High Availability for IBM i

A Technical Review of Software- and Hardware-Based Solutions

Abstract

This paper briefly describes what you need to know in order to make an informed decision with regard to IBM i high availability strategies so that your business requirements for Recovery Time Objective (RTO) and Recovery Point Objective (RPO) are not compromised. With the advent of the unified Power Systems platform, common storage platforms, and virtualization, maintaining high availability protection for an IBM i environment requires careful thought and clear understanding of its underlying architecture and technologies.

Executive Summary

On the IBM i platform, there are fundamentally two ways to protect your business from unplanned business outages due to server failure:

1. **Logical replication** to maintain a hot backup copy of the environment that is fully available to the business at the point of switchover
2. **Hardware HA** to maintain an offline backup copy of the environment that is ready and waiting to reconstruct the environment with the latest possible copy of the data

Both approaches are strategic for IBM i and Power Systems hardware, and both have advantages as well as challenges that must be overcome in order to meet the goals of no loss of data (RPO) and the fastest possible recovery (RTO).

Logical replication is the term used in this paper to define journal-based data resiliency. Every change to the data is logged, sent to the backup system, and immediately applied there. The OS journal function ensures that the journal record reflects the most recent state of the data.

The IBM i Remote Journaling feature ensures the data is securely written to the backup system so that the best possible RPO is achieved, while RTO will depend upon how quickly the high availability solution applies those journaled transactions, in the correct order, to the production database on the backup system.

Hardware HA is the term used in this paper to define the disk-based replication native to the IBM TotalStorage SAN system, which employs sector-based replication capability between two SAN environments. The IBM i Single-Level Store memory architecture employs a memory page–based system that maps to a string of disk sectors. When the memory page is written to disk, the first SAN makes a sector-by-sector copy on the second SAN. In order to preserve the unique architecture of IBM i, which treats memory and disk storage as one, the data being copied is encapsulated and treated as a special subset of the system address range. This encapsulation becomes the basis for many possible data resiliency choices for the customer to choose from. As such, in this paper, we will also treat IBM i–based XSM Geographic Mirror technology as a hardware HA solution.

The following discussion will show that the long-standing logical replication–based solution offered by Vision Solutions is not only still strategic, but also continues to be the chosen solution for a variety of technical and practical reasons.

Table of Contents

Overview

This paper is for businesses that may be considering a hardware HA solution based on XSM Geographic Mirror, PPRC Metro Mirror, or PPRC Global Mirror for their Power Systems for IBM i environment. The paper covers the following topics:

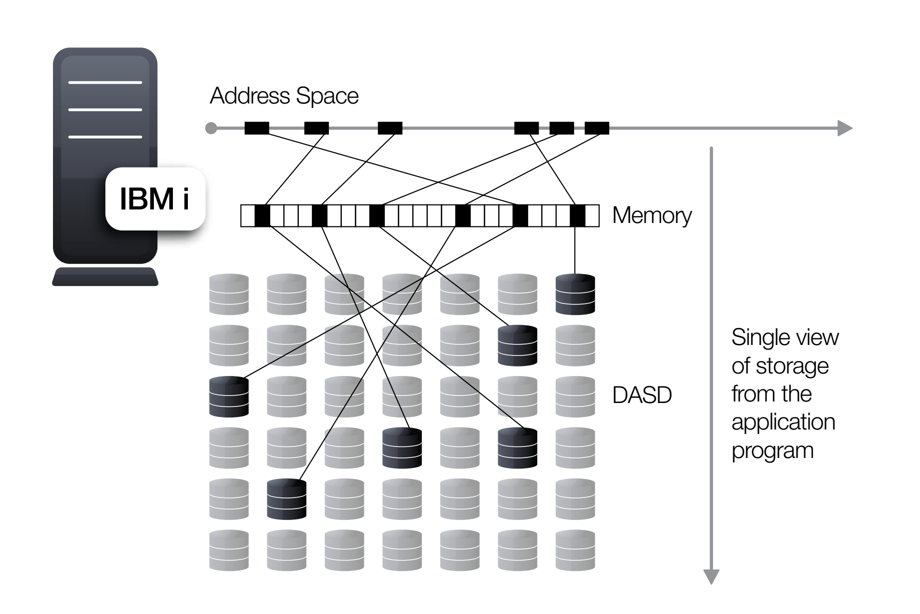
- A short explanation of the technologies used in each solution
- Potential limitations to each solution
- Questions that should be discussed with your HA consultant concerning the solutions and the ways each solution compares to or complements the others

The goal is to help determine the ideal solution based on business requirements and the estimated growth and flexibility of the technologies being considered.

Unique IBM i Platform Technologies

The IBM i architecture includes three unique concepts that must be kept in mind when designing a high availability solution.

Single-Level Store is a two-dimensional view of address spaces that points to “pages,” which can be located in memory or on physical disk. It is not necessary for the application requesting the data to be aware of the actual location of a page. Management of the page is handled by the operating system, a design aspect that allows for optimized performance of the disk subsystems. For more information on Single-Level Store, see http://en.wikipedia.org/wiki/Single-level_store.



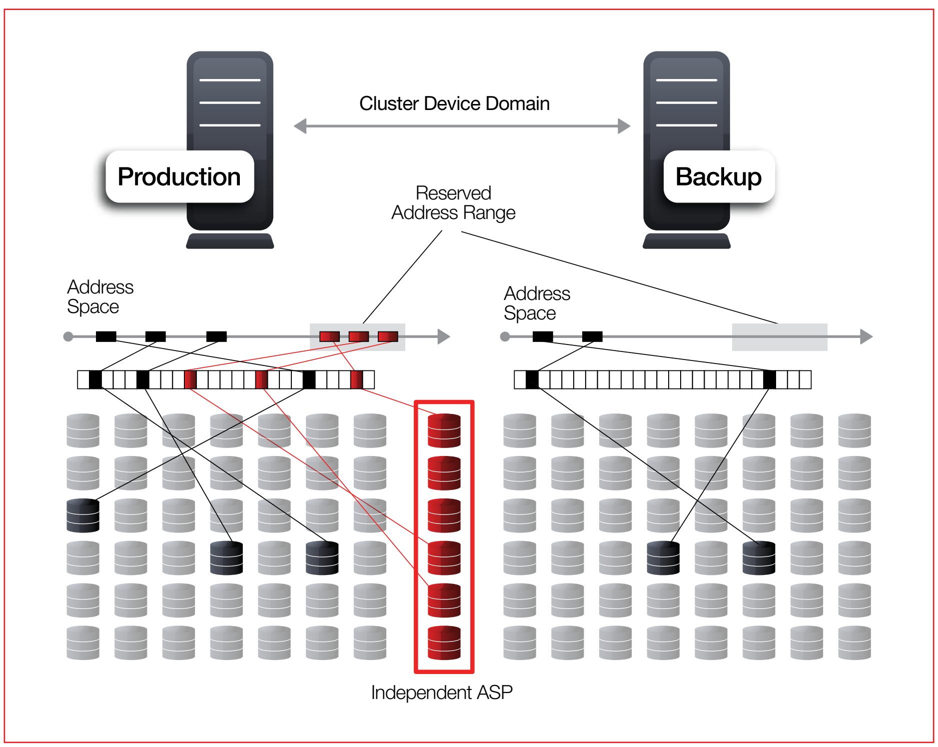
Single-Level Store

Switchable Independent Auxiliary Storage Pool (IASP) technology is the key to hardware HA on the Power Systems for IBM i because it provides a link between Single-Level Store technology and the physical hardware. An IASP is a collection of disk units that can be brought online or taken offline independent of the rest of the storage on a system.

