Introduction

Until recently, high availability solutions for IBM Power Systems servers running IBM i were reserved mostly for large enterprises. Now that high availability is dramatically easier to use and less expensive to own and manage, the landscape has changed. Thousands of small and mid-sized companies can now afford the "luxury" of real-time, offsite data protection, as well as rapid and complete data recovery.

Fortunately, this shift is occurring just as downtime is causing more of a disruption and expense to businesses than ever before. With technology costs dropping and downtime costs skyrocketing, companies of all sizes have a huge incentive to evaluate high availability technology.

This white paper will provide a review of the core causes and costs of both planned and unplanned downtime and will then provide a detailed discussion of current options for High Availability and Disaster Recovery solutions. Most importantly, as you read you will learn why true HA and DR protection are now within reach of even the smallest of businesses.

RPO vs. RTO

Before looking more closely at the cost factors of high availability (HA)—and why each has changed so significantly—it is helpful to first understand the concepts of recovery time objectives (RTO) and recovery-point objectives (RPO).

The graph in Figure 1 shows a variety of common IBM i business continuity technologies in which one axis indicates the time it takes to recover data after a failure/disaster (RTO), and the other axis indicates the completeness of data that is ultimately recovered (RPO).
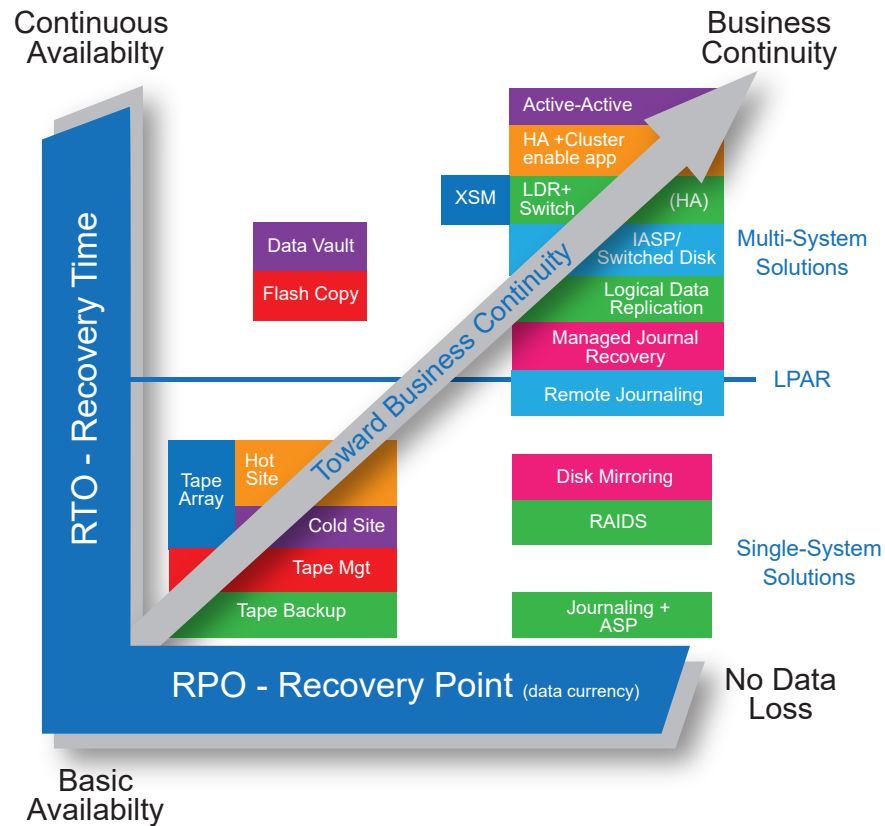

Figure 1: RPO and RTO and the spectrum of IBM i DR Solutions/strategies.

At the low end of the disaster recovery (DR) spectrum is tape backup (basic availability) and at the high end is high availability (HA)—a process more technically known as logical data replication-plus-switchover (LDR+Switch), which rapidly moves users and processes to a fully mirrored secondary server in order for it to assume all or most of the functions of the production server.

Unfortunately, the perception of many mid-size and small companies is that HA technology is so much more expensive than basic disaster recovery protection that it is considered "out of reach" in terms of both cost and complexity. But, in line with most other computing technologies, the range of options between the most basic DR protection and the high-end, fault tolerant, enterprise-scale solutions has increased and, overall, the cost of all the options has come down, radically in some cases.

Companies have a huge incentive to evaluate IBM i high availability technology.

Computer system downtime costs American businesses one million dollars per hour.

