HP StorageWorks Cluster Extension XP user guide

XP48
XP128
XP512
XP1024
XP12000

Product Version: 2.05.00

sixth edition (August 2004)

T1609-96004

This guide explains how to use the HP StorageWorks Cluster Extension XP software.

invent

Hewlett-Packard Company makes no warranty of any kind with regard to this material, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. Hewlett-Packard shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance, or use of this material.

All product names mentioned herein may be trademarks of their respective companies.

Hewlett-Packard Company shall not be liable for technical or editorial errors or omissions contained herein. The information is provided "as is" without warranty of any kind and is subject to change without notice. The warranties for Hewlett-Packard Company products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.

Printed in the U.S.A.

HP StorageWorks Cluster Extension XP User Guide

product version: 2.05.00
sixth edition (August 2004)
part number: T1609-96004

Contents

About this guide

This guide provides information about using and configuring HP StorageWorks Cluster Extension XP in an environment where clustered systems are connected to a disaster recovery array-based mirroring solution. Cluster Extension XP allows creation of dispersed multiplatform cluster configurations with the XP disk array. Cluster Extension XP enables cluster software to automatically failover applications where data is stored and continuously mirrored from a local to a remote disk array using HP StorageWorks Continuous Access XP. This guide describes the options you have to make your disaster tolerant environment as robust as possible to keep your data available at all times.

Because the XP family disk arrays supports a broad range of operating systems and cluster software, Cluster Extension XP can be integrated with almost any disk array-supported cluster software. This guide provides you with the information you need to create a two or more data center disaster tolerant environment utilizing the XP disk array and its Continuous Access XP remote mirroring feature.

Unless otherwise noted, the term *disk array* refers to these disk arrays:

HP Surestore Disk Array XP512
HP Surestore Disk Array XP48
HP StorageWorks Disk Array XP128
HP StorageWorks Disk Array XP1024
HP StorageWorks XP12000 Disk Array

Intended Audience

This guide is intended for system administrators who maintain the cluster environment and storage subsystems and have the following knowledge:

- A background in data processing and direct-access storage device subsystems and their basic functions.
- Familiarity with disk arrays and RAID technology.
- Familiarity with the operating system, including commands and utilities.
- A general understanding of cluster concepts and the cluster software used in the data center environment.
- Familiarity with related disk array software programs:

  HP StorageWorks Continuous Access XP
  HP StorageWorks RAID Manager XP

Disk array firmware and software dependencies

The features and behavior of failover operations depend on the XP firmware and RAID Manager XP versions. This guide describes Cluster Extension XP behavior based on features implemented in the latest XP firmware and RAID Manager XP versions.

Related information

For information about the disk arrays, please refer to the owner’s manuals.

For related product documentation, see the HP web site (www.hp.com):

HP StorageWorks RAID Manager XP: User’s Guide

HP StorageWorks Continuous Access XP: User’s Guide

HP StorageWorks Business Copy XP: User’s Guide

HP StorageWorks Command View XP: User’s Guide

HP StorageWorks Disk Array XP Operating System Configuration Guide: IBM AIX

HP StorageWorks Disk Array XP Operating System Configuration Guide: Sun Solaris

HP StorageWorks Disk Array XP Operating System Configuration Guide: Windows 2000/2003

- *HP StorageWorks Disk Array XP Operating System Configuration Guide: Linux*

For information about Serviceguard for Linux, see the HP High Availability web site:

docs.hp.com/hpux/ha/

For information about RS/6000 and HACMP, see the IBM web site:

www.rs6000.ibm.com/aix/library

For VERITAS Cluster Server information, see the VERITAS web site:

support.veritas.com

For Microsoft Cluster service information, see the Microsoft web site:

Windows 2000:
www.microsoft.com/windows2000/library/technologies/cluster/default.asp

Windows 2003:
www.microsoft.com/windowsserver2003/library/technologies/clustering/default.mspx

Terminology

This guide uses terminology to describe cluster-specific and disaster recovery-specific processes. Vendors of cluster software use different terms for the components of their cluster software. To standardize the usage among vendors, this guide uses the following terms:

application service This is the unit of granularity for a failover or failback operation. It includes all necessary resources that must be present and which the application depends on. For example, a file share must have a disk, a mount point (or drive letter) and an IP address to be considered an application service. A disk is a necessary resource for the application service. Depending on the cluster software, application services can depend on each other and run in parallel on the same system or on different systems.
Vendor equivalent terms
VCS: service group
HACMP: resource group
Microsoft Cluster service: resource group
SG-LX (Serviceguard): package

resource The smallest unit in an application service. It describes the necessary parts to build an application service. The implementation of such resources in cluster software is vendor-specific. Some vendors (such as IBM or HP) do not allow accessing the chains between dependent resources.
Vendor equivalent terms
VCS: resource
HACMP: resource group

Microsoft Cluster service: resource
SG-LX (Serviceguard): package

startup
shutdown

Startup and shutdown are also known as “bringing online” and “taking offline,” or “start” and “stop,” or “run” and “halt” in regards to an application service or resource. Only a few cluster software vendors (such as Veritas or Microsoft) offer starting and stopping of single resources.

Conventions

This guide uses the following text conventions.

Figure 1

Blue text represents a cross-reference. For the online version of this guide, the reference is linked to the target.

www.hp.com

Underlined, blue text represents a website on the Internet. For the online version of this guide, the reference is linked to the target.

literal

Bold text represents literal values that you type exactly as shown, as well as key and field names, menu items, buttons, file names, application names, and dialog box titles.

variable

Italic type indicates that you must supply a value. Italic type is also used for manual titles.

`input/output`

Monospace font denotes user input and system responses, such as output and messages.

Example

Denotes an example of input or output. The display shown in this guide may not match your configuration exactly.

[ ]

Indicates an optional parameter.

{ }	Indicates that you must specify at least one of the listed options.
|	Separates alternatives in a list of options.

HP storage website

For the most current information about HP StorageWorks XP products, visit the support website. Select the appropriate product or solution from this website:

http://h18006.www1.hp.com/storage/arraysystems.html

For information about product availability, configuration, and connectivity, consult your HP account representative.

HP authorized reseller

For the name of your nearest HP authorized reseller, you can obtain information by telephone:

United States	1-800-345-1518
Canada	1-800-263-5868
Or contact:	www.hp.com

Revision history

February 2001	First release.
March 2001	Added command-line interface chapter.
July 2001	Added MSCS support.
November 2001	Added quorum filter-service for MSCS on XP512/XP48.
May 2002	Updated content for version 1.03 of all Cluster Extension products.

	Updated content for version 1.04.00 of Cluster Extension for MSCS. Added support for Serviceguard on Linux. Updated content for version 1.1 of Cluster Extension XP quorum service with external arbitrator.
September 2002	Updated content for version 2.00. Changed product terminology from *MSCS* to *Microsoft Cluster service*. Added arguments for **clxchkmon**. Changed **LogLevel** values. Changed Windows log file directory location. Added message catalog.
December 2002	Updated content for version 2.01 for VCS and Serviceguard. Added rolling disaster protection features. Added GUI features.
January 2003	Updated content for version 2.01 for Windows GUI.
April 2003	Updated content for version 2.02. Added “Cluster Extension XP quorum service message catalog” (page 261).
November 2003	Updated for versions 2.02 and 2.03. Added SUSE Linux and Windows 2003 support. Removed XP256. Changed *MC/ServiceGuard* to *Serviceguard*.
March 2004	Modified document for version 2.04.00.
August 2004	New format applied. Modified document for version 2.05.00

Warranty statement

HP warrants that for a period of ninety calendar days from the date of purchase, as evidenced by a copy of the invoice, the media on which the Software is furnished (if any) will be free of defects in materials and workmanship under normal use.

DISCLAIMER. EXCEPT FOR THE FOREGOING AND TO THE EXTENT ALLOWED BY LOCAL LAW, THIS SOFTWARE IS PROVIDED TO YOU "AS IS" WITHOUT WARRANTIES OF ANY KIND, WHETHER ORAL OR WRITTEN, EXPRESS OR IMPLIED. HP SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OR CONDITIONS OF MERCHANTABILITY, SATISFACTORY QUALITY, NON-INFRINGEMENT, TITLE, ACCURACY OF INFORMATIONAL CONTENT, AND FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow exclusions of implied warranties or conditions, so the above exclusion may not apply to you to the extent prohibited by such local laws. You may have other rights that vary from country to country, state to state, or province to province.

WARNING! **YOU EXPRESSLY ACKNOWLEDGE AND AGREE THAT USE OF THE SOFTWARE IS AT YOUR SOLE RISK.** HP DOES NOT WARRANT THAT THE FUNCTIONS CONTAINED IN THE SOFTWARE WILL MEET YOUR REQUIREMENTS, OR THAT THE OPERATION OF THE SOFTWARE WILL BE UNINTERRUPTED, VIRUS-FREE OR ERROR-FREE, OR THAT DEFECTS IN THE SOFTWARE WILL BE CORRECTED. THE ENTIRE RISK AS TO THE RESULTS AND PERFORMANCE OF THE SOFTWARE IS ASSUMED BY YOU. HP DOES NOT WARRANT OR MAKE ANY REPRESENTATIONS REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE OR RELATED DOCUMENTATION IN TERMS OF THEIR CORRECTNESS, ACCURACY, RELIABILITY, CURRENTNESS, OR OTHERWISE. NO ORAL OR WRITTEN INFORMATION OR ADVICE GIVEN BY HP OR HP'S AUTHORIZED REPRESENTATIVES SHALL CREATE A WARRANTY.

LIMITATION OF LIABILITY. EXCEPT TO THE EXTENT PROHIBITED BY LOCAL LAW, IN NO EVENT INCLUDING NEGLIGENCE WILL HP OR ITS SUBSIDIARIES, AFFILIATES, DIRECTORS, OFFICERS, EMPLOYEES, AGENTS OR SUPPLIERS BE LIABLE FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, CONSEQUENTIAL, PUNITIVE OR OTHER DAMAGES (INCLUDING LOST PROFIT, LOST DATA, OR DOWNTIME COSTS), ARISING OUT OF THE USE, INABILITY TO USE, OR THE RESULTS OF USE OF THE SOFTWARE, WHETHER BASED IN WARRANTY, CONTRACT, TORT OR OTHER LEGAL THEORY, AND WHETHER OR NOT ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Your use of the Software is entirely at your own risk. Should the Software prove defective, you assume the entire cost of all service, repair or correction. Some jurisdictions do not allow the exclusion or limitation of liability for incidental or consequential damages, so the above limitation may not apply to you to the extent prohibited by such local laws.

NOTE. EXCEPT TO THE EXTENT ALLOWED BY LOCAL LAW, THESE WARRANTY TERMS DO NOT EXCLUDE, RESTRICT OR MODIFY, AND ARE IN ADDITION TO, THE MANDATORY STATUTORY RIGHTS APPLICABLE TO THE LICENSE OF THE SOFTWARE TO YOU; PROVIDED, HOWEVER, THAT THE CONVENTION ON CONTRACTS FOR THE INTERNATIONAL SALE OF GOODS IS SPECIFICALLY DISCLAIMED AND SHALL NOT GOVERN OR APPLY TO THE SOFTWARE PROVIDED IN CONNECTION WITH THIS WARRANTY STATEMENT.

1

Cluster Extension XP features

The quest to extend high availability over geographically dispersed locations has driven today's IT personnel to demand cluster solutions capable of recovering from even the most extensive disasters. HP StorageWorks Cluster Extension XP enables you to monitor HP StorageWorks Continuous Access XP-mirrored disk pairs and allows access to the remote data copy if the application becomes unavailable on the local site. If the application service is restarted on the remote site, after the local (primary) application service has been shut down, Cluster Extension XP uses its internal database to check whether the current disk states allow automatic access to your data based on consistency and concurrency considerations. Integrated in the cluster software or available as command-line interface for your own integration, Cluster Extension XP ensures that the data can be accessed if necessary.

Cluster Extension XP software provides these key features:

- integration into cluster software
- disaster tolerance through geographical dispersion
- automated redirection and monitoring of mirrored Continuous Access XP pairs
- command-line interface for easy integration

Integration into cluster software

The value of Cluster Extension XP is to provide tight integration into the cluster software, wherever possible. Cluster Extension XP is a resource to the clustered application service (like the disk or volume group) and must therefore be managed as such. The architecture of Cluster Extension XP allows integration into many cluster software products, including these:

- VERITAS Cluster Server (VCS)
- IBM HACMP
- Windows 2000 Advanced Server and Datacenter Server Cluster service
- Windows Server 2003 Enterprise Edition and Datacenter Edition
- Serviceguard for Linux (SG-LX)

For the current list of supported cluster software, contact your HP representative.

Disaster tolerance through geographical dispersion

Using two or more disk arrays with Continuous Access XP allows you to copy your most valuable data to a remote data center. Cluster Extension XP provides the cluster software with a mechanism to check and allow data access (in case the local application service must be transferred to a remote cluster system). The distance to your remote location is only limited by the technology your cluster software uses to communicate with each system in the cluster, the technology you use for physical data replication, and the degree of failover automation.

Disaster tolerance considerations

Application availability is essential for today's businesses. The capability to restore the application service after a failure of the server, storage or the whole data center in a timely fashion is a must and is considered as *disaster tolerance*. Complete data center failures can be caused by earthquakes or hurricanes but more often they are caused by power outages or fires.

To protect against such disasters, a single data center is not sufficient. Systems (storage as well as the servers) must be geographically distributed in order to build a disaster tolerant architecture which protects against planned and unplanned downtimes.

Of course, redundant network cards and storage host bus adapters are a basic requirement. The same applies for the power supplies of both the storage and the server. With this hardware in place, the external power service and the network must also be designed to provide no single point of failure (SPOF).

Today, data is the most valuable asset in your enterprise. The XP family of disk arrays provides a fully redundant architecture, and the flexibility to upgrade firmware online reduces the risk of unplanned and planned downtime. The disk array also provides the feature of remotely mirroring your data to a second disk array.

To have this expensive hardware in place must be compared to the risk of a true disaster. The costs pay off in a real disaster to ensure that the business critical applications are still accessible from another location.

Guidelines The following considerations, applied to the cluster environment, can ensure an application service survives a disaster with minimal downtime.

- geographical dispersion of hardware and applications
- redundant paths to access the network and storage
- alternative power sources
- redundant networks
- data replication

Several ways of implementing such disaster tolerant architectures are possible. All of those solutions can be covered by a clustered solution using the XP family of disk arrays and Continuous Access XP. Cluster Extension XP is needed to enable access to your critical data.

Disaster-tolerant architectures

With Cluster Extension XP, you can extend your cluster solution beyond the limitations of existing data center and campus-wide distances. Cluster Extension XP enables metropolitan-wide failover capabilities, and beyond.

Having a local disk array in each data center also means that the server does not have to write twice because Continuous Access XP mirrors each write-IO to the remote site and therefore relieves the server of the burden, preventing performance bottlenecks.

Disaster tolerant architectures using data replication over the network

Data replication over the network is a way to achieve disaster tolerance and is considered *logical replication*.

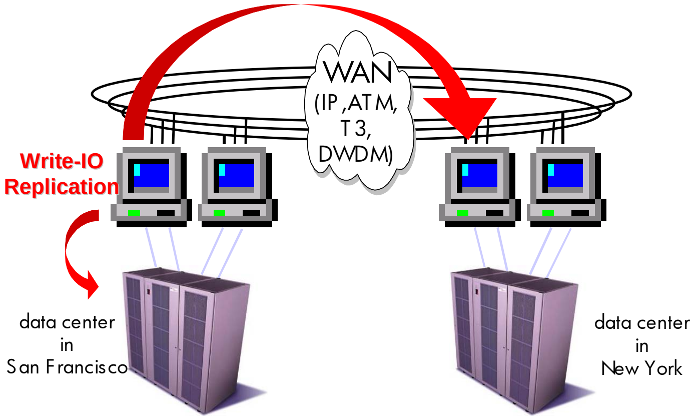


Figure 1. Logical replication over networks

Logical replication uses specific host-based software to write data to local disks and also to replicate that data to a remote system connected to an attached storage device. Because data is replicated over the network, there is no distance limitation for such solutions.

Logical replication techniques imply that the failover process is mainly manual. This means each site belongs to a different cluster, or only the primary site is clustered, while the secondary site acts as a standby system. It is also possible that no cluster software is involved and that only one system is available at each site.

Data replicated over the network can be at the granularity of a single volume, a file system, or a transaction.

All logical replication techniques have some significant disadvantages: The remote system is a standby system. That is, it must perform the same task as the primary system and cannot be used for any other purpose. If the standby system is activated, it must replay redo logs first and cannot automatically serve as a replication source (for example, Oracle’s standby database implementation).

Another significant disadvantage of such architectures is that the server must write every IO twice, once to the attached storage device and once to the remote system over the network. These replication techniques can only be implemented asynchronously; otherwise, the application experiences noticeable performance degradation.

Because of the nature of replication products, additional CPU power is necessary to mirror write requests.

Logical replication implies that all logs, which have not been shipped (or which are in transit) are lost in case of a disaster.

Disaster-tolerant architectures using Fibre Channel networks

Disaster tolerant architectures using Fibre Channel networks can be achieved by the use of *physical replication*.

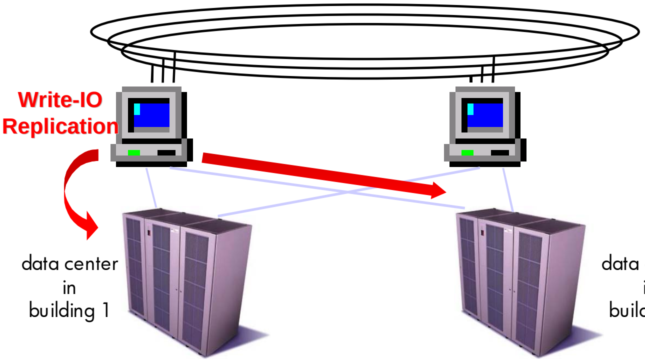


Figure 2. Physical replication using Fibre Channel

As with logical replication products, physical replication often uses host-based software to replicate data. Here, data is written to server-attached storage devices twice. Most of today's logical volume management products offer this feature.

Using Fibre Channel, you could use dual-attached storage devices, where one port is connected to the local server and one is connected to the remote server. To be able to access your data at the remote location in case of a disaster, each server must have a local and a remote storage device. The volume management software, then, must be set up to mirror each write request to both the local and the remote storage device. With the XP disk arrays, several servers can be connected to each disk array.

This solution is called *campus cluster*. A single cluster can be used and the failover process can be automated. With campus clusters, both sites can be active.

Data replicated via volume mirroring is based on the granularity of a single volume.

Campus cluster solutions are limited to the distance Fibre Channel supports today. While storage systems must be in a range of 500 meters (direct connect) or up to 10 kilometers (connected via Fibre Channel switches or Fibre Channel hubs), campus cluster solutions can only offer limited protection against natural disasters.

Another limiting factor is the cluster heartbeat protocol or the communications protocol used for cluster reformation processes. Those protocols are vendor-specific implementations and require private networks. This means, those protocols are not routable. The distance limitations of a private network depends on the supplied network infrastructure and latency issues of the heartbeat or cluster reformation protocol.

Another significant disadvantage of such architectures is that the server must write every IO twice: once to the locally attached storage device and once to the remote attached storage device. These replication techniques are implemented as software running on the server, which reduces the available compute power and degrades server performance.

Because of the nature of volume mirroring products, additional CPU power is necessary to mirror write requests across two host bus adapters.

Most of these products have another significant disadvantage. In case of a path failure, the whole volume must be copied to resynchronize the second volume with the current state of the first volume. If the storage device must be replaced, all volumes must be copied. This significantly affects server performance.

Disaster-tolerant architectures using disk array-based mirroring

Using Continuous Access XP-based mirroring is also considered *physical replication*. Continuous Access XP is disk array-based mirroring. As with campus clusters, such solutions require two or more disk arrays.

The key difference from the above-mentioned solutions is that the disk array keeps track of the data integrity of the mirrored disks. XP disk arrays offer RAID-1 or RAID-5 protection as a standard feature and allow online addition and replacement of disks, IO adapter cards, and memory. To provide copies of data, internal and external mirroring features are available. For disaster tolerant solutions, Continuous Access XP can mirror your data with no distance limitation.

ESCON or Fibre Channel protocol is used to transfer data between two disk arrays. Using converters, ESCON and Fibre Channel can be routed over IP networks and T3 to allow unlimited distance between the disk arrays. To replicate data over more than 0.5 km (Fibre Channel) or 3km (ESCON), special extenders or switches must be purchased.

The cluster solutions using Continuous Access XP-based disk mirroring are called *metropolitan clusters* or *geographically dispersed clusters*. Servers are members of the same cluster dispersed over two or more sites. Since the disk array controls the replication process, the server is relieved from writing any IO-request to the disk more than once.

Continuous Access XP-mirrored disks typically have a read/write-enabled primary disk and a read-only secondary disk. This leads to problems because current cluster software products cannot distinguish between write-protected and write-enabled disks.

Cluster software assumes that the application service has access to read/write-enabled data disks on any system that the application service has been configured to run. Since the secondary volume of a disk pair is not normally accessible, the failover process would typically involve manual intervention.

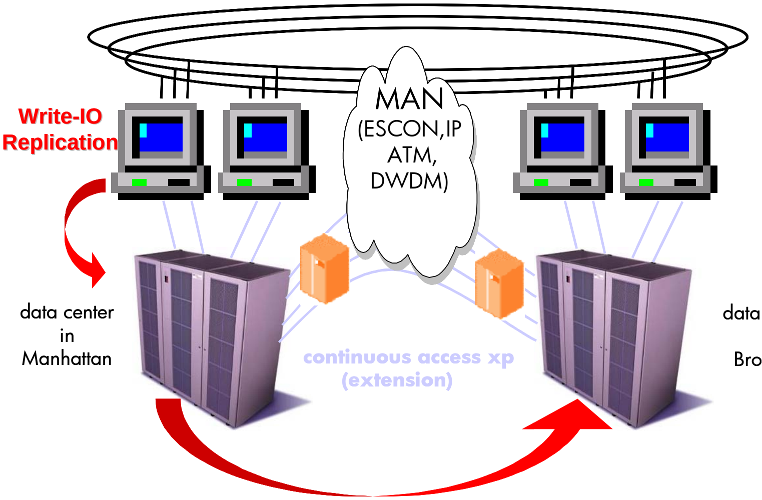


Figure 3. Physical replication using HP StorageWorks Continuous Access XP

Cluster Extension XP provides the software to enable automated failover and failback procedures integrated as a resource of the application service. Cluster Extension XP uses an internal database to decide whether the data on the failover site is safe to be accessed or not. Manual intervention is required if the current disk array states and the user settings conflict with the rules stored in the Cluster Extension XP internal database.

The limiting factor of metropolitan or geographically dispersed clusters is the cluster heartbeat protocol or the communications protocol used for the cluster reformation processes. Those protocols are vendor-specific implementations and typically require private networks. These protocols are not routable; a router cannot be used. The distance limitations of networks supporting a private network are dependent on the supplied network infrastructure and latency issues of the heartbeat or cluster reformation protocol.

To address these issues, cluster manager software can be used. This software offers disaster tolerance by managing two or more clusters from a single console or server and is considered a *continental cluster*. Depending on the implementation, *automated* or *semiautomated* failover processes between clusters are possible.

As mentioned above, metropolitan or geographically dispersed clusters as well as continental clusters require metropolitan area networks or wide area networks. In most cases, those network connections involve common carriers and special network equipment which can be very expensive. The reliability of a direct connection or a campus network can be degraded and involves more planning to deploy and maintain a disaster tolerant environment.

Using Continuous Access XP, data is accessible and consistent in every failover case and the resynchronization of a completely failed disk array can be done while the application is running with almost no impact to the server performance. This allows reestablishing disaster tolerance without application downtime.

Automated redirection and monitoring of mirrored Continuous Access XP pairs

Disk arrays with Continuous Access XP provide a unique feature that allows the redirection of the mirroring destination. This means Continuous Access XP almost instantaneously swaps the primary/secondary relationship of disk pairs if the application must access the secondary disk. This feature ensures that the disk pairs are always synchronized, ensuring that the failback process is as fast as the failover process. If the links between your disk arrays are broken, each array maintains a bitmap table to synchronize the changed, delta data if the links become available again. In a failover case, Cluster Extension XP takes the appropriate action for each link/array status and makes sure that your application service has the latest data.

Cluster Extension XP includes a pair/resync monitor to monitor the health of the links between your arrays. Furthermore, it detects a lost and later reestablished link and automatically resynchronizes the suspended disk pairs, ensuring the most current data is available on either site.

Rolling disaster protection

Rolling disaster protection minimizes the impact of downtime and ensures data integrity during recovery operations. Rolling disaster protection combines Continuous Access XP remotely mirrored disk pairs and internal Business Copy XP disk copies to protect data locally as well as remotely. In combination, these features support the highest data protection levels to prevent disastrous loss of data.

What is a rolling disaster?

A rolling disaster refers to catastrophic events or outages that affect the data stored on remote mirrored disk pairs. In a rolling disaster, data stored on remote mirrored disk pairs can be entirely lost during a recovery attempt.

In a rolling disaster, the mirrored disk pairs typically experience the following sequence of events:

1. The primary data center failed.

   The cluster software successfully transferred application execution to the remote data center.

2. The Continuous Access XP link failed.
3. The secondary volume of the disk pair is used to continue operation after failover while the CA link is not functional.

   The secondary volume represents the latest state of data, whereas the data on the primary volume is now out of date.

4. The primary data center is recovered and the Continuous Access XP link is restored.

5. A recovery operation is initiated to resynchronize (update) the original (primary) disk from the secondary disk.

   This is known as a restore operation after a disaster, or a restore after failover operation.

   The resynchronization/restore operation can take minutes to days depending on the amount of data that must be updated and transferred between the two disk arrays.

   During the recovery operation, data is vulnerable to the effects of a disaster or outage. During a resynchronization operation, data updates are sent in the order of changed tracks and not in the transactional order in which the data was originally written or acknowledged.

6. The secondary site fails during the resynchronization/restore operation.

   The restored data at the original, primary site becomes unusable.

Although resynchronization operations are possible while an application is running, resynchronization could lead to unrecoverable data if a rolling disaster occurs. This type of rolling disaster can occur in the following circumstances:

- during manual resynchronization attempts
- during failover operations using Cluster Extension XP in a cluster environment when the Cluster Extension XP **AutoRecover** object is set to **yes,** or where the pair/resync monitor is used with the **ResyncMonitorAutoRecover** object set to **yes**.

Recovering the disaster tolerant environment

To ensure survival of critical data during a resynchronization/restore operation, Cluster Extension XP supports the use of preconfigured Business Copy disks and allows suspending any number of Business Copy pairs that can be associated with the primary data disks. Cluster Extension XP recovers automatically, provided that at least one internal Business Copy mirror could be suspended to guarantee a recoverable state.

Cluster Extension XP also resumes internal Business Copy mirrors automatically, if specified, to allow the local site to keep an up-to-date image of the data.

This internal copy represents the state of the primary volume before the data center failure. This copy is needed to survive a possible failure of the secondary volume or disk array during the resynchronization operation. Although the data could be out of date, it represents the best starting point for the recovery effort, unlike the inconsistent data that results from a rolling disaster.

Recovery from a consistent, point-in-time copy ensures the integrity of data and eliminates the need for full tape restore procedures. Rolling disaster protection provides a rapid recovery method and so minimizes downtime.

Figure 5 (page 49) illustrates an example of a disaster-tolerant configuration.

To implement rolling disaster protection, see “Rolling disaster protection and Business Copy XP” (page 45).

Command-line interface for easy integration

Cluster Extension XP provides you with a command line interface to enable disaster tolerant environments even if no cluster software is available for your operating system. This feature is convenient if you use in-house-developed software to migrate application services from one system to another or if you want Cluster Extension XP to check the disk states to make sure you can automatically start your application service on the local disk array.

Graphical user interface

Cluster Extension XP for Microsoft Cluster service and VCS can be configured with the cluster software GUI. Both cluster software products provide a graphical user interface to set and change resource values. Cluster Extension XP offers full integration into the GUI so that you can utilize the capacity of your cluster software.

Quorum service *(Microsoft Cluster service only)*

Microsoft Cluster service depends on the cluster quorum disk resource to maintain a persistent log of cluster configuration changes and status, as well as a single point to resolve any possible events that could result in a split brain situation. The Cluster Extension XP quorum service adds an additional dimension to disaster tolerance by remotely mirroring the quorum disk resource, thus preventing it from being the single point of failure.

The quorum service allows the quorum disk resource to be mirrored between dispersed sites and supports the movement and failover of the quorum disk between the two sites without disrupting other cluster services. The external arbitrator (also included in Cluster Extension XP) solves the potential split brain syndrome as well. This feature significantly increases the availability of the critical quorum disk resource, thus reducing the possibility of cluster failure due to the loss of the quorum disk.

The quorum service and Cluster Extension have been certified to fulfill all requirements for Microsoft cluster. For certified configurations, see the Microsoft web site:

www.microsoft.com/windows/catalog/server

1. Click the Hardware tab.
2. Select Cluster Solutions from the left side menu.
3. Select Geographically Disposed Cluster Solution.

2

Cluster Extension XP processes and components

Cluster Extension XP is shipped in the appropriate format for each platform:

Platform	Implementation
VCS	agent
IBM HACMP	pre-event executable
Microsoft Cluster service	resource DLL, quorum service, and external arbitrator
SG-LX	function call/executable

Customized solutions to failover application services must implement Cluster Extension XP through its command-line interface prior to the disk activation procedure.

Cluster Extension XP environments

The ideal environment for a Cluster Extension XP configuration consists of at least four servers (two at each site) and separated redundant communications links for cluster heartbeats, client access and Continuous Access XP (Extension). All communications interfaces must be installed in pairs to serve as failover components, preventing single points of failure (SPOFs).

Recommendation Use load balancing and alternative pathing software for host-to-storage connections, such as HP StorageWorks Auto Path for IBM AIX or Secure Path for Linux and Windows operating systems. For Sun Solaris operating systems, VERITAS offers such software. These software products enable you to upgrade XP firmware while the application service is running.

Network communications links between the dispersed data centers must be redundant and physically routed differently. This prevents the “backhoe issue,” that is, where all links between data centers are cut together. This is especially important, since the cluster is more vulnerable to “split brain” syndromes. A split brain syndrome is where both data centers’ systems form new clusters which could allow access to both copies of the data. This can be prevented with physically separated network links and redundant network components. Cluster Extension XP allows you to configure the failover behavior in such a way that the application service startup procedure will be stopped if none of the remote cluster members can be reached. The default configuration of Cluster Extension XP expects the cluster software to deal with the “split brain” syndrome.

Since the disk array stores your most valuable data, this data must get across to the remote disk array. At least four Continuous Access XP links must be available when the disk arrays are connected directly and are configured for bidirectional takeover. For extended distances, extender components must be purchased. These components are able to bundle Continuous Access XP links. At least two links are necessary to provide redundancy and protection against single points of failure. Although communications links can cover considerable distances, each network segment must be extended to the dispersed data center in order to maintain a heartbeat among all servers.

Recommendation

Use four systems to give local application service failover among local cluster systems priority over remote, more time-consuming failover procedures. When failing over, Cluster Extension XP must reconfigure the disk arrays to change the mirroring direction. This takes more time than just checking for the correct disk array disk states. On the remote site, two systems should be available in the case the failover system experiences a hardware or power failure.

Figure 2 (page 24) depicts a preferred Cluster Extension XP configuration.

Caution

Cluster Extension XP works with only one system at each location, with a single I/O path between the server system and the disk array and a single link in each direction between disk arrays.

However, those configurations are not considered highly available, nor are they disaster tolerant. Therefore, Cluster Extension XP configurations with single points of failure are not supported by HP.

Cluster Extension XP execution

Cluster Extension XP requires cluster software to automatically fail over and fail back among systems on a local site or between sites. Cluster Extension XP must manipulate the application startup process before disk array disks are activated. Cluster Extension XP, therefore, must be integrated as first resource (in the order of resources). To activate Continuous Access XP paired disk devices, the paired disk devices must be in read/write mode. Continuous Access XP disks are usually in read/write mode on the primary disk only; the secondary disk is in read-only mode. In case of a failover, the direction of the mirrored pair is changed by Cluster Extension XP automatically. In case of a disaster, the disk array can have several different states for disks in a RAID Manager XP device group. Cluster Extension XP decides whether those disks can be activated.

Cluster Extension XP must be installed on any server in the cluster that can run the application service in the cluster.

Cluster Extension XP stores information about the application environment in an internal object database and uses RAID Manager XP to gather information about the state of the associated disk pairs. The information about the configured disk array environment and failover behavior is transferred either directly by the cluster software or by gathering from the user configuration file.

The internal object database provides Cluster Extension XP with knowledge about supported parameters, their formats, and default values.

Disk array disk states are stored in an internal object database and a rule engine is used to process those disk states. The rule engine matches current disk states and configuration parameters with a defined rule, stores it in the database, and invokes predefined actions. Those actions prepare the disk array disks to be activated, or it stops the application service startup process if the matching rule requires it to do so.

Continuous Access XP and RAID Manager XP

Continuous Access XP provides remote copy functionality for the disk arrays. Disk arrays can be mirrored to many different remote disk arrays.

Cluster Extension XP does not support two disk arrays as either primary or secondary disk arrays. Cluster Extension XP supports configurations where two (or more) disk arrays use one remote disk array as the failover site. In those cases, the disk array configuration can be considered as a logical one-to-one configuration.

Figure 4 (page 41) depicts an example of a supported configuration.

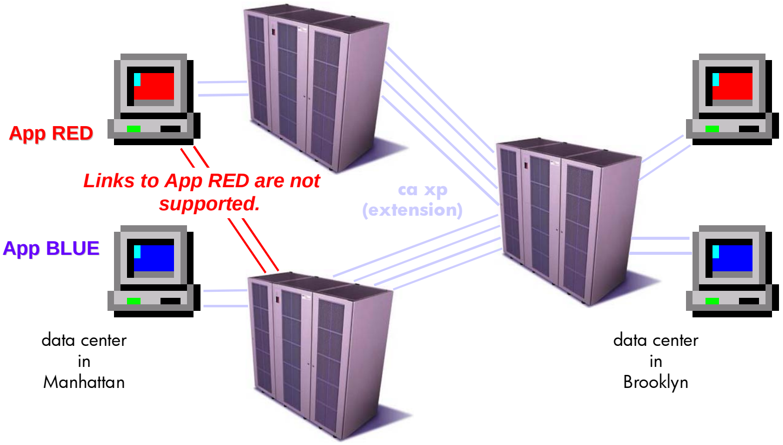


Figure 4. Supported XP disk array configuration

To control Continuous Access XP-mirrored disks from a server, RAID Manager XP must be installed on the server. A special disk, called a *command device*, must be configured to control the paired disks. The special disk must not be part of Microsoft Cluster service resources and cannot be paired. The command device, which is identified by a “CM” appended to the emulation type, can be assigned to a 36-Mbyte or greater CVS volume. RAID Manager XP uses the command device to communicate with the disk array controller (DKC).

Using Continuous Access XP Extension, consistency groups can be configured. Consistency groups are units in which the disk array keeps data consistent among paired disks.

Continuous Access XP links are unidirectional links. For disaster tolerant configurations, two links must be provided in each direction. Both sender (RCP) and receiver (LCP) ports must be configured on each redundant IO board used for Continuous Access XP.

Continuous Access XP offers two modes of replication:

- synchronous replication
- asynchronous replication

Synchronous replication

Using synchronous mode, all write requests from the server are first transferred to the remote disk array. After each IO has been mirrored in the cache area of the remote array, it is acknowledged to the local disk array. The write request is then acknowledged to the server.

Synchronous replication modes can be configured in the following fence levels:

NEVER Allows write requests even if the request cannot be replicated to the remote disk array. If a write request cannot be replicated the remote disk array, the area on the disk is marked in a bitmap table and transferred after a resynchronization request has been ordered.

STATUS This fence level is not supported by Cluster Extension XP.

DATA Prohibits write requests immediately if a link failure or disk failure occurs. The local disk array cannot replicate data to the remote disk array. Fence level **DATA** provides data concurrency at any time.

The preceding fence levels provide data integrity on a per disk basis, so a failure affecting a single disk pair does not lead to a halt of the replication activities of non-affected disk pairs.

Synchronous replication can affect the performance of the system if the distance between the disk arrays is significant.

Asynchronous replication

Continuous Access XP Extension offers a unique feature to replicate data asynchronously.

To keep replicated data consistent among two disk arrays, any incoming write request is ordered and numbered. The write request is then acknowledged to the server, offering the fastest response time for remote mirroring. Each write request is transferred to the remote disk array asynchronously. The remote array orders all write requests before they are destaged to the disk, keeping data consistent.

Asynchronous replication offers excellent performance for remote mirroring and provides data consistency on a group of disks (consistency groups) level.

RAID Manager XP instances

A RAID Manger instance is necessary to control pair operations and to gather disk array status information.

The RAID Manager XP instance numbers used for the **RaidManagerInstances** object must be the same among all systems using Cluster Extension XP.

Several RAID Manager XP instances can be configured to provide additional redundancy. Cluster Extension XP switches to the next available instance when an instance becomes unavailable.

The RAID Manager XP instances should be running at all times to provide the fastest failover capability. Cluster Extension XP provides scripts to include the RAID Manager XP startup procedure in the system startup file (for example, **/etc/inittab**). However, Cluster Extension XP starts the configured RAID Manager XP instances if it cannot find any running instance.

Quorum service

The Cluster Extension XP quorum service employs static RAID Manager API calls and therefore is not dependent on a RAID Manager instance.

RAID Manager XP device groups

A single device group must be configured for a service group (VCS), a resource group (HACMP), a cluster group (Microsoft Cluster service), or a package (SG-LX). This device group must include all disks being used for the application service.

The device group is the unit in which the failover/failback operation is being carried out. A device group can contain several volume groups.

Rolling disaster protection and Business Copy XP

To implement rolling disaster protection, you must create Business Copy XP disk pairs for the Continuous Access XP disk pairs locally. BC disk pairs used for rolling disaster protection must be created with the **–m noread** option of the **paircreate** command. This ensures that BC disks are unavailable to other services, because these disks are intended to be used for rolling disaster protection only. The BC SVOLs must be mapped to a backup server and not to the local cluster node. When Cluster Extension XP suspends the BC pairs, they become available to the local server, which could result in duplicated volume or disk group IDs or signatures.

To enable rolling disaster protection with Business Copy XP, set the following objects for data centers A and B:

- **BCEnabledA** page 79
- **BCEnabledB** page 79

When these objects are set to **YES**, rolling disaster protection is enabled and Cluster Extension XP checks whether the configured Business Copy XP disk pairs are in PAIR state. Before initiating the resynchronization operation, Cluster Extension XP suspends specified Business Copy XP disk pairs that are in PAIR state.

If the **BCEnabledA** and **BCEnabledB** objects are set to **YES**, you must configure specific Business Copy XP disk pairs by using MU (mirror unit) numbers. The MU number defines one of the many disk pair relationships you can create with Business Copy XP disk pairs. You can specify any number of MU numbers that are supported by the Business Copy XP software. Disk pair MU numbers are specified by the following objects for data centers A and B:

- **BCMuListA** page 79
- **BCMuListB** page 80

To enable resynchronization of Business Copy XP disk pairs that have been split by Cluster Extension XP, use the following objects for data centers A and B:

- **BCResyncEnabledA** page 79
- **BCResyncEnabledB** page 80

Cluster Extension XP maintains a list of all associated Business Copy XP disk pairs that were in PAIR state before a resynchronization attempt. If pairs were suspended, Cluster Extension XP automatically resynchronizes those disk pairs after the Continuous Access XP remote mirrored disk pairs have been paired. This feature supports automatic resynchronization of locally split BC disk pairs only. You must specify MU numbers for resynchronization by using the following objects for data centers A and B:

- **BCResyncMuListA** page 80
- **BCResyncMuListB** page 80

Caution *If rolling disaster protection is enabled and none of the Continuous Access XP mirrored disk pairs have a Business Copy disk pair that is in PAIR state, Cluster Extension XP returns a global error and you will not be able to activate the application service. Ensure that at least one Business Copy disk pair is in PAIR state.*

You can use the ***forceflag*** *to start the application service. See "Force flag". In this case, Cluster Extension XP disables rolling disaster protection.*

Integration with RAID Manager XP

Rolling disaster protection does not require Business Copy XP disk pairs to be defined in the RAID Manager XP *horcmX.conf* files that are used by Cluster Extension XP. Cluster Extension XP uses the MU number to monitor and control associated Business Copy XP pairs.

However, you must create a RAID Manager XP configuration file to control the Business Copy XP disk pairs, which are outside of Cluster Extension XP control.

The management of Business Copy XP disk pairs is independent of Cluster Extension XP/Continuous Access XP remotely mirrored disk pairs.

Cluster Extension XP uses the MU number to control the Business Copy disk pairs. Therefore, only the RAID Manager XP instances that are configured for Cluster Extension XP are required for rolling disaster protection.

The Rolling Disaster Protection feature cannot suspend Business Copy XP disk pairs on the XP family disk array in the remote data center if the RAID Manager XP instance in the remote data center is not running or not reachable.

Integration with automatic recovery

If the **AutoRecover** object is set to **YES**, Cluster Extension XP automatically resynchronizes the Continuous Access XP disk pairs to update the remote disks. If rolling disaster protection is enabled, it suspends the Business Copy XP disk pair that is attached to the remote Continuous Access XP disk.

