Improving Your HP e3000 Availability

Walter McCullough

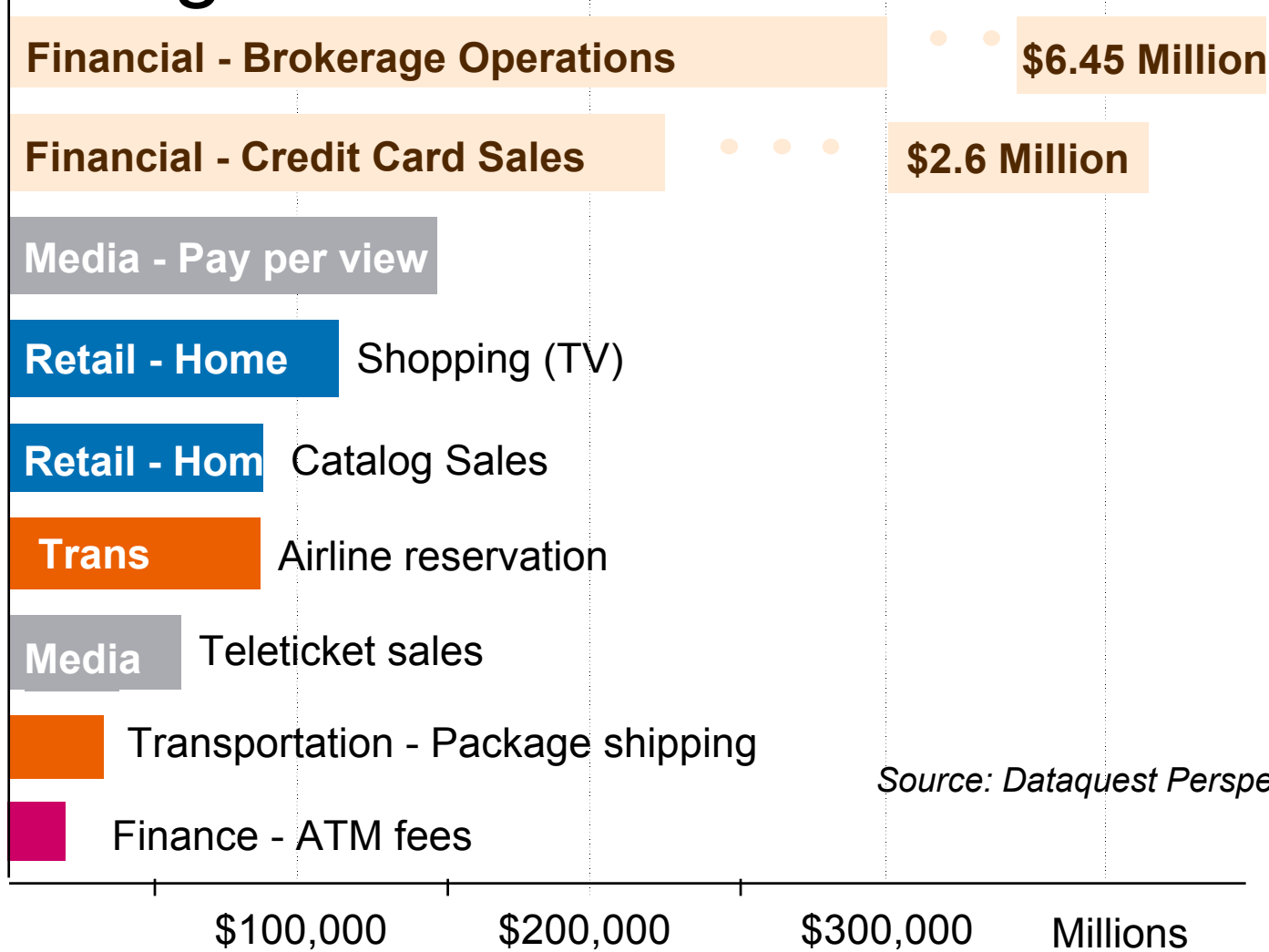MPE/iX High Availability R&D Architect

Hewlett-Packard

Goals

- Pass on knowledge to start your own availability improvement initiative

- Discuss Cluster/iX vs HAFO

- Go through a HAFO implementation

- Make sure you know where to get HP e3000 HA and storage related information

Agenda

- Cost of Downtime

- Reasons for Downtime

- Gathering Metrics

- High Availability Pyramid

- Types of Outages

- Typical Cluster/iX Environment

- Setting Up a FailOver/iX Environment

- Recover From an Event

- Q&A

Average Cost Per Hour of Downtime

Financial - Brokerage Operations   **$6.45 Million**

Financial - Credit Card Sales   **$2.6 Million**

Media - Pay per view

Retail - Home Shopping (TV)

Retail - Hom Catalog Sales

Trans Airline reservation

Media Teleticket sales

Transportation - Package shipping

Source: Dataquest Perspective, Sept. 1996

Finance - ATM fees

$100,000 $200,000 $300,000   Millions

Save What's Important First

- Create realistic outage scenarios

- Create a list of activities (and data)

- Prioritize that list based on customer impact and business impact

Survey Question

- Have you separated your applications/processes that can be offline for:

   A - One Hour

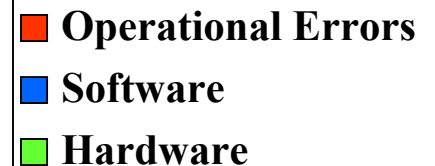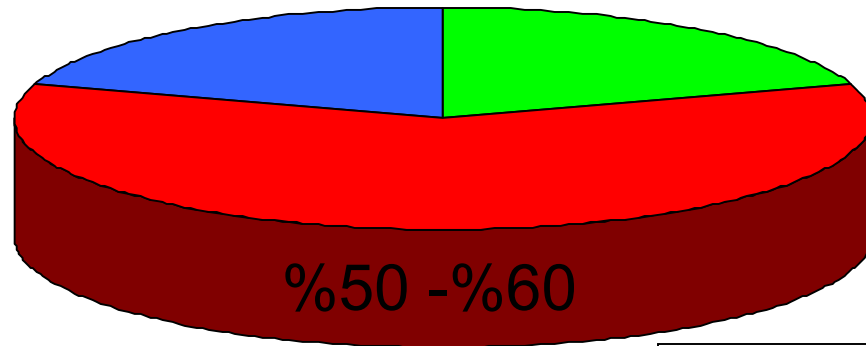   B - 4 Hours

