2019 Buyer's Guide to Storage for Backup

40 POINTS OF ASSESSMENT TO USE DURING THE RESEARCH AND BUYING PROCESS

Introduction

This **Buyer's Guide to Storage for Backup** provides insight into the key features and cost drivers to consider when choosing a storage solution for backup. It looks at the ways that features offered in leading solutions help enterprises meet their key short- and long-term objectives for data protection, data recovery, and cost. The guide is designed to be used by upper mid-market to enterprise IT leaders with more than 50TB of data who are looking to implement a new solution or to improve their existing backup installation.

Contents

A storage for backup solution needs to meet several critical evaluation criteria, including the ability to provide:

- Fast and efficient backup performance to meet backup windows
- Fast VM boots and restore times to meet Recovery Time Objectives (RTOs)
- Scalable backup storage that grows as data grows
- Overall total cost of ownership up front and over time

Unlike smaller organizations that can meet their business objectives with standard disk, organizations protecting more than 50TB of data require a solution that enables more advanced capabilities, such as version retention, non-disruptive scalability, deduplication, and sophisticated technical support. For these organizations, choosing the right storage infrastructure can mean the difference between meeting RPOs and RTOs or suffering from disruptive missed backup windows, slow recoveries, and massive storage cost overages.

This backup storage buying guide is comprised of four main sections:

1. **Introduction to Backup Storage** – A brief overview of fundamental backup and deduplication methodologies used in backup storage available in the market today
2. **Functional Requirements** – Understanding the fundamental product features that are required to meet an organization's backup, recovery, replication and retention goals and service-level agreements
3. **Cost Factors** – Summarizes the factors that drive short term acquisition cost and the total cost of ownership
4. **Support and Maintenance Costs** – Discusses the support capabilities, business practices, and cost models that should be considered

There are more than 40 points of assessment in checklists provided to use during the research and buying process.

Introduction to Storage Options for Enterprise Backup

The method used in storing backup data dramatically affects the length of time it takes to back up, retrieve, and replicate data, regardless of the backup application. Exceeding backup windows can tax networks and slow end-user application responsiveness, forcing the IT organization to prematurely shut off backups as the backups run into user production hours. This results in incomplete backups with no easy way to know what is and is not restorable.

Several classes of backup storage solutions have emerged to improve backup including: closed end-to-end systems, backup applications with deduplication to standard disk, deduplication storage appliances, and intelligent hyperconverged storage for backup.

Closed End-to-End Systems
Closed solutions contain backup software, storage, and compute in a single appliance, enabling organizations to use a single vendor for their entire backup infrastructure. These systems support a homogeneous backup environment, restricting the ecosystem to the proprietary backup solution. Note that enterprises with Oracle RMAN, SQL Server direct dumps, or other backup applications and utilities may need to look outside of the closed solutions to ensure effective protection of their enterprise applications.

Backup Applications with Deduplication to Standard Disk
For organizations requiring only a few retention versions of backups and minimal selected data, standard disk and cloud storage could meet their backup storage needs. Standard disk supports heterogeneous environments and is a commodity solution; however, since the deduplication is built into the backup application, only that backup application's data can be deduplicated. If management overhead or data growth is a concern, enterprises may be better served using hyperconverged storage solutions that reduce management overhead and cost.

Deduplication Storage Appliances
Traditional deduplication storage appliances are disk-based backup storage solutions that have a front-end server controller architecture and add disk shelves as data grows. Deduplication is performed inline on data ingestion assisted by software that runs on the backup server. Storage can be added behind the front-end server, though performance is restricted to the primary controller. Due to the inline nature of the deduplication process and the front-end controller, deduplication storage appliances supply a decrease in disk consumption, with a trade-off for backup performance, scalability, and cost. A secondary concern with deduplication storage appliances is in data recovery – solutions that only store data in a deduplicated form are forced to rebuild the data prior to recovery taking place resulting in long restore times. Deduplication appliances support heterogeneous backup application environments as they are target-side appliances.

