



Automated Instance Failover using the IBM DB2 High Availability Instance Configuration Utility (db2haicu)

April 2008

Authors:

Steven Raspudic, IBM Toronto Lab (stevera@ca.ibm.com)

Selvaprabhu Arumuggharaj, IBM Lenexa Lab (sprabhu@us.ibm.com)

Saurabh Sehgal, IBM Toronto Lab (ssehgal@ca.ibm.com)

Table of Contents

1. Introduction and Overview	3
2. Before You Begin	4
2.1 Knowledge Prerequisites	4
2.2 Hardware Configuration Used	4
2.3 Software Configuration Used	5
3. Overview of Important Concepts	5
3.1 The db2haicu Utility	5
3.2 HA Overview	5
3.3 Two-Node Shared Disk HA Topology	5
4. Setting up an Automated HA Topology for an Instance with No Databases	7
4.1 Topology Configuration	7
4.1.1 Basic Network Setup	9
4.1.2 DB2 Instance Setup	9
4.1.3 Cluster Preparation	11
4.1.4 Network Time Protocol	12
4.1.5 Client Reroute	12
4.2 The db2haicu Interactive Setup Mode	12
4.2.1 DB2 Operations That Will Add or Remove Cluster Elements	20
5. Setting up an Automated HA Topology with Pre-existing Databases	26
5.1 Topology Setup	26
5.2 Utilizing the db2haicu XML File	27
6. Post Configuration Testing	34
6.1 The 'Power off' Test	37
6.1.1 Reboot	38
6.2 Network Failures	39
6.3 Disk I/O Failures	39
6.4 DB2 Failures	40
6.5 Manual Instance Control	43
7. Maintenance	45
7.1 Disable High Availability	45
7.2 Db2haicu Maintenance Mode	45
8. Problem Determination	52
8.1 Restrictions and Supportability	52
8.2 Db2haicu '-delete' Option	52
8.3 Hostnames specified to db2haicu	53
8.4 Broadcast pings and netmon.cf	53
8.5 The syslog and db2diag.log	53
8.6 Data Capture/Gathering for TSA/RSCT	54
9. Summary	55

1. Introduction and Overview

This paper describes a distinct configuration of an automated IBM® DB2® for Linux®, UNIX®, and Windows® failover solution using shared disk storage. The configuration is based on the DB2 High Availability (HA) feature and the DB2 High Availability Instance Configuration Utility (db2haicu) available with DB2 Version 9.5.

Target Audience for this Paper:

- DB2 database administrators
- UNIX system administrators

2. Before you begin

Below you will find information on knowledge requirements, as well as hardware and software configurations used to set up the topology depicted in Sections 4. It is important that you read this section prior to beginning any setup.

2.1 Knowledge Prerequisites

- Basic understanding of DB2 9.5 and HA feature*
- Basic understanding of IBM Tivoli® System Automation for Multiplatforms (TSA) cluster manager software**
- Basic understanding of Linux operating system concepts

*Information on DB2 HA can be found here:

<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/topic/com.ibm.db2.luw.admin.ha.doc/doc/c0006354.html>

**Information on TSA can be found here:

<http://www.ibm.com/software/tivoli/products/sys-auto-linux/>

2.2 Hardware Configuration Used in Setup

For the topology covered in this paper, the following hardware configuration was used:

- Two machines each with:
 - o CPU = 2 AMD Opteron Processors 246, 2 GHz each
 - o Network Adapters = 1 IBM NetXtreme BCM5704 Gigabit Ethernet Adapter
 - o Memory = 2.5 GB
 - o IBM TotalStorage® FASTT200 High Availability (HA) Storage Server (Model 3542-2RU)

2.3 Software Versions Used in Setup

For the topology covered in this white paper, the following software configuration was used:

- DB2 9.5
- Linux Enterprise Server
 - o **Distribution Information:**
SUSE Linux Enterprise Server
VERSION = 10
PATCHLEVEL = 1
 - o **Kernel Information:**
2.6.16.46-0.12-smp #1

Note: The most current versions of DB2 9.5 and TSA that have been tested can be found at Supported Distributions - DB2 for Linux page.

<http://www.ibm.com/software/data/db2/linux/validate/>

3. Overview of Important Concepts

3.1 The db2haicu utility

Db2haicu is a tool available with DB2 9.5 and stands for the 'DB2 High Availability Instance Configuration Utility.' This utility takes in user input regarding the software and hardware environment of a DB2 instance, and configures the instance for High Availability using the TSA cluster manager. During this configuration process, all necessary resources, dependencies, and equivalencies are automatically defined to TSA. **Note:** TSA does not need to be manually installed on your system as it is pre-packaged with DB2 9.5.

Two input methods can be used to provide the necessary data to db2haicu. The first method is the interactive mode, where you are prompted for input at the command line. The second input method is the XML mode, where db2haicu can parse the necessary data from a user-defined XML file.

The db2haicu interactive mode is covered in Section 4 and the db2haicu XML mode is covered in Section 5.

3.2 HA Overview

The High Availability (HA) feature of DB2 9.5 allows a Database Administrator (DBA) to have redundant systems and a failover mechanism thereby improving the availability of database solution; in the event of an unplanned or planned outage, an instance can be failed over to the available redundant systems.

The db2haicu tool can be used to configure such an HA system. During the db2haicu configuration process, the necessary HA resources and their relationships are defined to the cluster manager. Failure events in the HA system can then be detected automatically and takeover operations can be run without manual intervention.

3.3 Two-Node Shared Disk HA topology

A two-node shared disk HA topology contains an active node (e.g., amdha01) to host the DB2 instance and data, and a passive node (e.g., amdha02) to which DB2 resources are moved in case of a failure event.

For this topology, the two nodes are connected to each other over a single public network and wired to a shared disk storage system. The term 'shared storage' is used to describe a system in which one or more disks are cabled and defined to both the active and passive nodes but only mounted at one node at a time, i.e., the node that hosts the active DB2 instance. At any given time, only one node has exclusive access to the shared storage disks. That is, only one node can effectively do input/output (I/O) to shared storage disks.

Two primary examples are fully worked through in this paper. First, a single network shared disk HA topology with no pre-existing databases is automated using the db2haicu interactive mode. This example will demonstrate how the creation of database objects will add mount

point resources to the cluster, and conversely, how the removal of databases objects will remove mount point resources from the cluster.

Secondly, a single network HA topology, automated using the db2haicu XML mode for DB2 instance with pre-existing databases is configured. This shows how db2haicu can automatically discover mount points for pre-existing databases and create the appropriate cluster resources.

4. Setting up an Automated Single Network HA Topology using the db2haicu Interactive Mode for DB2 Instance with No Pre-existing Databases

The configuration of an automated single network HA topology, as illustrated in Fig. 1, is described in the steps below.

Notes:

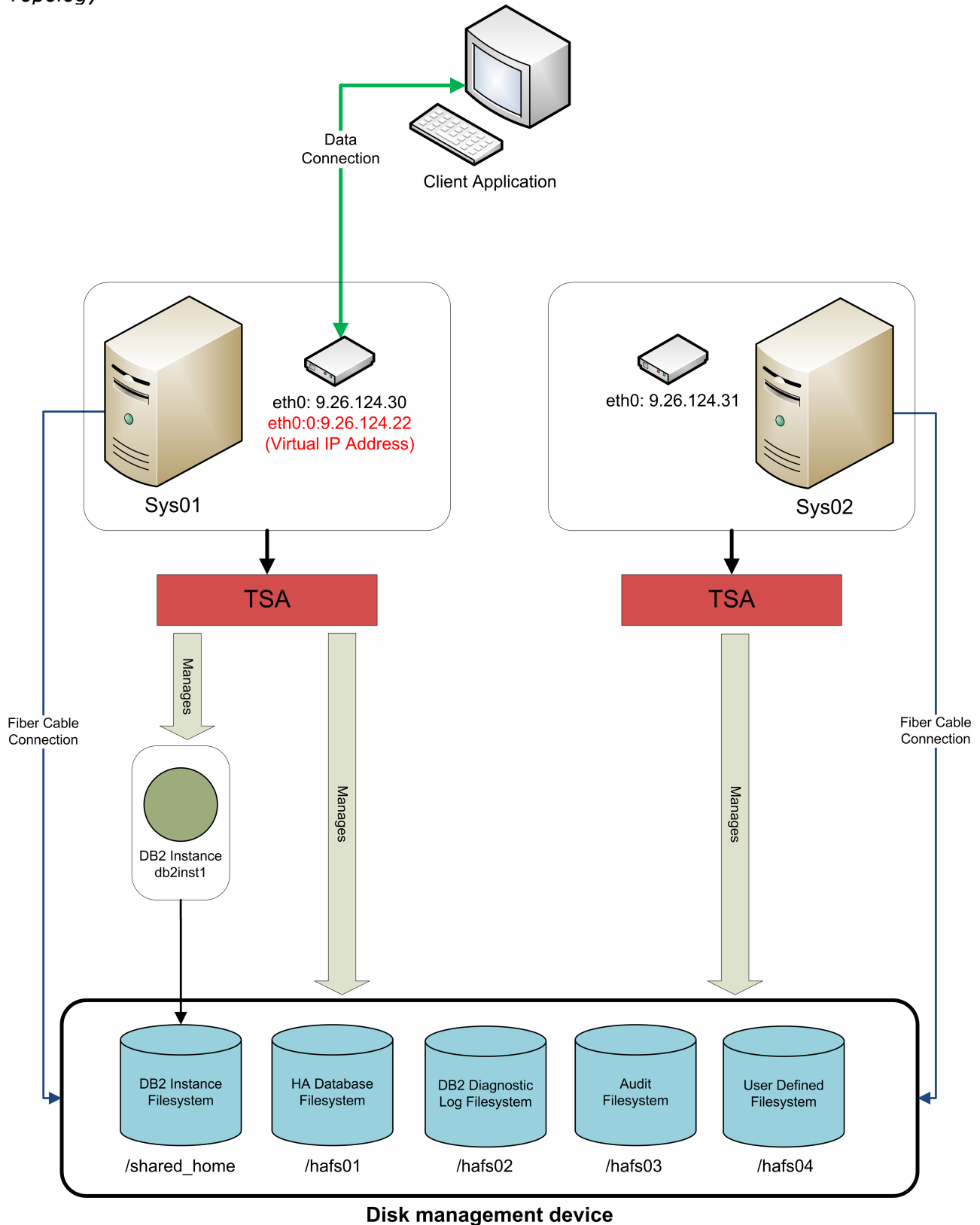
1. There are two parts to this configuration. The first part describes the preliminary steps needed to configure the shared disk HA topology. The second part describes the use of db2haicu's interactive mode to automate the topology for failovers.
2. The parameters used for various commands described below are based on the topology illustrated in Fig. 1. You must change the parameters to match your own specific environment.