   C - 1 Day

   D - 1 Week

Agenda

- Cost of Downtime

- **Reasons for Downtime**

- Gathering Metrics

- High Availability Pyramid

- Types of Outages

- Typical Cluster/iX Environment

- Setting Up a FailOver/iX Environment

- Recover From an Event

- Q&A

Components of Availability

- Technology

- Processes, practices and training

- Proper and adequate support agreements

Reasons for Downtime

%50 -%60

- ☐ **Operational Errors**
- ☐ **Software**
- ☐ **Hardware**

Agenda

- Cost of Downtime

- Reasons for Downtime

- **Gathering Metrics**

- High Availability Pyramid

- Types of Outages

- Typical Cluster/iX Environment

- Setting Up a FailOver/iX Environment

- Recover From an Event

- Q&A

Gather Metrics

- Where does it hurt and how long

- Helps reduce subjective nature of pain

- Focuses on solutions that make real improvements to HA

high availability
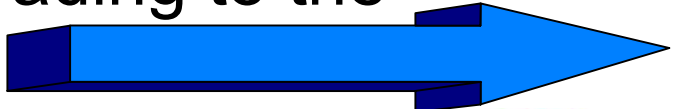
Annualized System Availability
Unplanned Events

Unplanned Event	Frequency of events experienced per year	# HRS down per event	Recovery time to ":" per year	Recovery time to application prompt
O/S Failures or Hangs				
SPU Hardware				
Networking				
Operator Error				
Total Unplanned Downtime				

Availability Needs
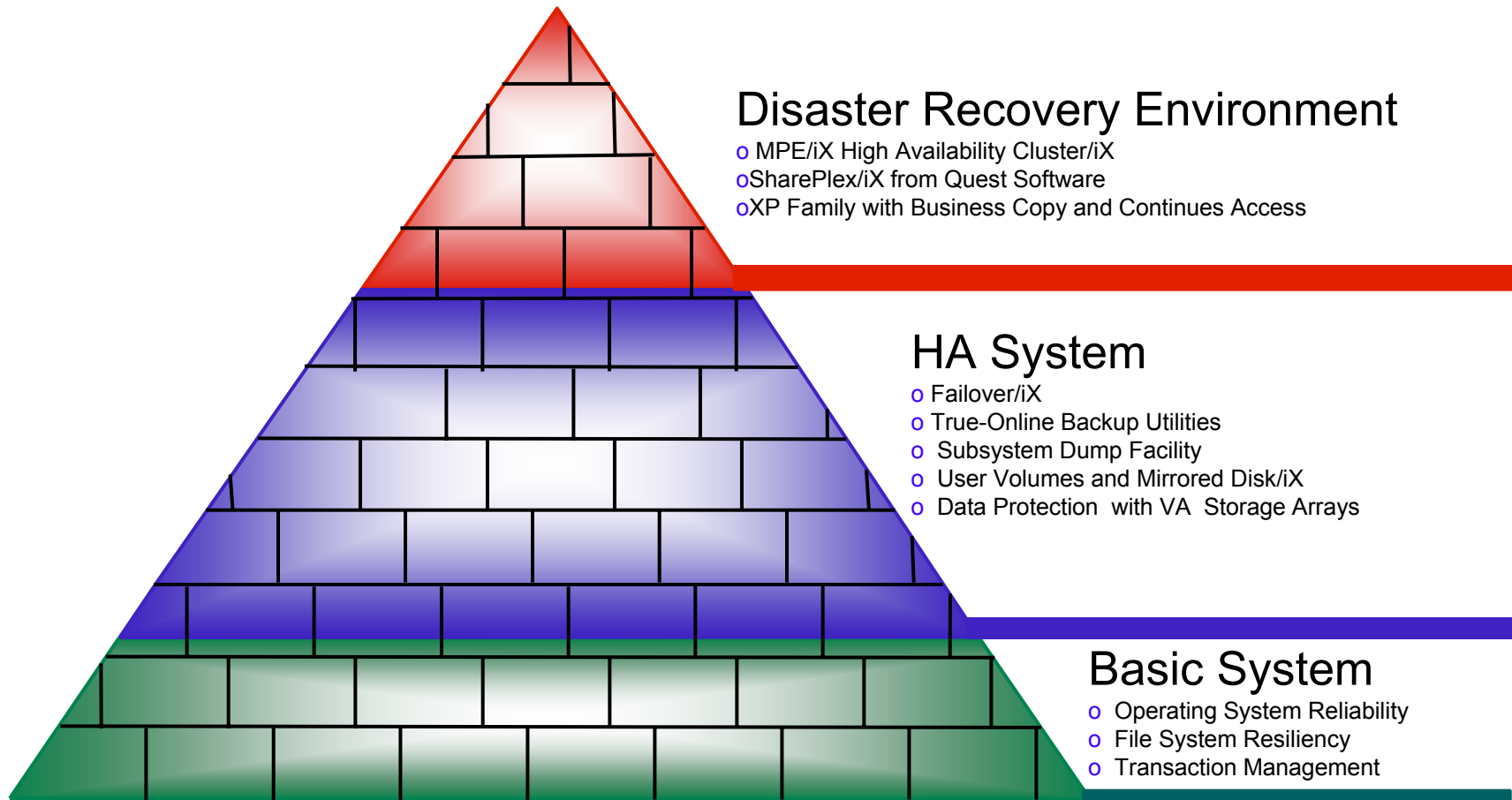## **Do I need the next level?**
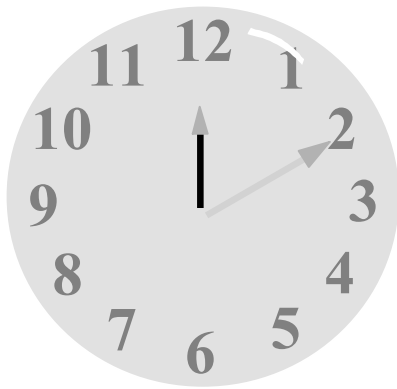
- Is availability necessary for meeting your operational requirements?

- Is your cost of downtime prohibitively expensive?

- Are your availability goals higher than what you are actually achieving?

If you answered **yes** to any of these questions, consider upgrading to the next layer of availability.

