UNIFIED DATA PROTECTION COMES OF AGE

JUNE 2013

THE NET NET

Traditional data protection is three decades old and is definitely showing its age. Poor management oversight, data growth, virtualization, data silos, and stricter SLAs all conspire to strain traditional backup to the breaking point.

Traditional backup usually follows a set pattern: full baseline backup, daily incremental backup, full weekly backup. When backup volumes were smaller and fewer, this process worked well enough. But a daily operation creates backup data that is missing up to 20 hours or more of current data input, making it impossible to restore to a meaningful RPO. The obvious solution is continuous backup with frequent snapshot recovery points. But this type of backup product can be expensive and resource-intensive, and IT often reserves it for a few Tier 1 transactional applications. But what happens to large and popular business applications such as email, back office files and content management systems? Failed backup and recovery can still devastate a business.

This article will look at why traditional backup is so difficult to do well these days, and why the risk and expense are so high.

Backup and Recovery Challenges

- *Lack of operational oversight* is a challenge for IT. Backup and data integrity verification are difficult to impossible, and there is little confidence in reaching Recovery Time and Recovery Point Objectives (RTO and RPO). Continuous data protection (CDP) with integrity verification provides more confidence but is resource-hungry and expensive to run.

- *Exponential data growth* results in poor backup and restore performance and requires frequent provisioning. IT is under big pressure to lower costs but they must protect growing data at the same time – and adding more storage and data protection is not cheap. Even backup to a public cloud steadily increases costs as the cloud stores more and more data.

- *Virtualization growth* has a serious impact on backup resources. Startups have introduced innovative products optimized for virtual networks, and mature data protection vendors added virtual support to their legacy backup applications. But the startups lacked the resources to develop for the physical environment as well, and well-established backup vendors tried to tack on new support code to decades-old applications.

- *Inefficient data silos* have been around for years, and are more costly and inefficient than ever. Separate backup products for virtual and physical networks only add to the silo problem. IT is left with limited backup scalability and poor support for multiple domains, and ROI continues to diminish. Silos also impact WAN transport costs. Limited bandwidth slows data movement to and from remote sites and the cloud, which affects the speed of replication and recovery. IT purchases additional WAN accelerators to provide fast remote transport for multiple point products, which adds to inefficiency and cost.

- *Service level agreements (SLAs)* are crucial to maintaining application availability. However, too many data protection products offer mediocre insight at best into application recoverability. And data silos running backup point products worsen the problem, making it very difficult to test and remediate recoverability.
- *Heavy management overhead* is epidemic in traditional backup environments. Numerous backup systems are costly to purchase, upgrade and manage. Backup customers expect features like automated scheduling, policies, flexible backup targets, and replication. But adding these features to multiple backup products across multiple data silos only increases expense and complexity. These capabilities are necessary to modern data protection but make the full backup infrastructure very difficult to manage (and afford).

New Architectures

Using aging backup products may be painful but there is a lot of inertia around replacing them IT knows very well they have to buy software, update hardware, buy third-party products to fill in the gaps, migrate old backup data so it can be restored, and integrate the whole package with existing network tools. Overwhelmed IT departments often decide just to give the old backup one more year, or buy a stop-gap tool just to get by. In the face of these pressures, the backup replacement had better be compelling and cost-effective enough to justify the switch.

This is where evolutionary backup technology appears front and center. Unified backup platforms are a strong trend because they extend unified data protection to applications, virtualized networks, physical networks, and multiple operating systems. They are based on the concept of snapshots and changed block tracking for near-continuous backup, global dedupe, automatic backup verification, and instantaneous restore.

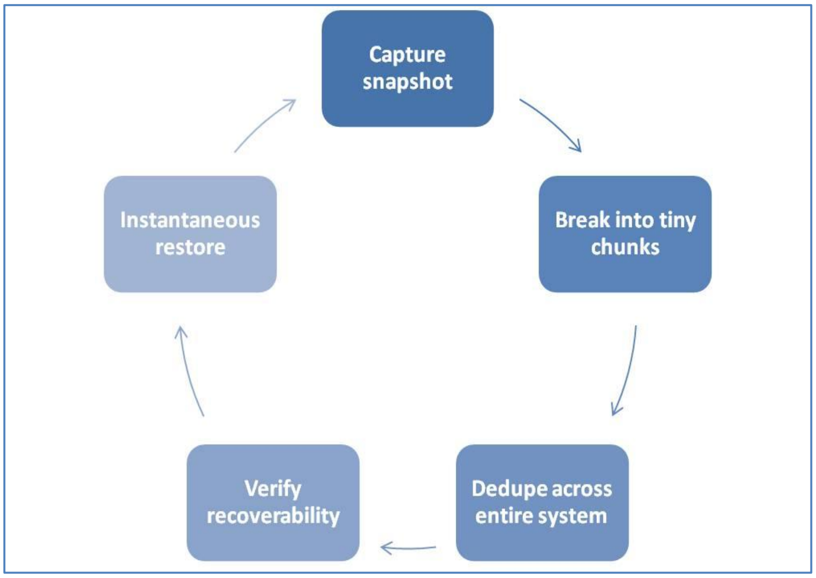

Figure 1. Optimal backup flow

Architectural Requirements for Next Gen Platforms

Let’s take a look at the requirements for a next-gen backup platform.

Requirement	Description
Near-Continuous Data Protection	Snapshots and changed block tracking (CBT) combine to produce continuous, incremental-forever backup. Incoming blocks are deduplicated across the entire backup infrastructure for real efficiency and capacity savings.
Extreme Recoverability	Recoverability depends on the integrity of the backup data and the ability to meet application RPO and RTO requirements.

	Recoverability should also include verification, where IT is confident that a backup has completed successfully with full data integrity.
Efficient Cross-Domain Support	New architectures provide common backup and recovery features and centralized management for physical and virtual systems, applications, and operating systems.
Cost-Effective Operations	Save money by accelerating backup times and frequencies, shrinking backup data sizes, and accelerating restores. The platforms also come with built-in replication capable of fast WAN speeds and should be highly scalable for leveraging backup operations across the data center.

DETAILING THE REQUIREMENTS

- **Near-Continuous Data Protection.** Traditional backup depends on making copies. A lot of copies. Most admins schedule incremental backup to control backup size and length but they must run full backups at least once a week, and on Tier 1 applications once to twice a day. In contrast, new data protection architecture takes a baseline image and combines it with CBT for incremental-forever backup. In addition, the changed blocks are deduplicated across the entire backup infrastructure for real efficiency and capacity savings.

- **Extreme Recoverability.** Recoverability depends on two factors: 1) how trustworthy is backup data integrity and 2) how fast can the system recover applications and data to a working state? In answer to the first question, backup verification is immensely important as recovery speed means nothing if the restored data is compromised. Verification should test that a backup completed correctly, even when backup is running every five minutes. Specific application verification is also a big plus in next gen platforms. Application testing gives IT and application admins high confidence in backup and restore integrity for critical systems. As for instantaneous recovery, the backup system should be able to priority-queue backup data for immediate restore. The system should also be capable of quickly restoring applications to working order while full data restoration occurs in the background. This capability can save hours to days of downtime with applications, a huge benefit to both IT and users.

- **Efficient Cross-Domain Support.** A number of backup applications support both physical and virtual operations, but the operations and management interfaces are different. This introduces backup system complexity and yet more management overhead. New architectures provide backup and recovery for physical and virtual systems, applications and operating systems – and they do it using shared backup features and a common management interface. Leveraging shared features such as global dedupe unifies backup operations across domains for greatly increased efficiency. An easy to use common management console is highly efficient and diminishes management overhead.

- **Cost-Effective Operations.** Next-gen platforms save money on backup resources by using CBT/incremental-forever to accelerate backup times, shrink capacity requirements and greatly accelerate restores. They should offer built-in fast and flexible replication with strong encryption options like AES-256. Modern platforms efficiently dedupe and compress data for fast data movement without the cost of additional equipment. Scalability is also important for optimizing backup and recovery costs across multiple applications, networks and operating systems. Petabyte storage levels along with deep data compression will yield high scalability, and connecting platform systems to one another will further scale storage load balances, common technology features, and centralized management.

Copyright© Taneja Group, Inc. 2013. All Rights Reserved.

Taneja Group Opinion

Near-continuous backup, near-immediate recovery, and broad domain support are not easy to do, and when you add in cost-effectiveness you have a real challenge on your hands.

Vendors are certainly trying because there are rich rewards for success in the data center. Many backup vendors offer scalable solutions with centralized management. Some vendors concentrate exclusively on virtual networks but the market leaders are attempting to protect a variety of environments in the backup infrastructure.

In our view Dell AppAssure 5 offers key features to accelerate backup and recovery and to protect data integrity, and they do so economically. In addition, AppAssure is hardware agnostic so that customers can leverage their existing storage assets, which provides CAPEX savings.

The product uses agents on protected systems for highly intelligent client-like abilities. The extreme efficiency of snapshots with CBT plus global dedupe provides near-continuous backup across Windows and Linux servers; VMware, Hyper-V and Xen virtual servers; and MS SQL Server, SharePoint and Exchange application servers. Automated recovery assurance protects backup integrity and recovery is very fast – allowing users to bring an application online without waiting for full data restore. Recovery also enables bare metal recovery and machine cloning.

Running multiple backup tools can be complex and costly, and extreme data growth and virtualized networks are worsening the problem. We strongly suggest that IT consider platform backup technology that unifies backup needs across the data center, and that is founded upon scalability, ease-of-use and exceptional efficiency. The more backup operations that IT brings under the platform's control, the greater the company's return on investment.

What about Agents?

For years debates have raged over the use of agents in backup software. Agentless solutions do have benefits such as no extra load on the protected server and no need for agent management features.

However, agent based solutions have strong advantages as well. Agents provide application awareness such as the ability to gather metadata about the application including application versions and components, configurations, and the ability to truncate database logs.

Intelligent modern agents act in concert with the host to quickly carry out backup and recovery operations directly on the protected machines.

Furthermore, even an agentless system has to assign resources to backup and recovery operations. In their case this intensive processing occurs in the host, which can require far more resources than distributing operations among agents.

NOTICE: The information and product recommendations made by Taneja Group are based upon public information and sources and may also include personal opinions both of Taneja Group and others, all of which we believe to be accurate and reliable. However, as market conditions change and not within our control, the information and recommendations are made without warranty of any kind. All product names used and mentioned herein are the trademarks of their respective owners. Taneja Group, Inc. assumes no responsibility or liability for any damages whatsoever (including incidental, consequential or otherwise), caused by your use of, or reliance upon, the information and recommendations presented herein, nor for any inadvertent errors that may appear in this document.