4.1 Topology Configuration

This topology makes use of two nodes: the active node (e.g., amdha01) to host the necessary DB2 mount points and the DB2 instance, and the passive node (e.g., amdha02) to which these mount points and the DB2 instance can be moved in case of a failure event.

The nodes are connected to each other using a *public* network. The public network is defined to host the virtual IP address that allows clients to connect to the active node.

Fig. 1 Automated Single Network HA Topology



4.1.1 Basic Network Setup

The two machines used for this topology contain one network interface each (eth0).

1. The eth0 network interfaces are connected to each other through the external network cloud forming the *public* network. The following are the static IP addresses of the eth0 adapters on the active and passive nodes:

```
Primary node (amdha01)  
eth0: 9.26.124.30 (255.255.255.0)
```

```
Standby node (amdha02)  
eth0: 9.26.124.31 (255.255.255.0)
```

2. Make sure that the active and passive node names are mapped to their corresponding public IP addresses in the /etc/hosts file:

```
9.26.124.30  amdha01 amdha01.torolab.ibm.com  
9.26.124.31  amdha02 amdha02.torolab.ibm.com
```

Defining the hostnames in the /etc/hosts file allows the nodes to identify each other in case of a DNS server failure. All cluster nodes should have the same entries on /etc/hosts file. In addition, they all should have fixed IP addresses. You should ensure that on the machine amdha01/amdha01.torolab.ibm.com the 'hostname' command returns the value 'amdha01' and on the machine amdha02/amdha02.torolab.ibm.com the 'hostname' command returns the value 'amdha02'.

3. The active and the passive machines should be able to ping each other on the public network. Issue the following commands on both nodes and make sure that they complete successfully:

```
% ping amdha01  
% ping amdha02
```

4.1.2 DB2 Instance Setup

DB2 instance owner should have the same user ID and group ID on all the nodes in the cluster domain. In addition, it is preferred that the DB2 instance owner have the same password on cluster nodes.

The 'id' operating system command can be used to verify the user ID and group ID for the instance owner on all cluster nodes.

```
For our case:  
User 'db2inst1' defined on /etc/passwd:  
db2inst1:x:20351:7777:DB2 Instance 1:/shared_home/db2inst1:/bin/ksh  
  
/shared_home/db2inst1> id  
uid=20351(db2inst1) gid=7777(pdxdb2) groups=7777(pdxdb2)
```

Mount Points:

The mount points for the shared disks must be defined to the operating systems being run on the active and the passive nodes (/etc/fstab file for Linux, /etc/filesystems for AIX®).

When you add new mount points for the cluster by adding entries to /etc/fstab on all cluster nodes, use the 'noauto' option to prevent the mount points from being automatically mounted on more than one machine in the cluster.

All the required file systems should be mounted only on the active node prior to configuring your instance for HA using db2haicu. These mount points should also be mountable from any other cluster node.

Permission for all mount points should be 755 (rwxr-xr-x),
owner = instance_owner, group = instance_owner's_group

For our case, /etc/fstab contained the following mount points to be used with the cluster. Note that all machines that are potential hosts of the mount points must have the identical entries in their own /etc/fstab. Here is an example of the needed /etc/fstab entries for a Linux machine.

# LUN	Mount Point	FileSystem Type	Automount?
/dev/sdd	/shared_home	ext3	noauto 0 0
/dev/sde	/hafs01	ext3	noauto 0 0
/dev/sdf	/hafs02	ext3	noauto 0 0
/dev/sdh	/hafs03	ext3	noauto 0 0
/dev/sdi	/hafs04	ext3	noauto 0 0
/dev/sdc	/hafs05	ext3	noauto 0 0

Mount Points for AIX

Note that an important difference between the Linux and AIX operating systems is that AIX uses a 'Volume Manager' to control disk volumes. Thus, the setup of disk volumes on AIX is somewhat different. Here are the key points to keep in mind.

The volume groups need to have the same major numbers on each machine and its failover:

1. The volume groups on the second node will be imported from the first (active) node. Ensure that the major numbers of all volume groups on the second node are the same as they are on the first (active) node. To verify that they are the same, perform the following steps:

- a. Determine the major number of the volume groups on the first (active) node using the command:

```
ls -al /dev/<volume group name>
```

The volume group major number is equivalent to the major device number of the special device file. For example, the output for the command:

```
ls -al /dev/vgdb2home
```

should be similar to this:

```
crw-rw---- 1 root system 52, 0 Aug 27 19:57 /dev/vgdb2home
```

In this example, the volume group major number is 52. Repeat these steps for the volume group vgp0.

b. Determine whether the same major number is available on the failover system. Issue the command `lvlstmajor` to determine the available major numbers.

c. If the major number is available on the failover node, import the volume group, specifying the required major numbers:

```
importvg -y vgdb2home -V <first major number> <disk>
```

d. If any required major number is not available on the failover node, you must reconfigure the volume group that is using the major number to make it available.

2. Configure the `vgdb2home` and `vgp0` volume groups not to come online automatically after a reboot:

```
chvg -a n vgdb2home
chvg -a n vgp0
```

3. Create the logical volume `/dev/lvdb2home` on the volume group `vgdb2home` and the logical volume `/dev/lvp0` on `vgp0`.

4. On both nodes, edit the `/etc/filesystems` file to include an entry for the mount points:

```
/shared_db2home:
dev =      /dev/lvdb2home
vfs =      jfs2
log =      INLINE
mount =    false
type =     bcu
check =    false
options =  rw
account =  false
```

```
/db2datap0:
dev =      /dev/lvp0
vfs =      jfs2
log =      INLINE
mount =    false
type =     bcu
check =    false
```

```
options =    rw
account =   false
```

4.1.3 Cluster Preparation

Before using the db2haicu tool, the primary and the standby nodes must be prepared with the proper security environment.

1) With root authority, issue the following command on all cluster nodes:

```
root@amdha01# /usr/sbin/rsct/bin/preprnode amdha01 amdha02
.....
root@amdha02# /usr/sbin/rsct/bin/preprnode amdha01 amdha02
```

This command only needs to be run *once* per node and *not* for every DB2 instance that is made highly available.

4.1.4 Network Time Protocol

It is recommended (but not mandatory) that the time and dates on cluster nodes be synchronized.

Synchronized clocks can make problem determination more straightforward because time ordering of events that appear in various log files can be performed without applying a delta time correction to account for time differences between machines. Note that the Network Time Protocol can be used for this purpose. Refer to your operating system documentation for information on how to configure NTP for your system.

4.1.5 Client Reroute

The client reroute feature allows a DB2 client application to recover from a lost database connection in case of a network failure. In the HA configuration discussed in this paper, we use a virtual IP address that allows clients to connect to the active node. Identify the IP address that will be created and failed over as part of the HA failover. In addition, identify the port number for the instance TCP/IP listener by checking the value of the SVCENAME DBM CFG parameter.

1) Issue the following command on the node hosting the instance directory to configure the virtual IP address for client reroute:

```
/shared_home/db2inst1/% db2 update alternate server for database \  
hadb using hostname 9.26.124.22 port 55445
```

In this example, 9.26.124.22 is the virtual IP address and port 55445 is the value used in the DBM CFG parameter 'SVCENAME'.

4.2 The db2haicu Interactive Setup Mode

After the preceding preliminary configuration steps are completed, the db2haicu tool can be used to automate HA failover.

The db2haicu must be run on the node hosting the DB2 instance. The details involving the process are outlined in the following section.

Note: The `...` above a db2haicu message indicates continuation from a message displayed in a previous step.

Creating a Cluster Domain

Log on to the machine hosting the DB2 instance and issue the `db2haicu` command:

The following welcome message will be displayed on the screen:

```
/shared_home/db2inst1% db2haicu
Welcome to the DB2 High Availability Instance Configuration Utility (db2haicu).

You can find detailed diagnostic information in the DB2 server diagnostic log file called
db2diag.log. Also, you can use the utility called db2pd to query the status of the cluster
domains you create.

For more information about configuring your clustered environment using db2haicu, see
the topic called 'DB2 High Availability Instance Configuration Utility (db2haicu)' in the
DB2 Information Center.

db2haicu determined the current DB2 database manager instance is db2inst1. The
cluster configuration that follows will apply to this instance.

db2haicu is collecting information on your current setup. This step may take some time
as db2haicu will need to activate all databases for the instance to discover all paths ...
When you use db2haicu to configure your clustered environment, you create cluster
domains. For more information, see the topic 'Creating a cluster domain with db2haicu' in
the DB2 Information Center. db2haicu is searching the current machine for an existing
active cluster domain ...
db2haicu did not find a cluster domain on this machine. db2haicu will now query the
system for information about cluster nodes to create a new cluster domain ...

db2haicu did not find a cluster domain on this machine. To continue configuring your
clustered environment for high availability, you must create a cluster domain; otherwise,
db2haicu will exit.

Create a domain and continue? [1]
1. Yes
2. No
```

Note that the number contained within square brackets is the default for that prompt; simply press Enter to select the default value.