Agenda

- Cost of Downtime

- Reasons for Downtime

- Gathering Metrics

- **High Availability Pyramid**

- Types of Outages

- Typical Cluster/iX Environment

- Setting Up a FailOver/iX Environment

- Recover From an Event

- Q&A

High Availability Pyramid

Disaster Recovery Environment

o MPE/iX High Availability Cluster/iX
oSharePlex/iX from Quest Software
oXP Family with Business Copy and Continues Access

HA System

o Failover/iX
o True-Online Backup Utilities
o Subsystem Dump Facility
o User Volumes and Mirrored Disk/iX
o Data Protection with VA Storage Arrays

Basic System

o Operating System Reliability
o File System Resiliency
o Transaction Management

Types of Outages

- Disk media failure

- Controller failure

- Wire crimping

- FC Switch failure

- HBA failure

- Software/application failure

- Operating system outage

- Site power outage

- Catastrophic equipment failure

- Disaster site

Cluster/iX Benefits and Costs

Important Caveat: Cluster/iX + HAFO = Unsupported

Benefits

- Geared for user requiring very limited downtime
  (1 to 10 minutes ie can't wait to reboot)

- Geared for users that have some control over their application (willing to make changes to their process)
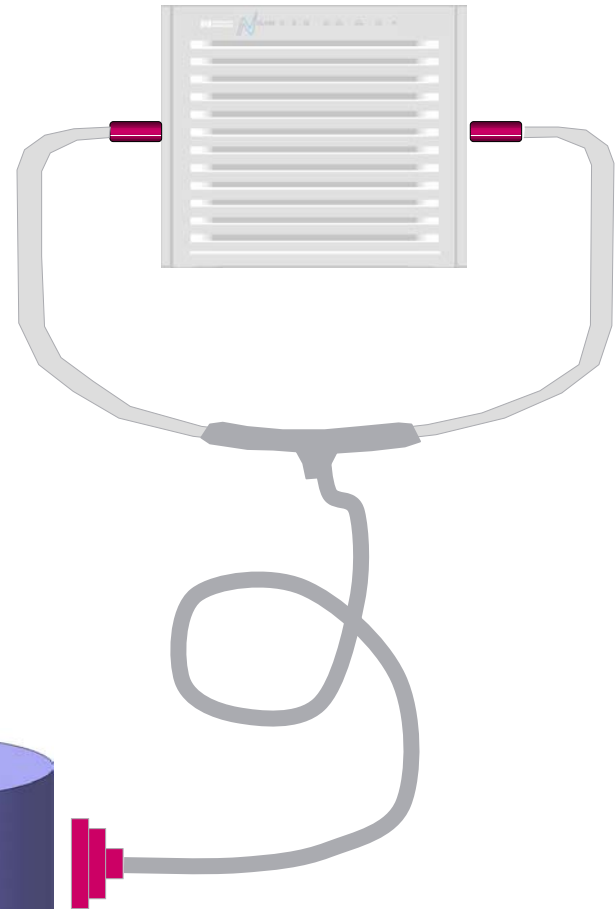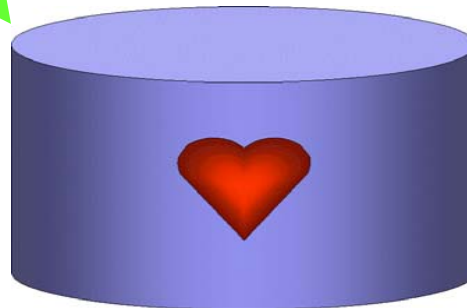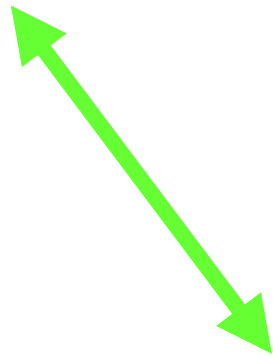
- Want single copy of data

Costs

- Requires 2 or more servers

- Customized scripting required

- Higher level of training involved to do the right things

Cluster/iX Solves

- Disk media failure (trick question)

- Controller failure

- Severed cable

- FC Switch failure

- HBA failure

- Software/application failure

- Operating system outage

★Site power outage
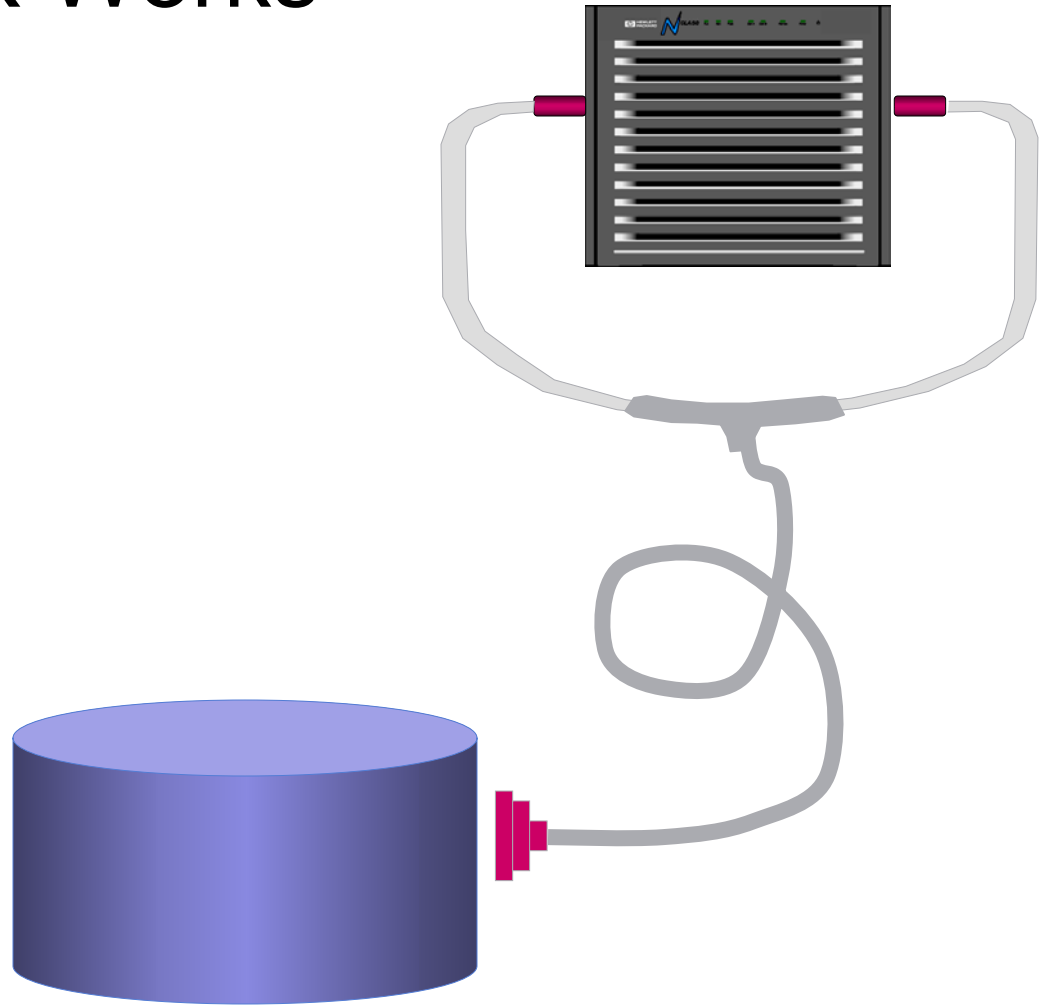
★Catastrophic equipment failure

