Pund-IT

PUND-IT, INC.

NOVEMBER 2, 2005

Weekly Review

Only the Shadow Knows

By David Hill, Mesabi Group LLC

Service level objectives (SLOs) for applications should specify both a recovery time objective (RTO) — the time required to restore an application to a working state after a downtime situation occurs — and recovery point objective (RPO) — the amount of data that is exposed to permanent loss. For an annual high availability (HA) of 99.999% (a.k.a. "five nines") for mission critical applications, the Storage Networking Association (SNIA) suggests an RPO of 1 minute and an RTO of 1.5 minutes. However, for very critical applications even five nines may not be low enough while for other applications four nines (recommended RPO of 10 minutes and RTO of 15 minutes) or even three nines (2 hours for both RPO and RTO) of availability may be sufficient.

Data protection technology that meets low RTO/RPO constraints delivers a high level of data protection. For *physical* data protection (for example, the failure of a disk drive), low RTO/RPO is standard and can be easily achieved. For ongoing business operations at individual locations, RAID technology is well-established. For disaster recovery at a remote location, remote mirroring has proven to be a successful approach. The same is not true for *logical* data protection. Logical data protection insures against logical failures — such as viruses, data base corruption, and inadvertent file or table deletions. Physical techniques such as remote mirroring cannot protect against data corruption, since the corruption simply propagates without regard to protective measures. Due to its inherent limitations, restoring from tape (or even from a virtual tape library) is unlikely to meet low RTO/RPO requirements. Snapshots provide logical protection, but standard snapshot approaches probably do not provide the necessary level of granularity.

Logical data corruption is therefore a *huge* data protection exposure problem that most IT organizations have yet to either recognize or address. And in organizations that depend on logical data, carefully-written SLOs are simply a wish list that does not yet translate into the proper level of overall protection required.

Enter Continuous Data Protection Stage Right

Continuous data protection (CDP) solutions propose to address logical data protection problems, though they may also provide an extra measure of physical protection. CDP solutions propose to restore data to *any* point in time by capturing a copy of each I/O as it occurs. That is, on the surface, they promise a RPO of zero (no loss of data). They cannot guarantee a zero RTO since IT administrators must perform a forensic analysis to determine the time to which the data should be restored. In addition, administrators may have to accept some data loss in order to restore to a point before the corruption occurred since it may be difficult to separate the good data from the

bad after the corruption point took place. However, CDP fulfills the requirements for low RTO/RPO for logical data protection.

Recent announcements by major vendors including EMC, HP, and IBM suggest that the market for CDP is real. Smaller suppliers — InMage, Kashya, Mendocino Software, Revivio, Storactive, TimeSpring, and XOsoft among others — illustrate the vitality and depth of CDP choices that IT managers have to choose from.

Table 1: A Sampler of Continuous Data Protection Suppliers

Vendor	Product	Product Focus	Technology Foundation
EMC	RestorePoint	Initial focus on Oracle and on Microsoft SQL Server	Uses a RecoverPoint engine to manage the metadata needed for CDP and works within EMC's Replication Manager
HP	RecoveryONE	OEM's Mendocino Software's block-based CDP software	Block-based data capture product that HP will support with professional services
IBM	Tivoli Continuous Data Protection (CDP) for Files	Users can self-restore files that they create.	Up to three copies of a protected file — one on the current system, one to a file server, and one to a Tivoli Storage Manager (TSM) backup server
InMage	DR-Scout VX	Disaster recovery (DR) solution for databases, e-mail, and file services	CDP is a part of an overall solution for DR that also includes pre-deployment modeling and trending as well as compression, encryption, and bandwidth management
Kashya	KBX5000 CDP	Provides both local (operational) and remote (DR) CDP for Oracle and SQL database environments	Application-aware I/O "bookmarks" for Oracle and VSS integration for SQL environments map application events to specific points in time
Mendocino Software	RecoveryONE	Continuous disk-to-disk data protection with what it calls "time slider" technology	It can provide readable, writable snapshots of any retroactively selected point in time.
Revivio	Continuous Protection System (CPS)	Restoration to any previous point in time for business applications	Appliance for either high-availability enterprise-class or workgroup class applications
Storactive	- LiveBackup - LiveServ	- For laptop and desktop PC's - For Exchange Servers	- Automatically tracks changes for end-users - Captures each new, modified, or deleted Exchange message in Real-Time
TimeSpring	TimeData Continuous Data Protection Software	Turn back time to the exact moment before a data loss or corruption occurred	Captures all changes to a secondary storage repository without server performance degradation or downtime

XOsoft	Enterprise Rewinder Product Suite for Continuous Data Protection	Fast recovery for corrupted databases for Exchange, SQL, and Oracle	Protects against isolated data corruption episodes and complements current operational backup and DR technologies

Source: Mesabi Group November 2005

Enter Continuous Shadowing Protection Stage Left

That would seem to be the end of the story. However, there is a class of products that does not strictly adhere to the general definition of CDP i.e., cannot guarantee any point-in-time recovery yet provides low RPO/RTO solutions for logical data protection. Since some of these products use tight interval snapshots, one school of thought has been to name the class as a form of snapshots. There are two problems with this. The first is that using frequent snapshots is only one of the technologies that can provide this type of solution. The second is that naming the category as a subclass of snapshots can lead to further confusion in mixing up standard uses of snapshots with specific low RPO/RTO uses. Also, a CDP solution, for example, Mendocino Software, may use snapshots as long as they can be created after the fact.

The other school of thought is to ignore the religious distinction and lump these solutions into CDP anyway. The problem with this is that one of the large players in this category is Microsoft, whose Data Protection Manager (DPM) is supported by many storage vendors, including CommVault Systems, Computer Associates, Dell, EqualLogic, HP, and Quantum. This class of products qualifies as what we consider "continual shadowing protection" (CSP) products. Continual suggests regular or frequent occurrence, but does not mean uninterruptible in the way that continuous does. Shadowing, in this situation, refers to screening or protecting data. Now the Society for the Prevention of Additional IT Jargon may have a fit, but this class already boasts Microsoft and Symantec as members. Mimosa Systems is another player in what we expect will be a rapidly growing space.

Table 1: A Sampler of Continual Shadowing Protection Suppliers

Vendor	Product	Product Focus	Technology Foundation
Microsoft	Data Protection Manager (DPM)	Disk-based backup and recovery for data on Windows file servers and network-attached storage devices	Snapshot-based (up to 64 with about a minimum hour granularity). Part of both Microsoft's Universal Distributed Storage plan and System Center family of products
Mimosa Systems	NearPoint	Continual application shadowing for Microsoft Exchange	Ships Exchange logs as soon as they are closed to an appliance; appliance serves both data protection and archiving needs
Symantec	Backup Exec 10d includes Backup Exec Continuous Data Protection Server (a.k.a. Panther)	Quick restoration of files by end users using a "Google-like" Web interface on a desktop or a laptop	Maps block-level granular changes on files to disk via VSS (Volume Shadow Copy Service) snapshots on a Veritas Backup Exec server

Source: Mesabi Group November 2005

Choosing Between a CDP and CSP Solution

There are several criteria to consider when choosing between a CDP and a CSP solution (as well as for a low RPO/RTO solution in general):

Application support — Does the solution support the type of "application" that is needed, such as Oracle or Microsoft Exchange?

Good enough — Choose a solution that is "good enough" to meet acceptable RTO/RPO requirements. If products in both categories meet the criteria, then common issues such ease of use, ease of integration, scalability, and cost) may help break the tie.

Supportability — What level of support is required? If a Microsoft "Good Application Housekeeping Seal of Approval" is required, the product is likely to fall within the CSP category.

Consistency group — Does the solution enable all of an application's data (even if scattered across multiple servers and disk volumes) to be managed as a single entity?

Durable restore points — What process is required to restore data to a point at which the application can use the data? How much data is at risk and how long does the process take?

Roadmap — Where is this product (and its vendor) going in the future and how does it fit into the organization's future plans?

Mission Accomplished?

The CDP and CSP market is in now in its wild-wild West phase. The next year should see the introduction of new products and the introductions of new functions and features in existing products. One area of focus will be tighter integration of products into an overall portfolio of data protection solutions. Another focus will be on achieving greater clarity on issues that are generating heat today, such as the ability to restore to a robust consistency point.

Well, should IT managers wait until the market sorts out and moves to the second generation products? That may be a grave mistake as it could perpetuate the exposure to logical data protection problems. Even though a great deal of sifting and sorting through numerous, sometimes confusing options is required, we believe that IT managers are well advised to consider and evaluate how the current generation of CDP and CSP solutions might fill their organizations' data protection needs.

David G. Hill is principal of the Mesabi Group (www.mesabigroup.com). The Mesabi Group focuses on the revolutions in Storage Networking and Storage Management, and helps clients make the best and most efficient use of information for business value.

Contact:
Pund-IT, Inc.
Phone: 510-909-0750
E-mail: charles@pund-it.com
Web: www.pund-it.com