We must now create a cluster domain.

1) Type '1' and press Enter at the following initial prompt.

```
...
Create a domain and continue? [1]
1. Yes
2. No
1
```

2) Enter a unique name for the domain you want to create (we use HA_domain) and the number of nodes contained in the domain (2 in our case).

```
...
Create a unique name for the new domain:
HA_domain
Nodes must now be added to the new domain.
How many cluster nodes will the domain HA_domain contain?
2
```

3) Follow the prompts to enter the names of the two cluster nodes and confirm domain creation.

```
...
Enter the host name of a machine to add to the domain:
amdha01
Enter the host name of a machine to add to the domain:
amdha02
```

db2haicu can now create a new domain containing the two machines that you specified. If you choose not to create a domain now, db2haicu will exit.

```
Create the domain now? [1]
1. Yes
2. No
1
Creating domain HA_domain in the cluster ...
Creating domain HA_domain in the cluster was successful.
```

Quorum Configuration

After the domain creation has completed, a quorum must be configured for the cluster domain. The supported quorum type for this solution is a 'network quorum'. A network quorum (or network tiebreaker) is a pingable IP address that is used to decide which node in the cluster will serve as the 'active' node during a site failure, and which nodes will be offline. Note that the machine hosting this IP address does not need any particular software or operating system level installed; its primary requirement is that it can be pinged from all nodes in the cluster, and must remain pingable in the case of cluster node failures.

You will be prompted by db2haicu to enter quorum configuration values:

...

You can now configure a quorum device for the domain. For more information, see the topic 'Quorum devices' in the DB2 Information Center. If you do not configure a quorum device for the domain, then a human operator will have to manually intervene if subsets of machines in the cluster lose connectivity.

Configure a quorum device for the domain called HA_domain? [1]

1. Yes
2. No

From the preceding prompt:

1) Type '1' and press Enter to create the quorum.

....

1

The following is a list of supported quorum device types:

1. Network Quorum

Enter the number corresponding to the quorum device type to be used: [1]

2) Type '1' and press Enter again to select the Network Quorum type. Then follow the prompt to enter the IP address you would like to use as a network tiebreaker.

...

Specify the network address of the quorum device:

9.26.4.5

Configuring quorum device for domain HA_domain ...

Configuring quorum device for domain HA_domain was successful.

Quorum configuration is now completed.

Network Setup

After the quorum configuration, you may define the public network of your system to db2haicu. If network failure detection is important to your configuration, you *must* follow the prompts and add the networks to the cluster at this point. The db2haicu tool automatically discovers all network interfaces.

Here is an example:

...

The cluster manager found two network interface cards on the machines in the domain. You can use db2haicu to create networks for these network interface cards. For more information, see the topic 'Creating networks with db2haicu' in the DB2 Information Center.

Create networks for these network interface cards? [1]

1. Yes
2. No

1

Enter the name of the network for the network interface card: eth0 on cluster node: amdha01

1. Create a new public network for this network interface card.
2. Create a new private network for this network interface card.

Enter selection:

1

Are you sure you want to add the network interface card eth0 on cluster node amdha01 to the network db2_public_network_0? [1]

1. Yes
2. No

1

Adding network interface card eth0 on cluster node amdha01 to the network db2_public_network_0 ...

Adding network interface card eth0 on cluster node amdha01 to the network db2_public_network_0 was successful.

Enter the name of the network for the network interface card: eth0 on cluster node: amdha02

1. db2_public_network_0
2. Create a new public network for this network interface card.
3. Create a new private network for this network interface card.

Enter selection:

1

Are you sure you want to add the network interface card eth0 on cluster node amdha02 to the network db2_public_network_0? [1]

1. Yes
2. No

1

Adding network interface card eth0 on cluster node amdha02 to the network db2_public_network_0 ...

Adding network interface card eth0 on cluster node amdha02 to the network db2_public_network_0 was successful.

Cluster Manager Selection

After the network definitions, db2haicu prompts you for the cluster manager software that you are using for the current HA setup. For our purpose, we select TSA:

...

The cluster manager name configuration parameter (high availability configuration parameter) is not set. For more information, see the topic 'cluster_mgr - Cluster manager name configuration parameter' in the DB2 Information Center. Do you want to set the high availability configuration parameter?

The following are valid settings for the high availability configuration parameter:

1. TSA
2. Vendor

Enter a value for the high availability configuration parameter: [1]

1

Failover Policy

Now you need to configure the failover policy for the instance db2inst1. The failover policy determines the machines on which the cluster manager will restart the database manager if the database manager goes offline unexpectedly.

For our purpose, we select option 3. Note that the failover policy is a powerful concept for larger clusters (with more nodes and more partitions), but for a simple two-node single partition setup (such as this one), it is generally best to select option 3.

...

The following are the available failover policies:

1. Local Restart -- during failover, the database manager will restart in place on the local machine
2. Round Robin -- during failover, the database manager will restart on any machine in the cluster domain
3. Active/Passive -- during failover, the database manager will restart on a specific machine
4. M+N -- during failover, the database partitions on one machine will failover to any other machine in the cluster domain (used with DPF instances)
5. Custom -- during failover, the database manager will restart on a machine from a user-specified list

Enter your selection:
3

Then db2haicu will prompt you to designate any noncritical mount points. For this case, we chose to designate only one noncritical mount point.

...

You can identify mount points that are noncritical for failover. For more information, see the topic 'Identifying mount points that are noncritical for failover' in the DB2 Information Center. Are there any mount points that you want to designate as noncritical? [2]

1. Yes
2. No

1

Enter the full path of the mount to be made non-critical:
/tmp

Note that you should add any mount points to the noncritical path list that you are sure that you never want to failover. This list should include any mount points specified in /etc/fstab that are local mount points and will never be failed over.

Active/Passive failover policy was chosen. You need to specify the host names of an active/passive pair.

Enter the host name for the active cluster node:
amdha01
Enter the host name for the passive cluster node:
amdha02

The db2haicu tool will now automatically add the DB2 partition instance to the specified cluster manager at this point.

Virtual IP Address Setup

Once the database partition has been added to the cluster, db2haicu will prompt you to create a virtual IP address:

```
...
Do you want to configure a virtual IP address for the DB2 partition: 0? [2]
1. Yes
2. No
```

1) At the preceding prompt, type '1' and press Enter to continue. Enter the value of the virtual IP address that you configured with the client reroute setup (Section 4.1.5).

```
....
1
Enter the virtual IP address:
9.26.124.22
...
```

2) Enter the subnet mask associated with the virtual IP address, followed by the network on which this IP address will reside. We assign this IP address to the public network to allow external clients to connect.

```
...
Enter the subnet mask for the virtual IP address 9.26.124.22:
[255.255.255.0]
255.255.255.0
Select the network for the virtual IP 9.26.124.22:
1. db2_public_network_0
Enter selection:
1
Adding virtual IP address 9.26.124.22 to the domain ...
Adding virtual IP address 9.26.124.22 to the domain was successful.
```

You must make sure that your IP address and subnet mask values are well formed and correspond with the subnet mask of the network you chose. All invalid inputs will be rejected. In such a case, you are advised to examine the IP addresses and netmasks of the NIC components of the network (using the 'ifconfig' command) and verify that the IP address and netmask specified are compatible with each of the NICs in the network. In addition, make sure that the IP address that you want to add is not already present on the network.

After the virtual IP address configuration, the Automated Cluster-controlled HA configuration is complete.

Note that to display the state of the cluster, you can issue the TSA command 'lssam' or the 'db2pd -ha' command. Issue the 'lssam' command to see the resources created during this process.

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

The command `'db2pd -ha'` can also be issued from the instance owner ID to examine the state of the resources:

```

/shared_home/db2inst1% db2pd -ha
      DB2 HA Status
Instance Information:
Instance Name           = db2inst1
Number Of Domains       = 1
Number Of RGS for instance = 1

Domain Information:
Domain Name             = HA_domain
Cluster Version         = 2.4.7.4
Cluster State           = Online
Number of nodes         = 2

Node Information:
Node Name               State
-----
amdha01                 Online
amdha02                 Online

Resource Group Information:
Resource Group Name     = db2_db2inst1_0-rg
Resource Group LockState = Unlocked
Resource Group OpState  = Online
Resource Group Nominal OpState = Online
Number of Group Resources = 3
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Member Resource Information:
Resource Name           = db2_db2inst1_0-rs
Resource State          = Online
Resource Type           = DB2 Partition
DB2 Partition Number    = 0
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name           = db2mnt-shared_home-rs
Resource State          = Online
Resource Type           = Mount
Mount Resource Path     = /shared_home
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name           = db2ip_9_26_124_22-rs
Resource State          = Online
Resource Type           = IP

Network Information:
Network Name            Number of Adapters
-----
db2_public_network_0   2

Node Name               Adapter Name
-----
amdha01                 eth0
amdha02                 eth0

Quorum Information:
Quorum Name             Quorum State
-----
db2_Quorum_Network_9_26_4_5:11_35_35  Online
Fail                    offline
Operator                Offline

```

During cluster domain creation, resource groups for the instance are created. Resource group to instance mapping is stored in a binary configuration file called db2ha.sys (under \$INSTANCEDIR/sqllib/cfg dir). We can dump the values using db2hareg utility.

