HP StorageWorks B-Series Remote Replication Solution white paper

Protecting your business with HP StorageWorks EVA and B-Series products

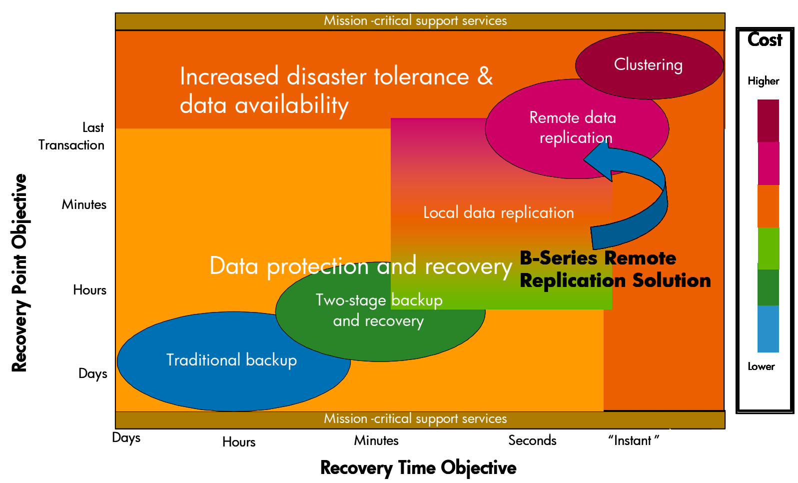
Protecting your business

In today's information-centric world, it is imperative that organizations protect their growing information assets to ensure business continuity, meet global data mobility requirements, and ensure compliance with new data retention and recovery regulations. Not having appropriate measures in place can result in lost productivity, lost revenues, lower customer satisfaction, and negative brand image as well as severe financial penalties.

However, what is "appropriate" for one organization may not be sufficient for another. The technical approach taken to address disaster recovery may vary, depending upon the type and size of the organization, the importance of the data, and the geographic location of the business.

To adequately architect Business Continuity (BC)/Disaster Recovery (DR) technology solutions that deliver the right level of service at the right cost, businesses should start by classifying and prioritizing their data and applications based on their importance to the business, from "business critical" to "not important," and establishing recovery objectives, in terms of recovery time (maximum time the application can be unavailable) and recovery point (level of data currency required) for each level. The HP StorageWorks continuum helps you select the appropriate technology solution that meets the various recovery objectives at the lowest possible cost, and thus minimize the impact to your operations (see Figure 1).

Figure 1.



Organizations need remote replication of data to a secondary site outside the "circle of disaster" to protect data and processes from major regional disasters like earthquakes, fires, or large scale power outages. Such solutions give your business the ability to recover data and applications quickly and avoid major business disruptions. The HP StorageWorks B-Series Remote Replication Solution is a leading solution from HP that provides a flexible and reliable infrastructure platform for protecting your information assets and minimizing business risk.

Solution overview

In a remote replication BC/DR solution, data is copied online, in real time, to a remote disk array over a storage area network (SAN) enabling recovery in case the local disk array or site fails.

HP has flexible solution options with its B-Series products allowing customers to choose which BC/DR configuration best fits their requirements. HP StorageWorks has Continuous Access solutions based on HP StorageWorks XP/Enterprise Virtual Array (EVA) disk arrays and FC direct/WDM/FCIP inter-site links. This white paper focuses on the HP StorageWorks Continuous Access EVA and FCIP as the inter-site link.

The solution consists of a minimum of these hardware and software components:

- HP StorageWorks EVA disk arrays with HP StorageWorks Continuous Access software (per site)
- B-Series SAN Infrastructure (directors, switches, and so on per site)
- B-Series Multi-Protocol Routers/Blades per site
- Inter-site links—FCIP, FC, or WDM as appropriate

HP StorageWorks Continuous Access provides:

- Real-time synchronous data replication between EVA storage arrays for disaster tolerance
- A new enhanced asynchronous replication scheme for longer distances (XCS V6.x)

Additionally, based on business requirements:

- Solutions can be deployed with Recovery Point Objectives (RPO) and Recovery Time Objectives (RTO) ranging from zero to several hours
- Clustering technologies can be deployed on primary/local and secondary/remote site if an RTO of near zero. These provide automatic application failover and failback capabilities.
- Solutions can be deployed using high availability and NSPOF inter-site links

Failover management can be simplified using HP StorageWorks storage management tools:

- HP StorageWorks Command View EVA
- HP StorageWorks Replication Solution Manager (RSM)

HP StorageWorks 400 Multi-Protocol Router configurations

In this solution, HP StorageWorks 400 Multi-Protocol (MP) Routers are deployed both in local and remote sites.

HP StorageWorks B-Series Multi-Protocol Router Blade configurations

In this solution, HP StorageWorks B-Series Multi-Protocol Router (MPR) Blades are deployed on the HP StorageWorks 4/256 SAN Director Chassis, providing similar capabilities as the HP 400 MP Router in an integrated form in both the local and remote sites.

HP StorageWorks 400 Multi-Protocol Router + HP StorageWorks B-Series Multi-Protocol Router Blade configurations

In this solution, HP customers can mix the standalone HP 400 MP Router and HP StorageWorks B-Series MPR blade on HP 4/256 SAN Director Chassis as they desire in local and remote sites.

The preceding configurations are discussed in detail in the Fabric design section.

Benefits and value

The B-Series Remote Replication Solution leverages a combination of HP products and technologies to ensure the highest level of data protection by:

- Mitigating risk:
  - Reduction of operational impact in the face of change, disruption, or disaster
  - Greater solution stability from network disruption by using a unique combination of routing and FCIP extension
  - Data encryption for improved security through the WAN
- Improving availability
  - Redundant fabric architectures for greater availability levels
  - Fault-tolerant director configurations and router blades
  - High-performance connectivity devices ensuring maximum application performance, network error recovery, and availability of remote storage applications
- Reducing cost
  - Reduction downtime cost
  - Lower capital expenses by offering choices of architectures and network devices
  - Lower operating costs by reducing bandwidth capacity with data compression, traffic management, and performance acceleration

Disaster recovery design basics

Several design considerations must be understood before a disaster recovery/remote replication solution is designed and implemented. This section helps to build a foundation in BC/DR basics so you can better understand your particular environment and business needs. The following topics are covered in this section:

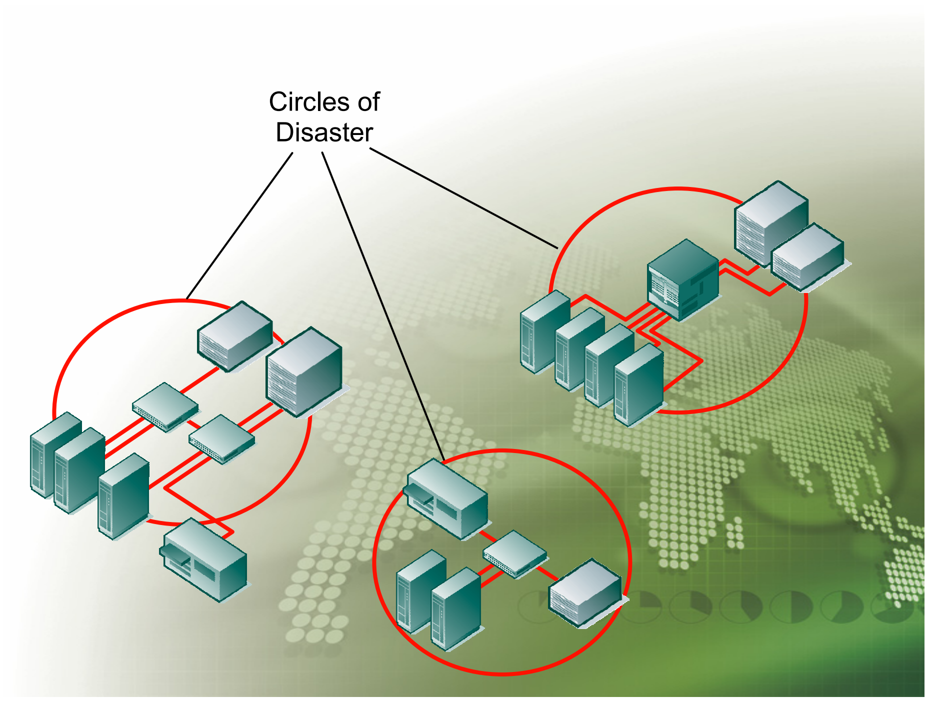
- Circle of disaster
- What data must be replicated?
- RPO and RTO
- Application throughput and response time
- Inter-site link latency
- Synchronous replication
- Asynchronous replication
- Inter-site link and bandwidth
- High availability and NSPOF inter-site links