IASPs can be implemented as either switchable or non-switchable. The switchable IASP requires the IBM i clustering framework to control the address range allocation of the IASP. When created, an IASP is assigned a range of system addresses for the entire storage space that is to be included within it. For switchable IASPs, this same address range is reserved on all of the systems that can access that IASP. Even though multiple systems can have access to the switchable IASP, the switchable IASP can only be “varied on” (that is, made active and available) on one system at a time.

When using IASP technology, switchable or non-switchable, you must be careful not to duplicate library/directory names between the system storage pool (*SYSBAS) and the IASP. Duplicate names will cause the IASP vary on to fail.

IASP technology supports many, but not all, object types found on the Power Systems for IBM i. Prior to investing in a hardware HA solution, ensure that all of the object types needed by your current suite of applications can be handled in an IASP.

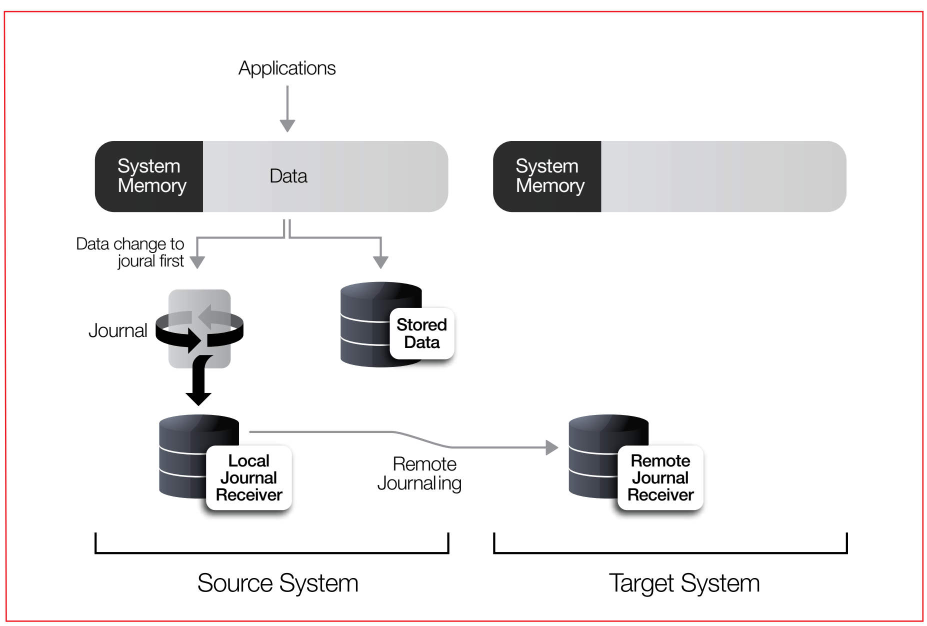


Journaling on the Power Systems for IBM i is an OS-level function similar to logging in other platforms. Data objects can be enrolled into journaling at the time of their creation, and, once journaled, all transactions to that data object are written first to the journal and then to the database. The operating system uses the journal for system recovery and commitment control functions.

The base journal function operates within a single system and is referred to as “local journaling.” Its primary use is in system recovery. In addition, remote journaling can be configured to maintain identical journal receivers on one or more remote systems. Its primary use is to maintain hot backup systems.

Remote journals can be set up to use synchronous communications or asynchronous communications. Synchronous remote journaling guarantees that the remote journal has the latest data changes. Prior to writing the data in the local journal and database on disk, the originating system must first receive an acknowledgement from the remote system that it has successfully received the journal entry. With asynchronous remote journaling, the local journal and database are written without waiting for remote journal confirmation. Synchronous and asynchronous remote journaling have their benefits and limitations, which will be described later in this document.

Both logical replication and hardware HA depend on journaling for data protection. Logical replication journals only the objects selected by the user to maintain a hot backup system. The hardware HA solution uses journaling for recovery of the environment on the backup system.



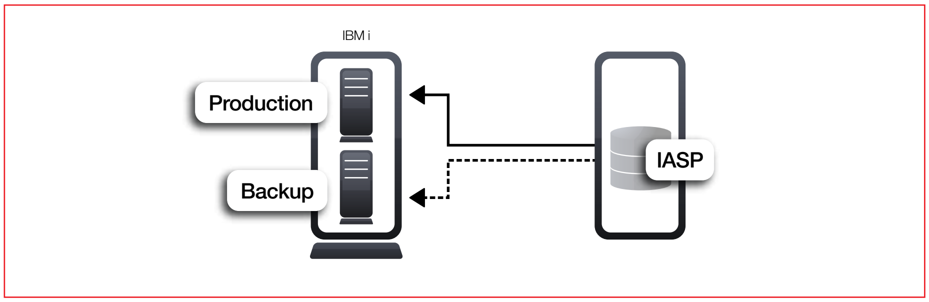
Journaling

Hardware HA Technologies

Hardware HA technologies maintain a single copy of the IASP on a backup system. In addition, three technology options exist for switching a single instance of the IASP (not a copy) from one system to another. They are worth discussing here because of their relevance to some of the topics we will cover later.

All technologies discussed use IBM i Cluster Resource Services to maintain cluster nodes; to monitor for node failure; and to provide start, stop, and switchover control of data resiliency technologies, including the IASP-based technologies discussed here.

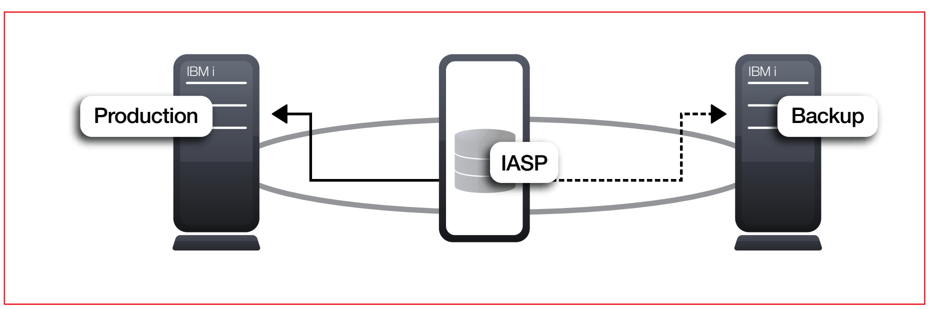
Switchable I/O Pools are internal DASD that can be defined as an IASP and configured to switch between two or more logical partitions in a single-system box.



Switchable I/O Pools Technology

Switchable Tower Technology assigns all of the internal DASD residing in a discrete storage tower to a single IASP. Two systems can connect to the switchable tower using an HSL cable that loops from one system to the tower and then to the other system and back. The tower and the switchable IASP contained in the tower can be switched between the two systems.

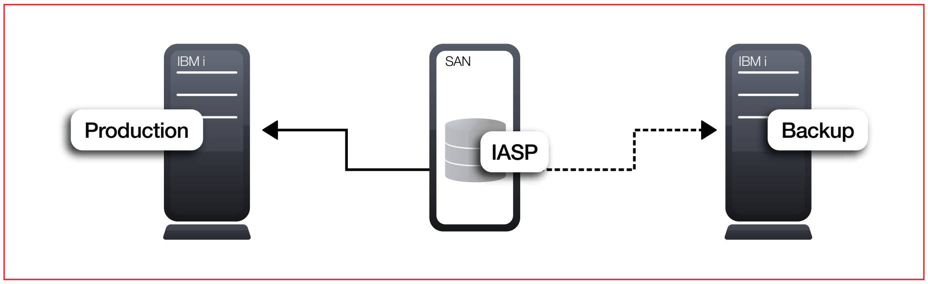
For purposes of forward planning, it is important to note that the current high-speed 12X cable technology that is replacing the HSL technology does not offer a switchable tower function.



