precisely

An Introduction to High Availability for Power Systems Running IBM i

Introduction

Every company faces critical hours when system downtime is unwelcome—whether it's planned or unplanned. One company's important hours might only be from 9-to-5, while for another it's 24x7. Increasingly, shops that were able to accommodate some periods of downtime for backups and system maintenance are finding that this window is quickly shrinking, or has disappeared altogether, because of increased demand for access to applications and data around the clock. Because of the need to keep systems available for increasing amounts of time, companies are realizing that a system failure or a site disaster would create enormous disruption and expense, especially if it went on for longer than a few hours. For many companies, exposure to this amount of potential downtime has become unacceptable. Shops that thought they weren't candidates for a high availability (HA) solution are now realizing an urgent need to start looking at their options.

This white paper provides an introduction to High Availability for Power Systems running IBM i. It is for companies that want to understand the technologies involved and evaluate whether such a solution can become a cost-justifiable component of their backup and recovery strategy.

One caveat: Depending on the size and complexity of your information systems, there are many other factors to consider besides the installation of an HA solution when trying to reduce your vulnerability to planned and unplanned downtime. HA clearly is a significant component in an overall data recovery/system availability strategy, but it often takes a variety of software and even hardware components to provide maximum protection against all exposures to downtime.

Before looking at the details of high availability, let's take a quick look at the cost of downtime and some of the primary strategies that are used to mitigate this cost.

The Cost of Downtime

Management is often amazed when they total all the direct and indirect costs of downtime for their company. At first, they may assume that if a system is down for several hours or even a day, it is certainly a big inconvenience but a tolerable risk—as long as this kind of downtime is a rare occurrence. However, once numbers are plugged into the following rule-of-thumb formula, they are often shocked.

Take the average sales lost during an hour of system downtime during business hours, then add the total hourly wage (including benefits) of all employees that are idle during that hour of downtime. Now multiply this figure by the estimated number of hours of system downtime during a year. Finally, multiply the result by two to take into account the costs of this lost employee productivity, lost business reputation, and lost business—both now and in the future—from your lost customers.

Unplanned Downtime vs. Planned Downtime

The IBM Power Systems server running IBM i is considered to be one of the most reliable business systems in the industry. IBM Technical Document 22053139 pins Mean Time Between Failure (MTBF) for this system at greater than 100,000 hours. This statistic only addresses the likelihood of a system failure and does not take into consideration problems that might occur outside the box.

In addition to system failures, other major contributors to downtime include:

- Power failures
- Network failures
- Site damage
- Human errors
- Malfeasance

Another sobering trend to consider is this: Courts are increasingly holding firms liable for losses caused by computer failures. According to Disaster Recovery Journal, litigation is becoming increasingly common and companies are feeling the financial strain of defending themselves against corporate lawsuits.

Despite the potentially dire consequences of unplanned downtime, less than 10% of all downtime can be attributed to unplanned events, and only a fraction of that is due to a site disaster. The other 90+%—the kind that companies face on a regular basis—is caused by system maintenance tasks, including:

> Regardless of the cause of downtime, what matters most is reducing or eliminating the risk of downtime during critical hours of operation.

- Tape backups (nightly, weekly, and monthly saves)
- File reorganization to reclaim disk space and improve performance
- IBM OS upgrades & PTFs
- New application software installations
- Software upgrades & data conversions
- Hardware upgrades
- System migrations

Regardless of the cause of downtime, what matters most is reducing or eliminating the risk of downtime during critical hours of operation.

Disaster Recovery Strategies

Tape Backup - Tape backup devices are still widely used for long term data retention. Tape-based strategies may also be relied upon by some small to medium sized businesses (SMBs) for disaster recovery. Tape DR strategies usually include periodic saves of the entire system, daily incremental tape saves of changed or otherwise critical data, and then storing these tapes safely offsite. Because of the reliability of IBM i servers, most companies think that this is sufficient. However, if a failure occurs that requires reloading entire applications from tape, it is not unusual for the data recovery time to be up to 48 hours or longer, depending on the time it takes to repair or replace hardware, retrieve the tapes from the vault, restore data from tape, and manually recreate all transactions since the last good tape save. And keep in mind that it is not unusual to run into media errors when restoring from tape.