If this remote Business Copy XP disk pair cannot be suspended because the remote RAID Manager XP instance is not running or cannot be reached, Cluster Extension XP continues the application service activation (online the Cluster Extension XP resource) without automatic resynchronization of the Continuous Access XP disk pair and without the suspension of the Business Copy XP disk pair.

In this case, the Continuous Access XP disk pair must be recovered manually.

Integration with the pair/resync monitor

If the **ResyncMonitor** object is set to **YES**, Business Copy XP disk pairs are not used when the pair/resync monitor automatically recovers suspended or failed Continuous Access XP disk pairs.

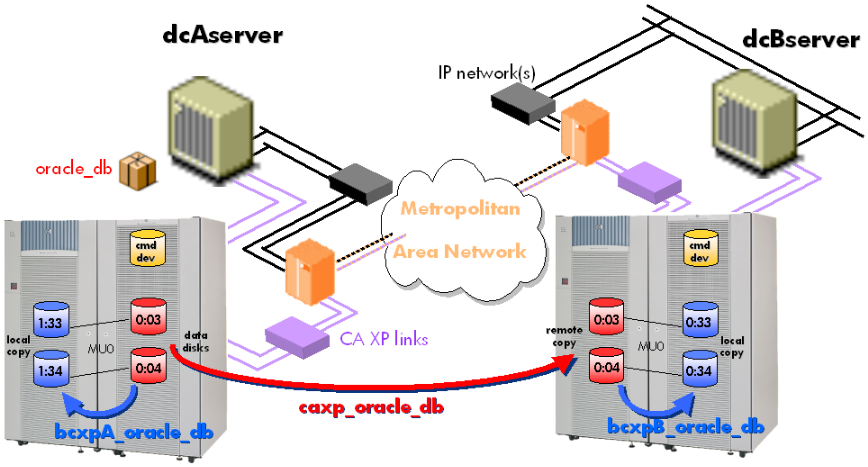
To protect the remote volume of an out-of-sync Continuous Access XP disk pair against rolling disasters, use the default settings for the pair/resync monitor. Resynchronize the Continuous Access XP disk pair manually after splitting off the Business Copy XP disk pair.

Restoring server operation

Rolling disaster protection automatically recovers the PAIR state of the Continuous Access XP disk pair of an application service. Before you failover (or failback) an application service from one data center to the other, you must restore the server operation. After you restart the server, also start the RAID Manager instance used to manage the Continuous Access XP disk pairs on those servers. This enables rolling disaster protection to work correctly during a recovery failover/failback operation.

Example

Figure 5 (page 49) depicts an example of a fully configured Cluster Extension XP environment that uses rolling disaster protection. The Business Copy disk pairs are specified as 0 in the Cluster Extension XP **BCMuListA** and **BCMuListB** objects.



RAID Manager XP configuration file	dcAserver	dcBserver
Instance 101 (horcm101.conf) manages both the CA disk pairs and BC disk pairs	HORCM_MON NONE ca101 1000 1000 HORCM_CMD /dev/sdb /dev/sdc HORCM_DEV caxp_oracle_db ca_003_003 CL1-A 1 7 caxp_oracle_db ca_004_004 CL1-A 2 0 bcxpA_oracle_db bc_003_133 CL1-A 1 7 0 bcxpA_oracle_db bc_004_134 CL1-A 2 0 0 HORCM_INST caxp_oracle_db dcBserver ca101 bcxpA_oracle_db dcAserver bc5	HORCM_MON NONE ca101 1000 1000 HORCM_CMD /dev/sdb /dev/sdc HORCM_DEV caxp_oracle_db ca_003_003 CL1-A 1 7 caxp_oracle_db ca_004_004 CL1-A 2 0 bcxpB_oracle_db bc_003_033 CL1-A 1 7 0 bcxpB_oracle_db bc_004_034 CL1-A 2 0 0 HORCM_INST caxp_oracle_db dcAserver ca101 bcxpB_oracle_db dcBserver bc5
Instance 5 (horcm5.conf) manages the BC disk pairs locally	HORCM_MON NONE bc5 -1 300 HORCM_CMD /dev/sdb /dev/sdc HORCM_DEV bcxpA_oracle_db bc_003_133 CL1-A 3 7 bcxpA_oracle_db bc_004_134 CL1-A 4 0 HORCM_INST bcxpA_oracle_db dcAserver ca101	HORCM_MON NONE bc5 -1 300 HORCM_CMD /dev/sdb /dev/sdc HORCM_DEV bcxpB_oracle_db bc_003_033 CL1-A 3 7 bcxpB_oracle_db bc_004_034 CL1-A 4 0 HORCM_INST bcxpB_oracle_db dcBserver ca101

Figure 5. Example of a disaster-tolerant configuration with rolling disaster protection

User configuration file

Cluster Extension XP provides a user configuration file to customize Cluster Extension XP failover/failback behavior. The user can specify all customizable objects of Cluster Extension XP with this file.

Related information “The user configuration file”

Pair/resync monitor

The pair/resync monitor **clxchkd** utility can be turned on or off with the **ResyncMonitor** object.

The pair/resync monitor can either monitor or both monitor and resynchronize the state of the RAID Manager XP device group for an application service. The cluster software must be able to stop the monitoring or resynchronization process if the application service is stopped.

If the **ResyncMonitorAutoRecover** object is set to **YES**, the monitor tries to resynchronize the remote disk based on the local disk. This occurs only if the disks are in a PVOL/SVOL or SVOL/PVOL relationship. If one or both disk peers are in the state SMPL or the device group state is mixed, automatic resynchronization is not initiated.

The monitor is started from Cluster Extension XP the first time that Cluster Extension XP checks the disk states. Any subsequent execution of the monitor program adds the RAID Manager XP device group to be monitored to the list of to-be-monitored device groups. The monitor interval specified with the **ResyncMonitorInterval** object is used to monitor the device group state. Do not set the monitor interval below the RAID Manager XP timeout parameter (**HORCM_MON** in the **horcm*X*.conf** file).

Caution *If the application service must be stopped, the cluster software or your customized solution to start and stop the application service must be able to stop the monitoring or resynchronization process. If this cannot be ensured, the use of the pair/resync monitor is not supported. It is highly recommended to disable application service failover for the time of the disk pair recovery (resynchronization). Cluster Extension XP assumes that if the monitor is enabled, immediate action will be taken to recover a reported suspended disk pair. If at any time the resynchronization process is running on both disk array sites, data corruption can occur.*

The **ResyncMonitorAutoRecover** option set to **YES** is supported with this monitor only if the minimum disk array firmware version is 01-11-xx

(XP512/XP48) or 21.01.xx (XP128/XP1024), and the minimum RAID Manager XP version is 01.04.00.

The pair/resync monitor uses the **syslog()** facility (Linux/UNIX) and the Event Log (Windows) to inform you if the link for the device group is broken. A broken link is recognized only if data will be written to disk; otherwise, the data is the same on the primary and secondary disk and therefore the device group state is reported as PAIR.

Force flag

The force flag forces Cluster Extension XP to skip the internal logic and enables write access to the local volume regardless of the disk pair state. This flag can be set when you are sure that the current site contains the latest data, even though a previous application service startup process failed because Cluster Extension XP discovered a disk pair status that could not be handled automatically.

To use the force flag feature, you must create a file called *application_name*.**forceflag** in the directory specified by the **ApplicationDir** object prior to starting the application service that uses Cluster Extension XP. Before you create this file, ensure that the application service is not running elsewhere.

This file will be removed after Cluster Extension XP detects the file.

You cannot use the force flag if the local disk state is **SVOL_COPY**, which indicates that a copy operation is in progress. A disk cannot be activated when a write operation is in progress to that disk; therefore, Cluster Extension XP returns a global error.

Using the force flag does not enable the automatic recovery features of Cluster Extension XP. After using the force flag, you must recover the suspended or broken disk pairs by using RAID Manager XP commands as described in “Recovery sequence” (page 228).

Pre-execution and post-execution programs

Cluster Extension XP can invoke pre-execution and post-execution programs prior to or after a Cluster Extension XP takeover function. Those programs can be any executable, and must be able to provide return codes to Cluster Extension XP. If the programs add significant execution time to the application service startup process, the timeout values for the startup process must be adjusted in the cluster software.

Cluster Extension XP transfers information as command-line arguments to the pre-execution and post-execution programs. Pre-executables and post-executables must be specified by full path in the **PreExecScript** and **PostExecScript** objects. If no executable is specified (empty value for the object), no preprocessing or postprocessing, respectively, is done.

The following arguments are transferred to the scripts in this order:

1. Name
2. Vgs
3. **RaidManagerInstances**
4. **DeviceGroup**
5. local device group state (check)

   Post-executable status after failover
6. local device group state (display)

   Post-executable status after failover
7. remote device group state (check)

   Post-executable status after failover
8. remote device group state (display)

   Post-executable status after failover
9. current fence level
10. disk array serial numbers (local)
11. reserved

12. reserved
13. disk array firmware version (local)
14. RAID Manager XP version (local)
15. application directory path (**ApplicationDir** object)
16. log file location (**LogDir** object)
17. **DC_A_Hosts** node names
18. **DC_B_Hosts** node names

The pre-executables and post-executables must return a return code. These return codes are used to determine whether a takeover function must be called.

Pre-executable return codes

0	**PRE_OK_TAKEOVER** Pre-executable ok and takeover action allowed.
1	**PRE_ERROR_GLOBAL** Pre-executable failed; no takeover; stop application service cluster-wide.
2	**PRE_ERROR_DC** Pre-executable failed; no takeover; stop application service in this data center.
3	**PRE_ERROR_LOCAL** Pre-executable failed; no takeover; stop application service on this system.
4	**PRE_ERROR_TAKEOVER** Pre-executable failed; takeover action allowed.
5	**PRE_OK_NOTKVR_NOPST** Pre-executable ok; no takeover; no post-exec.

Caution *If the pre-execution program returns 1, 2, 3 or 5, a properly configured post-executable will not be executed. If a takeover function fails, the post-executable will not be executed.*

Post-execution return codes

0	**POST_OK** Post-executable ok; continue.
1	**POST_ERROR_GLOBAL** Post-executable failed; stop application service cluster-wide.
2	**POST_ERROR_DC** Post-executable failed; stop application service in this data center.
3	**POST_ERROR_LOCAL** Post-executable failed; stop application service on this system.
4	**POST_ERROR_CONTINUE** Post-executable failed; continue without error.

Caution *Windows 2000 script and batch files return* ***0*** *if the program was successfully executed, even if you return a different return code.*

Cluster Extension XP log facility

The logging module of Cluster Extension XP provides log messages to the cluster software as well as to the Cluster Extension XP log file. The Cluster Extension XP log file includes disk status information.

The Cluster Extension XP log file is located in this directory:

Linux/UNIX

/var/opt/hpclx/log

Windows

By default, this location is defined as this value:

%ProgramFiles%\Hewlett-Packard\Cluster Extension XP\log\

For the quorum service, the log resides at this location:

%systemroot%\clxq.log

If the log file needs to be cleared and reset, for example, to reduce disk space usage, archive the log file and then delete it. A new log file is generated automatically.

Related information

For information about log levels, see “LogLevel”.

Error return codes

Cluster Extension XP provides the following error return codes for failover operations:

local error Prohibits an application service startup on the local system. This can be caused by the inability of Cluster Extension XP to enable disk access, or misconfiguration of the disk array environment.

data center error Prohibits an application service startup on any system in the local data center. This error is returned if the disk state indicates that it makes no sense to allow any other system connected to the same disk array to access the disks.

global error A global error is returned if the configuration or the disk state does not allow an automatic application service startup process. Manual intervention is required in such cases.

When Cluster Extension XP is integrated, an error message string and integer value are displayed. For the command-line interface, a return code is displayed. For more information, see “CLI Commands” on page 195.

Quorum service for Microsoft Cluster service

Microsoft Cluster service uses a single SCSI disk called the quorum disk to eliminate the potential split brain condition and to coordinate administrative actions performed by cluster nodes. The quorum disk represents a possible single point of failure (SPOF) because when it fails or communications is lost to it, the whole cluster will shut down. If this SPOF is not eliminated, the cluster is at risk even when geographically dispersed.

The quorum service resolves the SPOF problem with the quorum disk. It employs the HP StorageWorks Continuous Access XP (CA) technology to remotely mirror the quorum disk and extends the Microsoft Cluster service functions to manage this mirrored quorum disk.

The quorum service performs two major functions:

- To manage the mirrored quorum disk pair during regular cluster operations and failover.

  It detects quorum disk operations from Microsoft Cluster service and swaps the disk pair in a timely manner before Microsoft Cluster service moves the quorum ownership to the mirrored side or before a failover to the mirrored side occurs.

- To avert a split brain scenario should part of the cluster become completely isolated from the rest of the cluster.

  When cluster nodes on the mirrored side of a dispersed cluster lose all connections (including all heartbeats and CA links), the quorum service uses external arbitration to address the potential split brain problem.

Using the quorum service in a Microsoft Cluster service environment

In general, there are two types of disks used in a Microsoft Cluster service environment: the application data disk and the quorum disk. The quorum disk is a special case of an application disk.

While the Cluster Extension XP resource DLL provides disaster recovery (DR) and high availability (HA) support for Microsoft Cluster service application disks using the same CA technology, the quorum service focuses on protecting the quorum disk, which functions somewhat differently from data disks. In most circumstances, this protection is critical for the cluster's control resource. If the quorum disk fails, the whole cluster will be unavailable, regardless of how well the application data disks are protected.

The quorum service has few configuration options. After it is installed, the service works seamlessly with the Microsoft Cluster service and the cluster applications.

Install and use the quorum service together with the Cluster Extension XP resource type to provide complete coverage for the cluster.

Quorum processes

The quorum service creates a synchronous CA disk pair for the quorum disk to prevent it from being a SPOF. When the primary quorum disk fails, the quorum service allows the cluster to use the mirrored (secondary) quorum disk to continue cluster operations. Specifically, the service extends the Microsoft Cluster service quorum management protocol to:

- maintain the existing ownership in normal operations
- decide the new ownership when the original owning host node fails
- fail over ownership when the primary quorum disk fails in the context of mirrored quorum disks

One major task of the quorum service is to swap the quorum pair. Usually the primary disk has both read and write permissions. However, the mirrored disk on the secondary side has no write permission. To use the secondary quorum disk as a quorum, the service swaps the direction of the disk pair to make the quorum disk at the second site writable before Microsoft Cluster service starts to use it. The quorum service interacts with Microsoft Cluster service at the I/O operation level. When it finds the Microsoft Cluster service is going to use the secondary side quorum disk, it swaps the disk pair for Microsoft Cluster service first. It uses the RAID Manager XP library to communicate with the XP disk array.

The quorum service also moderates ownership of the quorum resource. More than one node can request ownership simultaneously. To coordinate those nodes between dispersed locations and to eliminate the dependency on the networks, the quorum service uses three sets of small volumes (disks) as control devices to manage quorum ownership and coordination. The first is always in a paired state; the second is not paired; and the third changes its paired state dynamically. Because they require little space, you can minimize resource requirements by creating CVS (custom volume size) disks for use as control devices.

Cluster operation continues even though access to the quorum disk is lost on cluster nodes that do not own the quorum disk. On those nodes, the quorum service (ClxQSvc) continues to monitor the quorum filter driver in case I/O access is lost to the XP disk system. However, logging to the clxq.log file will be suspended until the quorum filter driver detects the arrival of the missing quorum disk. This behavior follows the Microsoft Cluster service behavior, which means that the Microsoft Cluster service continues to run but the cluster node will not be able to become a quorum disk owner in case the original quorum disk owner fails. The quorum service and the quorum filter driver log this type of incident to the Event Log to inform you properly.

If the quorum service exits without being gracefully stopped by the Windows Service Control Manager (SCM), the SCM will automatically restart the quorum service under virtually all possible conditions.

The quorum service is set to wait for 30 seconds for Microsoft Cluster service startup activity. If the Cluster service does not go into START_PENDING state during this timeframe, the quorum service will stop automatically. The actual interval for Cluster Service startup checks is 1 second. It will continue checking every second until the Cluster service changes to START_PENDING state. START_PENDING is an internal state. The Service window will show STARTING as service status.

It is recommended that you start ClxQSvc (the quorum service) through its enforced dependency to the Cluster service. This means you should start the Cluster service first, which will cause the quorum service to start automatically before the Cluster service starts.

The wait timeout is the total amount of time (in seconds) that the quorum service will wait for the Cluster service to start up, before giving up and shutting itself down. This value should be sufficient for all cases. However, you can change the value during manual service startups as follows:

1. Open the Services window.
2. Right click on ClxQSvc.
3. Select Properties.
4. In the Start parameters: field, enter

   `/waitforclussvc <number in seconds>`
5. Press Start in the properties window.

The quorum service stops automatically if the Cluster service stops.

The quorum service checks during its initialization if the configured disk pairs for the quorum service control disk 1 (STATUS) and quorum disk are established. The quorum service will not start and does not allow the Cluster service to start if those disk pairs are not in PAIR state.

For emergency startup of the Cluster Service in case all quorum arbitration functions are unavailable or the disks cannot be paired, the quorum service can be started manually:

1. Open the Services window.
2. Right-Click on ClxQSvc.
3. Select Properties.
4. In the Start parameters: field, enter

   `/createsplitbrain`
5. Click the Start button in the properties window. This could create two separate clusters.

Caution *The creation of separate clusters can result in data loss as explained below.*

Clicking the Start button in the Services window can lead to a split brain syndrome, where the cluster runs the same applications on two different sets of disks. This can happen if the cluster is running or restarted in the remote data center and any number of cluster nodes are isolated from each other.

A split brain condition occurs when one site of the cluster loses all the connectivity with the second site and each site decides to form its own cluster. The serious consequence of the split brain scenario is the corruption of business data because client data will no longer be consistent. To eliminate the split brain syndrome in cases where there is a total loss of communications between sites, the quorum service employs an external arbitration mechanism.

The external arbitrator runs on a node on your intranet, external to either cluster site, but accessible by all cluster nodes. Assuming it is reachable by one data center, even when one cluster site of a cluster loses all connectivity with the other site, its major function is, upon request, to check the cluster status by communicating with the cluster nodes. By providing sufficient information to the cluster nodes, they can make the critical decision on whether to form a new cluster, thereby avoiding a split brain condition.

In addition to the external arbitrator, two processes on each cluster node work closely with the arbitrator. One process is created dynamically on a cluster node located at the site not holding the quorum disk when it detects that it has lost all connectivity to the site holding ownership of the quorum disk. Its major function is to communicate with the arbitrator to determine whether the cluster is still functioning. Based on the status of the cluster and which site owns the quorum disk, it may decide to form a new cluster, leave the cluster down, or join an existing cluster. This process is called the “cluster decision maker.”

The second process is created dynamically on the host node owning the quorum resource when the node detects that the mirror link between two sites goes down. The main purpose of this process is to shut down the cluster on the site that was controlling the quorum disk when that site became completely isolated from the external network. This prevents a potential split brain condition when the connectivity between the two sites is restored, and the formally isolated copy of the cluster suddenly becomes available to the external network. This process is called the “isolation checker.” The isolation checker will restart after it has finished and the CA XP link is still down. This can cause the Cluster service to stop if the network link between the arbitrator and the quorum owner in the cluster fails after all heartbeat network links and all CA XP links have failed.

If a broken CA XP link is detected by the quorum service, messages will be logged to the Event Log that show whether the cluster decision maker or the isolation checker is running.

Caution *Do not start the Cluster service while the cluster decision maker or the isolation checker is running. The cluster decision maker and the isolation checker will log start and stop messages with the results of their findings to the Event Log.*

The quorum service supports two ways to retrieve the required information about the external arbitrator, such as its IP address and port number. One is through a local configuration file, and the other is through the Active Directory server. During installation, the user is asked whether an Active Directory service is available. The quorum service will use the Active Directory if it is available. Otherwise, it will use its local configuration file generated by the installation process.

The quorum service is implemented as a Windows service. To ensure that the service is always available for the quorum disk, a startup dependency is imposed on the cluster service by the quorum service. The quorum service must start and remain functioning during the entire time that the cluster is running. If it is forced to stop, it will also stop the cluster service on that server, ensuring that the quorum disk pair is always properly managed by the quorum service.

3

User configuration file and Cluster Extension XP objects

Objects define the disk array environment and failover/failback behavior. Objects can be customized in the user configuration file or directly in the cluster software.

The user configuration file

Cluster Extension XP uses the user configuration file to gather application service-specific information. This file describes the dependencies between application services and RAID Manager XP device groups in one file for all application services in the cluster. This file must be copied to all nodes that use Cluster Extension XP.

The user configuration file must be placed in the configuration directory:

Linux
UNIX

/etc/opt/hpclx/conf

Windows

By default, this location is defined as this value:
%ProgramFiles%\Hewlett-Packard\Cluster Extension XP\conf

Related information

"Basic configuration example" (page 87)
"Creating and configuring the user configuration file" (page 194)

HACMP

The **UCF.cfg** file is required for IBM HACMP. A single **UCF.cfg** file must be maintained and copied to all systems using Cluster Extension XP. The **UCF.cfg** includes a "common" section to configure the Cluster Extension XP environment and an "application" section to configure the application service-dependent failover/failback behavior. The application section is a multitag component; the **APPLICATION** tag and application-related objects can appear numerous times in the **UCF.cfg**.

Related information

"User configuration file for HACMP" (page 101)

Microsoft Cluster service

Cluster Extension XP integration with Microsoft Cluster service does not require a user configuration file when the standard environment for Cluster Extension XP is used. The Cluster Extension XP objects that are integrated with Microsoft Cluster service are configurable as resource private properties in the cluster software.

Related information “Configuring Cluster Extension XP resources” (page 117)

VCS

Cluster Extension XP integration with VERITAS Cluster Server does not require a user configuration file when the standard environment for Cluster Extension XP is used. The Cluster Extension XP objects that are integrated with VERITAS Cluster Server are configurable as resource attributes in the cluster software.

Related information “Configuring the Cluster Extension XP resource” (page 154)

SG-LX

An environment configuration file is required for Serviceguard. The file must reside in the same directory as the package control file and is identified by the package name:

*package_name*_**clx.env**

The **APPLICATION** tag is required, although no value is required.

Related information “Configuration of the Cluster Extension XP environment” (page 174)

File structure

The configuration file comprises a common section and application sections. These sections are distinguished by control tags. Cluster Extension XP uses the following objects as control tags:

- **COMMON**
- **APPLICATION**

Objects have one of the following formats:

tag — a definition of an object, for example, **COMMON** or **APPLICATION**.

integer — a number, for example, a timeout value.

string — a name, which can include alphabetic and numeric characters and underscores, for example, an application startup value.

list — a list of space-separated strings, for example, a list of host names (lists of numbers are stored as lists of strings).

Specifying object values

When using the default configuration, you must provide values for these five objects:

DeviceGroup (page 82)
DC_A_Hosts (page 82)
DC_B_Hosts (page 82)
RaidManagerInstances (page 84)
XPSerialNumbers (page 86)

You do not need to change the default settings unless you want to change the degree of protection for your paired disks. If you change an object, you may need to change additional objects as well. For example, if you change the **FenceLevel** object to **DATA**, you might need to change the **DataLoseMirror** object also.

Objects are supported according to the requirements or capabilities of the cluster software, as listed in table 1 (page 69).

Table 1. Cluster Extension XP supported objects

Name	Page	CLI	HACMP	MS Cluster service	VCS	SG-LX
COMMON	71	●	●	●	●	●
LogDir	71	●	●	●	●	●
LogLevel	71	●	●	●	●	●
SearchObject	72		●			
VcsBinPath	72				●	
APPLICATION	74	●	●	●	●	●
ApplicationDir	74	●	●	●	●	
ApplicationStartup	75	●	●	●	●	●
AsyncTakeoverTimeout	77	●	●	●	●	●
AutoRecover	78	●	●	●	●	●
BCEnabledA	79	●	●	●	●	●
BCEnabledB	79	●	●	●	●	●
BCMuListA	79	●	●	●	●	●
BCMuListB	79	●	●	●	●	●
BCResyncEnabledA	79	●	●	●	●	●
BCResyncEnabledB	80	●	●	●	●	●
BCResyncMuListA	80	●	●	●	●	●
BCResyncMuListB	80	●	●	●	●	●
DataLoseDataCenter	80	●	●	●	●	●
DataLoseMirror	81	●	●	●	●	●
* **DC_A_Hosts**	82	●	●	●	●	●
* **DC_B_Hosts**	82	●	●	●	●	●

Table 1. Cluster Extension XP supported objects (Continued)

(continued)

Name	Page	CLI	HACMP	MS Cluster service	VCS	SG-LX
* **DeviceGroup**	82	•	•	•	•	•
FastFailbackEnabled	83				•	
FenceLevel	83	•	•	•	•	•
Filesystems	83	•	•			
PostExecCheck	84	•	•	•	•	•
PostExecScript	84	•	•	•	•	•
PreExecScript	84	•	•	•	•	•
* **RaidManagerInstances**	84	•	•	•	•	•
ResyncMonitor	85		•	•	•	•
ResyncMonitorAutoRecover	85		•	•	•	•
ResyncMonitorInterval	85		•	•	•	•
ResyncWaitTimeout	86	•	•	•	•	•
Vgs	86	•	•	•	•	•
* **XPSerialNumbers**	86	•	•	•	•	•

LEGEND

* Required

• Supported

COMMON section objects

The common part is used to set the environment of Cluster Extension XP.

The **COMMON** tag is a single-tag; it can appear in the configuration file only once. The common object does not require any value.

Objects of the type common can only appear once. Those objects must be placed after the **COMMON** tag in the configuration file.

If the default values fit your environment, there is no need to specify them in the file.

COMMON

Format tag

Description Distinguishes between general (common) and application-specific objects.

LogDir

Format string

Description *(Optional)* Defines the path to the Cluster Extension XP log file.

Default value *Linux/UNIX*
/var/opt/hpclx/log

Windows
%ProgramFiles%\Hewlett-Packard\Cluster Extension XP\log

LogLevel

Format string

Description *(Optional)* Defines the logging level used by Cluster Extension XP.

Valid values	**error** *(default)*	Logs only error messages for events that are nonrecoverable.
	warning	Logs **error** messages and **warning** messages for events that are recoverable.
	info	Logs **error** messages, **warning** messages, and additional information, such as disk status.
	debug	Logs **error** messages, **warning** messages, **info** messages, and messages that report on execution status, useful for troubleshooting.

SearchObject *HACMP only*

Format string

Description *(Optional)* Searches for the application service if the user configuration file specifies multiple applications. This object is not used for VCS, Microsoft Cluster service, or SG-LX.

Default value **Vgs**

VcsBinPath *VCS only*

Format string

Description *(Optional)* Defines the path to the VCS binaries. This object is not used for Microsoft Cluster service, SG-LX, or HACMP.

Default value **/opt/VRTSvcs/bin**

APPLICATION section objects

The application part defines the failover and failback behavior of Cluster Extension XP for each application service. **APPLICATION** is a multitag that can appear in the configuration file for each application service using Cluster Extension XP.

The **APPLICATION** object requires the name of the application service as its value. The objects specified after an **APPLICATION** tag must appear only once per application. As with the common part objects, the application part objects have predefined default values.

Cluster Extension XP also uses the following rules to define objects:

- If you use the default value, you do not have to specify the object.
- Cluster Extension XP uses objects depending on the setting of other objects. For example, if you set the **FenceLevel** object to **DATA**, Cluster Extension XP uses the values specified for the **DataLoseMirror** or **DataLoseDataCenter** object. However, these objects are ignored if the **FenceLevel** object is set to **NEVER**.
- The pre-execution and post-execution functions in Cluster Extension XP will not be processed if the associated object values are empty. (This is the default setting.)

CLI
HACMP
SG-LX

To set **APPLICATION** object values, use the user configuration file.

VCS

Use the VCS GUI to set **APPLICATION** object values.

Microsoft Cluster service

To set **APPLICATION** object values, use the Microsoft Cluster service Cluster Administrator GUI.

APPLICATION

Format tag

Description Distinguishes between general and application-specific objects. Specify the name of the application service. The format of its value is equivalent to a string value.

SG-LX

For Serviceguard, the tag is required; however, specifying a value is not necessary.

ApplicationDir

Format string

Description Specifies the directory where Cluster Extension XP searches for application-specific files, such as the force flag or online file.

If **ApplicationDir** is set to a nonexistent drive and **PairResyncMonitor** is not enabled, Cluster Extension is unable to create the online file and cannot put the resource online.

SG-LX

The value of **ApplicationDir** is derived from the package control file location.

Windows

If **ApplicationDir** is not set, Cluster Extension uses the local **%HPCLX_PATH%** values as defined in the registry.

Default values *Linux*
UNIX
/etc/opt/hpclx

Windows
%HPCLX_PATH%

Files

*resource_name***.createsplitbrain**
*resource_name***.forceflag**
*resource_name***.online**

If specified in a user configuration file, *resource_name* is the value of the **APPLICATION** tag; otherwise, *resource_name* is the value of the Cluster Extension XP resource name.

ApplicationStartup

Format

string

Description

(Optional) Specifies where a cluster group should be brought online.

The **ApplicationStartup** object can be customized to determine whether an application service starts locally or is transferred back to the remote data center (if possible) to start directly without waiting for resynchronization. This object is used only if an application service has already been transferred to the secondary site and no recovery procedure has been applied to the disk set (the disk pair has not been recovered and is not in **PAIR** state). This process is considered a failback attempt without prior disk pair recovery.

Cluster Extension XP can detect the most current copy of your data based on the disk state information. If Cluster Extension XP detects that the remote XP disk array has the most current data, it orders a resynchronization of the local disk from the remote disk, or it stops the startup process to enable the cluster software to fail back to the remote XP disk array.

If a resynchronization is ordered, Cluster Extension XP monitors the progress of the copy process. If the application service was running on a secondary XP disk array without replication link, a large number of records may need to be copied. If the copy process takes more time than the configured application startup timeout, the application startup will fail.

Microsoft Cluster service

If the **ApplicationStartup** resource property is set to **FASTFAILBACK** and the **FailoverThreshold** value is set to a number higher than the current number of clustered systems for the resource group, the resource group will restart on configured nodes until one of the following conditions is met:

- The resource is brought online in the remote data center.
- The resource failed because the **FailoverThreshold** value has been reached.
- The resource failed because the **FailoverPeriod** timeout value has been reached.

Caution *Disable subsequent automated failover procedures for recovery failback operations.*

Valid values **FASTFAILBACK** *(default)*

The cluster group will be brought online in the remote data center (if possible) without waiting for resynchronization. The application startup process will be stopped locally and Cluster Extension XP reports a data center error. Depending on the cluster software, the application service cannot start on any system in the local data center and the cluster software will transfer the application service back to the remote data center. Use this value to provide the highest application service uptime. Depending on the value configured for the **AutoRecover** object, Cluster Extension XP will attempt to update the former primary disk based on the secondary disk and swap the personalities of the disk pair so that the local disk will become the primary disk.

In a two-node cluster, this process will not work because the target failback system would not be available. In this case, the application service must be started manually, or the **ApplicationStartup** object should be set to **RESYNCWAIT**.

RESYNCWAIT Online local, cluster group must wait until the disk status is **PAIR**. Cluster Extension XP will initiate a resynchronization of the local disk based on the remote disk. The copy process will be monitored. If no copy progress was made after a monitoring interval expired, the copy process is considered failed and Cluster Extension XP returns a global error. If **RESYNCWAIT** has been specified for the **ApplicationStartup** object, the **ResyncWaitTimeout** object must be specified, in case Cluster Extension XP should wait for resynchronization changes for more or less than 90 seconds, which is the default.

AsyncTakeoverTimeout

Format integer

Description *(Optional)* Specifies the **horctakeover** command timeout in seconds. Must be adjusted based on disk mirroring link speed.

This object is used only if the **FenceLevel** object value is **ASYNC**.

The takeover operation for fence level **ASYNC** (Continuous Access XP Extension) offers the option to stop the data transfer process after a specified time value. This is used to allow access to the remote copy if the data transfer process has been stopped due to a Continuous Access XP-link failure. All data that has been copied up to the moment the timeout value has been reached is consistent and available to access at the secondary site.

Caution *Measure or calculate the full XP disk array cache copy time to use the gathered information for the **AsyncTakeoverTimeout** object. After a takeover command has been invoked, Continuous Access XP Extension copies the side file area residing in the XP disk array cache to the site where the takeover command has been issued (the secondary disks). The side file area cannot exceed the installed cache size. The maximum time for the **AsyncTakeoverTimeout** object is the time to fully copy the amount of cache size data. The takeover timeout value is used to terminate the copy process to provide access to the secondary disks, for example, if all links or the primary XP disk array are unavailable to copy the side file area. The copy time depends on the performance of the Continuous Access XP link between your sites. The takeover or resynchronization operation could take longer than the timeout value for application service startup in the cluster software. The application service startup might fail in this case. However, the takeover or resynchronization command will continue in the background.*

Default value **1800** *(default)*

AutoRecover

Format string

Description *(Optional)* Recovers a suspended or deleted disk pair when the resource is brought online at application service startup time.

If the **AutoRecover** object is set to **YES**, Cluster Extension XP will try to resynchronize the remote disk at application startup time. Cluster Extension XP will ignore the return code of the resynchronization command and allow access to the disk ensuring highest application availability.

If the resynchronization attempt fails, Cluster Extension XP will not fail. The internal logic will first apply the concurrency and consistency rules to allow access to the disk set.

If you configure fence level **DATA** for the device group and set the **FenceLevel** object to **DATA**, the **AutoRecover** object will change Cluster Extension XP's behavior. Cluster Extension XP will attempt to reestablish the **PAIR** state and wait for the **PAIR** state before it allows access to the disk. If the resynchronization or takeover process fails, Cluster Extension XP returns a global error.

Valid values **YES** *(default)*
NO

BCEnabledA

Format string

Description *(Optional)* Enables rolling disaster protection for data center A.

Valid values **YES**
NO *(default)*

BCEnabledB

Format string

Description *(Optional)* Enables rolling disaster protection for data center B.

Valid values **YES**
NO *(default)*

BCMuListA

Format list

Description *(Optional)* Space-separated list defines the MU number of the Business Copy XP disk pairs in data center A.

BCMuListB

Format list

Description *(Optional)* Space-separated list defines the MU number of the Business Copy XP disk pairs in data center B.

BCResyncEnabledA

Format string

Description *(Optional)* Enables automatic resynchronization of Business Copy XP disk pairs in data center A. The automatic resynchronization function is supported only when the split BC pair is located in the same data center where Cluster Extension XP is started.

Valid values **YES**
NO *(default)*

BCResyncEnabledB

Format string

Description *(Optional)* Enables automatic resynchronization of Business Copy XP disk pairs in data center B. The automatic resynchronization function is supported only when the split BC pair is located in the same data center where Cluster Extension XP is started.

Valid values **YES**
NO *(default)*

BCResyncMuListA

Format list

Description *(Optional)* Space-separated list defines the MU number of the Business Copy XP disk pairs in data center A.

BCResyncMuListB

Format list

Description *(Optional)* Space-separated list defines the MU number of the Business Copy XP disk pairs in data center B.

DataLoseDataCenter

Format string

Description *(Optional)* Specifies whether a resource should be brought online while the disk pair is (or will be) suspended or deleted and there is no connection (CA XP and IP network) to the remote data center.

Used only if the **FenceLevel** object value is **DATA**.

RAID Manager XP is able to access its remote peer to invoke takeover actions for Continuous Access XP device groups. It is also able to invoke a swap-takeover operation of the device group from the secondary site. If no configured remote RAID Manager XP instance replies to a request of the local RAID Manager XP instance (remote status EX_ENORMT), all network connections between the local and the remote data center are considered *DOWN*. If the swap-takeover operation leads into a suspended state for the device group, the Continuous Access XP links are considered *DOWN*.

Because redundant networks and Continuous Access XP links are necessary to build a disaster tolerant environment, this situation can be considered as a data center failure. The **DataLoseDataCenter** object is used to allow/prohibit automatic application service startup in this particular case.

The combination of setting the **DataLoseMirror** object to **YES** and the **DataLoseDataCenter** object to **NO** are contradictory.

Valid values **YES** *(default)*
NO

DataLoseMirror

Format string

Description *(Optional)* Specifies whether a resource should be brought online while the disk pair is suspended or deleted.

Used only if the **FenceLevel** object value is **DATA** and local and remote XP disk status information can be gathered. If the remote XP disk state information is not available (remote state **EX_ENORMT**), the setting of the **DataLoseDataCenter** object will be used.

Depending on the value configured for the **AutoRecover** object, Cluster Extension XP will attempt to recover the **PAIR** state for the device group. Cluster Extension XP waits until the **PAIR** state has been established. If this operation fails, Cluster Extension XP will return a global error. Because the **DATA** fence level ensures no loss of concurrency, manual intervention is required to recover the **PAIR** state. The **PAIR** state must be reestablished for all disks in the device group before you can start the application service.

The combination of setting the **DataLoseMirror** object to **YES** and the **DataLoseDataCenter** object to **NO** are contradictory.

Valid values **YES**
NO *(default)*

DC_A_Hosts *Required*

Format list

Description Space-separated list defines the cluster nodes in data center A.

VCS

This object is a string-vector element. Add a new element to the list for each system name.

DC_B_Hosts *Required*

Format list

Description Space-separated list defines the cluster nodes in data center B.

VCS

This object is a string-vector element. Add a new element to the list for each system name.

DeviceGroup *Required*

Format string

Description RAID Manager XP device group, containing the application service disk set.

Files *Linux*
UNIX
/etc/horcm*X*.conf

Windows
drive:**\winnt\horcm*X*.conf**
%system_root%\horcm*X*.conf

where *X* is the RAID Manager XP instance number.

FastFailbackEnabled *VCS only*

Format string

Description *(Optional)* Disables VCS service groups for the data center. This allows transferring the service group back to the remote data center immediately. To allow this operation, the VCS configuration file (**main.cf**) will be write enabled and saved later.

The service group will be disabled for all systems contained in either the **DC_A_Hosts** object or **DC_B_Hosts** object. Then, the VCS configuration file will be saved (dumped).

Valid values **YES** *(default)*
NO

FenceLevel

Format string

Description *(Optional)* The **FenceLevel** object specifies the fence level configured for the device group. Cluster Extension XP checks whether the current fence level reported by the XP disk array is the same as the configured (expected) fence level. This object is also used to make sure your configurations are supported based on consistency considerations. Different failover and recovery procedures are used for different fence levels.

If you change the **FenceLevel** object value, also review the values of these objects:

DataLoseMirror (page 81)
DataLoseDataCenter (page 80)
AsyncTakeoverTimeout (page 77)

Valid values **DATA**
NEVER *(default)*
ASYNC

Filesystems *CLI and HACMP only*

Format list

Description Space-separated list of file systems.

PostExecCheck

Format string

Description *(Optional)* The **PostExecCheck** object is used to configure Cluster Extension XP to gather XP disk pair status information after the takeover procedure. That information will be passed to the post-executable. In case of a remote data center failure, it could be time consuming to gather that information, especially if your post-executable does not need any XP status information. The arguments passed to the post-executable will include only the local disk status if the **PostExecCheck** object is set to **NO**. See "RAID Manager XP configuration" (page 90).

Valid values **YES**
NO *(default)*

PostExecScript

Format string

Description *(Optional)* Specifies an executable with its full path name to be invoked after the takeover action or failover procedure.

PreExecScript

Format string

Description *(Optional)* Specifies an executable with its full path name to be invoked before the takeover action or failover procedure.

RaidManagerInstances *Required*

Format list

Description A space-separated list of RAID Manager XP instances Cluster Extension XP can use to communicate with the disk array. The instance numbers must be the same among all cluster systems. Cluster Extension XP can alternate between the specified instances.

VCS

This object is a string-vector element. Add a new element to the list for each system name.

Files

Linux
UNIX
/etc/horcm*X*.conf

Windows
%systemroot%\horcm*X*.conf

where *X* is the RAID Manager XP instance number.

ResyncMonitor

Format string

Description *(Optional)* Starts the pair/resync monitor to monitor the disk pair status and resynchronize disk pairs if the **ResyncMonitorAutoRecover** attribute is set to **YES**.

Valid values

YES *(default: Microsoft Cluster service)*
NO *(default: HACMP; SG-LX; VCS)*

ResyncMonitorAutoRecover

Format string

Description *(Optional)* Automatically recovers disk pairs states if the disk pairs are monitored by the pair/resync monitor.

Valid values

YES
NO *(default)*

ResyncMonitorInterval

Format integer

Description *(Optional)* Specifies the monitor interval in seconds the pair/resync monitor will check the disk pair status.

Default value **60**

ResyncWaitTimeout

Format integer

Description *(Optional)* It is used to specify the timeout value in seconds for a disk pair resynchronization. It may take some time to resynchronize disks. The timer times out if there is no change in the percentage value of the copy status for the device group in the specified time interval. The timeout value is used if the **ApplicationStartup** object is set to **RESYNCWAIT**.

Default value **90**

Vgs *CLI and HACMP only*

Format list

Description List of volume groups

XPSerialNumbers *Required*

Format list

Description A space-separated list of at least two serial numbers must be specified: the serial numbers of the primary and secondary XP disk arrays. Cluster Extension XP checks whether the local disk array is contained in this list. Serial numbers of the disk arrays of the connected cluster nodes (at least two).

