (host/storage pair) that exhibit high latencies—either by putting them in their own fabric or on their own blade/switch—will contain the impact of the latencies to the fabric or blade/switch containing the high-latency device flows. Features such as integrated routing (Fibre Channel Routing) and local switching provide architectural-level solutions that limit the need for more complex monitoring and mitigation capabilities. However, using fabric design as a protection mechanism does require some knowledge of which devices are likely to exhibit latency.

Fabricwide parameters

The CLI command **configShow** can be used to list fabricwide configuration parameters on each switch. There is no need to change any of these parameters unless directed by a Brocade Support Representative, as misconfiguration can lead to fabric instability and major fabric disruptions.

Fill Word (Condor 2 8 Gbps platform only)

Prior to the introduction of 8 Gb, IDLEs were used for link initialization, as well as fill words after link initialization. To help reduce electrical noise in copper-based equipment, the use of ARB (FF) instead of IDLEs was standardized. Because this aspect of the standard was published after some vendors had already begun development of 8 Gb interfaces, not all equipment can support ARB (FF). IDLEs are still used with 1, 2, and 4 Gb interfaces. To accommodate the new specifications and different vendor implementations, Brocade developed a user-selectable method to set the fill words to either IDLEs or ARB (FF). Currently, setting the fill word can be done only via the CLI command **portCfgFillWord** (Ex: **portcfgfillword** [*slot/*]*port*, *mode)*. There are four modes:

MODE	MEANING
Mode 0	Use IDLEs in link initialization and IDLEs as fill word (default mode).
Mode 1	Use ARB (FF) in link initialization and ARB (FF) as fill words.
Mode 2	Use IDLEs in link initialization and ARB (FF) as fill words.
Mode 3	Try Mode 1 first; if it fails, then try Mode 2.

Traffic outside of frame traffic is made up of fill words: IDLEs or ARB (F0) or ARB (FF). Encoding errors on fill words are generally not considered impactful. This is why you may see very high counts of enc_out (encoding outside of the frame) and not have customer traffic affected. If many fill words are lost at once, the link may lose synchronization. On standard E_Ports, primitives are set to ARB, regardless of the portcfgfillword setting when not in R_RDY mode.

The recommended best practices are:

- Ensure that the fill word is configured to Mode 3.
- When connecting to a HDS storage device, set to Mode 2.
- When upgrading firmware, recheck the settings, since the fill word primitive has evolved over several Brocade FOS releases.

Bottleneck Detection

A bottleneck is a port in the fabric where frames cannot get through as fast as they should. In other words, a bottleneck is a port where the offered load is greater than the achieved egress throughput. Bottlenecks can cause undesirable degradation in throughput on various links. When a bottleneck occurs at one place, other points in the fabric can experience bottlenecks as the traffic backs up.

Bottleneck detection prevents degradation of throughput in the fabric and reduces the time it takes to troubleshoot network problems.

The bottleneck detection feature from Brocade detects two types of bottlenecks (as discussed in a previous section):

- Latency bottlenecks
- Congestion bottlenecks

A latency bottleneck is a port where the offered load exceeds the rate at which the other end of the link can continuously accept traffic, but it does not exceed the physical capacity of the link. This condition can be caused by a device attached to the fabric that is slow to process received frames and send back credit returns. A latency bottleneck due to such a device can spread through the fabric and slow down unrelated flows that share links with the slow flow.

By default, bottleneck detection detects latency bottlenecks that are severe enough that they cause 98 percent loss of throughput. This default value can be modified to a different percentage. A congestion bottleneck is a port that is unable to transmit frames at the offered rate, because the offered rate is greater than the physical data rate of the line. For example, this condition can be caused by trying to transfer data at 8 Gbps over a 4 Gbps ISL. Use the **bottleneckmon** CLI command to configure bottleneck monitoring and configure alert thresholds for congestion and latency bottlenecks.

Advanced settings allow you to refine the criteria for defining latency bottleneck conditions to allow for more (or less) sensitive monitoring at the subsecond level. For example, you use the advanced settings to change the default value of 98 percent for loss of throughput. If a bottleneck is reported, you can investigate and optimize the resource allocation for the fabric. Using the zone setup and Top Talkers, you can also determine which flows are destined to any affected F_Ports.