Switchable Tower Technology

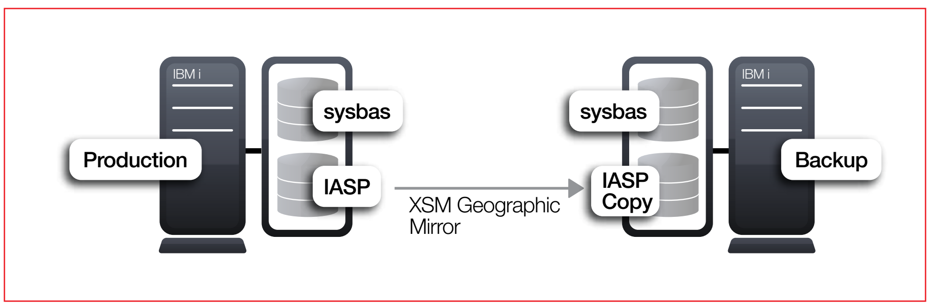
Switchable LUN Technology assigns DASD, allocated in terms of Logical Unit Numbers (LUNs), within a SAN unit to an IASP. Two systems can connect to the same SAN unit, and the LUNs that define the IASP can be switched between the two systems.

The following three technologies maintain a copy of the IASP on a backup system for recovery or for offline use.



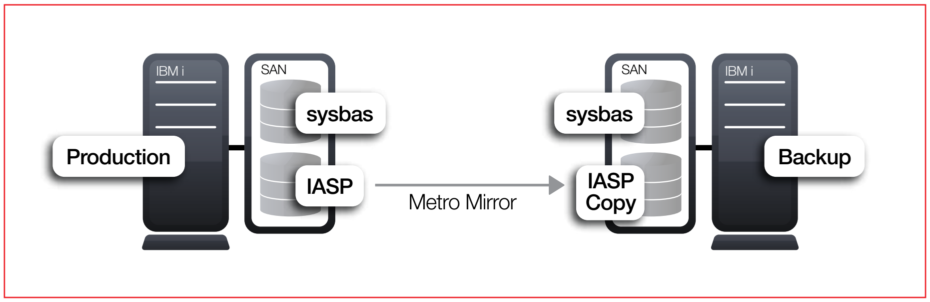
Switchable LUN Technology

XSM Geographic Mirror uses IBM i OS functions to maintain a second copy of an IASP on a second system in the cluster. The IASP can be varied on to either system but never to both systems at the same time. Geographic Mirror is a synchronous copy technology.



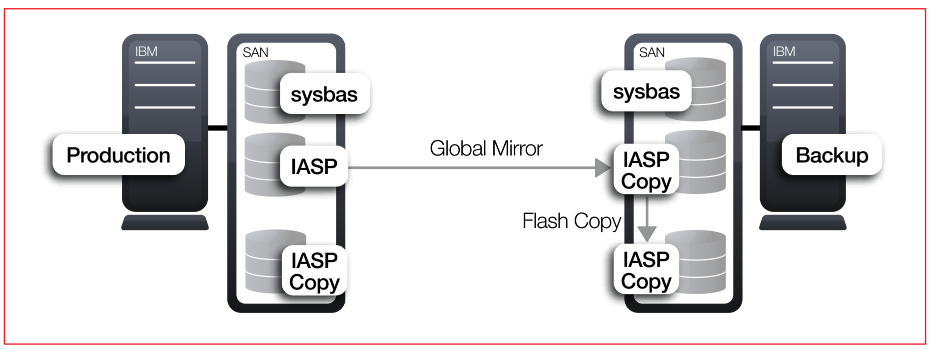
XSM Geographic Mirror

PPRC Metro Mirror uses IBM TotalStorage SAN functions to continuously copy the IASP on a source SAN to a target (backup) SAN. That target SAN is attached to a second system that is included in the cluster. When required, the target copy of the IASP can be reassigned as the source copy and varied on to the backup system. Metro Mirror is a synchronous copy technology.



PPRC Metro Mirror

PPRC Global Mirror is an asynchronous variation of Metro Mirror that is used when much greater distances exist between the source and target SANs, making synchronous replication impractical. Specifically, to overcome unacceptably long communication lag times, PPRC Global Mirror uses asynchronous replication across multiple communication paths and adds point-in-time copy (PTC) technology to maintain an additional, consistent copy of the IASP, with no sector writes out of order.



PPRC Global Mirror

Two Basic Technology Choices

With the uniqueness of IBM i identified and the various IBM i high availability technology options defined, we will now review the two basic methods of maintaining high availability across a system outage.

Hardware HA on Power Systems for IBM i is either a page-replication solution at the operating system level or a disk-sector replication solution that works at the TotalStorage SAN level. In both cases, with the correct version of the OS and the supported IBM TotalStorage hardware and licenses, data that is written to a defined set of drives is duplicated on another set of drives attached to another system. Hardware HA requires the use of switchable IASP technology. With hardware HA, local journaling of all objects in the IASP is strongly recommended to support recovery after a failure.

Logical replication is an application that runs above the operating system, monitoring and reading journals (logs) and then applying the changes defined in those journal entries to a second, target system. The target system may be a virtual system within the same computer frame (LPAR); a second, co-located physical system; or a separate physical system or LPAR on a system in a completely separate datacenter. With asynchronous remote journaling, there is no limitation to how far apart the source and target systems can be.

Each system in a logical replication environment requires its own instance of the IBM i OS, but the target hardware and/or OS instance do not need to be identical to those comprising the source system. This technology will operate with data in an IASP but does not require IASP technology to function. Journaling can be limited to selected objects required to keep the application running on a backup system if required. Remote journaling is strongly recommended for the most efficient transfer of journal entries to the remote system.

Technology Differences and Limitations

No matter the platform, when maintaining a backup system, either in the next room or across the country, there are challenges to overcome. Recovering from an outage in a reasonable amount of time (minimum RTO) with a amount of lost data (minimum RPO), all without damaging existing data, requires attention to details in the following areas:

- Bandwidth between systems
- Latency of synchronous communications
- Journaling protection for system recovery
- ASP vary on and recovery from system outage
- Target-side operations
- Target system availability and HA protection

As each topic is covered, both logical replication and hardware HA technologies will be addressed.

Finally, there are some practical issues to consider, including these:

- Sharing SANs between platforms
- IBM i OS release differences
- SYSBAS replication needs

Bandwidth Between Systems

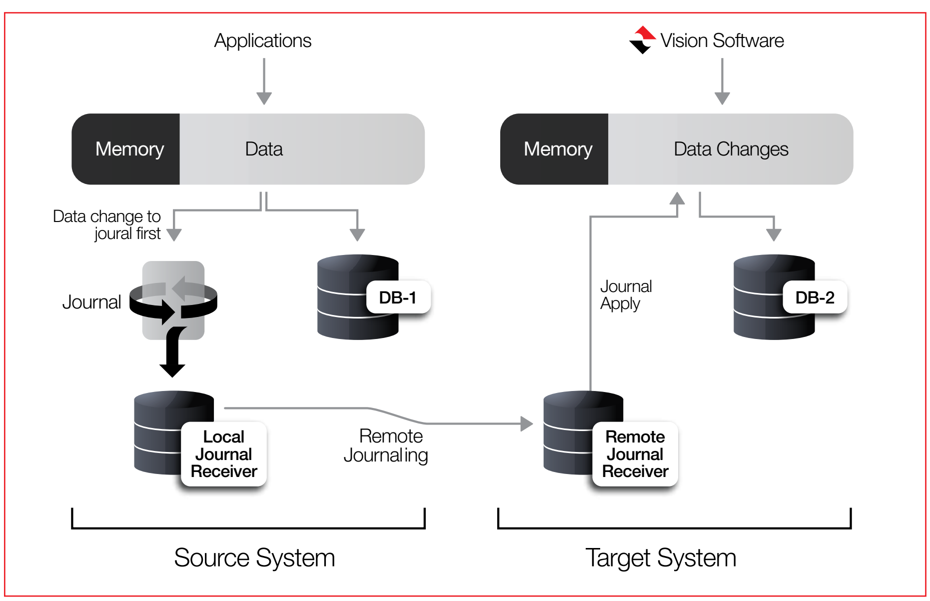
The amount of bandwidth required will depend on the replication technology that is implemented.