VCS

This object is a string-vector element. Add a new element to the list for each system name.

Basic configuration example

The following is an example of a basic **UCF.cfg** file.

```
#/etc/opt/hpclx/conf/UCF.cfg
#This is the Cluster Extension XP User Configuration File (UCF.cfg).
#The COMMON tag specifies the configuration for the
#Cluster Extension XP core environment
COMMON
LogLevel        info                    #default (not necessary)
APPLICATION     sap                     #the application service
 Vgs            sapdatavg saptmpvg      #the volume groups (not necessary)
 Filesystems    /sapdata /saptmp        #the filesystems
 DeviceGroup    sapdg                   #RM dev group for the app service
 RaidManagerInstances 22                #RM instance number for dev group
 DC_A_Hosts     host1a host2a           #Data center A
 DC_B_Hosts     host3b host4b           #Data center B
```

4

RAID Manager XP dependencies

Cluster Extension XP depends on HP StorageWorks RAID Manager XP and the cluster software it is integrated with.

Before you configure Cluster Extension XP, verify that the host and disk array systems are properly configured:

- The disk array and its remote peer have been properly configured.
- The host system recognizes the disk arrays.
- The HP StorageWorks Continuous Access XP links are bidirectional and working properly.
- You are familiar with the disk and volume configuration of the operating system.

RAID Manager XP configuration

To function properly, Cluster Extension XP requires at least one instance of RAID Manager XP. Cluster Extension XP starts the configured RAID Manager XP instance if it is not running. However, if the RAID Manager XP instance cannot be started or returns an error, Cluster Extension XP can switch to an alternate RAID Manager XP instance.

Ensure that the path to the RAID manager binary files is included in the **PATH** environment variable.

Recommendation Configure two RAID Manager XP instances per system and start those instances automatically at system boot time.

RAID Manager XP configuration file

The RAID Manager XP configuration file (**horcm*X*.conf**) is used to map device groups to the internal disk array disks. A device group is the common unit for failover operations initiated from the server side.

A RAID Manager XP configuration file consists of these four parts:

- **HORCM_MON**

  The monitor part defines the local network and port where the RAID Manager XP instance is listening for incoming requests from a remote instance. It also defines the polling interval and timeout value for request to other instances.

  The first entry defines the network that RAID Manager XP listens to. The default value is **NONE**. The default setting enables RAID Manager XP to listen on all configured networks.

  The timeout value is important to Cluster Extension XP. You can configure the time Cluster Extension XP will wait to receive information back from the remote site. The timeout interval applies for each remote instance configured in the **HORCM_INST** section of the RAID Manager XP configuration file. If the last instance configured in the **HORCM_INST** section is the only instance that will answer a request, it will take the number of seconds of the timeout value times the number of not responding remote instances until the request can be answered. This must be considered for the application service startup timeout value you can configure in your cluster software.

A general formula for this behavior in case of a complete site failure is the following:

$t_W = t_{HM}$ (in 10 msec) x $(n_{HI} + 1)$

t_W = wait time until remote error will be reported by local RAID Manager XP instance

t_{HM} = **HORCM_MON** timeout

n_{HI} = number of remote instances, specified per device group in **HORCM_INST**

Recommendation Reduce the default timeout value in conjunction with increasing numbers of different (at least two) network connections to the remote RAID Manager XP instance. The settings of these two parameters directly affect the timing of the failover behavior of Cluster Extension XP. Cluster Extension XP experiences the above mentioned wait time twice if all of the remote RAID Manager XP instances cannot be reached. If a post-executable is configured, a third wait time period is added.

- **HORCM_CMD**

  The command device part defines which raw disk device is used to communicate to the disk array. This device cannot be used for any data other than control data of the RAID Manager XP instance. Several command devices may be configured to provide alternate access paths to control Continuous Access XP pair operations.

  If command devices are configured in separate lines, RAID Manager XP interprets those devices as different disk arrays. Therefore, you can use one RAID Manager XP instance to control several XP disk arrays. Cluster Extension XP does not support this feature.

- **HORCM_DEV**

  The device group part maps device groups and device names to internal disks (LDevs) in the disk array. Failover operations are carried out for the device groups but can also be initiated for a single disk pair. The device groups and device names must be unique in the RAID Manager XP configuration file. However, device group names should be unique for the whole cluster environment to prevent any kind of user mistake.

For fence level **ASYNC**, the device group also represents a consistency group.

To combine CA disk pairs and BC disk pairs, you can use the MU number to specify internal BC disks.

Recommendation Use the local and remote LDEV and CU number as the device name to easily recognize configuration or mapping mistakes. For example, if the local LDEV number is 0a (hex), the local CU number is 0, the remote LDEV number is 1 (hex) and the remote CU is 3, a recommended device name would be **disk_00a_301**. This approach also ensures unique device names because the LDEV number together with the CU number is a unique disk identifier in a disk array.

Example

```
# pairdisplay –g testdg –fx –CLI
Group  PairVol    L/R Port# TID LU Seq#  LDEV# P/S Status Fence Seq#  P-LDEV# M
testdg disk_00a_301 L CL2-N 3   4  30061 00a   P-VOL PAIR NEVER 30071 301     -
testdg disk_00a_301 R CL2-N 5   1  30071 301   S-VOL PAIR NEVER -     00a     -
```

- **HORCM_INST**

  The remote RAID Manager XP instances part defines which remote system can be used to request information of the device group. For most failover operations, the remote RAID Manager XP instance is not necessary. However, it is used for pair consistency checks and considered important. A remote instance should be configured for each network available between the cluster nodes. The first and preferred network RAID Manager XP instances should communicate with each other in the cluster heartbeat network.

Network considerations

Since RAID Manager XP is an essential resource to Cluster Extension XP, it is highly recommended that you provide reliable network connections for RAID Manager XP communications. It is also recommended to use the heartbeat network (private network) for RAID Manager XP communications. As with the heartbeat network, alternative network paths are highly recommended. RAID Manager XP can be configured for the networks it uses for each device group within the **HORCM_INST** part of the RAID Manager XP configuration file.

Command device considerations

At least one command device must be configured for RAID Manager XP. RAID Manager XP offers the same command device being accessed by redundant paths. This feature should be used to prevent Cluster Extension XP from aborting if a single access path to the command device is missing.

Recommendation Set up a second command device to provide an alternative control to the paired disks.

Caution *If you use Auto Path for AIX to enable alternative pathing on IBM AIX together with the XP disk array, RAID Manager XP does not support Auto Path virtual paths for command devices.*

Start and stop the RAID Manager XP instances

The RAID Manager XP instances configured to be used for Cluster Extension XP should be started at system boot time to provide fastest access to disk status information.

Cluster Extension XP provides scripts (*Linux*/UNIX) or a service (Windows) to integrate RAID Manager XP instance startup into the system startup process. However, if the system cannot automatically start and monitor RAID Manager XP instances, RAID Manager XP can be started and stopped by executing the following commands:

Linux/UNIX	**horcmstart.sh** *instance_numbers* **horcmshutdown.sh** *instance_numbers*
Windows	**horcmstart** *instance_numbers* **horcmshutdown** *instance_numbers*

Starting RAID Manager XP without specifying an instance number will start instance 0 with the associated horcm.conf file. Zero (0) is not recommended as an instance number for a Cluster Extension XP RAID Manager XP instance.

Takeover basic functionality test

After RAID Manager XP has been configured for the device groups used by Cluster Extension XP, each device group must be verified to failover correctly between the disk arrays from each server in the cluster. Therefore, the device group must be in **PAIR** state already.

Caution *RAID Manager XP keeps configuration data of the XP disk array in system memory. Therefore, you must stop and restart RAID Manager XP instances on all systems if a configuration change has been applied to any of the involved XP disk arrays.*

To test the correct failover and failback behavior, log in to each system used with Cluster Extension XP and invoke the following commands if the local disk is the secondary (SVOL) disk:

Linux/UNIX

export HORCMINST=*instance_number*
pairdisplay –g *device_group_name* **–fx –CLI**
horctakeover –g *device_group_name* [**–t** *timeout*]

Windows

set HORCMINST=*instance _number*
pairdisplay –g *device_group_name* **–fx –CLI**
horctakeover –g *device_group_name* [**–t** *timeout*]

The output of the **pairdisplay** command indicates whether the local disk is the secondary (SVOL) disk and if so, the **horctakeover** command shows a SWAP-takeover as a result. If **pairdisplay** shows the local disk as primary (PVOL) disk, log in to a system connected to the secondary (SVOL) disk and invoke the **horctakeover** command there. If the **horctakeover** command does not result in a SWAP-takeover, refer to “Recovery procedures” (page 225) and “Troubleshooting” (page 203) to resolve the issue.

The **–t** option of the **horctakeover** command is only used for fence level **ASYNC**.

5

Integration with HACMP

Cluster Extension XP is integrated with the HACMP cluster software using the standard customization scheme provided by HACMP. This allows cluster administrators to configure the disk array-specific failover behavior as pre-event of the standard HACMP event **get_disk_vg_fs**.

Related information

For information about how to install Cluster Extension XP, see *HP StorageWorks Cluster Extension XP: Installation Guide*.

See the **readme** file on the product CD for supported configurations.

Configuring resources

The Cluster Extension XP objects must be configured using a user configuration file.

The Cluster Extension XP resource gathers all necessary information about the disk arrays if a resource group is brought online.

If configured, a pair/resync monitor is started to monitor the Cluster Extension XP resource. To use this monitor, HACMP must call a pre-event for the standard HACMP event **release_vg_fs**.

The Cluster Extension XP binary **clxhacmp** is called as a pre-event of the standard HACMP event **get_disk_vg_fs** in order to check the status of the RAID Manager XP device group and if necessary takes appropriate actions to allow access to these disks before HACMP is trying to access the disks of the particular resource group.

Procedure for HACMP

To integrate Cluster Extension XP into HACMP:

1. Create a new Custom Cluster Event.

   #smitty hacmp

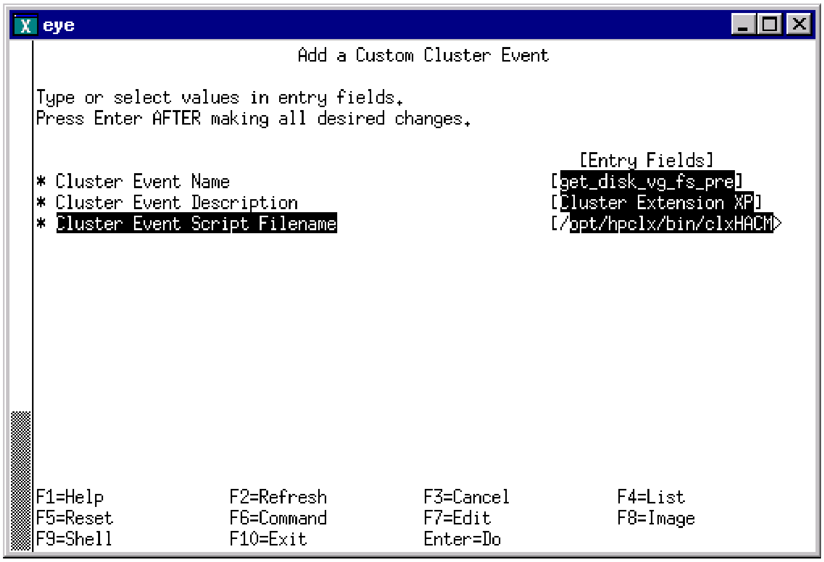
   Choose Cluster Configuration →
   Cluster Resources →
   Cluster Events →
   Define Custom Cluster Events →
   Add a Custom Cluster Event.

2. Enter values:

   Cluster Event Name: **get_disk_vg_fs_pre**

   Cluster Event Description: **Cluster Extension XP**

   Cluster Event Script File: **/opt/hpclx/bin/clxhacmp**



```
                         Add a Custom Cluster Event

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                     [Entry Fields]
* Cluster Event Name                                [get_disk_vg_fs_pre]
* Cluster Event Description                         [Cluster Extension XP]
* Cluster Event Script Filename                     [/opt/hpclx/bin/clxHACM>









F1=Help             F2=Refresh          F3=Cancel           F4=List
F5=Reset            F6=Command          F7=Edit             F8=Image
F9=Shell            F10=Exit            Enter=Do
```

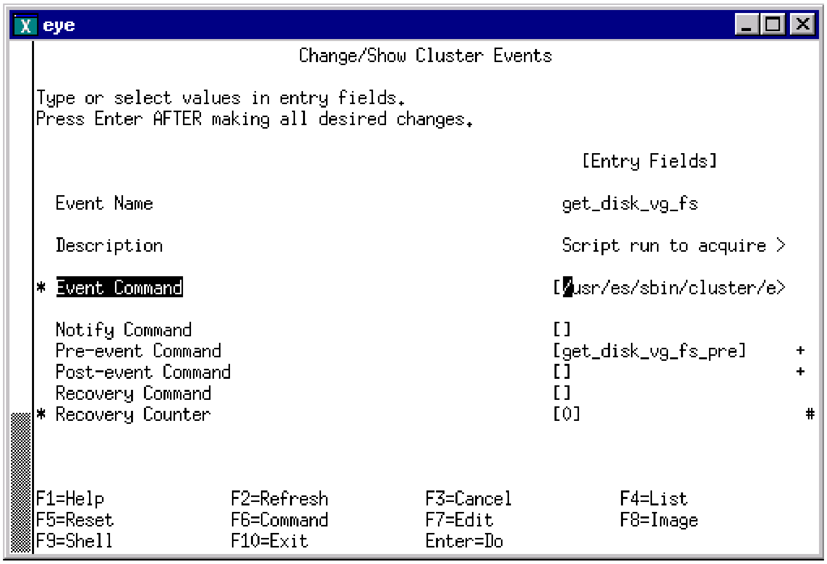
3. Configure the previously defined Custom Cluster Event as pre-event of **get_disk_vg_fs**.

   #smitty hacmp

4. Choose
   Cluster Configuration →
   Cluster Resources →
   Cluster Events →
   Change/Show Cluster Events.

   Select event **get_disk_vg_fs**.

   Define the previously defined custom event **get_disk_vg_fs_pre** as a pre-event of **get_disk_vg_fs**.



```
                     Change/Show Cluster Events

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                        [Entry Fields]

  Event Name                                          get_disk_vg_fs

  Description                                         Script run to acquire >

* Event Command                                      [/usr/es/sbin/cluster/e>

  Notify Command                                     []
  Pre-event Command                                  [get_disk_vg_fs_pre]     +
  Post-event Command                                 []                       +
  Recovery Command                                   []
* Recovery Counter                                   [0]                       #



F1=Help             F2=Refresh          F3=Cancel           F4=List
F5=Reset            F6=Command          F7=Edit             F8=Image
F9=Shell            F10=Exit            Enter=Do
```

Cluster Extension XP controls the disk pairs based on RAID Manager XP device groups. The volume group definition of the HACMP resource group is used to determine the corresponding RAID Manager XP device group. The mapping of the HACMP volume group configuration and the corresponding RAID Manager XP device group is done by the Cluster Extension XP user configuration file **/etc/opt/hpclx/config/UCF.cfg**. Because of this mapping mechanism, you must specify the volume groups owned by the HACMP resource groups in the user configuration file.

User configuration file for HACMP

Prior to configuring the objects in the user configuration file, review the Cluster Extension XP objects described in “User configuration file and Cluster Extension XP objects” (page 65).

Set the **ApplicationStartup** object to **RESYNCWAIT**, because HACMP does not offer the feature to disable resource groups on a particular system in order to move the resource group back to the most current copy of your data. If the **ApplicationStartup** object is set to **FASTFAILBACK** (default), the resource group will fail to be brought online in cases where the most current copy of your data resides in the disk array on the remote site. If you set the **ApplicationStartup** object to **FASTFAILBACK**, you must stop the resource group online process and either resynchronize your disk from the remote site or manually bring your resource group online at the remote site.

The **Vgs** object in the user configuration file of Cluster Extension XP is used to map the volume groups of the HACMP resource group to the corresponding **APPLICATION** object within the user configuration file. The **DeviceGroup** object within each **APPLICATION** section determines the RAID Manager XP device group needed to control all the shared disks of the HACMP resource group.

Figure 6 (page 102) shows two examples of possible mappings.

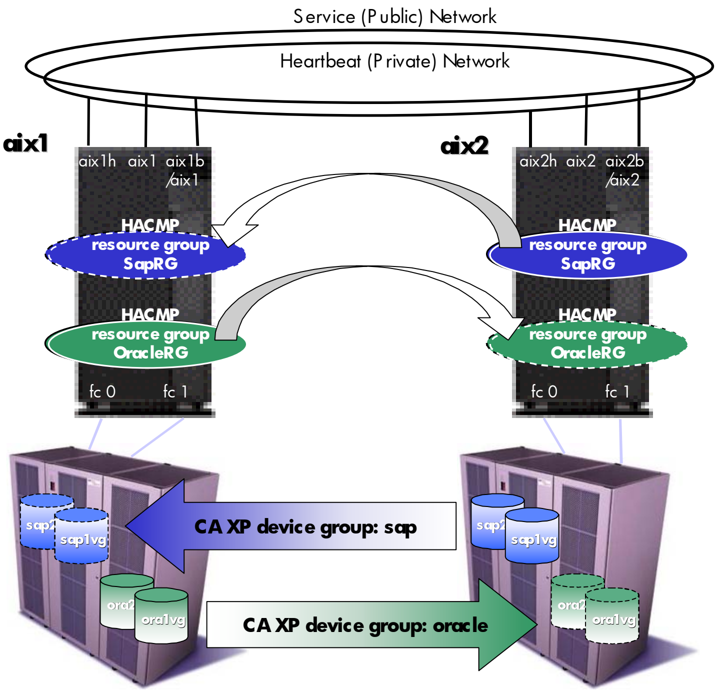

Service (Public) Network
Heartbeat (Private) Network
aix1
aix2
aix1h aix1 aix1b
/aix1
aix2h aix2 aix2b
/aix2
HACMP resource group SapRG
HACMP resource group SapRG
HACMP resource group OracleRG
HACMP resource group OracleRG
fc 0
fc 1
fc 0
fc 1
sap2
sap1vg
sap2
sap1vg
ora2
ora1vg
ora2
ora1vg
CA XP device group: sap
CA XP device group: oracle

Figure 6. Configuration example

Example 1 The application **OracleRG** corresponds to a HACMP resource group **OracleRG**, which consists of the volume groups **ora1vg** and **ora2vg**. The corresponding RAID Manager XP device group **oracle** controls all disks, which form the volume groups of the HACMP resource group. The resource group is configured to wait for a pair resynchronization in case you have not done any disk pair recovery after the resource group has been moved to an alternative system. It will now be brought online on the local system again (**ApplicationStartup** object is set to **RESYNCWAIT**). The **AutoRecover** object is set to **NO**, which means that you will not utilize Cluster Extension XP capabilities to automatically recover suspended disk pair states. The **DataLoseMirror** object and **DataLoseDataCenter** object are set to **NO**, which means Cluster Extension XP will not allow you to bring the resource group online if the disk pair is suspended or a takeover operation leads into a suspended disk pair.

Example 2 The application **SapRG** uses the device group **sap** to control all the disks of the corresponding HACMP resource group **SapRG**, which uses the volume groups **sap1vg** and **sap2vg**. The resource group is configured to failback to the remote system rather than waiting for a pair resynchronization in case you have not done any disk pair recovery after the resource group has been moved to alternative system. It will now be brought online on the local system again (**ApplicationStartup** object is set to **FASTFAILBACK** per default). This setup will lead into an error loop, since HACMP does not provide the feature to automatically failback after an error has been reported. The **AutoRecover** object is set to **NO** per default, which means that you will not utilize Cluster Extension XP capabilities to automatically recover suspended disk pair states.

```
COMMON
    LogDir        /var/opt/hpclx/log/ #default (optional)
    LogLevel      error             # error|info default: error (optional)
APPLICATION           OracleRG    # package/service group test_application
Vgs                   ora1vg ora2vg # HACMP specific, to map vg to OracleRG
ApplicationDir        /etc/opt/hpclx
XPSerialNumbers       30368 30380
RaidManagerInstances  11
DeviceGroup           oracle        # raid manager device group
FenceLevel            data          # values: data | never | async
ApplicationStartup    resyncwait    # values: fastfailback | resyncwait
AutoRecover           no            # possible values: yes | no
DataLoseMirror        no            # possible values: yes | no
DataLoseDataCenter    no            # possible values: yes | no
PreExecScript         /etc/opt/hpclx/ora_pre.sh
PostExecScript        /etc/opt/hpclx/ora_post.sh

APPLICATION           SapRG     # package/service group test_application
Vgs                   sap1vg sap2vg # HACMP specific, to map vg to SapRG
XPSerialNumbers       30368 30380
RaidManagerInstances  11
DeviceGroup           sap           # raid manager device group
FenceLevel            never         # possible values: data | never | async
```

Bringing a resource group online

Resource groups will usually be brought online automatically when the cluster is started on a particular system.

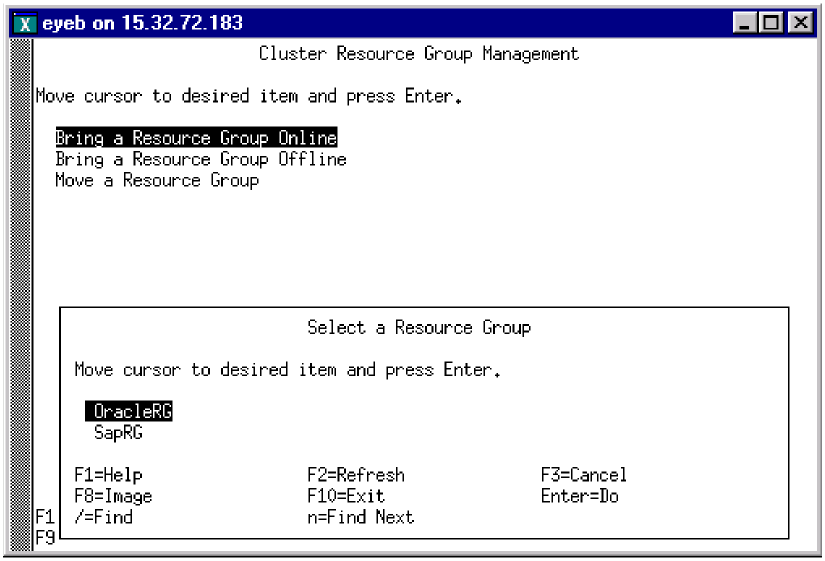
To bring a resource group online manually

1. Run SMIT (HACMP section).

   #smitty hacmp

2. Choose:
   Cluster System Management →
   Cluster Resource Group Management →
   Bring a Resource Group Online

3. Select the resource group.



```
                    Cluster Resource Group Management

Move cursor to desired item and press Enter.

  Bring a Resource Group Online
  Bring a Resource Group Offline
  Move a Resource Group



                         Select a Resource Group

    Move cursor to desired item and press Enter.

      OracleRG
      SapRG

    F1=Help                 F2=Refresh              F3=Cancel
    F8=Image                F10=Exit                Enter=Do
F1  /=Find                  n=Find Next
F9
```

Taking a resource group offline

Resource groups will usually be taken offline automatically when the cluster is stopped on a particular system.

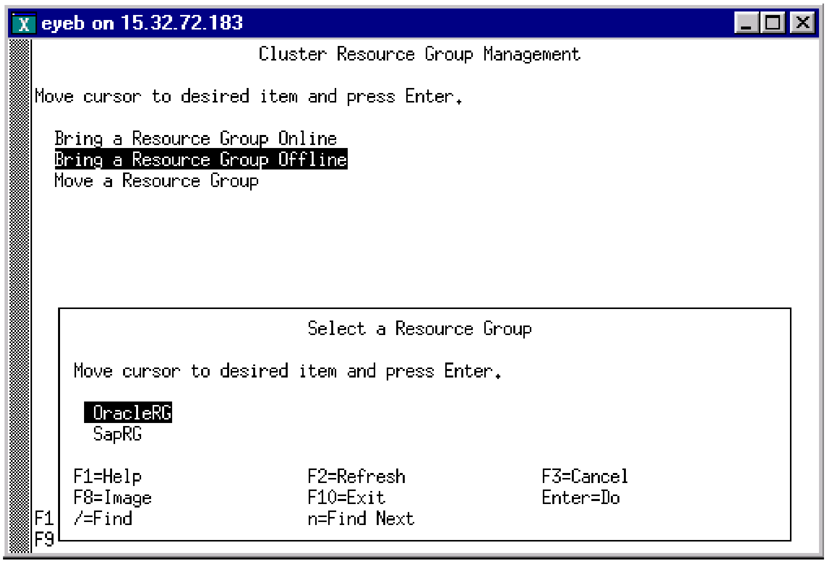
To bring a resource group offline manually:

1. Run SMIT (HACMP section).

   #smitty hacmp

2. Choose:
   Cluster System Management →
   Cluster Resource Group Management →
   Bring a Resource Group Offline

3. Select the resource group.



eyeb on 15.32.72.183

```
                        Cluster Resource Group Management

Move cursor to desired item and press Enter.

  Bring a Resource Group Online
  Bring a Resource Group Offline
  Move a Resource Group



    Select a Resource Group

    Move cursor to desired item and press Enter.

      OracleRG
      SapRG

    F1=Help                 F2=Refresh              F3=Cancel
    F8=Image                F10=Exit                Enter=Do
F1  /=Find                  n=Find Next
F9
```

Deleting Cluster Extension XP

To delete Cluster Extension XP, you must remove the pre-event entries for Cluster Extension XP from the **get_disk_vg_fs** and **release_vg_fs** events. This removes all references to Cluster Extension XP; then you can deinstall Cluster Extension XP.

Caution *Deleting the Cluster Extension XP integration from an online resource group or cluster does not remove the resource_name**.online** file and does not remove the device group from the list of monitored device groups if the pair/resync monitor is used to monitor the Continuous Access XP link. Therefore, the device group must be deleted from the list of monitored device groups manually using the **clxchkmon** command after deleting the Cluster Extension XP resource. See "Stopping the pair/resync monitor"* (page 198).

Caution *Failure to delete the monitored device group from the list of monitored device groups can cause data corruption if the **ResyncMonitorAutoRecover** attribute is set to **YES**.*

1. Enter the following command to start the deinstallation process.

   #smitty deinstall

2. Select Software and press F4 to select the Cluster Extension XP components you need to deinstall from the system. Then press Enter.
3. When the deinstallation process has finished, press F10 to exit SMIT.

Pair/resync monitor integration

The pair/resync monitor is used to detect and react on suspended Continuous Access links. It is activated if the **ResyncMonitor** object is set to **YES**. The automatic disk pair resynchronization feature is activated if the **ResyncMonitorAutoRecover** object value is **YES**, additionally.

When the HACMP resource group is taken offline, the monitor must be disabled for the RAID Manager XP device group used for this resource group.

Caution *If the resource group cannot be taken offline gracefully, the cluster administrator must disable monitoring of the device group for this HACMP resource group. To avoid data corruptions, this task must be part of the recovery procedure when Cluster Extension XP is deployed in the HACMP environment. See "Stopping the pair/resync monitor"* (page 198).

Ensure that the pair/resync monitor does not monitor and resynchronize the disk pair (device group) from both disk arrays sites.

To use the pair/resync monitor you must create a pre-event for the **release_vg_fs** event. If the resource group will be taken offline on a cluster system, the corresponding application/device group will be taken from the list of monitored device groups and the monitoring will be disabled.

Two steps must be performed to adjust the event handling. First, an additional Custom Cluster Event must be specified, and then this event must be configured as a pre-event for the standard event **release_vg_fs**.

To add a Custom Cluster Event to your HACMP cluster:

1. Run SMIT (HACMP section).

   #smitty hacmp

2. Choose
   Cluster Configuration →
   Cluster Resources →

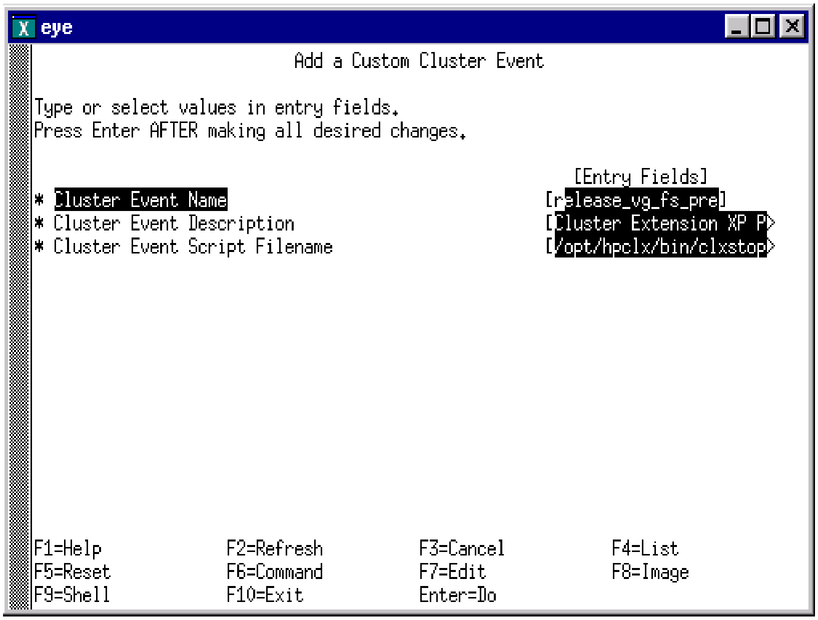
Cluster Events →
Define Custom Cluster Events →
Add a Custom Cluster Event.

3. Enter values.

   Cluster Event Name: **release_vg_fs_pre**

   Cluster Event Description: **Cluster Extension XP Pre-Event**

   Cluster Event Script File: **/opt/hpclx/bin/clxstopmonhacmp**



```
                        Add a Custom Cluster Event

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                      [Entry Fields]
* Cluster Event Name                                [release_vg_fs_pre]
* Cluster Event Description                         [Cluster Extension XP P>
* Cluster Event Script Filename                     [/opt/hpclx/bin/clxstop>
```

Define the previously defined Custom Cluster Event as a pre-event of the standard HACMP event release_vg_fs:

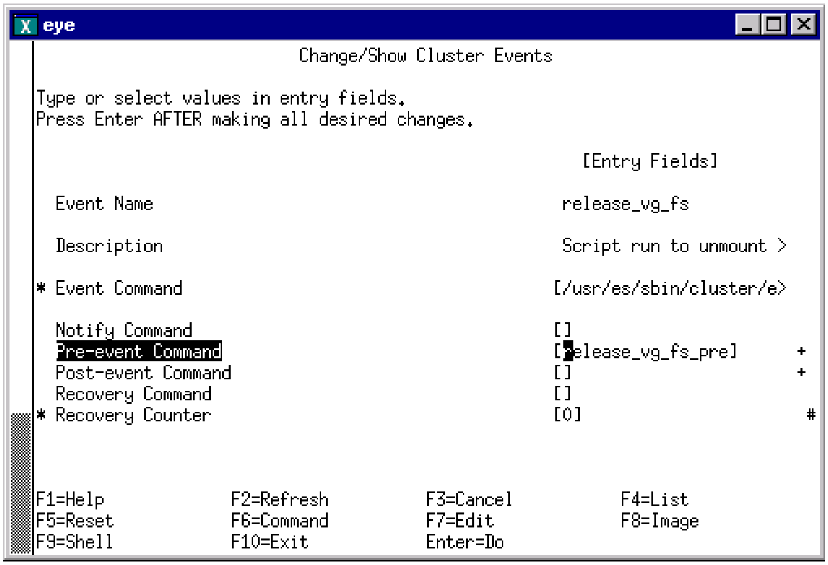
1. Run SMIT (HACMP section).

   #smitty hacmp

2. Choose:
   Cluster Configuration →
   Cluster Resources →
   Cluster Events →
   Change/Show Cluster Events

3. Select event **release_vg_fs**.

4. Define the previously defined Custom Cluster Event **release_vg_fs_pre** as pre-event of **release_vg_fs**.



```
                          Change/Show Cluster Events

Type or select values in entry fields.
Press Enter AFTER making all desired changes.

                                                      [Entry Fields]

  Event Name                                        release_vg_fs

  Description                                       Script run to unmount >

* Event Command                                    [/usr/es/sbin/cluster/e>

  Notify Command                                   []
  Pre-event Command                                [release_vg_fs_pre]      +
  Post-event Command                               []                       +
  Recovery Command                                 []
* Recovery Counter                                 [0]                       #


F1=Help             F2=Refresh          F3=Cancel           F4=List
F5=Reset            F6=Command          F7=Edit             F8=Image
F9=Shell            F10=Exit            Enter=Do
```

Timing considerations

Cluster Extension XP is designed to prefer XP disk array operations over cluster software operations. If Cluster Extension XP invokes disk pair resynchronization operations or gathers information about the remote XP disk array, Cluster Extension XP will wait until the requested status information is reported. This assumption has been made to clearly prioritize data integrity over cluster software's failover behavior.

In some cases, however, this behavior could lead into an HACMP error event (**config_too_long**). The default timeout value is 6 minutes. To increase the timeout, use the **chssys** command; you must stop the cluster to run this command.

Example

```
#chssys -s clstrmgr -a "-u 60000"
```

The timeouts described above can occur in the following situations:

- Cluster Extension XP uses RAID Manager XP instances to communicate with the remote XP disk array. Depending on the settings of the RAID Manager XP instance timeout parameter and the number of remote instances, the online operation could time out. This can happen if the local RAID Manager XP instance cannot reach the remote RAID Manager XP instance. See “RAID Manager XP Configuration” for further details.
- Cluster Extension XP tries to resynchronize disk pairs and waits until the RAID Manager XP device group is in PAIR state if the **ApplicationStartup** attribute is set to **RESYNCWAIT**. Depending on the RAID Manager XP version and the XP firmware version, this could be a full resynchronization and can take much longer than the online timeout interval. Even if the RAID Manager XP version and the XP firmware version allow a delta resynchronization, the delta between the primary and the secondary could be big enough for the copy process to exceed the online timeout value.
- If running in fence level **ASYNC**, the default value of the **AsyncTakeoverTimeout** can cause the resource group online process to fail because its value is set to a very high value. This is done

because of the fact that the takeover process for fence level **ASYNC** can take longer when slow communications links are in place.

To prevent takeover commands from being terminated by the takeover timeout before it is finished, the time to copy the installed XP disk array cache should be measured. To measure the copy time, use only the slowest link used for Continuous Access XP Extension. This ensures that the XP disk array cache can be transferred from the remote XP disk array, even in the event of a single surviving replication link between the XP disk arrays.

Because the failover environment is dispersed into two (or more) data centers, the failover time can not be expected to be the same as it would be in a single data center with a single shared disk device.

Failure behavior

Cluster Extension XP will run into an endless loop if a configuration error or if XP disk pairs states have been discovered, which does not allow automated actions. This is logged within the logfiles:

/var/opt/hpclx/log/clxhacmp.log

/tmp/hacmp.out

To return control back to the HACMP cluster software, the user must remove the lock file:

*application_dir/application_name***.LOCK**

Example **/etc/opt/hpclx/OracleRG.LOCK**

This process has been adopted from HACMP's behavior. HACMP will also run in an endless loop in the case of a failure until the user recovers all errors and manually starts the application. After all errors have been recovered, invoke the command **clruncmd** to return control back to the cluster software.

Restrictions for IBM HACMP with Cluster Extension XP

The following is a summary of restrictions that apply for HACMP configurations when Cluster Extension XP is used to enable failover between two XP disk arrays.

- The **FastFailbackEnabled** object is not used by the Cluster Extension XP integration with HACMP.
- Cluster Extension XP must not be used with concurrent resource group configurations (for example, parallel databases).
- Cluster Extension XP must not be used with raw devices without volume groups.
- Target mode SCSI: Serial networks within HACMP based on target mode SCSI (TMSCSI) are not supported.

6

Integration with Microsoft Cluster service

HP StorageWorks Cluster Extension XP for Microsoft Cluster service integrates with Microsoft Cluster service as a resource DLL. This allows cluster administrators to configure the disk array-specific failover behavior as easily as any other resource DLL in the Cluster service.

Related information

For information about how to install Cluster Extension XP, see the *HP StorageWorks Cluster Extension XP: Installation Guide*.

In addition to the Cluster Extension XP resource DLL, HP implemented a set of software that handles the specific needs of the quorum disk in the cluster group.

Configuring the quorum service

The Cluster Extension XP quorum service has few configurable parameters. Most parameters are set during the quorum service installation.

Please refer to Chapter 2, “Quorum processes” on page 60 for start up options and to Chapter 11 “Solving quorum service problems” on page 211 for configurable parameters.

Configuring Cluster Extension XP resources

Cluster Extension XP objects are configurable as properties of the specific Cluster Extension XP resource. The value of a Cluster Extension XP object can be compared with the value of a Microsoft Cluster service resource property value.

If you use the default values for Cluster Extension XP **COMMON** objects, no user configuration file is required.

The Cluster Extension XP resource gathers all necessary information about the resource group and the disk arrays if the Cluster Extension XP resource is brought online.

If configured, a pair/resync monitor is started. If the resource is not configured to use the pair/resync monitor, Cluster Extension XP creates a file:

> *resource_name***.online**

This file is created in the directory specified by the **ApplicationDir** resource property.

If the resource is taken offline, this file will be removed, or the device group associated with the resource group will be removed from the pair/resync monitor list. If the device group is the last monitored disk pair, the monitor will be stopped also.

Cluster Extension XP provides a RAID Manager XP service, which automatically starts RAID Manager XP instances at system boot time. This feature reduces resource group failover times because the Cluster Extension XP resource does not need to start the RAID Manager XP instances.

The Cluster Extension XP for Microsoft Cluster service resource can be configured using either the Microsoft Cluster service command-line interface or the Cluster Administrator GUI.

After installing Cluster Extension XP for Microsoft Cluster service, you create a new Cluster Extension XP resource by adding a resource of the type Cluster Extension XP to your desired resource group. The Cluster Extension XP resource type definition is preconfigured to fit most Continuous Access configurations.

Resource group and resource names

Cluster Extension XP resource names and resource group names must consist of one word. The pair/resync monitor is not able to interface with resources or resource groups that include space characters in their names.

Underscore or hyphen characters can be used to separate different words.

Example **CLX_resource**

Not **CLX resource**

Cluster Extension XP resource-specific parameters

Prior to configuring the resource properties, review the Cluster Extension XP objects described in “User configuration file and Cluster Extension XP objects” (page 65).

Setting non-Cluster Extension XP resource-specific parameters

Microsoft allows specific failover parameter and threshold values to be set by the user for a resource group as well as for a resource.

These values must be changed in a Cluster Extension XP environment to be able to take certain manual recovery actions in case of a disaster.

Resource properties and values

Cluster Extension XP requires that you change the values of the following resource properties provided by Microsoft Cluster service.

IsAlivePollInterval (page 121)
LooksAlivePollInterval (page 121)
PendingTimeout (page 121)
RestartAction (page 121)
RestartPeriod (page 122)
RestartThreshold (page 122)

Resource group properties and values

Cluster Extension XP requires that you change the values of the following group properties provided by Microsoft Cluster service.

AutoFailbackType (page 120)
FailoverPeriod (page 120)
FailoverThreshold (page 120)

AutoFailbackType

Format integer

Description Prevents automatic fail back of a resource group to its primary system. The resource group should be transferred back manually after the failure has been recovered. This allows for recovery of all possible failure sources and pair resynchronization (if necessary) while the application service is still running.

Required value **0**

FailoverPeriod

Format string

Description Determines time (in hours) over which the cluster service attempts to failover a resource group. This value is the Microsoft default value. See "Timing considerations for Microsoft Cluster service" (page 137).

Default value **6**

FailoverThreshold

Format integer

Description Determines number of failover attempts. The recommended setting allows the cluster service to transfer the resource group to each system once in case of subsequent system failure. Because of the nature of this parameter, it is possible that the resource group automatically restarts on a system several times if not all cluster systems are members of the cluster at that time.

If this value will be set to a number higher than the current number of clustered systems for the cluster group, the resource group will restart until either the **FailoverThreshold** value or the **FailoverPeriod** timeout value will be reached.

Valid values Number of systems in the resource group node list minus 1
10 *(default)*

IsAlivePollInterval

Format integer

Description Used to poll “Alive” state for the resource (in milliseconds). Decreasing this value will allow for faster resource failure detection but will also consume more system resources.

Default value **60000**

LooksAlivePollInterval

Format integer

Description Used to poll “Alive” state for the resource (in milliseconds). Cluster Extension XP calls the **IsAlive** routine with each **LookAlive** routine, automatically. Therefore, it does not make sense to set both interval values to different values. Decreasing this value will allow for faster resource failure detection but will also consume more system resources.

Default value **60000**

PendingTimeout

Format integer

Description Used to specify the timeout for status resolution (in seconds). This value is the Microsoft default value. See “Timing considerations for Microsoft Cluster service” (page 137).

Default value **180**

RestartAction

Format integer

Description Defines whether a resource can be automatically restarted after it has failed.

The value must “affect the group.” This ensures that the resource group fails over to another system in case of a resource is reported **FAILED**.

Valid values **0** (do not restart)
1 (restart)
2 (restart and affect the group) (default)

RestartPeriod

Format integer

Description Determines amount of time for restart (in seconds). This value is the Microsoft default value. It will not be used if the **RestartThreshold** is set to 0.

