IBM i

Availability
High availability overview

7.1

IBM i

Availability
High availability overview

7.1

Note

Before using this information and the product it supports, read the information in "Notices," on page 31.

This edition applies to IBM i 7.1 (product number 5770-SS1) and to all subsequent releases and modifications until otherwise indicated in new editions. This version does not run on all reduced instruction set computer (RISC) models nor does it run on CISC models.

This edition replaces SCnn-nnnn-nn.

Contents

High availability overview

Business continuity is the capability of a business to withstand outages and to operate important services normally and without interruption in accordance with predefined service-level agreements. To achieve a given level of business continuity that you want, a collection of services, software, hardware, and procedures must be selected, described in a documented plan, implemented, and practiced regularly. The business continuity solution must address the data, the operational environment, the applications, the application hosting environment, and the end-user interface. All must be available to deliver a good, complete business continuity solution.

Business continuity includes disaster recovery (DR) and high availability (HA), and can be defined as the ability to withstand all outages (planned, unplanned, and disasters) and to provide continuous processing for all important applications. The ultimate goal is for the outage time to be less than .001% of total service time. A high availability environment typically includes more demanding recovery time objectives (seconds to minutes) and more demanding recovery point objectives (zero user disruption) than a disaster recovery scenario.

High availability solutions provide fully automated failover to a backup system so that users and applications can continue working without disruption. HA solutions must have the ability to provide an immediate recovery point. At the same time, they must provide a recovery time capability that is significantly better than the recovery time that you experience in a non-HA solution topology.

What's new for IBM i 7.1

Read about new information for the High Availability overview topic collection.

What's new as of October 2016

IBM® PowerHA® for i enhanced advanced node failure detection to support a new representational state transfer (REST) interface. The Hardware Monitor Console (HMC) is being updated to replace the existing Common Information Model (CIM) server with a new representational state transfer (REST) based interface. HMC version V8R8.5.0 is the last version of HMC to support the CIM server, and is the first version of HMC to support all REST API functions that are required by IBM PowerHA for i licensed program. This function is provided through a new function PowerHA PTF.

Advanced node failure detection

IBM i cluster resource services can now use a Hardware Management Console (HMC) or a Virtual I/O Server (VIOS) partition to detect when a cluster node fails. This new capability allows more failure scenarios to be positively identified and avoids cluster partition situations.

See Advanced node failure detection for additional information about this topic.

Asynchronous delivery mode for geographic mirroring

Asynchronous Delivery Mode for Geographic Mirroring now supports a new asynchronous delivery mode which potentially increases the amount of latency (and thus distance) which can be tolerated by most applications using Geographic Mirroring.

See Geographic mirroring characteristics for additional information about this topic.

Logical unit level switching

Switched logical units allow data that is stored in the independent disk pool from logical units created in an IBM System Storage® DS8000® or DS6000™ to be switched between systems providing high availability.

See Switched logical unit characteristics for additional information about this topic.

PDF file for High availability overview

You can view and print a PDF file of this information.

To view or download the PDF version of this document, select High availability overview (about 415 KB).

You can view or download these related topic collection PDFs:

- High availability technologies (about 580 KB) contains the following topics:
  - Clusters technology
  - Cluster administrative domain
  - Switched disk pools
  - Switchable devices
  - Cross-site mirroring
    - Geographic mirroring
    - Metro mirror
    - Global mirror
  - FlashCopy®
  - High-availability management
- Implementing high availability (about 4,123 KB) contains the following topics:
  - Installing IBM PowerHA for i (iHASM) licensed program (5770-HAS)
  - Uninstalling IBM PowerHA for i (iHASM) licensed program (5770-HAS)
  - Implementing high availability with the solution-based approach
  - Implementing high availability with the task-based approach
  - Managing high availability
  - Troubleshooting high availability

Saving PDF files

To save a PDF on your workstation for viewing or printing:
1. Right-click the PDF link in your browser.
2. Click the option that saves the PDF locally.
3. Navigate to the directory in which you want to save the PDF.
4. Click **Save**.

Downloading Adobe Reader

You need Adobe Reader installed on your system to view or print these PDFs. You can download a free copy from the Adobe Web site (www.adobe.com/products/acrobat/readstep.html) .

Benefits of high availability

High availability protects companies from lost revenue when access to their data resources and critical business applications is disrupted.

The starting point for the selection of a high availability solution is to fully identify the set of availability problems that you are attempting to address. For business continuity, these problems can be collected into five major categories.

Planned outages

IBM i high availability can reduce the impact to your customers and users whenever you need to take systems or data offline to perform necessary maintenance tasks, such as nightly backups or the installation of new hardware or software.

As a business grows, uptime becomes increasingly important. The maintenance window for your systems can shrink dramatically. Scheduled downtime includes things such as tape backups, application upgrades, and operating system upgrades among other things. How many hours per week can the application be unavailable, and not impact your business? Planned outages are typically the most common event that a high availability solution is used for.

IBM i single system availability focuses on hardware and software concurrent maintenance and hardware redundancy, but there is a limit to what can be done on a single system level. Using IBM i high availability technologies, such as clusters and independent disk pools, you can switch production to a second system or have a second set of data available. These IBM i high availability solutions allow your business to continue while system maintenance is being performed. The impact of planned outages can be minimized using these high availability solutions.

Offline Saves to Tape
Saves to tape can be performed from a backup system that has a second copy of the user's data.

Application and Operating System fixes or upgrades
A rolling upgrade can be performed to allow fixes or upgrades to be installed. Fixes can be applied to the backup system while the primary system is running production. The workload can then be switched to the backup system and fixes can be applied to the original primary. After the upgrade has finished, production can be switched back to the original primary.

Hardware Maintenance
Changes that cannot be handled by concurrent hardware maintenance typically require downtime of the system. Having a high availability solution will allow production to be switched to a backup system and the hardware maintenance performed without impacting the business.

Related concepts:

"Outage coverage" on page 14
What kind of outage is the business trying to protect against? Backup window reduction, planned maintenance, unplanned outages, or site disasters are events to consider when choosing a high availability solution.

Related information:

Shortening planned outages

Unplanned outages

IBM i high availability solutions can provide protection from unplanned outages caused by human error, software problems, hardware failures, and environmental issues.

As a business grows, the protection from unplanned events becomes more critical. Unfortunately, unplanned events cannot be scheduled. The high availability requirement of the business should focus on the time frame that is most important to the business. The cost of being down at the most critical moment should be considered when selecting which high availability solution will be implemented and how the implementation is done.

Unplanned outages can be categorized by the following:

Human Error
Unfortunately human error is probably the biggest factor in unplanned outages. Procedures may not be followed correctly, warnings may be missed, education may be lacking, or there even may be communication problems and misunderstandings between groups. These can all lead to unplanned outages which impact the business.

Software Problems
Application, operating system, middleware, or database complexities can result in unplanned outages. Every business is unique and interaction issues between different software components can cause problems.

Hardware Failure
At some point in time, mechanical devices will fail. Electrical components are subject to environment changes such as heat, humidity, and electrostatic discharge that can cause premature failure. Cable damage can occur and connections may loosen.

Environmental Issues
Power failures, network failures and air conditioning can cause a single system to become unavailable. Redundant measures can be taken to help address some of these issues, but there is a limit to what can be done.

Recovery from unplanned outages in a high availability environment is failover to a backup system. While the problem is being diagnosed and fixed, the business can continue to operate on the backup server.

Related concepts:

"Outage coverage" on page 14
What kind of outage is the business trying to protect against? Backup window reduction, planned maintenance, unplanned outages, or site disasters are events to consider when choosing a high availability solution.

Related information:

Shortening unplanned outages

Preventing unplanned outages

Recovering recent changes after an unplanned outage

Recovering lost data after an unplanned outage

Disaster recovery

Disaster recovery addresses the set of resources, plans, services and procedures to recover and resume mission critical applications at a remote site in the event of a disaster.

As a business grows, recovery from a disaster by tapes at a remote site may not be feasible within the required time defined by the business. Every location, although different has some type of disaster to worry about. Fire, tornadoes, floods, earthquakes, and hurricanes can have far reaching geographical impacts. This drives remote disaster sites to be further and further apart. In some cases industry regulations can also determine the minimum distance between sites.

Some important questions about designing for disasters are:

- What is the monetary impact to the business in case of a disaster?
- How soon can the business be back in production?
- At what point in time can I recover to?

