

**Information Management** software



# **Continuous access to Read on Standby databases using Virtual IP addresses**

**January 2011**

## Table of contents

<b>1</b>	<b>Executive summary .....</b>	<b>1</b>
1.1	<i>HADR system configuration .....</i>	<i>1</i>
1.2	<i>System pre-requisites .....</i>	<i>1</i>
<b>2</b>	<b>Setup and configuration process.....</b>	<b>2</b>
2.1	<i>Configuring the HADR automation using db2haicu.....</i>	<i>2</i>
2.2	<i>Configuring the ROS HADR database.....</i>	<i>2</i>
2.3	<i>Creating the ROS VIP for the Standby HADR database .....</i>	<i>3</i>
2.4	<i>Configuring the clients .....</i>	<i>7</i>
<b>3</b>	<b>Conclusion and summary.....</b>	<b>8</b>
<b>4</b>	<b>Additional information.....</b>	<b>8</b>
<b>5</b>	<b>Contributors .....</b>	<b>8</b>
<b>Notices.....</b>		<b>9</b>
	<i>Trademarks .....</i>	<i>10</i>

# 1 Executive summary

---

The DB2® High Availability and Disaster Recover (HADR) feature is a database replication method that provides a high availability solution for both partial and complete site failures. HADR protects against data loss by replicating data changes from a source database, called the Primary, to a target database, called the Standby.

The HADR reads on standby (HADR ROS) feature, allows applications to access the current Standby databases for reads. With the HADR ROS feature, read/write applications can access the HADR Primary database, and read-only applications can access either the HADR Primary or the HADR Standby database. This enables you to offload some of the read-only workloads running on the Primary to the Standby.

This paper shows how you can use the HADR ROS feature in conjunction with Virtual IP addresses to ensure continued, automatic connectivity for read-only clients.

## 1.1 HADR system configuration

---

The HADR configuration used in this paper consists of two machines named grom01 and grom02. The two HADR instances have the same name, db2inst1 (db2inst1 is hosted by grom01, and the other db2inst1 instance is hosted by grom02.) The HADR database that is replicated between these two instances is named GSDB. In our scenario, we assume that there is a read/write workload which needs to be directed at all times to the current location of the HADR Primary database. We also assume that there is a read-only workload that should be directed at all times to the read-only copy (the Standby node hosting the HADR Standby ROS database.)

The routing of the read/write transaction workload is achieved through the usage of a virtual IP Address (VIP.) This VIP is collocated at all times (that is, the VIP is hosted on the same physical host) with the current location of the Primary HADR database. The same concept applies to the Standby HADR database. This paper explains how to create a second VIP that is collocated at all times with the current location of the Standby HADR database (the ROS VIP.) The paper then describes how to catalog the read-only copy of the database on the clients using the ROS VIP and how to redirect their read-only workloads to the Standby HADR database copy, regardless of the node which currently hosts that Standby HADR copy.

## 1.2 System pre-requisites

---

The information contained herein is applicable to DB2 V9.7 Fix Pack 1 or later, with the HADR reads on Standby feature enabled. The information is applicable to the AIX®, Linux®, and Solaris operating systems.

## 2 Setup and configuration process

---

In the following sections, you will perform the following tasks:

1. Configure the HADR automation using the db2haicu tool.
2. Configure the Read on Standby (ROS) HADR database.
3. Create the ROS Virtual IP address (VIP) for the Standby HADR database.
4. Catalog the clients against the ROS VIP.