Default value **900**

RestartThreshold

Format integer

Description Determines number of restart attempts. This value must be set to 0. It does not make sense to restart the Cluster Extension XP resource on the same system several times without fixing the problem.

Required value **0**

Adding a Cluster Extension XP resource

Using the Microsoft Cluster service command line to add a Cluster Extension XP resource:

Syntax **cluster resource** *resource_name* **/create /group:** *resource_group_name* **/type:"Cluster Extension XP"**

Example This example adds a Cluster Extension XP resource called **clx_fileshare** to the **CLX_SHARE** resource group.

```
C:\>cluster resource clx_fileshare /create
/group:CLX_SHARE /type:"Cluster Extension XP"
```

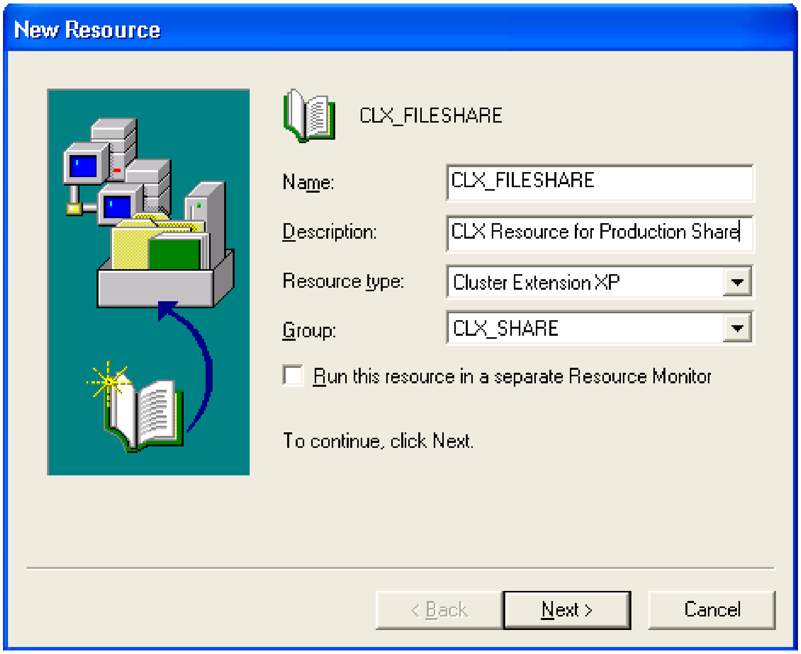
For information about naming restrictions, see “Resource group and resource names” (page 118).

Using the Microsoft Cluster service Cluster Administrator GUI to add a Cluster Extension XP resource:

1. Open Cluster Administrator.
2. From the File menu, choose New → Resource.
3. Enter values in fields:

   Name: **clx_fileshare**
   Description: as appropriate for the resource

Resource type: select **Cluster Extension XP** from the list
Group: select **CLX_SHARE** from the list



Name: CLX_FILESHARE

Description: CLX Resource for Production Share

Resource type: Cluster Extension XP

Group: CLX_SHARE

□ Run this resource in a separate Resource Monitor

To continue, click Next.

4. Choose the Next button.
5. Add or remove resource owners.
6. Choose the Next button.

   Do not add any dependencies.
7. Choose the Next button.
8. Modify resource property values of the new Cluster Extension XP resource **clx_fileshare** as needed.
9. Choose the Finish button to exit the wizard window.

Changing Cluster Extension XP resource properties

Cluster Extension XP resource properties can be changed any time. However, changes take effect only when the resource is brought online the next time.

To change Cluster Extension XP resource properties, you must take the resource offline.

Caution *Do not change any property of the Cluster Extension XP resource while the resource is running.*

Command-line syntax

The Microsoft Cluster service default properties for the resource can be changed, using the following syntax:

Syntax **cluster resource** *resource_name* **/privprop** *object_name*={*value* | **"***value1 value2 . . .***"**}

You can display all attributes of the Cluster Extension XP resource **clx_fileshare** with the following command:

Example **cluster resource clx_fileshare /privprop**

Example This example affects a Cluster Extension XP resource called **clx_fileshare** to change the default **FenceLevel** property:

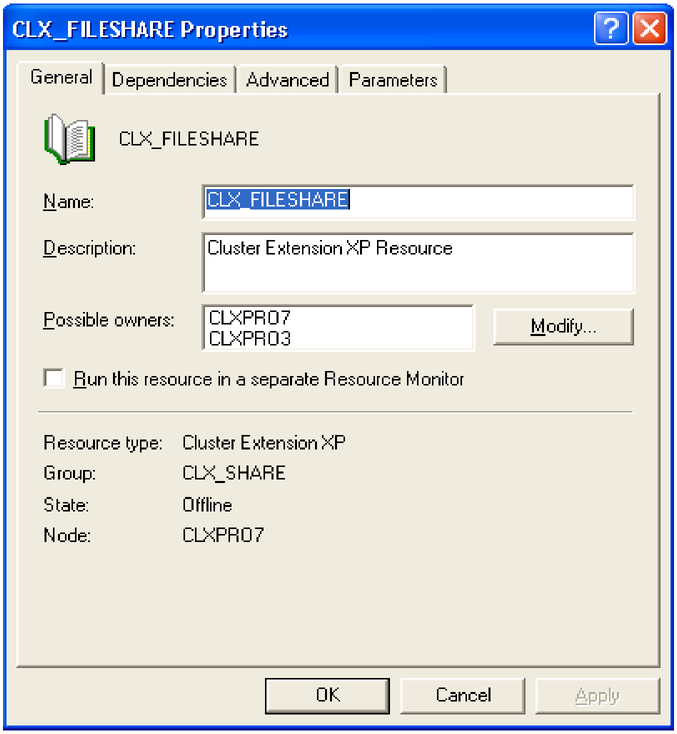
```
C:\>cluster resource clx_fileshare /privprop
FenceLevel=data
```

Example This example changes the RAID Manager XP instance used for the Cluster Extension XP resource **clx_fileshare** from 10 to 99 and then adds an additional instance (22) to provide redundancy.

```
C:\>cluster resource clx_fileshare /privprop
RaidManagerInstances="99 22"
```

To change property values of the Cluster Extension XP resource from the Microsoft Cluster service Cluster Administrator GUI:

1. Open Cluster Administrator.
2. Double-click on the Resource folder in the console-tree.
3. Double-click the **clx_fileshare** resource from the details pane.
4. Click on each tab and set the properties for the **clx_fileshare** resource.
5. Modify resource property values of the new Cluster Extension XP resource **clx_fileshare** as needed.
6. Choose the **APPLY** button to finish your modifications.
7. Choose the Finish button to exit the wizard window.



CLX_FILESHARE Properties

CLX_FILESHARE

Name: CLX_FILESHARE

Description: Cluster Extension XP Resource

Possible owners: CLXPR07
CLXPR03

☐ Run this resource in a separate Resource Monitor

Resource type: Cluster Extension XP

Group: CLX_SHARE

State: Offline

Node: CLXPR07

Advanced properties

The Parameters tab of the Cluster Extension XP resource offers the basic settings and is used to enter environment data, such as the disk array serial numbers and RAID Manager XP instances. The more advanced settings can be accessed through additional buttons in the Parameters tab.

Changing fence level-specific values

The fence level specific values of Cluster Extension XP can be changed by click on the **Advanced** button. The window allows changing these Cluster Extension XP objects:

AsyncTakeoverTimeout (page 77)
DataLoseDataCenter (page 80)
DataLoseMirror (page 81)

Changing failover and failback behavior

The failover/failback behavior of Cluster Extension XP can be changed by choosing the **Failover/Failback** button. The window allows changing these Cluster Extension XP objects:

ApplicationStartup (page 75)
AutoRecover (page 78)
ResyncWaitTimeout (page 86)

Activating the pair/resync monitor

The pair/resync monitor can be activated and deactivated by click on the **PAIR/RESYNC Mon** button. The window allows changing these objects:

ResyncMonitor (page 85)
ResyncMonitorAutoRecover (page 85)
ResyncMonitorInterval (page 85)

Configuring takeover actions

Pre-executables and post-executables can be defined to be executed before or after Cluster Extension XP invokes its takeover functions. The window allows changing these objects:

PostExecCheck (page 84)
PostExecScript (page 84)
PreExecScript (page 84)

Changing a resource name

Caution *Do not change the name of the Cluster Extension XP resource at any time. Changing the name of a Cluster Extension XP resource does not change or update the internal Cluster Extension XP database. Cluster Extension XP continues to use the old resource name. If a Cluster Extension XP resource name has been changed and a new Cluster Extension XP resource is created with the old resource name, the resources will impact each other. This could cause resources to go offline unexpectedly.*

To change a resource name, first delete the Cluster Extension XP resource and then create a new resource with a new name.

Adding dependencies on a Cluster Extension XP resource

Cluster Extension XP must be the first resource in the resource chain of a Microsoft Cluster service resource group.

All resources that depend on the disk resource, such as a file share, including all disk resources (physical disks) must be configured for dependency on the Cluster Extension XP resource.

To add a dependency on a Cluster Extension XP resource to another resource, use this syntax:

Syntax **cluster resource** *physical_disk_resource* /**adddependency:** *ClusterExtensionXP_resource*

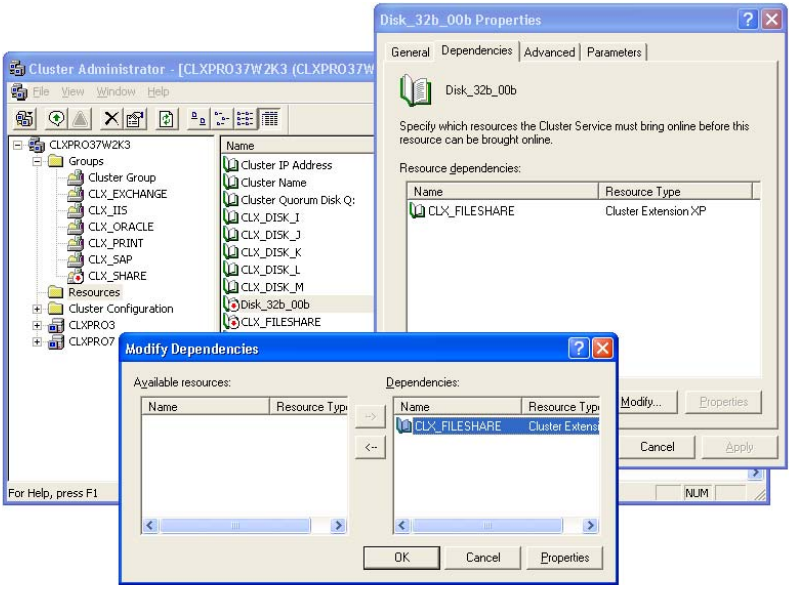
Example This command example adds a dependency on the Cluster Extension XP **clx_fileshare** resource to the physical disk resource **Disk_32b_00b**.

```
C:\>cluster resource Disk_32b_00b
/adddependency:clx_fileshare
```

To add a dependency on a Cluster Extension XP resource:

1. Open Cluster Administrator.
2. Double-click on the Resource folder in the console-tree.
3. Double-click the **Disk_32b_00b** resource from the details pane.
4. Click on the Dependencies tab and choose the Modify button.

5. Add the **clx_fileshare** resource to the Dependencies of the **Disk_32b_00b** resource.



Disk_32b_00b Properties

Disk_32b_00b

Specify which resources the Cluster Service must bring online before this resource can be brought online.

Resource dependencies:

Name	Resource Type
CLX_FILESHARE	Cluster Extension XP

Modify Dependencies

Available resources:

Name	Resource Type

Dependencies:

Name	Resource Type
CLX_FILESHARE	Cluster Extens

6. Choose the OK button to finish your modifications.

Bringing a Cluster Extension XP resource online

Resources are usually brought online automatically when the resource group is brought online. You might need to move the resource group to the node where you want to bring the resource online. Bringing a resource in the resource group online causes resources on which the resource depends to go online also.

The following commands are used to bring a Cluster Extension XP resource online:

Syntax **cluster resource** *ClusterExtensionXP_resource* /**online:** *system_name*

You can bring the Cluster Extension XP resource **clx_fileshare** in the resource group **CLX_SHARE** online with the following command:

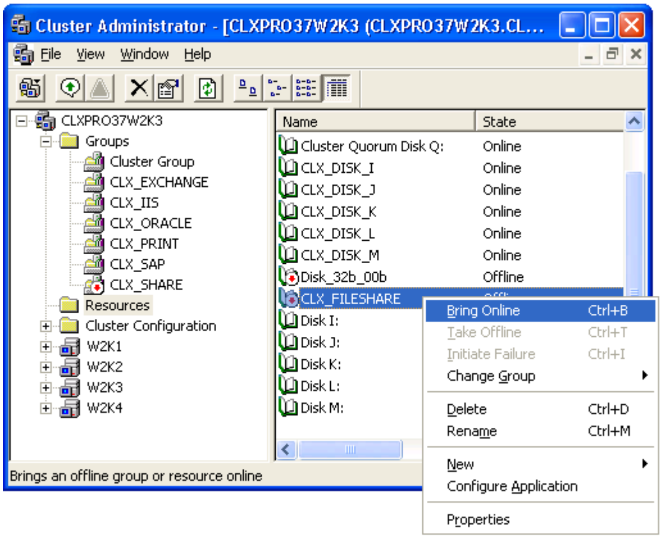
Example

```
C:\>cluster resource clx_fileshare /online:w2k1
```

To bring a Cluster Extension XP resource online from the Microsoft Cluster service Cluster Administrator GUI:

1. Open Cluster Administrator.

2. Double-click on the Resources folder in the console-tree. Then click the resource **clx_fileshare** in the details pane.



Name	State
Cluster Quorum Disk Q:	Online
CLX_DISK_I	Online
CLX_DISK_J	Online
CLX_DISK_K	Online
CLX_DISK_L	Online
CLX_DISK_M	Online
Disk_32b_00b	Offline
CLX_FILESHARE	Offli
Disk I:	
Disk J:	
Disk K:	
Disk L:	
Disk M:	

Brings an offline group or resource online

3. In the File menu, select Bring Online.

Taking a Cluster Extension XP resource offline

Resources are usually taken offline automatically when the resource group is taken offline. Taking a resource offline causes resources that depend on that resource to go offline also.

To take a Cluster Extension XP resource offline, use this syntax:

Syntax **cluster resource** *ClusterExtensionXP_resource* /**offline**

To take the Cluster Extension XP resource **clx_fileshare** offline, use this syntax:

Example

```
c:\>cluster resource clx_fileshare /offline
```

To take a Cluster Extension XP resource offline from the Microsoft Cluster service Cluster Administrator GUI:

1. Open Cluster Administrator.
2. Double-click on the Resource folder in the console-tree.
3. Choose the resource in the details pane.
4. From the File Menu, choose Take Offline.

Deleting a Cluster Extension XP resource

Deleting a running Cluster Extension XP resource causes the resource and its dependents to go offline.

To remove a Cluster Extension XP resource from an existing resource group, use this syntax:

Syntax **cluster resource** *ClusterExtensionXP_resource* /**delete**

Example

```
C:\>cluster resource clx_fileshare /delete
```

Caution *Deleting a running Cluster Extension XP resource does not remove the resource_name.****online*** *file and does not remove the device group from the list of monitored device groups if the pair/resync monitor is used to monitor the Continuous Access XP link. Therefore, the device group must be deleted from the list of monitored device groups manually using the* ***clxchkmon*** *command after deleting the Cluster Extension XP resource. See "Stopping the pair/resync monitor"* (page 198).

Caution *Failure to delete the monitored device group from the list of monitored device groups can cause data corruption if the* ***ResyncMonitorAutoRecover*** *attribute is set to* ***YES****.*

To delete a Cluster Extension XP resource from the Microsoft Cluster service Cluster Administrator GUI:

1. Open Cluster Administrator.
2. Double-click on the Resources folder in the console-tree.
3. Select the resource **clx_fileshare** in the details pane.
4. From the File Menu, choose Delete.

Pair/resync monitor integration

The pair/resync monitor detects and responds to suspended Continuous Access links if the **ResyncMonitor** object is set to **YES**. If the **ResyncMonitorAutoRecover** object is set to **YES**, automatic disk pair resynchronization is also activated.

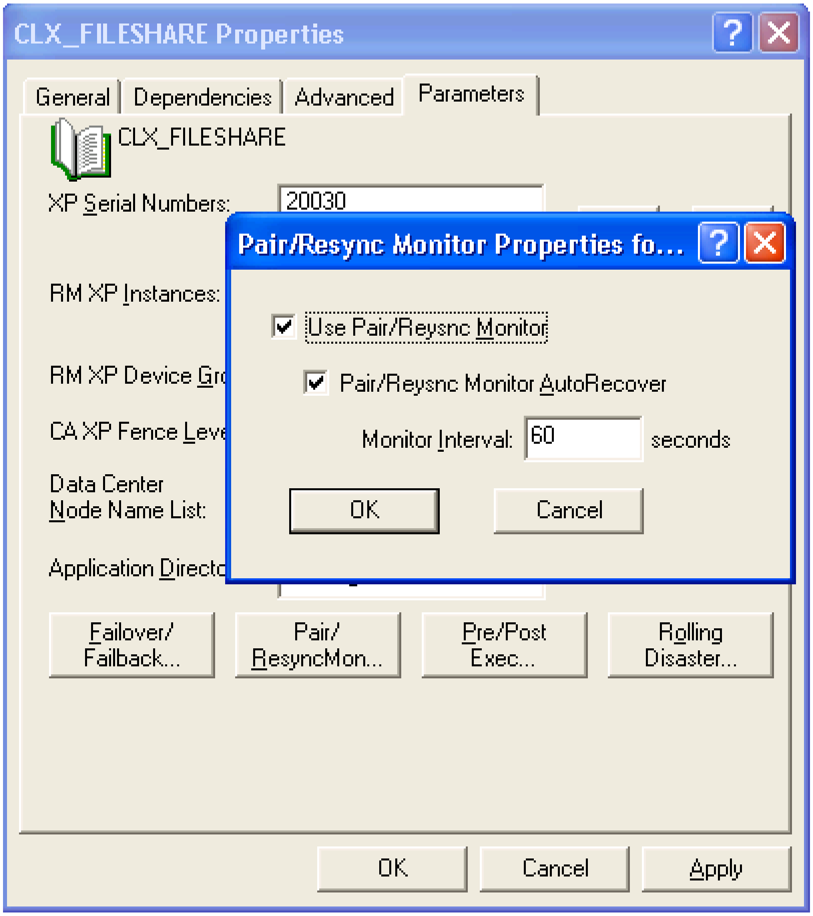
When the resource is taken offline, the monitor is stopped for the RAID Manager XP device group used for this resource.

Example Assuming that you have a Cluster Extension XP resource called **clx_fileshare**, you can use the following command to set the **ResyncMonitor** property from the Microsoft Cluster service command line:

```
C:\>cluster resource clx_fileshare /privprop ResyncMonitor=yes
```

To use the pair/resync monitor with the Cluster Extension XP resource **clx_fileshare**, you can take the following steps in the Microsoft Cluster service Cluster Administrator GUI:

1. Open Cluster Administrator.
2. Double-click on the Resource folder in the console-tree. Then double-click the **clx_fileshare** resource from the details pane. (You could also select the resource and click on Properties in the File menu.)
3. Click on the Parameter tab and click the **Pair/Resync Mon** button.
4. Check the **Use Pair/Resync Monitor** check box and change the monitor interval if necessary.
5. Click the OK button to finish your modifications.
6. Click the Finish button to exit the wizard window.



XP Serial Numbers: 20030

☑ Use Pair/Reysnc Monitor

☑ Pair/Reysnc Monitor AutoRecover

Monitor Interval: 60 seconds

Caution *If the resource group cannot be taken offline gracefully, the cluster administrator must disable monitoring of the device group for this resource. To avoid data corruption, this task must be part of the recovery procedure when Cluster Extension XP is deployed in the Microsoft Cluster service environment. See "Stopping the pair/resync monitor"* (page 198).

You must ensure that the pair/resync monitor does not monitor and resynchronize the disk pair (device group) from both disk arrays sites.

Timing considerations for Microsoft Cluster service

Cluster Extension XP gives priority to disk array operations over cluster software operations. If Cluster Extension XP invokes a disk pair resynchronization operation or gathers information about the remote disk array, Cluster Extension XP waits until the requested status information is reported. This ensures the priority of data integrity over cluster software failover behavior. However, this behavior can cause a failure of Cluster Extension XP resources in some cases:

- Cluster Extension XP uses RAID Manager XP instances to communicate with the remote disk array. Depending on the settings of the RAID Manager XP instance timeout parameter and the number of remote instances, the online operation could time out. This can occur if the local RAID Manager XP instance cannot reach the remote RAID Manager XP instance.
- Cluster Extension XP tries to resynchronize disk pairs and waits until the RAID Manager XP device group is in PAIR state if the ApplicationStartup resource property is set to **RESYNCWAIT**. In case of delta resynchronization, the delta between the primary and the secondary could be big enough for the copy process to exceed the resource **PendingTimeout** value.
- The **ResyncWaitTimeout** object can cause failed Cluster Extension XP resources when set above the resource **PendingTimeout** value.
- If running in fence level **ASYNC**, the default value of the **AsyncTakeoverTimeout** object can cause the resource to fail because its value exceeds the resource **PendingTimeout** value. The takeover process for fence level **ASYNC** can take much longer when slow communications links are in place.

  To prevent takeover commands from being terminated by the takeover timeout before completion, measure the time required to copy the installed disk array cache and adjust the resource **PendingTimeout** value according to the measured copy time. Use only the slowest link used for Continuous Access Extension to measure the copy time. This ensures that the disk array cache can be transferred from the remote disk array, even in the event of a single surviving replication link between the XP family disk arrays.

In general, because the failover environment is dispersed into two (or more) data centers, the failover time can not be expected to be the same as that in a single data center with a single shared disk device. Therefore, the resource group's **FailoverPeriod** value and the resource's values **RestartPeriod**, **PendingTimeout**, **LookAlive**, **IsAlive** of the Cluster Extension XP resource and the resource group using the Cluster Extension XP resource must be adjusted based on failover tests performed to verify the proper configuration setup.

The resource group's **FailoverPeriod** value must be higher than the resource's **RestartPeriod** value and both must be higher than the resource's **PendingTimeout** value.

Microsoft Cluster Server provides two parameters to adjust state change recognition/resolution:

- **IsAlive**
- **LookAlive**

Cluster Extension XP automatically calls the **IsAlive** function whenever the cluster service calls the **LookAlive** function. Therefore, both functions must be set to the same value.

Failure behavior with Microsoft Cluster service

Bouncing Resource Groups

Cluster Extension XP will alternate (start and fail) between local nodes if the **ApplicationStartup** object has been set to **FASTFAILBACK** and no remote system is available until the resource group restart limit has been reached. See **ApplicationStartup** (page 75).

The **FastFailbackEnabled** object is not used by the Cluster Extension XP integration with Microsoft Cluster service.

Unexpected offline conditions

In rare cases, ClusterExtensionXP resources go offline after the following conditions occur concurrently:

- ClusterExtensionXP resources have been online.
- The cluster resource monitor (**resrcmon.exe**) is restarted, for example, after the process has died or was killed.
- The CA XP links have failed.
- A remote RAID Manager XP instance is not available due to a full network outage.

Cluster Extension XP resources go offline because the primary volume state changes from PVOL_PAIR to PVOL_PSUE and the secondary volume state changes from SVOL_PAIR to EX_NORMT.

The state combination PVOL_PSUE and EX_NORMT is not designed to be handled automatically because the remote side (remote RAID Manager/ disk array), which has no state information available, could have more current data then the primary (PVOL_PSUE) site. In this particular case, the user is required to investigate data currency and to determine the appropriate action to be taken.

To bring the ClusterExtensionXP resources online:

1. Recover the CA XP link error and the network error, and restart RAID Manager on the remote site.
2. Manually resynchronize the affected disk pairs.
3. Bring the ClusterExtensionXP resources online.

or

1. Create the **forceflag** *resource_name***.forceflag** in the **ApplicationDir** path.

   Default:
   C:\Program Files\Hewlett Packard\Cluster Extension XP
2. Bring the ClusterExtensionXP resources online.
3. Depending on the attributes set for the resources, you might need to manually resynchronize the CA XP disk pairs.

Restrictions for Microsoft Cluster service with Cluster Extension XP

When Cluster Extension XP is used to enable failover between disk arrays, observe the following requirements for Microsoft Cluster service configuration:

- The Cluster Extension XP resource must be the first resource for all other disk resources.
- Only one Cluster Extension XP resource is allowed to be configured per cluster group.
- The resource's **PendingTimeout** value must be greater than the value specified for the **ResyncWaitTimeout** attribute.
- The resource's **PendingTimeout** value must be greater than twice the wait time of all remote RAID Manager XP instances times the number of remote systems. Otherwise, the Cluster Extension XP resource will fail to go online in case of a complete remote data center failure. If a post-executable is specified, the resource's **PendingTimeout** value must be greater than the number of remote systems times three times t_{WT}.

  t_{online} = resource online timeout

  $n_{remote\ sytems}$ = number of remote systems configured to run RAID Manager XP instances

  t_{WT} = wait time until remote error will be reported by local RAID Manager XP instance

  $t_{online} > n_{remote\ sytems}$ x 2 x t_{WT}

Related information

For configuration restrictions, review the Cluster Extension XP installation guide and the readme file on the installation CD-ROM.

Disaster-tolerant configuration example using a file share

The following example assumes that your environment consists of four systems (**w2k1**, **w2k2**, **w2k3** and **w2k4**), two XP disk arrays with serial numbers **35014** and **35013** and that you have configured **clxfileshare** as device group in the RAID Manager XP **c:\winnt\horcm101.conf** file and in the **c:\winnt\horcm102.conf** file. Furthermore, a pre-executable **clxpre.exe** will be invoked by Cluster Extension XP. You use the default failover behavior for the cluster group. The resource **CLX_FILESHARE** is part of the service group **CLX_SHARE** and must be brought online before the physical disk resources **Disk_32b_00b**.

The picture illustrates failover options and shows a second cluster group **CLX_IIS**.

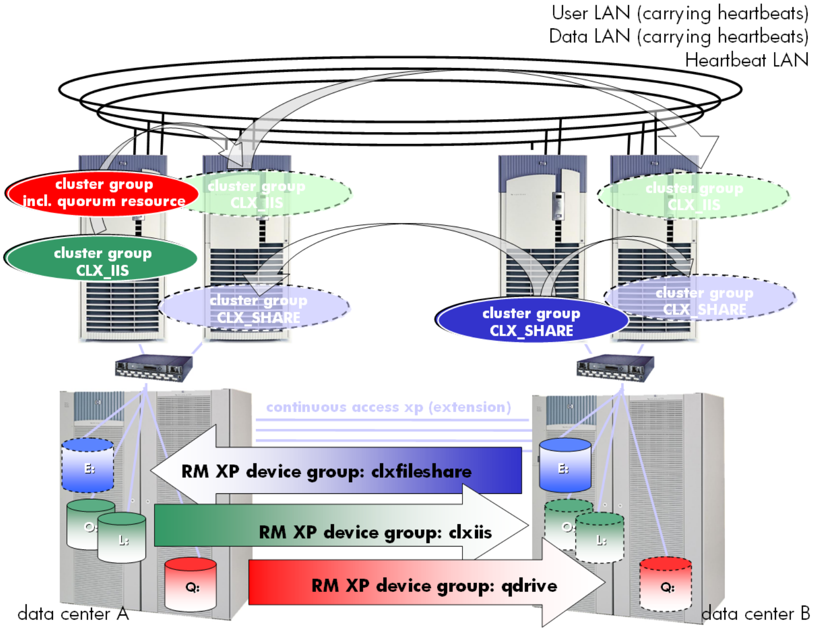
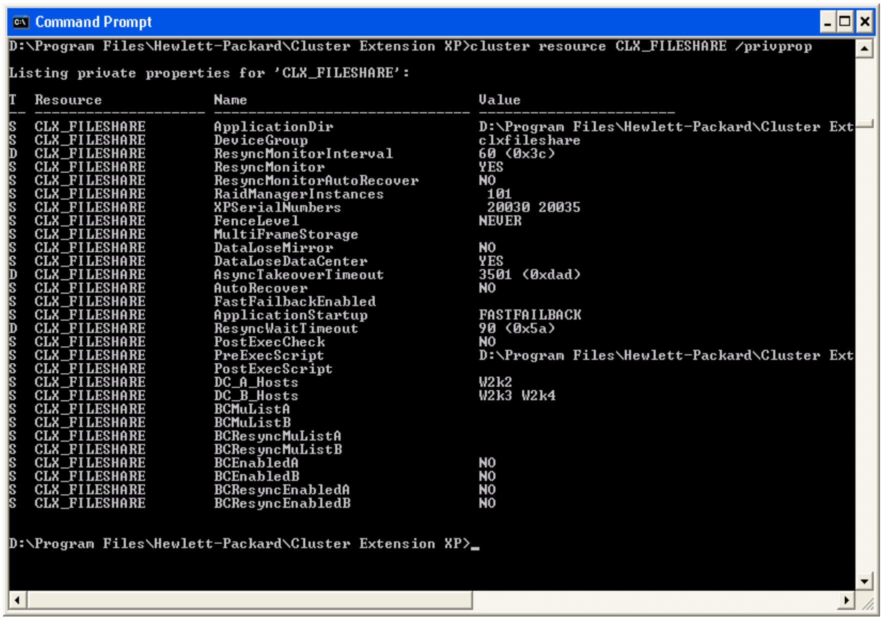

User LAN (carrying heartbeats)
Data LAN (carrying heartbeats)
Heartbeat LAN
cluster group incl. quorum resource
cluster group CLX_IIS
cluster group CLX_IIS
cluster group CLX_IIS
cluster group CLX_SHARE
cluster group CLX_SHARE
cluster group CLX_SHARE
continuous access xp (extension)
E:
O:
L:
Q:
E:
O:
L:
Q:
RM XP device group: clxfileshare
RM XP device group: clxiis
RM XP device group: qdrive
data center A
data center B

Figure 7. Resource group example (quorum service control disks not shown)

Example of the **CLX_FILESHARE** resource:



```
D:\Program Files\Hewlett-Packard\Cluster Extension XP>cluster resource CLX_FILESHARE /privprop

Listing private properties for 'CLX_FILESHARE':

T  Resource             Name                           Value
-- -------------------- ------------------------------ -----------------------
S  CLX_FILESHARE        ApplicationDir                 D:\Program Files\Hewlett-Packard\Cluster Ext
S  CLX_FILESHARE        DeviceGroup                    clxfileshare
D  CLX_FILESHARE        ResyncMonitorInterval          60 (0x3c)
S  CLX_FILESHARE        ResyncMonitor                  YES
S  CLX_FILESHARE        ResyncMonitorAutoRecover       NO
S  CLX_FILESHARE        RaidManagerInstances            101
S  CLX_FILESHARE        XPSerialNumbers                 20030 20035
S  CLX_FILESHARE        FenceLevel                     NEVER
S  CLX_FILESHARE        MultiFrameStorage
S  CLX_FILESHARE        DataLoseMirror                 NO
S  CLX_FILESHARE        DataLoseDataCenter             YES
D  CLX_FILESHARE        AsyncTakeoverTimeout           3501 (0xdad)
S  CLX_FILESHARE        AutoRecover                    NO
S  CLX_FILESHARE        FastFailbackEnabled
S  CLX_FILESHARE        ApplicationStartup             FASTFAILBACK
D  CLX_FILESHARE        ResyncWaitTimeout              90 (0x5a)
S  CLX_FILESHARE        PostExecCheck                  NO
S  CLX_FILESHARE        PreExecScript                  D:\Program Files\Hewlett-Packard\Cluster Ext
S  CLX_FILESHARE        PostExecScript
S  CLX_FILESHARE        DC_A_Hosts                     W2k2
S  CLX_FILESHARE        DC_B_Hosts                     W2k3 W2k4
S  CLX_FILESHARE        BCMuListA
S  CLX_FILESHARE        BCMuListB
S  CLX_FILESHARE        BCResyncMuListA
S  CLX_FILESHARE        BCResyncMuListB
S  CLX_FILESHARE        BCEnabledA                     NO
S  CLX_FILESHARE        BCEnabledB                     NO
S  CLX_FILESHARE        BCResyncEnabledA               NO
S  CLX_FILESHARE        BCResyncEnabledB               NO


D:\Program Files\Hewlett-Packard\Cluster Extension XP>_
```

Figure 8 shows an example **CLX_SHARE** resource group resource graph.

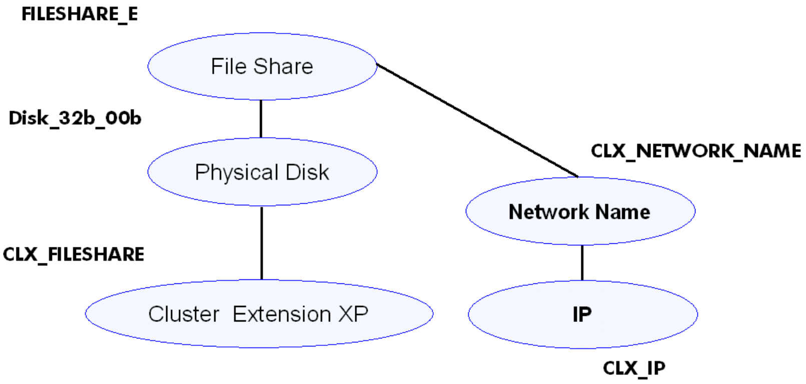


Figure 8. Cluster Extension XP resource tree for resource group CLX_SHARE

Cluster Extension XP is configured as a single resource to enable read/write access to the physical disk resource used for the **CLX_SHARE** cluster group. The physical disk resource depends on the Cluster Extension XP resource and can be brought online only when the Cluster Extension XP resource is already online. Independent on this resource tree the network card will be configured with the **CLX_SHARE** resource group's IP address and network name.

If all those resources have been brought online, the file share can be started.

To configure the Cluster Extension XP resource according to the configuration shown above, the following tasks are executed:

1. Log in to the **w2k3** system with the Administrator account.
2. Create the file share resource group with all above mentioned resources and its dependencies, except the Cluster Extension resource on **w2k3**.
3. Create a new resource of type Cluster Extension XP and add systems **w2k2**, **w2k3** and **w2k4** to its possible owners.
4. Change the restart behavior of the Cluster Extension XP resource so that the resource can be restarted and that the restart affects the group. Set the number of restarts to **0**.
5. Edit the properties in the Parameter tab window (or last window) to configure your Cluster Extension XP resource. Enter the RAID Manager XP instances, the RAID Manager XP device group, the XP serial numbers, **DC_A_Hosts**, and the **DC_B_Hosts**.
6. Click the Pre/Post Exec button and add **clxpre.exe** with its full path. (The clxpre.exe program is an example. It is not included in the Cluster Extension XP product.)
7. Add a dependency on the Cluster Extension XP resource **CLX_FILESHARE** to the physical disk resource **Disk_32b_00b**.
8. Check the cluster service settings, the group, and resource settings.

   ```
   C:\>cluster group CLX_SHARE /prop
   ```

   ```
   C:\>cluster resource CLX_FILESHARE /prop
   ```

9. Set the Cluster Extension XP resource property **RestartAction** to zero (0) or check the “Do not restart” check box in the resource’s Advanced tab window and check if the value has changed.

   ```
   C:\>cluster resource CLX_FILESHARE /prop RestartAction=0
   ```

   ```
   C:\>cluster resource CLX_FILESHARE /prop
   ```

10. Bring the resource group online on **w2k3** by using the Cluster Administrator GUI or command line interface:

    ```
    C:\>cluster group CLX_SHARE /online:w2k3
    ```

    Verify that the Cluster Extension XP resource and all other CLX_SHARE resource group resources are brought online.

    ```
    C:\>cluster group CLX_SHARE
    ```

11. Take the resource group offline and verify that all resources are stopped.

    ```
    C:\>cluster group CLX_SHARE /offline
    ```

    ```
    C:\>cluster group CLX_SHARE
    ```

12. Bring the resource group online again and verify that all resources are available.

    ```
    C:\>cluster group CLX_SHARE /online:w2k3
    ```

    ```
    C:\>cluster group CLX_SHARE
    ```

13. Check the cluster service settings of systems **w2k4**, the group and resource settings.

14. Move the resource group to system **w2k4** and verify that all resources are available.

    ```
    C:\>cluster group CLX_SHARE /moveto:w2k4
    ```

    ```
    C:\>cluster group CLX_SHARE
    ```

15. Check the cluster service settings of systems **w2k2**, the group and resource settings.

16. Move the resource group to system **w2k2** and verify that all resources are available.

    ```
    C:\>cluster group CLX_SHARE /moveto:w2k2
    ```

    ```
    C:\>cluster group CLX_SHARE
    ```

17. Check the cluster service settings of systems **w2k1**, the group and resource settings.

18. Take the resource group offline and verify that all resources are stopped.

```
C:\>cluster group CLX_SHARE /offline
```

```
C:\>cluster group CLX_SHARE
```

19. Change the Cluster Extension XP resource to be able to restart on another system.

```
C:\>cluster resource CLX_FILESHARE /prop RestartAction=2
```

```
C:\>cluster resource CLX_FILESHARE /prop
```

Administration

If the Cluster Extension log files need to be cleared and reset, for example, to reduce disk space usage, you can delete the files. Cluster Extension automatically creates new log files.

Recommendation Archive the log files before deleting them.

The system resources should be monitored on a regular basis as part of Windows 2000/Windows 2003 administration. If any system resource usage by the cluster service is reaching maximum levels, stop and then restart the cluster service. This action automatically fails over the resources and resets system resources.

An alternative method is to manually “move” all resources to another node in the cluster before stopping the cluster service. After all resources are successfully moved to another node, the Cluster service can be stopped and then restarted. Then manually “move” back all resources.

Related information Refer to Microsoft Cluster service documentation for information about how to stop a cluster service.

7

Integration with VCS

HP StorageWorks Cluster Extension XP for VCS provides a resource agent to VERITAS Cluster Server (VCS). This allows cluster administrators to configure the XP disk array-specific failover behavior as easily as any other resource in VCS. Cluster Extension XP objects are configured as attributes of a resource in VCS.

Related information

For information about how to install Cluster Extension XP, see *HP StorageWorks Cluster Extension XP: Installation Guide*.

See the **readme** file on the product CD for supported configurations.

Configuration of the Cluster Extension XP agent

The Cluster Extension XP agent is preconfigured to fit most of your cluster configurations. It comes with a sample configuration that can be modified to fit your VCS and disk array environments. Before configuring the Cluster Extension XP agent, review VCS resource attributes of the Cluster Extension XP resource type.

Disaster-tolerant configuration example using a web server

The example shown in figure 9 (page 151) assumes the following options and actions:

- four systems: **sunrise**, **dawn**, **sunset** and **dusk**
- two disk arrays with serial numbers **35014** and **35013**
- you have configured **web** as device group in the RAID Manager XP **/etc/horcm11.conf** file.
- a pre-executable **web_pre.sh** and a post-executable **web_post.sh** will be invoked by Cluster Extension XP.
- You use the default failover behavior for the service group. The resource **clx_web** is part of the service group **CLX_WEB_SERVER**, and must be brought online before the DiskGroup resources **webdg** and **httpddg**. The RAID Manger XP device group **web** includes all disks for the VxVM disk groups **webdg** and **httpddg** in the example shown below. The following figure illustrates failover options and shows a second service group, **CLX_ORACLE**.

Example of the clx_web resource

```
ClusterExtensionXP clx_web (
                XPSerialNumbers = { 35014, 35013 }
                RaidManagerInstances = { 11 }
                DeviceGroup = web
                PreExecScript = "/etc/opt/hpclx/web_pre.sh"
                PostExecScript = "/etc/opt/hpclx/web_post.sh"
                DC_A_Hosts = { sunrise, dawn }
                DC_B_Hosts = { sunset, dusk }
                )
```

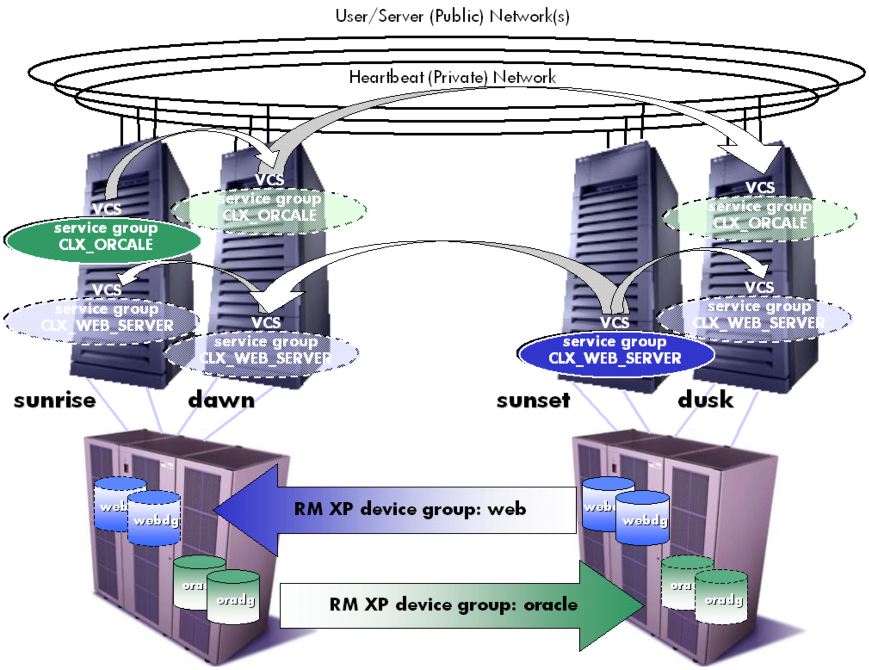


Figure 9. Configuration example

Figure 10 (page 152) shows an example resource graph of the **CLX_WEB_SERVER** service group.