Circle of disaster

A key element of an organization's BC/DR strategy is the circle of disaster. While replicating data to another disk system within your data center provides a certain level of protection, it certainly does nothing if an earthquake, hurricane, fire, or large-scale power outage takes out the data center. To adequately protect data from these types of regional disasters, organizations must replicate data to a secondary site that is geographically distant enough that any natural or man-made disaster cannot take out both sites at the same time. This distance between the primary/local and secondary/remote sites forms the radius of a circle called the "circle of disaster."

The nature of the business, regulations, and competitive advantages dictate the size of the circle of disaster. Some high-end multi-national financial institutions ensure that their local and remote sites are on two separate continents resulting in a circle of disaster that spans the entire globe. So, distance between the local and remote site is critical to reducing risk.

Figure 2.



What data must be replicated?

In an ideal world, organizations would choose to protect every bit of corporate data. However, such an all-encompassing data protection strategy will be prohibitively expensive. Therefore, businesses typically classify their corporate data and prioritize their applications based on their importance to the business, from "business critical" to "not important." Business-critical data should be replicated to protect the business from disasters or accidental data loss.

Recovery Point Objective (RPO) and Recovery Time Objective (RTO)

These two definitions are key to understanding the basics of BC/DR.

- **RPO**—The point in time to which data must be recovered after an outage as determined by the business unit. This indicates how much data loss can be tolerated.
- **RTO**—The period of time within which systems, applications must be recovered after an outage. This is a measure of high availability. It is the length of time the business can afford to spend returning an application to operation.

RPO and RTO are the business requirements that the IT manager must specify. They specify how fast application recovery must be accomplished and how much data loss can be tolerated.

After data classification has taken place, then the volumes that contain the business-critical data are set up to be replicated with the best RPO and RTO characteristics the organization can afford.

Ideally users do not want to experience any system or network failure in any situation, but the reality is that this level of service is either impossible or impractical to provide given the costs. As a result, customers must establish business rules for minimum recovery times and maximum data loss. Table 1 details some sample objectives for a few common application environments.

Table 1. Sample recovery objectives by application

Application or environment	RPO maximum	RTO maximum	Enabling technologies	Cost to implement
Email	30 min	30 min	Clusters	Moderate
File serving	30 min	30 min	Clusters	Moderate
Financial transactions	0 sec	10 min	Non-stop computing	High
Decision support	12 hours	4 hours	Tape backups	Low

An enterprise's ability to meet a given RTO is generally driven by the enabling technology since automation is the key to getting the remote site operational quickly after the primary or local site has failed. The RPO dictates the type of data replication that must be deployed. If RPO is zero, then synchronous replication is necessary. In case of asynchronous replication, RPO determines the minimum capacity and maximum latency of the inter-site link used to deliver writes to the recovery site.

Application throughput and response time

All applications that customers deploy have established targets for throughput and response time. Table 2 shows the typical requirements for a few prominent application categories.

Table 2. Target application response times

Application or environment	Average	Maximum
Email	20 ms	40 ms
File management	40 ms	60 ms
Transaction processing	15 ms	30 ms
Decision support	40 ms	60 ms

Sizing link bandwidth requirements for replication is critical to meeting the requirements of customer's application. With synchronous replication, inadequate bandwidth provisioning constricts the flow of writes, which can add significant queuing delays, causing unacceptable application response time. With asynchronous replication, inadequate bandwidth provisioning results in the inability to meet the desired RPO objective. Provision too much bandwidth and the cost of the replication solution can increase significantly.

Inter-site link latency

HP Continuous Access EVA can move data at extreme distances. However, the speed of light in fiber optic cables (5 microseconds per kilometer) causes inherent delays, called latency. At extreme distances, latency is the limiting factor in replication performance, regardless of bandwidth. Also, in real life, the quality of the inter-site link also affects the latency. Other contributors to latency are the number of hubs, switches, routers, and firewalls in the network route. Hence, it is essential that the inter-site link provided by the Telco vendor be tested to see what the "real" latency is. This should be clearly specified in the service level agreement with your Telco vendor.

There is a point of inter-site link latency after which an application performance becomes unacceptable (in synchronous replication) or the RPO objective is not met (in asynchronous replication). This point is individually determined for every customer's application environment and business needs.

Synchronous replication

An RPO of zero signifies that synchronous replication is needed. Synchronous replication offers the highest levels of data protection. In this, the disks in both the local and remote disk arrays are identical and concurrent at all times. Data is written simultaneously to the mirrored cache of the local disk array and the remote disk array, in real time, before the host application I/Os are completed, thus ensuring the highest possible data consistency. Hence, the local disk array waits for acknowledgment that the replication write has reached the cache of the remote disk array, before sending acknowledgment to the host. Size and latency of the bandwidth play an important role in the performance of the application with synchronous replication in terms of response time. In general, synchronous replication (and shorter RPOs) requires higher bandwidth.

Asynchronous replication

Asynchronous replication provides a faster I/O response time to servers by sending a write completion back to the host application upon receipt in the cache. RPO is usually greater than zero using asynchronous replication. In this, the local disk array acknowledges the write as complete to the host, as soon as the write is in local cache of the storage array. Data at the remote site will be consistent with the data at primary site, but it will lag behind the data at primary site. Size and latency of the bandwidth do not play an important role in the performance of the application with asynchronous replication in terms of application response time but they significantly affect the RPO of the customer data being replicated. From a business perspective, the RPO provides a tolerance for data loss. The lowest cost solution provides only enough bandwidth to exactly meet this objective.

Inter-site link and bandwidth

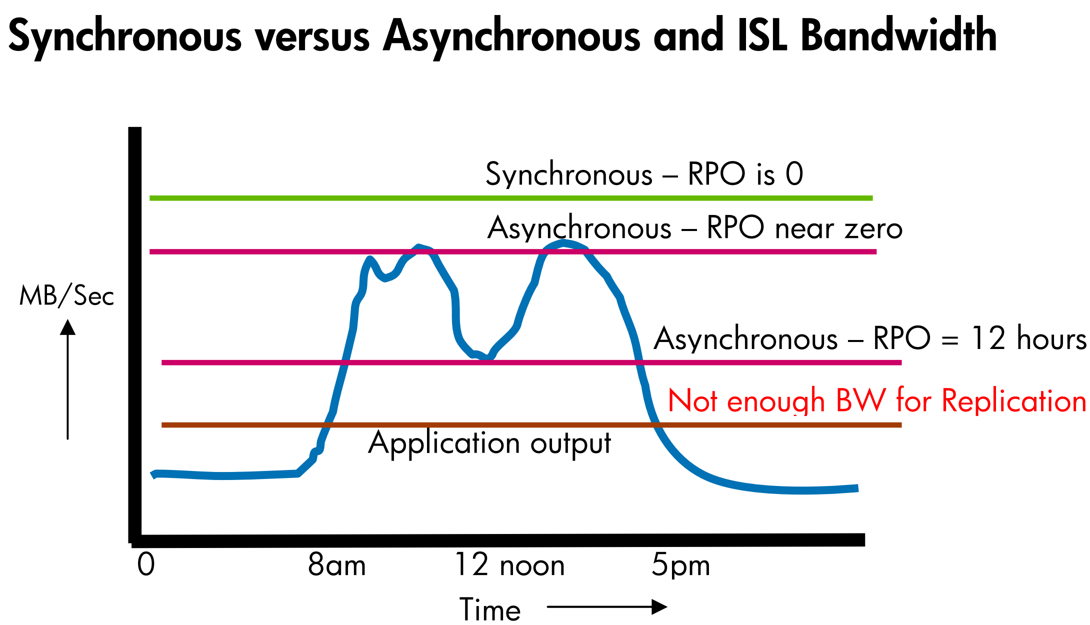
The location of the remote or secondary site determines the inter-site link technologies that meet your distance and performance requirements. Bandwidth is the amount of data that can be transferred from one computer to another in a certain amount of time. The following factors affect the bandwidth you need:

- **Distance**—Fibre Channel over IP (FCIP) supports the longest distances. You can replicate data across continents using FCIP inter-site link technology.
- **RPO**—As mentioned earlier, RPO is the amount of data loss that the business can tolerate as a result of a disaster. With a short RPO, new data must be copied over to the remote disk array quickly; hence, more bandwidth must be reserved for peak loading. Applications that require an RPO of zero (no data loss after a site failure) must replicate application data synchronously to meet this goal. Note that synchronous replication mode requires more bandwidth than asynchronous as shown in Figure 3.
- **Performance**—When you cannot adjust the distance between sites, you may be able to increase performance by increasing bandwidth. Data moves through the inter-site links at the same speed of light. Because the bits are longer in the low-bandwidth link (that is, they spend more time on the wire), the data takes more time to unload than the data in the high-bandwidth link. The same advantage applies to loading data into a high-bandwidth link compared to a low-bandwidth link.
- **Peak loads**—With synchronous replication, the inter-site link must accommodate the peak write rate and throughput of your applications. For asynchronous replication, the inter-site link must be sized up properly to meet the RPO objectives of your business. Usually, a lower bandwidth is necessary with asynchronous replication, but keep in mind, you can never achieve an RPO of zero with it.

- **Data change rate**—An enterprise should determine the rate of change for the logical volumes being replicated. This plays a critical role in determining the bandwidth required for long-distance replication. For example, if the source volume has an average change rate of 1 MB/s during peak times, the WAN will need to accommodate at least 1 MB/s. In fact, to accommodate unusually heavy loads and growth, additional bandwidth should be considered. However, if the data being replicated is highly compressible, then one could take into consideration the compression capabilities of the FCIP SAN extension hardware, to provision lesser bandwidth and thus save money on recurring bandwidth expenses.

Figure 3 describes how the RPO objective can be inversely proportional to the size of inter-site link bandwidth, that is, lower RPO means higher bandwidth.

Figure 3.



- **Data normalization**

  A frequently overlooked aspect of link sizing is normalization time. This is the time it takes to bring the remote site up to date with the primary initially or to resynchronize the primary after a site disaster. A couple of reasons why normalization is important during bandwidth sizing:
  - During the normalization period, availability is at risk until the remote site has a complete and up-to-date copy of the data.
  - Normalization can consume large portions of the available bandwidth causing degradation of other applications using the common interconnect.
  - It can take a long time for the initial copy to complete if the size of the primary data is large.

HP Storage Architects can help you size the right inter-site link and bandwidth for your business requirements.

High availability and NSPOF inter-site links

The business criticality of the data being replicated dictates how available the remote data replication solution must be. In the case of an active/active data center arrangement, where applications run on servers in two or more sites, a highly available system with considerable redundancy might be required. In this customer scenario, bidirectional disaster tolerance is usually required. Synchronous replication is typically deployed with an active/active system to meet the cluster's requirements. In the case of asynchronous replication, the availability requirements may not be as stringent. Usually, this type of an organization cannot afford the risk of a down backup system. Hence, the expense of a high-availability replication system is required.

Disaster recovery solution components

Creating a disaster recovery environment can be done with the existing FC SAN or from scratch. The environment requires several components for a successful implementation. This section describes the key components of a disaster recovery solution.

Inter-site link options

HP StorageWorks Continuous Access EVA software supports the following links and bandwidths:

Table 3. Supported links and bandwidths

Technology	Bandwidth (Mb/s)
4 Gb/s Fibre Channel	4000
2 Gb/s Fibre Channel	2000
1 Gb/s Fibre Channel	1000
1 GbE (Gigabit Ethernet) IP	1000
OC3 IP	155.5
E4 IP	139.3
100 Mb/s IP	100
T3 IP	44
E3 IP	34.304
10 Mb/s IP	10
E1 IP	2.048 (minimum supported bandwidth)
T1 IP	1.54 (not supported)

This table shows the link technologies and bandwidths that HP Continuous Access EVA supports for inter-site links. HP Continuous Access EVA supports bandwidths from 2 Mb/s to 4 GB.

Note
HP Replication Solutions Manager exhibits long management delays when managing remote arrays that are connected by way of low-bandwidth links with over 36 ms of one-way latency.

The IP inter-site link technology is typically purchased through your Telco or from an IP service provider.

HP StorageWorks 400 Multi-Protocol Router

The HP StorageWorks 400 MP Router enables organizations to derive more value from the SAN infrastructure by providing the ability to connect SAN islands and share resources, without the risk and complexity of merging them into a single large fabric, both for FCIP and native Fibre Channel connectivity. The resulting modular SAN design provides a more scalable, flexible, and stable infrastructure that is not limited by distance, maximizing the business value of the SAN infrastructure.

The HP 400 MP Router provides two types of SAN Services: (1) FC-FC Subnet Routing Service for SAN island consolidation and (2) FCIP Tunneling Service for SAN extension over distance. These services provide new options for connecting SAN islands and extending SAN benefits over multiple networks, to larger SAN sizes, and across longer distances. Key advantages of this approach include:

- Simplifying SAN design, implementation, and management through centralization.
- Providing a seamless and secure way to share resources across multiple SANs without the complexity of physically merging those SANs.
- Creating a more unified SAN environment with easier interconnection and support for SANs and SAN resources purchased from different storage vendors.
- Reducing disruptions created by events such as data migration, storage or server consolidation, migration to production environments, and application rebalancing between fabrics.
- Extending those benefits over long distances with high-performance FCIP services.

HP StorageWorks B-Series Directors with Multi-Protocol Router blade

The HP StorageWorks 4/256 SAN Director and Power Pack provide a high-performance, highly reliability, and flexible connectivity platform supporting the most demanding mission-critical data center applications. As an option for the industry-leading 4/256 SAN Director, the HP MPR blade is able to provide these SAN services with the performance and high-availability characteristics of the director platform. HP StorageWorks B-Series MPR blade provides the same functionality as the HP 400 MP Router. With up to 384 ports of connectivity in a single domain, the 4/256 SAN Director offers flexible scalability for large core-to-edge SAN architectures along with performance and high availability.

HP StorageWorks Enterprise Virtual Array (EVA)

The HP StorageWorks 4x00/6x00/8x00 Enterprise Virtual Arrays offer customers in the mid-range to enterprise market place leading high-performance, high-capacity, and high-availability "virtual" array storage solutions. Designed for the data center where there is a critical need for improved storage utilization and scalability, the EVA meets application-specific demands for consistent high transaction I/O for the customer, and provides easy capacity expansion, instantaneous replication, and simplified storage administration.

HP StorageWorks Continuous Access

HP StorageWorks Continuous Access EVA Software is an array-based application that provides disaster tolerant remote replication on the entire EVA product family. Continuous Access EVA provides the necessary components to solve enterprises business continuity objectives in a very cost-effective and easily deployable package. They can achieve a competitive advantage by combining disaster tolerant solutions and disaster tolerant managed services into their planning and daily routines, ensuring data security, availability, and integrity.

EVA4x00/6x00/8x00 customers can purchase an optional Continuous Access EVA license for their EVA units. Using Continuous Access EVA, customers can now even replicate data between all of their EVAs—EVA3000s, EVA5000s, EVA4x00s, EVA6x00s and EVA8x00s—and replicate data on high-performance enterprise drives to FATA drives.

The new Continuous Access EVA with XCS v6.x

New with the XCS v6.x firmware code, enhanced asynchronous replication is supported. Enhanced asynchronous replication is supported only between EVA 4x00/6x00/8x00 models, where both the arrays are running XCS v6.x firmware. Continuous Access EVA now uses a buffer-to-disk implementation that greatly increases the support for long-distance replication and latency. Note that there are many design factors (as described in this document) that influence the performance of replication when using replication. These factors include the I/O load, the available bandwidth, and link latency. The new enhanced asynchronous replication implementation insulates application users from latency delays caused by propagation delays and intermittent congestion on the inter-site link. In addition, enhancements to Continuous Access EVA have increased the capability to better load balance the replication traffic across links when two inter-site links are used.

Replication Solutions Manager (RSM)

HP Replication Solutions Manager integrates both local and remote replication into a brilliantly simple-to-use GUI and CLUI. RSM removes almost all of the complexity of configuring and managing replication environments through easy–to-use wizards and automation of common tasks.

HP StorageWorks Cluster Extension EVA software

HP StorageWorks Cluster Extension EVA software is an integrated solution that provides protection against system downtime with automatic hands-free failover/failback of application services and read/write enabling of remotely mirrored mid-range storage over metropolitan distances. It makes an RTO of zero or near zero a possibility.

HP MetroCluster/ContinentalCluster Continuous Access EVA

HP MetroCluster and HP ContinentalClusters are disaster-tolerant solutions for the HP-UX 11i environment. Both are built on the HP foundation clustering software, HP ServiceGuard, providing the highest level of availability and business continuity for enterprise data centers. Both are integrated with Continuous Access EVA. HP Continuous Access EVA supports stretched ServiceGuard cluster running on HP-UX 11iv1 or HP-UX 11iv2 Update 2. This software also provides automatic hands-free failover/failback of application services.

Fabric design

This section provides an overview of 400 MP Router and MP Router blade configurations based on the fabric architecture, SAN extension technology, and multi-protocol router type. For detailed configuration information, see the HP StorageWorks B-Series Remote Replication Solution Best Practices Guide at: www.hp.com/go/bseriesreplication

400 MP Router and MP Router blade SAN extension technology

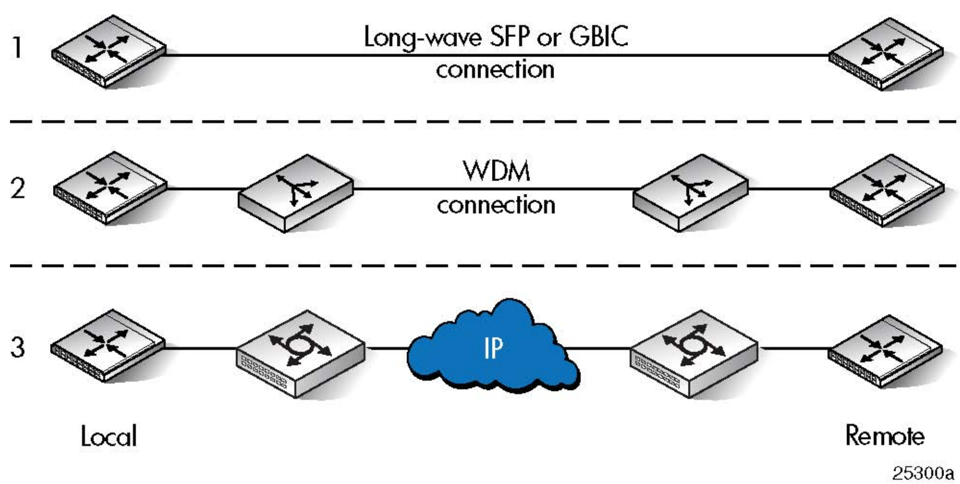
The 400 MP Router and MP Router blade are supported with four types of SAN extension technology:

- **Native Fibre Channel**—This choice is based on direct Fibre Channel extension using long-wavelength small form pluggable (SFPs) and Dark Fiber. Distance support is dependent on the SFP range, but is usually below 50 km. This offers the highest level of link performance.
- **Native Fibre Channel with Wavelength Division Multiplexing (WDM)**—SAN extension utilizing WDM (either DWDM or CWDM) through an existing optical metropolitan area network (MAN), which can handle multiple protocols. Distances range from 50 km up to 100s of km. CWDM is lower cost then DWDM, which offers longer distances. Both CWDM (8 wavelengths) and DWDM (32 wavelengths) offer the benefit of multiple wavelengths of light over a single fiber optic cable for increased link capacity.

- **Extension over FCIP**—Fibre Channel over IP (FCIP) is a primary technology for data replication over the IP WAN. The biggest advantage of an FCIP solution is the significant increase in distance over which data can be replicated. The overall cost of an FCIP solution is also generally lower, making FCIP-based replication a competitive solution, typically over longer distances. Performance is typically lower then native Fibre Channel. However, similar performance can be achieved with additional cost.
- **FC Routing**—Logically connects a subset of devices in one fabric to a subset of devices in another fabric with the administration and fault isolation benefits of separately managed fabrics. While FC Routing may not technically be an SAN extension technology, the capabilities and features of FC Routing eliminate or reduce many of the issues associated with merging physically separate SAN fabrics.

You can use the 400 MP Router and MP Router blades with any of these technologies to extend disaster tolerant applications such as HP Continuous Access EVA and XP.

Figure 4. SAN extension technologies



400 MP Router and MP Router blade configurations

There are two MP Router models available in the B-Series: the HP 400 MP Router and the MP Router blade. The blade is an option for the 4/256 SAN Director and inherits the high-availability attributes of the director platform. The HP 400 MP Router, on the other hand, provides the flexibility and affordability of a standalone device. The two models can be implemented in symmetric or asymmetric (mixed router types) configurations depending on cost and availability targets.

400 MP Router and MP Router blade native Fibre Channel configurations

The 400 MP Router and MP Router blade both support native FC extension using long-wave and extended-reach SFP transceivers for distances up to 10 km and 35 km, respectively. With WDM technology, distances up to 100 km @ 4 Gb/s and up to 500 km @ 1 Gb/s are supported.

In non-routed native FC configurations, local and remote fabrics merge when connected through long-wave or extended-reach SFP transceivers or WDM technologies. The connection is like any other ISL in a fabric connecting two switches. By using the 400 MP Router or MP Router blade FC routing feature along with native FC, the local and remote fabrics connect without merging. Using LSAN zones that contain both local and remote devices, the FC Routing capability of the MP routers will allow this subset of devices in the LSAN zone to communicate as if they were in the same fabric.

400 MP Router and MP Router blade FCIP configurations

You can use the 400 MP Router and MP Router blades' integrated FCIP, FC switching, and FC Routing capabilities for an efficient SAN extension implementation.