Logical Replication and Bandwidth

Logical replication uses remote journaling for transferring journal entries that consist of only the changes made to the objects that have been explicitly defined to the journal. Utilizing both user and system journals, logical replication allows for selective replication of only the objects needed for recovery. This selective replication minimizes both disk space requirements and communication bandwidth consumption. Only the contents of the journal receiver are transferred.

This selective replication is possible because not all data or objects are required to recover from a system failure. For example, many applications utilize temporary files, which are cleared or deleted at the start of a job and then created or used during the processing of the job. Replicating these temporary objects to the remote system is not necessary as they would just be deleted whenever the application is started on the backup system. Proper selection of objects to be journaled eliminates unnecessary replication, helping to optimize use of communication bandwidth and storage.

Remote Journaling includes many other features that reduce bandwidth requirements by minimizing and optimizing the data being transferred, such as bundling journal entries for better performance, streaming large block sizes when journal entry transmission is lagging behind production activity, and using IBM i Data Port Services to efficiently utilize multiple communication lines. Remote journaling is also inherently efficient because it runs at the operating system level and therefore has priority for accessing system resources. Remote journaling utilizes both synchronous and asynchronous communications.



Logical Replication with Remote Journaling

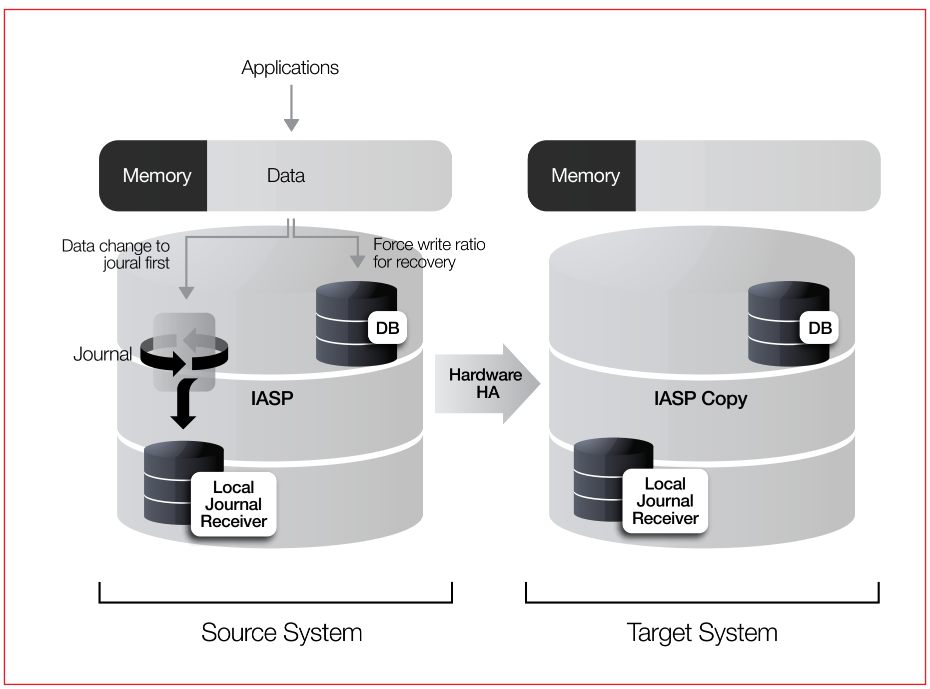
Hardware H A and Bandwidth

XSM Geographic Mirror transfers memory pages to the target system when a page has changed and is being written back to disk. The memory page size is 4KB, but starting with IBM i 6.1, some OS processes are taking advantage of 64KB page sizes. The advantage of memory page–based writes comes when a given page has a significant number of changes and remains pinned (locked) in memory. It is written only once—when the pin is removed or the page is forced to disk. Regardless of how much data actually changes, the entire page (currently 4KB for data) will be sent to the target. Geographic Mirror uses Data Port Services to efficiently use up to four communication lines in parallel when transferring data.

PPRC Metro Mirror and PPRC Global Mirror copy sectors of data that change on the source SAN external storage unit to a corresponding sector on a target SAN unit. For Power Systems for IBM i, each memory page that is written to disk will result in all sectors (eight sectors for a 4KB page) being rewritten. IBM TotalStorage bandwidth between SAN units is granular, allowing for additional fiber-channel cards to extend the number of connections to meet the bandwidth requirements.

In addition to the transfer of user, data to the target IASP, the local journal receivers must also be replicated to the IASP to support recovery in event of a failure (more about this later).

The following diagram shows the relationship between the local journal and the IASP. With hardware HA, everything in the IASP is copied.



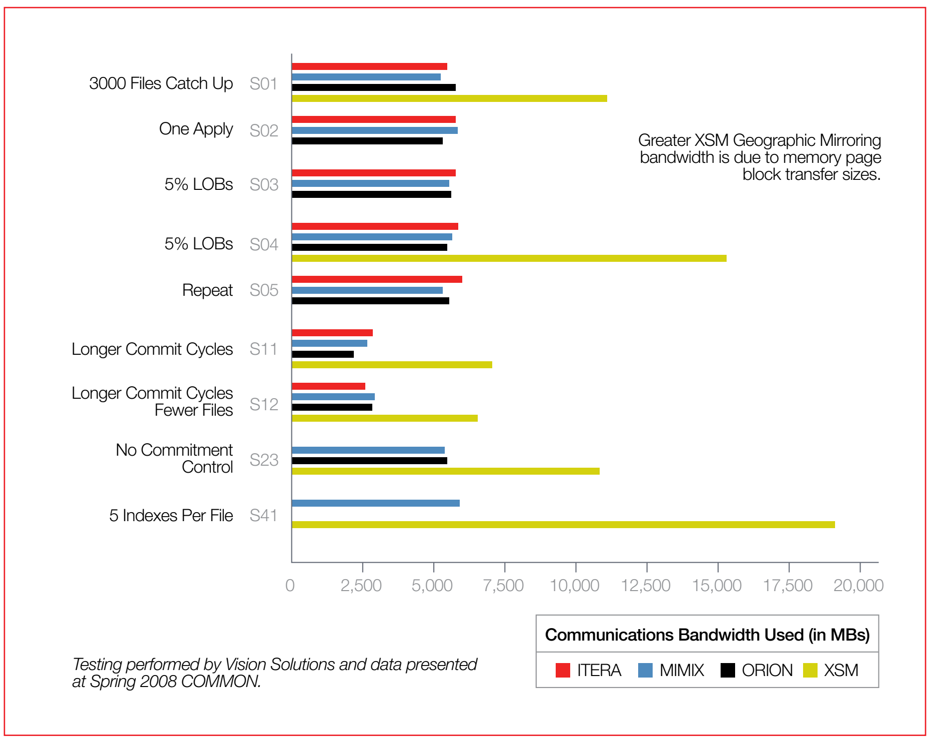
Hardware HA

Bandwidth-Related Decision Factors

When selecting any replication solution, it is important to determine the transaction volumes required, along with the available communications bandwidth. When the target system is located remotely, private communication lines are generally not an option, so the cost of communication bandwidth becomes an even more important factor.