Intelligent Hyperconverged Storage for Backup
Intelligent Hyperconverged Storage for Backup solutions are disk-based backup appliances with adaptive deduplication, purpose-built for backup. Adaptive deduplication leverages a linear scaling architecture optimized for performance, scalability and price. Backup ingestion is optimized using a landing zone disk cache, which maintains the most recent backup copy in an undeduplicated format for faster restores. Deduplication and replication is performed in parallel eliminating performance issues. Recovery is performed against full data copies that have

been stored in the backup application's native format, ensuring that RTOs are met. Intelligent Hyperconverged Storage for Backup supports heterogeneous environments, and deduplication is global.

Intelligent Hyperconverged Storage for Backup is best suited to environments storing more than six versions of backups to achieve gains in storage savings through adaptive deduplication and requiring short RTOs where disk cache storage of full backups is a direct recovery time advantage.

<div style="background:#eee;padding:1em">

Storage Options Consideration Checklist

☐ Is the environment protecting more than 50TB of data?

☐ Does the environment need to store more than six versions of any backup?

☐ Are scalability and backup performance important?

</div>

Support for Backup Applications

When choosing backup storage solutions, users should identify all applications being used to protect data within the enterprise and ensure the storage solution is capable of support all of them. The backup storage solution should offer flexibility by supporting backup applications for both physical and virtual servers as well as direct database dumps. While most solutions support standard data communication protocols and will present as a NAS device, several backup applications have enabled optimized communication APIs and protocols. Storage solutions should support a variety of protocols, CIFS, NFS, Veritas OST, Oracle RMAN Channels, and the Veeam Data Mover.

Specific Integration Opportunities

While many backup applications are suited to using the backup storage options as NAS devices or writing with straight CIFS, some have implemented performance optimization APIs. Some examples include:

- **Veritas Open Storage Technology (OST) –** Veritas Open Storage is a popular feature that allows for more integrated offsite data protection, and it is an important feature if you are planning on using NetBackup or Backup Exec. Identify solutions with integration into the Veritas OST API for faster backups and for unbalanced retention onsite versus offsite. Additional integration points include supporting NetBackup Storage Lifecycle Policies and NetBackup Auto Image Replication (AIR).

- **Veeam –** If an enterprise is using Veeam, it is critical to choose a backup storage solution that support Veeam's Scale-Out Backup Repository (SOBR). Administrators can direct all jobs to a single repository in a scale-out system, automating job management.

   Solutions that are integrated with the Veeam Data Mover will increase data write performance by using the Veeam protocol, providing up to 30% faster ingest performance than writing data via CIFS. This improves all backup and restore processes and can create a synthetic full backup directly on the system for increased performance.

   Most of Veeam's unique features such as Sure Backup, Virtual Lab, Instant VM Recovery, Copy and Replicate require an undeduplicated backup copy on disk, which can only be provided by intelligent hyperconverged storage for backup solutions.

- **Oracle RMAN Support of Oracle RMAN Channels** – Oracle RMAN is a critically important way to effectively protect Oracle databases. Backup storage solutions should support performance load balancing utilizing Oracle RMAN Channels. With Oracle RMAN Channels integration, customers get built-in performance load balancing, and global deduplication across all systems. These features are available in intelligent hyperconverged storage for backup solutions.

Backup Application Consideration Checklist

☐ Is more than one method of protecting data employed (heterogeneous vs homogeneous)?

☐ Will the environment support all backup methods simultaneously?

☐ Do you require integration with one of the vendor communication APIs, such as Veritas OST, Veeam SOBR, Veeam Data Mover, Oracle RMAN Channels?

☐ Is scale and performance a consideration?

Architecture – Scale-out vs Scale-up

Data growth is inevitable. The typical business or organization is seeing data growth of 30% a year, which doubles their total data every 2.5 years. This data growth presents challenges to IT to ensure their backup system can scale easily to support that growth.