- How much communication bandwidth can I afford?
- What disaster recovery solution is viable based on my distance requirements?

IBM i high availability solutions can be designed around the answers to these questions. This can be anything from making a single site more robust, contracting for use of a machine to restore tapes and run the business, or having a hot, up to date, backup at a remote site which is ready to take over production.

Related information:

Planning disaster recovery

Recovering your system

Backup window reduction

IBM i high availability solutions can reduce the time your system or services are unavailable during your backups. The time it takes to complete a backup from start to finish is called a backup window. The challenge is to back up everything in the window of time that you have.

The obvious techniques of reducing or eliminating the backup window involve either decreasing the time to perform the backup or decreasing the amount of data backed up. This includes the following:

Improved tape technologies
Faster and denser tape technologies can reduce the total backup time.

Parallel saves
Using multiple tape devices concurrently can reduce backup time by eliminating or reducing serial processing on a single device.

Saving to non-removable media
Saving to media that is faster than removable media, for example directly to direct access storage device (DASD), can reduce the backup window. Data can be migrated to removable media at a later time.

Data archiving
Data that is not needed for normal production can be archived and taken offline. It is brought online only when needed, perhaps for month-end or quarter-end processing. The daily backup window is reduced since the archived data is not included.

Saving only changed objects, daily backups exclude objects that have not changed during the course of the day.
The backup window can be dramatically reduced if the percentage of unchanged objects is relatively high.

Other save window reduction techniques leverage a second copy of the data (real or virtual). These techniques include:

Saving from a second system
Data resilience technologies, such as logical replication, that make available a second copy of the data can be used to shift the save window from the primary copy to the secondary copy. This technique can eliminate the backup window on the primary system. Therefore, it does not affect production since the backup processing is done on a second system.

Save while active
In a single system environment, the data is backed up using save processing while applications may be in production. To ensure the integrity and usability of the data, a checkpoint is achieved that ensures a point-in-time consistency. The object images at the checkpoint are saved, while allowing change operations to continue on the object itself. The saved objects are consistent with respect to one another so that you can restore the application environment to a known state. Save while active may also be deployed on a redundant copy achieved through logical replication. Employing such a technique can enable the save window to be eliminated effectively.

IBM System Storage FlashCopy
This technology uses the IBM System Storage function of FlashCopy on an independent disk pool basis. A point-in-time snapshot of the independent disk pool is taken on a single System Storage server. The copy of the independent disk pool is done within the System Storage server, and the host is not aware of the copy. Clustering enables bringing the copy on to the backup system for the purpose of doing saves or other offline processing. Clustering also manages bringing the second system back into the cluster in a nondisruptive fashion. Clustering supports multiple independent disk pools from the same system or multiple production systems being attached to the storage unit at the same time.

Related concepts:

"Outage coverage" on page 14
What kind of outage is the business trying to protect against? Backup window reduction, planned maintenance, unplanned outages, or site disasters are events to consider when choosing a high availability solution.

Related information:

Replication overview

Load balancing

IBM i high availability solutions can be used for load balancing. The most common technologies for workload balancing involve moving work to available resources. Contrast this with common performance management techniques that involve moving resources to work that does not achieve performance goals.

Example workload balancing technologies (each with its own HA implications) are:

Front end routers
These routers handle all incoming requests and then use an algorithm to distribute work more evenly across available servers. Algorithms may be as simple as sequential spreading (round robin) distribution or complex based on actual measured performance.

Multiple application servers
A user distributes work via some predefined configuration or policy across multiple application servers. Typically the association from requester to server is relatively static, but the requesters are distributed as evenly as possible across multiple servers.

Distributed, multi-part application
These applications work in response to end-user requests that actually flow across multiple servers. The way in which the work is distributed is transparent to the user. Each part of the application performs a predefined task and then passes the work on to the next server in sequence. The most common example of this type of workload balancing is a three-tiered application with a back-end database server.

Controlled application switchover
Work is initially distributed in some predetermined fashion across multiple servers. A server may host multiple applications, multiple instances of the same application, or both. If a given server becomes overloaded while other servers are running with excess capacity, the operations staff moves applications or instances of applications with associated data from the overloaded server to the under used server. Workload movement can be manual or automated based on a predetermined policy.

Related information:

TCP/IP routing and workload balancing

Creating peer CRGs

Components of high availability

High availability provides access to critical business applications and data in the event of a disruption in service. IBM i high availability solutions minimize and sometimes eliminate the effect of planned and unplanned outages and site-wide disasters for your business. The basis for IBM i high availability solutions is cluster technology.

A cluster is two or more systems (or operating system images) that share resources and processing and provide backup in the event of an outage. With clustering, high availability is viewed not as a series of identical copies of the same resource across these systems but rather a set of shared resources that continually provide essential services to users and applications.

Clustering does not provide a complete high availability solution all by itself, but it is the key technology on which all IBM i high availability solutions are based. Clustering infrastructure, called cluster resource services, provides the underlying mechanisms for creating and managing multiple systems and their resources as one unified computing entity. Clustering also monitors systems and resources defined in the high availability environment for failures and responds accordingly, depending on the type of outage. Clustering combines hardware and software to reduce the cost and effect of planned and unplanned outages by quickly restoring services when these outages occur. Although not instantaneous, cluster recovery time is rapid.

The following section defines the key components of a high availability solution.

Related tasks:

"Choosing a IBM i high availability solution" on page 21
After you have determined your business goals and requirements, you need to choose the right IBM i high availability solution that fits your business.

Application resilience

Application resilience can be classified by the effect to the user. Under an IBM i clustering infrastructure, application resiliency is controlled with an application Cluster Resource Group object (CRG). This CRG provides the mechanism, using an exit program, to control start, stop, restart, and switch of the application to back up systems. The entire application environment, including data replication and switchable devices can be controlled through the clustering infrastructure as a single entity.

Application resilience is classified into the following categories.

No application recovery
: After an outage, users must manually restart their applications. Based on the state of the data, users determine where to restart processing within the application.

Automatic application restart and manual repositioning within applications
: Applications that were active at the time of the outage are automatically restarted through the CRG exit program. The user must still determine where to resume within the application, based on the state of the data.

Automatic application restart and semi-automatic recovery
: In addition to the applications automatically restarting, the users are returned to some predetermined "restart point" within the application. The restart point may be, for example, a primary menu within the application. This is normally consistent with the state of the resilient application data, but the user might need to advance within the application to actually match the state of the data. Application changes are needed to save user state data. At sign on, the application detects the state of each user and determines if it needs to recover the application from the last saved state.

Automatic application restart and automatic recovery to last transaction boundary
: The user is repositioned within the application to the processing point that is consistent with the

last committed transaction. The application data and the application restart point match exactly. This category requires code changes in the application to save user states at the end of each commit cycle so the application knows where each user is in the application in case of a failure.

Full application resilience with automatic restart and transparent failover
In addition to being repositioned to the last committed transaction, the user continues to see exactly the same window with the same data as when the outage occurred. There is no data loss, signon is not required, and there is no perception of loss of server resources. The user perceives only a delay in response time. This category can only be obtained in an application with a client/server relationship.

Related concepts:

"Resilience requirements" on page 15
The business must identify what it is that needs to be protected when the system hosting the application experiences an outage. The resilience requirements are the set of applications, data and system environments required to be preserved across an outage of the production system. These entities remain available through a failover even when the system currently hosting them experiences an outage.

Related information:

Levels of application resiliency
Application resiliency can be customized to the level of resiliency that your business requires using the features of the IBM i clustering framework.

Making application programs resilient

Planning application resiliency

Data resilience

You can use a number of technologies to address the data resilience requirements described in the "Benefits of High Availability" section. Described below are the five key multisystem data resilience technologies. Keep in mind that multiple technologies can be used in combination to further strengthen your data resiliency.

Logical replication

Logical replication is a widely deployed multisystem data resiliency topology for high availability (HA) in the IBM i space. It is typically deployed through a product provided either by IBM or a high availability independent software vendor (ISV). Replication is run (through software methods) on objects. Changes to the objects (for example file, member, data area, or program) are replicated to a backup copy. The replication is near or in real time (synchronous remote journaling) for all journaled objects. Typically if the object such as a file is journaled, replication is handled at a record level. For such objects as user spaces that are not journaled, replication is handled typically at the object level. In this case, the entire object is replicated after each set of changes to the object is complete.

Most logical replication solutions allow for additional features beyond object replication. For example, you can achieve additional auditing capabilities, observe the replication status in real time, automatically add newly created objects to those being replicated, and replicate only a subset of objects in a given library or directory.

To build an efficient and reliable multisystem HA solution using logical replication, synchronous remote journaling as a transport mechanism is preferable. With remote journaling, IBM i continuously moves the newly arriving data in the journal receiver to the backup server journal receiver. At this point, a software solution is employed to "replay" these journal updates, placing them into the object on the backup server. After this environment is established, there are two separate yet identical objects, one on the primary server and one on the backup server.

With this solution in place, you can rapidly activate your production environment on the backup server by doing a role-swap operation. The figure below illustrates the basic mechanics in a logical replication environment.

A key advantage of this solution category is that the backup database file is live. That is, it can be accessed in real time for backup operations or for other read-only application types such as building reports. In addition, that normally means minimal recovery is needed when switching over to the backup copy.

The challenge with this solution category is the complexity that can be involved with setting up and maintaining the environment. One of the fundamental challenges lies in not strictly policing undisciplined modification of the live copies of objects residing on the backup server. Failure to properly enforce such a discipline can lead to instances in which users and programmers make changes against the live copy so that it no longer matches the production copy. If this happens, the primary and the backup versions of your files are no longer identical.

Another challenge associated with this approach is that objects that are not journaled must go through a check point, be saved, and then sent separately to the backup server. Therefore, the granularity of the real-time nature of the process may be limited to the granularity of the largest object being replicated for a given operation.

For example, a program updates a record residing within a journaled file. As part of the same operation, it also updates an object, such as a user space, that is not journaled. The backup copy becomes completely consistent when the user space is entirely replicated to the backup system. Practically speaking, if the primary system fails, and the user space object is not yet fully replicated, a manual recovery process is required to reconcile the state of the non-journaled user space to match the last valid operation whose data was completely replicated.

Another possible challenge associated with this approach lies in the latency of the replication process. This refers to the amount of lag time between the time at which changes are made on the source system and the time at which those changes become available on the backup system. Synchronous remote journal can mitigate this to a large extent. Regardless of the transmission mechanism used, you must adequately project your transmission volume and size your communication lines and speeds properly to help ensure that your environment can manage replication volumes when they reach their peak. In a high volume environment, replay backlog and latency may be an issue on the target side even if your transmission facilities are properly sized.

Switchable Device

A switchable device is a collection of hardware resources such as disk units, communication adapters, and tape devices that can be switched from one system to another. For data resilience, the disk units can be configured into a special class of auxiliary storage pool (ASP) that is independent of a particular host system. The practical outcome of this architecture is that switching an independent disk pool from one system to another involves less processing time than a full initial program load (IPL). The IBM i implementation of independent disk pools supports both directory objects (such as the integrated file system (IFS)) and library objects (such as database files). This is commonly referred to as switched disks.

The benefit of using independent disk pools for data resiliency lies in their operational simplicity. The single copy of data is always current, meaning there is no other copy with which to synchronize. No in-flight data, such as data that is transmitted asynchronously, can be lost, and there is minimal performance overhead. Role swapping or switching is relatively straight forward, although you might need to account for the time required to vary on the independent disk pool.

Another key benefit of using independent disk pools is zero-transmission latency which can affect any replication-based technology. The major effort associated with this solution involves setting up the direct-access storage device (DASD) configuration, the data, and application structure. Making an independent disk pool switchable is relatively simple.

Limitations are also associated with the independent disk pool solution. First, there is only one logical copy of the data in the independent disk pool. This can be a single point of failure, although the data should be protected using RAID 5, RAID 6 or mirroring. The data cannot be concurrently accessed from both hosts. Things such as read access or backup to tape operations cannot be done from the backup system. Certain object types, such as configuration objects, cannot be stored in an independent disk pool. You need another mechanism, such as periodic save and restore operations, clustering administrative domain or logical replication, to ensure that these objects are appropriately maintained.

Another limitation involves hardware associated restrictions. Examples include distance limits in the High Speed Link (HSL) loop technology and outages associated with certain hardware upgrades. The independent disk pool cannot be brought online to an earlier system. With this in mind, up-front system environment design and analysis are essential.

Switched logical unit (LUN) characteristics

Switched logical units allow data that is stored in the independent disk pool from logical units created in an IBM System Storage DS8000 or DS6000 to be switched between systems providing high availability.

A switched logical unit is an independent disk pool that is controlled by a device cluster resource group and can be switched between nodes within a cluster. When switched logical units are combined with IBM i clusters technology, you can create a simple and cost effective high availability solution for planned and some unplanned outages.

The device cluster resource group (CRG) controls the independent disk pool which can be switched automatically in the case of an unplanned outage, or it can be switched manually with a switchover.

A group of systems in a cluster can take advantage of the switchover capability to move access to the switched logical unit pool from system to system. A switchable logical unit must be located in an IBM System Storage DS8000 or DS6000 connected through a storage area network. Switched logical units operate similar to switched disks, but hardware is not switched between logical partitions. When the independent disk pool is switched the logical units within the IBM System Storage unit are reassigned from one logical partition to another.

Cross-site mirroring (XSM)

Geographic Mirroring

Geographic mirroring is a function of the IBM i operating system. All the data placed in the production copy of the independent disk pool is mirrored to a second independent disk pool on a second, perhaps remote, system.

The benefits of this solution are essentially the same as the basic switchable device solution with the added advantage of providing disaster recovery to a second copy at increased distance. The biggest benefit continues to be operational simplicity. The switching operations are essentially the same as that of the switchable device solution, except that you switch to the mirror copy of the independent disk pool, making this a straightforward HA solution to deploy and operate. As in the switchable device solution, objects not in the independent disk pool must be handled via some other mechanism and the independent disk pool cannot be brought online to an earlier system. Geographic mirroring also provides real-time replication support for hosted integrated environments such as Microsoft Windows and Linux. This is not generally possible through journal-based logical replication.

Since geographic mirroring is implemented as a function of the IBM i, a potential limitation of a geographic mirroring solution is performance impacts in certain workload environments.

When running input/output (I/O)-intensive batch jobs, some performance degradation on the primary system is possible. Also, be aware of the increased central processing unit (CPU) overhead required to support geographic mirroring, and the backup copy of the independent disk pool cannot be accessed while the data synchronization is in process. For example, if you want to back up to tape from the geographically mirrored copy, you must quiesce operations on the source system and detach the mirrored copy. Then you must vary on the detached copy of the independent disk pool on the backup system, perform the backup procedure, and then re-attach the independent disk pool to the original production host. Synchronization of the data that was changed while the independent disk pool was detached will then be performed. Your HA solution is running exposed, meaning there is no up-to-date second data set, while doing the backups and when synchronization is occurring. Using source and target side tracking will minimize this exposure.

Metro Mirror

Metro mirroring is a function of the IBM System Storage Server. The data that is stored in independent disk pools data is on disk units located in the System Storage Server. This solution involves replication at the hardware level to a second storage server using IBM System Storage Copy Services. An independent disk pool is the basic unit of storage for the System Storage Peer-to-Peer Remote Copy (PPRC) function. PPRC provides replication of the independent disk pool to another System Storage Server. IBM i provides a set of functions to combine the PPRC, independent disk pools, and IBM i cluster resource services for coordinated switchover and failover processing through a device cluster resource group (CRG).

You also have the ability to combine this solution with other System Storage-based copy services functions, including FlashCopy, for save window reduction.

Metro Mirror data transfer is done synchronously. You must also be aware of the distance limitations and bandwidth requirements associated with transmission times as with any solution when synchronous communications are used.

Global Mirror

Global Mirror uses the same base technology as Metro Mirror except the transmission of data is done in an asynchronous manner and FlashCopy to a third set of disks is required to maintain data consistency. Because this data transmission is asynchronous, there is no limit to how geographically dispersed the System Storage servers can be from each other.

Related concepts:

"Comparison of data resiliency technologies" on page 21
Data resiliency allows data to remain available to applications and users even though the system that originally hosted the data fails. Choosing the correct set of data resiliency technologies in the context of your overall business continuity strategy can be complex and difficult. It's important to understand the different data resilience solutions that can be used to enhance availability in multiple system environments. You can either choose a single solution or use a combination of these technologies to meet your needs. The following topics compare and contrast the different data resiliency technologies.

"Data resilience method comparison" on page 18
This table provides a brief description of the major characteristics of the solution that generates a copy of the data onto auxiliary storage.

"Resilience requirements" on page 15
The business must identify what it is that needs to be protected when the system hosting the application experiences an outage. The resilience requirements are the set of applications, data and system environments required to be preserved across an outage of the production system. These entities remain available through a failover even when the system currently hosting them experiences an outage.

Related information:

Planning data resiliency

Environment resilience

Environment resilience can be broken up into two sections, the physical environment, and the logical environment. The physical environment, which is really part of single system availability, focuses on things such as hardware redundancy, network topology, power infrastructure, and cooling capabilities. The logical environment is the application hosting and execution environment. It includes things like system settings, user profiles and system attributes that allow the user to run the application on multiple servers.

Physical Environment

The physical environment consists of single system availability features and the utilities required to adequately maintain a computer operating environment. These single system availability features are key to maintain a high availability environment. The system has many features to protect from hardware failures. The first component to protect is the disk subsystem. RAID 5, RAID 6, and disk mirroring are all offered protection mechanisms. One of these protection mechanisms is basically a requirement for any business.

Another component that should be protected is the network. This includes both redundant network adapters on the system, and multiple paths through the network over redundant network hardware for users and systems to use for communication.

The physical environment also includes the utility services needed to run the computer room. The system provides the capability to run on dual power cords. This means that each tower or rack has two power cables to plug into two different power outlets. This allows a computer room to have different breaker panels feeding each rack or tower. Due to the nature of public utility power, strong consideration should be given to protecting computer room power by an uninterruptible power supply or a generator.

Other considerations must be given to the physical room characteristics such as heating, cooling, air humidity, and air purity.

Logical Environment

The logical environment is the application runtime environment. This consists of the system attributes, system values, network configuration attributes, work management configuration and user profiles. These things must be the same for the application environment to operate the same way on the backup system as it does on the primary production system. Keeping these logical environmental values consistent across multiple systems can be done though a clustering administrative domain, logical replication, or a well defined manual process.

Related concepts:

"Resilience requirements" on page 15

The business must identify what it is that needs to be protected when the system hosting the application experiences an outage. The resilience requirements are the set of applications, data and system environments required to be preserved across an outage of the production system. These entities remain available through a failover even when the system currently hosting them experiences an outage.

Related information:

Planning environment resiliency

Simplicity

IBM i high availability addresses the three areas of customization, control, and automation with the goal of operational simplicity.

Customization

Every customer has a unique environment with unique requirements. The IBM i high availability architecture provides the framework from which each customer may design a solution based on their own application environment to meet their needs.

Control

The IBM PowerHA for i architecture provides for simple control over your high availability environment. With some level of customization, complete application environment activation, shut down, switchover, and failover is controlled through a simple to use clustering interface. The system operator now becomes the cluster operator.

Automation
High availability of the customer's production environment requires careful, coordinated operation of all aspects of the application in order to maintain resiliency and to quickly move from one server to another when a primary server goes down. The automation of the environment ensures that the pause in production is as short as possible. A major benefit of the automation capabilities in IBM PowerHA for i is the reduction of user error during failure scenarios. Reduced potential for user error improves the decision making process in case of a failure.

High availability criteria

IBM i high availability offers a choice of different technologies for data resiliency and application availability. Each of the different technologies has different characteristics. These characteristics should be matched with the unique requirements of each individual business application. The following parameters should be understood and considered when choosing which data resiliency technique is best for your business.

Budget

Each high availability solution has an associated cost. The cost for the solution must be compared to the benefit achieved for your business. When asked about a high availability solution, most customers will say that they want continuous availability with zero downtime. While this is technically possible, the cost of the protection offered by the solution may be too great.

The basic question behind how much resource should be given to a high availability solution is "What is the cost of an outage?" Backup sites, backup systems, backup copies of the application data have a cost, and an associated benefit for that cost. Until the actual cost of each unit of downtime is known, a true value cannot be assigned to the value of the additional benefit of the high availability solution to the customer.

Solution cost is the total cost of ownership which includes the initial cost to procure and deploy the solution, the ongoing costs to use the solution, and any cost/performance impacts. Cost is typically predicated on a thorough business impact analysis. The values are:

- Cost is not a factor.
- Cost has slight bearing on decision.
- Based on outage analysis, the solution cost must be contained within some budget.
- Cost is a significant factor in the decision.
- Unwilling or unable to spend anything on availability solution.

Uptime requirements

Up-time requirements refers to the total amount of time that the system is available for end-use applications. The value is stated as a percent of total scheduled working hours.

These are the uptime percentages and corresponding downtime values for customers that must be available all the time (24x365).

- Less than 90% (downtime of 876 or more hours (36 days)/year)
- 90 to 95% (downtime of 438 to 876 hours/year)
- 95 to 99% (downtime of 88 to 438 hours/year)
- 99.1 to 99.9% (downtime of 8.8 to 88 hours/year)
- 99.99% (downtime of about 50 minutes/year)

- 99.999% (downtime of about 5 minutes/year)

Typically the cost per outage hour is used as a determining factor in up-time requirements. When talking about unplanned outages, the uptime requirements must be based only off of the scheduled working hours. This means the cost of an outage should be calculated based on the worst possible time.

Outage coverage

What kind of outage is the business trying to protect against? Backup window reduction, planned maintenance, unplanned outages, or site disasters are events to consider when choosing a high availability solution.

Consideration must be given to the types of outages that you are trying to protect your business from.

Backup window reduction
In a single system environment, backing up the system is the most common contributor for planned system downtime. As the business need for application uptime increases, the amount of time to backup the data continues to get smaller. A high availability solution can give you the ability to perform offline saves. An offline save is the saving of application data from a backup copy. Each of the data resiliency technologies can offer different benefits for offline saving of data.

Planned maintenance
Planned maintenance is the time the system must be down to apply application, software, and hardware upgrades. When planned maintenance can no longer be scheduled around the scheduled working hours, a high availability solution can be implemented to allow for offline maintenance. With offline maintenance, the backup system is upgraded first. After the production environment is switched to the newly upgraded system, the old production system is then upgraded.

Unplanned outages
An unplanned outage is an outage that happens during scheduled working hours and can be due to human error, application/software failures, hardware failures, or utility failures and takes down the application environment. The high availability solution can switch the production environment to a backup.

Site disasters
A site disaster is typically thought of in terms of a natural disaster, and leads to the requirement of geographic dispersion between the systems in the high availability solution. In addition to natural disasters, there are also events such as chemical spills, terrorist attacks, and city wide loss of power that can impact your business site for a long period of time. The different high availability solutions have different time and distance characteristics. Consideration should be given to recovery time objectives (RTO) and if you need to run normal operations at the remote site, or just a subset of business processes.

Consideration should be given to the amount of disruption a user can tolerate. The application impact can be defined as the following:

- Not an issue. The availability of the application is the primary importance. Performance can be affected as long as availability solution delivers.
- Some performance degradation is acceptable
- Slight degradation in performance
- No perceived performance impact

Related concepts:

"Planned outages" on page 3
IBM i high availability can reduce the impact to your customers and users whenever you need to take systems or data offline to perform necessary maintenance tasks, such as nightly backups or the installation of new hardware or software.

“Unplanned outages” on page 3
IBM i high availability solutions can provide protection from unplanned outages caused by human error, software problems, hardware failures, and environmental issues.

Recovery time objective (RTO)

Recovery time objective (RTO) is the length of time that it takes to recover from an outage (scheduled, unscheduled, or disaster) and resume normal operations for an application or a set of applications.

The recovery time objective may be different for scheduled, unscheduled and disaster recovery outages. Different data resilience technologies will have differing RTO times. Possible values for RTO are:

- More than 4 days is acceptable
- 1 to 4 days
- Less than 24 hour
- Less than 4 hours
- Less than 1 hour
- Approaching zero (near immediate)

Recovery point objective (RPO)

Recovery point objective (RPO) is the point in time relative to the failure to which you need preservation of data. Data changes preceding the failure or disaster by at least this time period are preserved by recovery processing. Zero is a valid value and is equivalent to a "zero data loss" requirement.

RPO values are:

- Last save (weekly, daily, ...)
- Start of last shift (8 hours)
- Last major break (4 hours)
- Last batch of work (1 hour to tens of minutes)
- Last transaction (seconds to minutes)
- In-flight changes may be lost (power loss consistency)
- Zero data loss

Resilience requirements

The business must identify what it is that needs to be protected when the system hosting the application experiences an outage. The resilience requirements are the set of applications, data and system environments required to be preserved across an outage of the production system. These entities remain available through a failover even when the system currently hosting them experiences an outage.

The business choices are:

- Nothing needs to be made resilient
- Application data
- Application and system data
- Application programs
- Application state
- Application environment
- Preserve all communications and client connections

Related concepts:

"Application resilience" on page 7
Application resilience can be classified by the effect to the user. Under an IBM i clustering infrastructure, application resiliency is controlled with an application Cluster Resource Group object (CRG). This CRG provides the mechanism, using an exit program, to control start, stop, restart, and switch of the application to back up systems. The entire application environment, including data replication and switchable devices can be controlled through the clustering infrastructure as a single entity.

"Data resilience" on page 8
You can use a number of technologies to address the data resilience requirements described in the "Benefits of High Availability" section. Described below are the five key multisystem data resilience technologies. Keep in mind that multiple technologies can be used in combination to further strengthen your data resiliency.

"Environment resilience" on page 12
Environment resilience can be broken up into two sections, the physical environment, and the logical environment. The physical environment, which is really part of single system availability, focuses on things such as hardware redundancy, network topology, power infrastructure, and cooling capabilities. The logical environment is the application hosting and execution environment. It includes things like system settings, user profiles and system attributes that allow the user to run the application on multiple servers.

Automated failover and switchover

The business must define how much control is given up to automation during unplanned outages. IBM i high availability solutions have a customizable level of business interaction in failover processing. In case of a failure, the application can automatically failover to a backup system, including all application environment start.

Some customers want control over the failover processing. In this situation, the system will require a response for failover processing to occur. In a solution where user interaction is needed for failover, the think time (or time taken to make a decision to failover) is directly charged against the recovery time objective. The business must decide how much automation control will be given to the system during failover. The business should not take more time to make the decision to failover to the backup system, than it takes to actually do the failover.

Related concepts:

Switchover

Related information:

Failover

Distance requirements

Distance between systems, or geographic dispersion, has benefits but is gated by physical and practical limits. For a disaster recovery solution, there are always benefits in having geographic dispersion between the systems. Typically, the greater the distance between the systems, the greater the protection you will have from area wide disasters. However, this distance will come with application environment impacts.

When distance is added to a data replication solution, latency is introduced. Latency is the added time it takes for data to reach the target system. The further the systems are apart, the more latency (time) is added to the data transmission. There are two types of communication transmissions, synchronous and asynchronous.

Synchronous communications for data resiliency requires an acknowledgement from the target system that the data transmission has been received before continuing. This process guarantees no loss of in flight data from the source to the target in case of a failure. However, the latency, or time waiting for the acknowledgement can affect application performance.

Asynchronous communications for data resiliency does not require an acknowledgement from the target system to continue data transmission. Because this mechanism does not wait for a handshake, data sent close to the time of failure may be lost. This is known as *in flight data loss*.

The application, amount of data being sent, and geographic dispersion of the systems, will determine the needed transport mechanism for your high availability solution.

Related information:

Geographic mirroring

Planning cross-site mirroring

Scenario: Cross-site mirroring with metro mirror

Scenario: Cross-site mirroring with global mirror

Number of backup systems

Different data resilience technologies offer differing numbers of possible backup systems and copies of application data.

In a two system environment (single backup), planned maintenance will leave your business exposed. If a failure happens during this time frame, you will not have failover capability. In this situation, business continuity can be maintained by adding another backup system. The number of backup systems, and needed data sets will help determine the data resilience technology required for your business.

Access to a secondary copy of the data

Different data resilience technologies have different restrictions to the backup data set. Access to the backup data set requirements indicates the level of access that is required to secondary copies of the data for other work activity off-loaded from primary copies, such as saves and queries/reports. You should consider the frequency, duration, and what type of access is needed for the backup copy of the data.

Possible requirements can be:

- None
- During non-production periods
- Infrequent but during normal production for short (seconds to minutes) durations
- Infrequent but during normal production for long durations
- Frequently during production for short durations
- Frequently during production for long durations
- Nearly all the time (near continuous)

Related information:

Backup from a second copy

System performance

Implementing high availability may have performance implications. The requirements of the business may determine what data resilience technology is required.

Implementing high availability comes with a varying performance overhead. Journaling for logical replication and geographic mirror processing require system resources for normal runtime. In addition, synchronous remote journaling, geographic mirroring in synchronous delivery mode, and metro mirror technologies all run in a synchronous communication mode. This synchronous mode produces a latency based on distance and network topology, which will impact the application environment. The business requirements along with testing will help determine which solution is viable for the customer.

Geographic mirroring also supports an asynchronous delivery mode that may require additional resources such as CPU and main storage.

Switchover and failover processing are not instantaneous and also have an associated overhead. Each technology has different characteristics for bringing a dataset, or entire application environment online for processing.

Related information:

Managing system performance

System values: Performance overview

Data resilience method comparison

This table provides a brief description of the major characteristics of the solution that generates a copy of the data onto auxiliary storage.

Table 1. Comparison of data resilient technologies that can be used with clusters. Learn about characteristics of different data resiliency technologies to help you determine the best solution for your cluster.

Data resiliency technologies	Logical replication software product	Switched disks	Cross-site mirroring with geographic mirroring	Cross-site mirroring with IBM System Storage Copy Services	Switched logical unit
Primary use	HA and DR	HA	HA and DR	HA and DR	HA
Characteristic of replication mechanism	• Object based replication of selected objects • Replicate based on journal changes • Supports independent disk pools or *SYSBAS objects	• Single copy of data, switchable between systems • Independent disk pool data	• Memory page-level replication by IBM i • Independent disk pool data • Non-identical source and target independent disk pool configuration	• External DASD sector level replication • Physical copy of an Independent disk pool based on disk I/O (cache based)	• Single copy of data, switchable between systems • Independent disk pool data
Budget/cost factors	• HA independent software vendor (ISV) software • Duplicate DASD • Network bandwidth	• IBM PowerHA for i licensed program • IBM i Option 41 (HA Switchable Resources) • Single DASD copy	• IBM PowerHA for i licensed program • IBM i Option 41 (HA Switchable Resources) • Duplicate DASD • Network bandwidth	• IBM PowerHA for i licensed program • IBM i Option 41 (HA Switchable Resources) • IBM System Storage Server(s) • Metro or Global Mirror • Duplicate or triplicate DASD • Network bandwidth	• IBM PowerHA for i licensed program • IBM i Option 41 (HA Switchable Resources) • IBM System Storage Server(s) • Single DASD copy
Outage coverage	Backup window, planned, unplanned, disaster	Planned, unplanned	Backup window, planned, unplanned, disaster	Backup window, planned, unplanned, disaster	Planned, unplanned

Table 1. Comparison of data resilient technologies that can be used with clusters (continued). Learn about characteristics of different data resiliency technologies to help you determine the best solution for your cluster.

Data resiliency technologies	Logical replication software product	Switched disks	Cross-site mirroring with geographic mirroring	Cross-site mirroring with IBM System Storage Copy Services	Switched logical unit
Recovery time objective (RTO)	• Apply lag • Replication switch • Journal settings	• Hardware switch time • Independent disk pool vary on time • System-managed access-path protection (SMAPP) and journal settings	• Independent disk pool vary on time • SMAPP and journal settings	• Independent disk pool vary on time • SMAPP and journal settings	• Communication time with IBM System Storage Server • Independent disk pool vary on time • System-managed access-path protection (SMAPP) and journal settings
Recovery point objective (RPO)	• Mixed, audit and data journal • May lose async in-flight data and non-journaled objects • Data is journaled	• All data written to Independent disk pool • Data should be journaled • May lose memory resident data (non journaled)	• All data written to Independent disk pool • Data should be journaled • May lose memory resident data (non journaled)	• All data written to Independent disk pool • Data should be journaled • May lose memory resident data (non journaled)	• All data written to Independent disk pool • Data should be journaled • May lose memory resident data (non journaled)
Automated failover and switchover	IBM i Cluster Controlled	IBM i Cluster Controlled	IBM i Cluster Controlled	IBM i Cluster Controlled	IBM i Cluster Controlled
Distance (Geographic Dispersion)	• Synchronous is limited by performance impacts • Asynchronous is virtually unlimited	• HSL cable length limited • 15 Meter (Copper) • 250 Meter (Fiber)	• Synchronous delivery mode is limited by performance impacts • Asynchronous delivery mode is virtually unlimited	• Synchronous (Metro Mirror) is limited by performance impacts • Asynchronous (Global Mirror) is virtually unlimited	• Fiber channel cable length limited
Number of backup systems	• 1<=n<=127 (or BP max) • Can be combined with switched disks	n=1 physical backup system (any and all partitions on both physical systems)	1<=n<=3 physical backup systems (any and all partitions on all 4 physical systems)	1<=n<=3 (with cascading PPRC) (any and all partitions all 4 physical systems)	n=1 physical backup system (any and all partitions on both physical systems)

Table 1. Comparison of data resilient technologies that can be used with clusters (continued). Learn about characteristics of different data resiliency technologies to help you determine the best solution for your cluster.

Data resiliency technologies	**Logical replication software product**	**Switched disks**	**Cross-site mirroring with geographic mirroring**	**Cross-site mirroring with IBM System Storage Copy Services**	**Switched logical unit**
Access to secondary copy of data	• Typically read only • Lag time in data currency (apply process on target)	No concurrent access since there is only one copy of data	• No, requires detach and partial resynchronization • Second copy current at detach time	• No concurrent access • Copy current with Metro Mirror or latest consistency group with Global Mirror • Point in time access with FlashCopy function	No concurrent access since there is only one copy of data
Risks	• Loss of in-flight data for async journaling, and all non-journaled objects • Monitoring logical object replication environment	One set of data (single point of failure)	Synchronization might cause lengthy unprotection condition (not protected when synchronizing)	Additional environment complexity due to external storage boxes	One set of data (single point of failure)

Note: In some cases, the distance limits are stated as "virtually unlimited". While this is technically true, the actual distance limits are gated by response time degradation tolerances, throughput impacts, characteristics of the communications fabrics, and other factors.

Related concepts:

"Data resilience" on page 8
You can use a number of technologies to address the data resilience requirements described in the "Benefits of High Availability" section. Described below are the five key multisystem data resilience technologies. Keep in mind that multiple technologies can be used in combination to further strengthen your data resiliency.

"Comparison of data resiliency technologies" on page 21
Data resiliency allows data to remain available to applications and users even though the system that originally hosted the data fails. Choosing the correct set of data resiliency technologies in the context of your overall business continuity strategy can be complex and difficult. It's important to understand the different data resilience solutions that can be used to enhance availability in multiple system environments. You can either choose a single solution or use a combination of these technologies to meet your needs. The following topics compare and contrast the different data resiliency technologies.

System-managed access-path protection

Choosing a IBM i high availability solution

After you have determined your business goals and requirements, you need to choose the right IBM i high availability solution that fits your business.

Related concepts:

"Components of high availability" on page 7
High availability provides access to critical business applications and data in the event of a disruption in service. IBM i high availability solutions minimize and sometimes eliminate the effect of planned and unplanned outages and site-wide disasters for your business. The basis for IBM i high availability solutions is cluster technology.

Related information:

Managing a high availability solution

Managing cluster resource groups (CRG)

Levels of application resiliency

Application resiliency can be customized to the level of resiliency that your business requires using the features of the IBM i clustering framework.

Recovery Time Objective (RTO) for your business plays directly into the level of application resiliency that is needed. As it is defined in the Components of High Availability topic, there are different levels of application resiliency. These application resiliency levels range from no application recovery, where a system operator must start the application manually, to uninterrupted service, where the user may not even know that an outage even happened. Your business requirements for the application to be available to the user after a failure sets the requirements for how much automation the resilient application must recover in the event of a system failure.

The IBM i clustering framework gives the ability to automate the application recovery for different types of failures. The amount of automation possible depends on the amount of coding to automate manual procedures and the type of application your business is using. To maximize application resiliency, all manual switchover/failover steps must be automated with exit programs, and the application must be a client-server type application where the application availability is separated from the application data availability.

Comparison of data resiliency technologies

Data resiliency allows data to remain available to applications and users even though the system that originally hosted the data fails. Choosing the correct set of data resiliency technologies in the context of your overall business continuity strategy can be complex and difficult. It's important to understand the different data resilience solutions that can be used to enhance availability in multiple system environments. You can either choose a single solution or use a combination of these technologies to meet your needs. The following topics compare and contrast the different data resiliency technologies.

Related concepts:

"Data resilience" on page 8
You can use a number of technologies to address the data resilience requirements described in the "Benefits of High Availability" section. Described below are the five key multisystem data resilience technologies. Keep in mind that multiple technologies can be used in combination to further strengthen your data resiliency.

"Data resilience method comparison" on page 18
This table provides a brief description of the major characteristics of the solution that generates a copy of the data onto auxiliary storage.

Logical replication characteristics

Logical replication makes and keeps the objects on your production and backup systems identical. For journaled objects, the transactional operations on the source are duplicated on the target by applying journal changes. For non-journaled data, the changed data is saved and then written on the target. These apply processes on the target are provided by the logical replication.

Characteristics of logical replication

- Two or more copies of the data, eliminating single point of failure
- Offline saves and queries are allowed while maintaining data resiliency. For saves, the apply processes are suspended but the replication of changes to the target system continue during the save
- Geographically dispersed backup system(s) utilizing asynchronous remote journaling
- High Availability independent software vendor (ISV) or IBM iCluster for i product, with logical replication product utilizing remote journaling
- No in flight data loss with synchronous remote journaling for journaled objects. Synchronous remote journaling may limit the geographic dispersion of backup system(s)
- Possible in flight data loss with asynchronous remote journaling for journaled objects
- Data currency may be an issue. Although data is replicated real time or near real time, the logical replication process will lag behind the source system.
- Solution may offer synchronization issues. Not all objects can be journaled so separate technologies are used to replicate the entire data set
- Second set of disks are needed for each backup copy of data
- Can be used in conjunction with the IBM i switched disk technology
- Solution needs to be monitored for data synchronization issues
- System overhead exists for the necessary journal function on the source system
- System overhead exists for the data apply process on the target system
- Cost associated with a second set of disk
- Replication is at a data transaction level

Related information:

Planning for logical replication

Switched disk characteristics

Switched disk allow data that is stored in the independent disk pool to be switched between systems provides a level of data resiliency.

Characteristics for switched disks

- All data maintained in the independent disk pool can be switched and made available on backup system
- No data synchronization issues
- Single set of data minimizing cost for disk
- Potential single point of failure depending on the disk configuration
- Single site solution using HSL copper cables (15 meter maximum length)
- Switchover and failover include vary on time before independent disk pool data available
- Can be used in conjunction with the other technologies

Related information:

Switchable devices

Independent disk pools

Benefits of independent disk pools

Independent disk pool examples

Switched logical unit characteristics

Switched logical units allow data that is stored in the independent disk pool from logical units created in an IBM System Storage DS8000 or DS6000 to be switched between systems providing high availability.

Characteristics of Switched logical units

- All data maintained in the independent disk pool can be switched and made available on backup system.
- No data synchronization issues.
- Single set of data minimizing cost for disk.
- Single point of failure for data in the independent disk pool.
- Single site solution using a storage area network and IBM System Storage.
- Switchover and failover include vary on time before independent disk pool data available.
- Can be used with the other technologies

Geographic mirroring characteristics

Geographic mirroring enables you to mirror data on disks at sites that can be separated by a geographic distance.

Characteristics of geographic mirroring

- All data maintained in the independent disk pool will be replicated to a second copy of the data on a second system.
- Independent disk pool data synchronization is maintained by the system.
- Application can be switched to backup system and operate on the independent disk pool copy.
- Two copies of the data eliminating single point of failure.
- Second copy of data can be geographically dispersed.
- Data transmission is a synchronous delivery mode. No in flight data loss is possible. Synchronous delivery mode over a distance my impact application performance due to communication latency.
- Data transmission is an asynchronous delivery mode. In flight data loss is possible. There will be some overhead on the source system, however, individual transaction performance will not be impacted by the data transmission.
- Data transmission over 1 to 4 TCP/IP communication lines for throughput and redundancy.
- It is also recommended that a separate line be used for the clustering heartbeat since sharing the heartbeat with data port can cause contention and time outs.
- It is recommended that all of the data port lines in use be of the same speed since round robin will effectively cause data port to run the speed of the slowest line.
- Offline saves and queries to backup copy of the data while backup dataset is detached.
- Data resiliency not maintained while backup dataset is detached. Data resiliency is resumed after partial or full resynchronization has completed
- Can be used in conjunction with the IBM i switch disk technology.
- System performance overhead is associated with running geographic mirroring.
- It is strongly recommended that you configure separate main storage pools or user jobs that access independent disk pools in order to prevent those jobs from contending with other jobs on the system and using more main storage than desired. More specifically, independent disk pool jobs should not use the machine pool or base pool. If independent disk pool jobs use the same memory as jobs that are not accessing the independent disk pools, independent disk pool jobs can monopolize the memory pool, lock out other jobs, and in extreme situations deadlock the system. Exposure for this situation is greater when using geographic mirroring.
- Journaled objects in the independent disk pool will guarantee data update to target system.

- Simple monitoring of mirror process.
- Cost associated with a second set of disk.
- Replication is at a memory page level managed by IBM i.

Related information:

Geographic mirroring

Planning geographic mirroring

Managing geographic mirroring

Geographic mirroring messages

Scenario: Cross-site mirroring with geographic mirroring

Metro Mirror characteristics

A synchronous form of hardware replication managed by a System Storage Server.

Characteristics of Metro Mirror

- IBM System Storage Server solution integrated with PowerHA framework.
- Second copy of data can be geographically dispersed a short to medium distance.
- Two System Storage Servers or two datasets on the same System Storage Server are required.
- Cost is associated with a second set of disk.
- Offline saves and queries possible while replication is suspended or from a point in time copy of the data.
- Data resiliency not maintained while backup dataset is detached. Data resiliency is resumed after resynchronization has completed.
- Data transmission is a synchronous process. No in flight data loss is possible.
- Synchronous data replication process may impact application performance if communications bandwidth not properly sized or if the distance is too great.
- No system overhead to run Metro Mirror, it is handled by the storage server.
- Journaling the objects in the independent disk pool ensures that those changes are forced quickly to disk where they are then replicated to the target system.
- Replication of the independent disk pool data is at the disk sector level between the disks on the two Storage Servers. All objects in the independent disk pool will be synchronized.
- Multiple fiber channel communication lines available for redundancy and increased bandwidth.

Related information:

Metro mirror

Planning metro mirror

Managing metro mirror

Scenario: Cross-site mirroring with metro mirror

Global Mirror characteristics

An asynchronous form of hardware replication managed by a System Storage Server.

Characteristics of Global Mirror

- IBM System Storage Server solution integrated with PowerHA cluster framework.
- Second copy of data can be geographically dispersed over potentially large distances.
- Two System Storage Servers are required.
- Two copies of the data on the target System Storage Server required to ensure consistency of data across distances.
- Offline saves and queries possible from a point in time copy of the data, maintaining data resiliency.
- Data transmission is an asynchronous process. In flight data loss is possible.

- Asynchronous data replication process does not affect application performance.
- Replication of the independent disk pool data is at the disk sector level between the disks on the two Storage Servers. All objects in the independent disk pool will be synchronized.
- Cost is associated with a second and third set of disk.
- No system overhead to run Metro Mirror, it is handled by the storage server.
- Journaling the objects in the independent disk pool ensures that those changes are forced quickly to disk where they are then replicated to the target system.
- Multiple fiber channel communication lines available for redundancy and increased bandwidth.

Related information:

Global mirror

Planning global mirror

Managing global mirror

Scenario: Cross-site mirroring with global mirror

High availability management

To plan, configure, and manage a complete high availability solution requires a set of management tools and offerings. With IBM i systems, several choices exist for high availability management.

Depending on your needs and requirements, high availability management provides graphical interfaces, commands, and APIs that can be used to create and manage your environment. You can also choose to use an IBM business partner application. Each of these choices of high availability management tools has their advantages and limitations.

IBM PowerHA for i interfaces

IBM PowerHA for i , licensed program number (5770-HAS), is an end-to-end high availability offering. When combined with independent auxiliary storage pools (iASPs) and HA Switchable Resources (HASR - Option 41). It enables a complete solution to be deployed via IBM DS8000 storage server or internal disk. PowerHA provides several interfaces to configure and manage high availability solutions and technology.

The IBM PowerHA for i licensed program, is an end-to-end high availability offering. When combined with independent auxiliary storage pools (iASPs) and HA Switchable Resources (HASR - Option 41). It enables a complete solutions to be deployed via IBM DS8000 storage server or internal disk.

The IBM PowerHA for i licensed program provides two graphical interfaces that allows you to configure and manage a high availability solution. This product also provides corresponding commands and APIs for functions related to high availability technologies. With this licensed program, high availability administrators can create and manage a high availability solution to meet their business needs, using interfaces that fit their skills and preferences. You can also work with multiple interfaces seamlessly, using graphical interfaces for some tasks and commands and APIs for others.

The IBM PowerHA for i licensed program provides the following interfaces:

High Availability Solutions Manager graphical interface
This graphical interface allows you to select from several IBM i supported high availability solutions. This interface validates all technology requirements for your selected solution, configures your selected solution and the associated technologies and provides simplified management of all the high availability technologies that comprise your solution.

Cluster Resource Service graphical interface
This graphical interface provides an experienced user more flexibility in customizing a high availability solution. It allows you to configure and manage cluster technologies, such as CRGs. You can also configure some independent disk pools from this interface when they are used as part of a high availability solution.

IBM PowerHA for i commands
These commands provide similar functions but are available through a command-line interface.

IBM PowerHA for i APIs
These APIs allow you to work with functions related to independent disk pools, PowerHA version information, mirroring technologies, and cross-site mirroring.

Related information:
Installing IBM System i High Availability Solutions Manager licensed program

High Availability Solutions Manager graphical interface:

IBM PowerHA for i licensed program provides a solution-based approach to setting up and managing high availability with a graphical interface called High Availability Solutions Manager. This interface allows high availability administrators to select, configure, and manage a predefined high availability solution which are based on IBM i high availability technologies, such as independent disk pools and clusters.

The High Availability Solutions Manager graphical interface guides users through the process of selecting, configuring, and managing a high availability solution. The user must complete each step before continuing to subsequent steps. When PowerHA is installed, you can access the High Availability Solutions Manager graphical interface in the IBM Navigator for i console. The High Availability Solutions Manager graphical interface has the following features:
- Provides a flash demo that provides overview for each solution
- Provides a choice of several predefined IBM solutions using IBM i high availability technologies
- Verifies hardware and software requirements before setting up the selected high availability solution
- Provides a customized list of missing requirements
- Provides easy configuration of your selected high availability solution
- Provides simplified management of your selected high availability solution

The High Availability Solutions Manager graphical interface provides an easy-to-use, guided approach to setting up high availability. This interface ensures and validates prerequisites, configures all necessary technologies for the selected solution, and tests the set up. This management solution interface is best for smaller businesses who want simpler solutions that require fewer resources.

Cluster Resource Services graphical interface:

IBM PowerHA for i licensed program provides a graphical interface that lets you perform tasks with IBM i high availability technologies to configure and manage a high availability solution.

The Cluster Resource Services graphical interface allows you to build and customize a high availability solution that meets your needs. This interface provides a task-based approach for setting up and managing your high availability solution. Instead of a single predefined solution to choose, you can create a customized high availability solution by separately creating each element of the high availability solution. With the Cluster Resource Services graphical interface you can create and manage clusters, cluster resource groups, device domains, cluster administrative domains, and perform switchovers.

Depending on the type of high availability solution you are creating, you may need to configure additional technologies, such as geographic mirroring or independent disk pools, which are outside of the Cluster Resource Services graphical interface. You can also use a combination of commands and graphical interface functions when building and managing your high availability solution.

Related information:
Implementing high availability with the task-based approach

IBM PowerHA for i commands:

IBM PowerHA for i licensed program provides IBM i command line interfaces to configure and manage your high availability solution.

The PowerHA commands consists of the following categories:
- Cluster administrative domain commands
- Monitored resource entries commands
- Cluster commands
- Commands and APIs for working with copies of independent disk pools

Related information:
IBM System i High Availability Solutions Manager commands

IBM PowerHA for i APIs:

IBM PowerHA for i provides APIs that can be used to implement IBM System Storage mirroring technologies and cross-site mirroring functions that can be used by IBM i application providers or customers to enhance their application availability.

To use these APIs, you must have the IBM PowerHA for i licensed product installed on your systems in your high availability environment. The following APIs are provided:
- Change High Availability Version (QhaChangeHAVersion) API
- List High Availability Information (QhaListHAInfo) API
- Retrieve High Availability Information (QhaRetrieveHAInfo) API
- Retrieve ASP Copy Information (QyasRtvInf) API

Option 41 (HA Switchable Resources)

Option 41 (HA Switchable Resources) is required when using several IBM i high availability management interfaces and functions require its installation in order to be used.

Option 41 (High Availability Switchable Resources) is required if you plan to use the following interfaces:
- IBM PowerHA for i licensed program.
  - High Availability Solutions Manager graphical interface
  - Cluster Resource Services graphical interface
  - IBM PowerHA for i commands
  - IBM PowerHA for i APIs

Option 41 is also required for the following functions:
- Create and manage switched disk using device domains
- Create and manage cross-site mirroring using devices domains

Advanced node failure detection

Cluster Resource Services can now use Hardware Management Console (HMC) or a Virtual I/O Server (VIOS) partition to detect when a cluster node fails.

Detection of node failures can be accomplished using a Hardware Management Console (HMC) or Virtual I/O Server (VIOS) partition.

To add advanced node failure detection for HMC with a CIM server or IVM, which uses a CIM server, a cluster node that is configured with a cluster monitor must install the following software.
- IBM i option 33, IBM Portable Application Solutions Environment for i

- 5733-SC1, IBM Portable Utilities for i
- 5733-SC1 option 1, OpenSSH, OpenSSL, zlib
- 5770-UME, IBM Universal Manageability Enablement for i

To add advanced node failure detection for HMC with a representational state transfer (REST) server. A cluster node that is configured with a cluster monitor must have the following software installed.

- IBM i option 33, IBM Portable Application Solutions Environment for i
- IBM i option 3, Extended Base Directory Support
- 5733-SC1, IBM Portable Utilities for i (Only required for initial configuration of a cluster monitor.)
- 5733-SC1 option 1, OpenSSH, OpenSSL, zlib (Only required for initial configuration of a cluster monitor.)
- HMC version must be V8R8.5.0 or later. This is the first version of HMC to support the REST server.
- PowerHA new function cluster monitor HMC REST support PTFs

High availability function in the base operating system

Some cluster CL commands and all Cluster APIs exist in the base IBM i.

Cluster commands

The following cluster commands will remain in QSYS for debugging purposes and for deleting cluster-related objects:

- Delete Cluster Resource Group (DLTCRG) command
- Dump Cluster Trace (DMPCLUTRC) command
- Change Cluster Recovery (CHGCLURCY) command
- Start Clustered Hash Table Server (STRCHTSVR) command
- End Clustered Hash Table Server (ENDCHTSVR) command

Cluster APIs

You can write your own custom application to configure and manage your cluster by using Cluster APIs. These APIs take advantage of the technology provided by cluster resource services provided as a part of IBM i. New enhanced functions are included in the IBM PowerHA for i commands which are provided by the IBM PowerHA for i licensed program.

QUSRTOOL

In IBM i 6.1, a majority of the cluster resource services commands were moved from QSYS to theIBM PowerHA for i licensed program. A V5R4 version of the cluster resource services command source and the source for the command processing program is available in QUSRTOOL. These QUSRTOOL commands can be useful in some environments. See the member TCSTINFO in the file QUSRTOOL/QATTINFO for more information about these example commands. An example application CRG exit program source is also included in the QUSRTOOL library. The sample source code can be used as the basis for writing an exit program. Sample source, TCSTDTAEXT, in file QATTSYSC contains a source for a program to create the QCSTHAAPPI and QCSTHAAPPO data areas, and QACSTOSDS (object specifier) file.

To save space, the QUSRTOOL library is shipped with many save files. To convert the save files to source physical files, run these commands:

```
CALL QUSRTOOL/UNPACKAGE ('*ALL ' 1)
CRTLIB TOOLLIB TEXT('Commands from QUSRTOOL')
CRTCLPGM PGM(TOOLLIB/TCSTCRT) SRCFILE(QUSRTOOL/QATTCL)
CALL TOOLLIB/TCSTCRT ('TOOLLIB ')
```

These commands were created in the library TOOLLIB.

Note: Commands and programs in QUSRTOOL are provided 'AS IS'. Therefore, they are not subject to APARs.

Cluster middleware IBM Business Partners and available clustering products

In addition to IBM PowerHA for i, there are other cluster management products available.

IBM iCluster for i, as well as other products, provide software solutions for replication and cluster management functions. Most of these solutions are based on logical replication. Logical replication uses remote journal or similar technology to transfer object changes to a remote system, where they are applied to the target objects. In addition to PowerHA management solutions, you can purchase other cluster middleware products which use logical replication technology. Those products typically also include a management interface.

Related information:

Planning logical replication

Related information for High availability overview

Product manuals, IBM Redbooks® publications, Web sites, and other information center topic collections contain information that relates to the Implementing high availability topic collection. You also can find related information about implementing independent disk pools, cross-site mirroring, and disaster recovery. You can view or print any of the PDF files.

IBM Redbooks

- Implementing IBM PowerHA for i
- Availability Management: A Guide to planning and implementing Cross-Site Mirroring on System i5®
- Data Resilience Solutions for IBM i5/OS High Availability Clusters
- Clustering and IASPs for Higher Availability
- High Availability on the AS/400 System: A System Manager's Guide
- IBM eServer™ iSeries Independent ASPs: A Guide to Moving Applications to IASPs

Web sites

- http://www-03.ibm.com/systems/power/software/availability/i/index.html (www-03.ibm.com/systems/power/software/availability/i/index.html)

  IBM site for High Availability and Clusters

Other information

- Disk management
- Availability roadmap

Appendix. Notices

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.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan, Ltd.
3-2-12, Roppongi, Minato-ku, Tokyo 106-8711

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.

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.

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

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation

Software Interoperability Coordinator, Department YBWA
3605 Highway 52 N
Rochester, MN 55901
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement, IBM License Agreement for Machine Code, or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

All IBM prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs 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.

Each copy or any portion of these sample programs or any derivative work, must include a copyright notice as follows:

© (your company name) (year). Portions of this code are derived from IBM Corp. Sample Programs. © Copyright IBM Corp. _enter the year or years_.

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

Programming interface information

This High availability overview publication documents intended Programming Interfaces that allow the customer to write programs to obtain the services of IBM i.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., 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.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

INFINIBAND, InfiniBand Trade Association, and the INFINIBAND design marks are trademarks and/or service marks of the INFINIBAND Trade Association.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

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

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Red Hat, the Red Hat "Shadow Man" logo, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc., in the United States and other countries.

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

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

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

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

Terms and conditions

Permissions for the use of these publications is granted subject to the following terms and conditions.

Personal Use: You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative works of these publications, or any portion thereof, without the express consent of IBM.

Commercial Use: You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

Index

Readers' Comments — We'd Like to Hear from You

IBM i
Availability
High availability overview
7.1

We appreciate your comments about this publication. Please comment on specific errors or omissions, accuracy, organization, subject matter, or completeness of this book. The comments you send should pertain to only the information in this manual or product and the way in which the information is presented.

For technical questions and information about products and prices, please contact your IBM branch office, your IBM business partner, or your authorized remarketer.

When you send comments to IBM, you grant IBM a nonexclusive right to use or distribute your comments in any way it believes appropriate without incurring any obligation to you. IBM or any other organizations will only use the personal information that you supply to contact you about the issues that you state on this form.

Comments:

Thank you for your support.

Submit your comments using one of these channels:
- Send your comments to the address on the reverse side of this form.
- Send a fax to the following number: United States and Canada: 1-800-937-3430
  Other countries or regions: 1-507-253-5192
- Send your comments via email to: RCHCLERK@us.ibm.com

If you would like a response from IBM, please fill in the following information:

Name	Address
Company or Organization	
Phone No.	Email address

Readers' Comments — We'd Like to Hear from You

Cut or Fold Along Line

Fold and Tape | **Please do not staple** | Fold and Tape

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

BUSINESS REPLY MAIL

FIRST-CLASS MAIL PERMIT NO. 40 ARMONK, NEW YORK

POSTAGE WILL BE PAID BY ADDRESSEE

IBM CORPORATION
Attn Bldg 004-2 ID Clerk
3605 HWY 52 N
Rochester, MN 55901-7829

Fold and Tape | **Please do not staple** | Fold and Tape

Cut or Fold Along Line

Printed in USA