Even though tape backup remains a baseline data protection strategy, many companies have put a pencil to the real cost of this downtime, and as a result have introduced additional layers of protection to reduce data recovery time. Many options exist to reduce the recovery time; some of these include:

Journaling – An IBM i process that efficiently monitors any change made to data and objects. In the event of a system failure, journaling allows data to be recreated without the need to manually re-key it. It is available to users as a basic data change capture capability but becomes a foundational technology when fully utilized in sophisticated high availability solutions. More details about journaling later in this white paper.

Disk protection – Installing disk drives that perform parity protection or disk mirroring to help prevent the chance of data loss in the event of a disk drive failure.

Cloud Recovery Services – Protected third-party, cloud-based recovery sites, available on a subscription basis, where data changes made between tape saves are transmitted (data vaulting). Backup tapes can also be restored on a comparably configured system after the loss or failure of a production system.

High Availability (HA) – True HA on Power Systems running IBM i consists of replicating all changes to data and critical objects on the production system to a designated secondary system. If a failure or system maintenance event occurs, users are moved to this second 'mirrored' system where they can resume business without the loss of data. In general, high availability provides the most efficient way to mitigate most planned and unplanned downtime events.

Disaster Recovery as a Service (DRaaS) – Leveraging replication technology to mirror data and objects to cloud-based recovery systems that are ready to assume the production role in the event of a server failure.

The Components of High Availability

Every HA solution has four primary components:
1. System-to-system communications
2. Data replication processes
3. Monitoring functions
4. Switching / Role swapping capabilities

System-to-System Communications

The first step of the HA process is to establish communications between your production and backup Power Systems servers. Existing LANs and WANs are an easy way for two systems to communicate with each other, especially when moving large amounts of data between them. Setting up TCP/IP communications is fairly simple, but the challenge comes in determining the amount of bandwidth required between the two systems to handle the volume of data to be replicated.

Certainly, you can have the second system in the same location as the first and then directly connect the two. However, you may decide tolocate the second machine in another county or another state, which adds a significant disaster-recovery advantage. Extending this strategy even further, some companies use two systems in a local data center for fast switching and include a third off-site system in another FEMA zone to protect against a widespread geographical outage from a natural disaster like an earthquake or a hurricane. Using a cloud-based DRaaS solution, contracting with a hosting provider or building a dual location infrastructure are three common ways that businesses satisfy the need for separating the production and backup systems.

Data Replication Processes

Once communication is established between systems, the next component needed is an engine that replicates or mirrors transactions between the production and backup systems and does it as close to real time as possible. There are two basic designs for replicating changes to a backup system. Hardware replication writes a duplicate copy of the write operation to a second storage unit as a storage subsystem task. Logical replication captures data changes as they occur in system memory and performs the same operation in the backup system memory. All logical replication-based HA solutions for IBM i systems use the journaling function of IBM i operating system to monitor for changes to data. They either harvest the journal entries from the production system and use their own proprietary process to send these journal entries to the backup system (see Figure 1), or they configure the remote journal feature of journaling to send the journal entries to the backup system. In either case an HA solution based on logical replication applies the journal entry change to the data on the backup system to keep the data in sync across both machines.