Backup storage solutions can generally be split into two types of architectures: scale-up and scale-out.
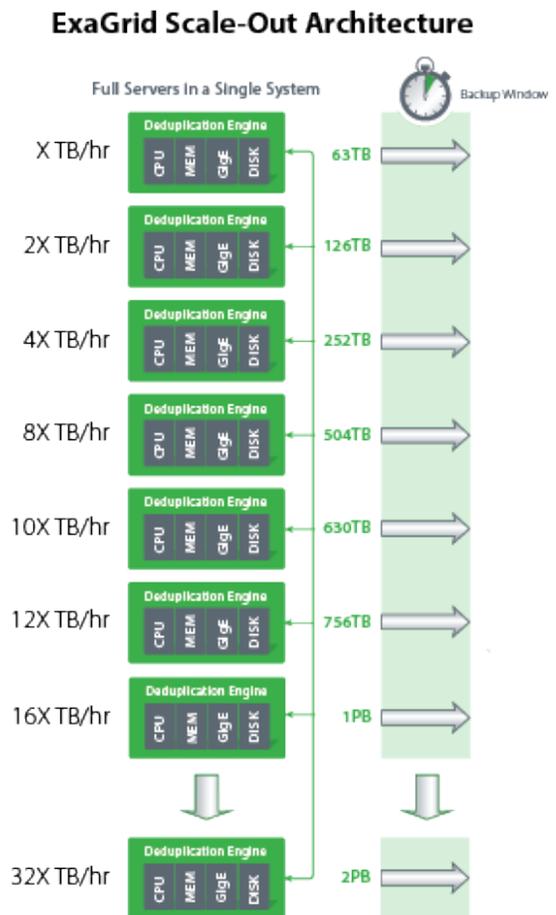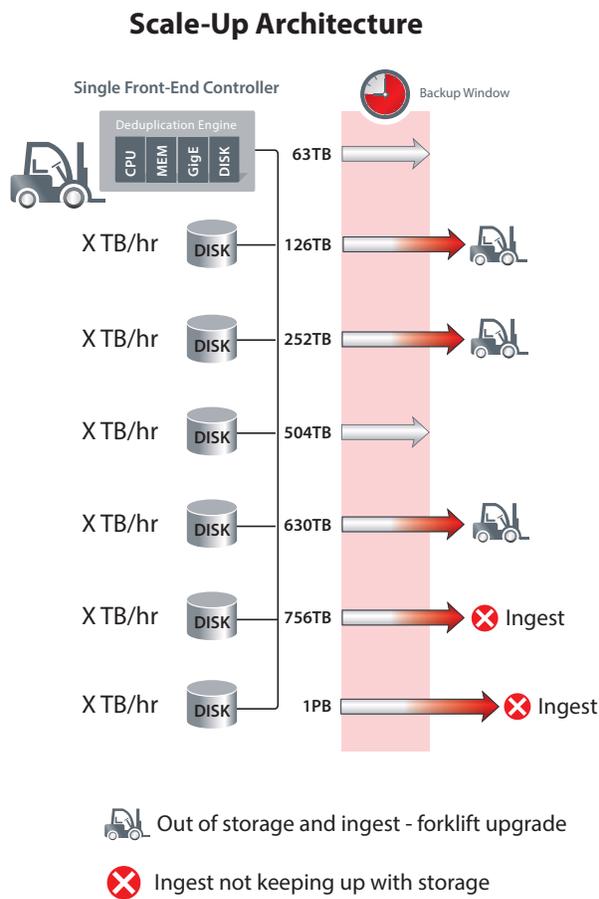
Scale-Up
A scale-up architecture is a solution with a single point of entry, often referred to as a controller. The front-end controller points to disk shelves for storage. As data grows, only disk shelves are added. At first, this is a low-cost method of protecting a larger amount of data. However, with a single front-end controller, memory and compute resources remain fixed even as throughput requirements increase. Backups take longer and longer to complete. When they exceed the acceptable backup window, a forklift upgrade is needed to replace the front-end controller with a bigger, faster, and more expensive one.

