High Availability Cluster/iX User's Manual

HP e3000 MPE/iX Computer Systems

Edition 1

Notice

The information contained in this document is subject to change without notice.

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

Hewlett-Packard assumes no responsibility for the use or reliability of its software on equipment that is not furnished by Hewlett-Packard.

This document contains proprietary information which is protected by copyright. All rights reserved. Reproduction, adaptation, or translation without prior written permission is prohibited, except as allowed under the copyright laws.

Acknowledgments

UNIX is a registered trademark of The Open Group.

1. High Availability Cluster/iX

1 High Availability Cluster/iX

This guide documents the High Availability Cluster/iX utility for the HP e3000 systems.

The Cluster/iX product allows multiple computers to physically connect to a single set of disks, which are configured using the Cluster/iX utility to control cluster volume sets. In the event that the owner computer encounters a failure, another computer in the cluster can take over ownership of the cluster volume set.

Once installed, the Cluster/iX utility provides continued access to data on high availability disk array products like the SureStore E Disk Array XP256, XP48, XP512, VA7100 and any other supported dual-ported arrays.

Product Description

The configuration consists of multiported arrays connected to multiple computers as shown in Figure 1-1, where each port has access to the drives on the other port. Any single computer in this topology has direct access to the cluster volume set but all the computers in the topology agree on a single owner. All access to the data goes through the computer that owns the volume set.

In the event that the owner computer encounters a failure, another computer in the cluster topology can take over ownership of the cluster volume set. The system volume set can not be configured as a cluster volume set.

Figure 1-1   Cluster/iX
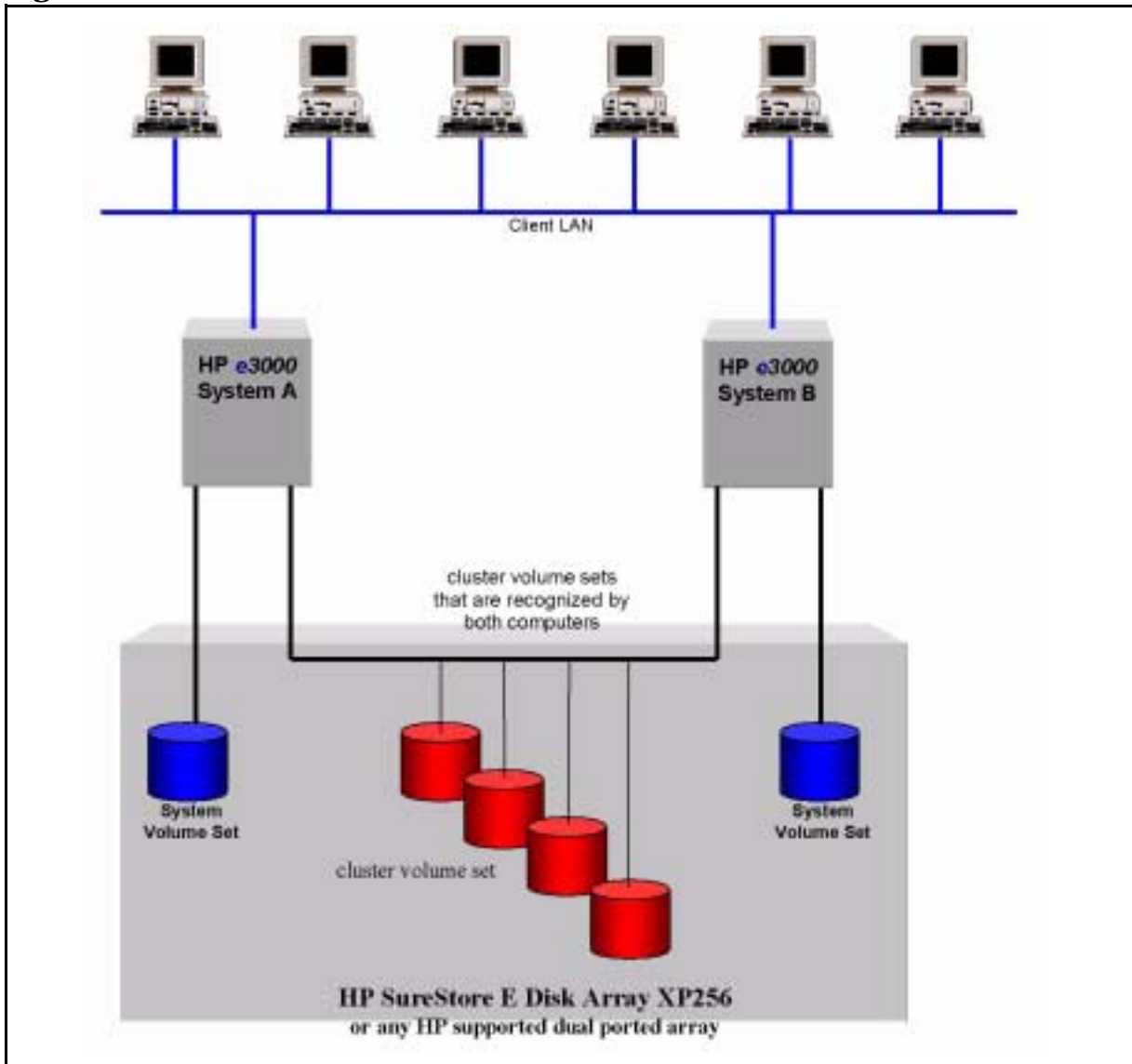
Benefits

The purpose of this product is to reduce planned or unplanned downtime caused by having a system outage while that system owns critical business data. The plan is to move or switch the data (user volume set) onto an up-and-working system where the users can then gain access to the critical data through the backup or secondary computer.

The Cluster/iX product will continue using the benefits of MPE's transaction management to rebuild system data structures and database facilities, like IMAGE, to ensure no or minimum loss of data.

Requirements

Hardware Requirements

The hardware consists of multiported arrays, like the XP family or 12H, connected to multiple computers where each port has access to the drives on the other port.

Limitations

- Clusteri/X does not migrate user entries, IP addresses nor user/group capabilites. The system implementor must script these features.

NOTE	The Heartbeat facility is no guarantee that the primary system is down. Care must be exercised in developing job scripts to cause a switchover to the secondary computer.

- Cluster/iX will not be supported on System Volume Sets.
- Only user volume sets are allowed to be cluster volume sets.
- Cluster/iX is not supported with path failover for the first release.
- Cluster/iX is not supported on Mirrored disks.
- Cluster/iX is not supported on split volume sets.
- Cluster/iX does not support data sharing.

Configuration

Once the Cluster/iX product is installed, perform the following steps to configure the systems and to convert the required user volume set to a cluster volume set:

1. Connect the primary and the secondary systems to a multiported volume set but only configure the ldevs for the primary system, leaving the configuration of the secondary system for later.

2. Verify that the MPE program **SXOutil** and the catalog SXCAT000 are properly installed in the group PUB.SYS for the primary and secondary systems, by using the LISTF command.

3. Using the cluster utility **SXOutil**, create (cmd: SXOMAKELABEL) the Cluster ID block on the primary system. This will convert your volume set into a cluster volume set. Please refer to the help facility within **SXOutil** for more details about the command.

4. VSCLOSE the newly created cluster volume set.

5. VSOPEN the cluster volume sets. Every cluster volume set can be owned by only one system belonging to the cluster at any time. The system that initiates creation of the cluster volume set will be the initial owner of the disk and is allowed to perform disk operations on these disks after mounting (VSOPENing).

6. Execute the DSTAT command to verify that the cluster volume sets are created properly. DSTAT will display all cluster volume sets as "**MASTER-SX**" & "**MEMBER-SX**".

7. Configure the ldevs on the secondary system by using SYSGEN or IOCONFIG. As each of the drives is recognized by the operating system, they will mount automatically as "**LONER**." Neither VSOPEN nor VSCLOSE should change the state from the secondary computer.

Secondary System

When the primary/host system experiences an outage, a switchover operation must be initiated. The currently active/secondary system needs to take ownership of all the cluster volume sets that the primary host/system owned. This is termed as "taking control or switchover" of a cluster volume set which is mounted as **loner** on the secondary system.