★Disaster site

Requires★
Continuous Access XP

How Cluster/iX Works

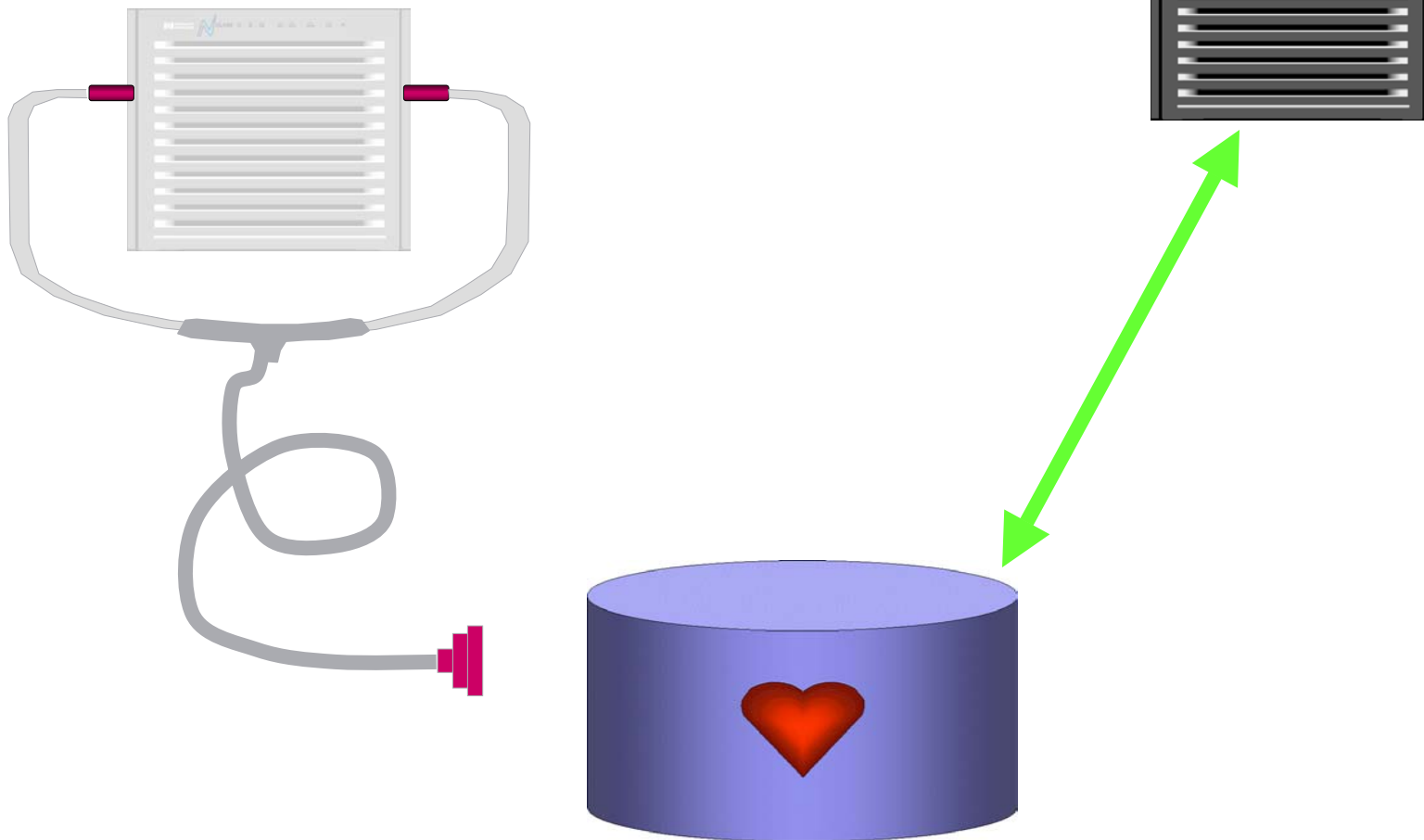# How Cluster/iX Works

How Cluster/iX Works

Physical Cluster/iX Environment
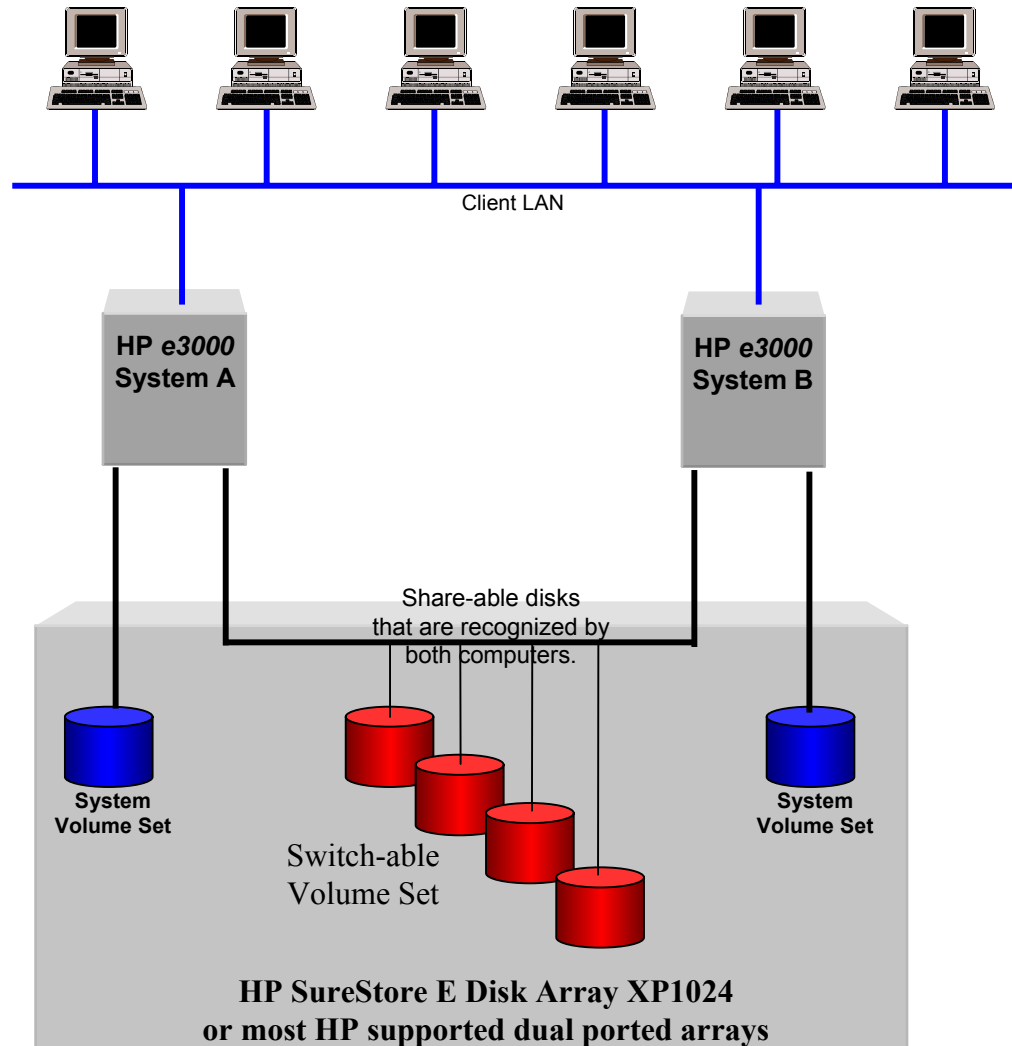
Client LAN

HP *e3000* System A

HP *e3000* System B

Share-able disks that are recognized by both computers.

System Volume Set

System Volume Set

Switch-able Volume Set

HP SureStore E Disk Array XP1024 or most HP supported dual ported arrays

Agenda

- Cost of Downtime

- Reasons for Downtime

- Gathering Metrics

- High Availability Pyramid

- Types of Outages

- Typical Cluster/iX Environment

- **Setting Up a FailOver/iX Environment**

- Recover From an Event

- Q&A

Failover/iX Solves

- Disk media failure (trick question)