In non-routed FCIP configurations, local and remote fabrics merge when connected through an IP network. The IP connection is like any other ISL in a fabric connecting two switches. By using the 400 MP Router or MP Router blade FC Routing feature along with FCIP, the local and remote fabrics connect without merging. Using LSAN zones that contain both local and remote devices the FC Routing capability of the MP routers will allow this subset of devices in the LSAN zone to communicate as if they were in the same fabric.

Note:
HP recommends using the FC Routing feature along with FCIP to avoid the issues associated with merging the local and remote fabrics.

400 MP Router and MP Router blade fabric architecture

There are four proven SAN extension solutions that can be implemented with various director/switch/router combinations depending on the size of the SAN and the availability level required:

- 2 Fabric architecture—Used for the low end of the spectrum with lower throughput and connectivity levels (smaller SANs)
- 4 Fabric architecture—Used for the low end of the spectrum with higher availability then the 2 Fabric solution
- 5 Fabric architecture—Used when larger scalability is required, and for I/O write-intensive situations
- 6 Fabric architecture—Provides the highest availability level

By using the B-Series directors and standalone switches capabilities, and the combination of FC Routing, switching, and SAN extension, the physical configuration can be reduced to fewer elements to drive additional simplification and ease of maintenance.

MP Router FCIP configurations

Figure 5 shows typical HP Continuous Access no single point of failure (NSPOF) configurations in which the MP Routers provide FC Routing and FCIP capabilities. Although the local and remote fabrics are connected through the IP networks as with any implementation of FCIP, with the FC Routing feature, the local and remote fabrics do not merge. This feature allows you to selectively share devices without the need to merge the fabrics when using FCIP and thus the administration and fault isolation benefits of separately managed fabrics are maintained.

There are five permutations of this configuration using any combination of MP Router blades and 400 MP Routers.

Figure 5. FC Routing and FCIP NSPOF configuration with direct connect devices to MP Router blades and 400 MP Routers

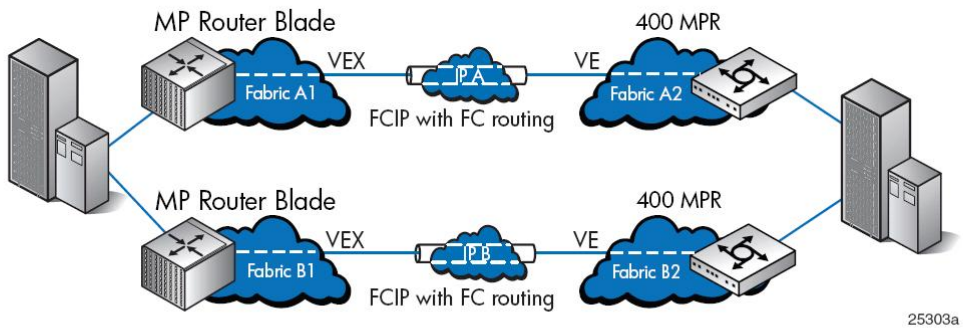


Figure 7 is an example of this topology using MPR 400 router to implement a Continuous Access four-fabric LSAN replication zone solution.

Design considerations when using FC Routing

There are some considerations that must be taken into account when designing an Extended Fibre Channel SAN using FCIP and Fibre Channel Routing. At the highest level for a two fabric solution, one fabric must be designated as the "Backbone Fabric" and the other fabric (at the other site) is an "Edge Fabric." The "Backbone Fabric" is the functional element that makes the logical connection between "Edge Fabrics" or between the "Edge Fabrics" or between the "Backbone Fabric" and an "Edge Fabric" and there are some limitations placed on the "Backbone Fabric."

In the example in Figure 5, the MPR FCIP port of the "Backbone Fabric" is configured as a VEX Port, and the "Edge Fabric" MPR FCIP port is configured as a VE Port. It is beyond the scope of this document to cover all the benefits, issues, and configurations of FC Routing, but there are two implications of this decision. The first is that there are scaling limitations placed on the "Backbone Fabric."

Table 4. "Backbone Fabric" versus "Edge Fabric" limits

	FOS Version	
	FOS 5.1	**FOS 5.2**
Max # of Device Ports per Backbone Fabric	256	512
Max # of Device Ports per Edge Fabric	1000	1200
Max # of Fibre Channel Switches per Backbone Fabric	10	10
Max # of Fibre Channel Switches per Edge Fabric	26	26

The second limitation for this simple two site example is which fabric contains the IP network. With the scaling limitations of the "Backbone Fabric" it might seem intuitive to select the smaller site to be the "Backbone Fabric," but then the larger site has the IP network as part of that fabric. The interface between the "Backbone Fabric" and "Edge Fabric" for a FCIP solution is a VEX Port on the MP Router and therefore the placement of the VEX port determines which Fabric(s) is the "Backbone Fabric." Because the VEX Port is the fabric boundary, the IP network is part of the Secondary or "Edge Fabric." The issue is that the IP network is the most unstable part of this configuration and therefore the fabric that contains the IP networks will see more disruptions than the fabric without the IP network.

The design needs to balance the scaling limits of the "Backbone Fabric" with the disruptions an IP network can have on the "Edge Fabric."

There are two alternate solutions to overcome these two issues. The more traditional solution has been to create a dedicated "Replication Fabric" or fabrics as shown in the five-fabric and six-fabric implementations in Figure 8 and Figure 9. The other possible solution is to use a dedicated "Backbone Fabric" solution as shown in the six-fabric with dedicated "Backbone Fabrics" implementation in Figure 10. This configuration is really a modification of the four-fabric solution where the functionality of the "Backbone Fabric" is removed from both the local and remote site fabrics and a new dedicated fabric is created for the "Backbone." Both solutions eliminate the scaling limitations on both sites and the IP network only affects the replication traffic. The dedicated "Backbone Fabric" has two advantages over the dedicated "Replication Fabric" solution. First it is more scalable, and second it allows devices other than the EVA storage connected to this dedicated "Replication Fabric" to communicate if needed.

Continuous Access EVA configurations

The 400 MP Router and MP Router blade support the HP standard Continuous Access EVA replication configurations. This includes the two-fabric, four-fabric, five-fabric, six-fabric, and six-fabric with dedicated "Backbone Fabric" implementations as shown in Figure 6, Figure 7, Figure 8, Figure 9, and Figure 10. These five configuration examples are all drawn showing 400 MP Routers. However, they could all be implemented using the MP Router blade or a combination of both.

Figure 6 shows a typical 2-port EVA6000/5000/4000/3000 configuration of a Continuous Access two-fabric LSAN replication zone solution. In this configuration, zoning is used to separate host traffic from replication traffic in the fabric. Standard B-Series zoning is used for one port from each EVA controller for local host access as in any other fabric. The other EVA port on each controller is dedicated to replication traffic and LSAN zoning is used to enable the FC Routing feature of the MP Router to allow these devices to communicate across the FCIP Link as if they were in the same fabric.

LSAN zoning is done using the same tools as standard B-Series zoning is with the exception that the zone **must** begin with "LSAN_" (case insensitive) and will include the Port WWNs of both the Local and Remote EVA replication ports from both fabrics and this LSAN zone is created on both the Local and Remote fabrics.

This is the lowest cost configuration. However, no path redundancy is provided and therefore this configuration does not provide any redundancy in the event of a fabric or WAN failure.

Figure 6. Continuous Access EVA two-fabric FCIP-router configuration

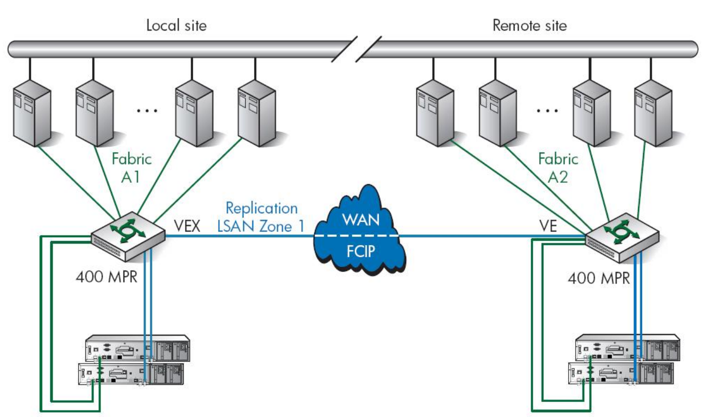


Figure 7 shows a typical 4-port EVA8000 configuration of a Continuous Access four-fabric LSAN replication zone solution. Like the previous configuration, zoning is used to separate host traffic from replication traffic in each fabric. In each fabric standard B-Series zoning is used for two ports from each EVA controller for local host access as in any other fabric. The other EVA ports on each controller are dedicated to replication traffic and LSAN zoning is used to enable the FC Routing feature of the MP Router to allow these devices to communicate across the FCIP Link as if they were in the same fabric.

This configuration is a better solution that the Continuous Access two-fabric LSAN replication zone solution in that there is NSOF as both the Local and Remote sites implemented dual redundant fabrics and a separate FCIP link for each fabric pair.

Figure 7. Continuous Access EVA four-fabric FCIP-router configuration

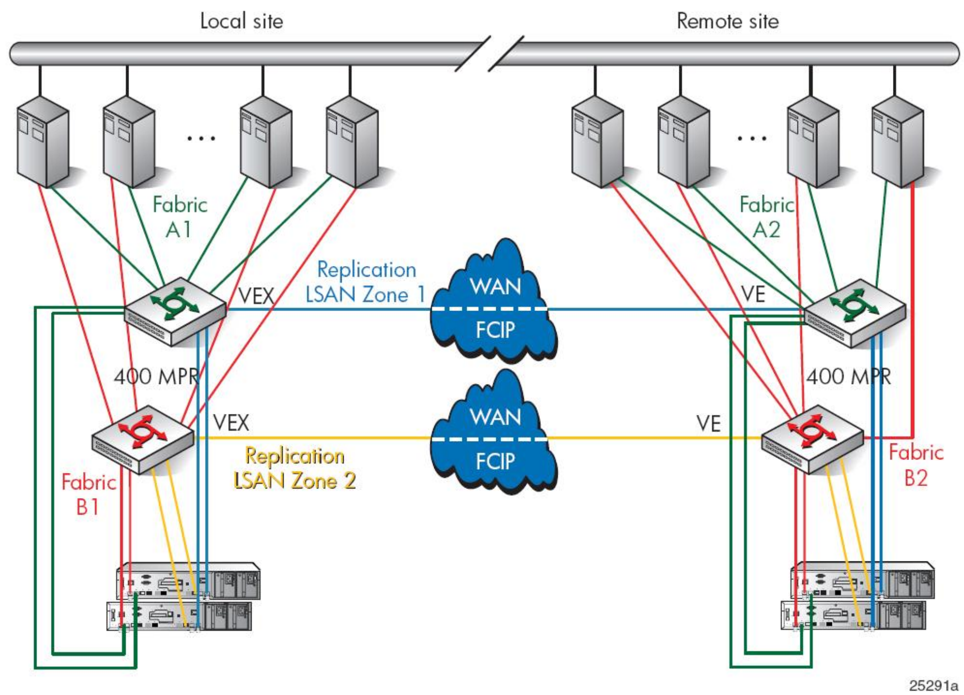


Figure 8 shows a typical 2-port EVA6000/5000/4000/3000 configuration of a Continuous Access five-fabric solution. In this configuration a dedicated fabric is implemented for replication traffic, thus eliminating the fabric merging issue with a traditional FCIP implement and the scaling limitation of an integrated "Backbone Fabric" solution.

In this configuration half the EVA controller ports are connected to a dedicated replication fabric. Since the only devices in this replication fabric are the replication ports of the local and remote site, FC Routing is not required. The benefit of the MP Router over traditional FCIP gateways is that it eliminates the need for a separate FC switch as the MP Router has both capabilities in one device.

Figure 8. Continuous Access EVA five-fabric FCIP-router configuration

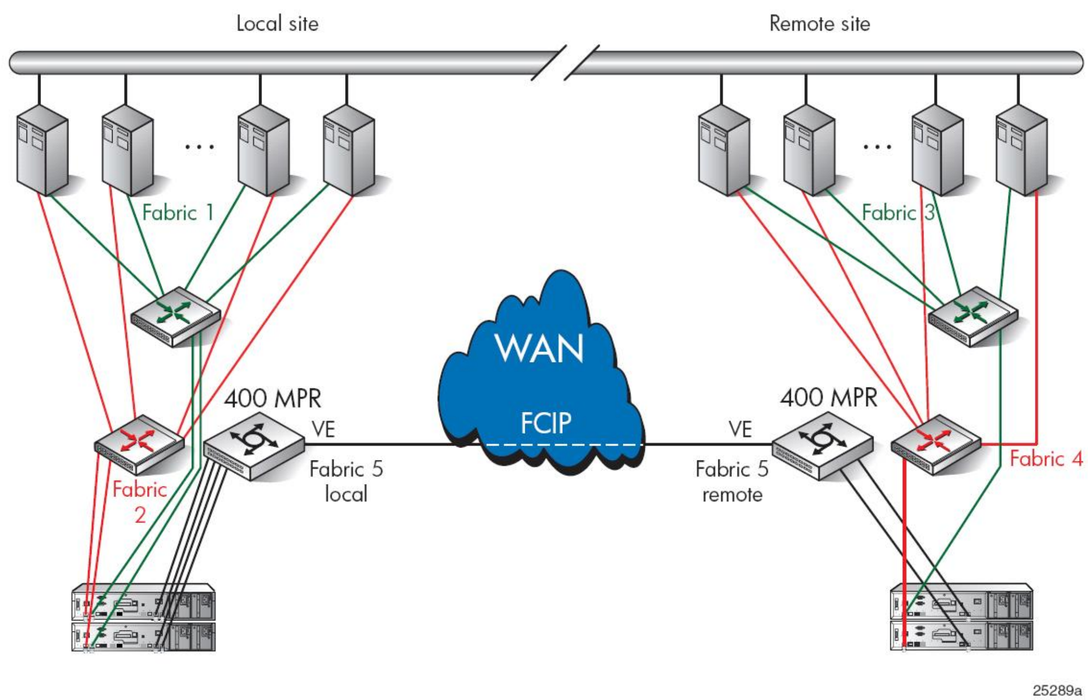


Figure 9 shows a combination of a 4-port EVA8000 and 2-port EVA6000/5000/4000/3000 configuration of a Continuous Access six-fabric solution. In this configuration two dedicated fabrics are implemented for replication traffic, thus eliminating the fabric merging issue with a traditional FCIP implement and the scaling limitation of an integrated "Backbone Fabric" solution. Another possible solution would be to implement a six-fabric with dedicated "Backbone Fabrics" as shown in Figure 10.

In this configuration half the EVA controller ports are connected to each dedicated replication fabric. Since the only devices in this replication fabric are the replication ports of the local and remote site, FC Routing is not required. The benefit of the MP Router over traditional FCIP gateways is that it eliminates the need for a separate FC switch as the MP Router has both capabilities in one device.

Note:
HP recommends that the primary or local site in this configuration be a 4-port EVA8000 and the secondary or remote site can be either a 4-port EVA8000 or 2-port EVA6000/4000/5000/3000.

Figure 9. Continuous Access EVA six-fabric FCIP-router configuration

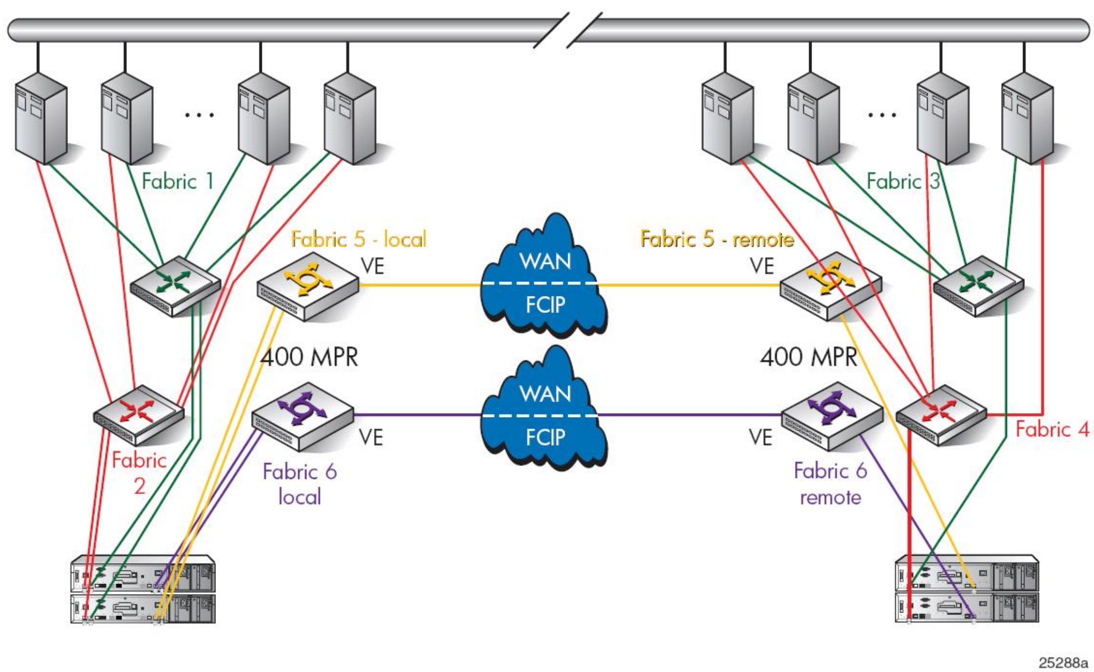
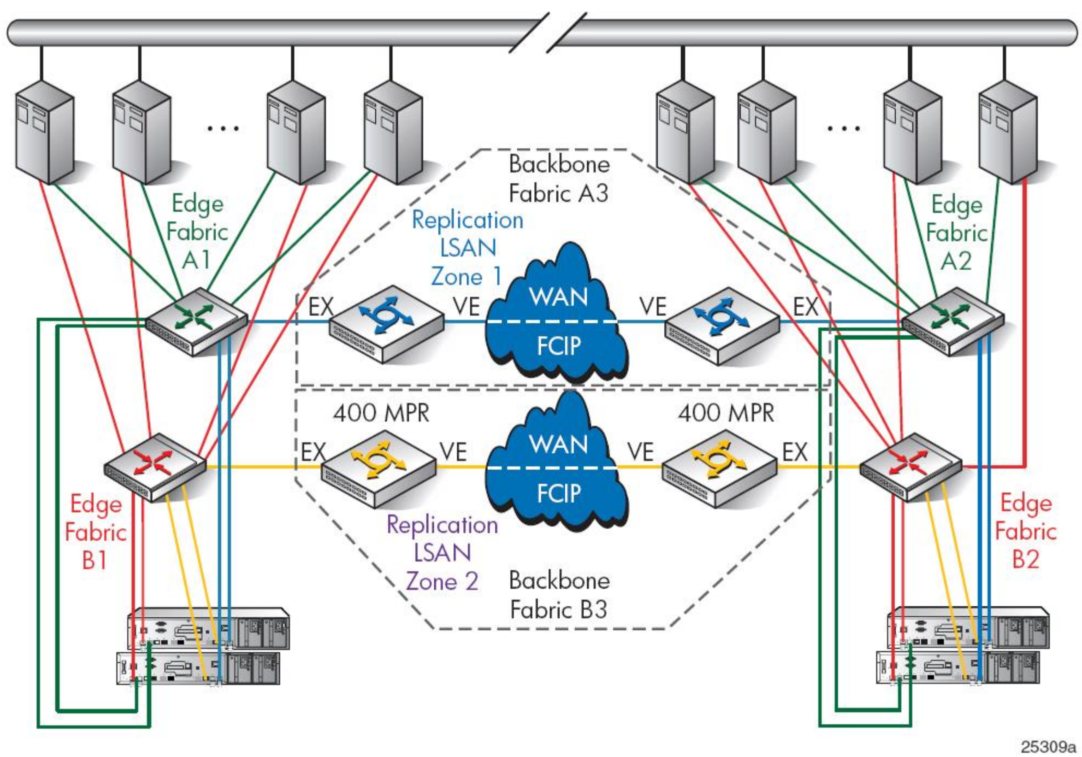


Figure 10 shows a typical 4-port EVA8000 configuration of a Continuous Access six-fabric with dedicated "Backbone Fabrics" LSAN replication zone solution. In this configuration half the EVA controller ports are connected to each of the dual redundant fabrics for each site.

This configuration using two dedicated "Backbone Fabrics" for FC Routing traffic between "Edge Fabrics" solves all the issues with traditional FCIP implementations and the scaling limitations of an integrated "Backbone Fabric" (Figure 6 and Figure 7) solution. FC Routing solves the issues associated with the merger of two physically separate fabrics. Since the "Backbone Fabric" is no longer used for local traffic, the scaling issues in terms of number of devices and switches in the fabric are mitigated. Also because the IP network is now part of the "Backbone Fabric" and not part of the local or remote fabrics, this minimizes fabric disruptions on these fabrics. And finally this solution allows for connecting more than two sites to the dedicated "Backbone Fabrics" and so the same "Backbone Fabric" could be used for a primary site to connect to multiple secondary sites.

This configuration also has the advantage of being able to share more devices than the dedicated replication fabrics. As the dedicated replication fabrics only have the EVA ports connected to them, only the storage controller can communicate over these fabrics. This enables a server to connect to storage resources on both the local and remote fabrics while maintaining the ease of management and fault isolation that two smaller physically separate fabrics have over one large merged fabric.

Figure 10. Continuous Access EVA six-fabric FCIP-router with dedicated "Backbone Fabrics" configuration



Customer scenarios

This section provides some examples of customer business requirements with deployment details and configuration. These solutions can be deployed with any of the fabric designs described in the preceding section. The technology focus in this white paper is on real-time array-based replication only.

Caution
The following tables detail examples of RPO, RTO, and bandwidth sizing based on a sample workload. Your specific application workload must be carefully analyzed by HP Storage Architects to determine the optimum size of inter-site link and specific disaster recovery parameters.

Microsoft Exchange

The following example configurations consist of Microsoft® Exchange Server 2003 remote replication solution deployed using HP StorageWorks EVA disk arrays, HP StorageWorks Continuous Access and HP StorageWorks B-Series routers. **Inter-site link latency and bandwidth are the significant factors in determining the performance of any Exchange replication solution (synchronous or asynchronous), and proper sizing of inter-site link will ensure best performance.** Microsoft recommends synchronous replication of Microsoft Exchange 2000 and 2003 servers (asynchronous replication is supported for Microsoft Exchange 2000 and 2003 but supportability is the responsibility of your specific hardware vendor). With the introduction of Microsoft Exchange 2007, both synchronous and asynchronous replication technologies are supported within the product. Synchronous replication solutions should be employed under the guidelines established for Exchange 2003. Synchronous replication of the Exchange 2007 production LUNs or LCR copies can be implemented for site-to-site disaster recovery and HP Cluster Extension EVA or XP can be implemented (in conjunction with HP Continuous Access) to automate failover and failback between data centers. Asynchronous replication is built into Exchange 2007 by way of Cluster Continuous Replication (a two node, Active Passive Majority Node Set cluster solution), or Local Continuous Replication (a single Server solution). Both imbedded technologies utilize host-based (asynchronous) replication on the local bus for LCR, and over Ethernet for iSCSI LCR and CCR. Replication is accomplished by creating a copy of databases and then seeding the databases by replay of transaction logs into the replica copies of data (known as replay). For more details on these technologies, refer to www.hp.com/go/hpcft.