This is accomplished by the SXOSETLABEL command within **SXOutil** and followed by issuing the VSOPEN command.

The currently active/secondary system overrides the "Ownership Label" on the disks owned by the primary/host system, and assigns its "Ownership Label" on the disks owned by the primary/host system. Only then can the active/secondary system mount or VSOPEN the cluster volume set. On the completion of the switchover operation, these volumes will be mounted as master and member volumes, respectively, on the currently active/secondary system.

Once the secondary system has taken control of the cluster volume sets, reboot the primary system and *verify* that the cluster volume sets are automatically mounted as **loner** within the primary/host system.

Configuration Restrictions

It is ****IMPORTANT**** to note that when a "Ownership Label" is being created or being reset to the secondary active system, the other system should be shutdown, or must have done a VSCLOSE of the volume set.

Recreating a volume set and converting to a cluster volume set should be done only when the primary system can "see" the volumes. Failure to do so will cause the secondary to have stale data on its system data structure, forcing a reboot to refresh the structures.

SXOutil

This phase of the Cluster/iX product consists of a utility program called **SXOutil** that is used to create and initialize a master/member of the cluster volume set. In addition, it provides other commands that grant access to the cluster volume set from any computer within the cluster configuration.

NOTE	The user requires CV (create volumes) or SM (system manager) capability to execute the commands within **SXOutil**.

The following commands are currently available in **SXOutil**. For a complete description of the command and its usage please refer to the help facility within **SXOutil**.

SXOMAKELABEL

The SXOMAKELABEL command is used to create the cluster data structure, which include "Ownership label," associated audit trail information and history records. "Ownership Label" consists of the unique HPSUSAN number of the system and the Root UFID of the System Volume Set. This command will set the "Ownership Label" to that of the computer where the command was issued. The "Ownership Label" is assigned only to the **MASTER** volume of the volume set created and not the member volumes.

Creation of "Ownership Label" prevents two systems from trying to mount the same volume set simultaneously. Thus, disk operations are allowed only on disk volumes owned by the system.

NOTE	The operator is responsible for verifying that this computer is the only one that has access to the volume set until after the volume set has been created and the data structures built using this command. After these structures are built, the system administrator can then configure ldevs on the other computers that will be in the cluster configuration.

SXOSETLABEL

The SXOSETLABEL command is used to assign/switch a user defined "Ownership Label" on the master volume of the volume set. The value of the HPSUSAN and Root UFID of the ldev 1 is the default value for the "Ownership Label."

This command can be initiated from any system belonging to the cluster. The default value of the "Ownership Label" will be assigned only to cluster volumes previously created by any system on this cluster. The **MASTER** volume should be in **LONER** state to execute this command. This operation is also known as switchover.

SXOREMOVELABEL

The SXOREMOVELABEL command removes the "Ownership Label" from the cluster volume set, thus converting it into non-cluster volume set.

SXOutil also allows the execution of all CI commands and other utilities within this utility.

In addition to the above commands, **SXOutil** provides the following commands:

- Listredo

- Redo

- Help

- Log

- Verify

- Exit

SXOBEAT

The SXOBEAT command is used to increment or change the heartbeat counter on the cluster volume set for any SXOLISTEN commands that may be monitoring it.

When SXOBEAT is used in a job script, the job will behave as if it is hung at that location until the job is aborted, or if the SXOBEAT command detects that the "Ownership Label" has changed.

Extraordinary steps should be taken in the event that SXOBEAT detects that the ownership has changed. These steps should included the immediate cessation of all I/O to the volume set and immediate closure of the volume set or an immediate reboot of the system.

SXOLISTEN

The `SXOLISTEN` command is used to monitor or listen to the heartheat counter on the cluster volume set. The command will pause until such time as the heartbeat stops. When it defects that a heartbeat has not happened within the tineframe passed in as a parameter, the command will fall through and continue processing the job script.

The Listen job script should be designed to alert the system administrator and to not assume that because the heartbeat stopped that the system is down. Other tests should be part of the decision process before having a job script make the final decision of taking over a cluster volume set.