Cluster Extension XP is configured as a single resource to enable read/write access to the disk groups used for the webserver service group. The DiskGroup resources depend on the Cluster Extension XP resource and the Mount resources can be brought online only when the DiskGroup resources and the Cluster Extension XP resource are already online. Independent of this resource tree, the network card will be configured with the webserver service group IP address.

When all these resources have been brought online, the web server can be started.

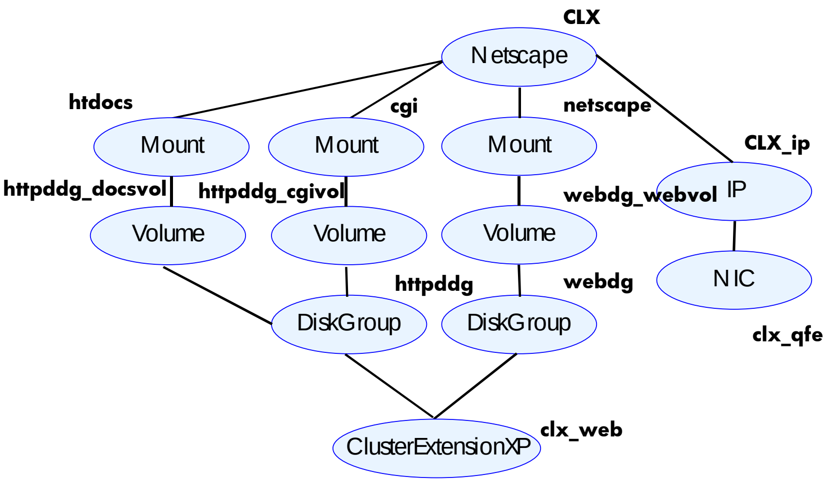


Figure 10. Example resource graph

To configure the Cluster Extension XP agent according to the configuration shown:

1. Log in to system **sunrise** as **root**.
2. Create the Cluster Extension XP resource (for example, **clx_web**) in the **$VCS_CONF/config/main.cf** file, using the example above.
3. Link the new resource as a child resource to all disk resources in the service group.
4. Edit the attributes in the file **$VCS_CONF/config/main.cf** to configure your Cluster Extension XP resource. Enter the RAID Manager XP instances, the RAID Manager XP device group, the XP serial numbers, **DC_A_Hosts**, and the **DC_B_Hosts**.
5. Verify the syntax of the file **$VCS_CONF/config/main.cf**.

   ```
   #hacf –verify $VCS_CONF/config
   ```
6. Start the VCS engine (**had**) on **sunrise**.

   ```
   #hastart
   ```
7. Verify that the Cluster Extension XP and all other webserver service group resources are brought online.

   ```
   #hagrp -display
   ```

8. Take the service group offline and verify that all resources are stopped.

   ```
   #hagrp –offline CLX_WEB_SERVER –sys sunrise
   ```

   ```
   #hagrp –display
   ```

9. Bring the service group online again and verify that all resources are available.

   ```
   #hagrp –online CLX_WEB_SERVER –sys sunrise
   ```

   ```
   #hagrp –display
   ```

10. Start the VCS engine on **dawn**.

    ```
    #hastart
    ```

11. Start the VCS engine on **sunset** and **dusk**, and switch the webserver service group to **dawn** and later to **sunset** and **dusk**. Before you switch the service group to the remote data center, make sure the Continuous Access XP links are configured for bidirectional mirroring and RAID Manager XP instances include the device group, configured for the webserver service group.

    ```
    #hagrp –switch CLX_WEB_SERVER –to system_name
    ```

12. Verify that all Cluster Extension XP and webserver service group resources are brought online.

    ```
    #hagrp -display
    ```

Configuring the Cluster Extension XP resource

For VCS, you can configure a Cluster Extension XP resource either using the VCS command-line interface or the VCS Cluster Manager GUI.

If you use the default values for Cluster Extension XP common objects, no user configuration file is required.

The Cluster Extension XP resource gathers all necessary information about the service group and the XP disk arrays if the Cluster Extension XP resource is brought online.

If configured, a pair/resync monitor is started to monitor the Cluster Extension XP resource. If the resource is not configured to use the pair/resync monitor, a file will be created in the directory specified by the **ApplicationDir** attribute:

*resource_name***.online**

If the resource is taken offline, the file will be removed or the device group associated with the service group will be removed from the pair/resync monitor list. If the device group is the last monitored disk pair, the monitor is stopped also.

The resource type definition file,**ClusterExtensionXPTypes.cf**, must be included in the VCS configuration file **main.cf**. The Cluster Extension XP resource type definition is preconfigured for the most typical cluster configurations. It comes with a sample configuration that can be modified to fit your VCS and disk array environment.

Cluster Extension resource types

Prior to configuring the objects in the user configuration file, review the Cluster Extension XP objects described in “User configuration file and Cluster Extension XP objects” (page 65).

Resource type definition

To configure a Cluster Extension XP resource, use the following object definitions.

```
type ClusterExtensionXP (
        static str ArgList[] = { ApplicationDir, DeviceGroup,
                              ResyncMonitorInterval, ResyncMonitor,
                              ResyncMonitorAutoRecover,
                              RaidManagerInstances,XPSerialNumbers,
                              FenceLevel,
                              DataLoseMirror, DataLoseDataCenter,
                              AsyncTakeoverTimeout,
                              AutoRecover, ApplicationStartup,
                              ResyncWaitTimeout,
                              FastFailbackEnabled, PostExecCheck,
                              PreExecScript, PostExecScript,
                              DC_A_Hosts, DC_B_Hosts,
                              BCMuListA, BCMuListB, BCResyncMuListA,
                              BCResyncMuListB,
                              BCEnabledA, BCEnabledB,
                              BCResyncEnabledA, BCResyncEnabledB }

        NameRule = ClusterExtensionXP_ + group.Name
        str ApplicationDir = "/etc/opt/hpclx/"
        str XPSerialNumbers[]
        str RaidManagerInstances[]
        str DeviceGroup
        str DC_A_Hosts[]
        str DC_B_Hosts[]
        str FenceLevel = never
        str DataLoseMirror = no
        str DataLoseDataCenter = yes
        int AsyncTakeoverTimeout = 1800
        str ApplicationStartup = fastfailback
        int ResyncWaitTimeout = 300
        str FastFailbackEnabled = yes
        str AutoRecover = no
        str ResyncMonitor = no
        str ResyncMonitorAutoRecover = no
        str ResyncMonitorInterval = 60
        str PreExecScript
        str PostExecScript
        str PostExecCheck = no
        str BCMuListA[]
        str BCMuListB[]
        str BCResyncMuListA[]
        str BCResyncMuListB[]
        str BCEnabledA = no
        str BCEnabledB = no
        str BCResyncEnabledA = no
        str BCResyncEnabledB = no
)
```

Adding a Cluster Extension XP resource

These procedures add a resource to an existing service group.

To add a Cluster Extension XP resource from the VCS command line:

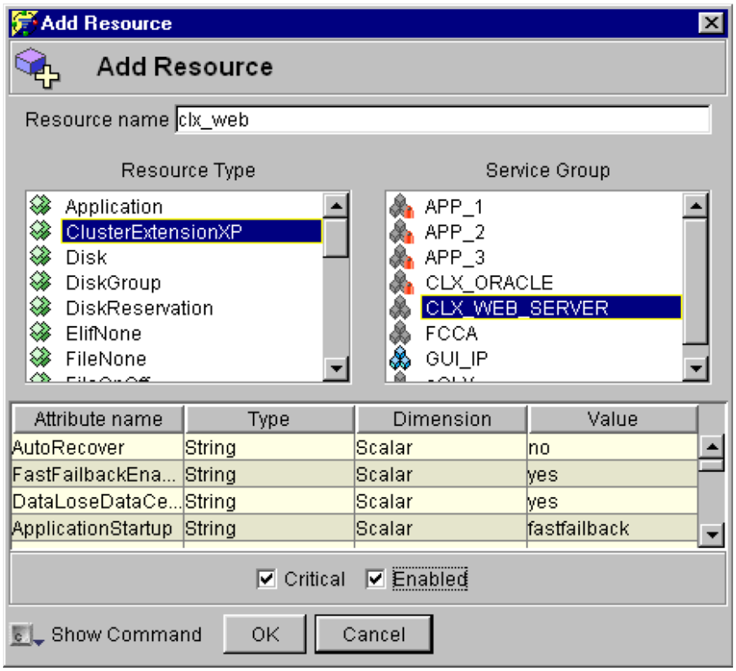
Syntax **hares –add** *resource_name* **ClusterExtensionXP** *service_group*

Example This example adds a Cluster Extension XP resource called **clx_web** to service group **CLX_WEB_SERVER**.

```
# hares -add clx_web ClusterExtensionXP CLX_WEB_SERVER
```

To add a Cluster Extension XP resource from the VCS Cluster Manager GUI:

1. Use the Cluster Explorer.
2. Click on the Add Resource icon in the Cluster Explorer toolbar.
3. Enter the resource name in the Resource name box.
4. Select **ClusterExtensionXP** from the list of Resource Types.
5. Select the service group you want to add the new Cluster Extension XP resource to from the Service Group box.
6. Modify the attribute values of the new Cluster Extension XP resource.
7. Click the check boxes for Critical and Enabled.
8. Choose OK.



Resource name: clx_web

Resource Type: ClusterExtensionXP

Service Group: CLX_WEB_SERVER

Attribute name	Type	Dimension	Value
AutoRecover	String	Scalar	no
FastFailbackEna...	String	Scalar	yes
DataLoseDataCe...	String	Scalar	yes
ApplicationStartup	String	Scalar	fastfailback

☑ Critical ☑ Enabled

Changing Cluster Extension XP attributes

Cluster Extension XP resource attributes can be changed after the configuration has been write enabled.

To change attribute values of the Cluster Extension XP resource the resource must be taken offline.

To change an attribute value from the VCS command line:

Syntax **hares –modify** *ClusterExtensionXP_resource* [**–add | –update**] *attribute value*

Example This example changes a Cluster Extension XP resource called **clx_web** to change the default **FenceLevel** attribute.

```
# hares -modify clx_web FenceLevel data
```

These commands change the RAID Manager XP instance used for the Cluster Extension XP resource **clx_web** and then adds an additional instance to provide redundancy.

```
# hares -display clx_web -attribute RaidManagerInstances

# hares -modify clx_web RaidManagerInstances -update 90

# hares -modify clx_web RaidManagerInstances -add 22
```

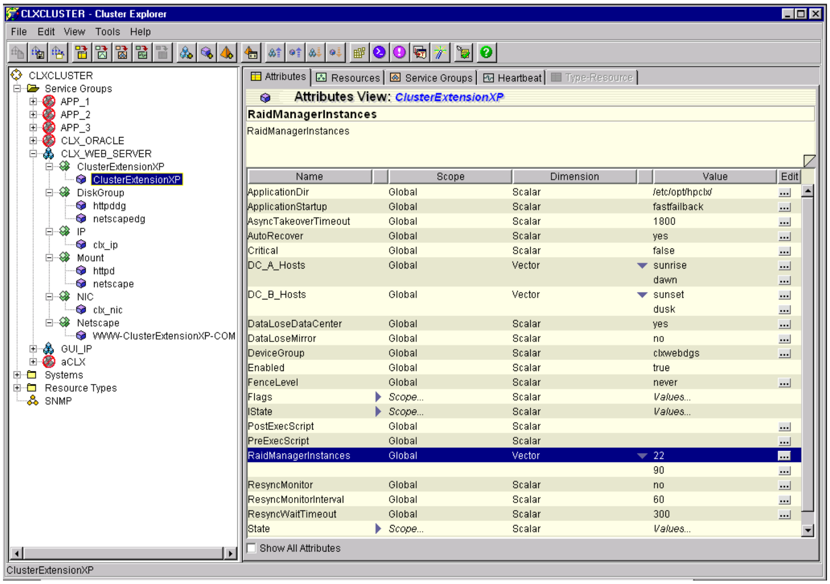
This example displays all attributes of the Cluster Extension XP resource **clx_web**.

```
# hares -display clx_web
```

To change attribute values from the VCS Cluster Manager GUI:

1. Use the Cluster Explorer.
2. Click the resource name.
3. Click the Attributes View Tab on the View Panel.
4. Click the Edit Iconin the Edit Column of the Attribute.

5. Enter changes to the attribute value. For nonscalar attributes, use + and x buttons to add or remove elements. Do not change the attributes scope to local. All Cluster Extension XP attributes are global in scope.



Name	Scope	Dimension	Value	Edit
ApplicationDir	Global	Scalar	/etc/opt/hpclx/	...
ApplicationStartup	Global	Scalar	fastfailback	...
AsyncTakeoverTimeout	Global	Scalar	1800	...
AutoRecover	Global	Scalar	yes	...
Critical	Global	Scalar	false	...
DC_A_Hosts	Global	Vector	sunrise	...
			dawn	...
DC_B_Hosts	Global	Vector	sunset	...
			dusk	...
DataLoseDataCenter	Global	Scalar	yes	...
DataLoseMirror	Global	Scalar	no	...
DeviceGroup	Global	Scalar	clxwebdgs	...
Enabled	Global	Scalar	true	
FenceLevel	Global	Scalar	never	...
Flags	*Scope...*	Scalar	*Values...*	
IState	*Scope...*	Scalar	*Values...*	
PostExecScript	Global	Scalar		...
PreExecScript	Global	Scalar		...
RaidManagerInstances	Global	Vector	22	...
			90	...
ResyncMonitor	Global	Scalar	no	...
ResyncMonitorInterval	Global	Scalar	60	...
ResyncWaitTimeout	Global	Scalar	300	...
State	*Scope...*	Scalar	*Values...*	

6. Choose OK.

Linking a Cluster Extension XP resource

Cluster Extension XP must be the first resource in the resource chain of a VCS service group.

All resources depending on the disk resource (for example, **Mount**) including the disk resources (**DiskGroup**, **Disk**, **DiskReservation**) must be parent resources to the Cluster Extension XP resource.

Caution *Cluster Extension XP does not support ServiceGroupHB resources in Continuous Access XP configurations because of the read/write mode differences between the primary and secondary disk in an XP disk array.*

To link other resources to the Cluster Extension XP resource:

Syntax **hares –link** *disk_group_resource ClusterExtensionXP_resource*

Example

```
# hares -link netscapedg clx_web_server
```

To link other resources from the VCS Cluster Manager GUI:

1. Use the Cluster Explorer
2. Click the Resources View Tab on the View Panel.
3. Click the resource icon of the resource that is to be the parent resource.
4. Move the yellow line to the resource that is to be the (child) Cluster Extension XP resource and click.
5. Click YES in the dialog box to confirm the dependency.

Bringing a Cluster Extension XP resource online

Resources are usually brought online automatically when the service group is brought online.

To bring a resource group manually online, the service group must be enabled on the system and the service group must be auto-enabled in the cluster. Finally, the resource must be enabled.

To enable and bring a Cluster Extension XP resource online from the command line:

Syntax **hares –modify** *ClusterExtensionXP_resource* **Enabled 1**

hares –online *ClusterExtensionXP_resource* **–sys** *system_name*

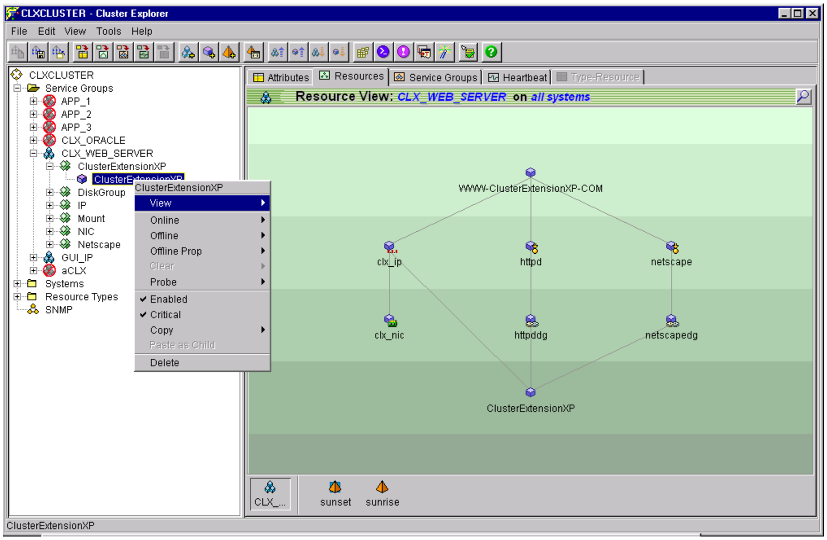
Example This example enables and brings the Cluster Extension XP resource **clx_web**.

```
# hares -modify clx_web Enabled 1

# hares -online clx_web -sys sunrise
```

To enable and bring a Cluster Extension XP resource online from the VCS Cluster Manager GUI:

1. Use the Cluster Explorer.
2. Right-click the resource name.
3. Select online and then the system, where you want to bring the resource online.
4. Click YES in the dialog box to confirm.



Taking a Cluster Extension XP resource offline

Resources are usually taken offline automatically when the service group is taken offline.

There are two ways to manually bring a resource group offline:

- Take only the specified resource offline.
- Propagate the offline request to all parent resources, which takes all parent resources offline before the specified resource.

To take a Cluster Extension XP resource offline or to propagate the offline request to all parent resources from the command line:

Syntax **hares –offline** *ClusterExtensionXP_resource* **–sys** *system_name*

hares –offprop *ClusterExtensionXP_resource* **–sys** *system_name*

Example This example takes the Cluster Extension XP resource **clx_web** offline or propagates the offline request to all its parent resources before taking it offline.

```
# hares -offline clx_web -sys sunrise
```

```
# hares -offprop clx_web -sys sunrise
```

To take a Cluster Extension XP resource offline from the VCS Cluster Manager GUI:

1. Use the Cluster Explorer.
2. Right-click the resource name.
3. Select Offline or Offline Prop and then the system where you want to bring the resource offline.
4. Click YES in the dialog box to confirm.

Deleting a Cluster Extension XP resource

These procedures remove a Cluster Extension XP resource from an existing service group.

To delete a resource from the VCS command line:

Syntax **hares –delete** *ClusterExtensionXP_resource*

Caution *Deleting a running Cluster Extension XP resource does not remove the resource_name**.online** file and does not remove the device group from the list of monitored device groups if the pair/resync monitor is used to monitor the Continuous Access XP link. Therefore, the device group must be deleted from the list of monitored device groups manually using the **clxchkmon** command after deleting the Cluster Extension XP resource. See "Stopping the pair/resync monitor"* (page 198).

Caution *Failure to delete the monitored device group from the list of monitored device groups can cause data corruption if the **ResyncMonitorAutoRecover** attribute is set to **YES**.*

To add a Cluster Extension XP resource to an existing service group from the VCS Cluster Manager GUI:

1. Use the Cluster Explorer.
2. Right-click the resource name.
3. Select Delete.
4. Click YES in the dialog box to confirm.

Pair/resync monitor integration

The pair/resync monitor is used to detect and react to suspended Continuous Access XP links. It is activated if the **ResyncMonitor** attribute value is set to **YES**. The automatic disk pair resynchronization feature is also activated if the **ResyncMonitorAutoRecover** attribute value is **YES**.

When the resource is taken offline, the monitor will be stopped for the RAID Manager XP device group used for this resource.

The pair/resync monitor will not be started when the **ResyncMonitor** attribute will be changed to **YES** while the resource is online. However, a running **ResyncMonitor** will be disabled for the resource if the **ResyncMonitor** attribute is changed to **NO** while the resource is online.

Caution *If the resource group cannot be taken offline gracefully, the cluster administrator must disable monitoring of the device group for this resource. To avoid data corruption, this task must be part of the recovery procedure when Cluster Extension XP is deployed in the VCS environment. See "Stopping the pair/resync monitor"* (page 198).

VCS automatically attempts to stop the pair/resync monitor for the resource if it is running on more than one system.

Caution *Ensure that the pair/resync monitor does not monitor and resynchronize the disk pair (device group) from both disk array sites.*

Timing considerations for VCS

Cluster Extension XP gives priority to XP disk array operations over cluster software operations; if Cluster Extension XP invokes disk pair resynchronization operations or gathers information about the remote XP disk array, Cluster Extension XP waits until the requested status information is reported. This feature prioritizes data integrity over the cluster software's failover behavior.

However, in some cases, this behavior could lead to failed Cluster Extension XP resources:

- Cluster Extension XP uses RAID Manager XP instances to communicate with the remote XP disk array. Depending on the settings of the RAID Manager XP instance timeout parameter and the number of remote instances the online operation could time out. This can happen if the local RAID Manager XP instance cannot reach the remote RAID Manager XP instance.
- Cluster Extension XP tries to resynchronize disk pairs and waits until the RAID Manager XP device group is in PAIR state if the ApplicationStartup attribute is set to **RESYNCWAIT**. Depending on the RAID Manager XP version and the XP firmware version this could be a full resynchronization and may take longer than the online timeout interval. Even if the RAID Manager XP version and the XP firmware version allow a delta resynchronization, the delta between the primary and the secondary could be big enough for the copy process to exceed the online timeout value.
- The **ResyncWaitTimeout** attribute can automatically lead into failed Cluster Extension XP resources when set higher than the online timeout interval.
- If running in fence level **ASYNC**, the default value of the **AsyncTakeoverTimeout** can cause the resource to fail because its value is set beyond the resource online timeout interval. This is done because the takeover process for fence level **ASYNC** can take much longer when slow communications links are in place.

  To prevent takeover commands from being terminated by the takeover timeout before finishing, the time to copy the installed XP disk array cache should be measured and the resource online timeout interval should be adjusted according to the measured copy time. When measuring the copy time, measure only the slowest link used for Continuous Access XP Extension. This ensures that the XP disk array cache can be transferred from the remote XP disk array, even in the event of a single surviving replication link between the XP disk arrays.

Because the failover environment is dispersed into two (or more) data centers, the failover time cannot be expected to be the same as it would be in a single data center with a single shared disk device. Therefore, the online timeout values, the monitor interval of the Cluster Extension XP resource, and the service group using the Cluster Extension XP resource should be adjusted based on failover tests performed to verify the proper configuration setup.

Enable/disable service groups

Based on the XP disk array status information, Cluster Extension XP can change the cluster software behavior to automatically failover (or failback) the service group faster to the remote data center.

For example, if the remote disk state is SVOL_SSUS and the SSWS flag has been set to indicate a prior takeover to the secondary disk set. If you have set the **ApplicationStartup** object to **FASTFAILBACK**, Cluster Extension XP would disable the service group for all systems in the respective data center and VCS would transfer the service group back to the remote site rather than waiting for a pair resynchronization to be finished before the service group could start on the local site.

This could happen only if you have not recovered the suspended disk pair after a prior takeover, where the PAIR state could not be maintained because of, for example, a Continuous Access XP link failure.

This feature reduces application downtime because the service group (and the application) will not be brought online on each system in the service group's system list. It will be moved to the first available system listed in the service group's system list, which is connected to the remote XP disk array.

This is done by enabling the VCS configuration file (**main.cf**) to be writable. The service group will be disabled for all systems contained in either the **DC_A_Hosts** object or **DC_B_Hosts** object. Then, the VCS configuration file will be saved (dumped).

This feature can be disabled. If the **FastFailbackEnabled** object is set to **NO**, the standard VCS process is used and the Cluster Extension XP resource fails on the local system (and so would the service group). VCS then tries to bring the service group online on the next system in the service group's system list (which should be a local system). This will fail because the state of the local XP disk array has not changed. The service group will fail until the service group is brought online on a system connected to the remote XP disk array. The service group online process will take longer and it will not access the VCS configuration file.

Restrictions for VCS with Cluster Extension XP

The following is a summary of restrictions which apply for VERITAS Cluster Server configurations when Cluster Extension XP is used to enable failover between two XP disk arrays.

- The Cluster Extension XP resource must be the first (child) resource for all other disk resources.
- Heartbeat disks cannot be used because of the P/SVOL read/write behavior of Continuous Access XP.
- No service group heartbeat disks are allowed in the service group. The ServiceGroupHB resource is not supported in Cluster Extension XP configurations because of the P/SVOL read/write behavior of Continuous Access XP.
- Only one Cluster Extension XP resource is allowed to be configured per service group.
- Cluster Extension XP must not be used with Parallel service groups. If Cluster Extension XP is used in a Parallel service group, all systems configured for this service group must be connected to the same XP disk arrays. A failover operation to the secondary XP disk array must be done manually only. In such case, all active service groups must be brought offline before any of those service groups can be brought online on the secondary XP disk array.
- The **ApplicationDir** attribute value must not be changed when the resource is online. The **ApplicationDir** attribute defines the location of the *application_dir*/*resource_name***.online** file. This file is created when the resource is brought online (if **ResyncMonitor** attribute is set to **NO**). The Cluster Extension XP resource monitors the file located in the location specified by **ApplicationDir**. Changing this attribute can cause the Cluster Extension XP resource to fail.
- The resource online timeout must be greater than the value specified for the **ResyncWaitTimeout** attribute.
- The resource online timeout should be greater than twice the wait time of all remote RAID Manager XP instances times the number of remote systems. Otherwise, the Cluster Extension XP resource will fail to go online in case of a complete remote data center failure. If a

post-executable is specified, the resource online timeout should be greater than the number of remote systems times three times $t_{WT.}$

t_{online} = resource online timeout

$n_{remote\ sytems}$ = number of remote systems configured to run RAID Manager XP instances

t_{WT} = wait time until remote error will be reported by local RAID Manager XP instance

$t_{online} > n_{remote\ sytems}$ x 2 x t_{WT}

Unexpected offline conditions

In rare cases, ClusterExtensionXP resources go offline after the following conditions occur concurrently:

- ClusterExtensionXP resources have been online.
- The cluster has been stopped forcibly (without taking the resources offline).
- The CA XP links have failed.
- A remote RAID Manager XP instance is not available due to a full network outage.

ClusterExtensionXP resources go offline because the primary volume state changes from PVOL_PAIR to PVOL_PSUE and the secondary volume state changes from SVOL_PAIR to EX_NORMT.

The state combination PVOL_PSUE and EX_NORMT is not designed to be handled automatically because the remote side (remote RAID Manager/ disk array), which has no state information available, could have more current data then the primary (PVOL_PSUE) site. In this particular case, the user is required to investigate data currency and to determine the appropriate action to be taken.

To bring the ClusterExtensionXP resources online:

1. Recover the CA XP link error and the network error, and restart RAID Manager on the remote site.
2. Manually resynchronize the affected disk pairs.
3. Bring the ClusterExtensionXP resources online.

or

1. Create the **forceflag** *resource_name***.forceflag** in the **ApplicationDir** path.

   Default: **/etc/opt/hpclx/**
2. Bring the ClusterExtensionXP resources online.
3. Depending on the attributes set for the resources, you might need to manually resynchronize the CA XP disk pairs.

8

Integration with Serviceguard for Linux

Cluster Extension XP is integrated with the Serviceguard cluster software on Linux using the customization scheme provided by Serviceguard. This allows cluster administrators to configure the disk array-specific failover behavior in an environment file (user configuration file) and enables Cluster Extension XP to be used by simply setting the Data Replication parameter in the standard package control file.

Related information

For information about how to install Cluster Extension XP, see *HP StorageWorks Cluster Extension XP: Installation Guide*.

See the **readme** file on the product CD for supported configurations.

Configuration of the Cluster Extension XP environment

The Cluster Extension XP objects must be configured using a user configuration file. This file must be located in the same directory as the package control file (also known as package directory). The Cluster Extension XP software gathers all necessary information about the disk arrays during the package startup process.

If configured, a pair/resync monitor is started to continuously monitor the Continuous Access XP disk pairs/mirroring link used by the package. To use this monitor, Cluster Extension XP must be configured as a Serviceguard service.

The Cluster Extension XP binary **clxmcsg** is called prior to the volume group activation in order to check the status of the RAID Manager XP device group. If necessary, appropriate actions are taken to allow access to these disks before Serviceguard accesses the disks of the particular package.

Disaster-tolerant configuration example using a web server

Prior to configuring the objects in the user configuration file, review the Cluster Extension XP objects described in "User configuration file and Cluster Extension XP objects" (page 65).

The Cluster Extension XP integration with Serviceguard allows only one **APPLICATION** tag specified per user configuration file. This means you must specify one user configuration file per Serviceguard package. The **APPLICATION** object in the user configuration file of Cluster Extension XP does not need to have a value specified. The **DeviceGroup** object within the **APPLICATION** section determines the RAID Manager XP device group needed to control all the shared disks of the Serviceguard package.

The example shown below assumes the following options and actions:

- Four systems for a cluster: **clxrh1**, **clxrh2**, **clxrh3** and **clxrh4**

  (**quorumserver** provides quorum but is not part of the actual cluster.)
- Two disk arrays with serial numbers **30047** and **30053**
- You have configured **clxwebvgs** as device group in the RAID Manager XP **/etc/horcm101.conf** file.
- A pre-executable **clxweb_pre_takeover.sh** and a post-executable **clxweb_post_takeover.sh** will be invoked by Cluster Extension XP. The executable files can be any script or program of your choice. Cluster Extension XP provides sample scripts in the **/etc/opt/hpclx/sample** directory.
- You use the default failover behavior for the package. The user configuration file **CLXWEB_clx.env** is a part of the package **CLXWEB** and must be located in the same directory as the package control file. The package using the package control script **CLXWEB.sh** checks the disk pair states before the volume groups **vgweb** and **vghtdocs** are activated and the webserver is started.The RAID Manger XP device group **clxwebvgs** includes all disks for the LVM volume groups **vgweb** and **vghtdocs** in the example shown below. The example illustrates failover options and shows several other packages and their RAID Manager XP device groups.

Example of the CLXWEB_clx.env user configuration file

```
COMMON
LogLevel              info         # values: error|info (optional)

APPLICATION           CLXWEB       # == PKGNAME of the package
XPSerialNumbers       30047 30053
RaidManagerInstances  101
DeviceGroup           clxwebvgs    # raid manager device group
DC_A_Hosts            clxrh1 clxrh2# systems in data center A
DC_B_Hosts            clxrh3 clxrh4# systems in data center B

#optional parameter   (only necessary if other than default)
FenceLevel            data         # values: data | never | async
ApplicationStartup    resyncwait    # values: fastfailback | resyncwait
AutoRecover           yes          # possible values: yes | no
DataLoseMirror        no           # possible values: yes | no
DataLoseDataCenter    no           # possible values: yes | no
PreExecScript         /etc/opt/hpclx/clxweb_pre_takeover.sh
PostExecScript        /etc/opt/hpclx/clxweb_post_takeover.sh
```

- The package is configured to wait for a pair resynchronization in case you have not done any disk pair recovery after the package has been failed over to an adoptive node **(ApplicationStartup** object is set to **RESYNCWAIT**).

- The **AutoRecover** object is set to **YES**, which means that you will utilize Cluster Extension XP capabilities to automatically recover suspended disk pair states.
- The **DataLoseMirror** object and **DataLoseDataCenter** object are set to **NO**, which means Cluster Extension XP will not allow you to start the package automatically if the disk pair is suspended or a takeover operation leads into a suspended disk pair.

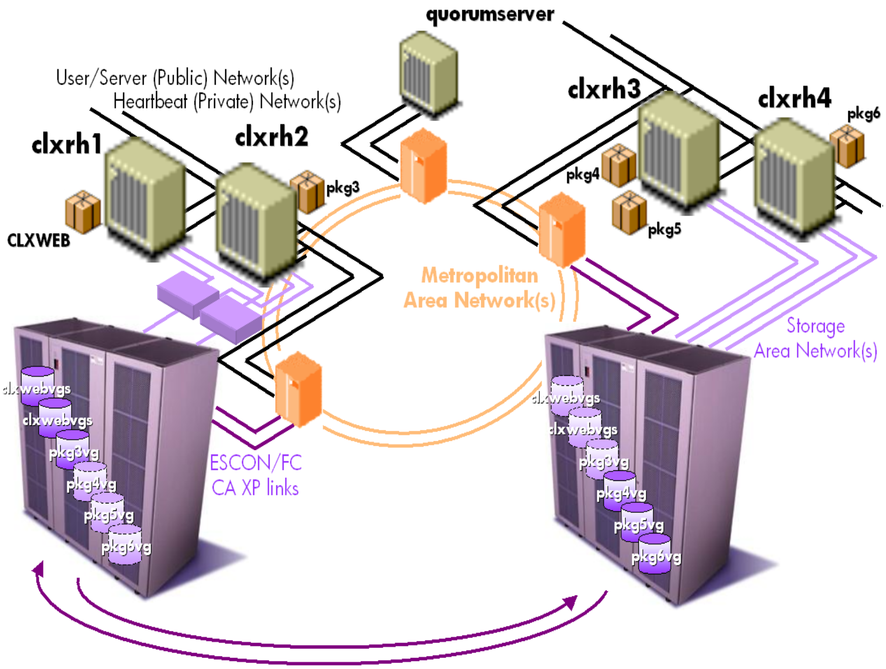


Figure 11. Configuration example

Cluster Extension XP is configured to enable read/write access to the disk groups used for the webserver's package. The activation of the volume groups depends on a successful return code from Cluster Extension XP. The mount points can be mounted only when the volume groups are active and Cluster Extension XP was able to allow read-write access to the disk set. After the file system for the webserver's executables and content data have been mounted and checked, the network card will be configured with the webserver's IP address.

When all these resources are up, the web server can be started as configured in the **customer_defined_run_cmd** section of the package control file. (The integration of application services can be purchased and delivered through HP Consulting.)

To configure the Cluster Extension XP environment according to the configuration shown:

The following procedure assumes that you have created the Serviceguard cluster configuration including the nodes mentioned earlier and that this configuration has been successfully applied to all nodes and the cluster is running.

1. Log in to system **clxrh1** as **root**.
2. Source the Serviceguard environment variables before you use them.

Example

```
#. /etc/cmcluster.conf
```

3. Create a package directory with the name of the package in the **$SGCONF** directory.

Example

```
#mkdir $SGCONF/CLXWEB
```

4. Create the package configuration file and edit its content.

Example

```
#cmmakepkg –p $SGCONF/CLXWEB/CLXWEB.config
```

5. Create the package control file and edit its content.

Example

```
#cmmakepkg –s $SGCONF/CLXWEB/CLXWEB.sh
```

   For details on how to create a Serviceguard package configuration and control file, see *Managing Serviceguard for Linux*.

6. Create the Cluster Extension XP user configuration file **CLXWEB_clx.env** in the **$SGCONF/CLXWEB/** package directory, using the example above or copy and edit the sample file **UCF.cfg** provided in the **/etc/opt/hpclx/sample** directory.

Example

```
#cp /etc/opt/hpclx/sample/UCF.cfg $SGCONF/CLXWEB/CLXWEB_clx.env
```

7. Uncomment the parameter **DATA_REP** in the package control file **$SGCONF/CLXWEB/ CLXWEB.sh** and change its value to **clx**.
8. Copy the content of the package directory to all other cluster members.

Example

```
#for i in clxrh2 clxrh3 clxrh4; do
rcp –rp $SGCONF/CLXWEB $i:$SGCONF/
done
```

9. Verify and apply the new package information to the cluster database.

Example

```
#cmapplyconf –v -P $SGCONF/CLXWEB/CLXWEB.config
```

10. Start the package on **clxrh1**.

Example

```
#cmrunpkg CLXWEB
```

11. Verify that the package has been started successfully.

Example

```
#cmviewcl –v –p CLXWEB
```

12. Halt the package and verify that all services for the package are halted.

Example

```
#cmhaltpkg CLXWEB
#cmviewcl –v –p CLXWEB
#clxchkmon –show
```

13. After the mount points for the shared directories have been created on all configured nodes you can start the package on **clxrh2**. The volume groups will be automatically imported during the first package start on each system.

Example

```
#cmrunpkg –n clxrh2 CLXWEB
```

14. Before you transfer (halt and run) the package to the remote data center, make sure that the Continuous Access XP links are configured for bidirectional mirroring and RAID Manager XP instances include the device group, configured for the webserver's package. Start the package on the remaining systems **clxrh3** and **clxrh4**. Verify that the disks used by your **CLXWEB** package are in PAIR state.

Example

```
#export HORCMINST=101
#pairdisplay –fcx –g clxwebvgs
```

Halt the package and verify that all services for the package are halted.

Example

```
#cmhaltpkg CLXWEB
#cmviewcl –v –p CLXWEB
```

Start the package on **clxrh3**. Verify that the package has been started successfully.

Example

```
#cmrunpkg –n clxrh3 CLXWEB
#cmviewcl –v –p CLXWEB
```

Verify that the disk pairs have swapped their personalities.

Example

```
#pairdisplay –fcx –g clxwebvgs
```

Halt the package and verify that all services for the package are halted.

Example

```
#cmhaltpkg CLXWEB
#cmviewcl –v –p CLXWEB
```

Start the package on **clxrh4**. Verify that the package has been started successfully.

Example

```
#cmrunpkg –n clxrh4 CLXWEB
#cmviewcl –v –p CLXWEB
```

Verify that the disk pair status has not changed and halt the package on **clxrh4**.

Example

```
#pairdisplay –fcx –g clxwebvgs
#cmhaltpkg CLXWEB
```

15. After you verified that the webserver package including Cluster Extension XP can be run on each system in the cluster you can move the package back to its primary system.

Example

```
#cmrunpkg –n clxrh1 CLXWEB
#cmviewcl –v –p CLXWEB
#pairdisplay –fcx –g clxwebvgs
```

Adding a Cluster Extension XP integration to an existing Serviceguard package

These procedures add a Cluster Extension XP integration to an existing package.

To add a Cluster Extension XP integration:

1. Copy the Cluster Extension XP user configuration file **UCF.cfg** provided in **/etc/opt/hpclx/sample** into the **$SGCONF/***my_package_name/* package directory and name it *my_package_name***_clx.env**, where *my_package_name* is the name of your Serviceguard package.
2. Change the parameter in the user configuration file based on your configuration.
3. Uncomment the parameter **DATA_REP** in the package control file **$SGCONF/***my_package_name/my_package_name***.sh** and change its value to **clx**.
4. Copy the content of the package directory to all other cluster members.

Example

```
#for i in system2 system3 system4; do
rcp –rp $SGCONF/my_package_name $i:$SGCONF/
done
#for i in system2 system3 system4; do
rcp –p $SGCONF/my_package_name/my_package_name.sh $i:$SGCONF/my_package_name/
rcp –p $SGCONF/my_package_name/my_package_name.config $i:$SGCONF/my_package_name/
rcp –p $SGCONF/my_package_name/my_package_name_clx.env $i:$SGCONF/my_package_name/
done
```

5. Verify the new package information to the cluster database.

   ```
   #cmapplyconf -v -C $SGCONF/your_cluster_config_filename.config \
   –P $SGCONF/my_package_name/my_package_name.config
   ```

Starting a Serviceguard package with Cluster Extension XP

Serviceguard packages usually start automatically when the system reboots or can be started with the command **cmrunpkg** *my_package_name*.

To start a package manually, the package must be enabled to run on the system.

1. Make sure that the mount points for your file systems exist on each system that is configured to run the package.
2. Start and stop the package on each system once.

To enable and start a package including Cluster Extension XP from the command line:

Syntax **cmmodpkg –e –n** *system_name my_ package_name*

cmrunpkg –v –n *system_name my_ package_name*

Example This example enables and starts the Serviceguard package **CLXWEB**.

```
#cmmodpkg –e -n clxrh1 CLXWEB
```

```
#cmrunpkg –v –n clxrh1 CLXWEB
```

To enable and start a Cluster Extension XP package from the Serviceguard Manager:

1. Use the Cluster Explorer.
2. Right-click the package name.
3. From the menu, select **run package on node**, and then select the system where you want to start the package.
4. Click YES in the dialog box to confirm.

Halting a Serviceguard package with Cluster Extension XP

Serviceguard packages can be stopped with the command **cmhaltpkg** *my_package_name*.

In case of a maintenance stop of your Serviceguard package, you must disable package switching.

To disable and halt a package including Cluster Extension XP from the command line:

Syntax **cmmodpkg –d –n** *system_name my_ package_name*

cmhaltpkg –v –n *system_name my_ package_name*

Example This example enables and starts the Serviceguard package **CLXWEB**.

```
#cmmodpkg –d -n clxrh1 CLXWEB
```

```
#cmhaltpkg –v –n clxrh1 CLXWEB
```

To disable and halt a Cluster Extension XP package from the Serviceguard Manager:

1. Use the Cluster Explorer.
2. Right-click the package name.
3. From the menu, select **halt package on node**, and then select the system where you want to halt the package.
4. Select disable package switching and then the system, for which you Click YES in the dialog box to confirm.

Deleting Cluster Extension XP from a Serviceguard package

These procedures remove Cluster Extension XP from an existing package.