- Controller failure

- Severed cable

- FC Switch failure

- HBA failure

- Software/application failure

- Operating system outage

- Site power outage

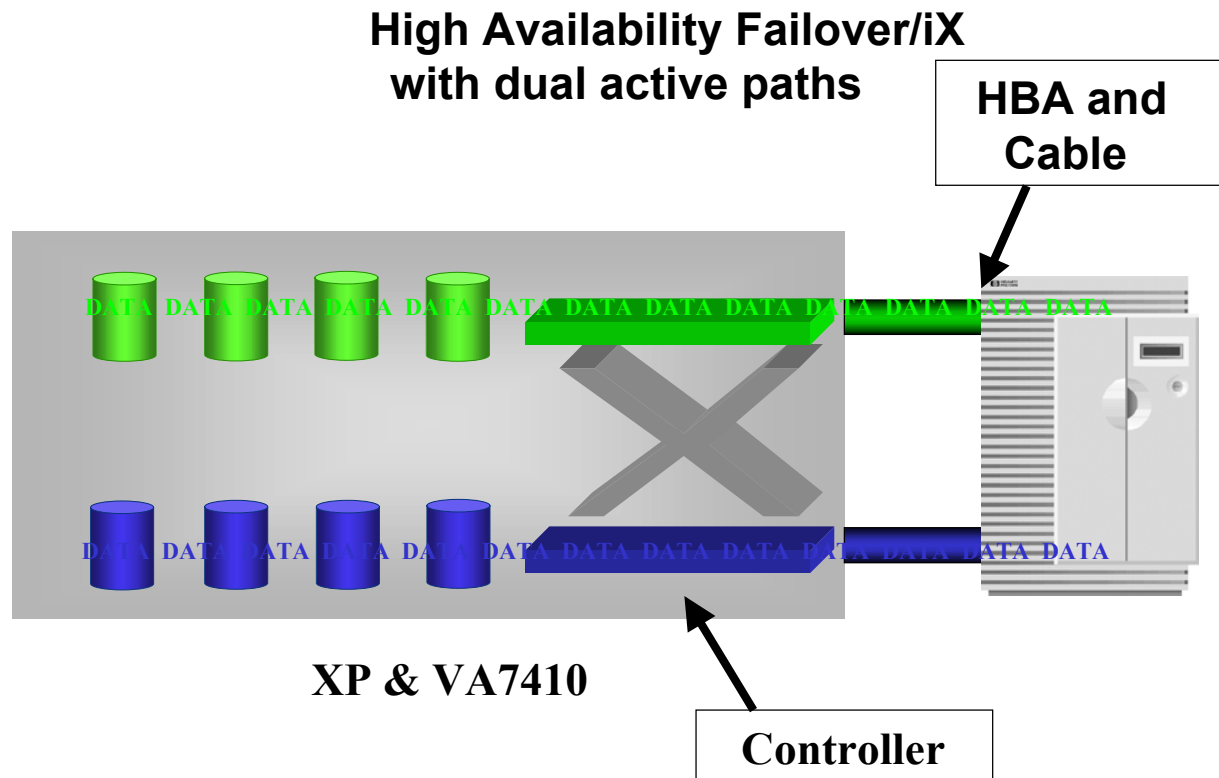- Catastrophic equipment failure

- Disaster site

Problem Description

**High Availability Failover/iX
with dual active paths**

HBA and
Cable

DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA

DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA

XP & VA7410

Controller

Problem Description

**High Availability Failover/iX
with dual active paths**

HBA and
Cable

DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA
DATA DATA DATA DATA
DATA DATA DATA
DATA DATA DATA DATA DATA DATA

XP & VA7410

Controller

Problem Description

SCSI HAFO (XP256 Only)

