
The Benefits of Continuous Data Protection (CDP) for IBM i and AIX Environments

New flexible technologies enable quick and easy recovery of data to any point in time.

Introduction

Downtime and data loss pose intolerable risks to every business today. From IT departments to the Board Room, managers have seen the importance of business uptime and data protection to continued success, productivity and profitability. This white paper will provide a road map to the most effective strategies and technologies to protect data and provide fast recovery should data be lost or corrupted due to accident or malicious action.

Planning for recovery—designing and implementing a solution to reduce the amount of recovery time needed after an interruption—is a pressing requirement for businesses of all sizes. In implementing an operational plan that ensures that both data and applications can be recovered quickly, IT managers are generally confronted with several challenges:

- How can I ensure my applications and data are continuously recoverable without impacting business operations?*
- Do I have data protection strategies available to me that meet my recovery point and recovery time objectives?*
- Can I afford to implement a comprehensive plan that covers both local and remote (disaster) recovery requirements?*
- Are there cost-effective alternatives that meet my requirements?*

Bottom Line: *Businesses face a variety of risks to their data such as accidentally deleted files, data corruption from viruses or hacker attacks, software/hardware failures, power outages or any of a wide range of natural disasters. Business and IT managers need a data protection and recovery strategy that keeps the organization's business doing business. For IBM i and AIX IT departments, this is a high priority.*

Tape Backups: First Line of Defense

If you're like most businesses, you're using some form of data protection today—probably tape-based backup. Periodically, someone shuts applications down to perform a backup to tape. Depending on the volume of data that is being copied, this may take several hours and requires manual intervention to set up the backup job, run it, confirm that it occurred, and then return the application to operation.

The backup copy may be kept locally in case data needs to be recovered in the near term, and eventually it may be moved to an offsite location for archival storage purposes. The reason to make and keep copies of your data is so that, in the event of some sort of event or catastrophe that deletes or destroys data, you have a clean copy safely tucked away to use for recovery purposes.

The two most important metrics for determining the optimal capabilities of any data protection strategy are the recovery point objective (RPO) and recovery time objective (RTO).

Recovery Time Objective (RTO). RTO defines how quickly you need to restore data and applications and have them fully functional again. The faster your RTO requirement, the closer you move to zero interruption in uptime and the highest level of data protection.

Recovery Point Objective (RPO). RPO defines the point at which the business absolutely cannot afford to lose data. It points to a place in each data stream where information must be available to put the data back in operation. Again, the closer you come to zero data loss and real-time access, the more continuous protection of data will be required.

Tape-only backups are no longer a feasible data protection strategy in today's business environment.

Tape is used for backup and archive because it is very inexpensive, but it is an old technology that has been available almost since the dawn of computing. There are several issues with tape-based backup:

- Tape-based backup is a time-intensive process that is potentially disruptive to your applications; this issue is commonly referred to as the backup window problem.
- Because of its impact on applications and resources, tape-based backups are usually not performed more than once a day, and often only once every several days, meaning that there are very few tape-based recovery points available for use over the course of a week.
- Because your data is changing very frequently (on the order of seconds or minutes), fewer recovery points mean you are risking the loss of large amounts of current data for a given recovery.
- Once it is clear that a recovery needs to occur, it takes time to perform recovery tasks including locating the correct tape, transporting it (if it's offsite), restoring it to disk and restarting the application with the recovered data.
- As storage media for backup data, tape is not entirely reliable; in fact, leading analyst groups such as the Gartner Group, the Enterprise Strategy Group and the Taneja Group state that as many as 1 in 4 backup tapes suffer from some sort of problem that precludes performing a recovery.

Transporting tapes to offsite facilities for archival purposes also has inherent risks. Recently publicized tape losses during physical transport (by truck) have hit large companies like Bank of America, Citigroup Inc., ChoicePoint Inc. and LexisNexis in the U.S. and resulted in the loss of hundreds of thousands of company records.

Replication of data across secure IP-based networks is a much faster, easier and safer way to transport data to offsite locations for archival storage purposes. If you are driven by either business or regulatory requirements to deploy a disaster recovery solution, a pure tape-based strategy can subject you to undue risk.

As all of these issues illustrate, tape-only backups are no longer a feasible data protection strategy in business environments that require frequent access and updates to critical business data.

Pinpoint the Most Valuable Data for Your Business

Most restore requests are driven by issues such as an inadvertently deleted file or data corruption that is introduced by a virus or a hacker.

It has been proven over time that most data recovery requests are for relatively recent data, and that there is a direct correlation between the age of data and the possibility that it would be required for restore purposes. Most restore requests are driven by issues such as an inadvertently deleted file or data corruption that is introduced by a virus or a hacker. Typically these problems are discovered within several hours or at most a few days from when they first occur, resulting in restore requests for more recent data.

In general, the only time you may need to restore data that has already been archived would be in the event of a disaster that physically destroys computer equipment and facilities, such as an earthquake or a tornado. While it pays to be prepared against these occurrences, they are very rare. The slope of the orange line in Figure 1 varies by company type, but it reflects the general relationship in all industries between the age of data and the chance that it would need to be restored.

Another key factor to note is that as data ages, it becomes less important to support the ability to restore to any point in time. Note the inflection point in the orange line in Figure 1 that occurs around Day 3. Restore requests for data drop off significantly after that point.

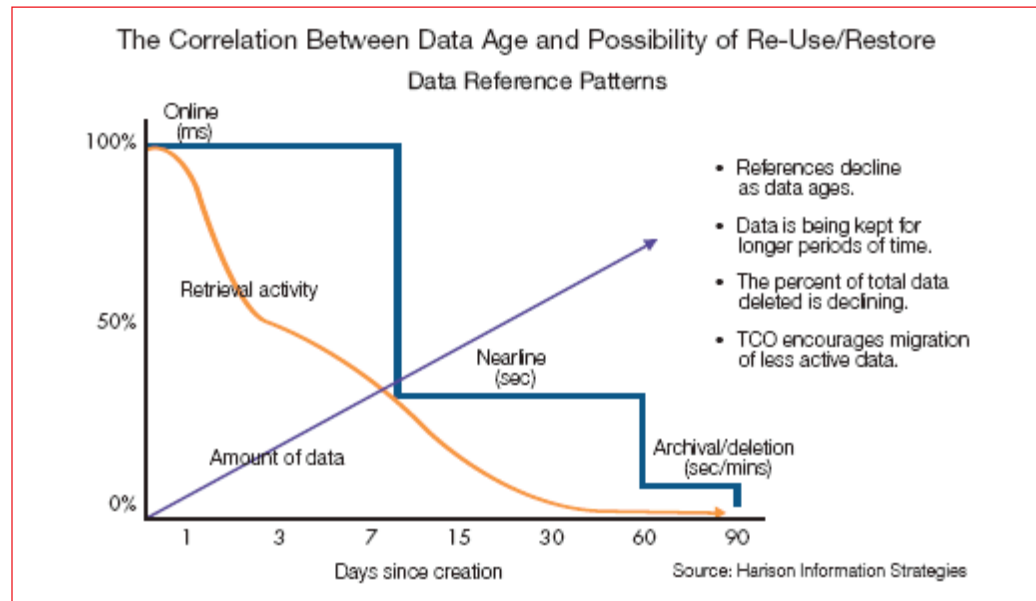


Figure 1. As data ages, it is less likely to be the focus of a restore request.

Matching Business Needs to Data Protection and Recovery Solutions

How do you best meet the data recovery requirements of each system in your organization and achieve the optimum RTO and RPO appropriate for your organization? Some organizations, or some particularly critical information within an organization, may require an exceptionally fast level of recoverability.

You may have different RTOs and RPOs for different types of business critical information. For example, a supply chain application that feeds a production plant may require a recovery time of only a few minutes with very minimal data loss. A payroll system that is updated weekly with only a few records may only require a recovery time of 12 hours and a recovery point of 24 hours or more before the impact will affect the business.

Any data protection strategy must ensure that information remains as accessible and available as needed to continue to drive revenue, profitability and productivity at acceptable levels no matter what planned or unplanned events occur. The data protection solution you choose should:

- Protect your data to a level that meets your business requirements and RTO and RPO.
- Manage business uptime as automatically as possible to streamline operations and save time.
- Assure the integrity and quality of your data during interruptions and when it returns to full operations.

Continuous Data Protection (CDP): A Breakthrough Innovation

The good news for businesses is that the technology for data protection is easier and more effective than ever before. Innovations have kept pace with the need to provide comprehensive data protection and make data recovery a quick and easy process. Perhaps the most exciting recent innovation in this area is the introduction of continuous data protection, or CDP.

CDP is a flexible disk-based technology that enables businesses to quickly and easily recover their data to any point in time. For example, it's not uncommon for a user to accidentally delete a critical file. Or for a virus to corrupt business data. These actions render the data unusable, even though the server or other hardware resources continue to work as expected. CDP enables you to recover a version of your data to a point in time just prior to the accidental deletion or virus corruption. This earlier version of the data can then be restored to the production environment.

Unlike tape backups, CDP technology does not require the interruption of applications to perform backups. It works continuously to back up your critical data to an alternate server so you can immediately recover data from any point in time. If you accidentally delete an important document or experience data corruption due to a virus or hacker attack, you can return to the point in time just before the problem occurred. Recovery occurs immediately with just the push of a button. Recovery for much larger amounts of data takes only minutes as shown in Figure 2. With CDP, both data protection and data recovery occur with only a fraction of the time and labor resources required by a tape-only strategy. It also eliminates the threat of major data loss posed by the infrequent recovery points of a tape-only strategy.

Continuous Data Protection incorporates several techniques from traditional backup, replication and snapshot solutions. How the CDP solution achieves its goals has much to do with its architecture and how it's configured. It's important to note the difference between "True CDP" and "Near CDP" solutions when evaluating your data protection strategy.

- **True CDP** – True CDP captures every data write and transfers them to a secondary disk. True CDP enables a data undo by allowing recovery to any point in time. This is especially beneficial for a data corruption issue, such as a virus. With true CDP, you can identify a tainted email, for example, then roll back to a point just prior to the time the email was received.

- **Near CDP** – Near-CDP differs from True-CDP in that you can only recover to specific points in time. For example, near-CDP will copy data when a file is saved or closed so the recovery point is only to the last known saved file. In some cases this could be several hours or more. In high transaction environments or environments with rigid compliance or governance regulations, this may not be sufficient.

The efficiencies and flexibility of CDP translate into much superior data protection and recovery, as well as cost savings realized through the elimination of both planned downtime for backups and lengthy, error-prone tape recovery processes.

Any Point-in-Time Data Recovery plus High Availability

“By 2011, some form of CDP will be deployed in 80 percent of the Fortune 2000.”
 —Gartner

High availability solutions are especially useful for data protection because they provide the power to keep your applications operating continuously, regardless of planned or unplanned downtime. HA solutions replicate your production environment to a backup server in real-time, and enable you to switch operations to the backup in the event of downtime. With recent innovations in automation and ease of use, they require little management or intervention to remain effective around the clock.

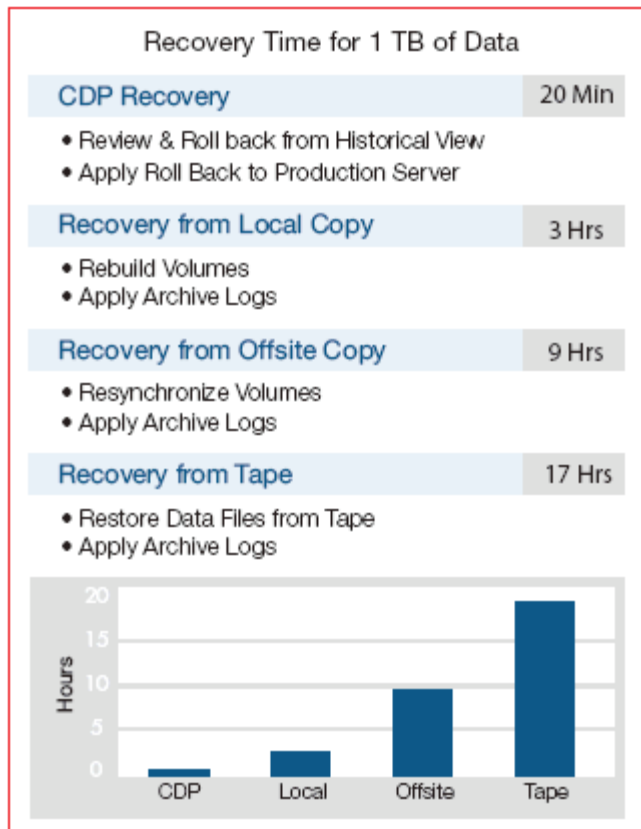


Figure 2. CDP offers significant savings in recovery time that can translate directly to cost savings.

But in certain data corruption situations, high availability solutions could use some help. Because HA solutions replicate data in real time to a backup server, they also replicate the data corruption to the backup environment. Options for recovering data that was accidentally deleted or corrupted run the spectrum from manually re-entering the work to utilizing advanced database technologies, such as journaling or logging, and generally require a tape restore.

However, when high availability is combined with CDP the issue of data corruption is no longer a problem. CDP enables you to simply return to any point in time previous to the corruption for instant recovery. The combination of HA and CDP provides seamless protection against data loss, data corruption and any type of IT or application downtime.

Conclusion

For businesses looking to take the next step in their data protection strategies, CDP is an essential consideration. Most IT analysts agree that businesses will be incorporating this strategy in the next few years as part of an integrated solution. CDP enables you to reverse data corruption in a fraction of the time and labor required for recovery from tape. It doesn't require planned downtime for backups and recovers data instantly at the push of a button. Whether as a standalone solution or integrated into an HA solution, CDP provides the easiest and most effective protection against the loss of critical business data.

Easy. Affordable. Innovative. Vision Solutions.

Vision Solutions, Inc. is the world's leading provider of high availability, disaster recovery, and systems and data management solutions for the IBM® Power Systems markets. With a portfolio that spans the industry's most innovative and trusted HA brands, Vision's iTERA™, MIMIX® and ORION™ solutions keep business-critical information continuously protected and available.

Affordable and easy to use, Vision products help to ensure business continuity, increase productivity, reduce operating costs, and satisfy compliance requirements. Vision also offers advanced cluster management, data management, and systems management solutions, and provides support for IBM i (formerly i5/OS), Windows® and AIX® operating environments.

As IBM's largest high availability Premier Business Partner, Vision Solutions oversees a global network of business partners and services and certified support professionals to help our customers achieve their business goals. Privately held by Thoma Cressey Bravo, Inc., Vision Solutions is headquartered in Irvine, California with offices worldwide.

For more information visit visionsolutions.com or call 800.957.4511.



iTERA HA
MIMIX HA
ORION HA
