



HACMP – High Availability

Introduction Presentation
February 2007

Introduction

- Scope
- HACMP – Concepts
- HACMP – Cluster Topologies
 - NFS
 - Cascading – 1 way
 - Cascading – 2 way
 - Rotating
 - Concurrent
- HACMP – Cluster Resources
- Mt Xia implementation of HACMP

Scope

- Scope
 - This presentation covers HACMP standards, design and concepts in the CGE environment.
 - This presentation is not designed as a tutorial on HACMP operations.

HACMP - Concepts

- High Availability Clustered Multi-Processing
 - The masking or elimination of planned or unplanned downtime
 - The elimination of single points of failure (SPOFs)
 - Computers
 - Disks
 - Networks
 - etc
 - Fault resilience, but not fault tolerance
 - A fault tolerant environment has no service interruption
 - A high availability environment has minimal service interruption

HACMP - SPOFs

- Identify and eliminate SPOF's by:
 - Node - Using multiple nodes
 - Power source - Using multiple circuits or UPS's
 - Network adapter - Using redundant network adapters
 - Network - Using multiple networks to connect nodes
 - TCP/IP subsystem - Using serial networks
 - Disk adapter - Using redundant disk adapters
 - Controller - Using redundant disk controllers
 - Disk - Using redundant hardware and disk mirroring
 - Application - Assigning a node for application takeover

HACMP Resource Groups

- Resource Groups are a set of cluster resources handled as one unit by HACMP, configured by the HACMP administrator.
- Attributes of a Resource Group
 - startup, fail-over and fall-back policies
 - settling time and the distribution policy
 - delayed fall-back timer that affects fall-back.
- Resource groups can be concurrent and non-concurrent.
- Resource Groups are used to coordinate the fail-over of:
 - Network Addresses
 - Disks
 - NFS mounts
 - Applications

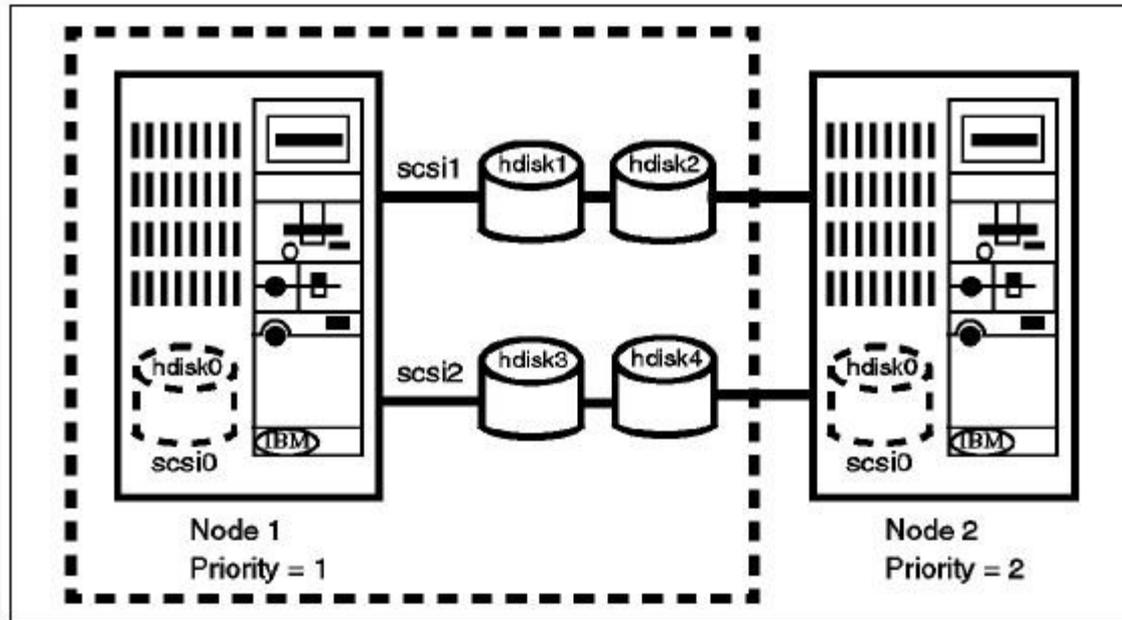
HACMP – Resource Group Fail-over

- Mechanisms implemented by HACMP for resource group fail-over :
 - Cascading
 - The active node with the highest priority controls the resource group
 - Concurrent
 - All active nodes have access to the resource group
 - Rotating
 - The node with the rotating resource group's associated service IP address controls the resource group