### 2.1 Configuring the HADR automation using db2haicu

---

To configure the automated HADR failover, perform the following tasks:

1. Setup the databases on the two systems as an HADR pair. For more information, see <http://publib.boulder.ibm.com/infocenter/db2luw/v9r7/topic/com.ibm.db2.luw.admin.ha.doc/doc/t0011725.html>.
2. Confirm that the HADR pair is in peer state using the **db2pd** `-hadr -db dbname` command (in our case, as instance owner, use the command `db2pd -hadr -db GSDB`).
3. Create the initial cluster using the db2haicu tool. Follow the instructions in the following document for this part of the configuration: <http://www.ibm.com/developerworks/data/library/long/dm-0907hadrd2haicu/>

**Note:** Any configuration supported by the db2haicu tool can be the basis for adding the ROS VIP address. Our example assumes a single network configuration automated with db2haicu, with a VIP address created by db2haicu and collocated with the Primary HADR database. To this configuration, we will add a VIP collocated with the HADR Standby database.

### 2.2 Configuring the ROS HADR database

---

To configure the ROS HADR database, perform the following steps:

1. Disable the automatic database maintenance on both HADR databases using the following command:

```
db2 update db cfg for GSDB using AUTO_MAINT OFF
```

2. Set the following registry variables on both systems using the following commands:

```
db2set DB2_HADR_ROS=YES
db2set DB2_STANDBY_ISO=UR
db2set DB2COMM=tcPIP
```

3. Enable the TCPIP listener on port 60001 (you can use any port, as long as it is unused and the same on each instance):

```
db2 update dbm cfg using SVCENAME 60001
```

4. Stop and restart the DB2 instances for the changes to take effect.

5. Confirm that the HADR pair is in peer state before proceeding to create the ROS VIP using the **db2pd** command (in our case, use `db2pd -hadr -db GSDB`).

## 2.3 Creating the ROS VIP for the Standby HADR database

---

To create the ROS VIP for the Standby HADR database, perform the following steps:

1. On both systems, as root, create a script called `hadrV97_monitor_shadow.ksh`. This script must meet the following requirements:
  - It must be executable by root.
  - It must be saved in the `/usr/sbin/rsct/sapolicies/db2/` directory on each machine.
  - It must have the same permissions as the `hadrV97_monitor.ksh` script which should already reside in the same directory.

The `hadrV97_monitor_shadow.ksh` script should have the following contents:

```
#!/bin/ksh -p
# THIS SCRIPT REQUIRES THE SAME INPUT PARAMETERS AS THE hadrV97_monitor.ksh SCRIPT
# For example: hadrV97_monitor_shadow.ksh $1 $2 $3
# $1 must be instance name at 'THIS' node
# $2 must be instance name at 'THE OTHER' node
# $3 must be HADR DBNAME

VIPhere=2

su - ${1?} -c "db2pd -hadr -db ${3?}" | egrep '(^Primary|^Standby)' | awk '{print $1 " " $2}' | read Role State

# ROS VIP is up when Standby is in Peer state
if [[ X${Role} == 'XStandby' && X${State} == 'XPeer' ]]; then
    VIPhere=1
fi

# ROS VIP is also up when Standby is in Disconnected Peer state
if [[ X${Role} == 'XStandby' && X${State} == 'XDisconnectedPeer' ]]; then
    VIPhere=1
fi

# ROS VIP is also up when Standby is in Remote Catchup state
if [[ X${Role} == 'XStandby' && X${State} == 'XRemoteCatchup' ]]; then
    VIPhere=1
fi

# ROS VIP is also up when Standby is in Remote Catchup Pending state
if [[ X${Role} == 'XStandby' && X${State} == 'XRemoteCatchupPending' ]]; then
    VIPhere=1
fi

logger -i -p notice -t $0 "$1,$2,$3: Role=${Role}, State=${State}, Returning $VIPhere"
return $VIPhere
```

2. Login as root on either system (grom01 or grom02 in our case.) In your root session, ensure that the CT\_MANAGEMENT\_SCOPE variable is set to 2. You can verify this using the following command (the # symbol represents the shell prompt and does not need to be typed in):

```
# export CT_MANAGEMENT_SCOPE=2
```

**Note:** You must be logged in as root to run the remainder of the commands. Prior to running each command, ensure that CT\_MANAGEMENT\_SCOPE is set to 2 using the command `echo $CT_MANAGEMENT_SCOPE`.