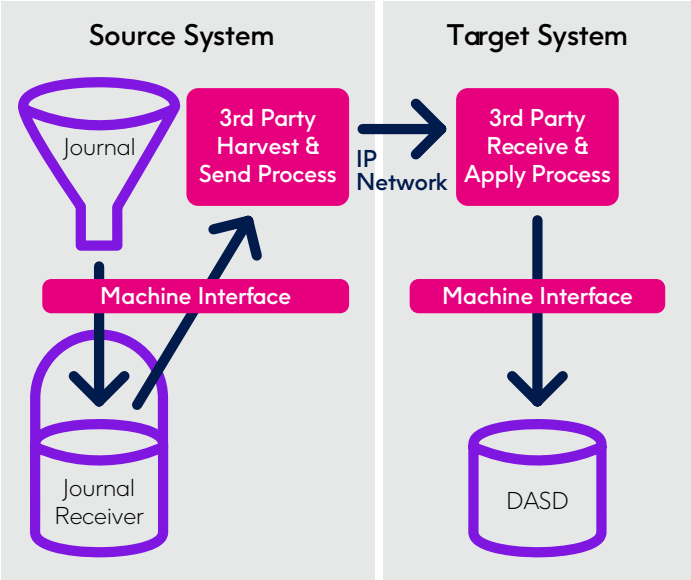
Taking a Closer Look at the Journaling Process

When you enable local journaling for an object, you essentially initiate a process that 'watches' the object. Journaling consists of two parts: the 'journal' and the 'journal receiver.' When any change occurs to th object that the journal is 'watching,' the journal writes everything about this change in a very efficient form in the journal receiver. Each change that is recorded is called a 'journal entry.' When the journal receiver grows to a predetermined number of entries it is available to be saved off-line and a new journal receiver is defined.

When journaling is used with a tape-based backup and recovery strategy, if a system failure occurs between tape saves, the journal receivers that were saved to tape can be restored and the journal entries within each can be retrieved and 'applied' to the data that is also restored from tape.

This reintegrates the data changes recorded in the journal entries wit the data file, which restores the data in that file nearly back to the state at the point of failure. The time required for this method of recovery will be measured in hours or days.

The high availability replication process uses journaling in a different way by sending journal entries to the backup system immediately, where they are applied as quickly as possible to duplicate copies of the objects to keep them current with the production system. This replication of the journal entries is accomplished with some HA products using a Harvest and Send process.



Remote Journaling Replication

The remote journaling function, an extension of local journaling in the IBM i operating system, transmits and writes an identical copy of a journal entry to a duplicate journal receiver on another connected system. It is used in conjunction with the journal-based engine that generates the journal entry.

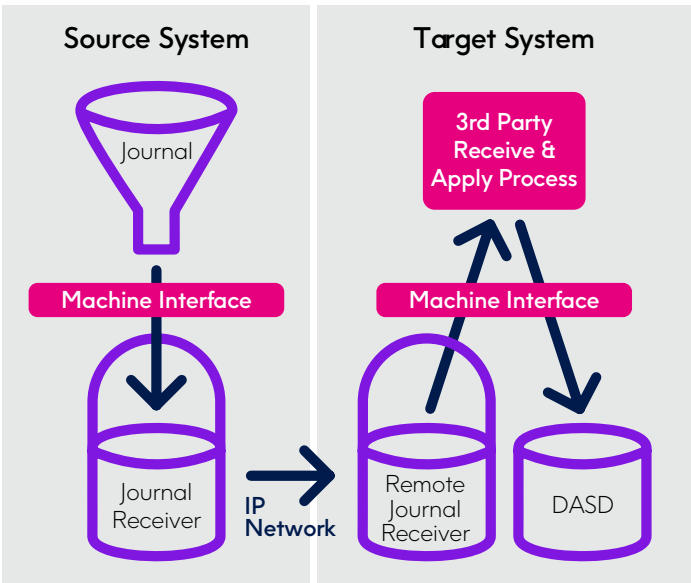


Figure 2: Replication based on Local Journaling

Figure 2 illustrates the remote journaling process, showing that as changes are made to application data, journaling detects these changes on the production (source) system, and as journal entries are made, remote journaling automatically replicates and transmits each journal entry to an identical journal receiver on the backup (target) system.

Once the journal entry is saved in the journal receiver on the backup system, a process within the HA software harvests the journal entry, validates the data, and then applies the changes to the data on that system, thus bringing it current with the production system.

Keep in mind that real-time mirroring of changes to objects by any kind of logical replication-based HA solution can only be done if the object can be journaled. Objects that can be journaled include database, IFS, data areas and data queue objects. However, an HA solution must be able to keep other system-critical objects updated on the backup system, including: program objects, spool files, user profiles, device configurations etc. Typically, these kinds of objects are replicated by monitoring the system audit journal for events and then a third-party product is needed to copy the changes to the backup system.

It is important that any object replication process, whether it uses journaling or not, should be able to continue object replication even if an object is renamed or moved, and it should never stop or slow the ongoing replication of objects if an object needs resynchronization with the production system for some reason.

It is important to note that before an ongoing high availability replication process is fully operational, all objects that could change need to be replicated and reside in total, on the backup system. This may be accomplished by a copy operation prior to activating replication or implement an "as-you-go" copy process. In addition to the data, if you intend to run your applications on your backup system in the event of a system failure, or during

planned maintenance, you will also need a current copy of all of the application's objects on the backup machine. If the business application is from a third-party, this may require a separate application license. Most software vendors grant additional licenses at no extra charge for this specific purpose.

Synchronous vs. Asynchronous Replication

Data can be replicated between systems either synchronously or asynchronously. If data is sent synchronously, then control is not returned to the job on the production system until it is confirmed that the data has been received on the backup system. When using local journaling to harvest and send data to the backup system, the "job" is the process of sending the data across the communication link. When using remote journaling, the "job" is the actual journaling process of saving the data change to both the local storage on the production system and saving the data change to the storage on the target system. This means that synchronous remote journaling will confirm that the journal entry is on the backup system before it is saved to storage on the production system.

With the tremendous computational speeds of systems today, synchronous communication operations will negatively impact system performance over even short distances, due to the speed of data across a wire or fiber which can be limited. Every HA solution has the same challenge. What is more important is that the bandwidth be adequate to get the data off of the production system as soon as possible. Think of bandwidth as the measurement of the size of the pipe, which is different from the speed of the data through the pipe. In practice, if the production and backup systems are further apart than in the same computer room or campus, in all likelihood the chosen method of transfer of the data will be asynchronous, whether using logical replication or hardware replication solutions.

The amount of bandwidth needed for HA depends on the amount of data that is replicated. This differs based on the HA solution implemented. Hardware HA will have a greater bandwidth requirement due to the technology which sends entire disk sectors across the communication link, both for the data object and for the journal entry associated with the change made to that object.

Monitoring Functions

Once data replication is in place between systems, you need a mechanism to continuously monitor for delays, outages and data integrity issues in the communication process.

Having a monitoring process in place that ensures replication integrity is important. Otherwise, in the event of a failure, your ability to reliably use the backup system could be compromised. The reason is that there are plenty of opportunities for problems to arise within the communication process, whether delays, retransmissions of data due to dirty networks or even

communication errors that escape detection by the communication network itself. If any problem arises, it should be apparent on the monitoring screens (or even alert a system operator if the severity warrants) and then automatically attempt to correct the problem.

Self-healing capabilities are crucial to reliability and ease of use for an HA solution for IBM i servers. For instance, an HA monitor should automatically determine if an object on the backup system is out of synchronization with the same object on the production system. If so, the monitor should self-initiate the process of re-synchronizing the object by recopying that object from the production system to the backup system and applying all necessary journal entries to bring it current. And it should do this without halting or slowing the ongoing replication of other objects.

Finally, the ability to manage your entire HA environment in minutes a day from a single point of control should be achievable. Consider the skillsets of your team and whether they require a 5250 interface, browser-based GUI, or both to manage most efficiently. Functionality that automatically resolves most problems as they arise, or that alerts on issues requiring attention, is also necessary.