Generally, the entries in the HA registry are only useful in cases of problem determination. It is advised to contact IBM Service in cases where a deeper understanding of the contents of this registry file is required.

4.2.1 DB2 Operations That Will Add or Remove Cluster Resource Elements

This section covers DB2 operations that add or remove resources from the cluster manager automatically.

The database manager communicates with the cluster manager whenever instance changes require cluster changes, such that users are freed from having to perform separate cluster operations after performing instance configuration changes.

Database Creation

Let us first create a database on one of the shared disks defined in /etc/fstab and mounted at the active node.

Note that we continue to use the HA instance created in Section 4.2 above. Create a database 'HADB' with the database directory and storage paths on /hafs01 using the following command.

```
db2 "create database hadb on /hafs01"
```

Once the database is created, issue the 'lssam' command as root to see the new resources created during this process.

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
| - Online IBM.Application:db2mnt-hafs01-rs
  | - Online IBM.Application:db2mnt-hafs01-rs:amdha01
  | - Offline IBM.Application:db2mnt-hafs01-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

Table Space Creation

Continuing with this example, we will connect to the database 'hadb' and create a new DMS table space with storage path on /hafs02.

```
db2 "create tablespace dms01 managed by database using (file
'/hafs02/db2inst1/CONT/dms01c1' 500) extentsize 4"
```

Issue 'Issam' command as root to see the new resources created during this process.

Issam snippet showing only the new resource

```
...
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Online IBM.Application:db2mnt-hafs02-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs02-rs:amdha02
...
```

Add New Container by Altering the Above Table Space

Adding a new container will cause a new mount resource to be created if the container path references a mount point not yet defined in the cluster.

```
db2 "alter tablespace dms01 add (file '/hafs03/db2inst1/CONT/dms01c2' 500)"
```

Issue 'Issam' command as root to see the new resources created during this process.

Issam snippet showing only the new resource created

```
...
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Online IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs03-rs:amdha02
...
```

Change Active Log Path

To change the active log path for this database, update NEWLOGPATH database configuration parameter.

```
db2 "update database cfg for hadb using newlogpath /hafs04/db2inst1/NEWLOGPATH"
```

Even though this operation requires all applications to disconnect from the database for the change to take effect, the cluster manager will go ahead and add the mount resource 'hafs04'.

Issue 'Issam' command as root to see all the resources created during the above sequence of steps.

```

Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Online IBM.Application:db2mnt-hafs01-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Online IBM.Application:db2mnt-hafs02-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Online IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Online IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02

```

So far, we saw how adding new database objects created new mount resources on the resource group. Let us see how dropping these objects will remove the appropriate resources from the cluster group.

Drop New Container by Altering the Table Space

When a container is dropped from the existing table space, the database manager communicates with the cluster manager requesting cluster changes. To drop a container, issue the command:

```
db2 "alter tablespace dms01 drop (file '/hafs03/db2inst1/CONT/dms01c2' )"
```

db2pd -ha will not show /hafs03 mount resource in the cluster domain after the above operation.

Drop Table Space

When a table space is dropped, the cluster manager will clean up its corresponding mount resource and it will be reflected by db2pd -ha (and lssam) outputs.

To drop a table space, issue the command:

```
db2 "drop tablespace dms01"
```

If there are any external dependencies on the mount resource (that is, dependencies that are not created by the database manager or do not belong to other DB2 instances), those dependencies and the mount resource are not removed.

The lssam output and db2pd -ha will no longer show /hafs02 mount resource on the cluster domain after the above operation.

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Online IBM.Application:db2mnt-hafs01-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Online IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

Drop Database

A database drop causes the mount resource dependencies associated with the table space containers to be removed from the DB2 resource if the container paths reference a mount point no longer being used by the instance. If there are no other dependencies on the mount resources, the mount resources themselves are removed from the cluster. After this operation, only the DB2 instance level resources are left to keep the instance online on the current active node.

Then issue 'lssam' command, and you should expect to see output similar to this:

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

Database Restore

When a database is restored on an HA configured instance, the database manager will create new mount resources if the container paths referencing the mount points are not yet defined in the cluster.

To restore 'HADB' database, issue the following command.

```
db2 "restore database hadb from /shared_home/db2inst1 dbpath on /hafs01"
```

Note: Backup image for HADB database is at /shared_home/db2inst1.

You can issue db2pd -ha to find all the resources for the current cluster group. You can also query the SYSIBMADM.DBPATHS view for all the paths using the following command.

```
db2 "select substr(type,1,15) as type, substr(path,1,60) as path
from sysibmadm.dbpaths"
```

SYSIBMADM.DBPATHS output after a database restore

TYPE	PATH
LOGPATH	/hafs04/db2inst1/NEWLOGPATH/NODE0000/
DB_STORAGE_PATH	/hafs01/
TBSP_CONTAINER	/hafs03/db2inst1/CONT/dms01c2
TBSP_CONTAINER	/hafs02/db2inst1/CONT/dms01c1
LOCAL_DB_DIRECT	/hafs01/db2inst1/NODE0000/sqlbdir/
DBPATH	/hafs01/db2inst1/NODE0000/SQL00001/

So far, we saw how database-related mount paths were discovered automatically.

Now let us demonstrate how instance-level mount paths are auto-discovered in a single partition environment.

Change DB2 Diagnostic Log Path

Update the database manager configuration parameter DIAGPATH to a different mount point using the following command.

```
db2 "Update database manager configuration using DIAGPATH
/hafs05/db2inst1/DB2DIAG"
```

The 'lssam' command shows the new mount resource created during this process.

```
...
|- Online IBM.Application:db2mnt-hafs05-rs
  |- Online IBM.Application:db2mnt-hafs05-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs05-rs:amdha02
```

In addition, db2pd -ha would show additional entries for mount point /hafs05.

Change Audit Log Path

Similarly, when you update either audit log path or archive audit data path with different mount points, the corresponding mount resources are also created in the cluster domain. If the mount points already exist in the current DB2 resource group, no new resources are created.

An example follows. If audit data path is also updated to the same mount point as DIAGPATH, then no new HA resource is created.

```
db2audit configure datapath /hafs05/db2inst1/AUDITPATH
```

However, if you change the DIAGPATH to a different mount point (for example, /hafs06), it will not remove /hafs05 from the cluster domain as audit data path has a dependency on the same mount resource.

For detailed information on the list of paths that are considered to be made HA, consult the IBM DB2 v9.5 Information Center, at

<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/index.jsp?topic=/com.ibm.db2.luw.admin.ha.doc/doc/c0052942.html>

5. Setting up an Automated Single Network HA Topology Using the db2haicu XML Mode for DB2 Instance with Pre-existing Databases

The configuration of an automated single network shared disk HA topology, as illustrated in Fig. 1, is automated using the db2haicu XML mode.

Note that steps 4.1.1 – 4.1.5 must be followed for this configuration as well.

5.1.6 HA Database Setup

Here we will see how to make an instance highly available with a pre-existing database. Let us say that the database 'HADB' was already created with the database directory on /hafs01 and storage paths on /hafs02 before we run db2haicu.

```
db2 "create database hadb on /hafs02 dbpath on /hafs01"
```

Note: /hafs01 and /hafs02 are mount points already defined on /etc/fstab.

DB2 diagnostic log was set to /hafs03 mount point:

```
db2 get dbm cfg |grep DIAG  
Diagnostic data directory path          (DIAGPATH) = /hafs03/db2inst1/DB2DIAG
```

Audit log path was set to /hafs04:

```
db2audit describe |grep Path  
Audit Data Path: "/hafs04/db2inst1/AUDITPATHLOG/"
```

In short, this is the usage of all mount points:

```
/hafs01 - Database directory path  
/hafs02 - Storage paths  
/hafs03 - DB2 diagnostic log path  
/hafs04 - Audit log path
```

5.2 Using the db2haicu XML File

A sample db2haicu XML file is shown below. It contains all the information that db2haicu needs to know in order to make an instance highly available with a pre-existing database 'HADB':

```
<DB2Cluster xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:noNamespaceSchemaLocation="db2ha.xsd" clusterManagerName="TSA"
version="1.0">

  <ClusterDomain domainName="HA_domain">
    <Quorum quorumDeviceProtocol="network" quorumDeviceName="9.26.4.5"/>

    <PhysicalNetwork physicalNetworkName="db2_public_network_0"
physicalNetworkProtocol="ip">
      <Interface interfaceName="eth0" clusterNodeName="amdha01">
        <IPAddress baseAddress="9.26.124.30" subnetMask="255.255.255.0"
networkName="db2_public_network_0"/>
      </Interface>
      <Interface interfaceName="eth0" clusterNodeName="amdha02">
        <IPAddress baseAddress="9.26.124.31" subnetMask="255.255.255.0"
networkName="db2_public_network_0"/>
      </Interface>
    </PhysicalNetwork>

    <ClusterNode clusterNodeName="amdha01"/>
    <ClusterNode clusterNodeName="amdha02"/>
  </ClusterDomain>

  <FailoverPolicy>
    <Mutual></Mutual>
  </FailoverPolicy>

  <DB2PartitionSet>
    <DB2Partition dbpartitionnum="0" instanceName="db2inst1">
      <VirtualIPAddress baseAddress="9.26.124.22" subnetMask="255.255.255.0"
networkName="db2_public_network_0"/>
      <MutualPair systemPairNode1="amdha01" systemPairNode2="amdha02" />
    </DB2Partition>
  </DB2PartitionSet>

  <HADBSet instanceName="db2inst1">
    <HADB databaseName="HADB" />
  </HADBSet>

</DB2Cluster>
```

The existing values in the preceding file can be replaced to reflect your own configuration and environment. Below is a brief description of what the different elements shown in the preceding XML file represent:

- The *<ClusterDomain>* element covers all cluster-wide information. This includes *quorum information, cluster node information, and cluster domain name.*

- The *<PhysicalNetwork>* sub-element of the ClusterDomain element includes all network information. This includes the name of the network and the network interface cards contained in it. We define our single public network using this element.
- The *<FailoverPolicy>* element specifies the failover order of the cluster nodes. Mutual is an Active/Passive pair.
- The *<DB2PartitionSet>* element covers the DB2 instance information. This includes the current DB2 instance *name*, the DB2 partition number, and the virtual IP address associated with the instance.
- The *<HADBSet>* element specifies the database name that is to be made highly available. It includes the current DB2 instance *name*.