Simple estimates on bandwidth requirements for logical replication can be identified by setting up journaling for all required objects and changing the journal every hour. The average and peak sizes of the journal receivers will be a good indicator of the data bandwidth required for replicating between source and target using logical replication technology.

The bandwidth requirements for remote journaling are generally much less than for hardware HA solutions. The following chart represents tests conducted by Vision Solutions and presented at COMMON 2008. It details the communication bandwidth requirements for various replication scenarios using Vision Solutions' logical replication products and XSM Geographic Mirror.



Communication Bandwidth Requirements

As can be seen from the chart above, the bandwidth requirements for logical replication are less than for hardware HA. The major difference is that logical replication sends over only the journal changes, whereas hardware HA technologies send the updated pages of data as well as the journal changes relating to the data.

It is important to determine the best way to measure the bandwidth requirements for any hardware replication solution you may be considering and to determine if it can be measured prior to the installation.

Latency with Synchronous Communications

Latency is the time required to send data across a communication line to a target system. Normal communications are asynchronous, meaning that once you send the data, you continue normal processing, letting the communication function report back only if there has been a problem with the data transfer.

Synchronous communications are used when it is vital that the data has been successfully and completely replicated to the target system before allowing normal processing to continue on the primary system. When the source and target systems are in the same data center, the impact of synchronous transfers upon performance of normal processing is detectable but may be insignificant, based on requirements of the application. However, starting at less than two kilometers between source and target, an increase in latency can be detected when using synchronous communications. This may or may not appear to be significant, but it will delay application performance.

Asynchronous remote journaling is appropriate for most environments. If there is any reason to consider synchronous remote journaling, a test can be performed by changing the journaling parameter to “synchronous” for a few minutes, hours, or days to determine whether synchronous remote journaling has an effect on the production application performance.

Hardware HA offers both synchronous and asynchronous transfer options. XSM Geographic Mirror and PPRC Metro Mirror are synchronous technologies. They both ensure that the copy of the data in the IASP is in sync with the source data before releasing the changed memory page. Specifically, before continuing processing, both technologies require verification that the changed page is received on either the target system or the SAN unit’s memory (XSM Geographic Mirror terms this “asynchronous geo-mirroring”) or that it has been written to disk on the target (XSM Geographic Mirror terms this “synchronous geo-mirroring”).

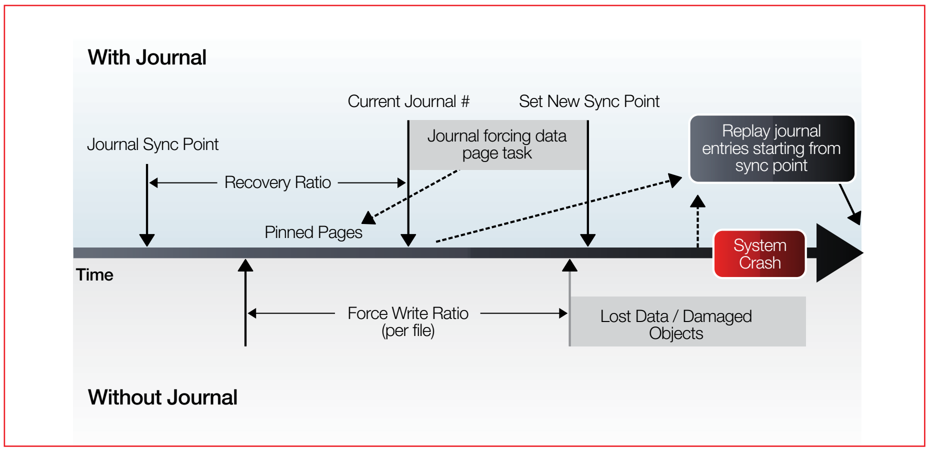
Because of this requirement, there are practical limits on the distance between source and target systems when using XSM Geographic Mirror or PPRC Metro Mirror. The IBM Redbook titled *Implementing PowerHA for IBM i (SG24-7405)*, mentions that PPRC Metro Mirror is recommended for distances of less than 50 kilometers. The same Redbook also recommends testing the environment to ensure acceptable application performance but offers no recommended distance limit for XSM Geographic Mirror.

Prior to using Metro Mirror or Geographic Mirror, you should determine what level of synchronous communications-induced application-processing delays can be tolerated.

Many disaster recovery or business continuity requirements are regulated with regard to the minimum distance between production and disaster recovery sites. Ensure the regulations for your industry are understood and that their defined minimum distance requirements can be achieved by the technology chosen without causing unacceptable delays to application processing.

Journaling Protection for System Recovery

Local journaling is a key function of IBM's Single-Level Store architecture. The main purpose is to maintain a log of all data changes for use in single-system recovery of data in the event of a system outage. Journaling capability ensures that the journal is always the most current representation of the data by pinning the pages that have changed data in memory until the corresponding journal entry has been written to the local journal receiver on disk. The following diagram provides a look at single-system recovery functions both with and without journaling.



Journal Protection

Note that local journaling not only forces the journal entry to disk ahead of the data page in memory, but also periodically sets a new journal sync point after forcing data pages to disk for all journal entries prior to that point. Then, during an abnormal IPL recovery from a system failure, the system replays all of the journal entries from that sync point to the point of failure, recovering all of the lost data up to the last journal entry.

How often journaling will force the data pages to disk is set using the Journal Recovery Ratio parameter. The value chosen is a tradeoff between runtime performance and abnormal IPL recovery time (RTO).

Logical replication takes full advantage of remote journaling to maintain what is often referred to as a "hot backup server." Therefore, the journal recovery operation is not part of the switch to the backup server. The backup server is, by definition, "hot"—that is, completely ready to run without additional journal recovery. However, recovery from journals is still necessary on the failed system prior to rejoining the high availability environment as a backup server.

Hardware HA technologies rely on local journal recovery to prevent damaged objects and lost data when there is a system failure. The recovery action takes place on the target system, after the primary system failure, using a copy of the local journal that is located in the IASP.

Journaling of the objects in the IASP and placing the local journal receiver in the IASP are not configured automatically by hardware HA technologies. But both IBM and Vision Solutions highly recommend implementing this configuration whenever using hardware HA. For more on this topic, see the IBM Redbook *Implementing PowerHA for IBM i (SG24-7405).*

To clarify, both logical replication and hardware HA technologies do rely on journaling for recovery operations. The difference between them is that, in the case of logical replication, the recovery operation is preemptive. The backup copy of the data is maintained continuously so that, at the point of failure, the most recent data is already available on the backup system (after any yet-unapplied journal entries have been processed). In contrast, for hardware HA technologies, the recovery operations start after the failure. The actual recovery steps in this case are discussed in the next topic.

When considering hardware HA solutions, confirm that journaling is essential to an implementation using XSM Geographic Mirror, PPRC Metro Mirror, and PPRC Global Mirror.

IASP Vary On and Recovery from System Outage

The hardware HA technologies of XSM Geographic Mirror, PPRC Metro Mirror, and PPRC Global Mirror operate on data maintained in an IASP with journaling implemented over that data. This IASP, consisting of two or three copies, can be varied on to only one system at a time.

The production-to-target system switchover process will detach the target IASP copy and then vary on the IASP. The vary on operation is essentially an IPL of the portion of the system identified by the IASP and a duplicate object violation check between SYSBAS and IASP. There are 34 sequential steps defined in the vary on process to handle both normal and abnormal IPLs. In the case of a system failure on the source, the vary on of the IASP on the target performs all of the steps that are necessary to recover the data stored in the IASP. In a sense, it is like performing an abnormal IPL of the failed source system on the backup system. The data and journal are both in the IASP; it just needs cleaning up before it can be used.

