

High Availability Options for SAP Using IBM® PowerHA® SystemMirror for i

Lilo Bucknell

Jenny Dervin

Luis BL Gonzalez-Suarez

Eric Kass

June 12, 2012



About the Authors

Lilo Bucknell was an IBM senior software engineer working in the IBM i PowerHA and clustering development area. She was with IBM for over 30 years. Previous assignments include the AS/400 System Design Control Group and an assignment at SAP in Walldorf, Germany. She is a certified SAP Basis Consultant. Lilo has recently retired from IBM.

Jenny Dervin is the architect for IBM PowerHA SystemMirror for i and the technology leader for high availability on IBM i. She has been with IBM for over 20 years. She first worked on AS/400 performance, including the Performance Tools product and also SAP benchmarking. She has been on the PowerHA development team for over 10 years.

Luis BL Gonzalez-Suarez is the SAP on IBM i Installation and Logistics team lead. Luis has been working on SAP at their Walldorf campus since 2004. Luis designed and developed the SAP Installer (SAPIinst) running natively on IBM i. Before coming to Germany, Luis worked on the DB2 on IBM i team in Rochester; there he designed the standard database feature and regression test infrastructure for the platform.

Eric Kass began working at IBM in 1997 on the communication and encryption components of the System License Internal Code (SLIC) of the IBM i operating system. Since 2000, Eric works in Germany on the SAP on IBM i porting team. Eric's leading projects include: heterogeneous application servers for SAP on IBM i, SAP running within the PASE subsystem of IBM i, the SAP on IBM i database driver, and SAP kernels optimized for IBM Power. Eric is the technical focal point for SAP on IBM i.

Special Thanks to:

Ron Schmerbauch, technical team leader of the SAP on IBM i team in Rochester. Ron helped with the coordination of this project and contributed to the content and review of this paper.

Kolby Hoelzle, now a member of the SAP on IBM i Center of Excellence within Rochester's Lab Based Services, spent previous years developing the initial SAP implementation in an Independent ASP. Kolby is the author of numerous papers about disaster recovery of SAP on IBM i. Many of Kolby's concepts are reflected in this paper.

Table of Contents

About the Authors	2
Table of Contents	3
Abstract.....	5
1 Introduction.....	6
1.1 IBM PowerHA SystemMirror for i.....	6
1.2 Traditional High Availability Options for SAP.....	7
1.2.1 2-Tier Implementations	7
1.2.2 3-Tier Implementations	9
2 IBM i High Availability for SAP Solutions	11
2.1 Standard Four Node HA Scenario	11
2.2 IBM i Two Node HA Scenario	12
3 A fully integrated SAP & PowerHA SystemMirror Solution.....	14
3.1 PowerHA Components of the Solution	15
3.1.1 Cluster	15
3.1.2 Independent ASP	15
3.1.3 Cluster Resource Groups	16
3.1.4 Administrative Domain	16
3.2 SAP Components of the Solution	17
3.2.1 Database	17
3.2.2 Central Services & Enqueue Replication Server	18
3.2.3 File System	18
3.2.4 Application Servers	18
4 References	19
4.1 PowerHA SystemMirror for i Cookbook	19

4.2	SAP on IBM i Reference Architecture.....	19
4.3	Implementing SAP in an IASP environment.....	19
4.4	Setup of an SAP Solution with Power HA.....	19
4.5	IBM i Lab Services	19
	Trademarks and Special Notices.....	20

Abstract

While most SAP on IBM i customers now implement some form of high availability solution, previous solutions only replicated the SAP database and file-system (IFS) data. Planned or unplanned outages still required an application restart, which is not transparent to the end users.

Introduced with SAP Central Services in SAP NetWeaver 7.10, SAP provided the infrastructure to implement a highly available landscape using PowerHA SystemMirror for i. In a high availability configuration, PowerHA SystemMirror for i manages a cluster of IBM i nodes within which SAP Central services provides application resiliency and PowerHA SystemMirror for i provides data resiliency for DB2.

This paper describes the different options for achieving a highly available SAP environment, utilizing PowerHA SystemMirror for i. Special attention is given to the new options which are based on the SAP high availability architecture. The concepts of this implementation are described along with various scenarios for managing the environment.