Switch / Role-swap

All the functions of an HA solution described up to this point exist primarily to minimize planned and unplanned downtime by quickly making a fully synchronized, fully functional backup system available to users. This is the basic functionality required of all high availability solutions.

The process of moving users to a backup system is called a switch or a role swap because the backup system essentially takes on the role of your production system during the time your actual production system is being maintained or repaired.

It is vital that once you have the components of data replication and system monitoring in place, that you regularly test the switch process to verify smooth execution of the process and the integrity of the data on the backup system.

A switch generally includes the following processes:

- Monitoring that all objects are currently synchronized between the two systems
- Ending all user and application jobs on the production system
- Ending the replication and monitoring jobs on the production system
- Designating the backup environment as the production envi-ronment
- Starting the replication and monitoring jobs on the backup system
- Starting user and application jobs on the backup system

Of course, once you have executed the switch and are running your business operations on the backup system, when you are ready to return to the production system you will need to reverse the process.

As part of the implementation of your HA solution, a switch should be performed. This allows you to identify special switch steps for your environment related to communications, system addressing, and the ending and restarting of user jobs, interfaces and HA components. The switch requirements of every system are unique due to different types of objects being replicated, as well as the relationship of jobs, objects, applications and interfaces on the particular system. Documenting these special steps or incorporating them into switch automation is key.

Keep in mind that a good high availability system will be able to keep a variety of objects synchronized in near real time—not just data. Again, you should be able to replicate user profiles, device configurations, spool files, and any other necessary objects. To have a successful switch, all necessary components must not only exist on the backup system, but they must be current.

The switch process of your high availability solution should have sufficient automation built into it so that during a switch, most—if not all— components needed for the backup system to assume the role of the production system are automatically activated. This includes all system addressing, as well as all replication and monitoring jobs. If everything is working properly and the process is fine-tuned, your users shouldn't have to wait long before they see a sign-in screen. As you can see, the switch process is another area where self-healing functions are a critical component of an efficient high availability solution.

> The key elements in any HA solution are replication performance, ease of use, and features that ensure switch confidence.

Evaluating High Availability Solution Vendors

Now that you understand the need for HA as well as a fundamental understanding of the necessary components of an HA solution for IBM i Power Systems, you have an idea of what to look for in a solution.

To recap, the key elements of an HA solution are:

- Efficient replication of objects in as close to real time as possible.
- An easy-to-use user interface that not only makes it simple to see components that aren't functioning properly or objects that are not in synchronization, but also automatically resets components and corrects out-of-sync conditions. This should only require minutes of operator attention each day.
- An easy-to-execute switch process that automates the processes of monitoring synchronization, ending necessary jobs on the production system, and starting all necessary jobs on the back-up system.

In addition to the above, vendors of high availability solutions should be able to provide you with names of customers who can confidently attest that they:

- Perform regular, successful tests of the switch-over process—ideally on a monthly basis.
- Have had their HA solution successfully switch to the backup system for maintenance purposes or after experiencing a failure on their production system.
- Consistently experience a high level of support from their HA vendor. For a solution as critical as high availability, the quality of customer support should be one of the deciding factors when choosing a solution.

In Summary

In today's world of ever-expanding business hours and increasing reliance on data and application availability, it is becoming easier for companies of all sizes to justify the cost of an HA solution.

When evaluating HA, it's important that you understand its critical components, including how the solution fits into your environment, the degree of automation that is built into the system, and what it will require of your IT staff resources to manage the solution.

precisely

About Precisely

Precisely is the global leader in data integrity, providing accuracy and consistency in data for 12,000 customers in more than 100 countries, including 90 percent of the Fortune 100. Precisely's data integration, data quality, location intelligence, and data enrichment products power better business decisions to create better outcomes.

Learn more at precisely.com.