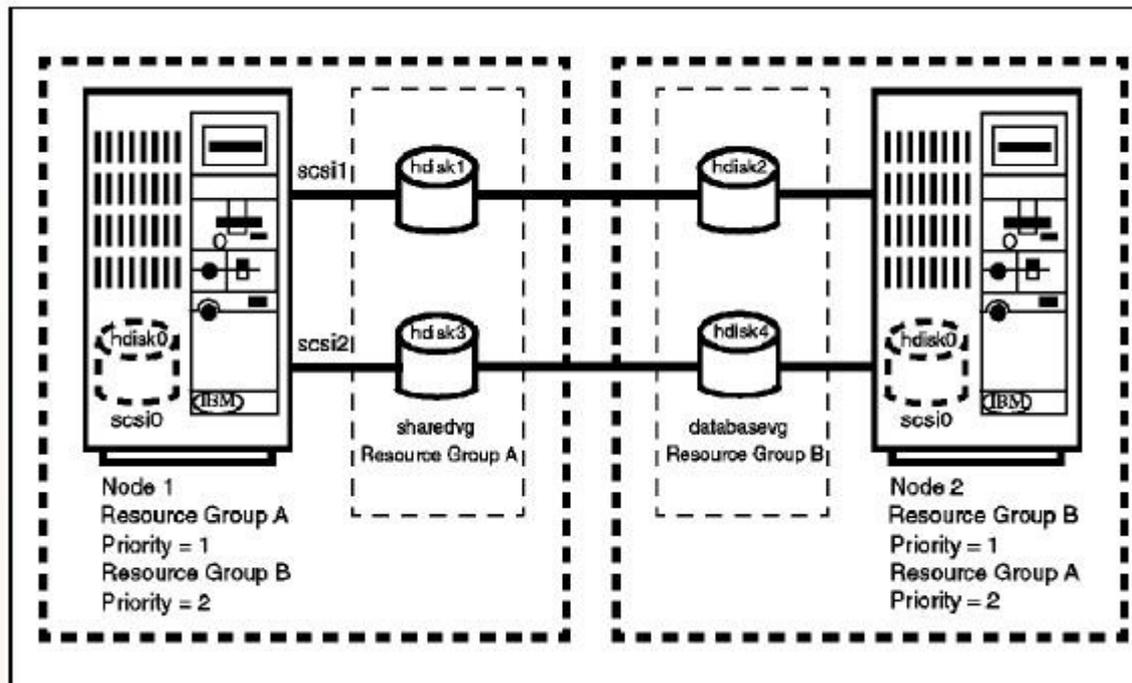
HACMP – Disk Access

- The first distinction that you need to make while designing a cluster is whether you need a non-concurrent or a concurrent shared disk access environment.
- The possible non-concurrent disk access configurations are:
 - Hot-standby
 - Rotating standby
 - Mutual takeover
 - Third-party takeover