3. Create two Tivoli SA MP resources (hadr\_shadow\_01 and hadr\_shadow\_02) to monitor the HADR state by running the following commands as root on either one of the nodes (grom01 or grom02) in the cluster domain.

In the following commands, make the following substitutions:

- Replace 'grom01' with the name of your first host.
- Replace 'grom02' with the name of your second host.
- Replace 'db2inst1' with your instance name on the first host. If the instance name on the standby host is different than the instance name on the primary host, replace the second instance name with the standby host's instance name.
- Replace 'GSDB' with the name of your database.

Press Enter at the end of each line. The backslash (\) character enables the shell to interpret multiple lines as a single command:

```
# mkrsrc IBM.Application Name="hadr_shadow_01" ResourceType=0 NodeNameList="{grom01}" \
UserName=root StartCommand="/bin/true" StopCommand="/bin/true" \
MonitorCommand="/usr/sbin/rsct/sapolicies/db2/hadrV97_monitor_shadow.ksh db2inst1 db2inst1 GSDB" \
MonitorCommandPeriod=5 MonitorCommandTimeout=29 RunCommandsSync=1
```

```
# mkrsrc IBM.Application Name="hadr_shadow_02" ResourceType=0 NodeNameList="{grom02}" \
UserName=root StartCommand="/bin/true" StopCommand="/bin/true" \
MonitorCommand="/usr/sbin/rsct/sapolicies/db2/hadrV97_monitor_shadow.ksh db2inst1 db2inst1 GSDB" \
MonitorCommandPeriod=5 MonitorCommandTimeout=29 RunCommandsSync=1
```

4. Create an equivalency of the hadr\_shadow\_01 and hadr\_shadow\_02 resources by running the following command as root. Replace 'grom01' with the name of your first hostname, and replace 'grom02' with the name of your second hostname:

```
# mkequ -p NoControl hadr_shadow_equiv \ IBM.Application:hadr_shadow_01:grom01,hadr_shadow_02:grom02
```

5. Create the VIP resource for the Standby HADR database. In the command that follows, replace 'grom01' with the name of your first hostname and replace 'grom02' with the name of your second hostname. In addition, you must select an IP address. Replace '9.26.124.81' and '255.255.255.0' with the IP address and netmask respectively that you will be associating with the Standby HADR database.

The IP address that you select must be reserved by the network administrator exclusively for the use of this configuration, and must not already be assigned to any existing computer on the network. Verify that the IP address selected for the ROS VIP is not pingable, reachable, or present on the network in any way. In addition, this IP address must be routable from each machine. This IP address cannot be used by any other database in the cluster; each database must have two unique IP addresses assigned to

it, one for the Primary HADR database and one for the Standby HADR database. The netmask must match exactly the netmask that is currently in use on each of the machines. If you are unclear about the networking and VIP requirements in your enterprise, discuss the choice of the desired VIP with your network administrators.

```
# mkrsrc IBM.ServiceIP ProtectionMode=1 Name=ros_ip_9_26_124_81-rs \  
IPAddress=9.26.124.81 NetMask=255.255.255.0 NodeNameList="{grom01,grom02}"
```

**Note:** The name 'ros\_ip\_9\_26\_124\_81-rs' is arbitrary but it is suggested that you follow a similar naming convention using the actual IP address that you choose. The name 'ros\_ip\_9\_26\_124\_81-rs' reappears in subsequent steps, replace it with the actual name that you have chosen for the ROS VIP resource.