Bottleneck detection was introduced in Brocade FOS 6.3.0 with monitoring for device latency conditions. It was then enhanced in Brocade FOS 6.4.0 with added support for congestion detection on both E_Ports and F_Ports. Brocade FOS 6.4 also added improved reporting options and simplified configuration capabilities. The Brocade FOS 6.3.1b release (and later) included enhancement in the algorithm for detecting device latency, making it more accurate. Bottleneck detection does not require a license and is supported on 4, 8, and 16 Gbps platforms.

Edge Hold Time

Edge hold time configuration is a new capability added in the Brocade FOS 6.3.1b release. There is no license required to configure the edge hold time setting.

Edge hold time is the maximum time a frame can wait after it is received on the ingress port and before it is delivered to the egress port. If the frame waits in the egress buffer for more than the configured hold time, the switch drops the frame, replenishes the sender's credit, and increments the counters sts_tx_timeout and er_c3_timeout on the TX and RX ports, respectively. The frame-timeout indicates a slow draining or a congestion or bottleneck in the fabric. Decreasing hold time on the edge switches may reduce frame drop counts in the core switches.

Choose one of the following options for configuring the edge hold time:

- 0: Low edge hold time of 80 milliseconds
- 1: Medium edge hold time of 220 milliseconds
- 2: Long edge hold time of 500 milliseconds. This is the default value.

Edge hold time can be configured non-disruptively using the **configure** command.

Debug Log Level Settings

Debug level settings can be changed on more than 150 modules for in-depth troubleshooting and diagnostics. Customers should leave the setting at factory default unless advised by Brocade support personnel. The data collected via the SupportSave process is sufficient for initial diagnosis.

For environments where an audit trail is necessary, login failures, zone configuration changes, firmware downloads, and other configuration changes—in other words, critical changes that have a serious effect on the operation and security of the switch—can be sent to the syslog server.

Auditable events are generated by the switch and streamed to an external host through a configured system message log daemon (syslog), and only the last 256 events are persistently stored on the switch. Audited events generated are specific to the particular switch and have no negative impact on its performance. In case the audit log is too verbose, and too many events are generated by the switch, the remote host's system message log may become a bottleneck, and audit events can be dropped by the switch.

Audit logging can be configured/displayed using the **auditcfg** command. This command allows you to set filters by configuring certain classes, to add or remove any of the classes in the filter list, to set severity levels for audit messages, and to enable or disable audit filters. Based on the configuration, certain classes are logged to syslog for auditing. Syslog configuration is required for logging audit messages. Use the **syslogdIpAdd** command to add the syslogd server IP address.

Brocade Fabric Watch

Brocade Fabric Watch is an optional (licensed) SAN health monitor that enables each switch to constantly monitor its SAN fabric for potential faults and automatically alerts for potential problems long before they become costly failures. This feature was enhanced in Brocade FOS 6.1.0 with the addition of port fencing. Port fencing allows a switch to monitor specific behaviors on the port and protect a switch by fencing the port when specified thresholds are exceeded. Fabric watch notifies the user of the action taken through one or more of the following mechanisms:

- Send an SNMP trap
- Log a RASlog message
- Send an e-mail alert
- Log a SYSlog message

Fabric Watch monitoring: Fabric Watch supports monitoring of different aspects of the system:

- The Fabric class groups areas of potential problems arising between devices, such as zone changes, fabric segmentation, E_Port down, fabric reconfiguration, domain ID changes, and fabric logins. A fabric-class alarm alerts you to problems or potential problems with interconnectivity.
- Performance monitoring groups areas that track the source and destination of traffic. Use the Performance monitor class thresholds and alarms to determine traffic load and flow and to reallocate

resources appropriately. The performance monitor class includes end-to-end monitors, frame monitors, and Top Talker monitors.