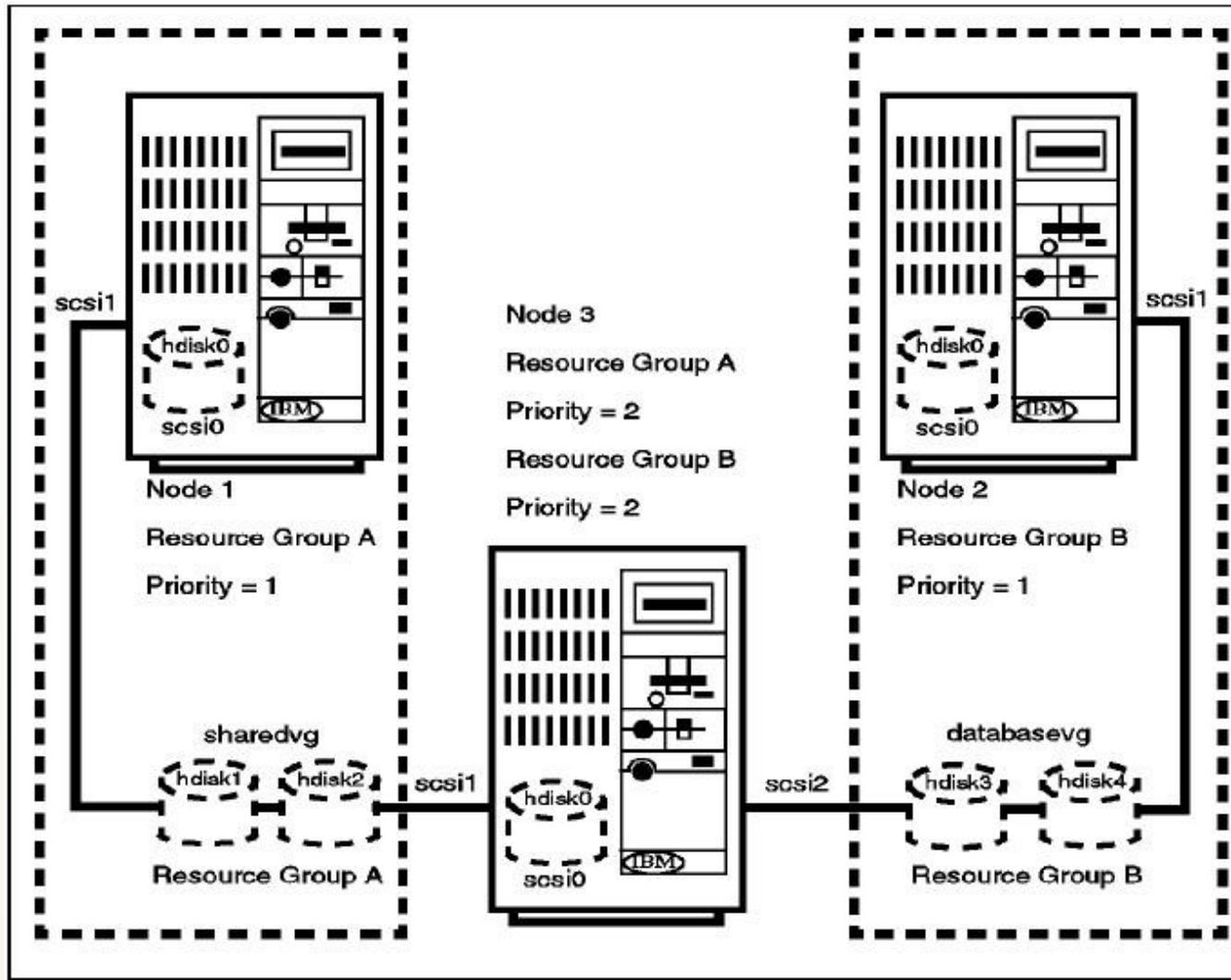
HACMP – Hot-standby configuration



HACMP – Mutual takeover



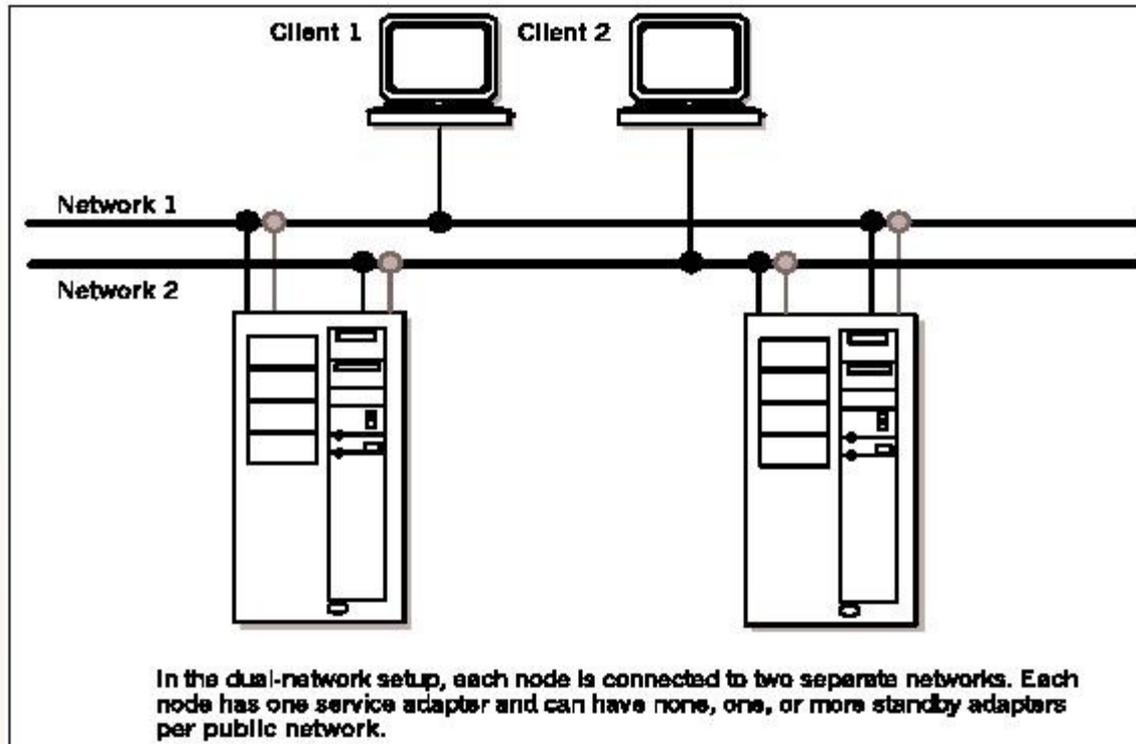
HACMP – Third party takeover



HACMP – IP Address takeover

- The goal of IP Address Takeover is to make the resource group's service address highly available and to give the clients the possibility of always connecting to the same IP address
 - Design the network topology.
 - Define a network mask.
 - Define IP addresses (adapter identifiers) for each node's service and standby adapters.
 - Define a boot address for each service adapter that can be taken over, if you are using IP address takeover or rotating resources.
 - Define an alternate hardware address for each service adapter that can have its IP address taken over, if you are using hardware address swapping.

HACMP – Dual Networks



HACMP – Network Adapter Labels

- Service Adapter Label
 - A node has one service adapter for each network to which it connects
- Standby Adapter(s) Label
 - Backup device for the service adapter
 - Zero or more standby adapters may exist
- Boot Adapter Label
 - Nodes use the boot label after a system reboot and before HACMP is started.
 - IPAT via IP address replacement
 - IPAT via IP Aliases

HACMP – Synchronization

- Numerous entities must be synchronized between nodes in a cluster, requires unique identifiers within the cluster