6. Create the resource group for the VIP using the following commands:

```
# mkrgrp hadr_ros_ip-rg  
# addrgmbr -g hadr_ros_ip-rg IBM.ServiceIP:ros_ip_9_26_124_81-rs  
# mkrel -p DependsOn -S IBM.ServiceIP:ros_ip_9_26_124_81-rs \  
-G IBM.Equivalency:hadr_shadow_equiv ros_do
```

7. Bring the HADR ROS IP resource group online using the following command:

```
# chrg -o Online hadr_ros_ip-rg
```

At this point, the ROS VIP address is online and should be collocated with the current location of the Standby HADR database. Verify this using the **lsnam** command. Your output should be similar to the output that follows. In particular, confirm that all states are normal (if any states are printed in red or yellow, then there is likely a configuration error.) Confirm also that the VIP for the ROS is online and that it represents the Standby HADR node.

```

root@grom01:/usr/sbin/rsct/sapolicies/db2# lssam
Online IBM.ResourceGroup:db2_db2inst1_db2inst1_GSDB-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_db2inst1_GSDB-rs
|   |- Online IBM.Application:db2_db2inst1_db2inst1_GSDB-rs:grom01
|   |- Offline IBM.Application:db2_db2inst1_db2inst1_GSDB-rs:grom02
|   |- Online IBM.ServiceIP:db2ip_9_26_124_80-rs
|       |- Online IBM.ServiceIP:db2ip_9_26_124_80-rs:grom01
|       |- Offline IBM.ServiceIP:db2ip_9_26_124_80-rs:grom02
Online IBM.ResourceGroup:db2_db2inst1_grom01_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_grom01_0-rs
|   |- Online IBM.Application:db2_db2inst1_grom01_0-rs:grom01
Online IBM.ResourceGroup:db2_db2inst1_grom02_0-rg Nominal=Online
|- Online IBM.Application:db2_db2inst1_grom02_0-rs
|   |- Online IBM.Application:db2_db2inst1_grom02_0-rs:grom02
Online IBM.ResourceGroup:hadr_ros_ip-rg Nominal=Online
|- Online IBM.ServiceIP:ros_ip_9_26_124_81-rs
|   |- Offline IBM.ServiceIP:ros_ip_9_26_124_81-rs:grom01
|   |- Online IBM.ServiceIP:ros_ip_9_26_124_81-rs:grom02
Online IBM.Equivalency:db2_db2inst1_db2inst1_GSDB-rg_group-equ
|- Online IBM.PeerNode:grom01:grom01
|- Online IBM.PeerNode:grom02:grom02
Online IBM.Equivalency:db2_db2inst1_grom01_0-rg_group-equ
|- Online IBM.PeerNode:grom01:grom01
Online IBM.Equivalency:db2_db2inst1_grom02_0-rg_group-equ
|- Online IBM.PeerNode:grom02:grom02
Online IBM.Equivalency:db2_private_network_0
|- Online IBM.NetworkInterface:eth1:grom01
|- Online IBM.NetworkInterface:eth0:grom02
Online IBM.Equivalency:db2_public_network_0
|- Online IBM.NetworkInterface:eth0:grom01
|- Online IBM.NetworkInterface:eth2:grom02
Online IBM.Equivalency:hadr_shadow_equiv
|- Offline IBM.Application:hadr_shadow_01:grom01
|- Online IBM.Application:hadr_shadow_02:grom02

```

If your output meets these conditions, the server side setup is now complete. Otherwise, go back and recheck your configuration steps. You can remove any partial or incorrectly configured resources that were just added using the following commands:

```

# chrg -o Offline hadr_ros_ip_rg
# rmrel hadr_shadow_equiv_ros_do
# rmrsrc -s "Name like 'ros_ip_%'" IBM.ServiceIP
# rmrsrc -s "Name like 'hadr_shadow%'" IBM.Application
# rrmrg hadr_ros_ip_rg
# rmequ hadr_shadow_equiv

```

After the server setup is complete, you are ready to test that the VIP addresses can switch to the correct server during takeover operations. As the instance owner, issue the **TAKEOVER HADR** command (in our case, db2 takeover hadr on database GSDB). After you run the command, the location of the Primary side VIP should move to grom02, and the location of the Standby side VIP should move to grom01. If this is not the case, you have likely chosen an invalid VIP. Remove the created resources using the commands above and recheck the chosen VIP. If the VIP switches correctly, you can then issue another **TAKEOVER HADR** command, the VIP addresses should switch back to their original settings.