- The Security class monitors different security violations on the switch and takes action based on the configured thresholds and their actions.
- Port monitoring monitors port statistics and takes action based on the configured thresholds and actions. You can configure thresholds per port type and apply the configuration to all ports of the specified type, using the **portThConfig** command. For example, Fabric Watch can monitor CRC errors (available in Brocade FOS 6.1.x), invalid words (available in Brocade FOS 6.1.x), and state changes (ports transitioning between offline and online, available in Brocade FOS 6.3). It is a recommended best practice to use Fabric Watch to detect frame timeouts, that is, frames that have been dropped because of severe latency conditions (the Fabric Watch "C3TX_TO" area available in Brocade FOS version 6.3 for 8 Gbps ports and available in Brocade FOS 6.3.1b/6.4.0 and later for 4 Gbps ports).
- The SFP class groups areas that monitor the physical aspects of an SFP, such as voltage, current, RXP, TXP, and state changes in physical ports, E_Ports, FOP_Ports, and FCU_Ports. An SFP class alarm alerts for an SFP media fault. The most common cause of credit loss is corruption to credit return messages (VC_RDY or R_RDY) due to faulty media. Credit corruption is tracked by an encoder out error, which is an invalid word error. Monitoring and mitigating invalid word issues protects against credit loss.
- System resource monitoring enables monitoring systemwide components, including temperature, power supplies, FANs, system RAM, flash, memory, CPU, and so forth.

Fabric Watch quarantine: Fabric Watch also provides a mechanism that quarantines the badly behaving component with the optional action of port fencing. Port fencing is available for each of the previously noted conditions and is recommended to automatically protect the fabric from these error conditions. The recommended thresholds are specified in "Appendix A: Configuring Port Fencing." Refer to the *Fabric Watch Administrator's Guide* and *Fabric Resiliency Best Practices Guide* for the recommended thresholds that have been tested and tuned to quarantine components that are misbehaving to the point at which they are likely to cause a fabric-wide impact. The thresholds do not falsely trigger on normally behaving components.

Zoning

Zoning is a fabric-based service that enables you to partition your SAN into logical groups of devices that can access each other. Zones provide controlled access to fabric segments and establish barriers between operating environments. A device in a zone can communicate only with other devices connected to the fabric within the same zone. A device not included in the zone is not available to members of that zone. When zoning is enabled, devices that are not included in *any* zone configuration are inaccessible to all other devices in the fabric.

There are two types of zoning: WWN zoning and port zoning. Registered State Change Notification (RSCN) messages are limited to the zone in which they occurred.

WWN zoning: WWN zoning permits connectivity between attached nodes based on WWN. The attached node can be moved anywhere in the fabric and remains in the same zone. WWN zoning is used in open systems environments and does not make sense for FICON channels and FICON control units.

Port zoning: Port zoning limits port connectivity based on port number—in other words, all devices connected to all the ports that are members of the zone can talk to each other. Port zoning is used in FICON configuration. Ports can easily be added to port zones, even if there is nothing attached to the port.

Mixed zoning: A zone can be configured containing members specified by a combination of ports or aliases and WWNs or aliases of WWNs.

Zone configuration is managed on a fabric basis. When a new switch is added to the fabric, it automatically takes on the zone configuration information from the fabric. Adding a new fabric that has no zone configuration information to an existing fabric is very similar to adding a new switch. All switches in the new fabric inherit the zone configuration data. If the existing fabric has an effective zone configuration, then the same configuration becomes the effective configuration for the new switches. If a new switch that is already configured for zoning is being added to the fabric,

the zone configuration should be cleared on that switch before connecting it to the zoned fabric. When a change in the configuration is saved, enabled, or disabled according to the transactional model, it is automatically distributed to all switches in the fabric (by closing the transaction), preventing a single point of failure for zone information.

Zone changes in a production fabric can result in a disruption of I/O under conditions when an RSCN is issued because of the zone change and the HBA is unable to process the RSCN fast enough. Although RSCNs are a normal part of a functioning SAN, the pause in I/O might not be acceptable. For these reasons, you should perform zone changes only when the resulting behavior is predictable and acceptable. Ensuring that the HBA drivers are current can shorten the response time in relation to the RSCN.

Advanced Zoning Considerations

Brocade Fabric OS supports the following types of zones for advanced functionality:

- Broadcast zones: A broadcast zone restricts broadcast packets to only those devices that are members of the broadcast zone. Fibre Channel allows sending broadcast frames to all Nx_Ports, if the frame is sent to a broadcast well-known address (FFFFFF); however, many target devices and HBAs cannot handle broadcast frames. To control which devices receive broadcast frames, you can create a special zone, called a broadcast zone, that restricts broadcast packets to only those devices that are members of the broadcast zone and are also in the same regular zone. Devices that are not members of the broadcast zone can send broadcast packets, even though they cannot receive them. Broadcast zones are supported starting with Brocade FOS 5.3 and onwards.
- Frame redirection zones: Frame Redirection provides a means to redirect traffic flow between a host and a target that use virtualization and encryption applications, such as the Brocade SAS blade and Brocade Data Migration Manager (DMM), so that those applications can perform without having to reconfigure the host and target. Frame redirection zones are supported starting with Brocade FOS 5.3 and onwards.
- LSAN zones: These provide device connectivity between fabrics without merging the fabrics. A logicas SAN (LSAN) consists of zones in two or more edge or backbone fabrics that contain the same devices. LSANs essentially provide selective device connectivity between fabrics without forcing you to merge those fabrics. FC routers provide multiple mechanisms to manage interfabric device connectivity through extensions to existing switch management interfaces. To share devices between any two fabrics, the LSAN zone must be created in both fabrics that contain the port WWNs of the devices to be shared. LSAN zones are supported starting with Brocade FOS 5.2 and onwards.
- QoS zones: A Quality of Service (QoS) zone is a special zone that indicates the priority of the traffic flow between a given host/target pair. The members of a QoS zone are the host/target pairs. The switch automatically sets the priority for the “host,target” pairs specified in the zones based on the priority level (H or L) in the zone name. QoS zones are regular zones with additional QoS attributes specified by adding a QOS prefix to the zone name. WWN-based QoS zones are supported starting with Brocade FOS 6.0, and support for D,I (Domain, Index) QoS zones was added in Brocade FOS 6.3.
- Traffic Isolation zones (TI zones): These isolate inter-switch traffic to a specific, dedicated path through the fabric. The Traffic Isolation zoning feature allows you to control the flow of inter-switch traffic by creating a dedicated path for traffic flowing from a specific set of source ports. Enhanced TI zones allow the same port to be part of multiple TI zones at the same time. A TI zone can be created using D,I notation only, except for TI zones in a backbone fabric, which use port WWNs. TI zones are supported from Brocade FOS 6.0 and support for enhanced TI Zones was added in Brocade FOS 6.3.

Refer to the *Brocade Fabric OS Administrator’s Guide* for zone naming requirements for advanced zones.

Zoning Recommendations

- Use single initiator single target or single initiator and multiple target zone sets. In a large fabric, zoning by single HBA requires the creation of possibly hundreds of zones; however, each zone contains only a few members. Zone changes affect the smallest possible number of devices, minimizing the impact of

an incorrect zone change. This zoning philosophy is the preferred method and avoids RSCN performance concerns with multiple initiators in the same zone.

- Define zones using device WWPNs (World Wide Port Names).
- Monitor zone database size using the **cfgSize** CLI command. 1 MB is the maximum supported size of a zone database.
- Periodically back up and clean up the zone database of entries that are no longer in the fabric.
- If using Brocade HBAs, use Dynamic Fabric Provisioning.
- The default zone setting (what happens when zoning is disabled) should be set to No Access, which means that devices will be isolated when zoning is disabled.
- Always zone using the highest Fabric OS level switch. Switches with earlier Fabric OS versions do not have the capability to view all the functionality that a newer version of Fabric OS provides, as functionality is backwards-compatible but not forwards-compatible.
- Zone using an enterprise-class platform rather than a switch. An enterprise-class platform has more resources to handle zoning changes and implementations.
- Before implementing a new zone, verify the zone via the Zone Analyzer from Web Tools to isolate any possible problems. This is especially useful as fabrics increase in size.
- Follow vendor guidelines for preventing the generation of duplicate WWNs in a virtual environment.
- Zoning changes affect the entire fabric. Thus, when you are executing fabric-level configuration tasks, allow time for the changes to propagate across the fabric before executing any subsequent commands. For a large fabric, you should wait several minutes between commands.

Firmware Management

Brocade offers customers a choice in selecting Brocade FOS releases with new features or versions with extensive field deployments. Prior to upgrading, check the release notes of the selected Brocade FOS release to see if all the switches in your fabric are supported. Review the following questions before determining which Brocade FOS release to use:

- Why am I upgrading?
- Is it part of my server/SAN/storage firmware upgrade to keep current support?
- Do I need to upgrade to address a Technical Support Bulletin or bug fix that I encountered?
- Do I need the new features to monitor some fabric anomalies?
- Are there new switches in the fabric that require the latest firmware?