All backup applications and large-brand deduplication appliances use the scale-up approach whether in software or in a hardware appliance. With all of these solutions, as data grows, the backup window does as well. (See the "Deduplication" section for additional considerations on scale-up resource consumption.)

Scale-Out
Scale-out architectures are full appliances (processor, memory, bandwidth, and disk) in a single system that can be appended to a system in a network of backup storage. As data grows, all resources are added including additional primary storage, bandwidth, processor, and memory as well as disk capacity. This keeps the backup window fixed in length regardless of data growth, which eliminates expensive forklift upgrades.

Scale-Up Architecture

Single Front-End Controller — Backup Window

Deduplication Engine: CPU | MEM | GigE | DISK — 63TB

X TB/hr — DISK — 126TB

X TB/hr — DISK — 252TB

X TB/hr — DISK — 504TB

X TB/hr — DISK — 630TB

X TB/hr — DISK — 756TB — Ingest

X TB/hr — DISK — 1PB — Ingest

Out of storage and ingest - forklift upgrade

Ingest not keeping up with storage

ExaGrid Scale-Out Architecture

Full Servers in a Single System — Backup Window

X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 63TB

2X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 126TB

4X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 252TB

8X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 504TB

10X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 630TB

12X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 756TB

16X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 1PB

32X TB/hr — Deduplication Engine: CPU | MEM | GigE | DISK — 2PB

Unlike the scale-up approach where you need to guess which sized front-end controller is required, the scale-out approach allows you to simply add the appropriate sized appliances as your data grows. When choosing a scale-out system, verify that any size or age appliance can be mixed in a single system, which allows IT departments to buy compute and capacity as they need it. This evergreen approach also eliminates product obsolescence and forklift upgrades.

Scale Considerations Checklist

☐ Understand the data growth rate in the protected environment.

☐ Look for solutions with seamless scalability that eliminate the need for forklift upgrades.

☐ Ensure the architecture will outlast the amortization term of the hardware purchase.

☐ Consider a scale-out implementation for low up-front cost and better performance over time.

☐ Ensure the environment meets the throughput requirements.

☐ Choose a solution that does not make previous model systems obsolete – Can the environment support different ages and models of the appliances or storage to enable scale-out over time without worrying about unsupported hardware?

Deduplication: Reducing the Total Disk Consumed

When choosing a backup technology, consider the way the technology uses data deduplication. All vendors need to offer data deduplication to make the cost of disk storage comparable to tape. Deduplication is not compression. **Compression** stores data at a reduced footprint by eliminating or reducing the white space not required to hold data. **Deduplication** is a method of identifying existing data and ensuring that the information is only stored once. Essentially, all deduplication technologies work by looking at data and then doing a compare to find changes at the byte, block, or zone level. As each backup is examined, only new data and the tracking of existing data is stored. Deduplication can be done inline as data moves to storage or after data has been written to the storage media.

There are several categories of deduplication to consider:

Block (exact match) vs Zone Deduplication (similarity detection)

Traditional approaches use block-level deduplication. This method divides data into 4KB to 1MB groups of bytes called blocks. Most use fixed length block and in more advanced cases some vary the length of the block size to improve matches. This method has two challenges: the compute required for processing and the capacity required for hash tables. For every 10TB of backup data, the hash table has one billion blocks. The hash table grows so large that it needs to be housed in a single front-end controller and grows by only adding additional disk shelves. The backup time also lengthens as data volumes increase as no additional resources except disk storage are added. As described above, an expensive, disruptive forklift upgrade to a new front-end controller is required.

In contrast, hyperconverged storage for backup solutions utilize zone-level deduplication, which breaks data into larger zones and then compares it at a sub zone level called "similarity detection." The tracking table is 1,000th the size of the hash table used by the block-level approach and allows for full scale-out appliances. As data grows, the environment can also grow along with all resources: processor, memory, and bandwidth as well as disk.

In summary, block-level deduplication drives a scale-up architecture that only adds disk as data grows, which expands the backup window; zone-level deduplication allows for adding full appliances in a scale-out approach.

Inline vs Parallel Deduplication

How deduplication is implemented changes everything about backup. If not implemented correctly, deduplication will create compute problems that dramatically slow backup performance and will be the difference between whether the backup time stays fixed or lengthens as data grows.

Most backup storage vendors deduplicate inline during the backup process. Because deduplication is compute-intensive it inherently slows backups, resulting in a longer backup

Compression vs Deduplication

Data deduplication enables the cost-effective use of disk because it reduces the amount of disk required by only storing unique bytes or blocks from backup to backup. Over an average backup retention period, deduplication will use about 1/10th to 1/50th of the disk capacity, depending on the mix of data types.

Compression takes out unused or white space and can only reduce data by at most 2:1 and therefore is not enough for backup data which keeps many copies. Compression can be used in addition to – but not in place of – data deduplication.

time. Some vendors install software on the backup servers to use additional compute resources to help alleviate the deduplication bottleneck, but this steals compute from the backup environment. If you calculate the published ingest performance and compare it to the specified full backup size, you will find that products with inline deduplication cannot keep up with themselves. All deduplication implementations within backup applications are inline, and all large brand deduplication appliances use the inline approach. All of these products slow down backups, resulting in a longer backup time.

Intelligent hyperconverged storage for backup, such as ExaGrid, implements deduplication in a different way. It provides the best possible reduction of data while still maintaining the highest possible backup performance. Each appliance uses a landing zone where backups land straight to disk without any inline processing, so backups are fast. Deduplication and offsite replication occur in parallel with the backup process to meet the most stringent RPO without impeding the backup process.

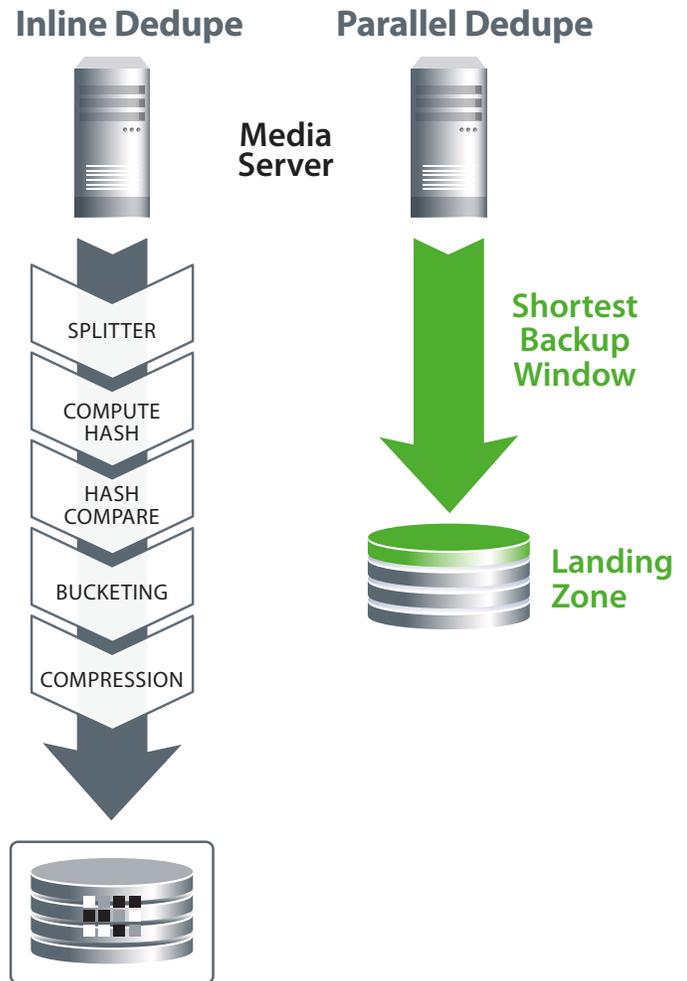**Deduplication's Impact on Restore Performance**
Although data deduplication reduces disk storage and replication bandwidth requirements, it can also create restore problems. Poorly implemented deduplication solutions increase consumption of compute resources and do not store a full copy of recent backup data. As a result, they have to reconstitute or rehydrate deduplicated data before restores can happen.