~		-
No Response	**Errors**	**Analog Anomalies**
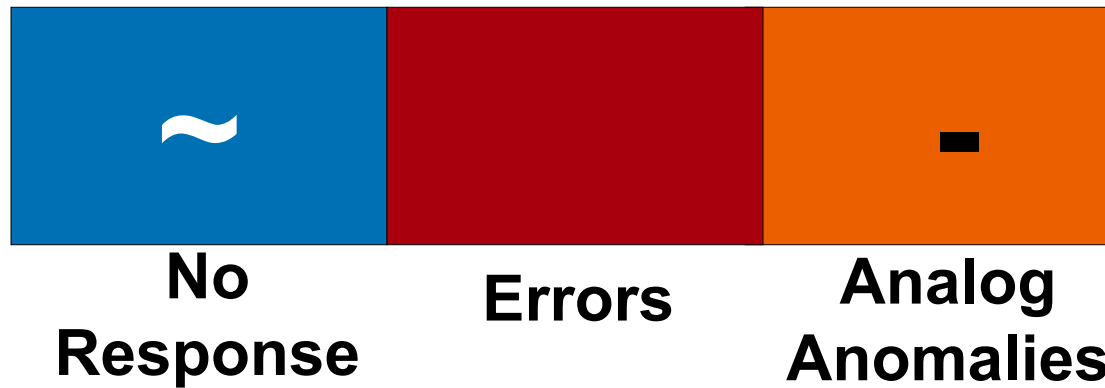
Problem Description

SCSI HAFO

(+Fabric Router solution)

No Response	Errors	Analog Anomalies
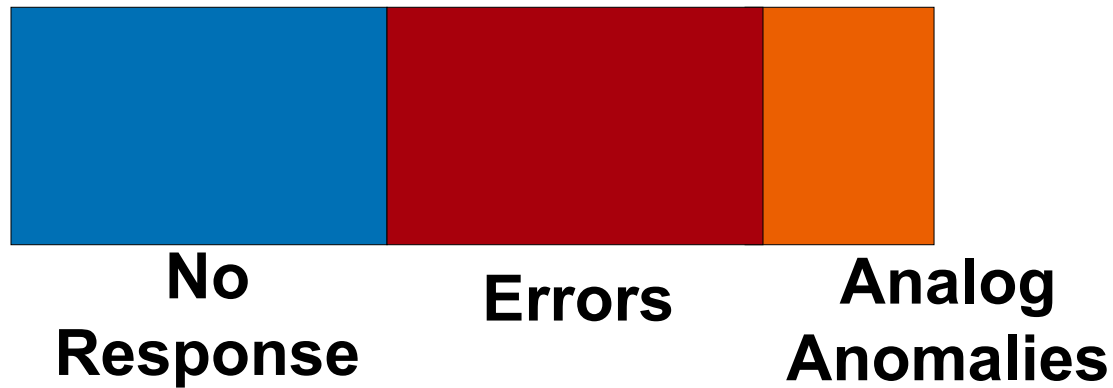
Problem Description

Native FC HAFO

No Response   **Errors**   **Analog Anomalies**
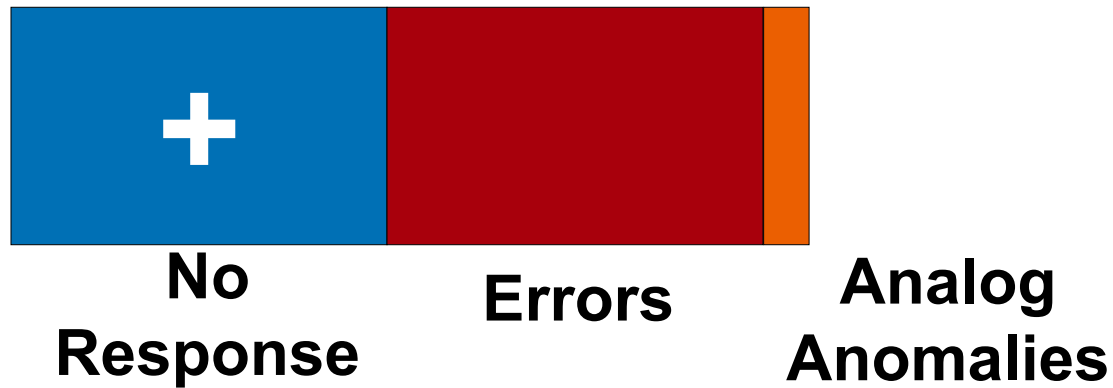
Software/Hardware Caveats

- HAFO and Cluster/iX are not supported (together)

- All Ldevs (luns) on a bus must be configured as HAFO protected

- All HAFO protected Ldevs must use similar connection strategy (switches and paths)

- HAFO is not a fault-tolerant solution (unplanned outages are converted to planned outages to fully recover)