Examples of Beat and Listen Job Scripts

```
!Job beat,manager.sys
!SXOutil
Verify off
Sxobeat RES_DBSET 50 2
Exit

!comment If beat terminates Emergency close volumes
!comment because the other machine accidentally took over.
!VSCLOSE RES_DBSET;now
!TELLOP        Help me mister wizard.
!TELLOP        Log onto secondary system and verify data base integrity
!Eoj
```

```
!Job listen,manager.
!Comment This job will check every 2 seconds for a heartbeat.
!Comment If after 2 minutes no heartbeat detected, takeover.
!SXOutil
Verify off
SXOListen RES_DBSET 50 2 120
Exit
!comment      Heartbeat stopped
!comment      Take over volume set
!SXOutil
verify off
SXOsetlabel RES_DBSET 50
Exit
!VSOPEN RES_DBSET
!Stream IPMigrat.Job.Sys
!Stream Launch.App.Sys
!TELLOP        Ready to go
!stream Jbeat.job.sys
!Eoj
```

DSTAT

DSTAT has been changed to display cluster volume type as shown here. In this example, SXO_SET1 is a cluster volume set.

```
:dstat all
LDEV-TYPE    SATUS      VOLUME     VOLUME SET - GEN
----------------------------------------------------
1-C2247M1    MASTER     MEMBER1    MPEXL_SYSTEM_VOLUME_SET-0
2-C3010M1    MEMBER     MEMBER     MPEXL_SYSTEM_VOLUME_SET-0
30-C5447A    MASTER-SX  MASTER1    SXO_SET1-0
31-C5447A    MEMBER-SX  MEMBER1    SXO_SET1-0
32-C5447A    MEMBER-SX  MEMBER2    SXO_SET1-0
33-C5447A    MEMBER-SX  MEMBER3    SXO_SET1-0
34-C5447A    MEMBER-SX  MEMBER4    SXO_SET1-0
35-C5447A    MEMBER-SX  MEMBER5    SXO_SET1-0
36-C5447A    MEMBER-SX  MEMBER6    SXO_SET1-0
37-C5447A    MEMBER-SX  MEMBER7    SXO_SET1-0
```

Error Handling

In addition to the normal command parsing errors and illegal command errors from the utility **SXOutil**, the following new errors pertaining to the creation and manipulation of the cluster block, are introduced:

MESSAGE: Cluster label exists already, cannot create another one

CAUSE: This error is returned while executing the command SXOMAKELABEL. This error indicates that an initial Cluster Ownership ID block has already been created by invoking the command SXOMAKELABEL.

ACTION: You only can reset this label to the current system by invoking the command SXSETLABEL. If there is a need to re-create the Cluster ID block, then issue the SXOREMOVELABEL command first.

MESSAGE: Volume mounted in invalid state for desired operation

CAUSE: The command failed because the volume is mounted in a wrong state.

ACTION: The volume should be in MASTER state in order to create (SXOMAKELABEL) the Cluster Ownership ID block as well as to purge (SXOREMOVELABEL) it. It should be in **LONER** state to reset (SXOSETLABEL) to the current system.

MESSAGE: Volume not a member of the set being initialized

CAUSE: There is a mismatch between the setname and the ldev being passed to the SXOMAKELABEL or SXOSETLABEL command.

ACTION: Cluster Ownership ID block can be created or reset only on the master volume of a volume set.

MESSAGE: Cluster ID is invalid for system volume set

CAUSE: Cluster Ownership ID block is valid only for user volume sets.

MESSAGE: Cannot RESET because no Cluster label is found

CAUSE: Cluster ID block needed to be created before it can be reset to the current system.

ACTION: Issue the command SXMAKELABEL before issuing SXOSETLABEL command.

MESSAGE: Cannot purge. Cluster label not initialized

CAUSE: Just a warning to indicate that the volume set is not switchable. Issue SXOMAKELABEL command to create the Cluster ID block before accessing the block using other commands.

MESSAGE: Insufficient capabilities to execute this command.

CAUSE: User does not have proper capability to execute this command.

ACTION: User requires either CV or SM capability to execute Cluster/iX commands.