

Evaluating a Recovery Management Solution

The top 10 metrics you should consider before buying.

Author
Siew Sim, CTO
Asempra Technologies, Inc.

Evaluating a Recovery Management Solution

In the American Heritage dictionary of the English language, recovery management is defined as the act, manner, or practice of managing a return to normal conditions. In many ways, synonymous with business continuity, recovery management in the realm of IT is indeed about returning systems, applications, and data back to “normal” conditions. When failure does occur, the goal is to bring IT business infrastructure back to its most recent consistent state and to restore business operations with efficiency and reliability, so as to minimize business downtime, and by extension, prevent significant financial loss.

A business can be disrupted by something as mundane as a missing file, a temporary power failure, or as extraordinary as a major natural disaster. The first step in recovery management is to protect (i.e., to be able to recover) critical business data, because while equipment can easily be replaced, lost information cannot. To handle these diverse set of failures, IT systems must be able to access and recover something as small as a singular email or file to something as large as the data center for an entire site. It must be able to recover data to a consistent state when the data is linked to active applications which are continually accessed by end-users. It is also important that a business should have the ability to recover its data to any point-in-time in the past, to a point prior to a failure, without missing any critical information.

Understandably, these goals are far easier to enunciate than to achieve, given the existing technological constraints. Yet, businesses are increasingly dependent on machines that must operate 24 x 7; the resultant build-up of business data remains firmly pegged to an exponential growth curve, making data protection a matter of increasingly vital importance, but which is paradoxically more technologically challenging than ever before. Current data protection and recovery tools offer only piecemeal solutions that are not only expensive to deploy, but in many cases fail to live up to today’s business requirements. The need for a new innovative and cost-effective recovery management solution is both acute and pressing. As a result, new recovery management technologies are coming to market. However, what are the real pain points in recovery management and what must a recovery management solution accomplish in order to resolve those problems? And how should IT evaluate the myriad of product options to find the solution that will work best for a specific environment?

The following sections address these questions by focusing on the ten metrics necessary to evaluate and objectively measure recovery management solutions.

The Evaluation Metrics

In order to evaluate a recovery management solution, one must have properly defined metrics. Data recovery service level agreements (SLAs) are traditionally measured by recovery time objectives (RTO) and recovery point objectives (RPO). RTO defines the time required to recover a set unit of missing data, and RPO defines the potential data loss – the time gap between the most recent application consistent recovery point and the physical failure point. RTO and RPO may be good objectives for setting SLAs with regard to data recovery, but they are not sufficient for measuring a recovery management solution. For example, a snapshot tool may recover a server’s data in minutes; however, a snapshot tool does not have the ability to recover a granular object. When one needs to locate a lost object from snapshots, the process is manual and the RTO could be many hours. In this case, RTO has nothing to do with the tool per se, inasmuch as it is entirely dependent on the manual process. While a data replication tool is capable of delivering zero or near zero RPO when a server fails, it is not capable of recovering business data if the data is corrupted, and the corrupted data is replicated. As a result, IT needs more comprehensive metrics to properly evaluate a recovery management solution. There are ten core metrics that fall into three categories – Recovery Time Characteristics, Recovered Data Characteristics, and Recovery Scalability Characteristics. The following sections explore these metrics in detail.

Recovery Time Characteristics

1. Recovery Time Objective (RTO)

When applying RTO as a measuring parameter for a recovery management solution, it defines how fast the solution is capable of recovering the data and application it is designed to protect. A block-level CDP and storage snapshot tool would recover a volume or a database in minutes or hours if the solution is capable of provisioning its secondary storage volume as the primary storage. A file-based CDP would require minutes to recover a file, and hours or even more than a day to recover a file system. The RTO of most recovery solutions depends on whether or not a verification process is needed prior to the recovery and the size of the data set to be recovered. The solution that can break established boundaries to provide instant recovery regardless of data set size would greatly reduce or eliminate business down time.

2. Recovery Time Granularity (RTG)

RTG determines the time spacing for selecting a recovery point; this is an important parameter for recovering logical failure. Unlike RPO, which determines the last recovery point prior to a physical failure, RTG defines recovery point selection beyond the last recovery point.

A data replication solution may have zero RPO for recovering from physical failure. When it encounters a logical failure (such as partial data corruption that is not detected for a period of time), the replication solution is not able to recover the data. As a result, the RTG would be undetermined. In this case, one would need another recovering solution with a good RTG to locate a recovery point in the past prior to - but also as close as possible to - the logical failure.

The RTG of different continuous data protection (CDP) solutions, such as block or file journaling, can be completely different; it depends mostly on how a particular CDP solution keeps track of data history. There are two classes of CDP, the real-time CDP solutions store their protected data in a time-based continuous data store, while the near-CDP solutions protect data continuously in real-time, but collect data into snapshots, which are then stored in the protected storage medium on a periodical basis. When a physical failure occurs, the RTG for real-time CDP solutions is in the order of seconds; whereas the RTG for near CDP solutions could be up to an hour (or hours), depending on the frequency at which the snapshots are taken.

Recovered Data Characteristics

3. Recovery Point Objective (RPO)

As a measuring parameter, RPO defines the minimum time gap between the last physical failure and the point-in-time where data can be recovered using a specific recovering solution under evaluation. Obviously, the smaller the time gap, the less data are lost. Since CDP (block or file journaling solutions) and data replication solutions continuously protect the changed data, their RPO capability is within a second. In contrast, the RPO of snapshot solutions is in the range of minutes to hours, depending on the snapshot duration.

4. Recovery Object Granularity (ROG)¹

ROG is for measuring the level of object granularity a recovery management solution is capable of recovering. Object granularity may be a storage volume, a file system, a database table, a transaction, a mailbox, a message, etc. For example, storage snapshot solution and block-level CDP solutions are capable of recovering data of ROG in volume only. Even though some block-level CDP solutions claim capability of recovering a database, the recovery process may be manual and labor intensive, in which case the ROG should not be database, it should be volume instead. The ROG of file-based journaling solutions ranges from files to directories.

5. Recovery Event Granularity (REG)

REG measures the capability of a recovery management solution to track events and to recover a failed application or missing data to a specific event. Most traditional data protection tools are not designed to explicitly track any event. The only implicit event most of these tools have awareness of is the time at which the data are backed up. In many cases, such an event is in itself an affirmative marker of consistency within the data, namely when the backup process had taken place while the application was in either shutdown or quiescent mode. Aside from the backup event however, there are many other events in an application that carry various degrees of significance, such as when a transaction is committed, when a message is accepted, when a file is closed, when a database is check-pointed, when a group of databases are consistent, when an application is upgraded, when a patch is applied, when a compliance rule is changed, when a specific IT operational event has occurred, when a fiscal quarter is closed, etc. These events are useful to an IT administrator during troubleshooting of failures when selecting a recovery point. When equipped with the capability to track consistency and business events, a real-time continuous protection solution would simplify a recovery process significantly.

6. Recovery Consistency Characteristics (RCC)

RCC defines the usability of recovered data by the associated application. RCC of a recovery management solution depends not only on how data are captured and stored, but also on the data type being protected.

When a volume-based hot snapshot is taken on a non-journal file system, the recovered file system would be inconsistent or corrupted, because not only could there be incomplete updates to the files, the file system structure may be in the process of being modified. When a hot snapshot is taken on a journal file system, the directory structure can be repaired during recovery but the file system as a whole is not consistent. This is due to the fact that journal file systems only journal the file system structure, they do not journal file content. While a journal file system is capable of self recovering its directory structure using its journal, the content of the active files cannot be consistently restored. Since a file system is actively modified by many applications simultaneously during runtime, the only time that a snapshot can be taken with consistency is when the file system is shutdown. Since block-

¹ David Freund. "Backup is dead. Long live backup!" InfoStor, August, 2004

level CDP or file journaling solutions are similar to taking hot snapshots continuously, the RCC of these solutions is inconsistent when applied to the file system. A real-time continuous protection solution can have strong consistent RCC only if it tracks consistency events and recovers each individual object in a file system to their own consistency point.

Unlike file systems, databases typically journal their own content; they also have built-in crash recovery to repair their content upon failures. During runtime, the on-disk image of a database is usually in a crash consistent state at best, with the state of its log ahead of the binary updates. Unless a database is placed in a quiescent mode, a snapshot of a database can only capture a crash consistent image. A database typically and periodically flushes its memory to the persistent storage to complete all the binary updates; this is known as a checkpoint event. At the time point of a checkpoint event, a database is in a strongly consistent state, with its log and binary in synchrony with one another. Since most databases store their contents in file systems today, the RCC of block-level CDP solutions is crash consistent at best. Block-level CDP solutions must be able to preserve write-order across multiple volumes if a database file spans multiple volumes. Without this, the RCC of these solutions has no consistency. A real-time continuous protection solution if capable of tracking database checkpoint would have strongly consistent RCC.

Recovery Scalability Characteristics

7. Recovery Service Scalability (RSS)

RSS is an important factor in evaluating a recovery management solution, simply because business data and applications are growing very rapidly. A recovery management solution must be able to scale with the applications and the data it protects. RSS is measured by service (number of applications or data sets the solution is capable of protecting) and capacity (the maximum size of the data it can store). Since CDP captures and processes data in real-time, a CDP solution must be able to keep up with the applications it protects. The ideal recovery management solution should be able to scale easily both in service and capacity through simple addition of processors and storage hardware without major reconfiguration and downtime. Most CDP solutions today are not GRID-based architecture; therefore RSS is usually of some concern.

8. Recovery Service Resiliency (RSR)

RSR defines how well a recovery management solution tolerates failures. A recovery management service must not cause an application to fail. It must be more reliable than the application it protects. When a recovery management service fails, the data service must fail over to another recovery management instance, such that an application would be continuously protected. A resilient recovery service should not corrupt its protected data; it must be able to self recover from its own failure. The self recovery should not be destructive, and there should be no impact to the applications it protects. A recovery management solution must be secured, such that individuals without the proper authorization cannot freely configure its policy. Unlike an application, a recovery management solution should not allow any individual to alter its protected data. Data history can only be purged by policies.

9. Recovery Location Scope (RLS)

RLS defines where the protected data must be presented when recovery takes place. Most data recovery management solutions by design require that the protected data be presented locally before it can be recovered back to the primary storage. The RLS of replication tools is also LAN because the recovery location must be where the replicated data reside. Internet-based data management services, on the other hand, protect and recover data over the Internet. As a result, their RLS is WAN. As businesses become more global, and government regulatory requirements for business continuity become more stringent, it is increasingly more important and valuable for a recovery management solution to be able to support both LAN and WAN RLS.

10. Recovery Management Cost (RMC)

RMC defines the cost efficiency of a recovery management solution. Data services such as backup, snapshot, replication, hierarchical data management, information lifecycle management, and archive are traditionally separate tools with very different architectures. This is simple because some of the tools are schedule-based with different interval requirements, while others are real-time, either synchronous or asynchronous. Typically, the tools that manage data history are not real-time while the ones that do not manage data history are real-time. As tools that manage history move towards capturing data in real-time, they would become more available to include the other real-time services.

For better RMC, some of the existing CDP solutions already combine backup and replication, and others provide both backup and hierarchical data management. Without service consolidation, IT administrators would have to manage several different tools manually. Consolidation of data services makes it easier for IT administrators to manage their data. In many cases, this is also a more cost-effective way to operate than via panoply of piece-meal solutions.

From the above metrics, one can readily see that although all CDP solutions (block-level, file journaling, and Asempra's Business Continuity Server) have one common characteristic, which is the ability to capture data in real-time, their resemblance stops there. Most of these solutions have significantly different recovery characteristics. The recovery characteristic depends on the data type captured by the CDP, and the way the protected data and metadata are stored by the CDP. In short, CDP is not the answer to all recovery problems; it is the specific technology capability behind each of these so-called CDP that makes the difference in an IT environment. Based on the needs of an IT environment, one CDP may work considerably better than another.

Given these metrics, is it possible to create a recovery management super solution that scores high in all measurable parameters? The answer is affirmative, if one would take a completely novel architecture-based approach to solving the problems. Such a super solution is capable of recovering data of any granularity almost instantly to any point-in-time. It is capable of recovering data objects ranging from an email to an entire data center and can guarantee overall data consistency. It is capable of distributing data to multiple locations and into different host servers, and able to recover data objects from different parts of the network. It performs real-time event journaling to capture data changes, application events, server events, and user defined events, then stores and indexes this information for use during recovery. Its architecture is based on Utility Computing (or GRID) that uses storage area network (SAN) as back-end storage to allow for unlimited service and data capacity scalability. It has the capability to self recover from failure and is designed to consolidate multiple data services into one, to include data protection, disaster recovery, archive services, and hierarchical policy management. Such a super solution exists today; it is the Business Continuity Server from Asempra.

Applying the Recovery Management Metrics

Evaluation metrics are only meaningful if you can validate them with real world examples. Therefore, it makes sense to present a practical example comparing the recovery management solutions that exist in the market today. The recovery management market has had multiple data protection and recovery solutions that have been in existence for years, but recently there have been a handful of emerging products and companies in this space that are altering the landscape. These products need to be included in the measurement of recovery management solutions in order to be comprehensive in our evaluation.

While several of these new products have taken a new storage approach which continually protects data at the block-level, one company has taken a much more innovative approach to solving the problem of recovery management. Asempra has recently announced the Business Continuity Server and we need to introduce this solution before evaluating it against the existing recovery management tools.

The Asempra Business Continuity Server

Asempra's Business Continuity Server™ (BCS) is a real-time recovery management platform that delivers transaction-aware continuous data protection with guaranteed data integrity and instant recovery. By leveraging existing infrastructure and reducing the management complexity, tools and storage required for today's data protection needs, the Business Continuity Server dramatically improves your data recovery capabilities while reducing your overall costs.

The BCS has a plethora of features, but let's focus on key differentiators that will apply to the recovery management metrics:

✓ **Continuous Protection of Every Transaction**

In order to intelligently protect corporate data and eliminate data loss, the Asempra BCS transparently and continuously protects your application data as transaction and checkpoint events occur in real-time. Backup windows are eliminated and recovery point objectives (RPO) approach zero data loss. The Business Continuity Server also reduces storage usage for data protection up to 90% or more by protecting I/O transactions at the byte-level.

✓ **Recovery Time Independent of Data Set Size**

Asempra's patented technology, Virtual On-Demand Recovery™, makes recovery independent of the size of the data set. Once recovery is initiated you can restart your application immediately – the Business Continuity Server recovers the data your application needs on demand, while it's running, enabling recovery time objectives (RTO) to be measured in seconds or minutes. By being transaction-aware, Asempra is able to provide a scalable solution that can recover anything from a single email message or file, to a complete server or database, and even an entire data center or region.

✓ **Verifiable Data Integrity**

With the data integrity and verification provided by the Asempra Business Continuity Server, your data is guaranteed to be recoverable. The Business Continuity Server stores your data in a continuous Write-Once Read-Many (WORM) format and verifies protected data exactly matches your application's primary data during every I/O transaction.

✓ **A Global Policy-Based Management Platform**

The Asempra Business Continuity Server software leverages your existing hardware, OS, and application environment to turn your infrastructure into a real-time recovery management platform. The Business Continuity Server also provides disaster recovery of your Microsoft Exchange, Windows file servers, and SQL database applications by protecting data over existing IP LAN and WAN networks. Integration with existing tape archival products is provided via a standard CIFS protocol in order to leverage your existing software to store data natively to tape.

Recovery Management Scorecard

The management complexity and cost of solving data protection and recovery issues today is rooted in the fact that it takes multiple tools to deliver a solution that still doesn't meet the new requirements of today's data center. This leaves IT professionals spending countless hours trying to integrate disparate tools and manually recovering data in an attempt to simulate a real-time infrastructure that is required to support their enterprise. As we discussed earlier, there are a myriad of protection and recovery tools to choose from so it made sense to come up with the core metrics necessary to enable IT management to evaluate which solutions would fit their environment best.

Now that we have a detailed understanding of the "Top Ten" metrics necessary to evaluate a recovery management solution, let's apply these metrics to solutions that exist in the market today. Practical application of the metrics enables not only a solidified understanding of the metrics, but also a better comprehension of available solutions and how they compare.

Recovery Management Requirement		asempira	Traditional DP	Block-level CDP	Block Replication	File-level CDP
Recovery Time Objective	RTO	Instant via VODR	Hours to Days	Minutes to Hours	Minutes to Hours	Minutes to Hours
Recovery Time Granularity	RTG	Second	24 hours	Seconds to Hours	None	Seconds to Hours
Recovery Point Objective	RPO	Near Zero	24 hours	Near Zero	Near Zero or Minutes	Near Zero
Recovery Object Granularity	ROG	Transaction Object	File or blocks	Storage blocks	Storage blocks	Files only
Recovery Event Granularity	REG	Fine-grained - consistency events	Coarse - shutdown event	None	None	None or Coarse
Recovery Consistency Characteristic	RCC	Strong consistency	Strong consistency	Crash consistent only	Crash consistent only	Crash consistent only
Recovery Location Scope	RLS	Unlimited LAN & WAN	LAN-only	LAN-only	LAN and WAN	LAN-only
Recovery Service Scalability	RSS	Unlimited service and data	Limited service	Limited	Limited - one-to-one	Limited
Recovery Service Resiliency	RSR	Multi-nodes fault-tolerant self-verifying	None	Active-passive cluster	None	Active-passive cluster
Recovery Management Cost	RMC	\$	\$\$\$\$	\$\$\$	\$\$\$\$\$	\$\$

Summary

In most industries today, the service level agreements for data protection and recovery have moved to a point where there is no time for backup windows, no tolerance for data loss, and very little margin for recovery downtime. Add to that the increased business demands for disaster recovery of mission and business critical data, along with new compliance requirements and you can quickly determine that the legacy tools of data protection and recovery are ill-equipped to handle today's requirements.

Given the realization that traditional data protection and recovery tools no longer suffice, and new technologies have emerged to address increased business expectations, the time has come to put in place the objective metrics needed to properly evaluate the array of recovery management solutions.

The ten metrics of recovery management above enable IT management to apply thoughtful consideration to their own internal business requirements against the products they are evaluating. When looking at recovery time, recovery data, and recovery scalability characteristics, we believe the Asempra Business Continuity Server provides a compelling recovery management solution that will exceed the service level expectations of your business. Designed to address today's demanding business needs and architected to scale for future growth, Asempra provides the most comprehensive enterprise-class recovery management solution.

About Asempra

Asempra Technologies is a leading provider of real-time recovery management solutions, which delivers continuous data protection with guaranteed data integrity and instant recovery. The Asempra Business Continuity Server™ leverages existing IT infrastructures and reduces the management complexity and tools required for today's data protection needs.

Compliments of:  **ASEMPRA**

 **ENTERPRISE**
Storage Solutions

3835R East Thousand Oaks BLVD. #315
Westlake Village, CA 91365
Tel 877.230.2837 / Fax 805.435.2500 / www.ess-direct.com