WHITE PAPER

ZONING IMPLEMENTATION STRATEGIES FOR BROCADE SAN FABRICS

A best-practices guideline for determining an effective zoning configuration for your SAN environment

BROCADE
The intelligent platform for networking storage

This document is designed to provide an overview of zoning practices for Storage Area Networks (SANs) based on Brocade® SilkWorm® 2xxx, 3xxx, and 12xxx fabric switches running Brocade Fabric OS v2.6.x, 3.x, and 4.x. In addition to explaining the concepts of soft and hard zoning, this guide describes multiple zoning strategies and their relative merits. A practical zoning naming convention is included, with a cookbook-style method for zone configuration and implementation.

This document is not intended to be an exhaustive guide for all possible zoning strategies and implementations. Because there are slight variations in the zoning functions available in different switch models and Fabric OS versions, you should refer to the appropriate Brocade product manual for details on the specific functions available.

Historical Perspective of Zoning

Before implementing a zoning strategy for your SAN environment, it helps to understand some of the historical context behind zoning practices. The original SCSI cable implemented a form of physical zoning. In this implementation, each SCSI bus was attached to a maximum of two initiators (typically host servers) and up to 15 targets (such as disk, tape, and printer devices). Only devices connected to the SCSI bus were able to access the other devices connected to the bus. As a result, the SCSI bus formed a zone that was enforced by physical device attachment.

When the initiator first accessed the bus, it executed a series of actions to initialize the bus and discover the attached devices. The first action was a SCSI reset, which informed all attached devices to reset to an initial condition and prepare for communication. The initiator then performed a SCSI probe sequence, during which it attempted to communicate with all devices attached to the bus. Because the initiator had no knowledge of connected devices, it had to "probe" each possible device, waiting for a timeout to occur before presuming that the device was not there. This process often led to a long initialization time.

Early SCSI implementations allowed only seven targets and two Logical Unit Numbers (LUNs) per target. Later, this limit was increased to 15 targets and 16 LUNs per target. However, even with a one-second timeout on the probe, this led to excessively long initialization sequences. (As a result, new methods were implemented on the host operating system to limit the targets and LUNs that would be probed during the initialization sequence.)

After the initiator discovered the devices, it issued a SCSI inquiry to each device in order to discover its properties. Again, as the number of devices increased, the time the initiator spent discovering the properties of all attached devices often became excessive. This long initialization time limited the number of devices that could be supported on a server.

As the environment grew in size, there was also a need for servers to share devices—especially tape devices. The physical SCSI bus permitted only one initiator per bus and allowed devices to connect to only a single bus. Some specialized hardware was available that permitted limited sharing of devices but, in general, the single bus attachment of initiators and targets limited the size of the environment.

Device Discovery in SANs

With the emergence of Fibre Channel SANs, the number of devices the host might be required to discover increases dramatically. The Fibre Channel standard enables the attachment of up to 16 million devices on the network and large SAN fabrics now consist of over 1500 devices (targets and initiators), with the number of supported LUNs per target exceeding 256. This change in scale has required a new method to discover the devices on the fabric in a more acceptable time frame.

To address this need, product designers introduced the fabric well known service "nameserver". Now, when a device initially accesses the fabric, it must perform a login process. During this process, the device registers with the nameserver, which records information about the device such as protocols supported, World Wide Name (WWN), network address, supported class of service, and so on (see Figure 1).

Figure 1.
Registering with the nameserver provides a variety of information about fabric devices.

```
Type  Pid      COS PortName                NodeName             TTL(sec)
 N    441100;  3;  21:00:00:e0:8b:05:8f:b7; 20:00:00:e0:8b:05:83:b7; na
    Fabric Port Name: 20:01:00:60:69:15:09:0e
 N    441200;  3;  21:00:00:e0:8b:05:95:b7; 20:00:00:e0:8b:05:95:b7; na
    Fabric Port Name: 20:02:00:60:69:15:09:0e
 N    441300;  3;  21:00:00:e0:8b:05:9c:b7; 20:00:00:e0:8b:06:d2:30; na
    Fabric Port Name: 20:03:00:60:69:15:09:0e
 N    441400;  3;  21:00:00:e0:8b:05:d8:e9; 20:00:00:e0:8b:05:9e:b7; na
    Fabric Port Name: 20:04:00:60:69:15:09:0e
 N    441500;  3;  21:00:00:e0:8b:05:52:b6; 20:00:00:e0:8b:05:90:b7; na
    Fabric Port Name: 20:05:00:60:69:15:09:0e
```