To configure the system with db2haicu XML mode:

- 1) Log on to the machine that hosts the instance directory.
- 2) Issue the following command:
`db2haicu -f path to XML file`

At this point, the XML file will be used to configure the instance. If an invalid input is encountered during the process, db2haicu will exit with a non-zero error code. After the XML file has been processed, all mount points defined to the DB2 database manager, as described in Section 5.1.6, are also defined to the cluster manager.

Sample output from running db2haicu in XML mode

```
/shared_home/db2inst1 > db2haicu -f db2ha_sample_sharedstorage_mutual.xml  
Welcome to the DB2 High Availability Instance Configuration Utility (db2haicu).
```

You can find detailed diagnostic information in the DB2 server diagnostic log file called db2diag.log. Also, you can use the utility called db2pd to query the status of the cluster domains you create.

For more information about configuring your clustered environment using db2haicu, see the topic called 'DB2 High Availability Instance Configuration Utility (db2haicu)' in the DB2 Information Center.

db2haicu determined the current DB2 database manager instance is db2inst1. The cluster configuration that follows will apply to this instance.

db2haicu is collecting information on your current setup. This step may take some time as db2haicu will need to activate all databases for the instance to discover all paths ...

Creating domain HA_domain in the cluster ...

Creating domain HA_domain in the cluster was successful.

Configuring quorum device for domain HA_domain ...

Configuring quorum device for domain HA_domain was successful.

Adding network interface card eth0 on cluster node amdha01 to the network db2_public_network_0 ...

Adding network interface card eth0 on cluster node amdha01 to the network db2_public_network_0 was successful.

Adding network interface card eth0 on cluster node amdha02 to the network db2_public_network_0 ...

Adding network interface card eth0 on cluster node amdha02 to the network db2_public_network_0 was successful.

Adding DB2 database partition 0 to the cluster ...

Adding DB2 database partition 0 to the cluster was successful.

Adding database HADB to the cluster domain ...

Adding database HADB to the cluster domain was successful.

All cluster configurations have been completed successfully. db2haicu exiting ...

Issue the 'lssam' command as root to see the resources created during this process.

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Online IBM.Application:db2mnt-hafs01-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Online IBM.Application:db2mnt-hafs02-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Online IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Online IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

The 'db2pd -ha' command can also be issued from the instance owner ID to display the state of the resources.

```
/shared_home/db2inst1> db2pd -ha
      DB2 HA Status
Instance Information:
Instance Name           = db2inst1
Number Of Domains      = 1
Number Of RGs for instance = 1

Domain Information:
Domain Name             = HA_domain
Cluster Version        = 2.4.7.4
Cluster State          = Online
Number of nodes        = 2

Node Information:
Node Name              State
-----
amdha01                Online
amdha02                Online

Resource Group Information:
Resource Group Name    = db2_db2inst1_0-rg
Resource Group LockState = Unlocked
Resource Group OpState = Online
Resource Group Nominal OpState = Online
Number of Group Resources = 7
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Member Resource Information:
Resource Name          = db2mnt-hafs02-rs
Resource State         = Online
Resource Type          = Mount
Mount Resource Path    = /hafs02
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name          = db2mnt-hafs01-rs
Resource State         = Online
Resource Type          = Mount
Mount Resource Path    = /hafs01
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name          = db2ip_9_26_124_22-rs
Resource State         = Online
Resource Type          = IP
continues...
```



```

Resource Name           = db2mnt-hafs03-rs
Resource State          = Online
Resource Type           = Mount
Mount Resource Path     = /hafs03
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name           = db2mnt-hafs04-rs
Resource State          = Online
Resource Type           = Mount
Mount Resource Path     = /hafs04
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name           = db2_db2inst1_0-rs
Resource State          = Online
Resource Type           = DB2 Partition
DB2 Partition Number    = 0
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Resource Name           = db2mnt-shared_home-rs
Resource State          = Online
Resource Type           = Mount
Mount Resource Path     = /shared_home
Number of Allowed Nodes = 2
  Allowed Nodes
  -----
  amdha01
  amdha02

Network Information:
Network Name           Number of Adapters
-----
db2_public_network_0   2

Node Name              Adapter Name
-----
amdha01                eth0
amdha02                eth0

Quorum Information:
Quorum Name           Quorum State
-----
db2_Quorum_Network_9_26_4_5:14_36_50  Online
Fail                   Offline
Operator               Offline

```

Note that the above XML file can also be used to automate HA failover for DB2 instances without any pre-existing databases by removing the HADBSet element from the XML file.

6. Post Configuration Testing

Once the db2haicu tool is run on the active cluster node, the setup is complete, and we can take our automated HA environment for a test run. Issue the 'lssam' command, and observe the output displayed on the screen. You will see output similar to this:

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Online IBM.Application:db2mnt-hafs01-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Online IBM.Application:db2mnt-hafs02-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Online IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Online IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

Below is a brief description of the resources listed in the preceding figure:

1) DB2 instance resource group:

db2_db2inst1_0-rg

Member Resources:

- | | | |
|----|-----------------------|------------------------------------|
| a. | db2_db2inst1_0-rs | (DB2 instance) |
| b. | db2ip_9_26_124_22-rs | (virtual IP address) |
| c. | db2mnt-shared_home-rs | (Instance failover home directory) |
| d. | db2mnt-hafs01-rs | (Shared mount point) |
| e. | db2mnt-hafs02-rs | (Shared mount point) |
| f. | db2mnt-hafs03-rs | (Shared mount point) |
| g. | db2mnt-hafs04-rs | (Shared mount point) |

The resources created by db2haicu during the configuration can be in one of the following states:

Online: The resource has been started and is functioning normally.

Offline: The resource has been successfully stopped.

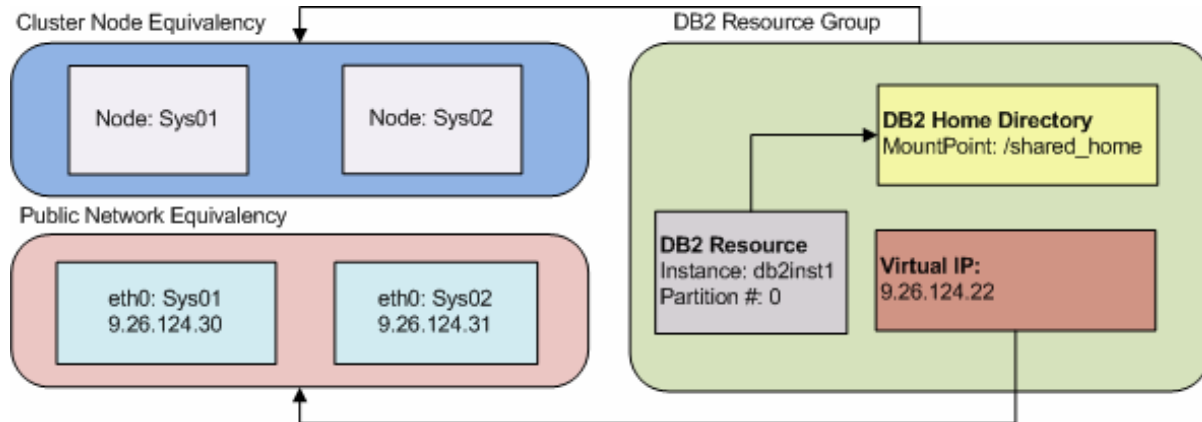
Failed Offline: The resource has malfunctioned.

For an exhaustive explanation of the possible resource states, consult Chapter 3 'OpState Attribute' in:

http://publib.boulder.ibm.com/tividd/td/ITSAFL/SC33-8272-02/en_US/PDF/HALBAU02.pdf

The relationship between the networks and the resource groups is illustrated in Fig. 2 for the single network HA topology.

Fig. 2. Resource groups created for a single network HA topology



6.1 The 'Power off' Test

Follow the instructions below to simulate active cluster node failure and understand the system state immediately following.