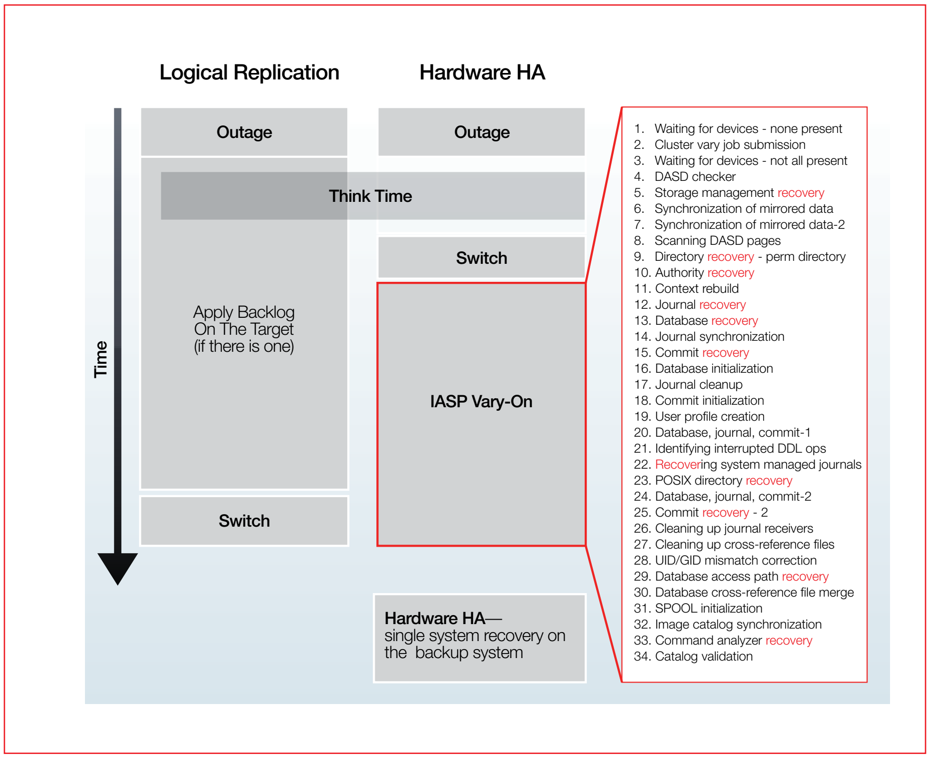
During a planned switch, the IASP vary on can be very predictable, based on the quantity of SYSBAS and IASP objects. The recovery steps are not a factor. However, during a failover, the recovery time cannot be predetermined.

Sequential Steps Performed by the Vary On Operation
(As displayed by the DSPASPSTS command in IBM i 6.1)

1. Waiting for devices - none present
2. Cluster vary job submission
3. Waiting for devices - not all present
4. DASD checker
5. Storage management recovery
6. Synchronization of mirrored data
7. Synchronization of mirrored data - 2
8. Scanning DASD pages
9. Directory recovery - perm directory
10. Authority recovery
11. Context rebuild
12. Journal recovery
13. Database recovery
14. Journal synchronization
15. Commit recovery
16. Database initialization
17. Journal cleanup
18. Commit initialization
19. User profile creation
20. Database, journal, commit - 1
21. Identifying interrupted DDL ops
22. Recovering system managed journals
23. POSIX directory recovery
24. Database, journal, commit - 2
25. Commit recovery - 2
26. Cleaning up journal receivers
27. Cleaning up cross-reference files
28. UID/GID mismatch correction
29. Database access path recovery
30. Database cross-reference file merge
31. Spool initialization
32. Image catalog synchronization
33. Command analyzer recovery
34. Catalog validation

The words “recovery” and “recovering” are highlighted to show the care taken to ensure that the vary on can recover from all possible states that the IASP may be in.

Logical replication does not require an IASP vary on as part of a switchover or failover involving an IASP since the IASP “copy” is actually a different IASP as far as IBM i is concerned. Thus, the target IASP is already varied on and available. The following chart defines the essential recovery steps for logical replication and hardware HA. Each of them has elements that contribute to meeting or missing the RTO of the business.



Switchover Steps Contrasted

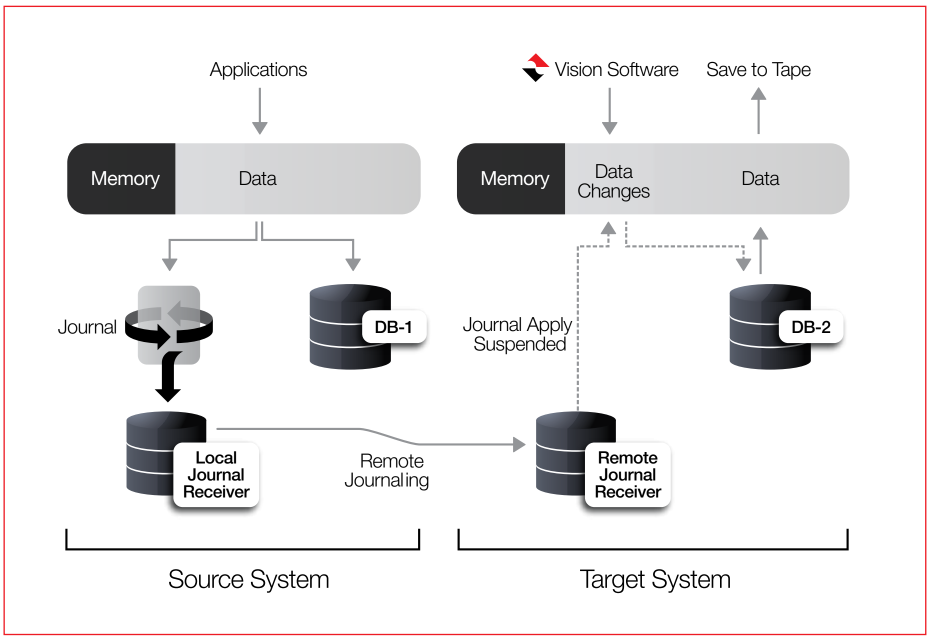
Understanding the time required for switchover or failover is critical to any recovery solution. Logical replication solutions provide a very good indicator for recovery based on apply lag and switch time. The switch time for logical replication should be nearly identical every time you switch. The current apply lag status is known and is often not an issue by the time the decision to switch is made.

Hardware replication does not provide any usable estimator for recovery processes. The switch process is nearly identical each time, but the recovery process during a failover is unknown because it cannot be determined what objects were being written at the point of failure. In essence, RTO for a system outage cannot be guaranteed.

Identify the options to determine your recovery times prior to a failure.

Target-Side Operations

Logical replication allows for real-time read access to the target-side data. When using logical replication with IASPs, the target side IASP is varied on and available at all times. Two common uses for this active target-side data are for queries and offline saves to tape. This has great value for many businesses. For offline tape save operations, the logical replication apply operation will be halted while the remote journal operation continues to send data changes to a buffer on the target system. This buffer of unapplied journal transactions is stored until the tape save operation is completed and the apply operations restarted. In the event of a system failure during this process, the tape save operations can be halted and the apply jobs restarted. Once all the journal entries stored in the buffer have been applied, switching can be performed. The latency (lag time) status is reported during the entire process.



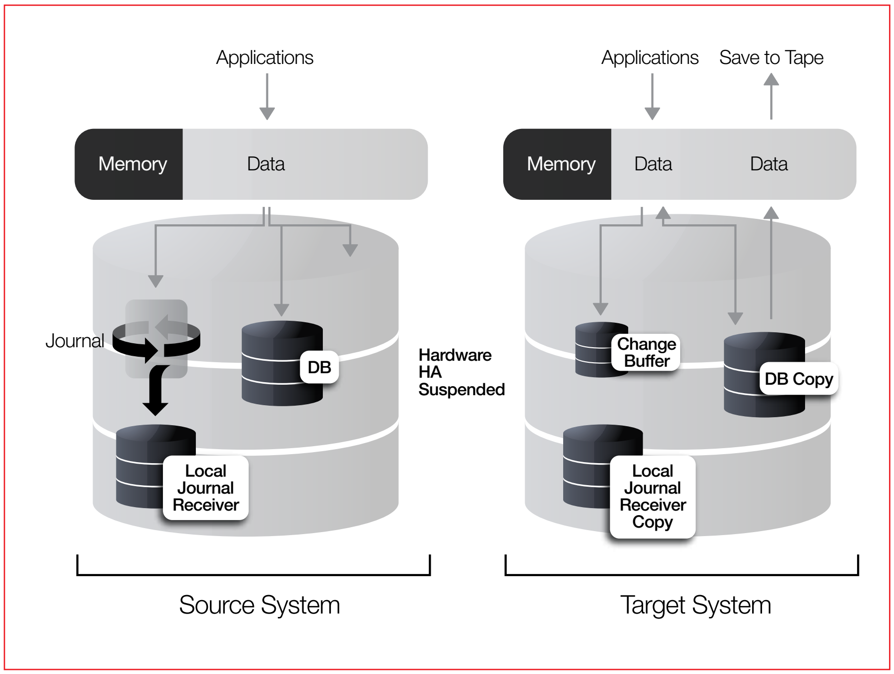
Logical Replication Detached on Target

Hardware HA technologies of XSM Geographic Mirror, PPRC Metro Mirror, and PPRC Global Mirror are based on IASPs, which can be accessed only when varied on and available. When replication is active, the IASP copy on the target system is “attached” and in effect is part of the source IASP environment. When attached, only the source system has access to the IASP.

It is possible to detach the IASP copy on the target, at which time the original and copy become two separate IASPs and can be varied on independently of each other. In this state, changes made to the target-side IASP are tracked, and the IASP can be restored to the original state when reattached to the source system. The changes made to the data on the source side are also buffered while the IASP copy of the target is detached.

During a reattach operation, the target-side changes are backed out and the source-side changes are applied prior to completion of the operation.

Since the buffering of both source data to send to the target and target-side changes to be returned to the previous state consist of memory maps, a large amount of data changes can be buffered. The relevant question is the length of time required after reattaching—that is, before the source and target are in sync and the data is protected against a possible system outage.



Hardware HA Detached on Target

It is important to define the practical limits for a detach operation in terms of how much data can realistically be changed on the target, how much data can realistically be changed on the source, and how much time will be required for the reattach operation to re-sync the IASP copy with the source, at which point protection is in place.

Target System Availability and HA Protection

It is important that the data be protected at the point of failure of the production server. In order to achieve this, the target system or target SAN must be available and operating as the backup system whenever the production environment is operational. The difficulties in achieving this include maintenance operations as well as unplanned system outages on the target system.

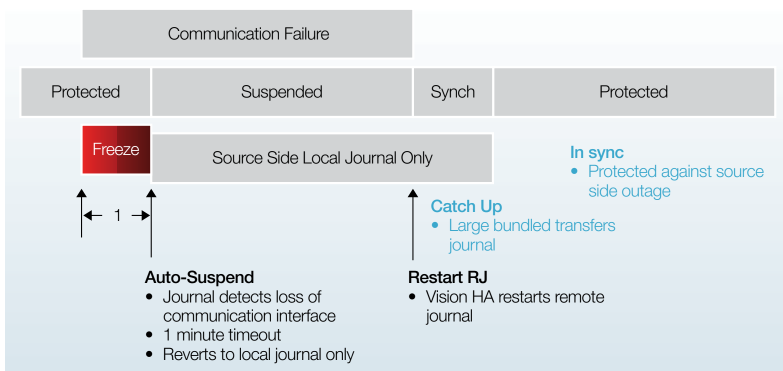
The following diagrams illustrate what happens to the protection of the environment during temporary loss of communication between source and target. Strategies must be considered (including additional backup servers) to provide protection coverage. In essence, the goal is to never rely on a single system not failing.

Unplanned Downtime Behaviors

Although planned outages can present challenges, unplanned downtime can be considerably more problematic.

Logical Replication

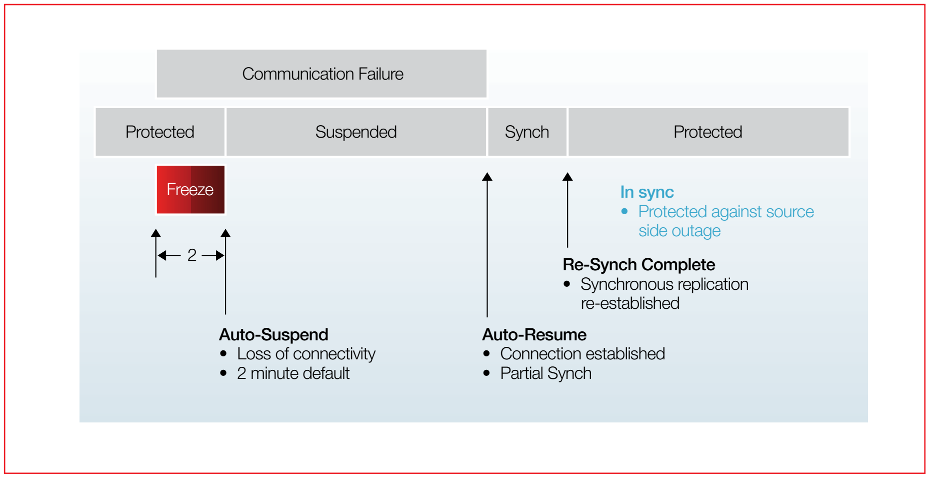
For unplanned loss of the target system or loss of communication to the target system, the journaling environment drops the remote journaling operation while continuing local journaling. This occurs when the journaling function detects that the TCP/IP communication interface has gone inactive for one minute. In a synchronous remote journaling environment, the application will hang until remote journaling is stopped. Vision Solutions' high availability solutions will restart remote journaling automatically when the target system is back up and communication is reestablished.



Logical Replication Communication Failure

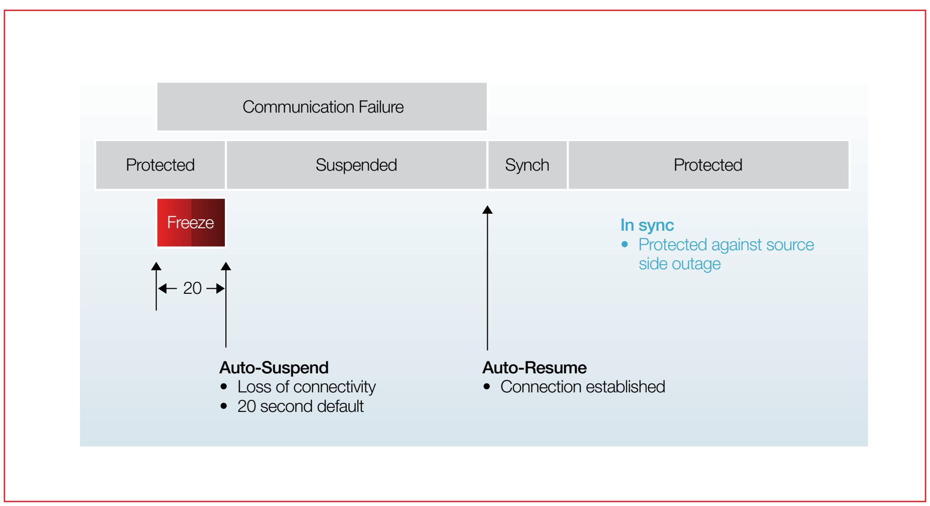
Hardware HA: XSM Geographic Mirror

For unplanned loss of the target system or loss of communication to the target system, XSM Geographic Mirror will attempt to maintain synchronization for two minutes prior to automatically detaching. During this time, the application will hang. After the detachment, data changes are buffered on the source system until the target is restored. When the target system is restored and communication is reestablished, data synchronization will begin for the memory pages tagged as having been written during that period.



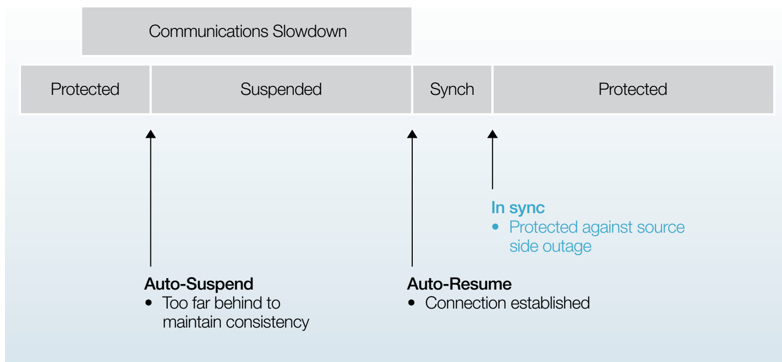
XSM Geographic Mirroring Communication Failure

Hardware HA: PPRC Metro Mirror
For unplanned loss of the target SAN or loss of communication to the target SAN, PPRC Metro Mirror will attempt to maintain synchronization for 20 seconds prior to automatically detaching. During this time, the application will hang. After the detachment, data changes will be buffered on the source system for recovery when the target becomes available.



PPRC Metro Mirror Communication Failure

Hardware HA: PPRC Global Mirror
PPRC Global Mirror is an asynchronous transfer technology. However, in the event of a communication slowdown between source and target SANs, Global Mirror will suspend replication when it reaches a threshold where the consistency of the data cannot be maintained.



PPRC Global Mirror Communication Slowdown

The details of the technology chosen make a difference as to how the high availability environment reacts in various situations. When considering any HA solution, ask the HA consultant about the effect of planned and unplanned outages on high availability protection.

Additional Practical Considerations

Along with the aforementioned points to consider when making a decision about an HA solution, there are a few other things to think about.

Maintaining Data Synchronization
It is sometimes stated by those not in favor of logical replication solutions that maintaining synchronization is a major problem. It is not, as long as the effort is undertaken to properly lock down the target system. While establishing a fully locked-down target can be difficult in some cases, doing so does resolve most problems relating to data going out of sync under logical replication. Indeed, with hardware HA solutions, the IASP is, by definition, locked down already. Any negative impacts to intended business operations incurred by locking down will be the same under either technology.

The question is not about the value of locking down the backup, but whether, when one has the choice under logical replication, the locked-down state of the target system should be compromised by design. Again, no such choice is available under hardware HA. In addition, when the target system cannot be locked down or when other operational considerations make it impossible to do so, Vision Solutions' logical replication solutions provide audit tools to monitor and correct out-of-sync situations.

Sharing SANs Between Platforms
In order to achieve its Single-Level Store architecture, IBM i attaches an 8-byte header to every sector on the disk. This header is used to rebuild the metadata in the unlikely event of a severe crash. SAN solutions under consideration for use in an IBM i environment will provide either a sector format that includes the 8-byte headers or a virtualization layer between the system and the SAN to store the headers somewhere else. IBM i blade servers, for example, utilize a virtualization layer between the blade and the storage subsystem. For larger systems, any virtualization layer should be performance benchmarked.

Sharing of the SAN subsystem between IBM i and other platforms may require more than just assigning LUNs.

IBM i OS Release Differences
IBM i adds major improvements at every release of the operating system. IBM i 6.1 saw a major improvement in a little-known problem relating to the handling of IASPs called “signature violations.” The cornerstone of IASP success is to guarantee that the source and target IASPs are in sync at all times. When an IASP copy is detached, a signature is written on the IASP that must match later. This is explained in the IBM Redbook Availability Management (SG24-6661).

The window of opportunity to run into a signature violation has proven to be great enough and the recovery inconvenient enough that at this time all new switchable IASP environments should be built on IBM i 6.1. The protection logic was redesigned in IBM i 6.1, eliminating the potential problem.

SYSBAS Replication Needs
SYSBAS is defined as all of the auxiliary storage pools (ASPs) that are not independent auxiliary storage pools (IASPs). While user data may reside in either location, environmental data and some object types cannot be located in an IASP. Vision Solutions' HA replicates objects found in SYSBAS, either through the user journal or through the system audit journal.

Hardware HA solutions on IBM i can only replicate data that is in an IASP. IBM does provide for replication of the following SYSBAS objects through the IBM i Administrative Domain technology:

- User profiles
- Classes
- Job descriptions
- ASP device descriptions
- System values
- Network attributes
- Environmental variables
- TCP/IP attributes
- Subsystem descriptions
- Network server descriptions
- NWS configurations
- NWSH device descriptions
- NWS storage spaces
- Tape device descriptions
- Optical device descriptions
- Ethernet line descriptions
- Token Ring line descriptions

There are operational challenges associated with Administrative Domain. The first is that all resources must be registered individually. The second challenge is the handling of new resources. Administrative Domain does not monitor or register new objects automatically. The third challenge is that deleted resources must be individually deleted on all systems, and their associated registration must also be deleted. For example, if you delete a user ID on one system, Administrative Domain will automatically recreate it based on the fact that it is still registered and exists on another system.

Vision Solutions provides utilities that monitor for new and deleted objects and can manage the registration of these changes automatically.

Finally, there are applications that do not support operation within an IASP. There are other applications that, in theory, might work well in an IASP environment but are not tested for this compatibility by their respective vendors, leaving it to the end user to “test via implementation.” Always check with your application vendors about any issues that would interfere with storing related data in IASPs.

Easy. Affordable. Innovative. Vision Solutions.

Vision Solutions, Inc. is the world's leading provider of high availability, disaster recovery and data management solutions for IBM® Power Systems. With a portfolio spanning the industry's most innovative and trusted HA technologies, including EchoStream™, EchoCluster™, iTERA™, MIMIX® and OMS/ODS™, Vision keeps critical business information continuously protected and available. Affordable and easy-to-use, Vision products ensure business continuity, increase productivity, reduce operating costs and satisfy compliance requirements. Vision also offers advanced cluster management and systems management solutions along with support for IBM i, Windows, and AIX operating environments.

As IBM's largest high availability Premier Business Partner (NYSE: IBM), Vision Solutions oversees a global network of partners and professionals to help customers achieve business goals. Vision Solutions is a proud member of IBM's Smart Business and SmarterPlanet initiatives. Privately held by Thoma Bravo, Inc., Vision Solutions is headquartered in Irvine, California with offices worldwide.

For more information, visit www.visionsolutions.com or call 800.957.4511

15300 Barranca Parkway
Irvine, CA 92618
800-957-4511
801-799-0300
visionsolutions.com