Caution *Deleting Cluster Extension XP from a running package does not stop the pair/resync monitor and does not remove the device group from the list of monitored device groups if the pair/resync monitor is used to monitor the Continuous Access XP link.The device group must be deleted from the list of monitored device groups manually using the* ***clxchkmon*** *command after deleting the Cluster Extension XP from the Serviceguard package. Therefore, removal of Cluster Extension XP from a running Serviceguard package is not supported. See "Stopping the pair/resync monitor"* (page 198).

To remove a device group from the pair/resync monitor manually you can use the following procedure:

Check whether the pair/resync monitor monitors the device group of your package by using the following command on all nodes in the cluster:

```
# clxchkmon –show
```

If the package name and device group are in the list, remove them from that node:

```
#clxchkmon –remove –n package_name \
–g device_group_name
```

Caution *Failure to delete the monitored device group from the list of monitored device groups can cause data corruption if the* ***ResyncMonitorAutoRecover*** *object is set to* ***YES****.*

To delete Cluster Extension XP:

1. Halt the *Serviceguard* package and disable package switching.

   #cmmodpkg –d *my_package_name*

   #cmhaltpkg *my_package_name*

2. Remove the Cluster Extension XP user configuration file *my_package_name***_clx.env** in the **$SGCONF/***my_package_name/* package directory, where *my_package_name* is the name of your Serviceguard package.

   Remove the file from all systems where the package was configured to run.

3. Change the parameter **DATA_REP** in the package control file **$SGCONF/***my_package_name***.sh** from **clx** to **none**.

4. Remove the lines starting with the parameter **NODE_NAME** in the package configuration file **$SGCONF/***my_package_name***.config**, which specifies the remote data center systems.

5. *(Optional)* If you added a Serviceguard service, which monitors the pair/resync monitor entry for the package you need to remove those service entries. If you specified several services you might need to renumber the service entries as well.

   Copy the content of the package directory to all other cluster members.

Example

```
#for i in system2 system3 system4; do
rcp -p $SGCONF/my_package_name/my_package_name.sh $i:$SGCONF/my_package_name/
rcp -p $SGCONF/my_package_name/my_package_name.config $i:$SGCONF/my_package_name/
done
```

Verify and apply the new package information to the cluster database.

#cmapplyconf –v -P $SGCONF/*my_package_name/my_package_name***.config**

For more details, see “Pair/resync monitor integration” (page 185).

Pair/resync monitor integration

The pair/resync monitor is used to detect and react on suspended Continuous Access XP links. It is activated if the **ResyncMonitor** object value is set to **YES**. The automatic disk pair resynchronization feature is also activated if the **ResyncMonitorAutoRecover** object value is **YES**.

When the Serviceguard package has been halted, the monitor will be stopped for the RAID Manager XP device group used for this package.

The pair/resync monitor does not allow online changes in the Serviceguard implementation of Cluster Extension XP.

The pair/resync monitor will not be started when the **ResyncMonitor** object is changed to **YES** while the package is running. A running pair/resync monitor will not be disabled if the **ResyncMonitor** object is changed to **NO** while the package is running.

Caution

If the Serviceguard package cannot be halted gracefully, the cluster administrator must disable monitoring of the device group for this package. To avoid data corruptions, this task must be part of the recovery procedure when Cluster Extension XP is deployed in the Serviceguard environment. See "Stopping the pair/resync monitor" (page 198).

Ensure that the pair/resync monitor does not monitor and resynchronize the disk pair (device group) from both disk array sites.

The pair/resync monitor will work correctly (started and stopped) with Serviceguard only if integrated as a service. The service will check that the package name and the device group are present in the pair/resync monitor table. The service will not restart or re-enter the package name/device group entry automatically if it has been removed from the pair/resync monitor.

The following procedure adds the pair/resync monitor feature to Cluster Extension XP for Serviceguard:

1. Halt the Serviceguard package and disable package switching

   #**cmmodpkg –d** *my_package_name*

   #**cmhaltpkg** *my_package_name*

2. Add or enable the parameter **ResyncMonitor** in the Cluster Extension XP user configuration file *my_package_name***_clx.env** in the **$SGCONF/***my_package_name*/ package directory, where *my_package_name* is the name of the Serviceguard package. Set this parameter to **YES**.

3. Uncomment and edit the **SERVICE** parameters in the package configuration file **$SGCONF/***my_package_name***.config** to use the service.

   ```
   SERVICE_NAME                   CLX_PAIR_MON_my_package_name
   SERVICE_FAIL_FAST_ENABLED      NO
   SERVICE_HALT_TIMEOUT           30
   ```

   For further information, see *Managing Serviceguard for Linux*.

4. Uncomment and edit the **SERVICE** parameters in the package control file **$SGCONF/***my_package_name***.sh**. If you want the package to fail if the pair/resync monitor fails or if somebody removes the package name/device group from the monitor, set the restart value to be empty or use **–r** *number_of_restarts_before_fail*.

   ```
   SERVICE_NAME[0]=”CLX_PAIR_MON_my_package_name”
   SERVICE_CMD[0]=”/opt/hpclx/bin/clxmsg service ${0} ${PACKAGE}”
   SERVICE_RESTART[0]=”-R”
   ```

5. Copy the content of the package directory to all other cluster members.

```
#for i in system2 system3 system4; do
rcp –rp $SGCONF/my_package_name $i:$SGCONF/
done
#for i in system2 system3 system4; do
rcp –p $SGCONF/my_package_name/my_package_name.sh
$i:$SGCONF/my_package_name/
rcp –p $SGCONF/my_package_name/my_package_name.config
$i:$SGCONF/my_package_name/
rcp –p $SGCONF/my_package_name/my_package_name_clx.env
$i:$SGCONF/my_package_name/
done
```

6. Verify and apply the new package information to the cluster database.

```
#cmapplyconf -v -C $SGCONF/your_cluster_config_filename.config \
-P $SGCONF/my_package_name/my_package_name.config
```

Timing considerations for Serviceguard

Cluster Extension XP gives priority to XP disk array operations over cluster software operations. If Cluster Extension XP invokes disk pair resynchronization operations or gathers information about the remote XP disk array, Cluster Extension XP waits until the requested status information is reported. This feature prioritizes data integrity over the cluster software's failover behavior.

However, in some cases, this behavior could lead to failed Serviceguard packages:

- Cluster Extension XP uses RAID Manager XP instances to communicate with the remote XP disk array. Depending on the settings of the RAID Manager XP instance timeout parameter and the number of remote instances the package start operation could time out. This can happen if the local RAID Manager XP instance cannot reach the remote RAID Manager XP instance. The package startup can be set to NO_TIMEOUT in the package configuration file if preferred.
- Cluster Extension XP tries to resynchronize disk pairs and waits until the RAID Manager XP device group is in PAIR state if the **ApplicationStartup** object is set to **RESYNCWAIT**. Even though RAID Manager XP and the XP firmware fully support delta resynchronization, the delta between the primary and the secondary disks could be big enough for the copy process to exceed the package startup timeout value, if not set to NO_TIMEOUT.
- The **ResyncWaitTimeout** object can automatically lead into failed Serviceguard packages when set higher than the package startup timeout value.
- If running in fence level **ASYNC**, the default value of the **AsyncTakeoverTimeout** can cause the package to fail because its value is set beyond usual recommended startup timeout values. This is done because the takeover process for fence level **ASYNC** can take much longer when slow communications links are in place.

  To prevent takeover commands from being terminated by the takeover timeout before finishing, the time to copy the installed XP disk array cache should be measured and the package startup timeout interval should be adjusted according to the measured copy time. When measuring the copy time, measure only the slowest link used for Continuous Access XP Extension. This ensures that the XP disk array cache can be transferred from the remote XP disk array, even in the event of a single surviving replication link between the XP disk arrays.

Because the failover environment is dispersed into two (or more) data centers, the failover time can not be expected to be the same as it would be in a single data center with a single shared disk device. Therefore, the package startup timeout value, the monitor interval of the RAID Manager XP device group, should be adjusted based on failover tests performed to verify the proper configuration setup.

9

Command-line interface (CLI)

Cluster Extension XP allows integration into almost any cluster software for commercial UNIX, Linux, and Windows operating systems. The command-line interface allows integration of Continuous Access XP configurations for these supported platforms:

- Sun Solaris 2.6
  Sun Solaris 7
  Sun Solaris 8
  Sun Solaris 9
- Windows 2000
- Windows Server 2003
- AIX 4.3.3
  AIX 5.1
- Linux Red Hat 7.1 and Kernel 2.4.2
- Linux Red Hat Advanced Server 2.1 and Kernel 2.4.9-e.3 or 2.4.9-e.25
- SUSE Linux Enterprise Server 8 powered by Unitedlinux 1.0 base version or Service Pack 2a

The Cluster Extension XP **clxrun** command can be used to check proper funtionality of Cluster Extension XP prior to integration with the cluster software.

Configuring the CLI

Using the Cluster Extension XP CLI requires the following configuration steps:

1. Create the Continuous Access XP environment.
2. Create the RAID Manager XP configuration.
3. Create and configure the user configuration file (page 194).

Creating the Continuous Access environment and configuring RAID Manager

HP support personnel are trained and authorized to set up Continuous Access XP (Extension). You can, however, configure and change XP disk pairs and RAID Manager XP instances using HP StorageWorks LUN Configuration Manager XP/HP StorageWorks CommandView XP and HP StorageWorks RAID Manager XP. For more information, refer to the documentation listed in “HP storage website” (page 14).

Timing considerations

Cluster Extension XP is designed to prioritize XP disk array operations over application service startup operations. If Cluster Extension XP invokes disk pair resynchronization operations or gathers information about the remote XP disk array, Cluster Extension XP will wait until the requested status information is reported. This prioritizes data integrity over application service startup and failover behavior.

Because the takeover timing depends on the configuration of your RAID Manager XP environment and the settings in **UCF.cfg**, these considerations must be evaluated:

- Cluster Extension XP uses RAID Manager XP instances to communicate with the remote XP disk array. Depending on the settings of the RAID Manager XP instance timeout parameter and the number of remote instances, the online operation could time out. This can also happen if **clxrun** is used in a script or called by another program and the local RAID Manager XP instance cannot reach the remote RAID Manager XP instance. See “RAID Manager XP configuration” on page 90 for further information.

- If the ApplicationStartup attribute is set to **RESYNCWAIT**, Cluster Extension XP tries to resynchronize disk pairs and waits until the RAID Manager XP device group is in PAIR state. In some versions of RAID Manager XP and XP firmware, a full resynchronization is done. Depending on the amount of data to be transferred, it could take hours to resynchronize. If this is the case, **clxrun** may take some time to return. Do not stop **clxrun**; use **pairdisplay** to check the status of the associated RAID Manager XP device groups.

  Even if the RAID Manager XP version and the XP firmware version allow a delta resynchronization, the amount of delta data to be transferred between the primary and the secondary could be long enough for the copy process to take awhile.

- If running in fence level **ASYNC**, the default value of the **AsyncTakeoverTimeout** is set to a very high number. This is done because the takeover process for fence level **ASYNC** can take much longer when slow communications links are in place; adjust this value after measuring the Continuous Access XP environment. Refer to “AsyncTakeoverTimeout” on page 77 for further details.

  To prevent takeover commands from being terminated by the takeover timeout before it is finished, the time to copy the installed XP family disk array cache should be measured and the resource online timeout interval should be adjusted according to the measured copy time. *Use only the slowest link used for Continuous Access XP Extension to measure the copy time.* This ensures that the XP disk array cache can be transferred from the remote XP disk array, even in the event of a single surviving replication link between the XP family disk arrays.

In general, because the failover environment is dispersed into two (or more) data centers, the failover time can not be expected to be the same as it would be in a single data center with a single shared disk device.

Restrictions for customized Cluster Extension XP implementations

The following are some restrictions which apply when using the Cluster Extension XP CLI.

- The Cluster Extension XP CLI call **clxrun** must be invoked before the associated disk resources are activated.
- Associated disk resources must not be activated on any other system. If other disk resources are activated, Cluster Extension XP may remove write-access rights for those disk devices (putting them in read-only mode).

Creating and configuring the user configuration file

The command line interface expects as an argument the name configured as the **APPLICATION** tag value. You do not need to specify the **SearchObject** object.

Example The following is an example of a customized user configuration file when using **clxrun**.

```
# /etc/opt/hpclx/conf/UCF.cfg
# This is the Cluster Extension XP User Configuration File (UCF.cfg).
# The COMMON tag specifies the configuration for the
# Cluster Extension XP core environment
#
COMMON
 LogLevel      info                 #show disk state info in the logs
# The APPLICATION tag specifies the configuration for the
# Cluster Extension XP failover behavior
APPLICATION    sap                   #the application service
 DeviceGroup   sapdg                #RM dev group for the app service
 RaidManagerInstances 22 90         #RM instance number for dev group
 XPSerialNumbers 34001 34005         #local and remote XP Serial
Numbers
 DC_A_Hosts    eserv1 eserv2         #data center A hostnames
 DC_B_Hosts    eserv3 eserv4         #data center B hostnames
 FenceLevel    data                  #FenceLevel changed from default
APPLICATION    netscape              #the application service
 DeviceGroup   netscapedg           #RM dev group for the app service
 RaidManagerInstances 22 90#RM instance number for dev group
 XPSerialNumbers 34001 34005         #local and remote XP Serial
Numbers
 DC_A_Hosts    eserv1 eserv2         #data center A hostnames
 DC_B_Hosts    eserv3 eserv4         #data center B hostnames
```

CLI Commands

Details of CLI command use are presented on the following pages:

clxrun: page 195
clxchkmon: page 197
clxqr: page 200

clxrun

Check disk set

Syntax

clxrun [**–forceflag**] *parameter*

Arguments

The **clxrun** program expects only one parameter as the default setting. This parameter is used to uniquely identify the application service in the **APPLICATION** section of the user configuration file.

clxrun first checks for the **–forceflag** option. When using **clxrun**, it is not necessary to create an *application_name***.forceflag** file. This option, however, must be specified first if used.

Caution

*The **forceflag** option is implemented as an emergency switch to manually activate your XP disk set. If the **forceflag** option has been specified, Cluster Extension XP will not check any consistency or concurrency rules before activating the XP disk set.*

Description

clxrun can be used to manually prepare the application service's disk set before an existing application service start procedure is invoked. When using **clxrun**, the status of the associated RAID Manager XP device group is checked to ensure that access to the disk set will occur under data consistency and concurrency situations only.

clxrun must be invoked before the application service disk set can be activated; it is considered an online-only program. However, the CLI features provide the same disaster recovery features as the integrated versions of Cluster Extension XP.

Execution of **clxrun** does not start the pair/resync monitor.

Return codes

clxrun exits with one of the following return codes:

0	**OK** Application service can be started.
1	**ERROR_GLOBAL** Application service should not start on any system in either site on either disk array.
2	**ERROR_DC** Application service should not start on any system in the local site on the local disk array.
3	**ERROR_LOCAL** Application service should not start on this system.

Example 1

```
# clxrun sap
```

In example 1, **clxrun** assumes that you have defined an **APPLICATION** tag named **sap** in the **UCF.cfg** file and you have specified all necessary objects including the **DeviceGroup** object to map the XP disk set to the application service **sap**. Cluster Extension XP will check the disk set mapped to the application service **sap**, run the necessary takeover procedure and return one of the return codes mentioned in the return code table.

Example 2

```
# clxrun -forceflag sap
```

In example 2, **clxrun** assumes that you have defined an **APPLICATION** tag named **sap** in the **UCF.cfg** file and you have specified all necessary objects including the **DeviceGroup** object to map the XP disk set to the application service **sap**. Cluster Extension XP will check the XP disk set mapped to the application service sap, and run the necessary takeover procedure to enable read/write access to the XP disk set.

clxchkmon

Pair/resync monitor access program

Syntax

clxchkmon [**–clx**]
[**–n** *resource_name* **–g** *device_group*]
[{ **–t** *monitor_interval* | **–autorecover** *mode* | **–remove** [**–force**] | **–show** | **–pid** | **–stopsrv** | **–log** {**error** | **warning** | **info** | **debug** } }]
[**–p** *port_number*]

–n *resource_name* Specifies the resource (application) name as used in Cluster Extension XP.

–g *device_group* Specifies a RAID Manager XP group name.

–t *monitor_interval* Specifies interval in seconds to update registered monitor resources.

–autorecover *mode* Specify **YES** to enable autorecovery, or **NO** to disable autorecovery for registered monitor resource.

–clx Executes the command only for Cluster Extension XP resources.

–remove Removes the resource from the monitor list.

–force Disables user confirmation to remove resource.

–show Displays monitored resources.

–pid Returns the process ID of the pair/resync monitor.

–stopsrv Stops the pair/resync monitor socket server.

–log Sets the log level for the pair/resync monitor.

–p *port_number* Specifies the port number to be used.

Description

The **clxchkmon** utility program allows starting and stopping of the resynchronization features and queries to gather state information of the monitored device groups.

To update or remove a specific resource, use **clxchkmon –n** *resource_name* **–g** *device_group*. If **–clx** is not specified, the command is applied only to non-Cluster Extension XP resources.

To update all non-Cluster Extension XP resources, use **clxchkmon –t**. To update Cluster Extension XP resources, use **clxchkmon –clx –t**.

Displaying resources

The following command displays all resources:

```
clxchkmon –show
```

The following command displays Cluster Extension XP resources only:

```
clxchkmon –clx –show
```

Removing resources

The following command removes only non-Cluster Extension XP resources:

```
clxchkmon –remove
```

The following command removes all Cluster Extension XP resources:

```
clxchkmon –clx –remove
```

Stopping the pair/resync monitor

The pair/resync monitor is stopped when all resources are removed from monitoring. To check whether the pair/resync monitor is running, execute the following command:

```
clxchkmon –show
```

Choose the application and device group combination you want to remove from the pair/resync monitor and remove it with the following command:

```
clxchkmon –n [ application_name | resource_group_name |
resource_name ] –g device_group_name –remove
```

where *application_name* | *resource_group_name* | *resource_name* is the resource name (as defined by the **APPLICATION** tag in the **UCF.cfg** file) of the Cluster Extension XP resource and should match the **clxchkmon** output. If the **–clx** option is not specified, the command is executed only for non-Cluster Extension XP resources.

Caution *If you respond Y (yes) to remove the combination, the resource will be removed from the list of resources to be monitored in the pair/resync monitor. If this is not an emergency removal attempt and the Cluster Extension XP resource is online, the procedure above will lead to a failed resource, which will offline all dependent resources and eventually force your application offline.*

Do not use this command to offline your Cluster Extension XP resources.

Return codes

clxchkmon exits with one of the following return codes:

0	Successful or device group is in PAIR state.
1	Device group is not in PAIR state.
2	Resource/device group is not registered with the pair/resync monitor.
3	Pair/resync monitor (**clxchkd**) not running.
4	Device group's pair status pending.
10	Pair/resync monitor internal error.
11	Invalid argument to pair/resync monitor.
12	Pair/resync monitor received signal (control-c) interrupt.
13	Unknown status for device group.
14	No port number specified in services file for clxmonitor.

16	Invalid use of the **–clx** flag on a non-Cluster Extension XP resource or Cluster Extension XP resource specified without the **–clx** flag.
100	RAID Manager XP error.

Related information

"Pair/resync monitor" (page 51)

clxqr

Clean up quorum service *Windows only*

Syntax

clxqr [**–splitpair** | **–cleanbit** | **–createpair**]

–splitpair	Splits the quorum service mirrored disk pairs and places them in SMPL (simplex) mode.
–cleanbit	Resets or initializes the state of the control devices.
–createpair	Reforms the quorum service mirrored disk pairs.

Description

The **clxqr** utility resets the state of Cluster Extension XP quorum service. Use this utility if the cluster group cannot be moved correctly and corruption of the quorum service metadata is suspected.

Related information

"Procedure for quorum service system cleanup" (page 233)

10

Troubleshooting

To troubleshoot problems with Cluster Extension XP, you must understand Continuous Access XP environments. Many issues can be isolated to incompatible disk pair states. Refer to the Continuous Access XP (Extension) and RAID Manager XP documentation before assuming that a problem has been caused by Cluster Extension XP.

Cluster Extension XP logs messages to the cluster-specific log location. However, it always keeps its own log file in its default log location.

Caution *Cluster Extension XP is not able to handle XP device group states automatically and correctly when they result from manual manipulations. Cluster Extension XP will try to automatically recover suspended RAID Manager XP device group states if the* ***AutoRecover*** *object is set to* ***YES****. However, if the recovery procedure experiences a problem Cluster Extension XP will not stop unless fence level* ***DATA*** *is used or the* ***ApplicationStartup*** *object is set to* ***RESYNCWAIT****. Therefore, ensure that the device group* ***PAIR*** *state has been recovered before the next failure occurs.*

Always disable automatic application service failover when resynchronizing disk pairs. A failure of the resynchronization source while resynchronizing can lead to unrecoverable data on the resynchronization target. The resynchronization process does not copy data in transactional order. For further information, see “Rolling disaster protection and Business Copy XP” (page 45).

Start errors

Errors can occur when the path to the RAID Manager XP binaries has not been set in the **PATH** environment variable.

If a user configuration file is not found in the correct directory location, Cluster Extension XP returns a local error.

Failover error handling

Cluster Extension XP automatically fails over application services if the system the application service is running on becomes unavailable. This also means that if a problem with the XP disk array state occurs, an application service startup process will be stopped. The behavior of Cluster Extension XP is highly configurable.

Depending on the customer setting, Cluster Extension XP is used to prevent application services from starting automatically under the wrong conditions.

Therefore, Cluster Extension XP will return local, data center-wide or even cluster-wide errors to prevent accidental access to the XP disk array disk set.

HACMP-specific error handling

Cluster Extension XP related messages are logged by HACMP to the following locations:

- **/usr/adm/cluster.log**

  General HACMP log file, which gives an overview of all events processed and whether they were successful or unsuccessful.

- **/tmp/hacmp.out**

  Detailed HACMP log file containing process logs of all event scripts. The output of Cluster Extension XP can also be found in this file.

The Cluster Extension XP log file is named **clxhacmp.log**.

Start errors

HACMP will go into a loop and wait until the problem is solved until the file **/etc/opt/hpclx/***application_name***.LOCK** has been removed. This process has been adopted from HACMP, which will also run in an endless loop in case of a failure until the user recovers all errors and starts the application manually. After all errors have been recovered, you can invoke the command **clruncmd** in order to return control back to the cluster software.

If the program is in a very early state of processing and experiences a problem before resolution of the application name, it may return an error return code. The **/etc/opt/hpclx/UNKNOWN.LOCK** file is created and must be removed after the problem has been resolved.

Failover errors

As mentioned above, the HACMP error handling of the Cluster Extension XP will create a **.LOCK** file for the resource group (for example, **/etc/opt/hpclx/OracleRG.LOCK**). Messages are logged to the log files **/var/opt/hpclx/log/clxhacmp.log** and **/tmp/hacmp.out**. The file can be removed after the problem has been solved. HACMP can then continue to start the resource group.

This file will be created for any error Cluster Extension XP returns. However, Cluster Extension XP will specify whether the error is a local, data center, or cluster-wide error.

The following example demonstrates the behavior of Cluster Extension XP for HACMP if a pair state is discovered (which does not allow for an automatic takeover operation by Cluster Extension XP). In this case, the pairs have been manually suspended. It is impossible for Cluster Extension XP to determine which copy of the mirrored data is the most current.

The output within **/tmp/hacmp.out** will show a similar message:

Example

```
clxHACMP:   > Fri Dec 15 16:35:19 NFT 2000
clxHACMP:   > Arguments: oracle ora1vg ora2vg 0 oracle PVOL_PSUS PSUS
SVOL_SSUS SSUS DATA 30368 30380 01-11-22/00 01.04.01
clxHACMP:   >    number of arguments: 14
clxHACMP:   >    1: oracle
clxHACMP:   >    2: ora1vg ora2vg
clxHACMP:   >    3: 0
clxHACMP:   >    4: oracle
clxHACMP:   >    5: PVOL_PSUS
clxHACMP:   >    6: PSUS
clxHACMP:   >    7: SVOL_SSUS
clxHACMP:   >    8: SSUS
clxHACMP:   >    9: DATA
clxHACMP:   >    10: 30368
clxHACMP:   >    11: 30380
clxHACMP:   >    12:
clxHACMP:   >    13: 01-11-22/00
clxHACMP:   >    14: 01.04.01
clxHACMP    > ===PRE===============================================
clxHACMP: pre-exec script successful (rc=0).
clxHACMP: ERROR - no takeover action found.
clxHACMP: ERROR - global cluster failure occurred - waiting!
clxHACMP: ERROR -
clxHACMP: ERROR -
================================================================
clxHACMP: ERROR - Cluster Extension XP takeover procedure FAILED.
clxHACMP: ERROR -
clxHACMP: ERROR - Pair state of device group "oracle" might be
clxHACMP: ERROR - incorrect. Manual checking and correction within
clxHACMP: ERROR - Continuous Aceess XP is required.
clxHACMP: ERROR - Remove file "/etc/opt/hpclx/OracleRG.LOCK" in order
clxHACMP: ERROR - to continue with HACMP specific recovery actions.

=================================================================
```

The last message is repeated every 5 minutes. Cluster Extension XP will stop any further processing until the you remove the *application_name***.LOCK** file to transfer control back to HACMP. This enables you to check the status of the data on each copy and decide whether it is safe to continue or not.

Depending on the amount of time needed for checking the configuration and the XP disk pair status, the HACMP timeout could be reached. This will automatically cause the event **config_too_long** to be called by HACMP. The following message will appear in the log file **/tmp/hacmp.out**:

```
WARNING: Cluster MYCLUSTER has been running recovery
program '/usr/es/sbin/cluster/events/node_up.rp' for
1110 seconds.

Please check cluster status.
```

If you think the Cluster Extension XP configuration is correct and the XP disk pair status allows you to manually continue to start the application, remove the application lock file **/etc/opt/hpclx/oracle.LOCK** mentioned in the error message above.

When this file has been removed, Cluster Extension XP transfers control back to HACMP. The event **get_disk_vg_fs** and all the subsequent events within the main event **node_up_local** will be processed. Because Cluster Extension XP as a pre-event of **get_disk_vg_fs** has produced an error, the main event **node_up_local** will fail as well. The following HACMP event **event_error** will be called:

```
node_up_local[30] [ 0 -ne 0 ]
node_up_local[8] exit 1
Dec 15 17:07:17 EVENT FAILED:1: node_up_local
node_up[326] [ 1 -ne 0 ]
node_up[328] cl_log 650 node_up: Failure occurred while processing
Resource Group OracleRG. Manual intervention required. node_up
OracleRG
***************************
Dec 15 2000 17:07:17 !!!!!!!!!! ERROR !!!!!!!!!!
***************************
Dec 15 2000 17:07:17 node_up: Failure occurred while processing
Resource Group OracleRG. Manual intervention required.
node_up[329] STATUS=1
node_up[337] [ AIX1 != AIX1 ]
node_up[356] exit 1
Dec 15 17:07:18 EVENT FAILED:1: node_up AIX1
```

To continue any further processing of HACMP, you must invoke the HACMP command **clruncmd** in order to recover from the status **event_error**:

Example

```
# clruncmd aix1
```

This will bring the cluster into normal status again. All subsequent events (for example, **node_up_complete**) will be processed.

Microsoft Cluster service-specific error handling

Cluster Extension XP related messages are logged by Microsoft Cluster service to the following locations:

%ClusterLog%\cluster.log

The Cluster Extension XP log file is named **clxmscs.log**.

The Cluster Extension XP quorum service log file is named **clxq.log** and resides in the **%SYSTEMROOT%** directory.

Solving quorum service problems

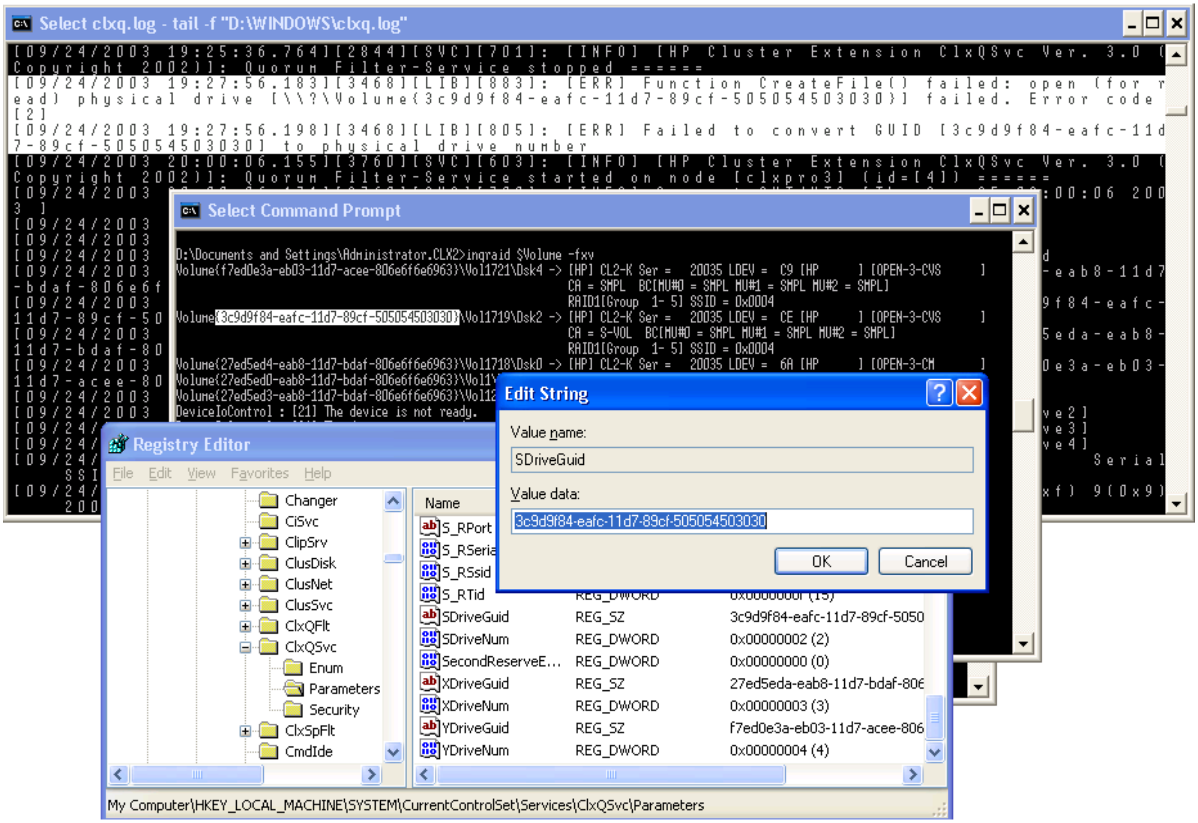
Use the following checklist to solve quorum service problems.

- If Microsoft Cluster service fails to start, or messages in the Event Log refer to the HP CLX quorum service or HP CLX quorum filter driver, a site failure may have occurred. See “Quorum service recovery” on page 230 in Appendix A for recovery instructions.
- If the **clxq.log** file shows problems converting GUIDs to physical drive numbers, the GUID may have been altered. Check the current GUID against the GUIDs in the registry key for the quorum service. You can find the current GUID by invoking the following Raid Manager XP command:

  inqraid $Volume -fvx

  If the disk is not shown, it doesn’t have a GUID. In this case the disk must be partitioned first. Refer to the Cluster Extension XP Installation Guide for details about the quorum service installation and prerequisites.

- If the GUID for any quorum service control disk is not the same as in the registry keys for GUIDs, you must update the registry key with the new GUID. If the registry key has changed for any of the quorum service control disks, the quorum service must be stopped and restarted.



Select clxq.log - tail -f "D:\WINDOWS\clxq.log"

```
[09/24/2003 19:25:36.764][2844][SVC][701]: [INFO] [HP Cluster Extension ClxQSvc Ver. 3.0 (
Copyright 2002)]: Quorum Filter-Service stopped ======
[09/24/2003 19:27:56.183][3468][LIB][883]: [ERR] Function CreateFile() failed: open (for r
ead) physical drive [\\?\Volume{3c9d9f84-eafc-11d7-89cf-505054503030}] failed. Error code
[2]
[09/24/2003 19:27:56.198][3468][LIB][805]: [ERR] Failed to convert GUID [3c9d9f84-eafc-11d
7-89cf-505054503030] to physical drive number
[09/24/2003 20:00:06.155][3760][SVC][603]: [INFO] [HP Cluster Extension ClxQSvc Ver. 3.0 (
Copyright 2002)]: Quorum Filter-Service started on node [clxpro3] (id=[4]) ======
```

Select Command Prompt

```
D:\Documents and Settings\Administrator.CLX2>inqraid $Volume -fxv
Volume{f7ed0e3a-eb03-11d7-acee-806e6f6e6963}\Vol1721\Dsk4 -> [HP] CL2-K Ser =   20035 LDEV =  C9 [HP      ] [OPEN-3-CVS     ]
                                                             CA = SMPL  BC[MU#0 = SMPL MU#1 = SMPL MU#2 = SMPL]
                                                             RAID1[Group  1- 5] SSID = 0x0004
Volume{3c9d9f84-eafc-11d7-89cf-505054503030}\Vol1719\Dsk2 -> [HP] CL2-K Ser =   20035 LDEV =  CE [HP      ] [OPEN-3-CVS     ]
                                                             CA = S-VOL  BC[MU#0 = SMPL MU#1 = SMPL MU#2 = SMPL]
                                                             RAID1[Group  1- 5] SSID = 0x0004
Volume{27ed5ed4-eab8-11d7-bdaf-806e6f6e6963}\Vol1718\Dsk0 -> [HP] CL2-K Ser =   20035 LDEV =  6A [HP      ] [OPEN-3-CM      ]
Volume{27ed5ed0-eab8-11d7-bdaf-806e6f6e6963}\Vol1
Volume{27ed5ed3-eab8-11d7-bdaf-806e6f6e6963}\Vol12
DeviceIoControl : [21] The device is not ready.
```

Edit String

Value name: SDriveGuid

Value data: 3c9d9f84-eafc-11d7-89cf-505054503030

Registry Editor

```
Name                 Type         Data
S_RPort
S_RSeria
S_RSsid
S_RTid
SDriveGuid           REG_SZ       3c9d9f84-eafc-11d7-89cf-5050
SDriveNum            REG_DWORD    0x00000002 (2)
SecondReserveE...    REG_DWORD    0x00000000 (0)
XDriveGuid           REG_SZ       27ed5eda-eab8-11d7-bdaf-806
XDriveNum            REG_DWORD    0x00000003 (3)
YDriveGuid           REG_SZ       f7ed0e3a-eb03-11d7-acee-806
YDriveNum            REG_DWORD    0x00000004 (4)
```

My Computer\HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\ClxQSvc\Parameters

- If the **clxq.log** file shows problems related to quorum service control disks or quorum disk pairing, pair the quorum disk and the quorum service control disks. The procedure is included in Appendix A: "Procedure for quorum service system cleanup" on page 233.

- The quorum disk may not appear in the **clxq.log** file if the path between your server system and the disk array has failed. Check the path between the server and the disk array before continuing to troubleshoot quorum service problems.

Emergency startup of Cluster service

If all quorum arbitration functions are unavailable or the disks cannot be paired, you can perform an emergency startup of the Cluster service:

1. Open the Services window.
2. Right-Click on ClxQSvc.
3. Select Properties.
4. In the Start Parameters field, enter

   `/createsplitbrain`
5. Click the start button in the Properties window. This will create two separate clusters.

Caution *The creation of separate clusters can result in data loss as explained below.*

Clicking the Start button in the Services window can lead to a split brain-syndrome, where the cluster runs the same applications on two different sets of disks. This can happen if the cluster is running or restarted in the remote data center and any number of cluster nodes are isolated from each other.

Resource start errors

Microsoft Cluster service configurations do not require a **UCF.cfg** file if the default common objects are used (recommended).

Microsoft Cluster service will fail the Cluster Extension XP resource on the local system if the **clxpcf** file is not present. If the program is in a very early state of processing, this operation might fail and Cluster Extension XP will not show the resource name in the error message. However, Microsoft Cluster service will fail the resource.

Failover errors

Cluster Extension XP's integration with Microsoft Cluster service returns a local error and fails the resource if a configuration error occurred. This could be a problem with the RAID Manager XP instance configuration or an error, which will probably allow starting the resource group on another system.

Cluster Extension XP resources return a data center error and fail the resource if the XP disk array status indicates that the problem experienced locally would not be solved on another system connected to the same XP disk array. This means all systems specified in the **DC_A_Hosts** resource property or the **DC_B_Hosts** resource property would fail to bring the resource group online.

Depending on the resource group and resource property values, the resource tries to start on different nodes several times. If the remote data center is down, this would look like the resource group alternates between the surviving systems. This happens until the above-mentioned resource and resource group property values are reached or restart of the resource is disabled by the user.

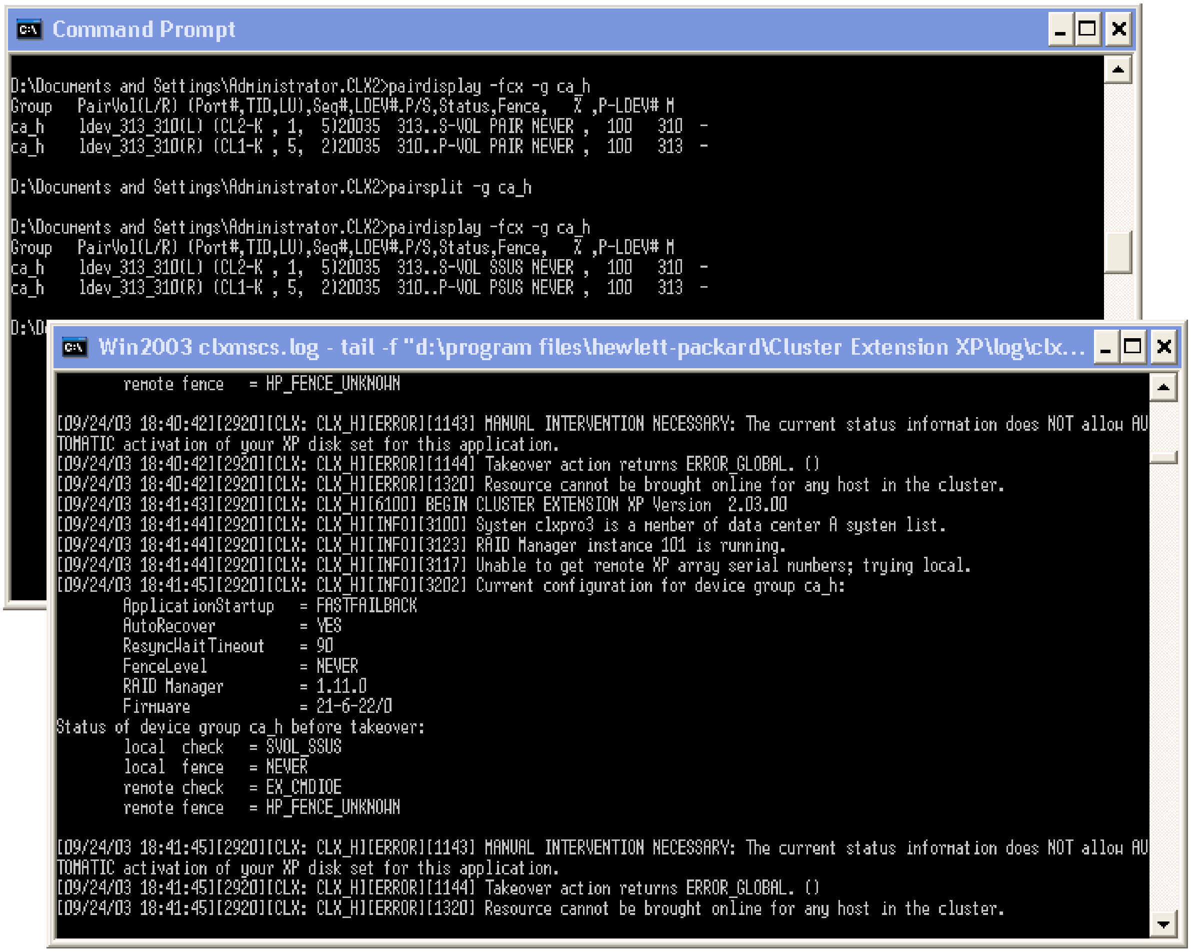
This could be also the case if the ApplicationStartup resource property has been set to **FASTFAILBACK**.

If an XP disk array state has been discovered that does not allow bringing the resource group online on any system in the cluster, a cluster error would be reported and the resource would fail on all systems. This could lead to the same behavior as described for a Cluster Extension XP data center error.

Such a state could be a SMPL state on both primary and secondary disks, a suspended (PSUS/SSUS) state on either site, or a state mismatch in the device group for this resource group. None of the above-mentioned scenarios will allow automatic recovery because the Cluster Extension XP resource cannot decide which copy of the data is the most current copy. In those cases, a storage or cluster administrator has to investigate what happened to the environment.

In any case, it is not recommended to restart a failed resource group without investigating the problem. A failed Cluster Extension XP resource indicates the need to check the status of the XP disk pair on each copy and decide whether it is safe to continue or not.