Switches can be upgraded sequentially or in parallel using Brocade Network Advisor or custom scripts. If you have a single resilient or dual redundant fabric, you can upgrade in parallel if application uptime is critical and can be assured.

Firmware Recommendations

- Upgrade firmware during non-business or peak hours if possible.
- If new features are not required, upgrade using the same release or minor release train.
- It is recommended that you not do a concurrent firmware upgrade on two switches that are physically E_Port connected.
- For a FICON environment, install firmware sequentially on switches in the FICON fabric. It is also recommended that the Control Unit Port (CUP) be varied offline before a FICON switch firmware upgrade.

NOTE: Refer to the IBM FICON qualification letter for the latest qualified release.

NOTE: Brocade Fabric OS (FOS) Target Path releases are recommended code levels for Brocade Fibre Channel switch platforms. These releases are guidelines to use when trying to determine the ideal version of Brocade FOS software, and they should be considered in conjunction with other requirements that may be unique to a particular environment. Refer to the Target Path section on the MyBrocade portal for the recommended Brocade FOS version appropriate for the environment.

ROUTING POLICIES

Data moves through a fabric from switch to switch and from storage to server along one or more paths that make up a route. Routing policies determine the path for each frame of data. Before the fabric can begin routing traffic, it must discover the route the frame should take to reach the intended destination. The routing policy configured on the switch determines the route selection, based on one of two user-selected routing policies:

- Port-based routing
- Exchange-based routing

Each switch can have its own routing policy, and different policies can exist in the same fabric.

NOTE: Setting either AP route policy is a disruptive process. Use the command **aptPolicy** to configure the desired routing policy. For most configurations, the default routing policy is optimal and provides the best performance. The routing policy should be changed only if there is a performance issue that is of concern, or if a particular fabric configuration or application requires it.

Port-Based Routing

The choice of routing path is based only on the incoming port and the destination domain. Thus, all the frames destined to a particular domain ingressing on a particular port follow the same route, as long as Dynamic Load Sharing is not enabled. This routing policy minimizes disruption caused by changes in the fabric (events not directly impacting the ports in the route); it represents a less efficient use of available bandwidth. To optimize port-based routing, Dynamic Load Sharing (DLS) can be enabled to balance the load across the available output ports within a domain.

Port-based routing is recommended for specific use cases:

- Some devices do not tolerate out-of-order exchanges; in such cases, use the port-based routing policy.
- FICON environments

Exchange-Based Routing

The choice of routing path is based on the Source ID (SID), Destination ID (DID), and Fibre Channel originator exchange ID (OXID), optimizing path utilization for the best performance. Thus, every exchange can take a different path through the fabric. Exchange-based routing requires the use of the DLS feature.

Exchange-based routing is also known as Dynamic Path Selection (DPS). DPS is where exchanges or communication between end-devices in a fabric are assigned to egress ports in ratios proportional to the potential bandwidth of the ISL or trunk group. When there are multiple paths to a destination, the input traffic is distributed across the different paths in proportion to the bandwidth available on each of the paths. This improves utilization of the available paths, thus reducing possible congestion on the paths. Every time there is a change in the network (which changes the available paths), the input traffic can be redistributed across the available paths. This is a non-disruptive process when the exchange-based routing policy is engaged.

NOTE: For Condor3 systems, one DPS entry can have up to 16 trunk groups. For Condor2 based systems, one DPS entry can have up to 8 trunk groups. To use 16 trunk groups on Condor2 based systems, two DPS entries can be created.

The trunking feature allows a group of physical links to merge into a single logical link, called a trunk group. Traffic is distributed dynamically and in order over this trunk group, achieving greater performance with fewer links, thus

optimizing the use of bandwidth. Within the trunk group, multiple physical ports appear as a single port, thus simplifying management. Refer to the *Brocade Fabric OS Administrator's Guide* for details on trunking.

Dynamic Load Sharing

With DLS enabled, Brocade Fabric OS balances ingress ports as evenly as possible across available ISL or trunk links.

The exchange-based routing policy depends on the Brocade Fabric OS DLS feature for dynamic routing path selection. When using the exchange-based routing policy, DLS is enabled by default and cannot be disabled. When the port-based policy is in force, DLS can be enabled to optimize routing. When DLS is enabled, it shares traffic among multiple equivalent paths between switches.