Factors of High Availability Costs

High availability is certainly not "cheap" when you consider all of the components that are needed. What has changed is how the cost of each of these factors—each for its own reasons—has dropped. Here are the major components that contribute to the cost of an HA solution:

Hardware—A second IBM Power Systems server is needed, with enough capacity to accommodate the storage of replicated data and potential production demands. For instance, depending on how fully you want to run your applications from your backup environment during planned and unplanned downtime, this server may need to handle the same scale of transaction volumes and devices supported by the production machine. If less than full capability is acceptable during downtime, adjustments can be made. But in the end, a second server, ready to run, is a must.

Communication Bandwidth—If the second Power Systems server is located off site, which is what is necessary to have true disaster recovery protection, then sufficient communication capacity (bandwidth) is needed to accommodate the amount of data flowing to it from the production machine. This includes the I/O processing capacity of the backup server and the communication lines between sites.

High Availability Software—This component executes, manages, and monitors the replication or mirroring of designated business-critical data to the backup server. It also provides the ability to efficiently move users and processes to the backup server during downtime events. In addition to the initial purchase cost for this software, annual maintenance contracts and installation and training costs must be considered.

High Availability Management—As with any other infrastructure software or system, some level of staff time is required each day to monitor and manage the data replication processes to ensure that the mirrored data is accurate and usable when needed. In part, the amount of time needed for this task depends on the scale of your environment. But the self-managing capabilities of the HA software can have an even bigger impact. Even large scale HA environments can be easy to manage, with the right software. So what has changed? Why should you re-consider whether you can justify investment in a true high availability solution? Here are five reasons:

Reason #1 – Decreasing Cost of Hardware

It's no secret Power Systems machines pack a lot of bang for the buck, and the current configurations and pricing models of the IBM Power Systems running IBM i make buying a second machine for high availability significantly less costly than it was, even just a few years ago. If you intend to replicate data only for disaster recovery purposes (not to run applications during downtime on the backup machine but only to be able to retrieve the data from it), an economical option is to buy a smaller model with enough power to handle replication. This at least keeps your "backup" current and on-disk, not just on tape.

Another alternative is to obtain a smaller machine with enough capacity to run critical applications for short amount of time, or for a limited number of high priority users. IBM offers several models of IBM Power Systems Capacity BackUp editions (CBU) that allow you to run them just as a replication target under normal conditions, but gear up to engage extra processing power when it is needed (i.e. after a switchover).

These servers come with a minimum amount of processing power to keep the costs associated with interactive users down. When needed, you can switch users to the CBU server and use its entire processing power, at no additional cost, until you can switch back to your primary production system.

If you are considering replacing your current system with a newer model, you might consider extending the value of your existing older server by using it as the backup server in a high availability environment, instead of selling it for pennies on the dollar. This can be a good "Bridge" solution, until a second HA or CBU edition server can be fit into your budget. But, on the downside of this option is the unavoidable fact that the server is older. If you have outgrown it for production purposes, it can't be expected to keep up with your growing HA needs for the long term. And, at some point, IBM will no longer support the highest IBM i release level your old server can handle, at which time you must upgrade.

Finally, there are also many options available today to obtain HA as a service from a Disaster Recovery company, value-added reseller (VAR), or even the HA software vendor itself. Some of these companies sell the entire HA package as a service, on a cost-per-month basis, effectively transforming your HA investment from a capital expenditure to a monthly service see. This may make both logistical and financial sense, if maintaining a second server at a second site is not possible or if service expenses make more sense for your finances than a capital purchase.

Reason #2 – Decreasing Cost of Communication Bandwidth

Many companies wisely choose to locate their backup server at an alternate site, from a few miles to hundreds of miles away from the production data center. This helps to assure continuous availability in the unlikely event of a site-loss disaster. Of course, the number and size of transactions determines the size of the communications pipeline required, in order to ensure that replication to the backup server does not fall behind production volume.

The good news here is that the cost of communications bandwidth has greatly decreased in recent years, which is allowing more companies to move up from high availability to full disaster recovery protection.

Reason #3 – Decreasing Cost of High Availability Software

The cost of HA software has changed in a number of ways. In general, as the state of the art has moved forward, vendors have begun to sell HA software products that offer all the core features needed by small and medium sized businesses, at prices significantly below those for software with the extended, high-end features that only the largest, enterprise-scale companies need. In other words, HA software is no longer "One-size-fits-all" in design or price. You have options. And combined with the lower costs for server hardware, the total investment required has really dropped to a fraction of what you would have expected even just two or three years ago.

HA Core Competencies

For the same reasons, though, you need to consider openly and realistically what level of HA and DR capability your company needs and can actually afford. You might be pleasantly surprised when you do the math. When you consider how much downtime is costing you already, the return on your investment, due to savings in planned downtime costs, may put a more capable, complete HA product within your reach. In other words, you have choices. But don't just look at price or just at features. Consider the overall value and ROI, even as you ensure you are getting the necessary HA functionality and support you truly need.