- 1) Turn the power off to the active cluster node (amdha01). For example, you may unplug the power cable.
- 2) Now issue the 'lssam' command to observe the state of the resources. All the resources will fail over to the other cluster node and remain online there. You should see output similar to this:

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Failed offline IBM.Application:db2_db2inst1_0-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Failed offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01 Node=Offline
  '- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Failed offline IBM.Application:db2mnt-hafs01-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Failed offline IBM.Application:db2mnt-hafs02-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Failed offline IBM.Application:db2mnt-hafs03-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Failed offline IBM.Application:db2mnt-hafs04-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Failed offline IBM.Application:db2mnt-shared_home-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-shared_home-rs:amdha02
```

The 'Failed Offline' state of all resources on amdha01 indicates a critical failure.

3) Power on the cluster node amdha01.

4) As soon as the node comes back online, 'failed offline' state will change to offline state.

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Offline IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Online IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Offline IBM.Application:db2mnt-hafs01-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Offline IBM.Application:db2mnt-hafs02-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Offline IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Offline IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Offline IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Online IBM.Application:db2mnt-shared_home-rs:amdha02
```

6.1.1 Reboot

The 'reboot' test is very similar to the power off test. For example, as the root user, you may issue the command 'reboot' at the active cluster node. This will cause an orderly shutdown of the machine at which the 'reboot' command was issued. During a reboot of the active cluster node (amdha01), resources will fail over to the other cluster node (amdha02) and remain 'online' there.

The output of the 'lssam' command would be similar to that depicted above.

Note that a power off or reboot of the non-active cluster node will not result in any perceived outage. However, the states of the resources at the non-active cluster node will be in a 'Failed Offline' state.

6.2 Network Failures

We discuss network failures in this section by simulating network interface malfunctions on the active cluster node.

- 1) Unplug the eth0 cable (public network cable) connected to the active node.
- 2) As this is a single network topology, the current active node will start to reboot, during which time the resources will fail over to the other cluster node.
- 3) Issue the 'lssam' or the 'db2pd -ha' command to examine the state of the resources. The resources should be in the following state:

```
Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Failed offline IBM.Application:db2_db2inst1_0-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Failed offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01 Node=Offline
  '- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Failed offline IBM.Application:db2mnt-hafs01-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Failed offline IBM.Application:db2mnt-hafs02-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Failed offline IBM.Application:db2mnt-hafs03-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Failed offline IBM.Application:db2mnt-hafs04-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Failed offline IBM.Application:db2mnt-shared_home-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-shared_home-rs:amdha02
```

Note: The node that is being rebooted will remain offline until the network cable is plugged back in.

To recover from this state:

- 1) Plug the eth0 cable back in.
- 2) Repeatedly issue the 'lssam' or the 'db2pd -ha' command and observe the system resources assume their normal states.

6.3 Disk I/O Failure

We discuss disk I/O failures in this section by simulating disk controller malfunctions on the active cluster node.

- 1) Gently pull the fiber-optic cable from the GBIC (Gigabit Interface Converter) on the RAID controller on the active cluster node (amdha01).
- 2) The disk mount will be detected as Offline, and the cluster manager will start to run a failover operation.
- 3) Issue the 'lssam' command to observe the state of the resources. All the resources will failover to the other cluster node and remain online there. You should see output similar to this:

```

Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Failed offline IBM.Application:db2_db2inst1_0-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Failed offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01 Node=Offline
  '- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Failed offline IBM.Application:db2mnt-hafs01-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Failed offline IBM.Application:db2mnt-hafs02-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Failed offline IBM.Application:db2mnt-hafs03-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Failed offline IBM.Application:db2mnt-hafs04-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Failed offline IBM.Application:db2mnt-shared_home-rs:amdha01 Node=Offline
  '- Online IBM.Application:db2mnt-shared_home-rs:amdha02

```

The 'Failed Offline' state of all resources indicates a critical failure.

To recover from this state:

- 1) Connect the fiber-optic cable back to the installed GBIC.
- 2) Repeatedly issue the 'lssam' or the 'db2pd -ha' command and observe the system resources assume their normal states.

6.4 DB2 Failures

A. Killing the DB2 Instance:

- 1) Issue the 'db2_kill' command on the machine where the resource is online.
- 2) Issue the 'lssam' or the 'db2pd -ha' command to examine the resources. You will see output similar to this:


```

Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Pending online IBM.Application:db2_db2inst1_0-rs
  |- Pending online IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Online IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Online IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02

```

3) The instance resource will be in the 'Pending Online' state.

4) Repeatedly issue the 'Issam' or the 'db2pd -ha' command. The cluster manager will automatically start the DB2 instance. This will result in the 'Pending Online' state changing to the 'Online' state.

B. Failing the DB2 Instance:

1) Log on to the machine where instance resource is online and rename the db2start executable:

```
/shared_home/db2inst1% mv $HOME/sql/lib/adm/db2star2 db2star2.mv
```

2) Issue the 'db2_kill' command.

3) The DB2 resource will assume the 'Pending Online' state. The cluster manager will try to start the DB2 instance indefinitely, but will fail because of the missing executable.

4) A timeout will occur, and any further start attempts on the DB2 resource will stop. The cluster manager will start to run a failover operation. This will be indicated by the 'Pending Online' state changing into the 'Failed Offline' state, as illustrated in the figure below:

Note: It might take 4-5 minutes for the DB2 resource timeout to occur

```

Pending online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Pending online IBM.Application:db2_db2inst1_0-rs
|   |- Failed offline IBM.Application:db2_db2inst1_0-rs:amdha01
|   |- Pending online IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
|   |- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
|   |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
|   |- Offline IBM.Application:db2mnt-hafs01-rs:amdha01
|   |- Online IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
|   |- Offline IBM.Application:db2mnt-hafs02-rs:amdha01
|   |- Online IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
|   |- Offline IBM.Application:db2mnt-hafs03-rs:amdha01
|   |- Online IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
|   |- Offline IBM.Application:db2mnt-hafs04-rs:amdha01

```

- 5) Then the cluster manager will attempt to start the instance on the failed over node. This is indicated by 'pending online' state for the instance resource on.
- 6) Log on to the machine where the instance resource is available (/shared_home file system), and rename the executable back to its original name. This will cause the cluster manager to start the instance successfully and the 'pending online' state will be changed to 'online' state:

```
/shared_home/db2inst1% mv $HOME/sql/lib/adm/db2star2.mv db2star2
```

- 7) Note that the state for amdha01 would still be 'Failed offline'. Issue the following command with root authority on the node where the instance resource is available (/shared_home file system) to reset the 'failed offline' flag.

```
resetsrc -s "Name like '<Standby DB2 instance resource name>' AND
NodeNameList in {'<standby node name>'} IBM.Application
```

In our case, the command will look like this:

```
resetsrc -s "Name like 'db2_db2inst1_0-rs' AND NodeNameList in {'amdha01'}"
IBM.Application
```

This command will reset the 'Failed Offline' flag for the instance resource on amdha01.

The 'lssam' output should look similar to this:

```

Online IBM.ResourceGroup:db2_db2inst1_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_0-rs
  |- Offline IBM.Application:db2_db2inst1_0-rs:amdha01
  '- Online IBM.Application:db2_db2inst1_0-rs:amdha02
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs
  |- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01
  '- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02
|- Online IBM.Application:db2mnt-hafs01-rs
  |- Offline IBM.Application:db2mnt-hafs01-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs01-rs:amdha02
|- Online IBM.Application:db2mnt-hafs02-rs
  |- Offline IBM.Application:db2mnt-hafs02-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs02-rs:amdha02
|- Online IBM.Application:db2mnt-hafs03-rs
  |- Offline IBM.Application:db2mnt-hafs03-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs03-rs:amdha02
|- Online IBM.Application:db2mnt-hafs04-rs
  |- Offline IBM.Application:db2mnt-hafs04-rs:amdha01
  '- Online IBM.Application:db2mnt-hafs04-rs:amdha02
'- Online IBM.Application:db2mnt-shared_home-rs
  |- Offline IBM.Application:db2mnt-shared_home-rs:amdha01
  '- Online IBM.Application:db2mnt-shared_home-rs:amdha02

```

6.5 Manual Instance Control (db2stop, db2start)

For various reasons such as planned maintenance, you might need to stop and start the instance.

1) Issue the 'db2stop' command on the node where all the resources are online. The following error will be encountered and the instance will *not* be stopped:

```

/shared_home/db2inst1> db2stop
01/31/2008 12:27:00  0  0  SQL1025N The database manager was not stopped because
databases are still active.
SQL1025N The database manager was not stopped because databases are still active.

```

2) Now issue the 'db2stop force' command on the instance. The command will complete successfully and the instance will be stopped.

```

/shared_home/db2inst1> db2stop force
01/31/2008 12:27:40  0  0  SQL1064N DB2STOP processing was successful.
SQL1064N DB2STOP processing was successful.

```

This will cause the cluster manager to lock the instance resource group. The resource group is not brought offline because that would cause the cluster manager to unmount the file systems associated with the mount resources in the resource group.

Note: Any failure on the node will not cause the cluster manager to initiate a failover at this point, as the resource group is in 'lock' state.