DLS recomputes load sharing when any of the following occurs:

- A switch boots up
- An E_Port goes offline and online
- An EX_Port goes offline
- A device goes offline
- There is a zone change in the fabric

DLS can be configured using the **dlsSet, dlsReset, dlsShow** commands from the CLI and GUI.

Lossless Dynamic Load Sharing

Lossless DLS enables DLS for optimal utilization of the ISLs without causing any frame loss. In other words, lossless DLS enables rebalancing port paths without causing input/output (I/O) failures. Lossless DLS was introduced starting with Brocade Fabric OS 6.2. Lossless mode ensures no frame loss during a rebalance and takes effect only if DLS is enabled. Note that "no frame loss" can be guaranteed only when a new additional path is used to do load rebalancing—"no frame loss" cannot be guaranteed on an existing data path that encounters the failure.

Lossless DLS can be enabled on a fabric topology in order to have zero frame drops during rebalance operations. If the end device also requires the order of frames to be maintained during the rebalance operation, then In-Order Delivery (IOD) must be enabled. However, this combination of lossless DLS and IOD is supported only in specific topologies, such as in a FICON environment.

Lossless DLS can be configured the **dlsSet, dlsReset, dlsShow** commands from CLI and GUI.

NOTE: In order to configure lossless DLS, the switches in the fabric must all have Brocade Fabric OS 6.3.0 installed, or they must all have Brocade Fabric OS 6.4.0 or later installed, to guarantee no frame loss. The lossless feature is disabled by default.

In-Order Delivery (IOD)

In a stable fabric, frames are always delivered in order, even when the traffic between switches is shared among multiple paths. However, when topology changes occur in the fabric (for example, if an ISL goes down), traffic is rerouted around the failure, and some frames can be delivered out of order. Most destination devices tolerate frames delivered out of order, but some do not.

By default, out-of-order frame-based delivery is allowed to minimize the number of frames dropped. Enabling IOD guarantees that frames are either delivered in order or dropped. You should enforce in-order frame delivery across topology changes, if the fabric contains destination devices that cannot tolerate occasional out-of-order frame delivery.

The order of delivery of frames is maintained within a switch and determined by the routing policy in effect. The frame delivery behaviors for each routing policy are as follows:

- Port-based routing: All frames received on an incoming port destined for a destination domain are guaranteed to exit the switch in the same order in which they were received.
- Exchange-based routing: All frames received on an incoming port for a given exchange are guaranteed to exit the switch in the same order in which they were received. Because different paths are chosen for different exchanges, this policy does not maintain the order of frames across exchanges.

In-Order Delivery can be configured using the **iodSet, iodReset and iodShow** commands.

NOTE: Some devices do not tolerate out-of-order exchanges; in such cases, use the port-based routing policy.

NOTE: The IOD capability can be enabled optionally for both port-based routing and exchange-based routing policies. In Brocade FOS versions prior to version 6.4.0, the lossless DLS feature was supported only for port-based routing, and IOD was always enabled.

Please refer to the *Brocade Fabric OS Administrator's Guide* for more details on routing policies.

FABRIC DIAGNOSTICS

Since the switch is the interconnection point between servers and storage, in any application anomalies that arise, the switch is normally blamed until proven otherwise. Brocade acknowledges the role the switches play in the fabric and has implemented hardware- and software-based diagnostic tools for improved problem determination and resolution.

Two areas customers struggle with are application performance due to high latency in the fabric and physical layer issues.

Device Latency

A device experiencing latencies 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 connected to the device.

Once it exhausts all available credits, the switch port connected to the device needs to hold additional outbound frames until a buffer credit is returned by the device. When a device does not respond in a timely fashion, the transmitting switch is forced to hold frames for longer periods of time, resulting in high buffer occupancy. This in turn 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 with devices attempting to send frames to devices attached to the switch with the high-latency device) and ultimately impacts the fabric.

NOTE: The impact to the overall fabric varies based on the severity of latency 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.