## 2.4 Configuring the clients

---

To configure the clients to connect using the ROS VIP, perform the following steps:

**Note:** In the example that follows, we assume that the VIP associated with the Primary HADR database is '9.26.124.80' and the VIP associated with the Standby HADR database is '9.26.124.81'. You can confirm your own values using the **lsnam** command. Both database managers are listening for client connections using port 60001 (this value is set using the database manager configuration value SVCENAME.)

1. On the database clients, login and catalog the Primary and Standby HADR databases using their associated VIPs using the following commands:

```
db2 catalog tcpip node HANODEP remote 9.26.124.80 server 60001
db2 catalog tcpip node HANODES remote 9.26.124.81 server 60001
```

2. On the database clients, catalog the Primary and Standby HADR databases using different aliases:

```
db2 catalog db GSDB as GSDBP at node HANODEP
db2 catalog db GSDB as GSDBS at node HANODES
```

3. Verify the connectivity to each of the database aliases by following these steps:
  - a. Attempt to connect to the database alias HANODEP. The connection is routed to the location of the Primary HADR GSDB database.
  - b. Attempt to connect to the database alias HANODES. The connection is routed to the location of the Standby HADR database.

**Note:** Client-side configuration steps depend on how the front-end applications access the HADR databases. For example, if you have a web application accessing the database using a middleware application server such as IBM WebSphere® Application Server (WAS), you would need to add a new JDBC driver source definition to access the Standby ROS database via the ROS VIP. Similarly, for an ODBC application, you would need to add a new Data Source Name (DSN) to the ODBC configuration file with the ROS VIP information.

Regardless of the current location of the Primary or Standby HADR database, connection requests to either of these aliases are routed appropriately; at all times, a connect request to the HANODEP alias directs the request to the Primary HADR database while a connect request to the HANODES alias directs the request to the Standby HADR database.

### 3 Conclusion and summary

---

This paper describes how you can setup your environment to provide seamless connectivity from your client applications to both the Primary HADR database and the Standby HADR copy.

To create this setup, we use a virtual IP address and associate it with the HADR ROS standby copy. Then, we catalog a database alias against this virtual IP address to ensure that clients can access that read-only copy, regardless of the physical location of the HADR ROS database.

In the event of a failover, whether it is a controlled failover using the **TAKEOVER HADR** command or a failover initiated by a system failure, this setup eliminates data access downtime because it does not require any database recataloging on the clients or any other user intervention during the failover. Applications can continue to automatically access the data in both HADR databases.

### 4 Additional information

---

- DB2 Best Practices  
<http://www.ibm.com/developerworks/db2/bestpractices/>
- DB2 9 for Linux, UNIX and Windows manuals  
<http://www.ibm.com/support/docview.wss?rs=71&uid=swg27009474>
- DB2 9.5 for Linux, UNIX, and Windows Information Center  
<http://publib.boulder.ibm.com/infocenter/db2luw/v9r5/index.jsp>
- DB2 9.7 for Linux, UNIX, and Windows Information Center  
<http://publib.boulder.ibm.com/infocenter/db2luw/v9r7/index.jsp>

### 5 Contributors

---

Steven Raspudic, DB2 Development

Aruna De Silva, DB2 Development

Louise McNicoll, DB2 Information Development

## Notices

---

© Copyright IBM Corporation 2011.

**IBM Canada**  
**8200 Warden Avenue**  
**Markham, ON**  
**L6G 1C7**  
**Canada**

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing*  
*IBM Corporation*  
*North Castle Drive*  
*Armonk, NY 10504-1785*  
*U.S.A.*

**The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims

related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

COPYRIGHT LICENSE:

This information contains sample programs and code in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

## **Trademarks**

---

IBM, the IBM logo, and [ibm.com](http://ibm.com)® are trademarks or registered trademarks of International Business Machines Corporation registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

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.