The 'Issam' output would show something similar to this:

```
Pending online IBM.ResourceGroup:db2_db2inst1_0-rg Request=Lock  
Nominal=Online
```

```
|- Offline IBM.Application:db2_db2inst1_0-rs  
  |- Offline IBM.Application:db2_db2inst1_0-rs:amdha01  
  '- Offline IBM.Application:db2_db2inst1_0-rs:amdha02  
|- Online IBM.ServiceIP:db2ip_9_26_124_22-rs  
  |- Online IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha01  
  '- Offline IBM.ServiceIP:db2ip_9_26_124_22-rs:amdha02  
|- Online IBM.Application:db2mnt-hafs01-rs  
  |- Online IBM.Application:db2mnt-hafs01-rs:amdha01  
  '- Offline IBM.Application:db2mnt-hafs01-rs:amdha02  
|- Online IBM.Application:db2mnt-hafs02-rs  
  |- Online IBM.Application:db2mnt-hafs02-rs:amdha01  
  '- Offline IBM.Application:db2mnt-hafs02-rs:amdha02  
|- Online IBM.Application:db2mnt-hafs03-rs  
  |- Online IBM.Application:db2mnt-hafs03-rs:amdha01  
  '- Offline IBM.Application:db2mnt-hafs03-rs:amdha02  
|- Online IBM.Application:db2mnt-hafs04-rs  
  |- Online IBM.Application:db2mnt-hafs04-rs:amdha01  
  '- Offline IBM.Application:db2mnt-hafs04-rs:amdha02  
'- Online IBM.Application:db2mnt-shared_home-rs  
  |- Online IBM.Application:db2mnt-shared_home-rs:amdha01  
  '- Offline IBM.Application:db2mnt-shared_home-rs:amdha02
```

The instance resource group will remain 'locked' until a successful 'db2start' command is issued.

3) Start the DB2 instance manually by issuing the 'db2start' command, once it completes, the lock will be removed. Henceforth, the cluster manager will detect any failures, and the instance resource group will be either failed over or restarted (as warranted) by the cluster manager.

7. Maintenance

7.1 Disabling High Availability

To disable the HA configuration for a particular instance, the 'db2haicu -disable' command can be used. After issuing this command, the system will not respond to any failures, and all resource groups for the instance will be locked. Any maintenance work can be performed in this state without risk of cluster manager intervention.

To enable HA, just issue the 'db2haicu' command again, and select Yes when prompted to continue.

Disabling High Availability for an HA DB2 instance

Welcome to the DB2 High Availability Instance Configuration Utility (db2haicu).

You can find detailed diagnostic information in the DB2 server diagnostic log file called db2diag.log. Also, you can use the utility called db2pd to query the status of the cluster domains you create.

For more information about configuring your clustered environment using db2haicu, see the topic called 'DB2 High Availability Instance Configuration Utility (db2haicu)' in the DB2 Information Center.

db2haicu determined the current DB2 database manager instance is db2inst1. The cluster configuration that follows will apply to this instance.

db2haicu is collecting information on your current setup. This step may take some time as db2haicu will need to activate all databases for the instance to discover all paths ...

Are you sure you want to disable high availability (HA) for the database instance db2inst1. This will lock all the resource groups for the instance and disable the HA configuration parameter. The instance will not failover if a system outage occurs while the instance is disabled. You will need to run db2haicu again to enable the instance for HA. Disable HA for the instance db2inst1? [1]

1. Yes

2. No

1

Disabling high availability for instance db2inst1 ...

Locking the resource group for DB2 database partition 0 ...

Locking the resource group for DB2 database partition 0 was successful.

Disabling high availability for instance db2inst1 was successful.

All cluster configurations have been completed successfully. db2haicu exiting ...

7.2 Db2haicu Maintenance Mode

When a system is already configured for High Availability, db2haicu runs in maintenance mode. Typing 'db2haicu' on the node will produce the menu shown below. This menu can be used to carry out various maintenance tasks and change any cluster-manager-specific, DB2-specific or network-specific values configured during the initial setup. Any time you want to close db2haicu without entering any value, you can do so by pressing Ctrl+C.

Db2haicu Maintenance Mode

```
/shared_home/db2inst1% db2haicu
Welcome to the DB2 High Availability Instance Configuration Utility (db2haicu).

You can find detailed diagnostic information in the DB2 server diagnostic log file called
db2diag.log. Also, you can use the utility called db2pd to query the status of the cluster
domains you create.

For more information about configuring your clustered environment using db2haicu, see the
topic called 'DB2 High Availability Instance Configuration Utility (db2haicu)' in the DB2
Information Center.

db2haicu determined the current DB2 database manager instance is db2inst1. The cluster
configuration that follows will apply to this instance.

db2haicu is collecting information on your current setup. This step may take some time as
db2haicu will need to activate all databases for the instance to discover all paths ...
When you use db2haicu to configure your clustered environment, you create cluster domains.
For more information, see the topic 'Creating a cluster domain with db2haicu' in the DB2
Information Center. db2haicu is searching the current machine for an existing active cluster
domain ...
db2haicu found a cluster domain called HA_domain on this machine. The cluster configuration
that follows will apply to this domain.

Select an administrative task by number from the list below:
  1. Add or remove cluster nodes.
  2. Add or remove a network interface.
  3. Add or remove a highly available database.
  4. Add or remove a mount point.
  5. Add or remove an IP address.
  6. Add or remove a non-critical path.
  7. Move DB2 database partitions and HADR databases for scheduled maintenance.
  8. Change failover policy for this instance.
  9. Create a new quorum device for the domain.
 10. Destroy the domain.
 11. Exit.
Enter your selection:
```

Add a New Mount Resource

Mount resources can also be manually added using the db2haicu tool. If the mount resource already exists in the current DB2 resource group, the database manager simply increments its internal use count.

To add an instance-level mount point, select option 4 of the db2haicu maintenance mode menu.

```
...
Enter your selection:
4
Do you want to add or remove mount points to or from the domain? [1]
1. Add
```

2. Remove

1

Enter the full path of the mount you wish to add to the cluster:

/hafs05

Adding mount point /hafs05 to the cluster ...

Adding mount point /hafs05 to the cluster was successful.

'Issam' would now show a new mount resource added to the DB2 resource group

```
'- Online IBM.Application:db2mnt-hafs05-rs
  |- Online IBM.Application:db2mnt-hafs05-rs:amdha01
  '- Offline IBM.Application:db2mnt-hafs05-rs:amdha02
```

If you try to add a mount point that is already part of a cluster, it will display the following message.

...

Enter the full path of the mount you wish to add to the cluster:

/hafs05

Adding mount point /hafs05 to the cluster ...

The mount point /hafs05 cannot be added because it is already being used in the cluster.

If you try to add a mount point that is not mounted yet or add a mount path that is mounted on the root (/) directory, db2haicu will not be able to resolve it and will display the following message.

...

Enter the full path of the mount you wish to add to the cluster:

/hafs06

Adding mount point /hafs06 to the cluster ...

Either db2haicu could not resolve the mount point / or the mount point / is on a file system that is not eligible to be added to the cluster.

To illustrate, issue the 'df' command at this node. You should see output similar to this:

```
root@amdha01:/# df -k /hafs06
Filesystem      1K-blocks  Used Available Use% Mounted on
/dev/sda1       63258244 12466732 47578112 21% /
```

As you can see, /hafs06 is just a simple directory on the root (/) directory and is thus ineligible for addition to the list of failover mount paths.

It is very important that you define the mount points on /etc/fstab on all cluster nodes and that you mount them only on the current active node prior to using them on a cluster domain. For a list of mount points that cannot be added manually to the cluster (such as the / root mount point), consult Section 8.1.

Remove a Mount Resource

Option 4 is used to remove an existing mount point (for example /hafs05).

```
...
Enter your selection:
4
Do you want to add or remove mount points to or from the domain? [1]
1. Add
2. Remove
2
Enter the full path of the mount you want to remove from the cluster:
/hafs05
Removing mount point /hafs05 from the cluster ...
Removing mount point /hafs05 from the cluster was successful.
```

Note that only mount points added manually using the db2haicu tool (as above), can be removed with the db2haicu option to remove a mount point. Mount points that are associated with HA databases and managed as part of the database operations cannot be removed using this operation. They are automatically removed once no longer required by the database.

Add a Noncritical Path

In addition to automatically discovered paths, the user can add arbitrary paths to the cluster. Not all paths associated with a database are required for its operation. For example, if the file for an inactive event monitor does not failover with an instance it will not impede the instance's functionality in any way. Paths such as this are expressed as noncritical path objects. Add file systems to the noncritical path list that you know to be local file systems only; and also add any other file systems that should not be failed over under any circumstances. If you have such a path, then the db2haicu option to add a noncritical path can be used.

```
...
Enter your selection:
6
Do you want to add or remove non-critical paths? [1]
1. Add
2. Remove
1
Enter the full path of the mount to be made non-critical:
/hafs06
Adding path /hafs06 to the non-critical path list ...
Adding path /hafs06 to the non-critical path list was successful.
```

In addition, any mount point entries in /etc/fstab that are known to not require failover (such as local mount paths not shared with any other machine in the cluster) should be added to the noncritical path list.

Remove a Highly Available Database

When a new database is created on an Automated Cluster-controlled HA configuration, it will automatically be made highly available. To make it non-HA, you can select option 3 in db2haicu maintenance mode.

For example, database 'sample' created with the database directory on /hafs05:

Selecting option 3 displays all the existing highly available databases for this instance:

```
...
The following databases are configured for high availability.
  Database: HADB
  Database: SAMPLE
Do you want to add or remove highly available databases to or from the cluster? [1]
1. Add
2. Remove
```

Select option 2 and enter the database name when prompted.

```
...
2
Enter the name of the database you would like to remove from the cluster.
sample
```

Db2haicu warns that the database paths will go offline for the database that is being made non-HA.

```
...
You have chosen to remove database SAMPLE from the cluster. This will cause the
database paths to go offline, which will unmount all file systems used to store their
data. Are you sure you wish to proceed? [2]
1. Yes
2. No
```

Select option 1 to remove the database from the cluster.

```
...
1
Removing database SAMPLE from the cluster ...
Removing database SAMPLE from the cluster successful.
Do you want to remove another database? [1]
1. Yes
2. No
```

Select option 2 to exit from db2haicu.