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 port type (E_Port or F_Port) and fail, often unpredictably and intermittently, making the failure even harder to diagnose. Faulty media involving F_Ports results in an impact to the end device attached to the F_Port and to devices communicating with this device. Failures on E_Ports can have an even greater impact. Many flows (host/target pairs) can simultaneously traverse a single E_Port. In large fabrics, this can be hundreds or even thousands of flows. In the event of a media failure involving one of these links, it is possible to disrupt some or all of the flows utilizing 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 is typically easy to detect and identify. When this occurs on an F_Port, the impact is specific to flows involving the F_Port. E_Ports are typically redundant, so severe failures on E_Ports typically only

result in a minor drop in bandwidth as the fabric automatically utilizes redundant paths. And the error reporting built into Brocade FOS readily identifies the failed link and port, allowing 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 can cause repeated errors, which might occur indefinitely or until the media fails completely. When these types of failures occur on E_Ports, the result can be devastating, as there can be repeated errors that impact many flows. This can result in significant impacts to applications that last for prolonged durations. Signatures of these types of failures include the following:

- CRC errors on frames
- Invalid words (includes encoder out errors)
- State changes (ports going offline/online repeatedly)
- Credit loss: Complete loss of credit on a VC on an E_Port prevents traffic from flowing on that VC, which results in frame loss and I/O failures for devices utilizing the VC.

Please refer to the *Fabric Resiliency Best Practices Guide* for details on using these tools to detect and mitigate application performance issues.

DATA COLLECTION FOR SUPPORT

When troubleshooting SAN fabric anomalies, it is important to capture all the required data and information before escalating the issue to the support provider. Here is an outline of information gathering steps to follow before escalating the issue:

1. Intermittent or continuous issue (no recovery or periodic recovery from the issue)? Specific time period (the observed issue only occurs at specific times or when certain tasks/jobs are executed)?
2. Recent changes to fabric/environment?
    a. This might be any change, no matter how small. Examples might be zone changes, adding of ISLs, disabling of ports, adding of devices, port configuration changes, firmware changes, device configuration changes, and so on.
3. What type of initiator (host) issue is observed?
    a. Provide the error(s) that the host is observing during the observed issue.
    b. What analysis of the error, if any, has the vendor provided?
4. What types of target (storage) errors are observed?
    a. Provide the error(s) that the host is experiencing during the observed issue.
    b. What analysis of the error, if any, has the vendor provided?
5. What is the connection location of the initiator and target (<slot>/port [area/index] and SAN switch name/number)?
6. If it is a performance issue, refer to “Other Host/Storage related issues” for collecting additional device information.
7. What should be sent to Brocade support personnel?
    a. A topology diagram of the SAN (refer to Brocade SAN Health)
    b. Switch model, serial number, and Brocade FOS version of the switch under investigation
    c. The **supportSave** command output
    d. Description of any troubleshooting steps already performed and their results
    e. Serial console and Telnet session logs
    f. Syslog message logs

APPENDIX A: CONFIGURING PORT FENCING

Use the **portFencing** CLI command to enable error reporting for the Fabric Watch port fencing feature on all ports of a specified type—and to configure the ports to report errors for a specific area. Supported port types include E_Ports, F_Ports, and physical ports. A specified port type can be configured to report errors for one or more areas.

Port fencing monitors ports for erratic behavior and disables a port if specified error conditions are met. The **portFencing** CLI command enables or disables the port fencing feature for an area of a class. You can customize or tune the threshold of an area using the **portthConfig** CLI command.

Use **portFencing** to configure port fencing for C3_TX_TO. For example:

```
portfencing --enable fop-port -area C3TX_TO
```

The same command can be used to configure port fencing on link reset. For example:

```
portfencing --enable fop-port -area LR
```

Use **portThconfig** to customize port fencing thresholds:

```
switch:admin> portthconfig --set port -area crc -highthreshold -value 2 -
trigger above -action email

switch:admin> portthconfig --set port -area crc -highthreshold -trigger below -
action email

switch:admin> portthconfig --set port -ar crc -lowthreshold -value 1 -trigger
above -action email

switch:admin> portthconfig --set port -ar crc -lowthreshold -trigger below -
action email
```

To apply the new custom settings so they become effective:

```
switch:admin> portthconfig --apply port -area crc -action cust -thresh_level
custom
```

To display the port threshold configuration for all port types and areas:

```
switch:admin> portthconfig --show
```

Refer to the *Brocade Fabric Watch Administrator's Guide* and *Fabric Resiliency Best Practices Guide* for the recommended thresholds that have been tested and tuned to quarantine misbehaving components.

Refer to the *Brocade Fabric OS Command Reference Guide* for CLI details.