- HAFO is dependent on performance expectations (false failovers)
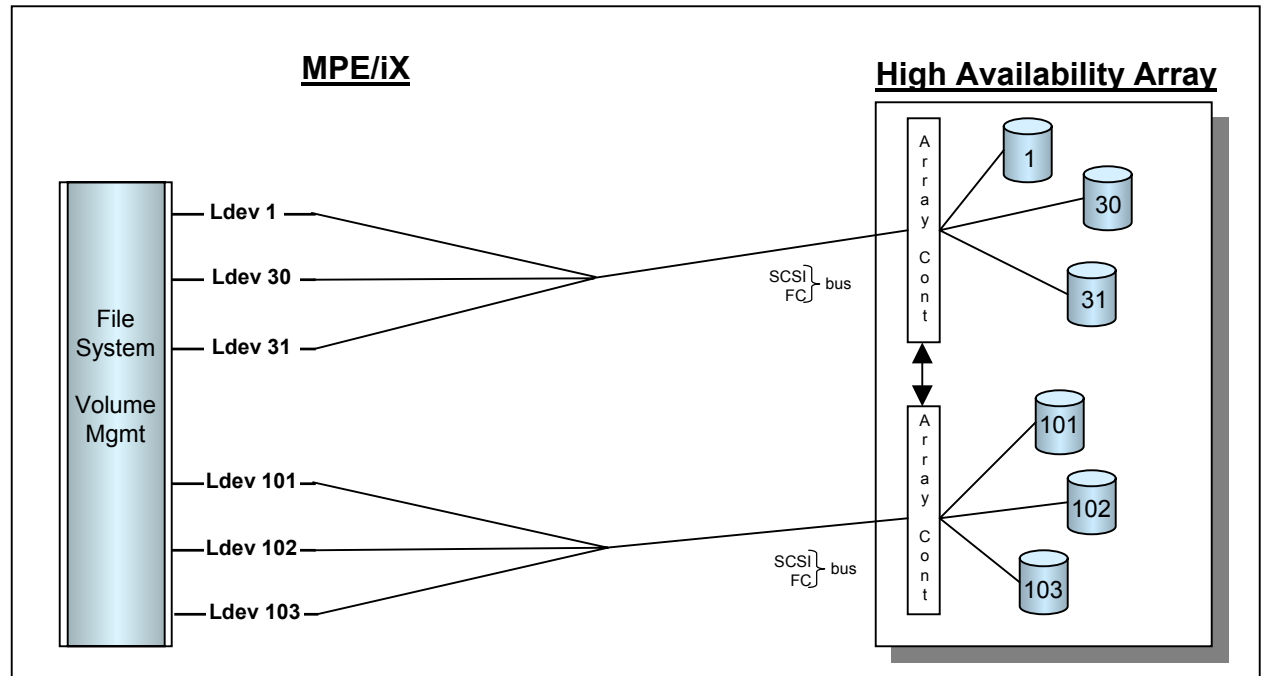
- Plan plan plan….

HAFO? Is it for you?

- HAFO is not the magic panacea of HA

- HAFO adds complexity to the operation of a system

- Unless there is a good understanding of the I/O characteristics of the system, you will introduce false failovers and reduce the I/O throughput of MPE

Configure the storage array

Create the Luns on the XP so that they are visible to MPE from both ports

Use Mapper or FCSCAN –h to verify Lun connection

MPE/iX

High Availability Array

File System

Volume Mgmt

Ldev 1

Ldev 30

Ldev 31

Ldev 101

Ldev 102

Ldev 103

SCSI FC bus

SCSI FC bus

Array Cont

Array Cont

1

30

31

101

102

103

Assigning Luns to Ports

Port address assigned by the path taken
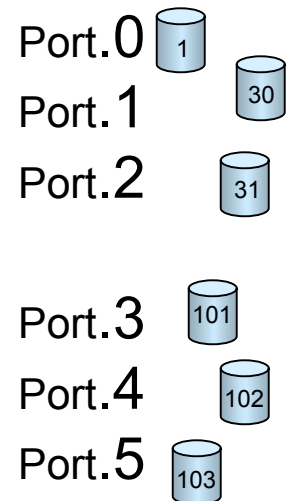
Example:

Primary Port ID is 36
Secondary Port ID is 28

To generate a path to the Ldev 1 take Port + Lun addr = 36.0

An alternate path to Ldev 1 is 28.0

Lun Addrs

Port.0   1
Port.1   30
Port.2   31

Port.3   101
Port.4   102
Port.5   103

Configure Ldevs

Using Sysgen

Create a configuration for Ldev 1-31 on path or port 36.

And configure Ldev 101-103 on path or port 28.

Example:

```
io>ad path=0/0.36.0 Ldev=1     id=HPDARRAY
io>ad path=0/0.36.1 Ldev=30    id=HPDARRAY
io>ad path=0/0.36.2 Ldev=31    id=HPDARRAY
io>ad path=0/0.28.3 Ldev=101   id=HPDARRAY
io>ad path=0/0.28.4 Ldev=102   id=HPDARRAY
io>ad path=0/0.28.5 Ldev=103   id=HPDARRAY
```

Keep and then reboot

Path 36

1

30

31

Path 28
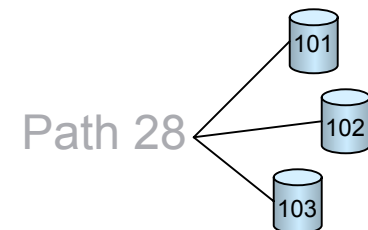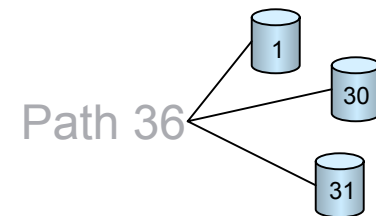
101

102

103
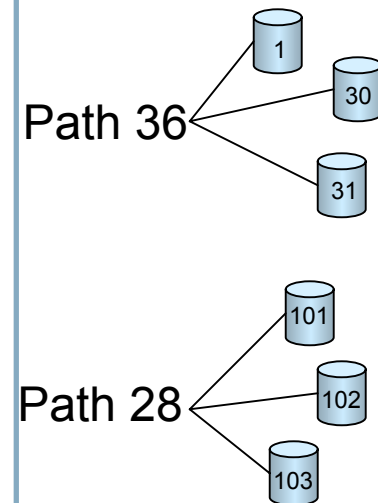
Protecting Ldevs with HAFO

Using Sysgen

Enter the HA menu and issue:

```
ha>ad 1    0/0.36.0 0/0.28
ha>ad 30   0/0.36.1 0/0.28
ha>ad 31   0/0.36.2 0/0.28
ha>ad 101 0/0.28.3 0/0.36
ha>ad 102 0/0.28.4 0/0.36
ha>ad 103 0/0.28.5 0/0.36
```

Hold and keep

Return to HA menu and issue DOHA

Path 36

1

30

31

Path 28

101

102

103

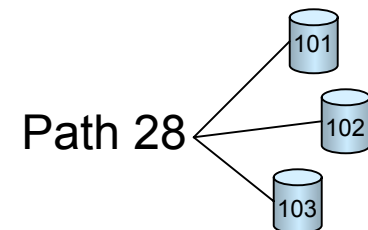# Creating User Volumes or Adding members to the system vol set

Using Volutil:

User Volumes

- Use the Newset and Newvol commands to create you user volume set.

- Then VSCLOSE and VSOPEN the volume set before using.

Adding to the System Volume Set

- Use the Newvol command.