- Disaster Recovery requires **enterprise wide uniqueTM** identifiers
- Identifiers include:
 - Volume Group names
 - Logical Volume names
 - JFS Log Logical Volumes
 - Filesystem mount points
 - Resource Group names
 - IP addresses, names, and aliases
 - Application Servers and startup scripts
 - User and group names, UID's, GID's

HACMP – Cascading

- Topology description of a 2-way cascading Production Application Server Cluster:
- Cluster Name (*example*)
 - mtxpaex12
- Participating Nodes (*example*)
 - dalapmtx01, dalapmtx02
- Networks
 - Application traffic
 - Management traffic
- Applications
 - Application Instance 1
 - Application Instance 2
- System Independent Resource Groups (*example*)
 - mtxpaex1
 - mtxpaex2

HACMP – Cascading – IP Names

- **Boot Names:**
 - dalapmtx01-bt01 (192.168.1.1) *non-routed*
 - dalapmtx01-bt02 (192.168.1.2) *non-routed*
 - dalapmtx02-bt01 (192.168.1.3) *non-routed*
 - dalapmtx02-bt02 (192.168.1.4) *non-routed*
- **Persistent Names:**
 - dalapmtx01-pers (10.1.1.101) *routed*
 - dalapmtx02-pers (10.1.1.102) *routed*
- **Management Names:**
 - dalapmtx01-man (10.254.1.101) *routed*
 - dalapmtx02-man (10.254.1.102) *routed*
- **Service Names**
 - dalapmtx01 (10.1.1.111) *routed*
 - dalapmtx01-rg01 (10.1.1.121) *routed*
 - dalapmtx02 (10.1.1.112) *routed*
 - dalapmtx02-rg01 (10.1.1.122) *routed*
- **Service Aliases**
 - mtxpaex1 (dalapmtx01-rg01)
 - mtxpaex2 (dalapmtx02-rg01)

HACMP – Cascading – IP Names – Node 1

- Node
 - dalapmtx01
- Boot Names:
 - dalapmtx01-bt01 (192.168.1.1) *non-routed*
 - dalapmtx01-bt02 (192.168.1.2) *non-routed*
- Persistent Names:
 - dalapmtx01-pers (10.1.1.101) *routed*
- Management Names:
 - dalapmtx01-man (10.254.1.101) *routed*
- Service Names
 - dalapmtx01 (10.1.1.111) *routed*
 - dalapmtx01-rg01 (10.1.1.121) *routed*
- Service Aliases
 - mtxpaex1 *known by user* (dalapmtx01-rg01)

HACMP – Cascading – IP Names – Node 2

- Node
 - dalapmtx02
- Boot Names:
 - dalapmtx02-bt01 (192.168.1.3) *non-routed*
 - dalapmtx02-bt02 (192.168.1.4) *non-routed*
- Persistent Names:
 - dalapmtx02-pers (10.1.1.102) *routed*
- Management Names:
 - dalapmtx02-man (10.254.1.102) *routed*
- Service Names
 - dalapmtx02 (10.1.1.112) *routed*
 - dalapmtx02-rg01 (10.1.1.122) *routed*
- Service Aliases
 - mtxpaex2 *known by user* (dalapmtx02-rg01)

HACMP – Network Adapters – Node 1

- Node
 - dalapmtx01
 - RG: mtxpaex1 (*normal operating mode*)
- Application Service Network Adapter (**en4**)
 - 192.168.1.1 dalapmtx01-bt01
 - 10.1.1.101 dalapmtx01-pers
 - 10.1.1.111 dalapmtx01
 - 10.1.1.121 dalapmtx01-rg01
- Management Network Adapter (**en5**)
 - 192.168.1.2 dalapmtx01-bt02
 - 10.254.1.101 dalapmtx01-man
- Application Service Access Name
 - mtxpaex1

HACMP – Network Adapters – Node 2

- Node
 - dalapmtx02
 - RG: mtxpaex2 (*normal operating mode*)
- Application Service Network Adapter (**en4**)
 - 192.168.1.3 dalapmtx02-bt01
 - 10.1.1.102 dalapmtx02-pers
 - 10.1.1.112 dalapmtx02
 - 10.1.1.122 dalapmtx02-rg01
- Management Network Adapter (**en5**)
 - 192.168.1.4 dalapmtx02-bt02
 - 10.254.1.102 dalapmtx02-man
- Application Service Access Name
 - mtxpaex2

HACMP – Node 1 – RG mtxpaex2 Fail-over

- Node
 - dalapmtx01
 - RG: mtxpaex1 (*normal operating mode*)
 - RG: mtxpaex2 (*fail over mode*)
- Application Service Network Adapter (**en4**)
 - 192.168.1.1 dalapmtx01-bt01
 - 10.1.1.101 dalapmtx01-pers
 - 10.1.1.111 dalapmtx01
 - 10.1.1.121 dalapmtx01-rg01
 - 10.1.1.122 dalapmtx02-rg02
- Management Network Adapter (**en5**)
 - 192.168.1.2 dalapmtx01-bt02
 - 10.254.1.101 dalapmtx01-man
- Application Service Access Name
 - mtxpaex1
 - mtxpaex2

HACMP – Node 2 – RG mtxpaex2 Fail-over

- Node
 - dalapmtx02
 - No Resource Groups
- Application Service Network Adapter (**en4**)
 - 192.168.1.3 dalapmtx02-bt01
 - 10.1.1.102 dalapmtx02-pers
 - 10.1.1.112 dalapmtx02
- Management Network Adapter (**en5**)
 - 192.168.1.4 dalapmtx02-bt02
 - 10.254.1.102 dalapmtx02-man
- Application Service Access Name
 - **NONE**

HACMP – Application Installation

- Volume Groups, Logical Volumes, JFS Log Logical Volumes
 - mtropaex100vg
 - mtropaex1ap01lv
 - mtropaex1jfs1lv
 - mtropaex200vg
 - mtropaex2ap01lv
 - mtropaex2jfs1lv
- Directory Mount Points
 - /mtropaex1
 - /mtropaex2
- Application Servers
 - mtropaex1_appl01
 - mtropaex2_appl01
- Application Server Start/Stop Scripts
 - mtropaex1_appl01_start.sh
 - mtropaex1_appl01_stop.sh
 - mtropaex2_appl01_start.sh
 - mtropaex2_appl01_stop.sh

HACMP – VG's, LV's, FS's – Node 1

- **Node dalapmtx01**
 - Resource Groups
 - mtzpaex1 (*normal operating mode*)
 - Volume Groups and Logical Volumes
 - mtzpaex100vg
 - mtzpaex1ap01lv
 - Mounted File Systems
 - /dev/mtzpaex1ap01lv /mtzpaex1

HACMP – VG's, LV's, FS's – Node 2

- **Node dalapmtx02**
 - Resource Groups
 - mtzpaex2 (*normal operating mode*)
 - Volume Groups and Logical Volumes
 - mtzpaex200vg
 - mtzpaex2ap01lv
 - Mounted File Systems
 - /dev/mtzpaex2ap01lv /mtzpaex2

HACMP – Node 1 – RG mtxpaex2 Fail-over

- **Node dalapmtx01**
 - Resource Groups
 - mtxpaex1 (*normal operating mode*)
 - mtxpaex2 (***fail-over mode***)
 - Volume Groups and Logical Volumes
 - mtxpaex100vg
 - mtxpaex1ap01lv
 - **mtxpaex200vg**
 - **mtxpaex2ap01lv**
 - Mounted File Systems
 - /dev/mtxpaex1ap01lv /mtxpaex1
 - **/dev/mtxpaex2ap01lv /mtxpaex2**

HACMP – Node 2 – RG mtxpaex2 Fail-over

- **Node dalapmtx02**
 - Resource Groups
 - **None**
 - Volume Groups and Logical Volumes
 - **None**
 - Mounted File Systems
 - **None**

HACMP – User Configuration

- User Names
 - Each user must have their own login ID, not shared.
 - Auditors interpretation of SOX, SAS70
- Service User Names
 - User names for application administration
 - Remote login disabled
 - Must login with user name, then “**su**” to service user name
- UID and GID Numbers
 - Must be enterprise wide unique values
 - Must be consistent between systems

```
echo “username” | sum -r
```

```
echo “groupname” | sum -r
```

HACMP Policies

- Any changes to the startup environment of a clustered machine requires fail-over testing
- Enterprise wide unique identifiers shall be used for all shared resources
- Enterprise wide unique user name, group name, UID, and GID number will be used
- Regularly scheduled HACMP verification for each cluster.
- Regularly scheduled audit of HACMP patch levels
- Policies are located at the following URL:

<http://www.mtxia.com/css/High%20Availability/hacmppolicy.shtml>

HACMP Guidelines

- Oracle database's will normally be implemented in a non-concurrent access cluster
- Application servers will normally be implemented in 2-way cascading clusters
- In virtualized environment, HACMP will normally be implemented in a 1-way cascading cluster with fail-over to a standby node built with minimal resources.
- Guidelines are located at the following URL:

<http://www.mtxia.com/css/High%20Availability/hacmpguide.shtml>

HACMP Standards

- Node names and Cluster names
- IP addresses, Host Names, and Aliases
- VG, Major Numbers, LV, JFS Log, and Mount Points
- Resource Group Names
- Application Server Names
- Application start and stop script names
- User names and UID numbers
- Group names and GID numbers
- HACMP filesets

- Standards are located at the following URL:

<http://www.mtxia.com/css/High%20Availability/hacmpstand.shtml>

HACMP Procedures

- SAN Storage requests
- DNS requests
- Network cabling
- TSM backups
- rootvg backups
- Disaster Recovery
- Fail-over Testing
- LVM Modifications
- Change Control
- User Modifications

- Procedures are located at the following URL:

<http://www.mtxia.com/css/High%20Availability/hacmproc.shtml>

Q&A