APPENDIX B: TERMINOLOGY

Term	Brief Description
48K	Brocade 48000 Director, 8-slot modular chassis
Base Switch	Base Switch of an enabled virtual fabric mode switch
DCX	Brocade DCX Backbone, 8-slot modular chassis
DCX-4S	Brocade DCX-4S Backbone, 4-slot modular chassis
Default switch	Default switch of an enabled virtual fabric mode switch
E_Port	A standard Fibre Channel mechanism that enables switches to network with each other
Edge Hold Time	Enables the switch to time out frames for F_Ports sooner than for E_Ports
EX_Port	A type of E_Port that connects a Fibre Channel router to an edge fabric
F_Port	A fabric port to which an N_Port is attached
FCIP	Fibre Channel over IP, which enables Fibre Channel traffic to flow over an IP link
FCR	Fibre Channel Routing, which enables multiple fabrics to share devices without having to merge the fabrics
ICL	Inter-Chassis Link, used for connecting modular switches without using front-end device ports
IFL	Inter-Fabric Link, a link between fabrics in a routed topology
ISL	Inter-Switch Link, used for connecting fixed port and modular switches
Logical switch	Logical switch of an enabled virtual fabric mode switch
Oversubscription	A condition in which more devices might need to access a resource than that resource can fully support
Port group	A set of sequential ports that are defined (for example, ports 0–3)
QoS	Quality of Service, a traffic shaping feature that allows the prioritization of data traffic based on the SID/DID of each frame
Redundant	Duplication of components, including an entire fabric, to avoid a single point of failure in the network (fabrics A & B are identical)
Resilient	Ability of a fabric to recover from failure, could be in a degraded state but functional (for example, ISL failure in a trunk group)
TI Zone	Traffic Isolation Zone, which controls the flow of interswitch traffic by creating a dedicated path for traffic flowing from a specific set of source ports
Trunk	Trunking that allows a group of physical links to merge into a single logical link, enabling traffic to be distributed dynamically at the frame level
VC	Virtual cChannels, which create multiple logical data paths across a single physical link or connection
VF	Virtual Fabrics, a suite of related features that enable customers to create a Logical Switch, create a Logical Fabric, or share devices in a Brocade Fibre Channel SAN

APPENDIX C: REFERENCES

Software and Hardware Product Documentation

- Brocade Fabric OS v7.0.x Release Notes
- Brocade Fabric OS Administrator's Guide, supporting Brocade Fabric OS v7.0.x
- Brocade Fabric OS Command Reference Manual, supporting Brocade Fabric OS v7.0.x
- Brocade Fabric Watch Administrator's Guide, supporting Brocade Fabric OS v7.0.x
- Brocade Access Gateway Administrator's Guide, supporting Brocade Fabric OS v7.0.x
- Brocade Fabric OS Troubleshooting and Diagnostics Guide, supporting Brocade Fabric OS v7.0.x
- Hardware Reference Guides and QuickStart Guides for backbone, director, switch, and blade platforms

Technical Briefs

www.brocade.com/sites/dotcom/data-center-best-practices/resource-center/index.page

www.brocade.com/products/all/san-backbones/product-details/dcx8510-backbone/specifications.page

Fabric Resiliency Best Practices Guide
www.brocade.com/sites/dotcom/data-center-best-practices/resource-center/index.page

Brocade Compatibility and Support

www.brocade.com/forms/getFile?p=documents/matrices/compatibility-matrix-fos-7x-mx.pdf

www.brocade.com/solutions-technology/enterprise/connectivity/mainframe/services.page

Brocade Scalability Guidelines

www.brocade.com/products/all/san-backbones/product-details/dcx8510-backbone/index.page
Document is located at bottom of page in the DCX 8510 Backbones Resources under Matrices.

Brocade SAN Health

www.brocade.com/services-support/drivers-downloads/san-health-diagnostics/overview.page

Brocade Bookshelf

- *Principles of SAN Design* (updated in 2007) by Josh Judd
- *Strategies for Data Protection* by Tom Clark
- *Securing Fibre Channel Fabrics* by Roger Bouchard
- The New Data Center by Tom Clark

Other

- www.snia.org/education/dictionary
- www.vmware.com/pdf/vsp_4_san_design_deploy.pdf
- www.vmware.com/files/pdf/vcb_best_practices.pdf
- www.knowledgebase.tolisgroup.com/?View=entry&EntryID=95