Enterprise Data Protection: A Four-Step Plan

WHITE PAPER

An Interactive Buyers' Guide and Checklist

What goes into an enterprise-wide data protection (DP) plan? More than you think. There are many factors involved in choosing a solution—including the variety and location of workloads, and criticality of each application and its data. Over time, many organizations have adopted multiple backup, disaster recovery, and archival solutions, but the constant addition of new applications may leave some data at risk. This guide provides you with a step-by-step approach to ensuring all workloads—whether on-premises or in the cloud—are protected efficiently.

Step 1: Assess your organization's backup needs

How much of an enterprise's data needs to be protected? Chances are that IT and business units want every bit of the data they have collected and generated protected and available on demand. However, it comes down to economics. A good data protection plan should optimize the data protection workflow by being flexible enough to adjust individual application data protection levels according to the importance of the application while keeping the budget in mind. And with enterprise data growing 40% to 60% annually, that optimization is increasingly important.

Start by evaluating which data sources you need to protect. Look closely at:

- Databases: Traditional SQL and non-SQL databases generally hold transaction data.
- Data Lakes: May have raw IoT or other data streams before they are parsed and processed.

VERITAS™ The truth in information.

TechTarget® | Custom Media

- Legacy: Mainframes (yes, running COBOL) still power many financial and business systems.
- Big Data Analytics: Hadoop, HBase, and other scale-out systems that power advanced analytics are soon to be a $200 billion-plus market.
- Virtual Machines: Both on-premises and cloud-based VMs are the majority of today's workloads.
- Containers: Most containers are transient, but they are increasingly creating persistent data.

Enterprises utilize a broad variety of workloads today—physical, virtualized, and legacy—often managed by different teams in different locations. An ideal data protection plan should encompass all of them seamlessly.

Step 2: Evaluate your backup strategy

There are many factors to consider when formulating a data protection strategy. Here are some that should be at the top of every organization's list:

1. Application Recovery Point Objectives/ Recovery Time Objectives Assessment

Every application is important, but the organization can withstand an outage for some applications for a longer time than others without suffering potentially catastrophic consequences. Each application's criticality – how much data loss and time to recovery – should be evaluated and each application prioritized.

2. Storage Options

Once you have established your RPOs (how much data you can bear to lose) and RTOs (how long an outage you can tolerate), you can then determine where protected datasets will reside. Choices abound, and many enterprises will choose to have multiple tiers of storage depending on the criticality of the workload being protected.

- Tape: While tape has its place—very low-cost media, off-site storage—many enterprises are moving away from it due to operational challenges. Disk-based, on-premises, long-term retention and archival storage often has similar total cost of ownership and greatly improves RTO and operational efficiencies.
- Backup Appliances: Purpose-built backup appliances alleviate many of the deployment burdens of a data protection solution and can also perform deduplication in the device itself.
- Long-Term Retention and Archival Appliances: Most enterprises have regulatory constraints that mandate some kind of long-term data retention (LTR)—and often those mandates include having the data stored on-premises. Purpose-built LTR appliances have similar performance characteristics to primary storage, yet are cost-optimized to compete with the economics of tape and cloud.

- Cloud Storage: As the number of virtualized workloads hosted at cloud providers like AWS, Azure, and Google increases, enterprises are increasingly turning to cloud storage to protect those workloads. Additionally, enterprises that lack an off-site location for protecting on-premises workloads can utilize cloud storage to meet that need.

Optimizing and managing all these data protection storage methodologies across the enterprise is an ongoing challenge, which can be eased by choosing the right recovery software solution.

Step 3: Select backup and recovery solution

When it comes to backup and recovery (B&R) solutions, more is not necessarily better. Having an array of point data protection solutions can strain resources, lead to errors, and leave new workloads unprotected. The ideal B&R solution will have support for every workload, can meet every desired RPO and RTO threshold, will support every desired storage approach, and offer a single pane of glass view into every workload under protection. However, in the real world, we necessarily have to make compromises. The list below helps you select solutions and move toward these lofty requirements.

☐ 1. Support Any Workload: Since the ideal is a single B&R solution for the entire enterprise, first on the list is support for a comprehensive array of workloads. In today's hybrid IT environment, that list must support:

- Client servers of all makes and models, whether Linux, Windows, or Unix-based.
- Popular hypervisors including VMware, Hyper-V, KVM and the virtual systems they encompass.
- Databases of all kinds, including Microsoft SQL Server, Oracle, DB2, Apache Hadoop and HBase, MariaDB, MySQL, NoSQL, PostgreSQL, and SAP HANA.
- All applications—whether commercial off-the-shelf (COTS) or home-grown.
- Data and applications that reside on cloud platforms including Amazon, Microsoft, Google, and IBM.

☐ 2. Agentless Architecture: An agentless architecture is key, allowing VM-level backups at the hypervisor level, eliminating the maintenance of in-guest clients while freeing up admins to focus on something more important. An agentless architecture also simplifies protection for all modern workloads such as Hadoop, and supports both client-direct backup and continuous data protection enterprise-wide.

☐ 3. Granular Restore: Granular restore saves time and increases recovery performance. Look for the ability to quickly search and select granular VMDK objects without the need to mount and browse the entire VM.

☐ 4. Intelligent Deduplication: As enterprises go beyond petabyte scale, deduplication becomes critical. Deduplication should occur before data is sent to the target as this will save not only storage but bandwidth costs as well.

☐ 5. Snapshot Integration: B&R should seamlessly integrate with various snapshot, replication, and NDMP facilities already present in storage devices as well as cloud snapshot technology.

☐ 6. Storage Tiering: Every data element has its own specific needs for availability and performance. Tiering should keep critical workload backups on high-performance storage to maintain availability, while archival data sets can be relegated to less expensive but lower-performing LTR storage.

☐ 7. Integrated Disaster Recovery: A data protection solution should automate disaster recovery (DR) by automatically replicating backup images and catalogs, as well as hardware snapshots, to other sites or the cloud.

☐ 8. **Flexible Delivery:** Different environments call for different deployments. For example, a B&R appliance with preinstalled software and storage can be deployed in minutes. Virtual appliances for remote offices can ease installation and maintenance burdens where there is little or no IT staff. All deployment modes should support Windows or Linux and provide automated client updates.

☐ 9. **Data Insight:** Since virtually all enterprise data is protected, the data protection solution should utilize metadata and other context to automatically classify data, for example by highlighting data subject to regulations like GDPR and HIPAA.

☐ 10. **Virtualization:** VMs now account for somewhere between 75% and 90% of all business workloads, so virtualization support is critical. To be effective, B&R must support all the major virtualization platforms, ideally with an agentless architecture. Virtualization support should include:

- Instant access and recovery for critical workloads.
- Changed block tracking to reduce the amount of data backed up, with automatic rehydration of backups to enable instant recovery.
- Support for and integration of hypervisor snapshots into the backup solutions.
- The ability to build policies with flexible search functions for granular recovery.

☐ 11. **Container Support:** Since businesses are increasingly adopting container technologies, such as Docker, for applications that create persistent data, support for containers must now be included in any enterprise data protection plan.

☐ 12. **Security and Compliance:** Key standards are FIPS 140-2 for encryption, role-based access control (RBAC) to help secure backups from prying eyes, and Security Technical Implementation Guides (STIG) hardened passwords for software and appliances.

☐ 13. **Target Types:** To enable tiered B&R, support must be included for the broadest array of targets as outlined in Step 2 above.

☐ 14. **Flexible Routes to Cloud:** Specific workloads may have different requirements for deduplicating the data before sending to the target (and rehydrating at the target) or sending the data directly to the cloud so rehydration is not needed. Both routes should be available.

☐ 15. **Automation:** B&R should include orchestration and automation by including decentralized, policy-based management from a single console, and provide a library of RESTful APIs to offer integrations with third party operations like ServiceNow. Additionally, self-service that lets users perform their own backup and restores—within their allowable access—further reduces reliance on IT resources and backup admins.

Step 4: Do Your Research

As enterprises become increasingly information-driven, the need to ensure data availability increases along with the pace of data growth. Where should IT, LOB execs, and backup admins turn to learn more about the world of data protection? Beyond discussions with peers, there is no shortage of research by reputable, independent organizations to gain more knowledge. Here is a short list to consider:

- ESG: Data Protection Predictions 2019
- IDC: Cloud Data Management and Protection
- Gartner: Reviews for Data Center Backup and Recovery Solutions
- Gartner: Magic Quadrant for Enterprise Information Archiving

Take Action

Protect your enterprise from the unforeseen. Ensure recovery of business-critical data across any workload at any scale, while eliminating the cost and complexity of point products, with Veritas NetBackup.

NetBackup is the industry's No. 1 choice in data protection, with the most exabytes under management and trusted by 97% of the Fortune 100 and 86% of the Fortune 500. NetBackup protects the largest and the most demanding multi-cloud and data center environments. It delivers breakthrough capabilities for virtualized and cloud-based deployments that go well beyond what traditional backup practices can achieve, and NetBackup meets all the criteria in the above checklist—and more.

With the ability to protect any workload to any storage device, NetBackup breaks the backup window with accelerators that back up only changed blocks and parallel streaming, enabling simultaneous backup of every node in a big data cluster.

NetBackup Appliances can be deployed in minutes, and NetBackup helps ensure that every workload—bare metal or virtualized, cloud or on-premises, even those in containers — is protected and can be restored, meeting even the most stringent RPOs and RTOs.

Click here to learn how Veritas NetBackup can transform your enterprise data protection strategy.