The following screens are examples of an incompatible XP disk pair state shown in the **clxmscs.log** file. The same messages can be found in the Microsoft Cluster service cluster log file if the Cluster Extension XP **LogLevel** object is set to **INFO**. This requires creating a **UCF.cfg** file.



```
D:\Documents and Settings\Administrator.CLX2>pairdisplay -fcx -g ca_h
Group   PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence,   % ,P-LDEV# M
ca_h    ldev_313_310(L) (CL2-K , 1,  5)20035  313..S-VOL PAIR NEVER ,  100   310  -
ca_h    ldev_313_310(R) (CL1-K , 5,  2)20035  310..P-VOL PAIR NEVER ,  100   313  -

D:\Documents and Settings\Administrator.CLX2>pairsplit -g ca_h

D:\Documents and Settings\Administrator.CLX2>pairdisplay -fcx -g ca_h
Group   PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence,   % ,P-LDEV# M
ca_h    ldev_313_310(L) (CL2-K , 1,  5)20035  313..S-VOL SSUS NEVER ,  100   310  -
ca_h    ldev_313_310(R) (CL1-K , 5,  2)20035  310..P-VOL PSUS NEVER ,  100   313  -

D:\D
```

Win2003 clxmscs.log - tail -f "d:\program files\hewlett-packard\Cluster Extension XP\log\clx...

```
        remote fence   = HP_FENCE_UNKNOWN

[09/24/03 18:40:42][2920][CLX: CLX_H][ERROR][1143] MANUAL INTERVENTION NECESSARY: The current status information does NOT allow AU
TOMATIC activation of your XP disk set for this application.
[09/24/03 18:40:42][2920][CLX: CLX_H][ERROR][1144] Takeover action returns ERROR_GLOBAL. ()
[09/24/03 18:40:42][2920][CLX: CLX_H][ERROR][1320] Resource cannot be brought online for any host in the cluster.
[09/24/03 18:41:43][2920][CLX: CLX_H][6100] BEGIN CLUSTER EXTENSION XP Version  2.03.00
[09/24/03 18:41:44][2920][CLX: CLX_H][INFO][3100] System clxpro3 is a member of data center A system list.
[09/24/03 18:41:44][2920][CLX: CLX_H][INFO][3123] RAID Manager instance 101 is running.
[09/24/03 18:41:44][2920][CLX: CLX_H][INFO][3117] Unable to get remote XP array serial numbers; trying local.
[09/24/03 18:41:45][2920][CLX: CLX_H][INFO][3202] Current configuration for device group ca_h:
        ApplicationStartup   = FASTFAILBACK
        AutoRecover          = YES
        ResyncWaitTimeout    = 90
        FenceLevel           = NEVER
        RAID Manager         = 1.11.0
        Firmware             = 21-6-22/0
Status of device group ca_h before takeover:
        local  check   = SVOL_SSUS
        local  fence   = NEVER
        remote check   = EX_CMDIOE
        remote fence   = HP_FENCE_UNKNOWN

[09/24/03 18:41:45][2920][CLX: CLX_H][ERROR][1143] MANUAL INTERVENTION NECESSARY: The current status information does NOT allow AU
TOMATIC activation of your XP disk set for this application.
[09/24/03 18:41:45][2920][CLX: CLX_H][ERROR][1144] Takeover action returns ERROR_GLOBAL. ()
[09/24/03 18:41:45][2920][CLX: CLX_H][ERROR][1320] Resource cannot be brought online for any host in the cluster.
```

VCS-specific error handling

Cluster Extension XP related messages are logged by VCS to the following locations:

- VCS 1.3.0 or later

  /var/VRTSvcs/log/engine_A.log
- VCS 1.1.2

  /var/VRTSvcs/log/engine.A_log

  General VCS engine log file, which gives an overview of all cluster-related activities and whether they were successful or unsuccessful.
- VCS 1.3.0 or later

  /var/VRTSvcs/log/ClusterExtensionXP_A.*log*
- VCS 1.1.2

  /var/VRTSvcs/log/ClusterExtensionXP.log_A

  Cluster Extension XP agent log file of VCS, which shows agent-related error information.

The Cluster Extension XP log file is named **clxvcs.log**.

Start errors

VCS will fail the resource and disable the service group on the local system if it the **clxpcf** file is not present. If the program is in a very early state of processing this operation might fail and Cluster Extension XP will not show the service group in the error message. However, VCS will fail the resource.

Failover errors

Cluster Extension XP's integration with VCS disables service groups on the local system if a configuration error occurs. In this case, Cluster Extension XP will return a local error.

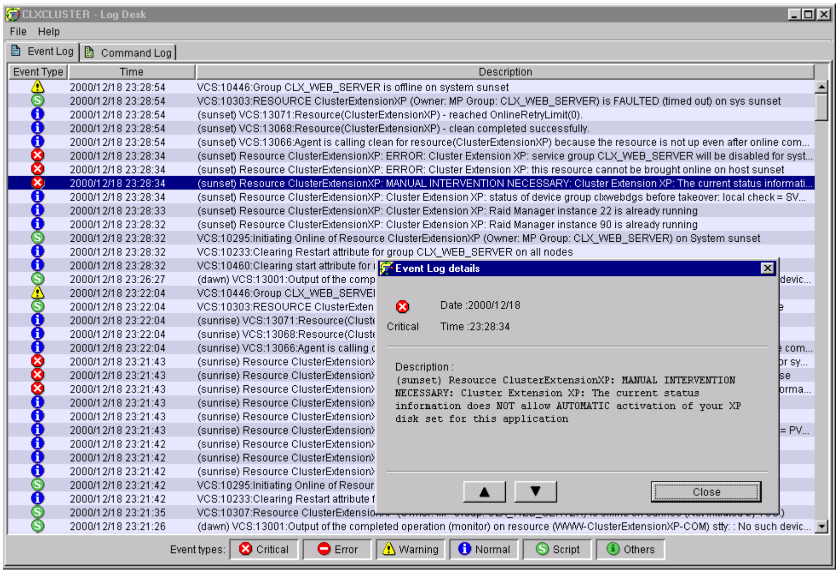
The service group is disabled in the data center if the XP disk array status indicates the problem experienced locally cannot be solved on another system connected to the same XP disk array. All systems specified in the **DC_A_Hosts** object or **DC_B_Hosts** object are disabled in order to bring the service group online.

This could be also the case if the **ApplicationStartup** object has been set to **FASTFAILBACK**.

If a XP disk array state has been discovered (which does not allow bringing the service group online on any system in the cluster), a cluster error is reported and all systems are disabled to bring the service group online. Such state could be a SMLP state on both primary and secondary disks, a suspended (PSUS/SSUS) state on either site, or a state mismatch in the device group for this service group. None of the scenarios allows automatic recovery because Cluster Extension XP cannot determine which copy of the data is the most current. In these cases, a storage or cluster administrator must investigate what happened to the environment.

Caution *It is not recommended to just enable the service group again and try to bring the prior failed service group online without investigating the problem. When a failed Cluster Extension XP resource occurs, check the status of the XP disk pair on each copy and decide whether it is safe to continue.*

Example The following screens show examples of an incompatible XP disk pair state shown in the VCS Cluster Manager Log Desk window.



Event Type	Time	Description
⚠	2000/12/18 23:28:54	VCS:10446:Group CLX_WEB_SERVER is offline on system sunset
S	2000/12/18 23:28:54	VCS:10303:RESOURCE ClusterExtensionXP (Owner: MP Group: CLX_WEB_SERVER) is FAULTED (timed out) on sys sunset
i	2000/12/18 23:28:54	(sunset) VCS:13071:Resource(ClusterExtensionXP) - reached OnlineRetryLimit(0).
i	2000/12/18 23:28:54	(sunset) VCS:13068:Resource(ClusterExtensionXP) - clean completed successfully.
i	2000/12/18 23:28:54	(sunset) VCS:13066:Agent is calling clean for resource(ClusterExtensionXP) because the resource is not up even after online com...
✖	2000/12/18 23:28:34	(sunset) Resource ClusterExtensionXP: ERROR: Cluster Extension XP: service group CLX_WEB_SERVER will be disabled for syst...
✖	2000/12/18 23:28:34	(sunset) Resource ClusterExtensionXP: ERROR: Cluster Extension XP: this resource cannot be brought online on host sunset
✖	2000/12/18 23:28:34	(sunset) Resource ClusterExtensionXP: MANUAL INTERVENTION NECESSARY: Cluster Extension XP: The current status informati...
i	2000/12/18 23:28:34	(sunset) Resource ClusterExtensionXP: Cluster Extension XP: status of device group clxwebdgs before takeover: local check = SV...
i	2000/12/18 23:28:33	(sunset) Resource ClusterExtensionXP: Cluster Extension XP: Raid Manager instance 22 is already running
i	2000/12/18 23:28:32	(sunset) Resource ClusterExtensionXP: Cluster Extension XP: Raid Manager instance 90 is already running
S	2000/12/18 23:28:32	VCS:10295:Initiating Online of Resource ClusterExtensionXP (Owner: MP Group: CLX_WEB_SERVER) on System sunset
i	2000/12/18 23:28:32	VCS:10233:Clearing Restart attribute for group CLX_WEB_SERVER on all nodes
i	2000/12/18 23:28:32	VCS:10460:Clearing start attribute for
S	2000/12/18 23:26:27	(dawn) VCS:13001:Output of the comp devic...
⚠	2000/12/18 23:22:04	VCS:10446:Group CLX_WEB_SERVE
S	2000/12/18 23:22:04	VCS:10303:RESOURCE ClusterExten e
i	2000/12/18 23:22:04	(sunrise) VCS:13071:Resource(Cluste
i	2000/12/18 23:22:04	(sunrise) VCS:13068:Resource(Cluste
i	2000/12/18 23:22:04	(sunrise) VCS:13066:Agent is calling c e com...
✖	2000/12/18 23:21:43	(sunrise) Resource ClusterExtension or sy...
✖	2000/12/18 23:21:43	(sunrise) Resource ClusterExtension se
✖	2000/12/18 23:21:43	(sunrise) Resource ClusterExtension orma...
i	2000/12/18 23:21:43	(sunrise) Resource ClusterExtension
i	2000/12/18 23:21:43	(sunrise) Resource ClusterExtension
i	2000/12/18 23:21:43	(sunrise) Resource ClusterExtension = PV...
i	2000/12/18 23:21:42	(sunrise) Resource ClusterExtension
i	2000/12/18 23:21:42	(sunrise) Resource ClusterExtension
i	2000/12/18 23:21:42	(sunrise) Resource ClusterExtension
S	2000/12/18 23:21:42	VCS:10295:Initiating Online of Resour
i	2000/12/18 23:21:42	VCS:10233:Clearing Restart attribute f
S	2000/12/18 23:21:35	VCS:10307:Resource ClusterExtensionXP (Owner: MP Group: CLX_WEB_SERVER) is online on sunrise (Not initiated by VCS)
S	2000/12/18 23:21:26	(dawn) VCS:13001:Output of the completed operation (monitor) on resource (WWW-ClusterExtensionXP-COM) stty: : No such devic...

Critical

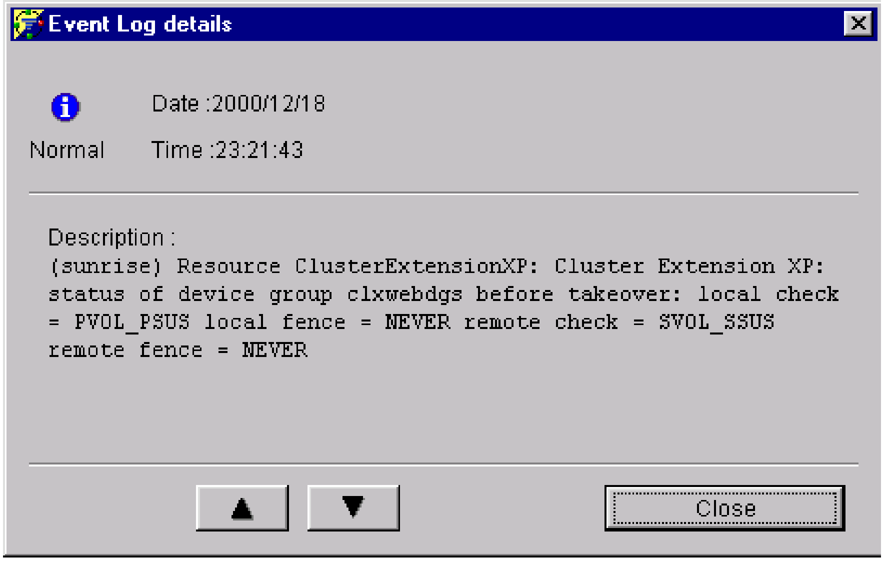
Date :2000/12/18

Time :23:28:34

Description :

```
(sunset) Resource ClusterExtensionXP: MANUAL INTERVENTION
NECESSARY: Cluster Extension XP: The current status
information does NOT allow AUTOMATIC activation of your XP
disk set for this application
```

The next screen shows detailed information for the current XP disk pair state, which will be displayed in the VCS Log Desk only if the Cluster Extension XP **LogLevel** object is set to **INFO**.



Date :2000/12/18

Time :23:21:43

Description :

```
(sunrise) Resource ClusterExtensionXP: Cluster Extension XP:
status of device group clxwebdgs before takeover: local check
= PVOL_PSUS local fence = NEVER remote check = SVOL_SSUS
remote fence = NEVER
```

Serviceguard (SG-LX)-specific error handling

Cluster Extension XP related messages are logged by Serviceguard to the following location:

/usr/local/cmcluster/conf/*package_name*/*package_control_file_name.log*

This is the default, recommended log file location, which gives an overview of all functions processed and whether they were successful or unsuccessful.

The Cluster Extension XP log file is named **clxmcsg.log**.

Start errors

Serviceguard will fail in the Cluster Extension XP function and disable the package on the local system if the **clxpcf** file is not present. If the program is in a very early state of processing, this operation may fail and Cluster Extension XP will not show the package name in the error message. However, Serviceguard will fail the package.

Failover errors

Cluster Extension XP's integration with Serviceguard disables packages on the local system if a configuration error occurred. In this case, Cluster Extension XP will return a local error.

The package will stop after a startup attempt on any system in the same data center if the disk array status indicates the problem experienced locally would not be solved on another system connected to the same disk array. Cluster Extension XP will return a data center error.

This could be also the case if the **ApplicationStartup** object has been set to **FASTFAILBACK**.

If a disk array state has been discovered that does not allow starting the package on any system in the cluster, a cluster error is reported and none of the systems will be allowed to run the package.

Such a state could be an SMLP state on both primary and secondary disks, a suspended (PSUS/SSUS) state on either site, or a state mismatch in the device group for this package. None of the scenarios allows automatic recovery because Cluster Extension XP cannot determine which copy of the data is the most current copy. In these cases, a storage or cluster administrator must investigate what happened to the environment.

Caution *Do not enable the package again and try to start the prior failed package without investigating the problem. When a package using Cluster Extension XP fails, check the status of the XP disk pair on each copy and decide whether it is safe to continue.*

Example The following shows an example of a package log file with a configuration error that causes the package to fail with a local error. The XP disk pair is configured to use fence level **NEVER**, but the **FenceLevel** object has been specified in the environment file to be **DATA**. (The **LogLevel** object has been set to **info**; otherwise, only the error messages would be shown.)

```
########### Node "matt": Starting package at Mon Mar 18 23:53:02 PST 2002 ###########
Mar 18 23:53:02 - Node "matt": This package is configured with remote data replication.
Mar 18 23:53:02 - Node "matt": Cluster Extension XP: onlining resource
Mar 18 23:53:02 - Node "matt": Cluster Extension XP: Raid Manager instance 101 is running
Mar 18 23:53:03 - Node "matt": ERROR: Cluster Extension XP: fence level mismatch: configured=data
current=NEVER
Mar 18 23:53:03 - Node "matt": ERROR: Cluster Extension XP: package CLX_APP1 cannot be started on this host
```

Example The next example shows a package log file with an incompatible XP disk pair state that causes the package to fail with a global error. (The **LogLevel** object has been set to **info**; otherwise, only the error messages would be shown.)

```
########### Node "matt": Starting package at Tue Mar 19 00:02:28 PST 2002 ###########
Mar 19 00:02:28 - Node "matt": This package is configured with remote data replication.
Mar 19 00:02:28 - Node "matt": Cluster Extension XP: onlining resource
Mar 19 00:02:28 - Node "matt": Cluster Extension XP: Raid Manager instance 101 is running
Mar 19 00:02:29 - Node "matt": Cluster Extension XP: current configuration for device group vgapp1:
ApplicationStartup = fastfailback
AutoRecover = yes
ResyncTimeout = 10
FenceLevel = never
Mar 19 00:02:29 - Node "matt": Cluster Extension XP: status of device group vgapp1 before takeover:
local check = PVOL_PSUS
local fence = NEVER
remote check = SVOL_SSUS
remote fence = NEVER
Mar 19 00:02:29 - Node "matt": MANUAL INTERVENTION NECESSARY: Cluster Extension XP: The current status
information does NOT allow AUTOMATIC activation of your Cluster Extension XP disk set for this application
Mar 19 00:02:29 - Node "matt": ERROR: Cluster Extension XP: takeover action returned ERROR_GLOBAL (1)
Mar 19 00:02:29 - Node "matt": ERROR: Cluster Extension XP: package CLX_APP1 cannot be started for any host
in the cluster
```

Pair/resync monitor messages in syslog/errorlog/messages/Event Log

Using the pair/resync monitor will cause a message in the system log file of your operating system (for any non-PAIR state of the device group, being monitored).

Those messages might indicate:

- The RAID Manager XP instance is not running or cannot be used to gather device group state information.
- The device group is not in the PAIR state.

This could be caused by Continuous Access XP link failures or manual manipulation of the disk pair state.

Recommendation Recover the PAIR state immediately because replication of your data is not possible.

Check monitored XP disk pairs by invoking the following command from the command line:

clxchkmon –n *application_name* **–g** *device_group* **–show**

Recommendation Disable application service failover for the time of the XP disk pair recovery (resynchronization). Cluster Extension XP's logic assumes that if the monitor is enabled, immediate action will be taken to recover a suspended XP disk pair.

Problem **Resource XYZ: Cluster Extension XP: device group XYZ is not in PAIR state**

This message appears even though the device group is in PAIR state.

Solution If you are using the pair/resync monitor, the **ResyncMonitorInterval** must be less than or equal to the resource monitor interval for the Cluster Extension XP resource to prevent erroneous logging.

The **ResyncMonitorInterval** in Cluster Extension XP defines when the pair/resync monitor checks the actual device group state. This state will be valid and shown until the next update (**ResyncMonitorInterval**) occurs. If the actual XP disk pair state changes between two **ResyncMonitorInterval**(s), the PAIR state shown by the pair/resync monitor will not be correct.

The resource monitor checks the status of the Cluster Extension XP resource at the resource monitor interval of the cluster software. The Cluster Extension XP resource reports the status of the device group at that interval based on the current state in the pair/resync monitor.

If the **ResyncMonitorInterval** is set to a higher value than the resource monitor interval for the Cluster Extension XP resource, the pair/resync monitor will update the device group state less often.

However, the Cluster Extension XP resource logs messages only if the device group is not in PAIR state or if a RAID Manager XP error occurred (for example, if RAID Manager XP is not running).

Example Set the ClusterExtensionXP agent's **MonitorInterval** attribute to 60 seconds (the default value); then set the Cluster Extension XP resource **ResyncMonitorInterval** attribute to less than 60 seconds.

A

Recovery procedures

RAID Manager XP version 01.04.xx introduces bitmap tables with the XP512/XP48 firmware version 01-11-xx/xx. This means a **horctakeover** command leads to a suspend state, where the XP disk array keeps track of changes on the disks rather than splitting the disk pairs if the link is down. RAID Manager XP 01.04.xx also has improved resynchronization commands. All this minimizes recovery effort dramatically.

However, earlier versions work with Cluster Extension XP. They differ in their behavior from the behavior described in this guide. With earlier software versions, the disk pairs will always be deleted, leaving a simplex disk (SMPL) behind if the remote RAID Manager XP instance cannot be reached. This also means that the recovery time is much longer, since the disk pair must be reestablished by a full copy process.

This recovery guideline is not intended to be sufficient for a full disaster recovery process. A disaster recovery process may include exchange of hardware. This appendix provides important information in case the Continuous Access XP link has been recovered and the disk pairs must be reestablished.

XP disk pair states

The following information is provided for a basic understanding of XP disk pair state information. The XP disk pair state transition process is complex; therefore, the following information is not intended to be complete.

Table 2. Pair states for the XP family of disk arrays

State	Description
PVOL	The primary (master) disk of a disk pair.
SVOL	The secondary (slave) disk of a disk pair.
SMPL	The disk has no pair affinity to any other disk. (This could be shown in **pairdisplay** outputs for your Continuous Access XP disk if you accidentally exported the Business Copy XP environment variable **HORCC_MRCF**. In such case, the MU (mirror unit) number field will not be empty.)
PAIR	The disk is either primary disk or secondary disk. If both (PVOL and SVOL) disks are in PAIR state, Continuous Access XP updates the secondary disk based on the primary disk. If you see only one disk in PAIR state (while the second disk is in another state), one of the following has occurred: • The pair affinity on only one site of the disk pair was deleted. • A takeover command has been invoked on the secondary site, while no data has been written to the primary site and the Continuous Access XP link was down. • A takeover command has been invoked on the primary site with the fence level configured to DATA to release the fenced disk, while the Continuous Access XP link was down. The secondary disk would stay in PAIR state.)

(continued)

Table 2. Pair states for the XP family of disk arrays (Continued)

State	Description
PSUS	The pair affinity has been manually suspended or a takeover operation has been invoked on the secondary site with the fence level configured to **NEVER** (In this case the secondary disk would have the state SSUS **-** SSWS.)
SSUS	The pair affinity has been manually suspended or a takeover operation has been invoked on the secondary site. In this case the secondary disk would have the state SSWS if you invoke **pairdisplay** with the **–fc** option. In fence level **ASYNC**, the disk could also show PFUL or PFUS when using the **–fc** option.
SSUS - SSWS	Only the secondary disk could show SSUS. With the **–fc** option of **pairdisplay**, you can check whether somebody manually suspended the pair or a takeover command had been invoked. A prior takeover command is indicated by the SSWS state. In this case, the secondary disk is mandatory and a resynchronization can be done only from the SVOL site.
PSUE	The disk is in a failure mode. Either the Continuous Access XP link is down, or the disk must be replaced.
PDUB	The disk is in a failure mode. Either the Continuous Access XP link is down, or the disk must be replaced. This is a special state of PSUE. If you have configured several disks into a LUSE configuration, where several LDevs are combined to create an extended size disk and one or more disks are in an error condition, this state will be shown.
PFUL	This state is used to indicate that a threshold of the side file area in the XP disk array cache has been reached. This state can be seen with fence level ASYNC only. Refer to the HP Continuous Access XP documentation for further information.
PFUS	This state is used to indicate that the side file has been full and the XP disk array was not able to transfer the cache content to the remote XP disk array for a certain time. The XP disk pair has been suspended to continue processing host IO. This state can be seen with fence level ASYNC only. Refer to the HP Continuous Access XP documentation for further information.

Recovery sequence

To recover from a certain server or Continuous Access XP link failure, follow this recovery sequence.

1. Start the RAID Manager XP instances on both local and remote server.

   Linux/UNIX

   ```
   export HORCMINST=instance_number
   horcmstart.sh instance_number
   ```

   Windows 2000

   ```
   set HORCMINST=instance_number
   HORCMSTART instance_number
   ```

2. Gather general pair status information:

   ```
   pairdisplay –g device_group
   ```

3. Pair status information after a failed swap-takeover (the SVOL state is SSWS):

   ```
   pairdisplay –g device_group –fc
   ```

   To recover from these states, invoke the following command from the SVOL side:

   ```
   pairresync –swaps –c 15 –g device_group
   ```

   If the pair needs to be used on the old primary side the following commands must be invoked from the primary side:

   ```
   pairresync –swapp –c 15 –g device_group
   horctakeover –g device_group
   ```

4. Pair status information after a pvol-takeover (local PVOL PSUS; remote SVOL PAIR):

   ```
   pairdisplay –g device_group –fc
   ```

   To recover from these states, invoke the following command from the PVOL side:

   ```
   pairresync –c 15 –g device_group
   ```

Caution *The application must be shut down and the file systems unmounted before a fenced disk in fence level DATA can be set in read/write mode again. After the PVOL-takeover, the file system must be checked before it can be mounted. Any other recovery procedure could lead to unrecoverable file systems.*

- If a **horctakeover** command results in SVOL, or PVOL becomes SMPL and none of the disks in the device group has been written to, you can recover from this situation by splitting the remaining PVOL or SVOL to SMPL:

  pairsplit [**–S** | **–R**] **–g** *device_group*

  After splitting the pair, the pair can be recreated without copying its content using:

  paircreate –nocopy –c 15 –f *fence_level* **–g** *device_group* **–v**[**r**|**l**]

- If a **horctakeover** command results in SVOL or PVOL becomes SMPL and data was written to one of the disks in the device group, you can recover from this situation by splitting the remaining PVOL or SVOL to SMPL:

  pairsplit [**–S** | **–R**] **–g** *device_group*

  After splitting the pair, the pair can be recreated with a full copy using:

  paircreate –c 15 –f *fence_level* **–g** *device_group* **–v**[**r**|**l**]

  To ensure that a certain pair state has been established, invoke the event wait command:

  pairevtwait –g *device_group* **–t** *time_to_wait* **–s** *pair_state*

Quorum service recovery *(Microsoft Cluster service only)*

A site failure is characterized by a data center incurring a failure or disruption of all critical, on-site components or a data center losing all connectivity to an associated data center at a dispersed site. Power failures and natural disasters are common causes for this kind of failure.

To detect a failure, check the quorum service log file (**clxq.log**). A probable failure is indicated by the following message:

```
812  Quorum state uncertain. Cluster will be shut down
```

Confirm the site failure or communications loss manually.

Use following procedures to allow the cluster to proceed on the remaining site.

Related information **clxqr** command (page 200)

Caution *Ensure that the remote site remains down until communications are restored and the remote site can be rebooted to join the local cluster. Data corruption can occur if both sites have running clusters while there is no communication between sites. Before manually restarting the cluster at any site, the administrator must ensure that only one site or no site is in operation and providing services.*

Single site failure recovery

In the event of a single site failure and depending on the nature of that disruption, the external arbitrator will automatically maintain the remaining site, providing continuous service to clients. However, in order to preserve the data integrity, manual cleanup on the failed site is required before restoring the site.

If the primary site fails, with the support of the external arbitrator the secondary site will become the primary site to continue to service connected clients. If the secondary site fails, the external arbitrator will leave the primary site running. Regardless of which site fails, follow the steps below to restore the failed site.

At the failed site:

1. If the nodes are powered on, ensure that the cluster service is stopped. If the nodes are powered off, leave them off.
2. Restore disk array operation and the Continuous Access links.
3. Verify that the pair status of the quorum disk pair and the first quorum service control disk has been automatically recovered. This means the quorum disk pair is in PAIR state as well as the first quorum service control disk (status disk). If not, execute:

   clxqr –splitpair

Caution *Do not start the Cluster service.*

   The quorum service at the primary site will automatically attempt to create the quorum disk pair and the first quorum service control disk pairs (status disks) and start copying data to this site. No action is needed.

4. When the copy process is finished, ensure that all network links are up; then power on all nodes on this site to rejoin the cluster or start the Cluster service.

After finishing the steps at the failed site, continue recovery at the current primary site:

5. Check to determine that the quorum and status disks are in the PVOL-PAIR state. If not, pause the quorum service **ClxQSvc** on the host node that owns the quorum disk. From a command prompt, execute:

   clxqr –splitpair

   clxqr –createpair

   When the pair copy process is finished, resume the **ClxQSvc** service.

If any error occurs during this procedure, follow the instructions in “Procedure for quorum service system cleanup” (page 233) to restart the quorum service.

Restore any of the data disk pairs managed by the Cluster Extension XP resource in the Cluster service, if necessary.

After restoration of the cluster, if the primary and secondary role was changed due to the failure, you can revert ownership roles back to their original conditions by using Cluster Administrator to move the cluster group, including the quorum disk and the resource groups, to their original nodes.

Failure recovery if both sites have failed

If an external arbitrator was not deployed or, for some reason, failed to work during a site failure, the Cluster Extension XP quorum service will shut down both sites to preserve data integrity. To restore both sites, take the following steps to restore one site at a time.

1. Determine which site contains the application data that you want to preserve. This site will become your new primary site.
2. Make sure all nodes are powered off at both sites. Recover both disk arrays.
3. If the quorum disk pair and the first quorum service control disk pairs are not already in a SMPL state, split the quorum disk pair and the first quorum service control disk pair.

   At the new primary site, from a command prompt, execute:

   clxqr –splitpair

   clxqr –createpair
4. When the pair copy process is finished, make sure all network connections for all nodes are resumed. Start the cluster nodes for this site, one at a time. The first node started will take ownership of the quorum disk.
5. To continue the recovery, follow the procedure for “Single site failure recovery” (page 230).

If the disk pairs cannot be recreated because the disk array has been destroyed, use the emergency startup procedure explained in “Solving quorum service problems” on page 211 in Chapter 10.

Procedure for quorum service system cleanup

Use this procedure to reset the state of the quorum service if the cluster group cannot be moved correctly and corruption of the quorum service is suspected.

1. Ensure that all user applications on the cluster have been stopped.
2. Pause the quorum service on all nodes in the cluster.

   Check each node in the cluster, and make sure the quorum service is paused.
3. Run the following command on one running node in each data center to clean the pair status.

   clxqr –splitpair

   Ignore any warning or errors.

   This command ensures that both sides of mirror pairs used by the quorum service are in SMPL mode.

   To run this command, at least one node must be up on each site.
4. Run the following command on one node that is running in each data center:

   clxqr –cleanbit

   This command ensures that all bits (status and lock) are cleaned.

   To run this command, at least one node must be up on each site.
5. Run the following command only on the node that currently owns the quorum resource (the node that has the cluster group online).

   clxqr –createpair

   This command creates the quorum disk pair and the quorum service control disk 1 (status disk) pair.
6. Wait for the pair copies to be completed.
7. Resume the Cluster Extension XP quorum service on all nodes, one at a time, starting with the node that has the cluster group online.

B

Cluster Extension XP resource message catalog

```
## The following error messages are logged to the clx{integration}.log
## log file. Where possible, these message are shown in the cluster-specific
## logging location also.
##
## Integrations are vcs, mcsg, mscs, hacmp, and run.
## The MSCS integration does NOT include the quorum service logging.
## All messages from the quorum service are logged to clxq.log and clxqarb.log.
##
## Strings: 100 -- 999
##

100     Program ended with return code: OPERATION_SUCCESSFUL.
101     Program ended with return code: INTERNAL_ERROR.
102     Pair/Resync Monitor started. Process ID = [%s1].
```

```
103      Can't add first object. Exiting...



##
## Log Level ERROR messages: 1000 -- 1999
##                   common: 1000 -- 1099
##   Cluster Extension core: 1100 -- 1499
##      pair/resync monitor: 1500 -- 1999
##

1000      File [%s1] open failed.
1001      File [%s1] access failed.
1002      File [%s1] not found or does not have execute permissions.
1003      File [%s1] not found.



1101      Failover does not lead into maintaining the PAIR Status of related disks
- stopped startup process.
1103      No device group found.
1104      Cannot find resource group for [%s1].
1105      Object index number [%s1] was not found in object database.
1106      Object name [%s1] was not found in object database.
1107      Object name [%s1] value not set.
1108      Cannot set value for object name [%s1].
1109      Object name [%s1] has invalid tag: [%s2] Expecting: [%s3].
1110      Object name [%s1] has invalid type: [%s2] Expecting: [%s3].
```

1111 Firmware version number has incorrect format.

1112 RAID Manager version number has incorrect format.

1113 Argument value is not numeric.

1114 Invalid return code.

1115 Device [%s1] is not supported by Cluster Extension XP - exiting.

1117 System [%s1] configured in both data center A and B system lists.

1118 System [%s1] is not configured in neither data center A's nor data center B's system list.

1119 Unexpected [%s1] at line #[%s2] in product configuration file.

1120 Invalid object type [%s1] at line #[%s2] in product configuration file.

1121 Invalid object tag [%s1] at line #[%s2] in product configuration file.

1122 Missing APPLICATION name at line #[%s1] in user configuration file.

1123 No value for object name [%s1] at line #[%s2] in user configuration file.

1124 Cannot set environment variable [%s1] with value [%s2].

1125 Cannot determine command device.

1126 Cannot execute shell command [%s1].

1127 Cannot set any RAID Manager instance environment variable.

1129 Device group state check failed locally.

1130 No device group found.

1131 Waiting for resynchronization to complete failed.

1132 Splitting the device group failed.

1133 Disk pair resynchronization failed for device group [%s1]. Check the device group state and recover the state manually.

1134 Multiple XP arrays configuration not supported.

1135 Fence Level for device group [%s1] could not be determined. Please check if the device group is in Pair state.

1136 Fence level mismatch: configured=[%s1] current=[%s2].

1137 Unsupported cluster software: [%s1].

1138 Cannot start any RAID Manager instances.

1139 Unable to get local RAID Manager and array firmware versions.

1140 The installed RAID Manager software is not a Hewlett-Packard product and is not supported by Cluster Extension XP.

1141 Check RAID Manager device group state for [%s1].

1142 Cannot find takeover function [%s1].

1143 MANUAL INTERVENTION NECESSARY: The current status information does NOT allow AUTOMATIC activation of your XP disk set for this application.

1144 Takeover action returns [%s1]. ([%s2])

1145 Initialization failed.

1146 Unable to get XP array serial numbers. Please check local and remote Raid Manager configuration.

1147 XP array serial number [%s1] not found in user configuration file.

1148 Unable to add serial number [%s1] to object database.

1149 Pre-exec script [%s1] not found.

1150 Pre-exec script FAILED (rc=[%s1]).

1151 Pre-exec failed; no takeover; stop application resource cluster-wide.

1152 Pre-exec failed; no takeover; stop application resource at all nodes in this data center.

1153 Pre-exec failed; no takeover; stop application resource on this node.

1154 Unknown pre-exec return code (rc=[%s1]).

1155 Pre-exec successful; no takeover; no post-exec (rc=[%s1]).

1156 Pre-exec script FAILED; continuing with takeover (rc=PRE_ERROR_TAKEOVER).

1157 Post-exec script [%s1] not found.

1158 Post-exec script FAILED (rc=[%s1]).

1159 Post-exec failed; stop application resource cluster-wide.

1160 Post-exec failed; stop application resource at all nodes in data center.

1161 Post-exec failed; stop application resource at this node.

1162 Post-exec failed; continuing to online application resource on this node.

1163 Unknown post-exec return code (rc=[%s1]).

1164 Improper number of arguments is passed to CLX.

1167 Invalid argument ([%s1]) specified - exiting.

1168 Pair/Resync Monitor utility ([%s1]clxchkmon) not found or does not have execute permissions; check that Pair/Resync Monitor is properly installed, or if Pair/Resync Monitor is not installed, change the [%s2] attribute to \"no\" to disable Pair/Resync Monitor option.

1169 Device group [%s1] is not registered with Pair/Resync Monitor.

1170 Pair/Resync Monitor is not running.

1171 Cannot remove device group [%s1] from Pair/Resync Monitor; failed to take resource offline. Remove the resource group from the Pair/Resync Monitor using the clxchkmon -remove command.

1174 Invalid cluster return code.

1200 \n==\n\nCluster Extension XP takeover procedure FAILED.\n\nPair state of device group [%s1] might be incorrect.\n\nManual checking and correction within CA is required.\n\n==\n

```
1201     Local error occurred; CLX stops application resource at this node!

1202     Data center error occurred; CLX stops application resource in data center!

1203     Global error occurred; CLX stops application resource cluster-wide!

1204     Cannot unset environment variable [%s1].

1205     Invalid value [%s1] - Possible value [%s2].

1206     Cannot create temporary directory [%s1].

1207     Cannot generate temporary file name [%s1].

1208     Invalid argument order.

1210     Failed to take resource offline.

1211     Cannot create directory [%s1].  Using default [%s2].

1300     Cannot remove online file ([%s1]).

1301     Online file ([%s1]) created; removing device group [%s2] from Pair/Resync
Monitor.

1302     [%s1] opton is set to \"no\"; device group [%s2] is registered with
Pair/Resync Monitor but online file does not exist; creating online file ([%s3]).

1303     Online file ([%s1]) exists and device group [%s2] is registered with
Pair/Resync Monitor; [%s3] option is set to \"no\"; removing resource from
Pair/Resync Monitor.

1304     Device group [%s1] is not in PAIR state.

1306     Cannot remove device group [%s1] from Pair/Resync Monitor.

1307     Pair/Resync Monitor is reporting RAID Manager XP errors for device group
[%s1].

1308     Cannot get status for device group [%s1] from the Pair/Resync Monitor.

1309     Unknown resource state reported by RAID Manager XP for device group [%s1].

1310     Resource is not registered with Pair/Resync Monitor but online file
([%s1]) exists; [%s2] attribute may have been enable subsequent to the resource
being brought online; either run the Pair/Resync Monitor manually or bring the
resource offline then online to register the resource with the Pair/Resync Monitor
or disable the [%s3] attribute.

1311     Pair/Resync Monitor is not running but online file ([%s1]) exists; [%s2]
attribute may have been enabled subsequent to the resource being brought online;
either start the Pair/Resync Monitor manually or bring the resource offline then
online to start the Pair/Resync Monitor or disable the [%s3] attribute.
```

```
1312      Cannot create online file [%s1]; continuing to use Pair/Resync Monitor.

1313      Cannot create directory [%s1] for online file [%s2].online.

1314      Cannot create online file [%s1].

1315      Creating online file [%s1].

1316      Resource cannot be brought online on this host.

1317      Resource cannot be brought online on host [%s1].

1318      Resource cannot be brought online for any host in the data center.

1319      Resource cannot be brought online for any host in data center [%s1].

1320      Resource cannot be brought online for any host in the cluster.



# BC ERRORs

1321      Business Copy MU# [%s1] does not exist

1322      Cannot insert into list object: [%s1]

1323    Rolling Disaster Protection enabled, but cannot split any of the specified
Business Copies; none are in PAIR status; create Business Copy(s) of device group
"[%s1]" and restart the application resource "[%s2]".  Or set the force flag to
disable the Rolling Disaster Protection feature.

1324    Rolling Disaster Protection enabled, but cannot split any of the specified
Business Copies for device group "[%s1]" because cannot reach remote host; check to
make sure RAID Manager is running on the remote data center and restart the resource
group "[%s2]".  Or set the force flag to disable the Rolling Disaster Protection
feature.

1325Invalid arguments passed to clxcmd module.



#AutoPass

1330  AutoPass has detected that this version of Cluster Extension XP does not have
a valid license.  You will not be able to online CLX resources until you obtain a
valid license.  Run [%s1] to obtain or import a permanent license.

1331  Unable to load AutoPass library [%s1].  Please make sure AutoPass is correctly
installed.

1332  Failed to find AutoPass API function.  Please make sure AutoPass is correctly
installed.
```

```
1333  Dumping AutoPass ERR obj:\n

1334  Error Number    = [%s1]\n

1335  Error Source    = [%s1]\n

1336  Error Message   = [%s1]\n

1337  Failed to decrypt success return code.

1338  Invalid command line option.

1339   AutoPass did not find any licenses to be removed.

1340  Could not create ClxAutoPass daemon file [%s1].

1341  ClxAutoPass failed to start the AutoPass GUI.  Please make sure AutoPass and
the java runtime environment is installed correctly and the path environment
variable includes the directory at which the java executable resides.



1500      Out of memory resources.

1501      Could not initialize WS2_32.DLL, WSAStartup failed.

1502      Unable to find port number of clxmonitor in services file.


1505      Establish connection to host [%s1] failed.

1506      Could not set environment variable: [%s1]

1507      Cannot get shared memory.

1508      Shared memory reached maximum size.

1509      Create process to execute command "[%s1]" failed.

1510      Cannot download shared memory: no server is running.

1511      Semaphore initialization failed.

1512      Semaphore not initialized.

1513      Semaphore operation error.

1514      Creation of the main semaphore failed.
```

```
1515      Set security descriptor to default failed.

1516      First load shared memory failed.

1517      Setting environment variable [GUARDIAN_RESYNC_CLXCHKD_CHILD_DAEMON]
failed.

1518      Setting the log level to [%s1] failed.








##

## Log Level WARNING messages: 2000 -- 2999

##                      common: 2000 -- 2099

##      Cluster Extension core: 2100 -- 2499

##         pair/resync monitor: 2500 -- 2999

##



2100      Object name [%s1] value not set.  Using default [%s2].

2101      Disk pair resynchronization failed for device group [%s1].  Check the
device group state and recover the state manually.

2102      Unknown RAID Manager error code [%s1].

2103      Device group state check failed.  Trying local.

2104      ResyncMonitor was configured to "yes", but due to an error, the resource
group [%s1] was not added to the Pair/Resync Monitor. Restart the HACMP resource
group [%s1] to use the Pair/Resync Monitor for this resource.

2105      Waiting for Pair state timeout.

2106      Resyncwait timeout has been reached.  Synchronization percentage has not
changed since last interval!  Please adjust the ResyncWaitTimeout value and check
the CA link utilization.

2300      Cannot find online file ([%s1]).
```

```
# BC Warnings

2400    Business Copy MU# [%s1] of device group [%s2] is not in BCMuList[%s3];
BCResyncMuList[%s3] should be a sub-set of BCMuList[%s3]

2401     Invalid Business Copy MU# [%s1] of device group [%s2] in list [%s3]; must
be in (0 - [%s4])

2402     Unable to split Business Copy MU# [%s1] of device group [%s2] in data
center [%s3]; check status

2403     Unable to resync Business Copy MU# [%s1] for device group [%s2] in data
center [%s3]; check status

2404     Unable to split any Business Copy MU for device group [%s1] in data center
[%s2]

2405     Unable to resync any Business Copy MU for device group [%s1] in data
center [%s2]

2406     Trying to resync previously split Business Copy pairs...

2407     Successful split of Business Copy MU# [%s1] of device group [%s2] in data
center [%s3]; manual resync might be necessary; check log file [%s4]clxrun.log for
successful resync if not using log level INFO.

2408     Duplicate Business Copy MU# [%s1] in list [%s2].

2409     Check state of Business Copy MU#: [%s1] of device group [%s2] in data
center [%s3]

2410Failed to find enviornment variable: [%s1].  Please make sure the enviornment
is set for the current process.

2411Failed to translate enviornment variable to actual path for the application
directory.  Using default path: [%s1] for the application directory.  In order for
the enviornment variable to be translated correctly, a single restart of the
cluster service might be necessary.



#AutoPass warnings

2450  AutoPass has detected that this version of Cluster Extension XP is using a
temporary license.  You have [%s1] day(s) left on your temporary license.  You will
not be able to online CLX resources if your license expires.  Run [%s2] to obtain or
import a permanent license.

2451  Invalid log level.  Setting log level to INFO.
```

2452 Failed to spawn ClxAutoPass daemon process to remind user to get a permanent license; No reminders will be sent.

2453 AutoPass has detected that this version of Cluster Extension XP is using a temporary license. Please refer to log file [%s1] for details.