When an initiator accesses the network, it can query the nameserver for all attached devices and their capabilities. As a result, the need for the initiator to probe the network for devices is eliminated—significantly reducing the time to initialize the environment. However, this approach creates a new issue. Unlike the SCSI bus that physically isolated initiators and targets, the fabric provides universal access for all initiators and targets. Any initiator can probe the fabric for all targets before proceeding to initialize the targets and perform I/O. This action can be disastrous if another initiator is already utilizing the target and has stored data on it, such as with a disk or tape device. To solve this problem, zoning was developed to provide a virtual SCSI bus that limits the number of devices that an initiator can discover and access.

The Principles of Zoning

A zone is a list of devices in the fabric that need to communicate with each other. And zoning is essentially a filtering process that limits the information the nameserver returns to the initiator during a query. When an initiator queries the nameserver for available devices in the fabric, the nameserver searches for all zones that contain the initiator. It then creates a list of all the members of the zones and compares it to the contents of the nameserver. Only devices that appear in both the list and the nameserver are returned to the initiator as available devices. (This practice is equivalent to a SCSI probe sequence in the original SCSI bus.) The initiator then uses the list of returned devices—along with the returned access parameters—to query the devices for specific operational capabilities.

Devices that are not part of the zone are omitted from the query response. If the initiator is not part of a zone or has no targets included in its zone, the nameserver response lists no devices. Note that this process does not prohibit access to the device, because the nameserver does not control access to fabric devices. If the initiator has prior knowledge of the device, or probes the fabric for devices, it can still gain access. As a result, reliable zoning operation depends on the cooperation of all initiators in the fabric to access only devices that are returned in the query from the nameserver. To provide device access control, Brocade fabric switches support hardware enforcement of the zoning configuration.

Devices can be members of more than one zone. This enables the creation of zones that contain some, but not all, of the same members. This type of configuration is referred to as overlapping zones. In this case, a concept of "most permissible access" is employed. As long as at least one zone contains both the initiator and target, the target is made available to the initiator, even though other zones containing the initiator do not contain the target device.

Fabric Zone Management

To manage fabric zoning information, you can use a command line interface, the Brocade WEB TOOLS GUI utility, or the Brocade Fabric Access API via third-party management applications. There are three steps for creating a zoning configuration:

1. Include one or more members (devices) in a zone.
2. Include one or more zones in a configuration.
3. Define one or more configurations, making only one the effective configuration for the fabric.

Note that zoning configurations are managed on a fabric basis. Brocade Fabric OS automatically distributes zoning configuration data to all switches in the fabric, which prevents a single point of failure. Although it is possible to manage zoning configurations from any switch, the best practice is to select one switch for all zoning administration. You should typically select a core switch or the switch with the most recent version of Fabric OS.

The Effect of Zoning on SCN Delivery

Whenever a change occurs in the nameserver—such as a device addition to, or removal from, the fabric—Fabric OS generates a State Change Notification (SCN). In the absence of zoning, an SCN is sent to all devices in the fabric, with each device querying the nameserver to determine how the membership of the fabric has changed. This process occurs even if the device change does not affect the device being notified. Especially in large fabrics, this can result in a significant amount of fabric service traffic (although typically for only a short time).

For instance, if a new initiator joins the fabric, there is little reason to notify all the other initiators of the change since initiators do not usually communicate with each other. An equivalent situation exists with targets. Because targets generally do not communicate with each other, they have little use in being notified about the addition or removal of another target.

Although all devices are supposed to handle SCN traffic without affecting normal operation, this is not always the case. Thus, the overall stability of the fabric increases after zoning is implemented: the fabric can restrict SCN delivery to only those devices in a zone with the added device. The same holds true when a device is removed from the fabric. The list of available devices is restricted to just those of interest to the initiator, rather than all devices in the fabric.