The following table exemplifies sample configurations for Exchange 2003 and 2007 Synchronous replication solutions.

Note
Round trip latency requirement of 20 ms for transaction logs

Table 5.

Application	RPO	RTO	Average application response time	Round trip Inter-site link latency	Distance between sites	Type of data replication	Local or Primary site	Remote site	Inter-site link bandwidth	Auto failover
Microsoft Exchange 2,500 users	0	15 min	20 ms	10 ms	500 Km	HP Continuous Access Synchronous (Full Replication)	HP EVA 4x00, 2 X HP MPR 400	HP EVA 6x00, 2 X HP MPR 400	OC3 (155 Mb/s)	Yes (CLX)
Microsoft Exchange 5,000 users	0	15 min	20 ms	8 ms	750 Km (max)	HP Continuous Access Synchronous (Full Replication)	HP EVA 8x00 2C12D, 2 X HP MPR 400	HP EVA 8x00 2C12D, 2 X HP MPR 400	1 Gb/s	Yes (CLX)

Oracle 10g with ASM

The following example configurations consist of Oracle® 10g with ASM deployed using HP StorageWorks EVA disk arrays, HP StorageWorks Continuous Access, and HP StorageWorks B-Series routers. The database is 1 TB with approximately 1,000 users generating OLTP type access. In large enterprise environments with demanding requirements for replication, HP EVA disk arrays along with HP Continuous Access for EVA software provide an effective hardware/array-based replication solution for Oracle 10g with ASM. As data is replicated at the block level between the EVA arrays, it reduces the time it takes to transfer and recover data, getting your business back online faster in the face of a disaster or data loss. For asynchronous replication, RPO depends on network bandwidth, latency, application load, and replication log size and utilization.

Table 6.

Application	RPO	RTO	Average application response time	Round trip inter-site link latency	Distance between sites	Type of data replication	Local or Primary site	Remote site	Inter-site link bandwidth	Auto failover
Oracle 10g with ASM, I TB database, OLTP Nature	0	30 min	35 ms	20	500 Km	HP Continuous Access Synchronous	HP EVA 8x00, HP MPR 400	HP EVA 8x00, HP MPR 400	OC3 (155 Mb/s)	No
Oracle 10g with ASM, I TB database, OLTP Nature	30 min	30 min	25 ms	40	1000 Km	HP Continuous Access Asynchronous	EVA 8x00, MPR 400	HP EVA 8x00, HP MPR 400	OC3 (155 Mb/s)	No

Required hardware and software

	Model	Part number
	HP StorageWorks Replication Solutions Manager Media Kit	T3680E
HP StorageWorks EVA4000	HP Continuous Access EVA4000 1TB LTU	T5356A
	HP Continuous Access EVA4000 Unlimited use per EVA LTU (recommended)	T5357A
	HP Continuous Access EVA4000 Upgrade to Unlimited LTU	T5358A
HP StorageWorks EVA6000	HP Continuous Access EVA6000 1TB LTU	T5359A
	HP Continuous Access EVA6000 Unlimited use per EVA LTU (recommended)	T5360A
	HP Continuous Access EVA6000 Upgrade to Unlimited LTU	T5361A
HP StorageWorks EVA8000	HP Continuous Access EVA8000 1TB LTU	T5362A
	HP Continuous Access EVA8000 Unlimited use per EVA LTU (recommended)	T5363A
	HP Continuous Access EVA8000 Upgrade to Unlimited LTU	T5364A
HP StorageWorks 400 MPR base	18 ports (16 Fibre Channel and 2 Gigabit Ethernet) multi-protocol router providing 2 types of SAN services:	AG458A
B-series FCIP MPR LTU	Optional software license to activate the FCIP SAN services	T4425A
HP StorageWorks 4/256 SAN Director	256-port capable Fibre Channel Director with 2 control processors, 2 power supplies, rack rails, Zoning, Web tools. Does not include Port blades or SFPs	A7988A
HP StorageWorks Multi-protocol Router Blade for B-Series	18 ports (16 Fibre Channel and 2 Gigabit Ethernet) multi-protocol router providing A maximum of two BP blades supported per Director chassis	AG461A
Director FCIP LTU	HP StorageWorks B-Series FCIP Blade LTU: Optional software license to activate the FCIP SAN services in MP blade. It includes the Encryption Services License.	T4427A
HP StorageWorks Cluster Extension EVA	HP StorageWorks Cluster Extension EVA Windows® LTU	T3667A
	HP StorageWorks Cluster Extension EVA Linux LTU	T4393A
HP StorageWorks Linux Cluster for CLX	HP ServiceGuard Extended Distance Cluster for Linux	T2376A

Ordering information

For information on how to buy HP products: http://welcome.hp.com/country/us/en/howtobuy.html

For details on how to order HP StorageWorks EVA disk arrays: www.hp.com/go/eva

Why HP

With our extensive experience in business continuity, storage, and information lifecycle management solutions, HP serves as a single source of accountability for getting the job done—from planning to deployment to ongoing management, with the business continuity solutions that match your budget and business needs.

- **Breadth of offerings**—HP is your one-stop shop for all your solutions needs—from storage to services and networking. HP is the only vendor you need to work with to achieve your IT goals.
- **Proven track record**—As a leader in business continuity and availability for more than 25 years, HP has helped enterprise customers build and manage highly available IT infrastructures.
- **Expertise**—Our certified business continuity consultants, availability specialists, solution architects, and project managers are among the most highly certified, technically proficient in the industry. Our global team includes more than 5,000 high-availability experts. We have more than 50 recovery centers and 13 mobile units operating across 37countries.
- **Global resources**—HP has one of the largest IT customer support organizations in the world, serving 160 countries along with leading high-availability labs for designing and testing solutions on four continents.
- **Comprehensive, integrated approach**—We look at business continuity and availability in an integrated way, and address all points of operational risk across.

With thousands of Business Continuity/Disaster Recovery (BC/DR) customers around the world, HP has the experience and expertise to help you architect and deploy a robust solution that meets your recovery objectives, cost-effectively. Whether your company is in the financial, insurance, healthcare, transportation, or manufacturing markets and whether your company is a Fortune 100 or mid-sized enterprise, you can rely on the fully tested and qualified business protection solutions from HP to ensure availability and protection of your information assets.

For more information

- HP StorageWorks Continuous Access Planning Guide
  www.hp.com/go/caeva
- HP StorageWorks SAN Design Guide
  www.hp.com/go/sandesignguide
- HP StorageWorks Customer Focused Testing
  www.hp.com/go/hpcft
- HP StorageWorks EVA documentation
  www.hp.com/go/eva
- HP StorageWorks CA documentation
  www.hp.com/go/caeva
- HP StorageWorks CLX documentation
  www.hp.com/go/clxeva
- HP StorageWorks MetroCluster/ContinentalCluster documentation
  www.hp.com/go/dt
- HP StorageWorks SAN documentation
  www.hp.com/go/san
- HP StorageWorks B-Series Remote Replication Solution Best Practices Guide
  www.hp.com/go/bseriesreplication
- Disaster Proof video and other information
  www.hp.com/go/DisasterProof

4AA1-2524ENW, June 2007

invent