2454 Failed to kill all [%s1] processes.

2455 Could not kill process: [%s1], with pid: [%s2], Error: [%s3]

2456 Call to OpenProcess failed, Error: [%s1]

2500 Add entry failed: ID [%s1] Name [%s2] RM device group [%s3].

2501 Add entry failed: Name [%s1] RM device group [%s2].

2502 Update entry failed: ID [%s1] Name [%s2] RM device group [%s3].

2503 Update entry failed: Name [%s1] RM device group [%s2].

2504 Update all entries check intervals failed.

2505 Show entry failed: ID [%s1] Name [%s2] RM device group [%s3].

2506 Show entry failed: Name [%s1] RM device group [%s2].

2507 Show all entries failed.

2508 Delete entry failed: ID [%s1] Name [%s2] RM device group [%s3].

2509 Delete entry failed: Name [%s1] RM device group [%s2].

2510 Delete all entries failed.

2511 Update entry failed: could not set pair check start time for ID [%s1] Name [%s2] RM device group [%s3].

2512 Update entry failed: could not set pair check start time for Name [%s1] RM device group [%s2].

2513 Update entry failed: could not set autorecover for ID [%s1] Name [%s2] RM device group [%s3].

2514 Update entry failed: could not set autorecover for Name [%s1] RM device group [%s2].

2515 Update entry failed: could not set volume status for ID [%s1] Name [%s2] RM device group [%s3].

2516 Update entry failed: could not set volume status for Name [%s1] RM device group [%s2].

2517 Update entry failed: could not set execute program flag for ID [%s1] Name [%s2] RM device group [%s3].

2518 Update entry failed: could not set execute program flag for Name [%s1] RM device group [%s2].

2519 Get all entries from shared memory failed.

2520 Host [%s1] is not in clxhosts list file.

2521 Pair resynchronization command terminated abnormally with [%s1].

2522 [[%s1]][[%s2]]: [%s3] resource's volumes are not in PAIR state. Attempt to resynchronize the volumes of RM device group [%s4]. Please, check the volume states!

2523 [[%s1]][[%s2]]: [%s3] resource's volumes are not in PAIR state or RAID Manager instance returned an error. Please, check the volume states and if necessary resynchronize the volumes of RM device group [%s4] soon!

2524 [[%s1]][[%s2]]: Pair/Resync Monitor cannot check the volume states for resource [%s3] using RM device group [%s4]. Please check the configuration for Raid Manager instance: [%s5]!

2525 [[%s1]][[%s2]]: [%s3] resource's volumes are not in PAIR state with volumes of RM device group [%s4]. Please check the configuration for Raid Manager instance: [%s5]!

2526 [[%s1]][[%s2]]: [%s3] resource's volumes of RM device group [%s4] are not in PAIR state and cannot communicate with or retrieve status information from remote Raid Manager instance : [%s5]!

2527 [[%s1]][[%s2]]: Cannot communicate with or retrieve status information from local/remote Raid Manager instance [%s5], [%s3] resource's volumes of RM device group [%s4] may not be in PAIR state. Please check the configuration for Raid Manager instance: [%s5]!

2528 [[%s1]][[%s2]]: Cannot communicate with or retrieve status information from local Raid Manager instance [%s5], [%s3] resource's volumes of RM device group [%s4] may not be in PAIR state. Please check the configuration for Raid Manager instance: [%s5]!

```
2539      Pair/Resync Monitor server did not respond to stop request of the
Pair/Resync Monitor daemon. Please stop (kill) any existing clxchksrv process
manually.



2600      Socket timeout.






##

## Log Level INFO messages: 3000 -- 3999

##                  common: 3000 -- 3099

##  Cluster Extension core: 3100 -- 3499

##     pair/resync monitor: 3500 -- 3999

##



3100      System [%s1] is a member of data center [%s2] system list.

3101      Device group state is PAIR.

3102      Synchronization percentage completed: [%s1].

3103      Timed out... synchronization percentage has not changed since last
interval!

3104      RAID Manager instance [%s1] not running.

3105      Current configuration for device group [%s1].

3106      Pair/Resync Monitor enabled.

3107      Pair/Resync Monitor with Auto Recovery enabled.

3108     Pair/Resync Monitor or utility ([%s1]) not found or does not have execute
permissions; check that Pair/Resync Monitor is properly installed, or if
Pair/Resync Monitor is not installed, change "[%s2]" attribute to "no" to disable
Cluster Extension XP's Pair/Resync Monitor option.

3109      Cannot start Pair/Resync Monitor.
```

3110 Invalid arguments for Pair/Resync Monitor.

3111 Cannot register device group [%s1] with Pair/Resync Monitor.

3112 Pair/Resync Monitor is disabled for device group [%s1].

3113 Calling post-exec script [%s1].

3114 Post-exec script successful (rc=[%s1]).

3115 Calling pre-exec script [%s1].

3116 Pre-exec script successful (rc=[%s1]).

3117 Unable to get remote XP array serial numbers; trying local.

3118 Online successful.

3120 Bringing application resource online.

3121 Resource taken offline successfully.

3122 RAID Manager instance [%s1] started successfully.

3123 RAID Manager instance [%s1] is running.

3124 RAID Manager instance [%s1] start failed; fork() was unsuccessful.

3125 RAID Manager instance [%s1] start failed because of an error.

3126 RAID Manager instance [%s1] caught a signal.

3127 RAID Manager caught a signal.

3129 Cannot get local host name.

3130 Cannot get operating system information.

3131 Host [%s1] found in data center [%s2].

3132 Takeover command returns [%s1].

3200 Status of device group [%s1] after takeover:\n

3201 Status after takeover:\n

3202 Current configuration for device group [%s1]:\n

3203 Current configuration:\n

3204 ApplicationStartup   = [%s1]\n

3205 AutoRecover   = [%s1]\n

3206 ResyncWaitTimeout   = [%s1]\n

3207 FenceLevel   = [%s1]\n

3208 DataLoseMirror   = [%s1]\n

3209 DataLoseDataCenter   = [%s1]\n

3210 AsyncTakeoverTimeout = [%s1]\n

3211 RAID Manager   = [%s1]\n

3212 Firmware   = [%s1]\n

3213 Status of device group [%s1] before takeover:\n

3214 Status before takeover:\n

3215 local check   = [%s1]\n

3216 local display = [%s1]\n

3217 local fence   = [%s1]\n

3218 remote check   = [%s1]\n

3219 remote display = [%s1]\n

3220 remote fence   = [%s1]\n

3221 Takeover action was successful.

3300 Cannot register device group [%s1] with Pair/Resync Monitor; creating online file [%s2].

3302 Successful resync of Business Copy MU# [%s1] of device group [%s2] in data center [%s3]

3303 Data center [%s1] MU#: [%s2] Status = [%s3]

3305 EnableNonHPDevice configuration option is enabled for device [%s1].

3306 Business Copy status before takeover in data center [%s1]\n

3307 Business Copy status after takeover in data center [%s1]\n

```
3308      Business Copy MU#'s specified to be split: [%s1]

3309      Business Copy MU#'s to split in PAIR status: [%s1]

3310      Business Copy MU#'s to split in invalid status: [%s1]

3311      Business Copy MU#'s specified to resync: [%s1]

3312      Business Copy MU#'s to resync in PAIR status: [%s1]

3313      Business Copy MU#'s to resync in invalid status: [%s1]

3314     At least one Business Copy specified to be split is in PAIR status; ready
to split.

3315      None of the Business Copies specified to be split are in PAIR status;
cannot split.

3316      At least one Business Copy specified to be resynced is in PAIR status;
ready to resync.

3317      Either no Business Copies were specified to be resynced or none of the
specified Business Copies to be resynced are in PAIR status; cannot proceed with
resync.



#AutoPass

3329  AutoPass has detected a valid license for this version of Cluster Extension
XP.

3330   AutoPass license info:\n

3331  LicenseString         = [%s1]\n

3332  PasswordType          = [%s1]\n

3333  Feature ID            = [%s1]\n

3334  Feature Version       = [%s1]\n

3335  Feature Description   = [%s1]\n

3336  Product Number        = [%s1]\n

3337  IP Address            = [%s1]\n

3338  LTU                   = [%s1]\n

3339  Capacity              = [%s1]\n
```

```
3340  LockedField           = [%s1]\n

3341  Future Date           = [%s1]\n

3342  Expiration Date       = [%s1]\n

3343  InstantOn Duration    = [%s1]\n

3344  IODaysRemaining       = [%s1]\n

3345  HostID                = [%s1]\n

3346  DeviceID              = [%s1]\n

3347  MACAddress            = [%s1]\n

3348  ProductBundle         = [%s1]\n

3349  ClusterInfo           = [%s1]\n

3350  Annotation            = [%s1]\n

3351  Created Time          = [%s1]\n

3352  InstantOnStartDate    = [%s1]\n

3353  LicenseEncryptionType = [%s1]\n

3354  AutoPass InstantOn installed successfully.

3355  AutoPass successfully removed license(s)

3356  Successfully started ClxAutoPass daemon.

3357   Successfully added password from license file [%s1].

3358  Successfully killed all [%s1] processes.

3359  Successfully killed process: [%s1], with pid: [%s2]


3500      Add entry: ID [%s1] Name [%s2] RM device group [%s3].

3501      Add entry: Name [%s1] RM device group [%s2].

3502      Update entry: ID [%s1] Name [%s2] RM device group [%s3].

3503      Update entry: Name [%s1] RM device group [%s2].

3504      All entries check intervals updated: internal [%s1].
```

```
3505      Entry deleted: ID [%s1] Name [%s2] RM device group [%s3].

3506      Entry deleted: Name [%s1] RM device group [%s2].

3507      All [%s1] entries deleted.

3508     Update entry: set pair check start time for ID [%s1] Name [%s2] RM device
group [%s3].

3509      Update entry: set pair check start time for Name [%s1] RM device group
[%s2].

3510      Update entry: set autorecover for ID [%s1] Name [%s2] RM device group
[%s3].

3511      Update entry: set autorecover for Name [%s1] RM device group [%s2].

3512      Update entry: set volume status for ID [%s1] Name [%s2] RM device group
[%s3].

3513      Update entry: set volume status for Name [%s1] RM device group [%s2].

3514     Update entry: set execute program flag for ID [%s1] Name [%s2] RM device
group [%s3].

3515      Update entry: set execute program flag for Name [%s1] RM device group
[%s2].

3516      No entries found.

3517      The log level for the Pair/Resync Monitor has been changed to [%s1].

3518      [[%s1]][[%s2]]: Name [%s3] RM device group [%s4] is in PAIR state.

3519      Pair/Resync Monitor server is halting after a stop request from the
Pair/Resync Monitor daemon.




##

## Log Level ALWAYS messages: 6000 -- 6999

##                    common: 6000 -- 6099

##    Cluster Extension core: 6100 -- 6499

##       pair/resync monitor: 6500 -- 6999
```

```
6100      BEGIN CLUSTER EXTENSION XP [%s1]

6101      Force flag is set; user enabled startup action.

6102      END CLUSTER EXTENSION XP




##

## Message numbers 9000 to 9999 are reserved for cluster-specific messages

## only.  That is, there are no equivalent generic messages.  You should

## create a message number between 9000 and 9999 and add your cluster

## offset to it to derive the message number.

## That is, there are no messages actually numbered "9000" to "9999".

##




##

## MSCS cluster-specific messages start at offset 10000 (MSCS_MSG_OFFSET)

##

#BC WARNING

12407    Successful split of Business Copy MU# [%s1] of device group [%s2] in data
center [%s3]; manual resync might be necessary; check log file [%s4] for successful
resync if not using log level INFO.

#BC ERROR

11323    Rolling Disaster Protection enabled, but cannot split any of the specified
Business Copies; none are in PAIR status; create Business Copy(s) of device group
"[%s1]" and restart the MSCS application resource "[%s2]".  Or set the force flag to
disable the Rolling Disaster Protection feature.




#AutoPass msgs for mscs
```

```
11330  AutoPass has detected that this version of Cluster Extension XP does not have
a valid license.  You will not be able to online CLX resources until you obtain a
valid license.  Run [%s1] to obtain or import a permanent license and then restart
the cluster service.  If a permanent or temporary license has already been
installed, please restart the cluster service for this to be reflected.

12450  AutoPass has detected that this version of Cluster Extension XP is using a
temporary license.  You have [%s1] day(s) left on your temporary license.  You will
not be able to online CLX resources if your license expires.  Run [%s2] to obtain or
import a permanent license and then restart the application resource.  If a
permanent license has already been installed, please restart the application
resource for this to be reflected.

12453  AutoPass has detected that this version of Cluster Extension XP is using a
temporary license.  Please refer to log file [%s1] for details.  If a permanent or
temporary license has already been installed, please restart the application
resource for this to be reflected.




##

## HACMP cluster-specific messages start at offset 20000 (HACMP_MSG_OFFSET)

##

#

# "HACMP only" messages

#



22522      [[%s1]][[%s2]]:  [%s3] resource group's volumes are not in PAIR state.
Attempt to resynchronize the volumes of RM device group [%s4]. Please check the
volume states!

22523     [[%s1]][[%s2]]:  [%s3] resource group's volumes are not in PAIR state or
RAID Manager instance returned an error. Please, check the volume states and if
necessary resynchronize the volumes of RM device group [%s4] soon!

22524      [[%s1]][[%s2]]:  Pair/Resync Monitor cannot check the volume states for
resource group [%s3] using RM device group [%s4]. Please check the configuration
for Raid Manager instance: [%s5]!

22525      [[%s1]][[%s2]]:  [%s3] resource group's volumes are not in PAIR state
with volumes of RM device group [%s4]. Please check the configuration for Raid
Manager instance: [%s5]!
```

```
22526     [[%s1]][[%s2]]:  [%s3] resource group's volumes of RM device group [%s4]
are not in PAIR state and cannot communicate with or retrieve status information
from remote Raid Manager instance : [%s5]!

22527      [[%s1]][[%s2]]:  Cannot communicate with or retrieve status information
from local/remote Raid Manager instance [%s5], [%s3] resource group's volumes of RM
device group [%s4] may not be in PAIR state. Please check the configuration for Raid
Manager instance: [%s5]!

22528      [[%s1]][[%s2]]:  Cannot communicate with or retrieve status information
from local Raid Manager instance [%s5], [%s3] resource group's volumes of RM device
group [%s4] may not be in PAIR state. Please check the configuration for Raid
Manager instance: [%s5]!



29002     CLX was unable to resolve resource group name to stop the Pair/Resync
Monitor.  Please check the UCF.cfg file and include ALL necessary volume groups for
each APPLICATION.  Remove the resource group from the Pair/Resync Monitor using the
clxchkmon -remove command.  Ignore this message if this resource group is not
CLX-aware.

29003     Lock file has been removed.

29004     Lock file [%s1] has been created.

29005     Ignoring HACMP dummy call.

29006     \nUsage: clxhacmp HACMPEvent Filesystems VolumeGroups\n\n

29007     Resource [[%s1]] device group [[%s2]] removed from Pair/Resync Monitor
successfully.



#BC WARNING

22407     Successful split of Business Copy MU# [%s1] of device group [%s2] in data
center [%s3]; manual resync might be necessary; check log file [%s4]clxhacmp.log
for successful resync if not using log level INFO.

#BC ERROR

21323    Rolling Disaster Protection enabled, but cannot split any of the specified
Business Copies; none are in PAIR status; create Business Copy(s) of device group
"[%s1]" and restart the HACMP resource group "[%s2]".  Or set the force flag to
disable the Rolling Disaster Protection feature.



##

## MCSG cluster-specific messages start at offset 30000 (MCSG_MSG_OFFSET)
```

```
##

31151     Pre-exec failed; no takeover; stop application package cluster-wide.

31152     Pre-exec failed; no takeover; stop application package in this data
center.

31153     Pre-exec failed; no takeover; stop application package on this node.

31159     Post-exec failed; stop application package cluster-wide.

31160     Post-exec failed; stop application package in data center.

31161     Post-exec failed; stop application package on this node.

31162     Post-exec failed; continuing to start application package on this node.

31171    Cannot remove device group [%s1] from Pair/Resync Monitor; failed to halt
package.  Remove the resource group from the Pair/Resync Monitor using the
clxchkmon -remove command.

# For MCSG message 31201 and 31202 are the same because ERROR_LOCAL==ERROR_DC.

31201     Package cannot be started on this node.

31202     Package cannot be started on this node.

31203     Package cannot be started for any node in the cluster.



32522      [[%s1]][[%s2]]:  [%s3] package's volumes are not in PAIR state. Attempt
to resynchronize the volumes of RM device group [%s4]. Please, check the volume
states!

32523      [[%s1]][[%s2]]:  [%s3] package's volumes are not in PAIR state or RAID
Manager instance returned an error. Please, check the volume states and if
necessary resynchronize the volumes of RM device group [%s4] soon!

32524      [[%s1]][[%s2]]:  Pair/Resync Monitor cannot check the volume states for
package [%s3] using RM device group [%s4]. Please check the configuration for Raid
Manager instance: [%s5]!

32525      [[%s1]][[%s2]]:  [%s3] package's volumes are not in PAIR state with
volumes of RM device group [%s4]. Please check the configuration for Raid Manager
instance: [%s5]!

32526     [[%s1]][[%s2]]:  [%s3] package's volumes of RM device group [%s4] are not
in PAIR state and cannot communicate with or retrieve status information from
remote Raid Manager instance : [%s5]!
```

```
32527      [[%s1]][[%s2]]:  Cannot communicate with or retrieve status information
from local/remote Raid Manager instance [%s5], [%s3] package's volumes of RM device
group [%s4] may not be in PAIR state. Please check the configuration for Raid
Manager instance: [%s5]!

32528      [[%s1]][[%s2]]:  Cannot communicate with or retrieve status information
from local Raid Manager instance [%s5], [%s3] package's volumes of RM device group
[%s4] may not be in PAIR state. Please check the configuration for Raid Manager
instance: [%s5]!




33118     Package started successfully.

33120     Starting package.

33121     Package halted successfully.



#BC WARNING

32407     Successful split of Business Copy MU# [%s1] of device group [%s2] in data
center [%s3]; manual resync might be necessary; check log file [%s4]clxmcsg.log for
successful resync if not using log level INFO.

#BC ERROR

31323    Rolling Disaster Protection enabled, but cannot split any of the specified
Business Copies; none are in PAIR status; create Business Copy(s) of device group
"[%s1]" and restart the MCSG application package "[%s2]".  Or set the force flag to
disable the Rolling Disaster Protection feature.



#

# "MCSG only" messages

#

39001    ResyncMonitor configuration parameter is either not set or is not enabled;
CLX Service requires the Pair/Resync Monitor to be enabled.

39002     Cannot run Pair/Resync Monitor command; CLX Service for package [%s1]
exiting.

39003     Device group [%s1] or Package [%s2] is no longer registered with
Pair/Resync Monitor. Check if Pair/Resync Monitor is running.
```

```
39004     Internal error; CLX Service for package [%s1] exiting.

39005     Starting CLX Service for package [%s1].

39006     No operation specified - exiting.

39007     Halting package.

39008     Cntrl-C Interrupt; CLX Service exiting.

39009     Service halt signal received; CLX Service exiting.

39010    \nUsage: clxmcsg <service|start|stop> <config_file> <package_name>\n
<config_file>  - Package configuration file (*.cntl).\n      <package_name> - Name
of package (application) in <package_name>_clx.env\n\n




##

## VCS cluster-specific messages start at offset 40000 (VCS_MSG_OFFSET)

##

41201     Application resource cannot be brought online on this host.

41202     Application resource cannot be brought online on any host in the data
center.

41203     Application resource cannot be brought online on any host in the cluster.



#AutoPass msgs for VCS

41330  AutoPass has detected that this version of Cluster Extension XP does not have
a valid license.  You will not be able to online CLX resources until you obtain a
valid license.  Run [%s1] to obtain or import a permanent license and then restart
the Cluster Extension XP agent.  If a permanent or temporary license has already
been installed, please restart the Cluster Extension XP agent for this to be
reflected.

42450  AutoPass has detected that this version of Cluster Extension XP is using a
temporary license.  You have [%s1] day(s) left on your temporary license.  You will
not be able to online CLX resources if your license expires.  Run [%s2] to obtain or
import a permanent license and then restart the Cluster Extension XP agent.  If a
permanent license has already been installed, please restart the Cluster Extension
XP agent for this to be reflected.
```

```
42453  AutoPass has detected that this version of Cluster Extension XP is using a
temporary license.  Please refer to log file [%s1] for details.  If a permanent or
temporary license has already been installed, please restart the Cluster Extension
XP agent for this to be reflected.




# BC WARNING

42407    Successful split of Business Copy MU# [%s1] of device group [%s2] in data
center [%s3]; manual resync might be necessary; check log file [%s4]clxvcs.log for
successful resync if not using log level INFO.

# BC ERROR

41323   Rolling Disaster Protection enabled, but cannot split any of the specified
Business Copies; none are in PAIR status; create Business Copy(s) of device group
"[%s1]" and restart the VCS service group "[%s2]".  Or set the force flag to disable
the Rolling Disaster Protection feature.



#

# "VCS only" messages

#

49004    Service group [%s1] will be disabled on system [%s2] in data center [%s3].

49005     Service group [%s1] will be disabled on all system in data center [%s2].

49006     Service group [%s1] will be disabled on all system in cluster.

49010     Parameter [%s1] not defined or not set; cannot disable service group for
data center [%s2].

49015     Agent monitoring function called with invalid argument.

49016     Resource type definition for the Cluster Extension XP resource is out of
date. Please restart the cluster server on all nodes to update the resource type
definitions.



##

## CLXRUN cluster-specific messages start at offset 50000 (CLXRUN_MSG_OFFSET)

##
```

```
59001     User did not confirm forceflag override.

59002     \nUsage: clxrun [-version] [-forceflag] <app_name> \n       -version
Displays CLX version\n       -forceflag  Force startup action\n       <app_name>
Application name configured\n                   in the user configuration file
(UCF.cfg).\n\n
```

C

Cluster Extension XP quorum service message catalog

```
#
# HP StorageWorks Cluster Extension Quorum Service Message File
#
#
# CLXQ service messages
#
600  CLX state unclear! Run clxqr.exe to check/repair and then
re-start.
601  Can not use WinSock 2.2 to get hostname.
602  Failed to get the hostname.
603  [%s1]: Quorum Filter-Service started on node [%s2] (id=[%s3])
======
604  Could not open handle to driver [%s1]. Will re-try after a while.
605  Could not open handle to driver [%s1] for three times. Exit.
606  HP CLX Quorum Filter Driver found, version [%s1]
607  Failed to obtain HP CLX Quorum Filter Driver version. Error code
[%s1]
608  Filter Timeout value is set to 8 seconds
609  Failed to set Filter Timeout value. Using default value (4
seconds)
611  Failed to set Filter Passthrough value. Use default
613  Service main loop, RECV_MSG failed. Error code [%s1]
629  Received WRITE_IO
631  No quorum device found
632  Do not have enough buffer space
633  One of the buffer parameters does not match internal tracking
634  There were no queued IOs or there is a recv_msg already
outstanding
635  Unknown error
636  Data size error, expected size/size got [%s1/%s2]
637  Sent back result
638  CLXQ SERVICE is stopped
639  Failed to send STOP_IO to the driver
640  Succeeded to send STOP_IO to the driver
641  SetServiceStatus failed
642  ClxQSvc lost connection with driver many times. Check array. From
now on, no warning/error messages will be logged.
```

```
643  ClxQSvc resumes connection with driver. Normal logging is
resumed.
700  Current GMT/UTC [%s1]
701  [%s1]: Quorum Filter-Service stopped ======
702  Received STOP request from SCM
703  Received PAUSE request from SCM
704  Received RESUME request from SCM
705  Received INTERROGATE request from SCM
706  Received SHUTDOWN request from SCM.
707  Received an unknown Control Code [%s1] from SCM. Ignore it
710  Waiting for cluster service to start...
711  Cluster service attempts to start
712  Detected cluster service status: [%s1]
713  Shutting down ClxQSvc...
714  User has specified the createsplitbrain option. There is a
potential for a split brain condition if the remote site is isolated
and running.
715  The wait time for cluster service startup is [%s1] seconds
716  The Cluster Decision Maker is running, therefore ClxQSvc will not
check the status of ClusSvc until the Cluster Decision Maker is
finished.
717  CA XP link is down, but either the Isolation Checker is already
running or the createsplitbrain flag has been set, so will not start
the Isolation Checker.
720  Disk [%s1] is not in SMPL state
721  Could not create thread to monitor: [%s1]
722  [%s1] disk pair status is [%s2]
723  HCT Testing is enabled.  There will NOT be checks to make sure
the Cluster Service is running.
#
# CLXQ library messages
#
800  Failed to get registry [%s1] value from Registry. Use default
value [%s2]
801  Failed to get registry [%s1] value from Registry.
802
803  Failed to get environment variable [%s1]
804  Failed to convert port name [%s1] to port number
805  Failed to convert GUID [%s1] to physical drive number
806  External Arbitration is enabled
807  External Arbitration is disabled
808  Active Directory Server's name/IP [%s1/%s2]
809  Active Directory is not available
810  Failed to swap quorum disk pair
811  S-VOL side node will form a cluster
812  Quorum state uncertain. Cluster will be shut down
813  Failed to check status of quorum disk pair
815  Control disk 1 pair (status disk) is not in PVOL_PAIR state
817  CA XP link for disk mirroring has been resumed
818  Quorum disk pair status is [%s1/%s2]
819  Quorum disk is not in PAIR. CA XP link may have failed!
820  CA XP link is up. Trying to re-create quorum disk pair
822  Failed to obtain locks (failed after many retries)
823  Failed to unlock local lock
824  Failed to swap Control disk 1 pair (status disk)
825  Failed to unlock global lock
838  Invalid port/tid/lun values: port/tid/lun values are not within
the range [%s1]
839  RMLib function attachcmddev() failed. Error [%s1]
840  RMLib function paircreate() failed. Error [%s1]
841  RMLib function detachcmddev() failed. Error [%s1]
842  RMLib function pairvolstat() failed. Error [%s1]
843  RMLib function pairsplit() with [%s1] option failed. Error [%s2]
844  Disks are not in pair
847  RMLib function pairsplit() using S_SWAPS failed. Error [%s1]
848  RMLib function pairresync() using R_SWAPS failed. Error [%s1]
849  Invalid drive number [%s1]
852  Function CreateFile() for open failed. Error code [%s1]
853  This device [%s1] is not a disk
860  Starting Cluster Decision Maker ...
861  Cluster Decision Maker started successfully
862  Failed to start Cluster Decision Maker. Error code [%s1]
```

```
863  External Arbitration disabled. No Cluster Decision Maker will be
started
864  Starting Isolation Checker ...
865  Isolation Checker started successfully
866  Failed to start Isolation Checker. Error code [%s1]
867  External Arbitration disabled. No Isolation Checker will be
started
868  Invalid node id
869  Failed to write local lock value to disk
870  Failed to read local lock value from disk
871  Failed to unlock local lock. Some node breaks in the lock
872  Function CreateFile() failed: open physical drive [%s1] for
checking device geometry failed
873  Function DeviceIO() failed: get physical drive %d disk geometry
failed
874  Function VirtualAlloc() failed. Error code [%s1]
875  Function VirtualFree() failed: decommit memory failed with error
code [%s1]
876  Function VirtualFree() failed: release memory failed with error
code [%s1]
877  Invalid sector number [%s1]
880  Function ReadFile() failed: read physical drive [%s1] failed.
Error code [%s2]
881  Function CreateFile() failed: open (for write) physical drive
[%s1] failed. Error code [%s2]
882  Function WriteFile() failed: write physical drive [%s1] failed.
Error code [%s2]
883  Function CreateFile() failed: open (for read) physical drive
[%s1] failed. Error code [%s2]
884  Failed to get disk [%s1]'s numldev
885  CA XP link is down. Let the P-VOL side obtain the global lock
886  Function getldev() failed with error [%s1]
887  CA XP link is down. Unlock succeeded.
888  Function clxq_PairStatus() failed
889  Function RegOpenKeyEx() to open key [%s1] failed. Error code
[%s2]
890  Function RegQueryValueEx() to query value [%s1] failed. Error
code [%s2]
891  Failed to open drive [%s1]. Error code [%s2]
892  Function IOCTL_DISK_GET_DRIVE_LAYOUT to get [%s1] layout failed.
Error code [%s2]
893  Function IOCTL_DISK_SET_DRIVE_LAYOUT to set [%s1] layout failed.
Error code [%s2]
894  Disk [%s1] is not in PAIR state
895  Failed to check disk [%s1] status
896  Init value on Local Lock disk is incorrect
897  Init check succeeded
899  Disk extent is too small (less than 512*15 bytes)
900  [%s1]
910 Function DeviceIoControl() to device [%s1] failed. Error code
[%s2]
911 Failed to connect to SCM. Error code [%s1]
912 Failed to open service [%s1]. Error code [%s2]
913 Failed to stop service [%s1]. Error code [%s2]
914 Succeeded to stop service [%s1]
915 Check pair [%s1] status failed
917 Global lock is too old (It has been held for more than [%s1]
seconds). Cleaning it ....
918 Cleaning global lock failed
#
# CDM messages
#
920 [%s1]: Cluster Decision Maker is starting ...
921 Unable to initilize WinSocket Library
922 Local hostname [%s1]
924 Wait for [%s1] seconds before consulting the external arbitrator
925 Function CreateProcess() failed. Error code [%s1]
926 Send request to external arbitrator failed. Retrying [%s1]
times...
928 Send request to external arbitrator failed.
929 Cluster is on the same side as this node
930 Failed to start cluster service to join the existing cluster
931 Succeeded in joining the existing cluster
```

```
932 Cluster is running on the remote/opposite side of this node. CLX
will not start the Cluster service on this node.
933 Do not know which side the cluster is
934 Failed to start cluster service to form a new cluster
935 Forming a new cluster succeeded
936 Either Active Directory is not available or failed to get external
arbitrator information from Active Directory
937 Failed to get external arbitrator information from either Active
Directory or local config file
938 Node names [%s1] and [%s2] are unknown
939 Cluster service is already up on local node
940 Failed to query service [%s1]. Error code [%s2]
941 Failed to start service [%s1]. Error code [%s2]
942 Cluster service started
943 Failed to get device [%s1]'s pair status
944 Failed to clean Control Device 1's status information
945 Openning socket [%s1]
946 Listen on socket failed. Error code [%s1]
947 Open stream socket failed. Error code [%s1]
948 Wait for reply on any socket
950 Reply's token is different from request's token [%s1]
952 Function recv() failed. Error code [%s1]
953 Function accept() failed. Error code [%s1]
954 Function receive() timed out after [%s1] seconds
955 Function select() failed. Error code [%s1]
957 Open local config file [%s1] failed
959 Failed to get port from service. Error code [%s1]. Use the default
value [%s2]
960 Failed to get service_name from config file
961 Failed to get arbiter_ip from config file
962 Failed to get cluster_name from config file
963 Failed to get cluster_ip from config file
965 Found no nodes from config file
967 Failed to get any address from config file
968 CDM: HP CLX Cluster Decision Maker was finished.
#
# IC messages
#
969 Failed to start HP CLX Isolation Checker
970 [%s1] : Isolation Checker is starting ...
971 Failed to get node location information from config file
972 Wait for [%s1] seconds before checking isolation
973 Connection to arbiter is still up. This side is not isolated. Do
nothing
974 Connection to arbiter is down. This side is isolated. Shut down
cluster
975 Failed to check the connection to arbiter
976 HP CLX Isolation Checker has determined that this node is NOT
isolated.
977 HP CLX Isolation Checker has determined that this node is
isolated.
978 Node name [%s1] is unknown
980 Failed to get node [%s1]'s state
981 Heartbeat link is down
982 Heartbeat link is still up
983 Node [%s1] is down
984 Node [%s1] is NOT down
985 Function GetClusterNodeState() failed. Error code [%s1]
986 Function CloseClusterNode() failed. Error code [%s1]
987 Function OpenClusterNode() failed. Error code [%s1]
988 Function CloseCluster() failed. Error code [%s1]
989 Function OpenCluster() failed. Error code [%s1]
990 Get cluster node state failed. Wait and retry ....
991 Failed to find node state after many retries
992 Process Send terminated with errors
995 Failed to stop cluster service on node [%s1]
996 Stopping my cluster service ...
997 Cluster service on node [%s1] is already stopped
998 Failed to change service [%s1] failure action. Error code [%s2]
999 Failed to change service [%s1] param online. Error code [%s2]
#
# External Arbitrator Service
#
```

```
1000 [%s1]: External Arbiter Service stopped =====
1001 [%s1]: External Arbiter is starting ...
1002 Failed to get service name and IP address from local config file
1003 Failed to create [%s1] thread. Error code [%s2]
1004 External Arbiter was started successfully
1005 Thread [%s1] was exited unexpectedly
1006 There are other errors with thread rather than thread exit. Error
code [%s1]
1007 External Arbiter was stopped unexpectedly
1008 Failed to bind to local address. Error code [%s1]
1009 Receiver is waiting for more buffer resource ...
1010 Receiver is waiting for requests ...
1015 Cluster [%s1] is up
1016 Quorum is on node [%s1]
1017 Cluster [%s1] is down
1018 Cluster [%s1] state is unknown
1019 Trying to re-connect ...
1020 Trying to re-connect failed many times
1021 Function OpenClusterGroup() failed. Error code [%s1]
1022 Function GetClusterGroupState() failed. Error code [%s1]
1028 Function CloseClusterGroup() failed.
1029 Function CloseCluster() failed.
1030 Wait and retry to check Quorum location
1031 Failed to locate cluster after many retries
1032 Failed to send back result using current IP/port. Try to send
using another IP and port numbers
1033 Failed to send back result using all available IP/ports
1034 Failed to get service_name from config file
1035 Failed to get arbiter_ip from config file
1036 Processing a PING request from cluster nodes ...
1037 Starting the External Arbitrator Monitor ...
1038 The External Arbitrator Monitor was started successfully
1039 Failed to start the External Arbitrator Monitor. Error code [%s1]
1040 The External Arbitrator is not configurated. No need to start its
Monitor
#
# External arbitrator monitor
#
1050 Checking Arbitrator status ...
1051 Arbitrator is up and functional
1052 Arbitrator is down
1053 Arbitrator Monitor failed to check Arbitrator status [%s1]
1054 Arbitrator is not configurated. No need to check its status
1055 Arbitrator Monitor was finished successfully
1056 Arbitrator received an invalid request [%s1]
1057 Checking if the cluster [%s1/%s2] has been checked recently in
the history log ...
1058 Arbitrator fails to receive messages many times. It is aborting
...
1059 Arbitrator waits and retries to receive a message
```

Quorum service Event Log messages

```
# Event log messages from quorum filter
001 Could not obtain SCSI inquiry information from device %2.
002  Driver could not obtain the serial number information of the
quorum disk. Please repair the %1 installation.
003  The QuorumSerialNumber registry parameter is invalid. Please
reinstall the %1 software.
004  HP CLX successfully inserted HP CLX filter driver above MS
Cluster Service cluster disk with serial number %2.
005  The HP CLX filter driver lost communication with HP CLX Quorum
Service. The CLX Quorum Service might have stopped. This cluster node
might not be able to become a quorum owner.
006  The MSCS quorum  disk has failed or was removed. HP CLX cannot
manage cluster quorum. Check IO path between this system and the HP
disk array!
007  Driver failed creating the parent object with error %2.
008  Driver failed creating the quorum filter object with error %2.
009  HP CLX failed to insert HP CLX filter driver above MS Cluster
Service cluster disk. Check if the quorum disk exists. Check IO path
between this system and the HP disk array!
#Event Log messages from ClxQSvc
256 HP CLX Quorum Service was started successfully.
For more information, see HP CLX Quorum log file %1.
257 HP CLX Quorum Service has been stopped.
For more information, see HP CLX Quorum log file %1.
258 For more detailed information, see HP CLX Quorum log file %1.
259 An error/warning occurred. For more detailed information, see HP
CLX Quorum log file %1.
260 HP CLX Quorum Service lost communication with HP CLX filter driver
after many recovery attempts. The service will try to re-establish
communication in the background, but any further logging to clxq.log
will be suspended. This cluster node will not be able to be a quorum
owner! Check IO path between this system and the HP disk array!
261 HP CLX Quorum Service resumed communication with HP CLX filter
driver. This cluster node is able to be a quorum owner again. Normal
logging resumes.
262 HP CLX Quorum Service has been paused.  Until the service is
resumed, there is potential for a split brain condition.
263 HP CLX Quorum Service has been resumed.  There is no longer
potential for a split brain condition.
264 HP CLX Cluster Decision Maker has been started. The Cluster
service will be shut down because of a potential partition of the
cluster (split brain syndrome). All communication paths will be
examined for possible partitioned clusters before the Cluster
Service will be started. !!! WARNING !!!!!! Please do NOT start the
Cluster service before the CLX Cluster Decision Maker has fininshed
!!!Please recover the CA XP link and the cluster heartbeat
connections.
265 HP CLX Cluster Decision Maker has determined that the Cluster
service is running in the remote data center. The Cluster service will
NOT be started because of a potential partition of the cluster (split
brain syndrome). Please recover the CA XP link and the cluster
heartbeat connections before you attempt to restart the Cluster
service.
266 HP CLX Cluster Decision Maker has determined that the Cluster
service is running in the local data center. The Cluster service will
join the local cluster. Please recover the CA XP link.
267 HP CLX Cluster Decision Maker has determined that the Cluster
service is not running in the local or remote data centers.  The
Cluster service will be started locally. Please recover the CA XP
link.
```

```
268 HP CLX Isolation Checker has been started. All communication paths
will be examined for possible partitioned clusters. Please recover
the CA XP link.
269 HP CLX Isolation Checker has determined that this node is
isolated. The Cluster service will be shut down because of a potential
partition of the cluster (split brain syndrome).
Please recover the CA XP link, the cluster heartbeat connections and
the arbitrator network connection before you attempt to restart the
Cluster service.
270 HP CLX Isolation Checker has determined that this node is NOT
isolated. Please recover the CA XP link.
```

Glossary

“backhoe” issue Simultaneous destruction of redundant communications links.

BC HP StorageWorks Business Copy XP

CA The HP StorageWorks Continuous Access XP software program that lets you create and maintain duplicate copies of the data store on a local disk array.

CLI Command-line interface.

CU Control unit.

DWDM Dense wavelength division multiplexing.

ESCON IBM Enterprise Systems Connectivity.

failover The transfer of control of an application or service from one node to another node after a failure.

fence level A level for selecting rejection of a write I/O request from the host according to the condition of mirroring consistency.

GUI Graphical user interface.

HACMP IBM High Availability Cluster Multi-Processing for AIX software.

heartbeat A periodic synchronization signal issued by cluster software or hardware to indicate that a node is an active member of the cluster.

LDEV Logical device. An LDEV is created when a RAID group is carved into pieces according to the selected host emulation mode. The number of resulting LDEVs depends on the selected emulation mode. The term LDEV is often used synonymously with the term volume.

MU Mirror unit. The MU number is assigned when you create a BC pair. The MU number can be specified in the RAID Manager XP horcmx.conf file, in the BC area of CommandView XP, or on the disk array remote console.

P-VOL The primary or main volume that contains the data to be copied.

primary site Data center location that owns the Cluster Group (quorum resource).

quorum In Microsoft Cluster service, a cluster resource that has been configured to control the cluster, maintaining essential cluster data and recovery information. In the event of a node failure, the quorum acts as a tie-breaker and is transferred to a surviving node to ensure that data remains consistent within the cluster.

secondary site Data center location with the mirror copy of the quorum disk pair.

“split brain” syndrome A state of data corruption can occur if a cluster is reformed as subclusters of nodes at each site, and each subcluster assumes authority, starting the same set of applications and modifying the same data.

SPOF Single point of failure.

S-VOL Secondary or remote volume. The copy volume that receives the data from the primary volume.

VCS VERITAS Cluster Server.

Index