- Shutdown and reboot to add HAFO protection

Path 36

1

30

31

Path 28

101

102

103
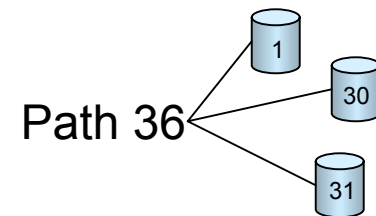
Important HAFO Commands

ADDCONF

addconf (ad) <Ldev> <path> <altpath> <timeout>

Example:

> **ha>ad 450 0/6/2/1.3.3 0/6/2/0 True**

New Feature

Timeout parm defaults to true. This parm allows for storage and server configurations where very poor I/O performance has been identified as the cause of false failovers. Turning this to false disable the timeout detection but increases the risk of encountering other undetectable timer related hangs/problems

Important HAFO Commands

LISTCONF

```
ha>LISTCONF


 Ldev      Primary Path           Alternate Path      Timeout

 =====  ====================   ===================   =======

   350 0/4/0/0.70954.23        0/6/0/0.73289         True

   351 0/4/0/0.70954.24        0/6/0/0.73289         True

   352 0/6/0/0.73289.25        0/4/0/0.70954         False

   353 0/6/0/0.73289.26        0/4/0/0.70954         False

   450 0/6/2/1.3.3             0/6/2/0               True

   451 0/6/2/1.3.4             0/6/2/0               True

   452 0/6/2/0.3.5             0/6/2/1               False

   453 0/6/2/0.3.6             0/6/2/1               False
```

Important HAFO Commands

DOHA

```
ha> doha

Start of validation for all HAFO configured devices.

=====================================================

VALIDATING ** Ldev: 50 Pri path: 8.15.0 Alt path: 48

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Ldev 50 configuration Validated Successfully

VALIDATING ** Ldev: 51 Pri path: 8.15.1 Alt path: 48

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Ldev 51 configuration Validated Successfully

End of validation for all HAFO configured devices.

=====================================================
```

Important HAFO Commands

GONEXT

`ha> Go <Ldev>`

After the problem has been repaired, issue the GoNext command to put the Ldevs back on their primary paths.

This command should not be used to cause the Ldev to go back to a known bad path. To do so may result in a reboot to reinitialize the MPE I/O configuration.

Agenda

- Cost of Downtime

- Reasons for Downtime

- Gathering Metrics

- High Availability Pyramid

- Types of Outages

- Typical Cluster/iX Environment

- Setting Up a FailOver/iX Environment

- Recover From an Event

- Q&A

HAFO Event

```
HIGH AVAILABILITY FAILOVER IS STARTED FOR Ldev# IN DISK ARRAY. NO DATA LOSS
OR CORRUPTION. SYSTEM OPERATION WILL CONTINUE.PLEASE PLACE SERVICE CALL SOON.

ACKNOWLEDGE HAFO FAILOVER IN DISK ARRAY FOR Ldev# (Y/N)?
```

Reply to the message

```
:HASTAT
 High Availability Failover Device Status

 Ldev    Primary Path           Alternate Path      Pri. Status      Alt. Status
 =====  ====================   ==================   ==============   ===============

  350 0/4/0/0.70954.23        0/6/0/0.73289        Array Failure    Ready
  351 0/4/0/0.70954.24        0/6/0/0.73289        Array Failure    Ready
  352 0/6/0/0.73289.25        0/4/0/0.70954        Ready            Validated
  353 0/6/0/0.73289.26        0/4/0/0.70954        Ready            Validated
  450 0/6/2/1.3.3             0/6/2/0              Ready            Validated
  451 0/6/2/1.3.4             0/6/2/0              Timeout/No Reply Ready
  452 0/6/2/0.3.5             0/6/2/1              Ready            Validated
  453 0/6/2/0.3.6             0/6/2/1              Ready            Validated
```
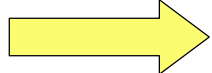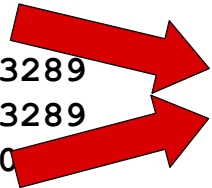
HAFO Event

- Ldev 350 and 351 have encountered an array error and have switched over successfully.

- Ldev 451 switched over because of an I/O timeout.

Reply to the message

```
:HASTAT
 High Availability Failover Device Status


 Ldev    Primary Path            Alternate Path       Pri. Status       Alt. Status
 =====  ====================    ==================   ==============   ===============

  350  0/4/0/0.70954.23        0/6/0/0.73289        Array Failure     Ready
  351  0/4/0/0.70954.24        0/6/0/0.73289        Array Failure     Ready
  352  0/6/0/0.73289.25        0/4/0/0.70           Ready             Validated
  353  0/6/0/0.73289.26        0/4/0/0.70954        Ready             Validated
  450  0/6/2/1.3.3             0/6/2/0              Ready             Validated
  451  0/6/2/1.3.4             0/6/2/0              Timeout/No Reply  Ready
  452  0/6/2/0.3.5             0/6/2/1              Ready             Validated
  453  0/6/2/0.3.6             0/6/2/1              Ready             Validated
```

Recovering from a HAFO Event

Array Failure Error

- This is a failure in the path of the Ldev and could be either the HBA or array controller or any component in between. Diagnose this problem as you would any other hardware component by collecting system and diagnostic log information.

  Only after repairing the part should you use the GoNext command.

- Ldev 350 and 351 have encountered an array error and have switched over successfully.

- Ldev 451 switched over because of an I/O timeout.

Recovering from a HAFO Event

Timeout Failover

- Treat this as if it was a hardware failure. Collect system and diagnostic log information. This information along with performance data is needed to prove that the I/O timeout is due only to the fact that the storage array can't keep up with the I/O load of MPE and is not another cause masquerading as Timeout.

  Only after proper diagnostic should you attempt to delete Ldev 451 (DELCONF command) and add it back into HAFO using the timeout parameter set to FALSE. Remember, setting the parm to False decreases HAFO protection.

- Ldev 350 and 351 have encountered an array error and have switched over successfully.

- Ldev 451 switched over because of an I/O timeout.

Important websites

MPE HA Website

http://jazz.external.hp.com/mpeha/

HAFO Documentation

http://www.docs.hp.com/mpeix/pdf/32650-90911.pdf

Cluster/iX Manual

http://www.docs.hp.com/mpeix/pdf/B9480-90001.pdf

MPE HA & Storage Matrix

http://jazz.external.hp.com/mpeha/ha_Xmatrix111703.pdf