



Welcome to Capgemini Energy, a business service company that will aim to deliver your immense depth of utility industry knowledge and business service skills to energy and utility customers across North America.



## HACMP – High Availability

Introduction Presentation

April 28, 2005

# Introduction

- Scope
- HACMP – Concepts
- HACMP – Cluster Topology
  - NFS
  - Cascading – 1 way
  - Cascading – 2 way
  - Rotating
  - Concurrent
- HACMP – Cluster Resources
- CGE 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, fallover and fallback policies
  - settling time and the distribution policy
  - delayed fallback timer that affects fallback.
- Resource groups can be concurrent and non-concurrent.
- Resource Groups are used to coordinate the failover of:
  - Network Addresses
  - Disks
  - NFS mounts
  - Applications

# HACMP – Resource Group Failover

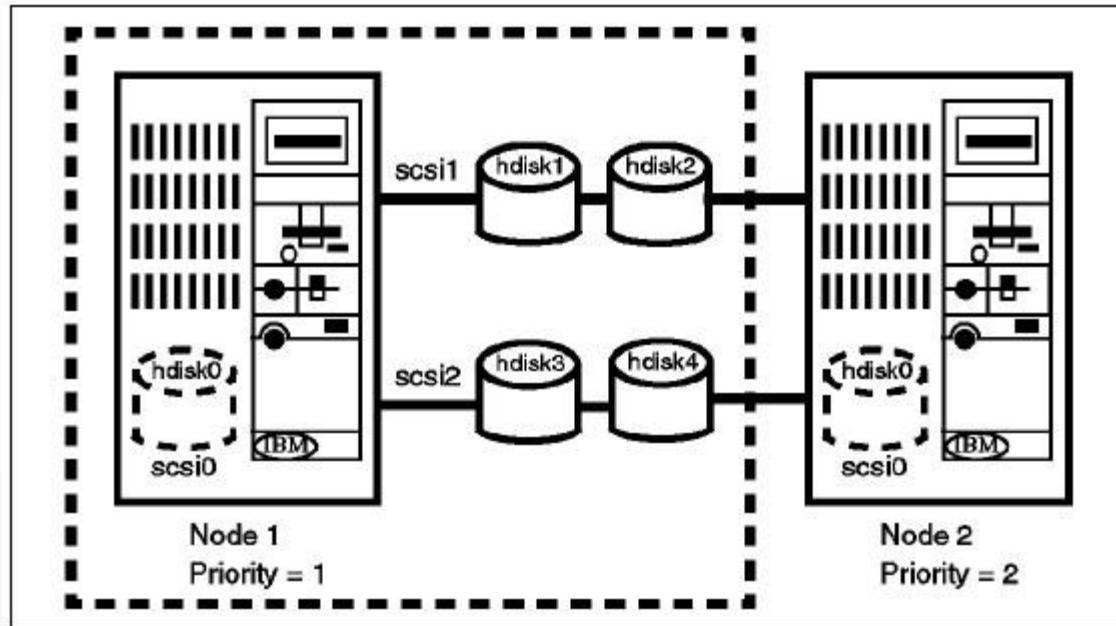
- Mechanisms implemented by HACMP for resource group failover :
  - Cascading
    - The active node with the highest priority controls the resource group
  - Concurrent
    - All active nodes have access to the resource group
  - Rotating
    - The node with the rotating resource group's associated service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

# HACMP – Disk Access

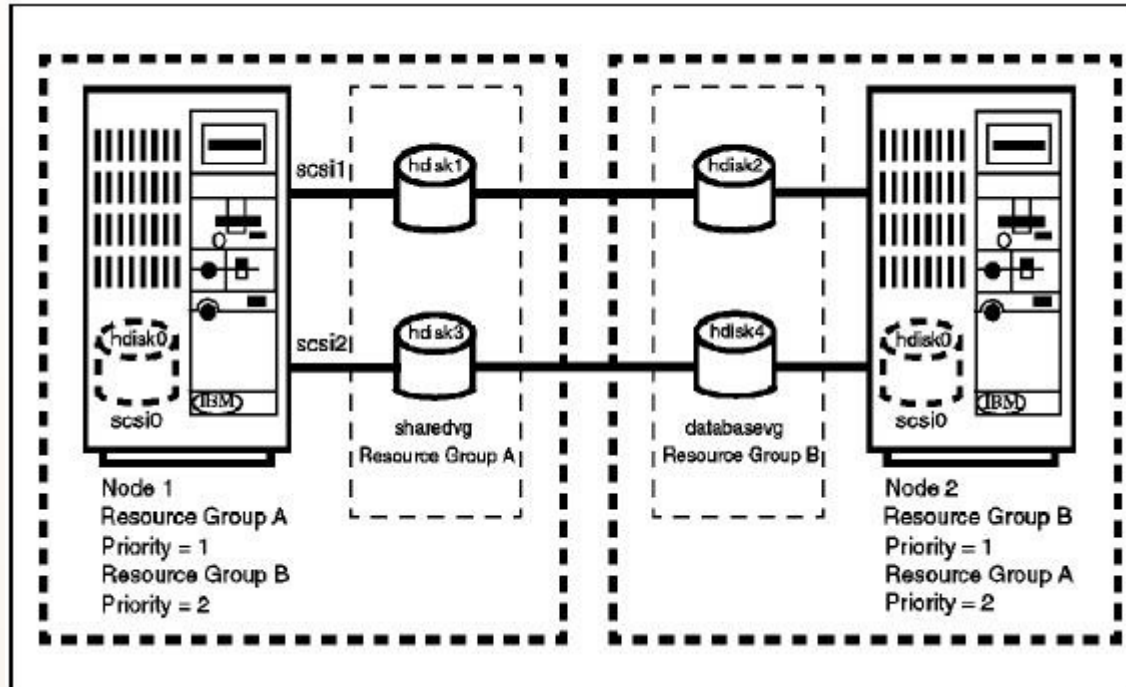
- The first distinction that you need to make while designing a cluster is whether you need a non-concurrent or a concurrent shared disk access environment.
- The possible non-concurrent disk access configurations are:
  - Hot-standby
  - Rotating standby
  - Mutual takeover
  - Third-party takeover



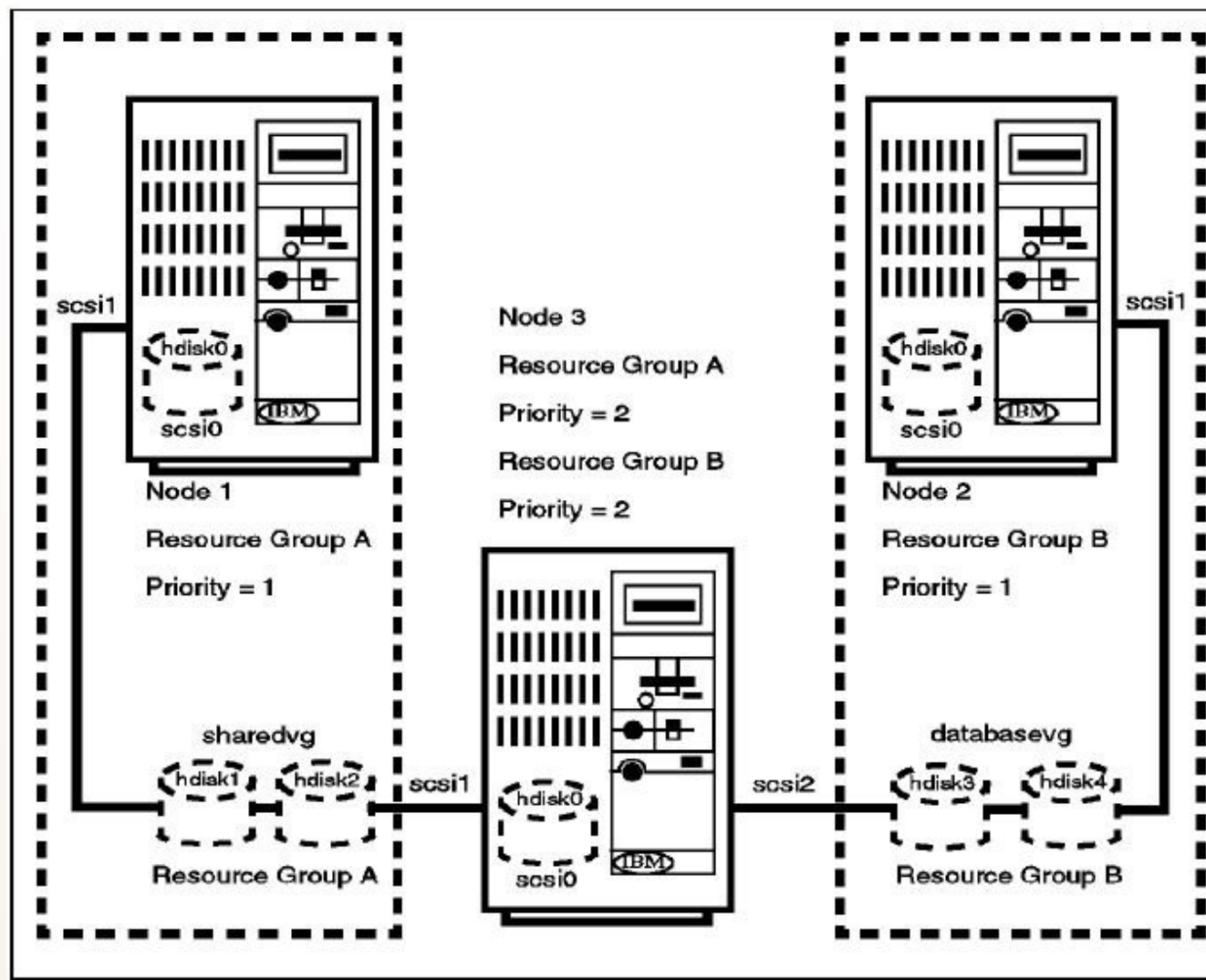
# HACMP – Hot-standby configuration



# HACMP – Mutual takeover



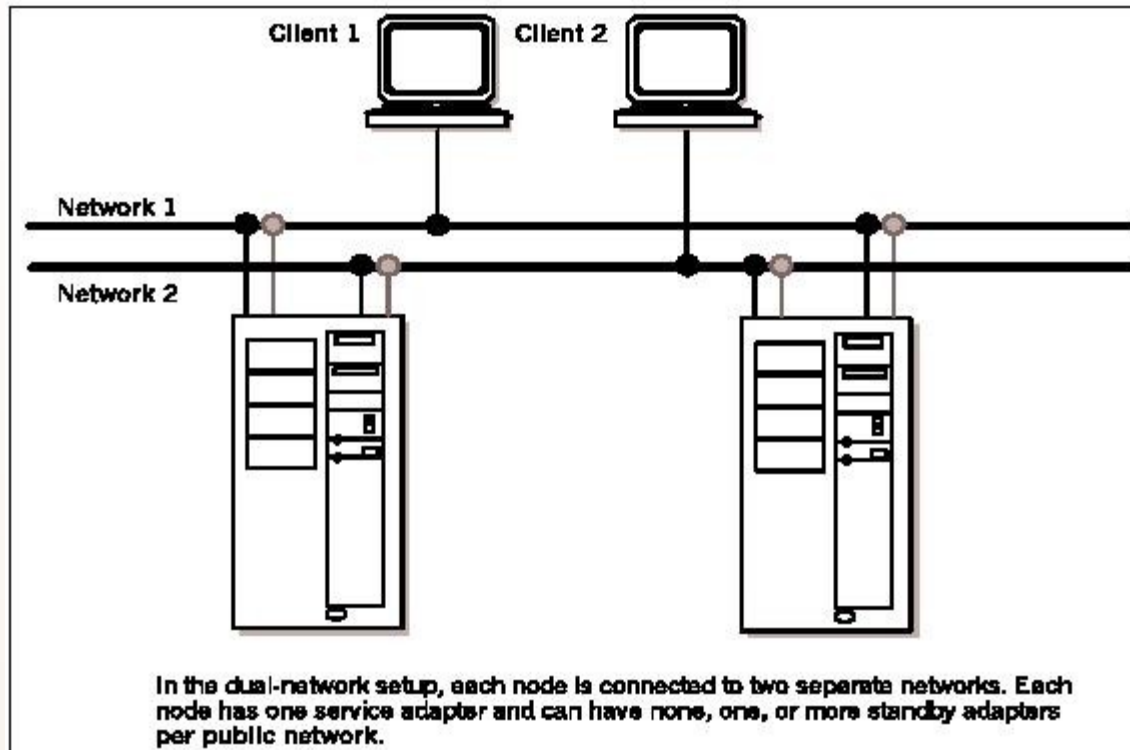
# HACMP – Third party takeover



# HACMP – IP Address takeover

- The goal of IP Address Takeover is to make the server'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 unique 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 – Cluster Topology

- Several different HACMP cluster topologies are utilized in CGE's environment
  - NFS
  - Cascading – 1 way
  - Cascading – 2 way
  - Rotating
  - Concurrent



# HACMP – Configured Clusters

- CGE has numerous HACMP clusters configured and running in both production and development environments. Those clusters are:
  - Production Atlas Database – 2 way cascading
  - Production EGATE Database – Concurrent Access
  - Production EGATE Appl Servers – Rotating
  - Devl EGATE Database – Concurrent Access
  - Devl EGATE Appl Servers – Cascading
  - Devl Atlas Database – NFS and Cascading
  - Devl Maximo Appl – 1 way cascading
  - Devl Middleware POC – Concurrent Access
  - Devl EGATE Database POC – Concurrent Access

# HACMP – NFS

- CGE currently has one NFS cluster running the development environment for the Atlas database:
- Cluster Name
  - ddcadat1
- Participating Nodes
  - ddcadat101, ddcadat102
- Resource Groups
  - ddcadat101rg1, ddcadat102rg1
- Service Labels
  - ddcadat101, ddcadat101\_nas, ddcadat102, ddcadat102\_nas
- Application Servers
  - ddcadat101server1, ddcadat102server1

# HACMP – Cascading

- CGE currently has several cascading clusters running a variety of applications. A topology description of the Production Atlas Database cluster follows:
- Cluster Name
  - delphicl01
- Participating Nodes
  - delphi\_e, delphi\_f
- Resource Groups
  - delphi\_e\_rg, delphi\_f\_rg
- Service Labels
  - delphi\_e, delphi\_f
- Application Servers
  - oracle\_e, tivoli\_e, dpcheck\_e
  - oracle\_f, tivoli\_f, dpcheck\_f

# HACMP – Cascading

- Topology description of the development EGATE Application Server Cluster:
- Cluster Name
  - ddcaaegaapps
- Participating Nodes
  - ddcaaega03, ddcaaega04, ddcaaega05
- Resource Groups
  - ega03rg, ega04rg, ega05rg
- Service Labels
  - ddcaaega03, ddcaaega04, ddcaaega05
- Application Servers
  - egate\_03, tivoli03, dpcheck\_03, tws\_03
  - egate\_04, tivoli04, dpcheck\_04
  - egate\_05, tivoli05

# HACMP – Cascading

- Topology description of the 1-way cascading development  
Maximo Application Server Cluster:
- Cluster Name
  - maximo
- Participating Nodes
  - ddcadtps01, ddcadtps02
- Resource Groups
  - tps\_appl01
- Service Labels
  - ddcadtps01, ddcadtps02
- Application Servers
  - none

# HACMP – Rotating

- CGE currently has one rotating cluster running a variety of applications:
- Cluster Name
  - mdctxuapps
- Participating Nodes
  - mdctxuapp80, mdctxuapp81, mdctuxapp82, mdctxuapp83
- Resource Groups
  - app80rg1, app81rg1, app82rg1
- Service Labels
  - mdctxuapp80, mdctxuapp81, mdctxuapp82
- Application Servers
  - app80server1, tivoli80
  - app81server1, tivoli81, tws\_81
  - app82server1, tivoli82

# HACMP – Concurrent

- CGE currently has several concurrent access clusters running Oracle. A topology description of the production EGATE database cluster follows:
- Cluster Name
  - egateopscl01
- Participating Nodes
  - mdctxudb85, mdctxudb86
- Resource Groups
  - mdctxudb85\_ip, mdctxudb86\_ip, opsconrg
- Service Labels
  - mdctxudb85, mdctxudb86
- Application Servers
  - Oracle, TSM

# HACMP – Concurrent

- A topology description of the development EGATE database POC cluster:
- Cluster Name
  - ddcaaegaops
- Participating Nodes
  - ddcpocega01, ddcpocega02
- Resource Groups
  - ega01\_ip, ega02\_ip, ora99conc
- Service Labels
  - ddcpocega02, ddcpocega02
- Application Servers
  - none



# HACMP – Concurrent

- A topology description of the development middleware POC cluster:
- Cluster Name
  - orapoc01
- Participating Nodes
  - ddcpocdb01, ddcpocdb02
- Resource Groups
  - db01rg1, db02rg1, ora02rg1
- Service Labels
  - ddcpocdb01, ddcpocdb02
- Application Servers
  - none

# HACMP Policies

- Any changes to the startup environment of a clustered machine requires failover 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://eperf.tu.com/GlobalSysAdmin/Power5/policies.shtml>

# HACMP Guidelines

- HACMP v5.2 is Current desired Version
- Oracle database's will normally be implemented in a non-concurrent access cluster
  - Current exception to this guideline is eGate
- Application servers will normally be implemented in 2-way cascading clusters
- In the new P590 Virtual environment, HACMP will be implemented in a 1-way cascading cluster with failover to a standby node built with minimal resources.
- Guidelines are located at the following URL:
  - <http://eperf.tu.com/GlobalSysAdmin/Power5/guidelines.shtml>

# HACMP Standards

- Machine Naming
- IP addresses, names, aliases
- VG, Major Numbers, LV, JFS Log, and Mount Points
- Cluster Names
- 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://eperf.tu.com/GlobalSysAdmin/Power5/standards.shtml>

# HACMP Procedures

- SAN Storage requests
  - DNS requests
  - Network cabling
  - TSM backups
  - rootvg backups
  - Disaster Recovery
  - Failover Testing
  - LVM Modifications
  - Change Control
  - User Modifications
- 
- Procedures are located at the following URL:
    - <http://eperf.tu.com/GlobalSysAdmin/Power5/procedures.shtml>

# Q&A