Possible Zoning Implementations

You can implement zoning at various levels—in hosts, most storage units, and the switch fabric:

- Host-based zoning can include WWN or LUN mapping, and is typically known as "persistent binding."
- Storage units typically implement LUN-based zoning, commonly referred to as "LUN masking."
- Fabric switches implement nameserver-based zoning where zone members are identified by their WWN or port location in the fabric.

Host-Based Zoning (Persistent Binding and LUN Mapping)

Host-based zoning is usually referred to as persistent binding or LUN mapping, and is perhaps the least implemented form of zoning. Because it requires the host configuration to be correct in order to avoid zoning conflicts, this form of zoning creates a greater opportunity for administrative errors and conflicting access to targets. Moreover, zoning interfaces vary among different host operating systems and HBAs—increasing the possibility for administrative errors. If a host is not configured with the zoning software, it can access all devices in the fabric and create an even higher probability of data corruption.

Host-based zoning is often used when clusters are implemented to control the mapping of devices to specific SCSI target IDs. However, it should never be the *only* form of zoning. Augmenting host-based zoning with storage- and fabric-based zoning is the only acceptable method to reliably control device access and data security.

Storage-Based Zoning (LUN Masking)

Storage-based zoning is usually referred to as LUN masking, and its basic function is to limit access to the LUNs on the storage port to the specific WWN of the server HBA. This form of zoning should be implemented in all SANs. It functions during the probe portion of the SCSI initialization where the server probes the storage port for a list of available LUNs and their properties. The storage system then compares the WWN of the requesting HBA to the defined zone list, returning the LUNs assigned to the WWN. Any other LUNs on that storage port are not made available to the server. To ensure reliability, this function must be implemented in the storage unit.

Fabric-Based Zoning

Fabric-based zoning is commonly referred to as nameserver-based or "soft" zoning. When a device queries the fabric nameserver, the nameserver determines which zone(s) the device belongs to. It then returns to the requesting device information about all members of the zone(s) present in the fabric. Devices in the zone can be identified by World Wide Node Name, World Wide Port Name, or domain/port of the switch the device is connected to. This form of zoning should be implemented in all SANs.

Fabric-Based Zoning Strategies

Because there are a number of implementation strategies for fabric-based zoning, it is perhaps the most controversial aspect of zoning. Although each strategy provides a different level of fabric security and stability, all strategies should work well in most cases. The primary forms of fabric-based zoning are:

- No fabric zoning
- Zoning by application
- Zoning by operating system
- Zoning by port allocation
- Single HBA zoning

The decision on which zoning strategy to use is critical, because you must verify the appropriate level of support from your server, HBA, switch, and storage vendors.

No Fabric Zoning

No fabric zoning is the least desirable strategy because of the unrestricted availability devices have to the fabric. All devices have full knowledge of the fabric membership and potentially have access to all the devices. Moreover, any device attached to the fabric—intentionally or maliciously—has unrestricted access to the fabric. In a large fabric, the delivery of SCNs can disrupt some HBAs because of the large size of the query returned. Only in rare cases should you not use fabric-based zoning—and then you should use an alternate form of zoning to provide adequate fabric security and stability.

Zoning by Application

In most environments, the concern for availability is at the application level, not the server level. This concern most likely requires zoning multiple operating systems into the same zones. However, some operating systems have difficulty co-existing in the same zone. In addition, a minor server in the application suite potentially has the ability

server. Zoning by application can also result in a zone with a large number of members that would provide a greater possibility of administrative errors. As a result, you should use this zoning strategy only after careful testing in a non-production environment.

Zoning by Operating System

The issues with zoning by operating system are much the same as zoning by application. In a large site, the zone might become very large and complex. When zone changes are made, they typically do not involve a particular server type, but rather applications. In some cases, operating system clusters might be involved. If members of different clusters can see storage assigned to another cluster, they might attempt to own the other cluster's storage and compromise the stability of the clusters. As a result, you should use this zoning strategy only after careful testing in a non-production environment.

Zoning by Port Allocation

You can implement zoning based on switch port rather than the WWN of the attached device. This strategy provides good security in the fabric but requires reliable processes to prevent incorrect devices from being attached to the wrong ports. You should normally avoid this form of zoning unless you have rigidly enforced processes for port and device allocation in the fabric.

