

PUND-IT RESEARCH

Weekly Review

September 27, 2006

In This Issue:

**Winning Strategies – Playing the Data Protection
Management Game More Effectively**

Pund-IT, Inc.
2776 Sulphur Drive
Hayward, CA
U.S.A. 94541

Phone: 510-909-0750
Fax: 510-886-4937
charles@pund-it.com
www.pund-it.com

Winning Strategies – Playing the Data Protection Management Game More Effectively

By David G. Hill, The Mesabi Group

Data protection is a must-do-well responsibility for all IT organizations. Even the temporary loss of availability of data to key applications may have unpleasant consequences ranging from the loss of revenue to reduced productivity, and the permanent loss of key data is unthinkable. So, to no one's surprise, IT organizations invest considerable sums in data protection technologies.

Although the data protection industry is undergoing a tectonic-like change (in the sense of magnitude of transformation; not the rapidity of the change), IT organizations still depend upon traditional backup/restore software-based processes as the backbone. As the old saying goes, the more things change, the more they stay the same. And the fact that a virtual tape library (VTL) may front-end a physical tape library does not alter basic backup/restore processes.

Yet in a large number of organizations these processes are at best frayed and at worst broken. In order to determine how well protected their data is, companies must answer two questions about the state of the processes.

1. Can IT guarantee that all data that needs to be restored after a production data loss is backed up all the time?
2. Can IT guarantee that all data that needs to be restored will be able to be restored in a timely manner that is consistent with the capabilities of the data protection technologies that are used?

If the answer to both questions is not an unequivocal yes, then the data protection processes are not delivering the necessary level of service.

The Blame Game

Where does the fault lie if the answers to the two questions are unacceptable? Some direct blame at tape technologies in general, yet we believe that is misplaced. Even if a VTL is used to complement a physical tape library, that does not solve the process problem if one already exists. Yes, performance, reliability, and manageability issues may be topics for discussion (and a VTL may very well provide benefits in those situations), but the main source of the problem is further upstream. All tape (and complementary disk-based) technologies do is serve as targets for backups and sources for restores. They do not affect fundamental data protection processes.

Can blame then be placed on the backup/restore software itself? The answer is generally no. Most such software is very robust and sophisticated. Backup/restore software copies data that is to be protected from a source and writes it to a target. The software reverses the process (in essence) when data restoration is called for. Now issues such as performance and manageability may differentiate backup/restore software, but the process culprit is typically not the backup/restore software. That software simply does what it is told — backup or restore — and does so to the best of its ability.

No, the real culprit for poorly performing data protection processes, primarily backup/restore processes, is complexity. One simple source of complexity is the never-ending growth of data. When backup/restore processes have been fully optimized performance-wise, if additional data is added to a backup job, it logically takes longer to run. Additionally, requirements to keep applications up longer (a growing necessity among increasingly globally-focused businesses) shortens the time available to do backups. So while the "days" are getting shorter, the amount of work that needs doing each "day" is increasing. Consequently, delays due to restarting a failed backup or coping with network congestion that prevents a running backup job from completing in the allo-

cated time has to be addressed as quickly as possible. The tradeoff between having an application up that is not fully protected or having unplanned downtime for an application while a backup job runs is unpalatable.

A second cause of complexity is mixing heterogeneous products (more than one type of backup/restore software, more than one type of operating system, and more than one type of storage). The amazing thing is that — despite all the problems inherent in such environments — backup/restore processes continue to run.

The New Monkey Wrench in the Works — Compliance

Risk management responsibility should be enough incentive to inspire any enterprise to deal effectively with data protection process complexity. Yet historical ways of doing business (“That’s the way we have always done things around here”) as well as cost pressures on IT organizations may have dulled active responses to the complexity problem. After all, the base objective is to provide a satisfactory level of data protection while minimizing costs. So long as the data protection processes can limp along at a minimal level of adequacy, many in IT management are willing to turn their time and attention to more pressing matters.

The demand for compliance by government and regulatory agencies not only adds extra added incentive, but also changes the essential data protection game. Compliance attracts the attention of senior management anxious to avoid fines or even jail time, and any IT-related activity that attracts senior management also garners the attention of IT management. The objective now is ensuring that data protection is performed completely and accurately. Cost remains a constraint (no organization wants to overspend), but it is no longer the objective (as extra funds can be found to do what is necessary). What are some of the key factors driving compliance?

- ***Requirements for Regulatory Compliance***

The Sarbanes-Oxley Act requires publicly-traded companies to monitor and verify the authenticity of financial records. CEOs and CFOs of such companies are held personally accountable to the extent that their personal freedom could be taken away. Consequently, these executives realize clearly that the Act is a mandate, not an option.

The Act requires that there be no destruction, alteration, or falsification of financially-related records. Each company must establish and maintain an adequate internal control structure and procedures for financial reporting. That includes an assessment of the effectiveness of that approach. Data protection performs a key role in that internal control structure. Compliance requires enterprise to be able to completely and accurately recover all required data in the event of a logical or physical failure. No loss of data that would invalidate the integrity of the financial reports is acceptable. That means that IT has to be able to give an unequivocal yes to the two questions asked earlier about is all the data protected and can it all be restored.

- ***Self-Regulation Is Also a Necessity***

Many organizations may feel that they are not subject to regulatory compliance or, if they are, that by meeting the letter of the regulatory law that they have done all that they need to do for regulating themselves. That is a dangerous misconception.

Enterprises need to put in place the necessary self-disciplines to make sure that they can deliver the proper level of risk management service, say for litigation support, as part of the self-regulation of a trade association, or simply as good business practice. For example, an enterprise should have a uniform policy for retention (and destruction as appropriate) of all e-mails. All available e-mails should be available quickly with completeness and accuracy that can be attested to by an outside auditor. Performing eDiscovery quickly at a judge’s request is a lot better than responding to a discovery order and having to dredge up and examine all tapes even from however long ago and wherever located. In itself, that process can be time-consuming and expensive. Failure to

find information that should be there can be very expensive — as numerous firms have found out to their dismay.

- ***The Focus on Risk Management Is Now both Internal and External***

Data protection used to focus only on internal needs, such as returning a down application to working status as quickly as possible. (Even though the customers who use the system may be external, the service level objectives, such as recovery point objective and recovery time objective, are internal.) With compliance, including self-regulation, the focus is on presenting an image of the enterprise to the outside world. Since failure to comply implies an inefficiently or feloniously run organization, that ups the ante for companies to manage their data protection infrastructure more effectively.

Data Protection Management Rides to the Rescue

To deal with all the problems with the data protection processes, many companies need additional help in better managing data protection processes. Data protection management (DPM) is the name for that category of products that help manage data protection environments. DPM products *do not do* data protection, but do enable the better management of data protection processes that perform the actual data protection. These include backup/restore software and continuous data protection (CDP) appliances as well as the other elements of the IT infrastructure that make up the data protection “ecosystem.”

The word ecosystem implies that there are interrelationships among the various components — or domains —of the IT infrastructure, including servers, networks, storage, applications, operating systems, file systems, and databases. For example, if a network is congested and backup I/O traffic cannot transverse the network in the allocated time, a backup job may not be able complete within the planned backup time window.

This example illustrates the need for IT management to have both timely and actionable information. Information must be timely to either prevent service-level-impacting events or, failing that, minimize the damage of the service-level-impacting events that have already occurred. Actionable means that the problem can be alleviated — either on a one-time basis or permanently. DPM delivers the reporting, monitoring, and troubleshooting capabilities that IT needs to manage data protection processes more effectively.

Among the types of issues that data protection management products can address are:

- *Ensure completeness of data protection coverage* — for example, by determining if any servers have not been backed up successfully and whether there are any servers for whom backup has not been attempted at all
- *Speed-up response to real-time data protection problems* — facilitating the troubleshooting process to identify and rectify potential or actual data protection service-level impacting events.
- *Carryout long-term backup window problem analysis* — performing a pattern analysis (e.g., determine from historical information the slowest, fastest, and most unreliable components of the data protection infrastructure) in order to see if there are any systemic issues, such as repeating problems or bottlenecks, that need to be addressed.
- *Perform preventive maintenance through predictive analysis to prevent unnecessary negative service level impacts* — using historical information to do a trend analysis to determine when elements of the data protection environment will exceed a predetermined threshold, such as when pieces of tape media will run out.

Dramatis Personae

The vendors that focus on data protection management are younger companies (Table 1). That does not mean that they do have relationships with some of the largest storage vendors. APTARE has a strategic partnership with Hitachi Data Systems (HDS). Bocada lists partners including: CA,

EMC Legato, HP, IBM Tivoli, NetApp, Sun, and Symantec. Illuminator is a NetApp Advantage Partner. Servergraph counts IBM/Tivoli and Symantec among its partners. Tek-Tools claims Brocade, CA, EMC, and NetApp among its technology partners. WysDM claims EMC among its channel partners, and NetApp, Oracle, and Sun as technology partners.

Table 1: Sampler of Data Protection Management Vendors

Vendor	Product	Product Focus	Technology Foundation
APTARE	APTARE StorageConsole	Seeks to be an early warning system to enable an organization to manage its entire data backup and recovery infrastructure.	A Web-based tool enables real-time threshold event management as well as manages a data protection and problem management database for historical analysis of past events as well as predictive intelligence about future data protection events.
Bocada	Bocada Enterprise 4	Focuses on providing information to ensure that data protection services meet service level objectives, including assured recoverability and regulatory compliance	Agentless software that aggregates, organizes, and presents views of an enterprise's data protection services across dispersed and heterogeneous environments
Illuminator	Restore-Illuminator	Focuses on the key issue of recoverability of enterprise applications rather than just the optimization of backup processes	Its software features the ability to tightly integrate metadata from applications, servers, storage, and backup applications in order to discover, analyze, report on, and resolve recovery gaps that might impact service levels
Servergraph	Servergraph Product Suite	Focuses on graphics and drill-down capabilities to help isolate and fix backup environment problems	Its agentless software not only can detect errors, but can be configured to take corrective action to events and thresholds as prescribed by an administrator, i.e., self-healing
Tek-Tools	BackupProfiler	Focuses on monitoring and reporting to ensure that no backups are missed	Its cross-vendor backup resource management software provides a consolidation view of all backups across all servers for both reporting and analysis
WysDM	WysDM for Backups WysDM for Fileservers	Focuses on making the management of data processes smart so that enterprises can move from reactive to proactive management of those processes.	Its software uses a predictive analysis engine in conjunction with a "data mine" repository and does continuous data collection to enable cross-domain (servers, networks, and storage) correlation for root cause analysis to find the true cause of a potential or real data protection problem

Source: Mesabi Group, September 2006

Mission Accomplished?

One might ask that if the need for data protection management was so critical for so long, why it was ignored before these companies burst upon the scene? One answer is that IT organizations expected their backup/recovery software vendors to do the job. Although those tools might provide some help in homogenous IT environments, DPM was not really their job. A second answer is that backups might be best viewed as application silos rather than as an overall service function. A third answer might be that until recently the pressures of continued data growth and the resulting complexity had not yet reached a critical stage.

But whatever the reason, these are now mere excuses that no longer matter. Compliance is the tipping point, and a side benefit is that in adhering to compliance rules, that enterprises also enjoy a level of data protection that they should have been getting all along anyway.

Monitoring and reporting is the first step in achieving efficient data protection management. Absorbing DPM capabilities and making use of them is the first task for IT. Then IT may very well want to turn to a deeper analysis of the information to be able to find and resolve the real problem that causes real or potential problems (rather than having to deal with a set of cascading alerts where the causal needle is hard to detect in the infrastructure haystack). A future goal may be to enable self-healing or other advanced processes, but IT organizations have more than enough DPM tools to work with and would be well-advised to get started now rather than wait.

© 2006 The Mesabi Group. All rights reserved.

About The Mesabi Group

The Mesabi Group helps organizations make their complex storage, storage management, and interrelated IT infrastructure decisions easier by making the choices simpler and clearer to understand.