'Issam' would still show hafs05 mount resource in the resource group because the other HA database 'HADB' still has dependency on this mount resource.

Now, in the above maintenance task, 'sample' database had database directory and table space containers on /hafs05 and did not have any dependencies on other mount resources.

Add/Remove IP Address

Sometimes, you might want to change the virtual IP address for the cluster. To do so, select option 5 to first remove it and then add it with the new virtual IP address.

```
...
Enter your selection:
5
Do you want to add or remove IP addresses to or from the cluster? [1]
1. Add
2. Remove
2
Enter the virtual IP address to be removed from the cluster:
9.26.124.22
Are you sure you want to remove the virtual IP address 9.26.124.22 from the cluster?
[2]
1. Yes
2. No
1
Removing virtual IP address 9.26.124.22 from the cluster ...
Removing virtual IP address 9.26.124.22 from the cluster was successful.
```

To add a new virtual IP address, select option 5 and enter values for all the prompts.

```
...
Enter your selection:
5
Do you want to add or remove IP addresses to or from the cluster? [1]
1. Add
2. Remove
1
Enter the virtual IP address:
9.26.124.83
Enter the subnet mask for the virtual IP address 9.26.124.83: [255.255.255.0]
255.255.255.0
Select the network for the virtual IP 9.26.124.83:
1. db2_public_network_0
Enter selection:
1
Adding virtual IP address 9.26.124.83 to the domain ...
Adding virtual IP address 9.26.124.83 to the domain was successful.
```

'Issam' would show a new virtual IP resource.

Note that if the Virtual IP address is changed, it is likely that the client reroute settings should be updated as well to reflect the new IP address in use.

Moving the Database Partition for a Scheduled Maintenance

There might be situations when a DBA or system administrator wants to perform a manual failover to other cluster nodes.

To accomplish this, you should use the TSA command `rgreq`. The `db2haicu` move option should only be used in cases where a non-shared home directory is used.

Log on to the machine and type in the following command to perform a manual failover.

```
rgreq -o move -n <nodename> <DB2 instance resource group>
```

In our case, the command will look like this:

```
rgreq -o move -n amdha01 db2_db2inst1_0-rg
```

Once the failover has been completed successfully, the `'ssam'` or the `'db2pd -ha'` commands will reflect the changes. The virtual IP address will also be moved to the new active node as part of the failover process, and any associated mount points will be moved as well.

You may also select the option to destroy a domain. If you select option 10 in `db2haicu` maintenance mode to destroy the domain, it will fail with the following message if any resource groups were present in the domain.

```
...  
Enter your selection:  
10  
The following resource groups were found in the domain HA_domain:  
Resource Group: db2_db2inst1_0-rg  
You will need to cleanup all cluster objects defined in the domain HA_domain, before  
the domain can be removed.
```

Generally, `'db2haicu -delete'` should be run to remove all cluster objects associated with this instance. The `db2haicu` option to destroy the domain should only be used in cases where no cluster objects are currently present within the domain.

8. Problem Determination

8.1 Restrictions and Supportability

1. File systems restrictions

The database manager automatically determines whether a file system used by DB2 needs to be defined to the cluster as a mount resource or not based on the file system type. There are some restrictions on what type of file systems can or cannot be made highly available.

Only file systems that can be made highly available are local file systems:

- jfs
- jfs2
- ext2
- ext3
- reiserfs
- xfs

These file systems are not made highly available:

- Shared file systems such as NFS
- Clustered file systems such as GPFS, CFS
- Any file system mounted on root (/) directory

2. No Federated support

CREATE/DROP WRAPPER statements do not add or remove cluster manager mount resources for wrapper library paths.

3. No db2relocatedb support

No explicit support for db2relocatedb is provided in this solution. The db2haicu utility has to be rerun to create the mount resources for the new database paths and remove mount resources no longer being used.

4. If a multiple domain XML file is provided to db2haicu the portions that apply to the domain running on the local node will be the only ones processed.

5. The DB2 High Availability feature does not support multiple instances in the same resource group. Furthermore, DB2 resource groups should not depend on each other. Any such relationships may result in unwanted and unpredictable behavior between the cluster manager and DB2.

6. If a database had to be manually cleaned up by removing all the storage paths and the database directory, the cluster manager would not remove the appropriate mount resources. The db2haicu tool must be used to remove the highly available database (option 3), or to remove the entire cluster and recreate it (option 1).

8.2 Db2haicu '-delete' Option

Db2haicu can also be run with the '- delete' option. This option removes a system's entire HA configuration and deletes all resources in the cluster for the instance in question. If no other instance is using the domain at the time, the domain is deleted as well.

It is recommended to run db2haicu with the delete option on an instance before it is made highly available. This makes sure that we are starting from scratch and not building on top of leftover resources.

For example, when running db2haicu with an XML file, any invalid attribute in the file will cause db2haicu to exit with a non-zero error code. However, before db2haicu is run again with the corrected XML file, one can run the `-delete` option to make sure that any temporary resources created during the initial run are cleaned up.

8.3 Hostnames specified to db2haicu

Note that the cluster nodenames specified to db2haicu should match exactly the value returned from the `'hostname'` command issued at the machine in question. For example, if using db2haicu to create a two node domain containing the machines `'node01'` and `'node02'`, logon to each of `'node01'` and `'node02'` and record the value returned by the `'hostname'` command. It is exactly this value that should be supplied to db2haicu when prompted to enter the host name of a machine to add to the domain.

8.4 Broadcast pings and netmon.cf

Broadcast pings should be enabled between all nodes in the cluster domain. In cases where broadcast pings are disabled, then a netmon.cf file must be setup for the domain, as described here:

http://publib.boulder.ibm.com/infocenter/clresctr/vxrx/index.jsp?topic=/com.ibm.cluster.rct.doc/rsct_14010/b15dia00/b15dia0031.html

8.5 The `'syslog'` and the DB2 Server Diagnostic Log File (`db2diag.log`)

The DB2 High Availability (HA) feature provides some diagnostics through the db2pd utility. The db2pd `-ha` option is independent of any other option specified to db2pd.

The information contained in the db2pd output for the HA feature is retrieved from the cluster manager. The DB2 HA feature can only communicate with the active cluster domain on the cluster node where it is invoked. All options will output the name of the active cluster domain to which the local cluster node belongs, as well as the domain's current state.

For debugging and troubleshooting purposes, the necessary data is logged in two files: the syslog, and the DB2 server diagnostic log file (`db2diag.log`).

Any DB2 instance and database-related errors are logged in the `db2diag.log` file. The default location of this file is `$HOME/sqllib/db2dump/db2diag.log`, where `$HOME` is the DB2 instance home directory. This location can be changed with the following command:

```
db2 update dbm cfg using DIAGPATH <new diagnostic log location>
```

In addition, there are 5 diagnostic levels that can be set to control the amount of data logged. These range from 0-4, where level 0 indicates the logging of only the most critical errors, and level 4 indicates the maximum amount of logging possible. Diagnostic level 3 is

recommended to be set on both the primary and the standby instances. The command to change the diagnostic level of an instance is:

```
db2 update dbm cfg using DIAGLEVEL <Diagnostic level number>
```

The syslog or the system log is used to log information from the cluster manager. The location and level of the logged messages can be defined in the /etc/syslog.conf file. For example, the following statement in syslog.conf will enable logging of all messages in the /var/log/messages file:

```
*.* -/var/log/messages
```

8.6 Data Capture/Gathering for TSA/RSCT

The labor-intensive task of capturing data can be significantly reduced by using a script called 'getsadata.sh'. This data-capture script also runs a utility called ctsnap that captures data specific to the RSCT component. It is used to gather detailed information about the domain and the current SA policy, including the policy scripts. It will also dump adapter details, format and collect traces, run and collect ctsnap data, and syslogs (AIX) and /var/log/messages files on Linux.

More TSA problem determination information can be obtained from the following link:

<http://www-1.ibm.com/support/docview.wss?rs=820&context=SSRM2X&uid=swg21285496>

9. Summary

We have demonstrated how the DB2 Version 9.5 Integrated HA feature can be used to simplify configuration, management, and usage of High Availability for a single partition DB2 instance shared between two machines.

For further information, consult the following URLs:

<http://www.ibm.com/software/tivoli/products/sys-auto-multi/>

<http://publib.boulder.ibm.com/tividd/td/IBMTivoliSystemAutomationforMultiplatforms2.3.html>

<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/index.jsp?topic=/com.ibm.db2.luw.admin.ha.doc/doc/c0052942.html>



© Copyright IBM Corporation, 2008
All Rights Reserved.
IBM Canada
8200 Warden Avenue
Markham, ON
L6G 1C7
Canada

Neither this document nor any part of it may be copied or reproduced in any form or by any means or translated into another language, without the prior consent of the above-mentioned copyright owner.

IBM makes no warranties or representations with respect to the content hereof and specifically disclaims any implied warranties of merchantability or fitness for any particular purpose. IBM assumes no responsibility for any errors that may appear in this document. The information contained in this document is subject to change without any notice. IBM reserves the right to make any such changes without obligation to notify any person of such revision or changes. IBM makes no commitment to keep the information contained herein up to date.

The information in this document concerning non-IBM products was obtained from the supplier(s) of those products. IBM has not tested such products and cannot confirm the accuracy of the performance, compatibility or any other claims related to non-IBM products. Questions about the capabilities of non-IBM products should be addressed to the supplier(s) of those products.

IBM, the IBM logo, AIX, DB2, Tivoli, and TotalStorage are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

Windows is a trademark of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates.