Best Practices

10

This chapter describes “best practices” for implementing heterogeneous Storage Area Networks. The information contained in this chapter should be used as a guide for constructing your SAN. Although every attempt has been made to provide a best practice recommendation, some aspects of SAN implementation are a matter of preference. Also, the physical location of servers, storage, computer labs, or specific building layout and location may dictate particular aspects of your SAN implementation. In part, this is an expected reality and is often easily accommodated, given the inherent flexibility in implementing SANs and Fibre Channel technology.

Rather than just present a list of best practices, the information has been organized into these sections:

- Planning a SAN
- Configuring a SAN
- Upgrading a SAN
- Migrating SAN Topologies
- Merging SAN Fabrics
- Troubleshooting

Much of what is presented here is the result of the actual experiences of building large SANs within the internal HP engineering environment and at customer sites.

Although this chapter does describe portions of the design process in the planning phase below, it is not meant to convey the entire SAN design process. Contact an HP Enterprise Storage Consultant or the Professional Services organizations for assistance and consultation on designing SANs. HP Storage Services may be contacted through this link:

http://h18005.www1.hp.com/services/storage/index.html

Note: Much of the information in this chapter applies equally to SANs with the B-Series, M-Series, or C-Series Fabric product lines of switches. Any reference to specific switch features pertains only to the B-Series product line.

Planning a SAN

Proper planning considers both present and future requirements. This can be accomplished by over-planning your initial SAN capacity and connectivity requirements to accommodate expected future needs. Whether using an HP standard topology or designing your own topology, select a design that not only offers the best implementation for present usage, but also allows you to expand your SAN over time.

It is important that you allocate an adequate amount of time to plan your SAN. In general, the more detail you can define in the planning phase, the greater the benefit you will realize during the configuration phase.

Consider each of these items during the planning phase:

- Deployment Strategy: You can choose to deploy separate smaller SANs or SAN Islands with the idea of increasing capacity by growing the SANs independently or by interconnecting the independent SANs in the future. Smaller SANs are easier to construct, larger SANs offer economies of scale from an operational standpoint, but take longer and are more complex to build.
- Topology Design: Consider the topology design compared to the ease of migrating to another, higher capacity design. In most cases this can be accommodated; however, it is always preferable to choose an initial design that can grow, without the need to transition to a different topology.
- Experience Level: If you are just beginning deployment of SAN technology, consider starting with a smaller implementation. As you gain experience, deploy larger SANs.
- SAN Management Strategy: Refer to *Chapter 6, SAN Fabric Management Tools* and Chapter 6, SAN Storage Management Tools for information about SAN management tools. After reviewing this chapter, define the management strategy and the specific tools that you will utilize to manage your SAN.
- Technology Advances: The ideal design considers expected future technological advances, and can easily accommodate the resultant changes. Plan for flexibility in your initial design. Higher port count Fibre Channel switches and faster interconnect speeds are an inevitable evolution of Fibre Channel technology. Ensure that your initial plan addresses and can accommodate expected changes such as these.
- Document the Design: This is one of the most important aspects of the planning process. This allows you to fully review and evaluate the design beforehand, evaluate trade-offs, make changes, and effectively communicate specific plans to all groups affected. The other important benefit of documenting your design is that during the later phases of implementation, the documentation serves as the roadmap for the actual implementation.

HP recommends, at a minimum, that you document the following before beginning the actual implementation:

1. Topology Map–Shows the logical SAN topology and fabric interconnect scheme; conveys the overall design from a strategic standpoint, and can also serve to convey how future growth and technological advances will be accommodated.
2. Configuration Layout–Shows the physical layout of the entire implementation. More detailed then the topology map, the layout is used during implementation to verify the correct connectivity. This is also extremely helpful if troubleshooting is required in later phases.
3. Storage Map–Defines the storage system arrangement and configuration in the SAN, and storageset settings such as SSP and RAID levels. This map effectively defines how all of the storage is configured in the SAN.

4. Zoning Map–Defines the inter-node communication access within the SAN. This map defines which nodes or user ports are allowed to communicate with each other in the SAN.

General Planning Considerations

It is difficult to make general recommendations about the choice of a specific SAN topology. There are so many variables in large installations that each new configuration requires substantial customized design work. The following suggestions provide background information for designs that meet typical large SAN requirements and that are compatible with the future direction of StorageWorks SAN technology.

Advantages of Dual Fabric SANs

Most large SANs should have two independent fabrics. Each fabric operates independently, and the failure of one fabric does not cause a complete loss of SAN communication.

The reliability of modern electronic hardware is so high that it is difficult to make meaningful predictions of failure rates. Software is used in all components, but it is difficult to estimate the likelihood of software failures. Operator errors are the most likely cause of problems, and the frequency of operator errors depends strongly on operational discipline and employee morale, both of which are very difficult to quantify. All of these potential failure points are minimized by the use of multiple fabrics.

The advantage of dual fabric designs is that they support path failover technology. Path failover is available in most operating systems that are supported in HP SANs. Two host bus adapters are used in each server, and if the communication path from one HBA to the storage system fails, then the I/O traffic is re-routed through the other HBA.[1]

The two fabrics should be similar in size and topology. This minimizes the risk of asymmetrical performance under certain workloads, and minimizes the total cost of the SAN. Failover software does not support the concept of primary and secondary fabrics.

It should be noted that there is not an automatic increase in cost caused by the use of two separate fabrics. For example, two switches in a single fabric give about two dozen usable ports (depending on the topology). Two separate fabrics, each with a single switch, gives 32 ports at the same cost.

Many of the SAN illustrations in this document show only a single fabric. This is because most of the design and compatibility requirements apply to each fabric as a complete unit. However, practical SAN designs should have two or more fabrics, each satisfying the configuration rules described in this guide.

Data Access Patterns

There are several supported HP SAN topologies, suitable for a wide range of applications from small to very large systems. For small installations, the topology may be chosen to maximize connectivity or to minimize cost. SAN performance is not likely to be an issue for a small installation, because of the very high I/O throughput that is provided by basic Fibre Channel SAN components.

1. Failover can also be useful in SANs with only one fabric. This protects against HBA failures and certain extremely unlikely potential problems in array controllers. In general, failover technology should be used in SAN configurations that have two fabrics.