However, zoning by port allocation has some positive features. For instance, when your replace a storage port, server HBA, or tape drive, the change of WWN for the new device is of no consequence. As long as the new device is connected to the original port, it continues to have the same access rights. You can pre-associate the ports on the edge switches to storage ports and establish control of the fan-in ratio. Following this technique, you cannot overload any storage port by associating too many servers with it.

Single HBA Zoning

Single HBA zoning most closely recreates the original SCSI bus. Each zone created has only one HBA (initiator) in the zone, with each target device then added to the zone. Typically, you should create a zone for the HBA and then add disk storage ports. If the HBA accesses tape devices, you should create a second zone for the HBA and associated tape devices. In the case of clustered systems, it might be appropriate to include an HBA from each of the cluster members in the zone. This technique is equivalent to having a shared SCSI bus between cluster members, and presumes that the clustering software provides a way to manage access to the shared devices.

In a large fabric, this requires the creation of possibly hundreds of zones. However, each zone would contain only a few members. As a result, zone changes would affect the smallest possible number of devices, minimizing the impact of an incorrect zone change. Because of these advantages, the single HBA zoning strategy is considered a best practice.

Implementing Zoning

The first step in implementing zoning is to establish a naming convention that is consistent, that produces meaningful names, and that you can use consistently. Other key aspects include implementing zoning in a new fabric, and implementing zoning in an existing fabric that does not conform to the single HBA zoning method.

Simple Naming Conventions

Three types of devices typically have an alias: server HBAs, storage ports, and tape ports. New devices such as virtualization devices and SAN management appliances might also require an alias. Any device that acts as either a SCSI initiator or SCSI target has an alias. In turn, each alias has three components: device type, device identifier, and port location. A suggested naming scheme for device types might be:

- **SRV**: Server
- **STO**: Storage
- **TPE**: Tape
- **VRA**: Virtualization appliance

The device identifier can be the host name for the server, the serial number, or the frame identifier for the storage. The port location should allow administrators to physically find the device port where the fiber is attached. A suggestion would be the PCI slot on a server, the fiber port on a storage frame, or the port on a tape drive. If a unit has only one port, you can omit this field or use identifier "SNG" for single attachment. For example:

- **SRV_MAILSERVER_SLT5**: A server with hostname "mailserver" in PCI slot 5
- **TPA_LTO9_SNG**: A tape with LTO drive number 9, single-attached
- **STO_DSK3456_5C**: A storage unit with serial ID 3456 on the fifth card in port C

A field to denote the fabric in dual-fabric environments might also be useful. Although shorter names are easier to remember and less prone to typing errors, you should not sacrifice meaning for brevity.

In a similar manner, you should name zones for the initiators they contain. For the server alias described above, the zone would be ZNE_MAILSERVER_SLT5. This clearly identifies the server HBA associated with the zone.

Configuration naming is more flexible. In general, you should name one configuration "PROD fabricname" with "fabricname" representing the name of the fabric. The purpose of the PROD configuration is to easily identify the configuration that can be implemented and provide the most generic services. If you use other configurations for specialized purposes, you can use names such as "BACKUP_A", "RECOVERY_2", or "TEST_18jun02".

Implementing Zoning in a New Fabric

A new fabric has no attached devices. However, a zone configuration must contain as least one zone, and the zone must contain at least one member. The easiest method to fulfill this requirement is to use the switch WWN to create an alias. The command "switchshow" provides the WWN of the switch. You can use this to create an alias with the switch name. Next, you can create a zone according to the naming convention.

Finally, you can create the configuration. For a switch with the name of "switch22", you would issue the following commands:

- Alicreate "switch22","20:00:60:60:12:34:56"
- Zonecreate "ZNE_switch22","switch22"
- Cfgcreate "PROD_faba","ZNE_switch22"
- Cfgenable "PROD_faba"
- Cfgsave

This sequence enables zoning in the fabric. Next, attach all the devices and allow them to join the fabric. The WEB TOOLS utility is the most expeditious way to create the alias and zone entries while minimizing the chance of typing errors. Create aliases for all of the devices, then create the zones for the SCSI initiator devices.

The final step is to add the SCSI target devices into the appropriate zones. If initiators are accessing disks that are on the same target port, be sure to implement some form of LUN masking to avoid concurrent access by multiple initiators to the same target disk. Failure to do so can result in data corruption. Next, create the configuration by adding zones. After this step is complete, enable the configuration with "cfgenable", then save it with "cfgsave".

At this point, the initiators might not be able to see their devices, because the HBA drivers and operating systems vary significantly in their ability to dynamically add devices. You should consult the appropriate product manuals for your systems and HBAs to determine how to add the devices to your systems. In most cases, a reboot of the server usually works.

Implementing Zoning in an Existing Fabric

The fact that zoning is already implemented in the fabric eliminates the first step of creating zones and a configuration. The step of creating aliases and zones, then adding them to the existing configuration is the same. It is acceptable to have overlapping zones. In fact, the single HBA zones should overlap one or more of the existing zones.

After the single HBA zones are in place, you can remove the previous zones and any aliases that are no longer needed. With careful execution, you can make this non-disruptive to the devices on the fabric. However, there might be a momentary disruption of data flow when the new zoning configuration is enabled due to the SCN. Properly configured devices should handle this without any problems.

Deviations from Single HBA Zoning

Certain cases might require you to deviate from the single HBA zoning strategy—such as clustered systems, SAN management appliances, and initiators that must communicate with each other because another protocol such as IP is being utilized. When this need arises, you should understand all the ramifications and be sure to test your strategy in a non-production environment.

Hardware-Enforced Zoning in Brocade Fabrics

All zoning employs the nameserver to limit the information returned to an initiator in response to a nameserver query. This practice is referred to as "soft" zoning. If an initiator has knowledge of the network address of a target device, it does not need to query the nameserver to access it. This allows undesired access to a target device and the potential for data corruption. To help prevent this, Brocade switches augment nameserver-based zoning with hardware enforcement. Nameserver or soft zoning is always active whenever a zone configuration is in effect. However, there are certain conditions that determine when hardware enforcement is active.

When hardware-enforced zoning ("hard" zoning) is active, Brocade switches monitor the communications and block any frames that do not comply with the effective zone configuration. This blocking is performed at the transmit side of the port where the destination device is located. A zone configuration is typically comprised of many zones, and the hardware enforcement decision is made on a zone-by-zone basis (the exact methodology varies based on switch model). Fabric OS determines the use of hardware enforcement whenever the fabric membership or zone configuration changes.

To better understand this concept, you should familiarize yourself with the following definitions:

- **Domain/port zone**: All members of this zone are specified with domain/port pairs only. This is sometimes referred to as hard port zoning.
- **WWN zone**: All members of this zone are specified with WWN numbers only. They can be node or port versions of the WWN.
- **Mixed zone**: This zone contains members specified by either domain/port or WWN, and it contains at least one member specified by each method.
- **Overlapping zones**: This involves two or more zones containing a common member that can be specified by domain/port or WWN.

Fabrics Composed of Only SilkWorm 2xxx Switches

SilkWorm 2xxx switches enable hardware-enforced zoning only on domain/port zones. Any WWNs or mixed zones are not hardware-enforced. Any domain/port zone that overlaps with a mixed or WWN zone is not hardware-enforced. Note that an overlap occurs when a member specified by WWN is connected to a port in a domain/port zone. The domain/port zone loses its hardware enforcement, even though a review of the zone configuration would not indicate this (see Figure 2).

```
zonecreate "domport","5,1;5,8;12,1;33,9" (domain/port only zone,
hardware-enforced)
zonecreate "wwn","21:00:00:e0:8b:05:8f:b7;21:00:00:e0:8b:05:95:b7"
(World Wide Name only zone, not hardware-enforced)
zonecreate
"mixed","5,1;12,9;21:00:00:e0:8b:05:8f:b7;21:00:00:e0:8b:05:95:b7"
(mixed zone, not hardware-enforced)
```

Figure 2.
Hardware-enforced zoning on domain/port zones.

Fabrics Composed of Only SilkWorm 3xxx and 12xxx Switches

SilkWorm 3xxx and 12xxx switches enable hardware-enforced zoning on domain/port zones and WWN zones. Mixed zones are not hardware-enforced. Overlapping like zone types does not result in the loss of hardware enforcement, but overlapping with a different zone type does result in the loss of hardware enforcement. As is the case with the SilkWorm 2xxx switches, connecting a device specified by WWN into a port specified in a domain/port zone results in the loss of the hardware enforcement in both zones (see Figure 3).

Figure 3.
Connecting unmatched devices in a zone results in the loss of hardware enforcement.

```
zonecreate "domport","5,1;5,8;12,23;33,9" (domain/port only zone,
hardware-enforced)
zonecreate "wwn","21:00:00:e0:8b:05:8f:b7;21:00:00:e0:8b:05:95:b7"
(World Wide Name only zone, hardware-enforced)
zonecreate
"mixed","5,1;12,9;21:00:00:e0:8b:05:8f:b7;21:00:00:e0:8b:05:95:b7"
(mixed zone, not hardware-enforced)
```

Mixed Fabrics Composed of SilkWorm 2xxx and 3xxx/12xxx Switches

In mixed fabrics, each switch type continues to enable hardware enforcement of zoning according to the conditions described above. The challenge is that you must know which switch type the device is attached to, in order to determine whether hardware enforcement is active.

Special Note Regarding SilkWorm 3xxx/12xxx Switches

Even when a zone is not hardware-enforced, SilkWorm 3xxx/12xxx switches still have an additional level of zoning protection in the hardware. Before an initiator can communicate with a target, it must establish communication with the target device. First, you must execute one or more of the Fibre Channel primitive commands PLOGI, PDISC, or ADISC. The switches trap the primitive commands in the hardware and forward them to the nameserver. The nameserver then compares the initiator and target in the primitive commands with the current zone configuration. If the current zone configuration does not permit the devices to communicate, the switch issues a reject to the initiator, effectively blocking communication. Although the switch does not block individual frames in a previously established connection, this action prevents the establishment of a new connection between devices.

Guidelines for Ensuring Hardware-Enforced Zoning

After you have established a policy requiring hardware enforcement, hardware enforcement should always occur. From the preceding sections, it should be clear that determining the applicability of hardware enforcement can sometimes be difficult. There are three basic cases for ensuring hardware enforcement:

- For SilkWorm 2xxx switch fabrics, use domain/port identifiers only. Never identify a zone member by WWN.
- For SilkWorm 3xxx/12xxx switch fabrics, use either WWN or domain/port zones. Never use both, and never used mixed zones.
- For fabrics with SilkWorm 2xxx and 3xxx or 12xxx switches, preferably use domain/port identifiers. If you want to use WWN identifiers, utilize the following methodology. Always place disk and tape devices (targets) on the SilkWorm 3xxx/12xxx switches, and use WWN zoning only. If all the targets are on the SilkWorm 3xxx/12xxx switches, hardware-enforced zoning is active to protect them. This does not hardware-protect the initiators on the SilkWorm 2xxx switches, but they usually do not offer resources anyway, and will reject probe attempts by other initiators.

Additional Zoning Information

In case you need additional help in determining the most effective zoning strategy for your particular environment, Brocade offers the following technical resources.

Books

Building SANs with Brocade Fabric Switches (**www.brocade.com/san/san_book.jsp**)

Online Resource Library

Brocade SAN Info Center Web site (**www.brocade.com/san**)

Technical Papers

The following technical papers are located at **www.brocade.com/san/technical_guides.jsp**:

Brocade Guide to Understanding Zoning

Brocade SAN Design Guide

Brocade SilkWorm 12000: Design, Deployment, and Management Guide

Building and Scaling Brocade SAN Fabrics: Design and Best Practices Guide

Brocade Product Manuals

Brocade Zoning Version 2.6 Publication Number 53-0000202-02

Brocade Zoning Version 3.0 Publication Number 53-0000135-03

Brocade Zoning Version 4.0 Publication Number 53-0000187-02

Corporate Headquarters
San Jose, CA USA
T: (408) 487-8000
info@brocade.com

European Headquarters
Geneva, Switzerland
T: +41 22 799 56 40
europe-info@brocade.com

Asia Pacific Headquarters
Tokyo, Japan
T: +81-3-5402-5300
apac-info@brocade.com

Latin America Headquarters
Miami, FL USA
T: 305-716-4165
latinam-sales@brocade.com