1 Introduction

SAP has been a key ISV application on IBM i since 1995, when its SAP R/3 Enterprise Resource Planning (ERP) software for OS/400 was released. Since then, IBM i has proven itself to be a reliable and scalable platform for SAP applications. With many reliability and availability functions built into the operating system, customers running SAP solutions can expect minimal downtimes. There are many IBM i customers utilizing SAP to run their business.

With the advent of 24x7 uptime requirements resulting from globalization and the ever-expanding world economy, high availability becomes a key component of the customer environment to protect from both planned and unplanned outages. A growing majority of our SAP customers are choosing PowerHA SystemMirror for i for their high availability requirements. As SAP has evolved, so has the need for even more robust environments. SAP has improved their high availability capabilities and PowerHA SystemMirror for i can exploit those capabilities to provide an environment where seamless switchovers of the entire SAP system can occur. This paper will describe the various options available as well as demonstrate how to manage this new type of environment.

The intent of this paper is to explain some of the high availability features of both PowerHA SystemMirror for i and SAP; covering the high availability advantages available with various SAP configurations; and describing different scenarios to take full advantage of the high availability capabilities of both PowerHA SystemMirror for i and SAP.

It is assumed that the reader has some existing knowledge of both SAP and PowerHA SystemMirror for i. Additional information on PowerHA SystemMirror for i can be found in the PowerHA SystemMirror for i Redbook at

<http://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/sg247994.html?Open>.

Additional information on SAP can be found in the SAP on IBM i Reference Architecture whitepaper at <http://scn.sap.com/docs/DOC-14443>. IBM Lab Services has been involved in many implementations of SAP and PowerHA SystemMirror on IBM i, and is an excellent source if implementation assistance is needed. More information on Lab Services can be found at <http://www-03.ibm.com/systems/services/labservices>.

1.1 IBM PowerHA SystemMirror for i

PowerHA SystemMirror for i is the premier IBM i high availability offering from IBM. PowerHA SystemMirror offers hardware-based replication options, where the replication is managed at a disk I/O level, either by IBM i or by an external storage server. PowerHA SystemMirror is an integrated solution and can be managed completely from IBM i. In the 6.1 release, it was known simply as PowerHA for i, and was renamed to PowerHA SystemMirror for i in the 7.1 release.

PowerHA SystemMirror for i offers several different replication offerings. Geographic mirroring is an IBM i replication technology which can be used on any type of storage, internal, external, direct attached, or virtual. The initial offering included synchronous geographic mirroring which supported limited distance between the production and backup systems. In 7.1, the offering was expanded to include asynchronous geographic mirroring. For customers using external storage, specifically DS8000®, SVC, or the Storwize® V7000, PowerHA SystemMirror for i offers integrated solutions utilizing their replication technologies.

Metro Mirror is the synchronous offering and Global Mirror is an asynchronous offering. For customers looking for a combination of local high availability plus disaster recovery, switched disk between partitions can be used in conjunction with geographic mirroring for internal storage, and LUN (logical unit) level switching can be combined with either Metro Mirror or Global Mirror. Both switched disk and LUN level switching involve having one copy of the data which can be switched between partitions for internal storage or systems for external storage. LUN-level switching is currently only available for the DS8000 storage servers, not SVC or the Storwize V7000.

PowerHA SystemMirror for i is built upon the concept of independent auxiliary storage pools (IASPs). The data to be replicated is located in the IASP, which is created from specific disk drives or LUNs. Whenever a disk write occurs within the source IASP, the same write is replicated on the target IASP. Since the target IASP is not available to users while replication is active, there is a guarantee that the source and target IASPs are identical for synchronous replication. For asynchronous replication, there may be a delay, but the writes are still guaranteed to be in order.

Objects which cannot be located in an IASP can be synchronized across the production and backup systems via another PowerHA technology called Administrative Domain. Although not all object types can be synchronized using this technology, the most common and frequently used objects are supported.

Using the combination of hardware replication within the IASP and administrative domain to synchronize the application environment, PowerHA SystemMirror for i provides a robust high availability environment which is easy to configure and maintain.

SAP environments have worked well within the PowerHA architecture for years, and running SAP within an IASP is well documented (see Appendix for reference). With the introduction of SAP's new high availability architecture, SAP and PowerHA can be even more closely tied, providing even better availability of both the database and the application servers, resulting in significantly improved availability of the entire SAP implementation.

1.2 Traditional High Availability Options for SAP

SAP is a very versatile product, and can be configured in many different ways, depending on the customer requirements. In each of these environments, high availability is recommended.

1.2.1 2-Tier Implementations

The traditional 2-tier environment for SAP has all components contained on one host: the SAP application server and all components of the *central instance*, the database and the file system. This configuration is very simple to install and operate but if any component fails, the SAP application is down.

The figure below depicts the most simple of SAP configurations whereby the previously mentioned components reside on one host. The Central Instance here is the only instance and performs application server processing as well as lock management (Enqueue) and message handling (Message Server).

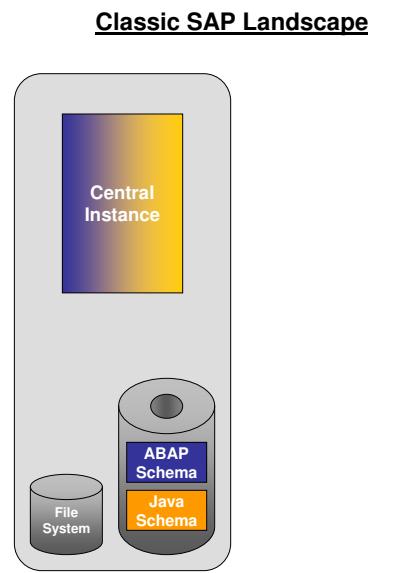


Figure 1 - Classic SAP Landscape

It is possible to make this 2-tier environment recoverable from a disaster by constructing two equivalent configurations of the above on two separate hosts or nodes and mirroring the database and file system between the nodes as shown in Figure 2. The word node introduced here defines a standalone server or logical partition (LPAR) which participates as a member in a high availability configuration. Member nodes of a high availability configuration are referred to as a cluster (of nodes).

Classic SAP on IBM i Landscape: Disaster Recovery with Power HA

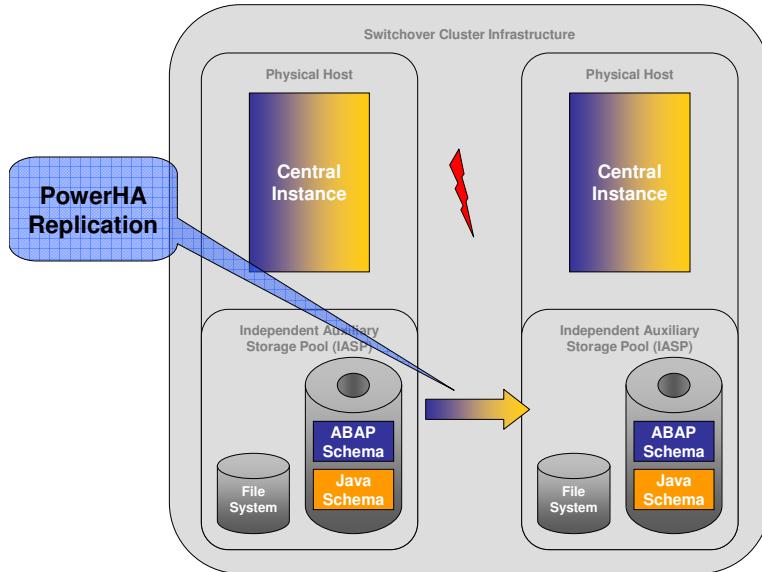


Figure 2 - Disaster Recovery with PowerHA

With this HA implementation shown in Figure 2, the SAP database is located in an IASP and is replicated from the source to the backup node. In the case of a planned or unplanned outage of the production node, the database can then be accessed from the backup node. The main drawback with this scenario is that the central instance will need to be restarted to access the new database location.

1.2.2 3-Tier Implementations

A basic 3-tier implementation provides another level of availability. With this implementation, the application servers are placed on multiple nodes other than the node of the central instance. If one application server is unavailable, work is switched to another.

Classic SAP 3-tier Landscape

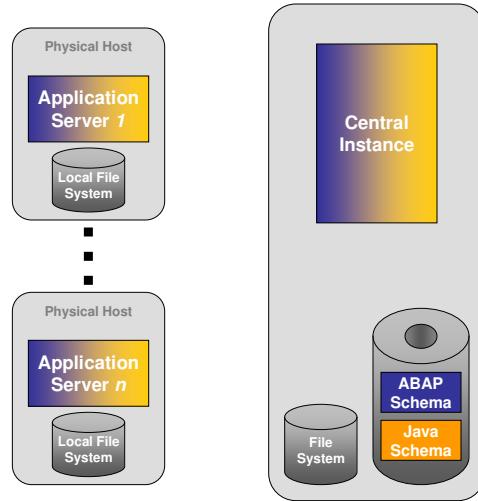


Figure 3 - Classic SAP 3-tier Landscape

It is again possible to replicate the database, providing a level of disaster recovery in this scenario. However, since the central instance is not replicated, the application servers will again need to be restarted in a planned or unplanned switch.

2 IBM i High Availability for SAP Solutions

2.1 Standard Four Node HA Scenario

Starting with NetWeaver 7.10, SAP defines a standard configuration for high availability scenarios that is more distributed than the classic central instance and application server configuration. In this new configuration the components that SAP refers to as the *single points of failure* (SPoF) are redundant.

SAP Central Services for Java (SCS) and ABAP SAP Central Services (ASCS) are primarily the combination of the SAP Message server and the SAP enqueue server; both formerly residing within the SAP central instance. The message server is the communication hub for distributed components of an SAP system. The enqueue server is the central application lock manager providing flow control and contention management between SAP applications running in parallel within multiple work-processes on multiple nodes.

The motive for separating message and enqueue services into a separate instance was specifically to facilitate high availability by; firstly, providing the architecture to support a replica of the enqueue server state; and secondly, the near continuous availability of enqueue and message services by way of backup SCS and ASCS instances. The enqueue replication server (ERS) running on a backup node actively records the state of the active enqueue server.

In Figure 4 below, five *SPoF* are identified; reading from left to right: SAP Central Services, Web Dispatchers, application servers, database, and file system.

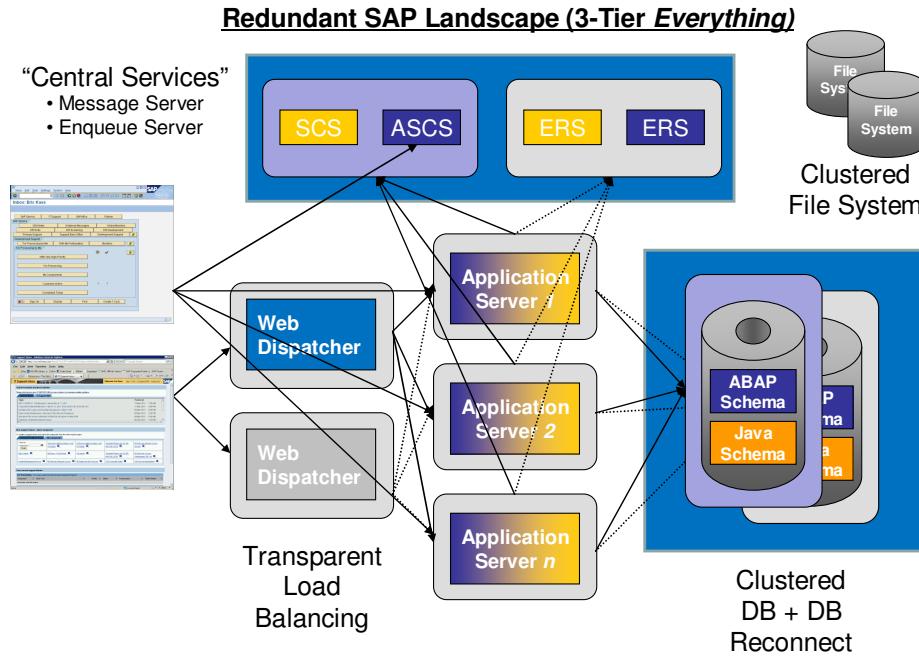


Figure 4 - Redundant SAP Landscape (3-Tier *Everything*)

SAP GUI clients (far left) communicate to one of many application servers (middle) directly or perhaps by first load-balancing via the Central Service's message server (top center). Web based clients (bottom left) communicate to one of a set of redundant Web Dispatcher (middle); they in turn communicate with the application servers. The application servers maintain persistent connections to the clustered database server, clustered file system, and the clustered SAP central services.

If any of the clustered components switch or fail-over, the application servers automatically re-establish connections to the failing component.

The above scenario is typically installed on multiple nodes – either physical servers or partitions; isolating each component to a node.. With IBM i, though, it is plausible to install an SAP high availability scenario completely on two nodes.

2.2 IBM i Two Node HA Scenario

IBM i subsystem technology offers the unique ability of running multiple SAP components on a single Logical Partition (LPAR) or physical system. The SAP Installer (SAPIInst) automatically configures each SAP component into a separate subsystem in a way that PowerHA SystemMirror can manage multiple independent and redundant SAP components residing on just two physical servers. This configuration is shown in figure 5.

As in the multi-node configuration described in section 2.1, three resources per SAP system are configured as cluster resource groups; SAP central services, database, and file system. Any number of application servers can be configured on each of the two nodes.

In figure 5, application servers are running on both nodes concurrently. When any of the clustered resources: central services, database, or file system are switched between nodes – independently or together, application servers on both nodes will automatically re-establish connections to the switched resources.

Using this two node configuration, hardware or operating system maintenance on a node can be accomplished by first moving all clustered resources to the *other* node. The process of moving involves ending SAP components, switching the database node, and then restarting the SAP components on the other node. The effect of switching in this fashion will be that users are logged out for the duration until SAP components are restarted and available. Once the node is freed, maintenance can occur.

Redundant SAP Landscape in 2-Tier

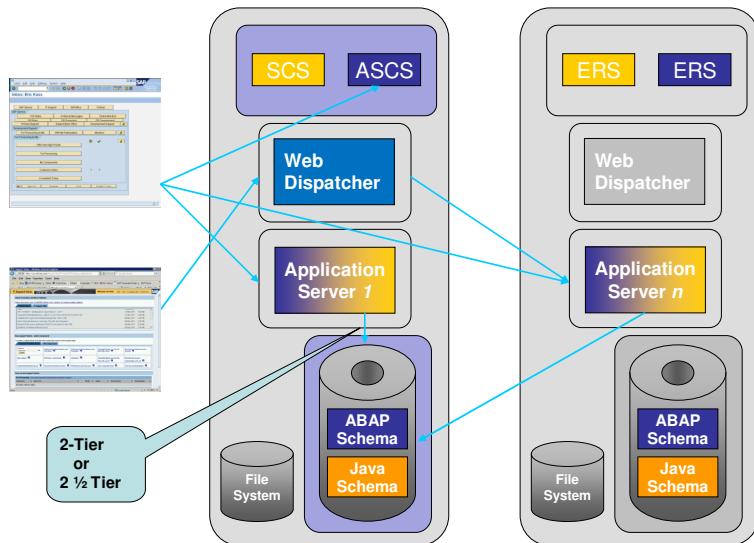


Figure 5 - Redundant SAP Landscape in 2-Tier

3 A fully integrated SAP & PowerHA SystemMirror Solution

Establishing a highly available SAP landscape requires removing any single points of failure (SPoF). To this end, it is convenient to conceptually separate the SAP landscape into components called resources. Each resource is a required component and must be redundant. Some components like application servers can simply exist in multiple. Other components, such as the database or file system, exist either as a clustered or switchable resource. A highly available SAP environment relies on the cluster services of the underlying operating system upon which it runs. On IBM i, the cluster services are provided by PowerHA SystemMirror for i. The redundant switchable resources in the SAP environment are controlled by PowerHA SystemMirror for i.

Using the new high availability characteristics of SAP and PowerHA SystemMirror for i, it is possible to create a highly available SAP environment which does not require the applications to end during a planned or unplanned outage.

The high availability solution described conceptually in this paper and in more detail in SAP note 1635602 separates the data and application components of the SAP environment across IBM i partitions and SAP instances to eliminate SPoF within the runtime environment.

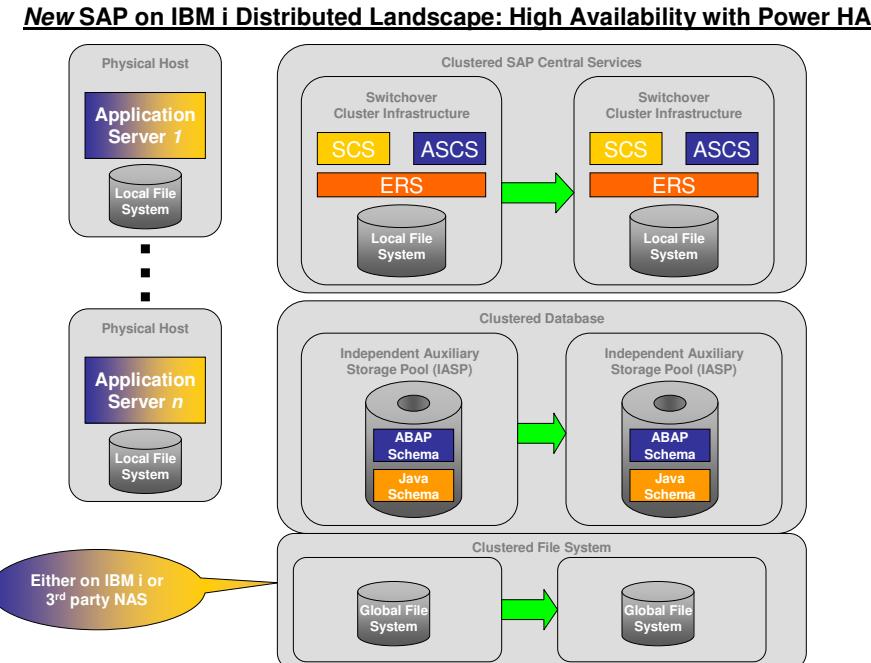


Figure 6 - SAP with PowerHA SystemMirror for i

For the data component, the SAP database exists in an IASP on an IBM i node. One of the PowerHA SystemMirror for i replication technologies is used to replicate the database to a second node. This provides the capability to switch between the primary and backup node.

Application high availability with PowerHA SystemMirror for i is achieved by implementing the SAP Central Services (SCS/ASCS) instance and the corresponding enqueue replication server (ERS) instance in two IBM i nodes, one node designated as the primary SCS/ASCS and the other as the backup. The SCS/ASCS runs on the current primary node, the ERS runs on the current backup node.

When combined with redundant SCS/ASCS instances, multiple remote SAP application server instances inherently provide application server redundancy. Multiple instances also provide application high availability during kernel updates. In figure 6, multiple application servers are configured, and login groups managed by SAP central services could be used to assign users to a defined set of application servers.

3.1 PowerHA Components of the Solution

The following sections describe at a high level the concepts that make up a complete HA solution for SAP.

3.1.1 Cluster

The cluster is the underlying PowerHA structure which provides communication between the nodes (partitions) in a cluster solution. A cluster node will receive notification of cluster "events", that is, messages upon which a node may or may not need to take action, such as a node (partition) failure.

3.1.2 Independent ASP

Auxiliary storage on IBM i refers to the disk units attached to the system; from the object-based perspective of the IBM i operating system, storage where objects remain persistent. The spinning or solid state disks can be attached directly to internal storage controllers (SAS), attached to a SAN via fiber-channel, or attached through an IBM VIOS (Virtual I/O Server) partition.

Auxiliary storage is organized into storage pools on IBM i. An auxiliary storage pool (ASP) can contain any number of disks and provides a means of data separation similar to that offered by a partition on Windows. A typical use for a secondary ASP is as a location for database journals (logs). When configured as a separate ASP utilizing other disks in other redundancy sets than the primary ASP, database recovery is possible in the event of a total loss of the primary ASP.

An Independent ASP is a special type of auxiliary storage pool that can be physically or logically switched between IBM i hosts configured to share the storage. SAP Note 56820 describes how to implement SAP in an IASP environment.

An IASP is a building block upon which database high availability on IBM i can be based. An IASP can be mirrored using hardware based replication and switched between nodes controlled by PowerHA SystemMirror for i. SAP Note 1635602 contains examples for configuring hardware replication through geographic mirroring. Notice in Figure 6 that the SAP global file system (Section 3.2.3) is contained within another IASP. Defining a file system IASP as a user defined file system type (UDFS) IASP minimizes switchover times since only the file system and no IBM i objects need to be made available when the storage is mounted (aka varied on).

3.1.3 Cluster Resource Groups

A cluster resource group (CRG) is an IBM i system object which is used to configure the HA characteristics for specific resources in the solution. The CRG manages one resource and specifies a set of nodes that serve resource in a highly available fashion. Only one node in the CRG is designated the primary for the CRG at any point in time. Two types of CRGs are used within a clustered SAP environment. The independent auxiliary storage pool (IASP) resources are managed by a device CRG. The SAP central services (ASCS/SCS) are managed by a application CRG.

In the example shown in Figure 6, there are three CRGs defined: a device CRG for data base resiliency; an application CRG for application resiliency for the SCS/ASCS and ERS instances; and another device CRG for resiliency of the global file system.

3.1.3.1 Exit Program

A CRG may have an exit program defined. IBM and SAP developed an exit program to be used with all of the CRGs managing SAP resources. The exit program is called for specific cluster events and is used to start and stop the SCS/ASCS and ERS instances in a synchronized manner. Exit program data determines the specific actions to be taken for each cluster event. The exit program is called SAPCRGEXIT and is delivered with the SAP host agent starting with version 7.20. The SAP host agent can be installed on any host, regardless of the version levels of any other SAP components. The exit program data for the application CRG is described in SAP note 1635602.

3.1.3.2 Takeover IP Address

A takeover IP address, otherwise known as a virtual IP address, is associated with the resource defined by the cluster resource group (CRG). When the CRG moves the resource it manages (by making another node the primary), the takeover IP address is moved too. A takeover IP address is often given a *virtual host name* by defining it in a host table or DNS (domain name server).

One takeover IP address can be defined for each CRG. One takeover IP address is then associated with the SAP database IASP, another for each SAP central service instance (ASCS/SCS), and another for the file system IASP. PowerHA SystemMirror for i takes care of starting the IP address on the current primary node of the CRG, and when moving a resource to a different node; takes care of ending the IP address on the previous primary node and starting it on the new primary node.

3.1.4 Administrative Domain

The administrative domain is used to synchronize objects which cannot be placed into an IASP. Many object types are supported in administrative domain, and the administrator can choose the specific objects to be synchronized. Once an object is added to the administrative domain, then any changes to that object on any system in the administrative domain will be synchronized across the other nodes. For SAP environments, this includes: user profiles, system values, job descriptions, and subsystem descriptions.

3.2 SAP Components of the Solution

The SAP components participating in providing a high available solution are shown in Figure 7. When operating an HA scenario, PowerHA SystemMirror for i controls the availability of these basic SAP components.

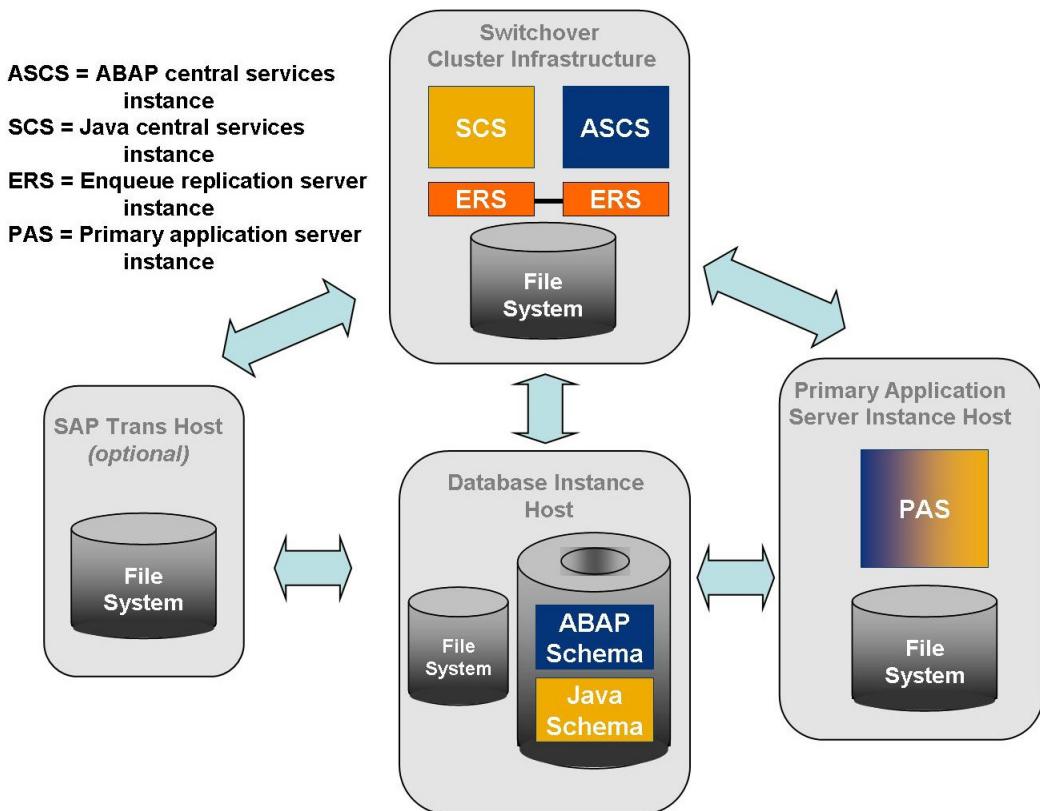


Figure 7 - SAP Components of an HA Solution

3.2.1 Database

The SAP database schemas for ABAP and Java and respective journals are stored in an IASP. A device cluster resource group (CRG) is used to manage the replication of the IASP and control on which node in the cluster the IASP is available.

3.2.1.1 Virtual Host Names

SAP uses the concept of virtual host names to transparently access services in an HA environment. A virtual host name corresponds directly to a takeover IP address. By accessing services by virtual host name, the application does not need to "know" on which node a particular service is actually running. In one example of an SAP landscape there might be a virtual host defined for the database (i.e. SAPDBHOST) and another virtual host defined for the SAP central service (i.e. rdisp/mshost, and enqueue/serverhost).

3.2.2 Central Services & Enqueue Replication Server

Availability to SAP Central Services and Enqueue Replication Server is managed via an application CRG and exit program provided by SAP. The exit program will monitor the 'wellness' of the SAP application and notify PowerHA of problems.

The second virtual host definition is the SCS/ASCS virtual host definition. The IP for this virtual host is defined in the application CRG as its failover IP address. The application CRG is the entity that provides the starting and stopping of the IP address on the appropriate primary or backup, depending on what cluster event it is acting upon. In this way, the central and ERS instances can switch transparently between systems without impacting the availability of the services to the data base and to the work processes.

3.2.3 File System

The SAP system relies on a shared Posix type file system referred to as the global file system in SAP literature.. For the most part, the global file system contains SAP configuration profiles (`/usr/sap/<SID>/SYS/profile`), a copy of the runtime executables (binary or java), and a global exchange directory (`/usr/sap/<SID>/SYS/global`).

In a clustered landscape, the file system must be configured to withstand a temporary loss of a file system server. Some options are to have the file system provided by an external clustered NFS server, or to have the file system provided by an IASP similar to that of the database (Refer to section 3.1.2).

3.2.4 Application Servers

Application servers are not replicated, but can be redundant. The example includes redundant application servers. The message server can control load balancing. Users are connected to specific application servers and would need to reconnect if that application server becomes unavailable. However, the users would be oblivious to any issues with the other SAP resources.

4 References

4.1 PowerHA SystemMirror for i Cookbook

<http://www.redbooks.ibm.com/Redbooks.nsf/RedpieceAbstracts/sg247994.html?Open>

4.2 SAP on IBM i Reference Architecture

<http://scn.sap.com/docs/DOC-14443>

4.3 Implementing SAP in an IASP environment

SAP Note **568820**: *IBM i : Implementing an Independent ASP (iASP) System*

4.4 Setup of an SAP Solution with Power HA

SAP Note **1635602**: *SAP High Availability on IBM i Configuration Guide*

4.5 IBM i Lab Services

<http://www-03.ibm.com/systems/services/labservices>

Trademarks and Special Notices

© Copyright IBM Corporation 2012. All rights Reserved.

References in this document to IBM products or services do not imply that IBM intends to make them available in every country.

IBM, IBM Systems Director and the IBM logo are trademarks of International Business Machines Corporation in the United States, other countries, or both.

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

ABAP, SAP NetWeaver, SAP Solution Manager, SAP, and SAP logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries.

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

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-IBM products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by IBM. Sources for non-IBM list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. IBM has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-IBM products. Questions on the capability of non-IBM products should be addressed to the supplier of those products.

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

Contact your local IBM office or IBM authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in IBM product announcements. The information is presented here to communicate IBM's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.