Large installations must be designed to maximize performance and minimize cost, to support current and future connectivity requirements, and to enable eventual migration to new technologies. Several factors must be taken into consideration to meet these requirements. The factors are categorized into three different data access patterns, one-to-one, many-to-one, and any-to-any.

- **One-to-one**

  The communication paths within the fabric are used in different ways, depending on the relationship between the servers and the storage systems. In some cases, each specific server stores data on only one or two storage systems. In this case, only a few specific storage systems service all I/O requests from a server, and there is little or no communication between the servers or between the storage systems. A given fabric port sends requests to one (or two) specific fabric ports. This is the traditional server-storage relationship. Many systems still operate this way today.

  From the viewpoint of the fabric, the I/O traffic has a "one-to-one" pattern, and the traffic pattern is stable. Each server sends I/Os to a small, specific set of storage systems, and each storage system is associated with only a handful of servers. Only significant changes to the configuration by the system manager will change the connection pattern.

- **Many-to-one**

  Multiple servers accessing data stored in a single centralized pool is another data access pattern. This is a common situation when high performance storage systems have enough capacity to handle a number of servers. In this environment, there is a "many-to-one" I/O traffic pattern on the SAN fabric, and the traffic pattern is stable. Each server sends I/O requests to a small set of storage systems, but each storage system may service a large number of servers. The connection pattern changes only when significant changes to the configuration are made by the system manager.

- **Any-to-any (or many-to-many)**

  In a third case, application servers access data that is distributed across many storage systems. This case may develop in several situations. The latest HP storage arrays may handle a large number of servers. (Refer to the configuration rules in this Guide for detailed information.) A system manager may decide to distribute information over a wide set of storage systems, thus requiring each application to access multiple storage systems. This situation can arise when host-based mirroring is used. Another possibility is that it may be easier to manage the data if it is partitioned and stored on multiple storage systems. For example, Accounting Department data might be stored on one storage system, and Personnel Records data on another. A server requiring access to both data types generates I/O requests to both storage systems.

  Another important situation where data is distributed across a range of storage systems is when the HP VersaStor virtualization technology is used. VersaStor distributes data over all the available storage systems in a SAN.[2] In this case, I/O requests from a given application server are handled by one or more storage systems, in a pattern that is controlled by the virtualization management appliance. In this environment, many servers access many storage systems, which is a "many-to-many" pattern. Management traffic may occur between servers, storage systems, and management appliances.

  From the viewpoint of the SAN fabric, any port may send traffic to any other port, which is an "any-to-any" pattern. Furthermore, since the virtualization manager performs dynamic reallocation of storage system capacity, the traffic patterns vary continuously without manual intervention.

2. The specific configuration details are controlled by management options.

The optimum SAN configuration depends on the I/O traffic, whether it be one-to-one, many-to-one, or any-to-any pattern.

Core and Edge Switch Concept

In the future, most large SANs will support any-to-any traffic patterns. The remainder of this chapter focuses on this problem.

The optimum fabric configuration uses a high performance “core” surrounded by a number of “edge switches.” The core provides roughly equal connection performance between any pair of ports. The edge switches provide port aggregation to match the performance requirements of the servers and storage systems to the performance of the core.

The figure below shows a large configuration that uses the core and edge switch approach. Using 16-port switches, the core is a 32 port fat tree. Four ISLs go between each switch pair in the fat tree. Two ISLs connect each edge switch to the core.

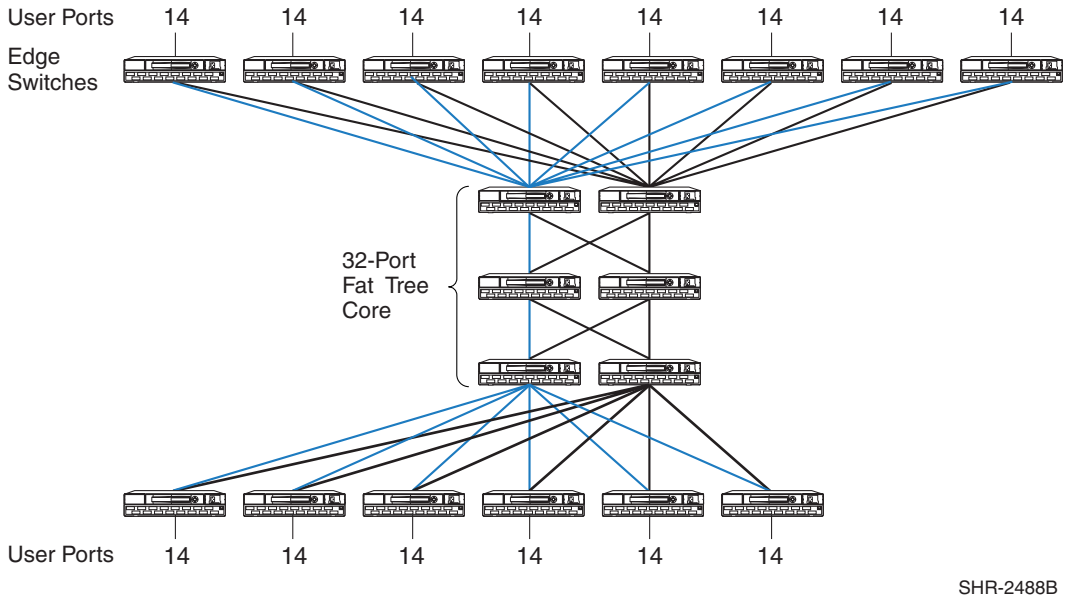

Figure 59: Example of Core Switch Plus Edge Switch Configuration

Fabric Core Options

The simplest fabric core is a single switch. Fibre Channel switches support simultaneous full bandwidth connections between any combination of port pairs. A single switch fabric core guarantees support for any-to-any traffic.

Any combination of switches has less performance than a single switch, and the difference depends on the fabric topology. The best-performing topology is the “fat tree”, which has enough Inter-Switch Links (ISLs) to provide, on the average, full bandwidth connections between any combination of port pairs. While it is possible to construct workloads that force traffic contention on the ISLs of a fat tree, which reduces the throughput, fat tree fabric core topologies provide full-bandwidth any-to-any communication, on the average, for random traffic patterns.

A related topology is the “skinny tree”, which has fewer ISLs and fewer switches. This topology introduces an unavoidable performance limit to the fabric. In many cases this limit is beyond what is required by the application servers. The process to upgrade a skinny tree topology to a fat tree topology is fairly straightforward, involving the addition of switches and ISLs to the existing tree

Edge Switch Options

The simplest edge switch is a single switch with one ISL connecting it to the fabric core. Each edge switch provides “User Ports” for connecting servers and storage systems.

The single ISL is a potential bottleneck. All the I/O traffic from the servers or storage systems connected to the edge switch must pass through just one ISL. More ISLs can be provided. Several combinations of ISL and user ports may be used. For example, with sixteen port switches, the ISL to user port ratio could be 1:15, 2:14, 3:13, 4:12, etc. Each of these combinations represents a “port aggregation ratio.” The ratios are 1:15, 1:7, 3:13, 1:3, etc.

The workload of the servers and storage systems attached to an edge switch determines the required port aggregation ratio for the switch. For lightly loaded application servers, a 1:15 port aggregation ratio may be adequate. Heavily loaded servers may require a 1:7 or 1:3 ratio. Extremely high performance servers, such as high-end HP Alpha systems, may be able to completely “fill up” a Fibre Channel connection. In this case, there is no advantage to using an edge switch, and the server should be connected directly to the fabric core. Storage systems may also be able to support a full bandwidth Fibre Channel connection.

To select the appropriate port aggregation ratio, refer to the I/O requirements of your applications and servers. This information is available for many situations by using the Active Answers application sizing tools. In other cases, measurements of an existing system may be required to determine the workload.

Designing a Subsettable SAN

In many cases, the growth pattern for a storage installation is difficult or impossible to predict. Global economic growth, conditions in a given business market, the growth rate of your company, and internal reorganizations or reallocations of computing resources may all have a significant impact on the requirements that must be met by the SAN.

To accommodate this unpredictable variability, the SAN designer should plan for growth within a predefined design. The initial installation should be a subset of a larger pre-designed configuration.

The “core plus edge switch” approach supports this strategy for SAN design.

When the time comes to expand an existing installation, the system manager can make incremental changes to the configuration rather than a complete reconfiguration of the entire Fibre Channel fabric. Changes to the fabric core are isolated from the edge switches, which minimizes the impact of changes required to support core growth. Changes to a given server’s connection to an edge switch are isolated from the core, which minimizes the impact of server-related changes. Furthermore, since two or more fabrics are in use, server I/O traffic may be temporarily forced to a single fabric while the other fabric is undergoing modification.

Start with a single switch core for a moderate sized initial installation,. When needed, the core can be expanded by replacing the switch with one that has more ports, or by reconfiguring the core to a skinny tree or fat tree topology. An existing fat tree core may be expanded by replacing it with a fat tree made up of switches with more ports, or by reconfiguring it to a wider fat tree configuration.

Use a generous estimate of the required I/O performance when selecting edge switches. A port aggregation ratio of 1:7 or 1:3 is adequate for most applications. Increasing bandwidth is a simple, localized modification, if it turns out that more is required.

The initial design should include spare ports on the core to support the future addition of edge switches. For example, consider a configuration that uses sixteen port switches, a single switch core, and edge switches with a port aggregation ratio is 1:3. This design supports up to four edge switches and 48 user ports. This would be a suitable solution for a system where 36 ports are required now, requiring three edge switches. Future growth to 48 ports can be accommodated by adding another edge switch.

SAN Design Summary of Recommendations

Large SANs should include the following features.

- Multiple independent fabrics.
- Core plus edge switch topology.
- Appropriate port aggregation ratio, depending on application server requirements.
- Appropriate core design, depending on number of ports required.
- Subsettable design, with initial installation suitable for current needs.

By following these guidelines for SAN planning, your design will be suitable for supporting future storage technology and future growth in your storage environment.

Configuring a SAN

Once you have completed the planning phase you can begin to configure your SAN. As described in the planning phase, it is important that you document the configuration. During the configuration phase, you should be recording the details of the actual physical configuration.

- Recording. As you construct the SAN, record the cable connections and mark this information on the configuration layout diagram. Record the WWN of all nodes and devices and identify where they physically reside. It is recommended that you place a label on each Fibre Channel HBA with the WWN clearly identified. HP storage systems are pre-labeled with this information; however, you may wish to place an additional label on the front of the unit in plain view.
- Cabling. Define a system for cable labeling. Even a small SAN can include a very high number of fiber optic interconnect cables. Label both ends of each cable with the same unique cable number or color code scheme. This will allow you to quickly identify each cable uniquely. Also consider placing a label at each end of the cables that identifies connection points at both ends, such as “TO” and “FROM”. Use label types that are easy to create and read, and ensure they are attached securely to the cable.
- Protect unused or open switch ports with port plugs. Never leave ports exposed.
- Cable Dressing. Use care when routing fiber optic cable and ensure that you do not exceed the recommended minimum bend radius. For single-mode and multi-mode fiber cable the minimum bend radius is 25 mm. Where cables are bundled or hanging unsupported, use velcro tie wraps to group and support the cables. Never use plastic tie wraps as they can damage the internal fiber core if over-tightened.
- Cable Symmetry. When connecting cables, consider slot/port-numbering symmetry. Be consistent across similar servers with cabling in terms of HBA slot placement and cabling to switches. If configuring with two SAN fabrics and multi-pathing, connect HBA 1 to SAN fabric 1, HBA 2 to SAN fabric 2, etc. Cable symmetry is not a requirement, but serves as an aid to troubleshooting if this is eventually required.
- Configure Fibre Channel Switches. Although all HP Fibre Channel switches are pre-configured, verify that all Fibre Channel switches in the fabric have the same parameter settings and that each has a unique domain ID.

  Label switches using a relevant naming scheme particular to the topology. For example, if implementing a ring topology, label each switch in the ring as Ring1, Ring2. Although not an absolute requirement in all configurations, it is highly recommended that all switches utilize the same switch firmware revision. Different switch code revisions running in the same fabric are supported during a rolling upgrade. This is considered a temporarily acceptable situation for the duration of the code update.
- Configure Servers. For each platform or operating system type, utilize the appropriate HP StorageWorks platform kit to ensure that the required server drivers and configuration settings are loaded. Ensure that servers are configured with the proper operating system versions and all required updates.

  Use a numbering type scheme for naming multiple servers of the same type, such as NT01 and NT02 for Windows NT servers.
- Configure Storage. Use the storage map created in the planning phase to configure each of the storage systems. Verify server-to-storage connectivity, and access one server at a time.

When initially defining storagesets, always disable all access first, and then enable the desired individual access. For Enterprise/Modular RAID Array storage systems, define connection names to be consistent with zoning alias names. Be consistent with connection names relative to storage port and controller connection. Choose a scheme that is easily understood and quickly conveys the physical connectivity.

- Define Zones. Use the zoning map to configure zones. Consider starting with small zones that allow a smaller logical subset of a larger physical SAN to be tested initially.

  Always save old zoning configurations before and after making any zoning change. If possible, it is recommended that no zoning changes be made when an individual switch normally configured in the fabric is temporarily not available.

  You can zone by operating system or by storage system. Zoning by operating systems is useful when the operating systems are accessing storagesets that are localized to specific raid arrays. For example, NT1, NT2 and NT3 have access to storage on ARRAY1, and VMS1, VMS2 and VMS3 have access to storage on ARRAY2.

ZONE NAME	**NT_ZONE**	**VMS_ZONE**
Members	NT1	VMS1
	NT2	VMS2
	NT3	VMS3
	ARRAY1	ARRAY2

  ARRAY1 will only have host connections for the NT1, NT2 and NT3 servers and ARRAY2 will only have host connections for the VMS1, VMS2 and VMS3 servers.

  Zoning by storage system will limit the connections to the G80 to those systems actually having storagesets on them. This is useful when the storagesets for a specific system are on multiple storage systems.

  In the above example, we add 3 more NT servers and another storage system to the NT zone:

ZONE NAME	**NT_ZONE**	**VMS_ZONE**
Members	NT1	VMS1
	NT2	VMS2
	NT3	VMS3
	ARRAY1	ARRAY2
	NT4	
	NT5	
	NT6	
	ARRAY3	

  Both Array1 and Array2 will have host connections from all 6 NT systems. This may not be a problem in a small SAN, but as the SAN grows the connections will increase. Also, we do not know which of the NT servers are accessing storage on ARRAY1, and which ones are accessing storage on ARRAY2.

  If we zone by storage system we get:

ZONE NAME	**ARRAY1_ZONE**	**ARRAY3_ZONE**	**ARRAY2_ZONE**
Members	NT1	NT4	VMS1
	NT2	NT5	VMS2
	NT3	NT6	VMS3
	ARRAY1	ARRAY3	ARRAY2

  Zoning this way also makes it much easier to troubleshoot, especially if servers access storage on multiple arrays. We could have a zone that looks like this:

ARRAY1_ZONE	ARRAY3_ZONE	ARRAY2_ZONE
ARRAY1	ARRAY3	ARRAY2
NT1	NT1	NT4
NT2	VMS2	NT5
VMS2	VMS3	VMS1
VMS3	NT5	NT2
NT6	NT6	NT6

This way it is more apparent that NT1 is only accessing storage on ARRAY1 and ARRAY3. If part of storage can not be seen then it is easy to locate the source of the problem.

Due to some zoning restrictions, you may need more than one zone for a particular ARRAY. If ARRAY1 also has IBM AIX servers, we must zone that separately.

ARRAY1_ZONE1

ARRAY1

AIX_1

AIX_2

Zone and Zone Alias Names

When setting up zoning, use meaningful names for zones and zone aliases and be consistent with the naming convention throughout the fabric.

Servers are identified by the WWN of the host bus adapter. Name these by using the system name and the host bus adapter number. For example, server NT1 with one Fibre Channel HBA would have an alias of NT1_HBA1. Server NT1 with a second HBA would have an alias of NT1_HBA2

RA8000 storage systems in a transparent failover configuration will have two WWN's on the fabric, one for port 1 and one for port 2. Give each RA8000 a unique number. RA8000 number 1 could have aliases of R1_P1 (port 1) and R1_P2 (port 2)

For a multiple-bus failover configuration the RA8000 will present 4 WWNS to the fabric. If you have a multi-path NSPOF configuration, two of the WWN's will be in one fabric, the other two will be in the second fabric. Name the ports using an alias such as R2_A1 (Controller A Port 1), R2_A2 (Controller A Port 2), R2_B1 (Controller B Port1), and R2_B2 (Controller B Port 2).

Ports A1 and B2 will be cabled to the first fabric. Ports A2 and B1 will be cabled to the second fabric. The aliases in fabric 1 will be R1_A1 and R1_B2, the aliases in the second fabric will be R1_A2 and R1_B1. Keep the ports and HBAs the same throughout the setup. For example, always have HBA 1, R1_A1 and R1_B2 in fabric1 and HBA 2, R1_A2 and R1_B1 in the second fabric.

Using this convention conveys the failover mode that the RA8000 is configured for. Any alias with a P1 or P2 is in transparent mode, any alias with A1, A2, B1, or B2 is in multiple-bus mode.

Define RA8000 host connection names for the adapter WWN's in the same manner as you defined the alias name in the fabric. For example, the fabric alias name for NT1, HBA1 will be NT1_HBA1. The host connections on the RA8000 controller should match this as closely as possible.

Example:

Alias NT1_HBA1 in the fabric would have host connection names on the RA8000 of:

```
NT1-P1  WINNT   THIS   1      081200 OL this  30
HOST_ID=2000-0000-C922-8ADC     ADAPTER_ID=1000-0000-C922-8ADC

NT1-P2  WINNT   OTHER  2      081200 OL other 130
HOST_ID=2000-0000-C922-8ADC    ADAPTER_ID=1000-0000-C922-8ADC
```

Note: While storage system connection names are not case sensitive, switch alias names are. That means that the switch might have a alias name of TRU64_1 and another alias name of Tru64_1 that refer to two different sets of things.

Upgrading a SAN

Upgrading a Fibre Channel Switch

See the Installation and Hardware Guide for your switch.

Scaling a SAN

The information in this section applies to all SAN topologies, whether a custom design or HP defined.

- Replace 8-port switches with 16-port switches.
- Add additional switches, up to the limits specified for a single fabric in Chapter 3, "SAN Fabric Design Rules".
- Add a second fabric as a high availability no single point of failure solution.
- Deploy multiple independent SANs.
- Migrate to a different topology (see below).

Scaling Specific SAN Topologies

The information in this section is specific to the HP-defined topologies. Refer to the Fibre Channel switch replacement procedure elsewhere in this chapter for information about preventing fabric segmentation when adding new switches to an existing fabric.

Whenever you are expanding a topology, ensure that the new switch and device connectivity is consistent with the original SAN topology design requirements and goals. Avoid making changes to the topology that may serve to disrupt the original topology design goals. If you need to make topology changes based on a change in data access requirements, consider migrating to a different topology that is better suited to meet these needs. It is important in any expansion that the original data access needs be maintained.

If you have implemented a high availability fabric design (refer to Chapter 2, "SAN Topologies"), it may be possible to expand your SAN in a non-disruptive manner. It is highly recommended, however, as a precaution, that all data be backed up and that I/O activity quiesed when adding new switches to the fabric.

Cascaded Fabric

Expand an existing cascaded fabric by connecting a new switch to an available port on an existing switch. If there are no available ports, remove a device or set of devices from an existing switch, connect the new switch to those ports, and connect the device or devices to the new switch.

Meshed Fabric

Expand an existing meshed fabric by connecting a new switch to available ports on an existing switch. If there are no available ports, remove a device or set of devices from an existing switch, connect the new switch to those ports, and connect the device or devices to the new switch. To maintain the meshed topology, you must ensure that there are multiple paths (ISLs) connecting the new switch to the existing meshed fabric.

Ring Fabric

Expand an existing ring fabric by breaking the ring and inserting another switch into the ring.

Add new switches cascaded off of the ring, up to the maximum number of switches supported in a single fabric. When expanding outside of the ring, ensure that no two devices that need to communicate are more then seven hops apart.

Tree Backbone Fabric

Add edge switches. Expand an existing Tree Backbone SAN fabric by adding additional edge switches. Connect these edge switches to available ports on the one or two backbone switches.

Add a second backbone switch (if your current design only contains one). Connect all of the edge switches to the new backbone switch.

Migrating SAN Topologies

This section describes how you can convert from one topology type to another if required. HP highly recommends that you thoroughly review your initial design to ensure that it meets your present and future requirements in order to avoid having to modify your initial topology design. There may be situations, however, based on changes in business requirements, that require you to consider converting to another topology type. For those circumstances, information is provided below that can help you gain an understanding of how the different topologies can be converted.

As described in the planning phase, it is important that the SAN fabric topology be well documented. If you are required to change from one topology type to another, use the existing topology diagrams to determine the most efficient manner in which to modify the topology. Create a new diagram that details the desired final connectivity scheme and use this as a map for the topology migration or conversion.

If you have implemented a high availability fabric design, depending on the specific cabling changes required, it may be possible to migrate your SAN in a non-disruptive manner. It is highly recommended, however, as a precaution, that all data be backed up and that I/O activity be quiesced when migrating or reconfiguring any portions of the fabric.

If you have implemented a two-fabric, no single point of failure (NSPOF) SAN, you have the ability to failover over all operations to one fabric while you reconfigure the other fabric. This makes it possible to perform a totally non-disruptive topology migration.

- As a general rule, migrations that only require the addition or re-cabling of ISLs are less disruptive then migrations that require devices be moved from one switch to another. When planning a migration, try to avoid or minimize scenarios that require moving devices from one switch to another.
- Cascaded to a Meshed Fabric. Whether you have implemented a linear cascade or branched cascade of switches from one top switch, additional ISLs are required to connect all switches together as required in a mesh fabric design. Proper planning requires that you carefully calculate the number of additional ports that are needed for the additional ISLs. This may require that devices be moved from one switch to another.
- Cascaded to Ring Fabric. If you have implemented a linear cascade, connect the last switch in the cascade to the first switch to create a ring fabric. For a branched cascade, extensive ISL re-cabling may be required.
- Cascade to Tree Backbone Fabric. Whether you have implemented a linear cascade or branched cascade, determine which switch or switches will be utilized as the backbone (typically no devices) and which switches will be edge switches. Connect all edge switches to the backbone switch or switches; connect all devices to the edge switches. This conversion is less disruptive if you add new switches to the fabric for the backbone and use all of the existing switches as edge switches. In this case, you can simply connect one end of the existing ISLs to the new backbone switches.
- Meshed to Ring Fabric. A meshed fabric can be converted to a ring fabric by simply removing the cross-connected ISLs, leaving the outer connected ISLs connected as a ring. The available ports can be utilized as additional redundant ring ISLs or for additional devices.
- Meshed to Tree Backbone Fabric. Determine which switch or switches will be utilized as the backbone (typically no devices) and which switches will be edge switches. Connect all edge switches to the backbone switch or switches; connect all devices to the edge switches. This conversion is less disruptive if you add new switches to the fabric for the backbone and use all of the existing switches as edge switches. In this case, you can simply connect one end of the existing ISLs to the new backbone switches.

- Ring to Meshed Fabric. If you have implemented two ISLs between all switches in the ring, move one end from an ISL between any two switches to the appropriate switch based on the final mesh design. Repeat this for all of the second ISLs between any two switches. There may be an optimal place to "break" the ring relative to re-cabling. Evaluate different scenarios prior to performing the actual conversion.
- Ring to Tree Backbone Fabric. Determine which switch or switches will be utilized as the backbone (typically no devices) and which switches will be edge switches. Connect all edge switches to the backbone switch or switches; connect all devices to the edge switches. This conversion is less disruptive if you add new switches to the fabric for the backbone and use all of the existing switches as edge switches. In this case, you can simply connect one end of the existing ISLs to the new backbone switches. It is also less disruptive if you have implemented 2 ISLs between all switches in the ring in your original design.

Zoning Rules and Guidelines

Configure the zones based on the 'zoning map' prepared during the SAN planning stage. There are several possible supported ways to configure zones and let us examine briefly these before looking at suggested guidelines for each of the product series.

First of all, one need to understand the distinction between configuring the zones and zoning enforcement within the switches and any correlation between the two. This may vary from product to product and hence the following paragraphs describe some general description about each of them, followed by specific details on all HP supported switches.

Note: When enabling a new configuration, it is strongly recommended that the fabric be quiesced. Zone membership should not be changed for devices that are actively performing I/O in a fabric. Once the new zoning is enabled, a state change notification is sent to all the nodes that have registered to receive the state change.

Zoning enforcement

Generally there are three types of zoning enforcement/authorization techniques in use in FC switches today.

Access Authorization

Access authorization provides frame level access control in hardware and verifies SID-DID combination of each frame and allows the frame to be delivered to the destination if that is a valid combination per the zone definition. This is definitely more secure and generally referred to as "hard zoning", and requires hardware resources at the ASIC level to implement this.

Discovery authentication

Zoning enforcement or protection of unauthorized access is only provided during access to the Name Service directory where the switch or fabric presents only a partial list of devices from the NS directory corresponding to the partition in which the requesting device is part of. This type of zoning enforcement is generally referred to as "soft zoning". While this is secure enough in most of the cases, it is prone to security threats if malicious hosts attempt to access unauthorized devices violating FC-Protocols.

Login Authentication

Some switches enforce authentication during fc protocol login frame level like PLOGI/ACC/ADISC/PDISC etc, in addition to providing discovery authentication. For example, if a host sends a PLOGI to a device that is not part of its zone, this frame gets dropped before reaching the destination. This type of enforcement has some additional protection compared to discovery authentication, but still not the same as access authorization at every frame level.

Zoning Configuration

Domain/port numbers

Zones can be defined using switch domain ID and port number combination to uniquely identify zone members. Advantage of this type is primarily ease of configuration and zoning definition remains intact even when an HBA or target controller is replaced with another having a different "world wide name". Disadvantage is that there is no flexibility to move around devices in the fabric and as soon as any device is moved to a different port in the switch/fabric it will not be part of the zone any more.

WWN

Zones can be defined using device WWN to uniquely identify zone members. Advantage of this type of definition is zoning definition remains intact even when the device is moved to a different port/switch in the fabric. Disadvantage is that whenever an HBA is replaced with another, having a different WWN, zoning definition has to be changed accordingly.

Mixture of both

Zones can be defined using combination of switch domain ID /port number as well as WWNs to uniquely identify zone members. Advantages and disadvantages as described in the above two methods are applicable to individual zone elements based on their definition.

Note: Movement of devices within the fabric, as described above depending on zoning definition, is only applicable from a zoning perspective. However, there can be other restrictions that will not let movement of devices within the fabric, irrespective of zoning type in effect. For example, some OSs like HP-UX which create device filenames based on 24-bit fabric address will not allow moving the device to a different port since it will change the 24-bit port address and hence will be treated as a different device.

There should not be any dependency between the way zones are configured and the way zones are enforced -meaning it should be possible to have any combination of zoning configuration/zoning enforcement from the above definitions. Due to implementation limits certain switch products impose restrictions the way zones are defined and the way zones are enforced.

The following tables and paragraphs detail how zoning is implemented followed by suggested guidelines for HP supported switches.

B-Series Product Line Switches

Table 44: Zone Types on HP B-Series Product Line Switches

Switch Models	Configuration	Enforcement	Comments
CPQ StorageWorks SAN Switch-8, SAN Switch 8-EL; SAN Switch-16 SAN Switch 16-EL; SAN Switch Integrated/32, SAN Switch Integrated/64 (FC-6164) (All 1Gb switches)	Define zones using all domain#,port# Define zones using only WWNs Define zones using combination of domain/port numbers and WWNs	Access authorization at frame level in hardware Discovery authentication Name Servers (NS) directory based Discovery based authentication	HARD zoning SOFT zoning SOFT zoning
CPQ StorageWorks SAN switch 2/8-EL, SAN switch 2/16-EL, **HP reseller** FC-8B, FC-16B(SAN switch 2/16); **HP StorageWorks** Core switch 2/64; SAN switch 2/32 (All 2Gb switches)	Define zones using all domain#,port# Define zones using only WWNs Define zones using combination of domain/port numbers and WWNs	Access authorization at frame level in hardware Access authorization at frame level in hardware Name service plus login authentication	HARD zoning HARD zoning SOFT+ (NS authentication plus login protection)
QuickLoop Zoning (all QL supported switch models)	Define zones using ALPAs, Domain/port numbers or combination of the above	Implemented in hardware tables, access prevented by hardware between unauthorized devices	HARD zoning

Maximum Zone Size

Generally the supportable 'maximum number of zones' and 'maximum members in a zone' are very large and are usually constrained by memory usage. These numbers are far larger than the maximum devices that could be connected in fabric configurations currently supported and hence usually there are no limitations on zone sizes.

However, there is an exception in pure hardware enforced zoning environment on all the above 2Gb switch models where it's likely that we exceed some preset architectural limits in which case those ports transition from HARD enforcement to SOFT type.

The current B-Series switches have a limitation of 64 unique SID entries per quad (pre-defined groups of 4 ports) and whenever this limit is exceeded the affected port/ports will transition from hard to soft enforcement.

This transition is completely transparent to fabric operations, though switch administrator may see warning messages displayed in switch logs. However, data integrity is completely preserved during this transition and HP validated this in large SAN configurations.

The following CLI output indicates a port transitioning to soft zoning:

```
WARNING ZONE-ZONEGROUPADDFAIL, 3, WARNING - port 7 Out of CAM
entries

WARNING ZONE-SOFTZONING, 3, WARNING - port 7: zoning enforcement
changed to SOFT
```

These two messages are related and indicate that the zoning configuration has outgrown internally preset architectural limits, thereby forcing the mentioned port be switched from hardware-enforced zoning to software-enforced zoning. It is important to note that only this specific port has turned "soft" and all other members that were zoned with the relevant port still remain hardware-enforced.These warning messages could be seen either statically at zoning configuration/setup time (in case if port-level zoning) or dynamically at run time (in case of WWN zoning).

The command "portzoneshow" will display the status of all ports as follows:

Hard - hardware enforcement

Soft - Name Server plus ASIC assisted authentication

All - no zoning enforcement

Zoning Guidelines (B-Series switches)

The following are suggested best practices only. However, other zoning configuration methods and zone types as appropriate for each switch are also supported.

- Define zones using WWNs always. All switch models support this type of definition, irrespective of zoning enforcement technique they use, whether it is hardware enforced or name server based or combination of both. Use port WWNs and not node WWNs.
- Exception: For all 1Gb fabric switches, define zones using domain/port numbers for selecting hardware enforced zoning
- Define zones for all devices in a fabric whenever any zone is defined. In other words do not define zoning partially for few devices in the fabric and leave others un-zoned.
- Overlapped zones can be defined and there is no upper limit on the number of zones and number of members in a zone
- Configure zones based on operating environment, on a “per OS” basis,
  See the SAN/Platform zoning requirements for individual storage arrays for exact details, as defined in Chapter 4.
- Switch zoning provides security at the port level only and for maximum security in a SAN environment, it's required to use array based LUN security- Secure Manager for XP/VA arrays and SSP (Selective Storage Presentation) for HSG/HSV array controllers.
- To minimize/avoid "soft" port transition/s in pure hardware enforced zoning environment(2Gb SAN fabric switches)
- Maintain locality as defined in your SAN design but avoid hosts/targets on the same quad. Quad is a group of pre-defined consecutive 4 ports (0-3,4-7,8-11,12-15 etc).
- Maintain a connectivity model that populates each quad with the members of the same zone or in other words avoid members of different zones on the same quad particularly when each of them are part of bigger zones. For example, if we have an UNIX zone and an WINDOWS zone, populate all UNIX zone members on one quad and WINDOWS members on a different quad.
- Minimize zone entries by including hosts and targets that practically need to talk to each other. For example, instead of combining all hosts of the same OS type into one zone, consider making smaller zones with only hosts and targets that need to talk to each other.

- Switch CLI command "portzoneshow" can be used to display and verify the individual status of each port whether it's "hard" or "soft" at any given time.

M-Series Product Line Switches

Table 45: Zone Types on M-Series Product Line Switches

Switch Models	Configuration	Enforcement	Comments
HP Surestore FC-64 CPQ StorageWorks SAN Director 64 1Gb director class switches	Define zones using all domain#,port#	Discovery authentication Name Servers (NS) directory based	SOFT zoning
	Define zones using only WWNs	Discovery authentication Name Servers (NS) directory based	SOFT zoning
	Define zones using combination of domain/port numbers and WWNs	Discovery authentication Name Server (NS) directory based	SOFT zoning
HP StorageWorks Edge Switch 2/16 Edge Switch 2/24 Edge Switch 2/32 Director 2/64 Director 2/140 (All 2Gb switches)	Define zones using all domain#,port#	Discovery authentication Name Server (NS) directory based	SOFT zoning
	Define zones using only WWNs	Discovery authentication Name Server (NS) directory based	SOFT zoning
	Define zones using combination of domain/port numbers and WWNs	Discovery authentication Name Server (NS) directory based	SOFT zoning

Note: In Open Fabric mode (which is the default mode), director/edge switches allow only WWN based zoning configuration.

Look for HARD zoning support in future versions of firmware for 2Gb products - hardware has built in support already.

Maximum Zone Size

The supportable 'maximum number of zones' and 'maximum members in a zone' are very large and are usually constrained by memory usage. These numbers are far larger than the maximum devices that could be connected in fabric configurations currently supported and hence there are no limitations on zone sizes.

Zoning Guidelines (M-Series switches)

The following are suggested best practices only. However, other zoning configuration methods and zone types as appropriate for each switch are also supported.

- Define zones using WWNs always. All switch models support this type of definition. Use port WWNs and not node WWNs.

- Define zones for all devices in a fabric whenever any zone is defined. In other words do not define zoning partially for few devices in the fabric and leave others un-zoned.
- Overlapped zones can be defined and there is no upper limit on the number of zones and number of members in a zone
- Configure zones based on operating environment, on a “per OS” basis.
  See the SAN/Platform zoning requirements for individual storage arrays for exact details, as defined in Chapter 4.
- Switch zoning provides security at the port level only and for maximum security in a SAN environment, it's required to use array based LUN security- Secure Manager for XP/VA arrays and SSP (Selective Storage Presentation) for HSG/HSV array controllers.

C-Series Product Line Switches

Table 46: Zone Types on HP C-Series Product Line Switches

Switch Models	Configuration	Enforcement	Comments
■ Cisco MDS 9509 Multilayer Director Switch ■ Cisco MDS 9216 Multilayer Fabric Switch	Define zones using all domain#,port# Define zones using only WWNs Define zones using combination of domain/port numbers and WWNs	Access authorization at frame level in hardware	HARD zoning

Special considerations in zoning (for all switch models)

- In high availability environments like HP-UX service guard, it is required to have homogeneous OS environments on a storage array port and this can be achieved by securing LUNs using array secure manager software and also by properly configuring zones.
- Software environments like OVSAM and CommandView for XP/VA do not impose any restrictions on switch zoning. The same supportability exists in these environments as well.
- For SANS with Data Protection (tape back up) might require separate rules and refer to "New HP SAN Backup Support Guide - Version 4.0, October 2002: (PDF)" on SPOCK at the following URL

  http://hpso.rose.hp.com/spock/documents/backupSAN_configGuide_v4.pdf

Merging SAN Fabrics

This section describes the process for merging two (or more) independent fabrics into a single, larger fabric. This is typically done when you:

- have grown independent SAN islands to the point where more resources are needed
- wish to share the resources in two or more fabrics
- wish to make information in one SAN available to servers in another SAN

With support for longer distances you may also desire to connect geographically separated SAN islands together into a single SAN, spanning across very long distances.

Although StorageWorks SAN designs and components allow versatile configurations, HP highly recommends that you thoroughly review all SANs to ensure they will meet existing SAN rules after they are merged into a single fabric. The newly created fabric should not exceed any existing SAN rules.

Merging fabrics can be a complicated process, especially if the fabrics are large. The procedures in the document require a complete understanding of fabrics, zoning commands, and rules. They also require that the user understand how to use the telnet commands as well as the web-based GUI.

It is important to consider not only current SAN configurations but any future SAN needs that may be required. Most difficulties related to merging SANs are due to the fact that not enough planning was put into future SAN considerations at the time the initial SAN was designed and built. Another problem is that the SANs being merged may be implemented differently.

When fabrics discover each other they must go through basic login procedures, or sanity checks, to determine if they are compatible to work as one fabric. If the discovery process determines they are not compatible then the fabric will segment. This means that although they are physically connected, they will still run as separate fabrics.

When zoned fabrics merge they append their zone configuration database to include each fabric's zone configurations. If a non-zoned fabric merges with a zoned fabric, all zoning information is proliferated to the non-zoned fabric switches. If there was a zone configuration enabled at the time of the merge, then that zone configuration will be enabled on the non-zoned fabric switches as well. This means that any devices that were in the non-zoned fabric will be not accessible until they are added into the current enabled configuration.

Please review these causes of SAN segmenting prior to physically connecting multiple fabrics together.

- *The name of a zone object in one fabric should not be used for a different type of zone object in the other fabric (Zone type mismatch).* In other words, if you create a zone name on Fabric A, that same name should not be an alias or configuration name in Fabric B; otherwise the fabrics will not merge.
- *The definition of a zone object in one fabric is different from its definition in the other fabric (Zone content mismatch).* If an alias, zone or configuration name is the same on both Fabric A and B but the content or definition of that object is different between the fabrics the fabrics will not merge.
- *Zoning is enabled in both fabrics and the zone configurations that are enabled are different (Zone configuration mismatch).* Because of this mismatch the switches within each fabric are not going to assume one fabric has the correct zone configuration enabled. The fabrics will not merge until one of the merging fabrics has its zone configuration disabled.

- *Not only must each switch within a fabric have a unique domain ID but each switch within the multiple fabrics of the enterprise should have a unique id as well.* For example, If Fabric A has five switches with domain IDs 1 through 5 and Fabric B has five switches with the same domain IDs these two fabrics will not merge until all switches within both fabrics have a unique domain ID.

Note: If you use port level zoning, changing the domain ID's may affect access to devices. Port level zones are based on the domain ID and the port number.

Note: When enabling a new configuration, it is strongly recommended that the fabric be quiesced. Zone membership should not be changed for devices that are actively performing I/O in a fabric. Once the new zoning is enabled, a state change notification is sent to all the nodes that have registered to receive the state change.

Merging fabric together can be accomplished by simply disabling the effective configuration on one fabric, then plugging both fabrics together. The problem with this method is that once you disable the effective configuration, you open up that fabric so all servers will see all storage. Also once you plug the fabrics together, devices from the second fabric will not be accessible until you add them into the effective configuration.

To merge these two fabrics without having to disable the effective configuration for the entire fabric, it is necessary to disable at least one switch in each fabric or have a spare switch available. This will be the switch used for merging the zones and creating the new configuration. Keep in mind that there can be multiple defined configurations, but only one can be the effective or enabled configuration.

Troubleshooting

The following section describes troubleshooting steps for isolating problems related to storage access. When initially building a SAN, lack of access either to individual storagesets or entire storage systems is not uncommon. This can usually be traced to an incorrect device setting or an inadvertent cabling or configuration setup error in the initial hardware configuration. The steps listed will assist you in isolating access problems.

1. On the server:
   a. From the server, determine if lack of access is to all of the storage (the entire storage system) or only to a portion of the storage (specific storagesets). If there is no access to only a portion of the storage system, refer to step 3.
   b. If access is not available to the entire storage system, verify from the server that the correct driver versions are loaded and that all parameters for the driver are correct. For multi-path applications, verify that the multi-path software is set up correctly.
   c. Verify that all Fibre Channel cables are plugged in and that all green indicator LEDs are on.
   d. Examine the event or error logs on the system.
2. On the Fibre Channel switch to which the server is connected:
   a. Verify the appropriate cable connection and that the port Link LED is on.
   b. Execute commands on the switch and verify that the server HBA is logged into the fabric correctly. Verify the correct port connection: F-Port, L-Port public, or L-Port private (refer to the specific HBA for more information on the correct login port types).

      F-Port: Tru64 UNIX, HP OpenVMS, HP-UX Fabric, Linux, Microsoft Windows NT, Windows 2000, SGI IRIX, and Sun Solaris.

      L-Port, 1 public: Novell NetWare.

      L-Port, x private, x phantom: HP-UX FC-AL.
   c. Verify all switch configuration and parameter settings.
   d. Verify that the switch is in the fabric and not segmented.
   e. Verify that all E-Ports are online.
3. On the Fibre Channel switch to which the storage is connected:
   a. Verify the appropriate cable connection and that the port green indicator LED is on.
   b. Execute commands on the switch and verify that the server HBA is logged into the fabric correctly. Verify the correct port connection: F-Port or L-Port private.

      F-Port: MA6000, MA/RA8000, EMA/ESA12000, EMA16000 set to FABRIC Topology.

      L-Port, x private, x phantom: MA6000, MA/RA8000, EMA/ESA12000, EMA16000 set to LOOP_HARD Topology.
   c. For MA6000, MA/RA8000, EMA/ESA12000, EMA16000:

      Verify the connections to the storage system. Execute a “show connections” command at the CLI and verify that the server connection is “online.” Verify the connections are named correctly.

4. On the storage system:
   a. Verify correct controller settings and configuration, “show this” and “show other.”
   b. Verify that the controller ports are online and configured for the correct topology setting.
   c. Verify that the storagesets are online to the appropriate controller without errors.
   d. Verify that the storagesets are correctly configured and enabled for access, “show unit dn.”
   e. Verify that unit offset parameters are correct. Also verify that the appropriate storage controller port is indicated in the connection name that will be accessed by the unit you have enabled.
   f. Verify that the connection OS parameter type is set correctly for the operating system that is using the connection.
5. General Fibre Channel switch verification:
   a. If zoning is in effect, verify that the effective zone matches the enabled zone.
   b. Verify that all zone definitions are correct.
   c. Verify that zoning alias/nick names are assigned to the correct WWNs.
   d. Verify that the servers and storage being accessed are in the same zone. If zoning is in effect, the WWN must be in a zone that is in the enabled configuration or it will not have access to the fabric.
   e. From the switch GUI, examine the name server table. Verify that the appropriate WWNs are listed and what zones they are in. Verify that the zones required are in the enabled configuration.
   f. Fabric segmentation occurs when you connect together two switches or two fabrics and one of the following mismatch conditions exists between them:
      — Zoning configuration mismatch
      — Zoning type mismatch
      — Zoning content mismatch
      — Switch configuration parameter mismatches

Note: All switches in a fabric must have the same switch parameter settings with the exception of the following parameters:
switch name
IP address
domain ID

If you are experiencing fabric segmentation, carefully review and compare these settings in each of the two switches or fabrics.

6. QuickLoop verification:

Note: QuickLoop is only required for HP-UX private loop attachment.

   a. Verify that the QuickLoop license is installed.
   b. Verify that the switch ports are set to QuickLoop mode.
   c. If using QuickLoop with two Fibre Channel switches, verify that the switches are in a QuickLoop partnership.