Also, it is particularly valuable to talk to customers of the different HA vendors and, if possible, to talk to a company who has used more than one HA product. There are still distinct differences between HA products and companies. For example, there are big differences between the size and expertise of the software development labs at the different HA vendors. Be sure the vendor you choose has a track record of leading advances in HA development, so that the investment you make today will continue to deliver ROI well into the future.

Reason #4 – Decreasing Cost of HA System Management

With many software applications, it often costs more to manage the software than to purchase it. The same can apply to high availability software, but just as power and capability are increasing, with some HA products, the management time required is actually decreasing, due to smart and powerful automation of the most time consuming tasks. High Availability products that require an operator to pore through reports and manually find and repair objects that are out of synchronization can tie up staff for 20 or more hours per week.

Unexpected glitches arise during almost every routine or periodic maintenance procedure.

New-generation products with the latest autonomic (self-healing, self-managing, self-configuring) technologies can reduce the amount of labor needed to monitor/ manage the product to half an hour or less per day. Of course, the more time it takes, the higher your total cost of ownership.

The Autonomics Difference

Here are typical ways that next-generation self-healing, self-managing HA products save time:

- Object auditing processes are performed automatically by the product and the results of the audits are shown on easy-to-read screens. It is critical that operators can quickly see the status of replicated object integrity. If a problem requires operator attention, the source of the problem must be quickly determined. An operator should never spend time searching for the source of problems within the various journaling and communications screens of the IBM i operating system, nor have to write special queries or programs to verify HA replication integrity.

- If an object needs resynchronization for any reason, the problem is automatically detected and corrected by the software. This self-healing process is even more effective if it is able to detect and repair discrepancies at the record level instead of re-copying entire objects when synchronization problems are detected.

- Objects don't lose synchronization when they are renamed or moved to a different library.

- Objects don't have to be manually synchronized when they are first created.

Self-Monitoring Capabilities

Computer technicians generally have skill sets that allow them to analyze business problems, design solutions to solve those problems, and maintain the systems that support their solutions. The more time they have available to devote to these tasks, the more effective the IT organization will be. When HA is smart enough to manage itself, in the area of replication synchronization and automated failover switching for example, then fewer technician labor hours are needed to support these tasks. HA solutions that are self-configuring, self-monitoring, and self-healing save labor time, allowing your technicians to focus on the strategic objectives of the business.

Vendor Implementation Support

Implementing a HA solution often requires specialized IT skills that may not be available within your organization, simply because they are not needed on a daily basis. Vision Solutions offers implementation and support services that reduce the costs and time associated with developing and implementing your HA solution, and make it easier for technicians to gain the basic knowledge needed to keep the solution running properly. Very often, an HA solution can be up and running in only days, and both onsite or remote installation services are available.

Ease of Use

Costs have also been decreased though advancements that have made availability systems more intuitive with browser-based UIs and content sensitive help. Management can take advantage of at-a-glance capabilities with a single monitor screen for multiple installations. And email alerts allow you to receive system status information while you're away from the console.

Reason #5 – The Rising Cost of Downtime

Now that you have a better understanding of the how the costs of high availability solutions have changed, let's take a look at the costs and causes of downtime to see how quickly your investment in HA can be recouped.

Time windows when system access can be restricted in order to perform maintenance tasks are shrinking. For many IT shops, the luxury of scheduled downtime has disappeared altogether. This can be primarily attributed to three factors that keep stretching the length of the business day:

1 Economic conditions dictate that companies can't afford to buy additional systems, so new, expanding production and reporting workloads are moved to "off-hours" to maximize the utilization of existing systems. But as a result, the time available for non-production maintenance work, such as software updates and tape backups, is reduced.

2 As businesses grow, they move from operating regionally to nationally, or from nationally to internationally. Or operations move from one eight-hour shift to two. Or business partnerships and supply chain operations require many companies keep their systems available.

3 To offer Internet-based retail websites, companies can instantly require 24 x 7 systems availability, just to be able to stay in business.

The Falling Cost of HA Ownership

How it adds up…

Given all of the ways that the cost of high availability has decreased, the total cost of ownership is significantly less than it was just a few years ago:

• A software license for full-featured HA typically costs 30% to 70% less than it did a few years ago.

• Reduced software license cost means annual support is also significantly less. Over five years, this can save you tens of thousands of dollars.

• The labor cost to monitor and manage HA is now typically only half an hour or less, instead of a half-time to full-time person.

• In general, a second Power Systems server for HA purposes costs many thousands less than it did just a few years ago.

• The cost of communication bandwidth has plummeted. In 2000, the cost of bandwidth was many times higher than it is today.

• Some service companies offer a turnkey HA package. This shifts HA from a capital expenditure to a monthly expense.

In the past, discussions about downtime used to be about planning for site disasters or system failures. In reality, the largest durations of downtime are attributed to system maintenance tasks. In fact, only five to ten percent of downtime is caused by unplanned events and only ten percent of that (about one percent of the total) is due to site disaters.

The other 90+ percent comes from the following:

• Data backups (nightly, weekly, and monthly saves)

• Reorganization of files to reclaim disk space and improve performance

• Vendor software upgrades and data conversions

• IBM OS release upgrades and PTFs

• New application software installations

• Hardware upgrades

• System migrations

Every hour that a system is unavailable—whether from planned or unplanned events—causes significant costs to be incurred to a business—often far more than you think. Plug your numbers into the following back-of-the-envelope formula to get a general idea of the total annual direct and indirect cost of downtime:

1 Take the value of the business lost during an hour of system downtime (whether from planned or unplanned downtime), then add the total hourly wage (including all benefits) of all employees that are idle during that hour of downtime.

2 Now multiply this figure by the estimated number of hours of planned system downtime during a year.

3 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.

Despite the fact that the largest cause of downtime is from planned events, and even though the IBM Power Systems server is considered one of the most reliable systems available (some studies have put its reliability at 99.95%), it is vital to put unplanned events into the equation. Simply stated, unplanned events that impede access to business-critical systems for an extended period can cost your business dearly and can even spell doom for a business.

Will a competitor take better care of your customer than you did?

Conclusion

It is a fortunate change of events for smaller companies that the powerful business continuity technology of high availability is no longer reserved for the largest enterprises. Because of dropping hardware and communications costs, the decreased cost of HA software, and a host of self-healing and self managing capabilities making HA simple to run, plus new options to purchase HA as a cost-per-month service, thousands of companies that used to precariously rely on tape backups as their sole disaster recovery strategy can now easily acquire more robust, full featured HA and DR protection, for far less than ever before.

IBM Power Systems High Availability Core Competencies

Important questions to ask vendors and their customers when evaluating HA software:

☐ On average, how long does it take each day to monitor and manage the system? It should rarely take more than half an hour a day. Be sure to verify this from several customer references.

☐ How well does the HA solution audit objects to ensure both environments are truly in sync? Audits for data should include a continuous, record-by-record comparison between the production and backup environments. The auditing process should include all necessary objects including IFS.

☐ Is the HA solution self-correcting and self-healing (does it have autonomic capabilities) when object discrepancies are found? If so, can objects be automatically corrected at the record level?

☐ Are the replication of large objects, spooled files, IFS, device configurations, data areas, scheduled jobs, data queues, job queues, and the jobs within job queues supported? The replication of IFS should occur incrementally at the record level where possible.

☐ Does the product provide automated object replication management (i.e. automatic object creation, deletion, moving, renaming)?

☐ How does the solution utilize IBM i Journaling? Are the journaling processes optimized for maximum efficiency?

☐ Is the switchover process (moving users and processes to the backup server during downtime on the production server) easy to execute as well as fast and reliable?

Customer references should be provided by an HA vendor in which the switchover process is successfully tested at least once per quarter (once per month is better).

Customer references should verify that a complete switchover is performed; in other words, the customer runs their mirrored application(s) on the backup environment for an extended time.

The HA software should execute the actual switchover of the systems within minutes (of course, additional time will likely be needed to physically move users, communications, and interfaces).

One or more customers should be able to attest to a successful "failover" (system failure on the production environment causing an automated switchover). Analysis tools should be included in the software that give operators confidence that all necessary replicated components are current.

☐ Can customer references confidently endorse the quality of technical support and services that they receive? How many support calls are made on a weekly and monthly basis?

Easy. Affordable. Innovative.
Vision Solutions.

Vision Solutions is a leading provider of business resilience solutions – high availability, disaster recovery, migration, data replication and security – for IBM Power Systems. For more than 25 years, customers and partners have trusted Vision to protect and modernize their environments, whether on-premises or in the cloud.

Visit visionsolutions.com and follow us on social media, including Twitter, Facebook and LinkedIn.

Find us on:

Facebook: http://www.facebook.com/VisionSolutionsInc
Twitter: http://twitter.com/VSI_Power
YouTube: http://www.youtube.com/VisionSolutionsInc
Vision Solutions Blog: http://www.visionsolutions.com/blog

15300 Barranca Parkway
Irvine, CA 92618
1.949.253.6500
1.800.683.4667

visionsolutions.com

visionsolutions.com