Rehydration slows restores, VM boots, and offsite copies due to the compute-intensive nature of data reconstruction. Recovery which would have taken minutes when completed from an undeduplicated format can take hours to retrieve from deduplicated form.

Since more than 95% of restores and 100% of instant VM recoveries and tape copies are needed from the most recent backup, enterprises should implement solutions that keep the most recent full backup intact and deduplicate older backups.

To avoid the problem of slow recoveries, it is necessary to retain the most recent backups in their complete, non-deduplicated form for fast restores, VM boots, and tape copies. There are two common ways to achieve this: running your own local disk pools in front of secondary storage and deduplication appliances or using an intelligent hyperconverged storage for backup solution with adaptive deduplication and a landing zone.

Adaptive deduplication with a landing zone maintains the most recent backups in the backup application's native format, undeduplicated to eliminate the lengthy data rehydration process.

Inline Dedupe

Parallel Dedupe

Media Server

SPLITTER

COMPUTE HASH

HASH COMPARE

BUCKETING

COMPRESSION

Shortest Backup Window

Landing Zone

Rehydration

*Storing only dedupli-
cated data on disk
requires that the
backup application
reassemble the data
prior to the data
being recoverable.
This is referred to
as rehydration. This
introduces a significant
delay in recovery times.*

*Most environments
need VM boot times of
single-digit minutes;
however, with a pool
of deduplicated data,
a VM boot can take
hours due to the time
it takes to rehydrate
the data.*

*All deduplication in
backup applications,
as well as the large-
brand deduplication
appliances, store only
deduplicated data. All
of these products are
very slow for restores,
offsite tape copies,
and VM boots.*

*Consider an
intelligent backup
appliance that includes
a landing zone,
like ExaGrid, which
eliminates the rehydra-
tion requirement.*

The result is that data retrieval is completed in minutes versus hours, often 20 times faster data restores than any other solution, including those that deduplicate within backup applications or target-side deduplication appliances. A repository of all deduplicated data sits behind the non-deduplicated data in the landing zone for cost-efficient storage of long-term retention.

Deduplication Considerations Checklist

☐ Does the environment require deduplication or is the retention low?

☐ Will inline deduplication impact the backup performance?

☐ Do you require a landing zone or disk pool of undeduplicated data to meet RTOs?

Restoration: Meeting the Primary Goal of Backup

The most critical aspect of any enterprise backup implementation is the ability to recover data and systems in the timeframe required to keep your business up and running. RTOs reflect the maximum length of time you are willing to be down for any given system outage or data access delay. It is unlikely that all systems and data in your organization are subject to the same RTO – not all recoveries are of equal importance.

Enterprises must consider RTOs for each tier of data type and system – ranging from mission critical systems often requiring near-zero RTOs to passive systems, such as archival data, where RTOs can be days or weeks. When segmenting or tiering recovery requirements, organizations should look at backup storage solutions that deliver what is needed to meet RTOs:

- Instant VM Recoveries
- Large database restores (e.g. Oracle RMAN or SQL databases)
- Latest data version recoveries
- Using storage as a short-term host for data /VMs on the appliance
- Recovering recent versions of files
- Recovering older or archived versions of files

Why is recovery a critical decision point when choosing backup storage?
The method in which backup data is stored dramatically impacts the length of time it takes to retrieve and restore data. Especially when deduplication exists.

Recovery Checklist

☐ Determine whether the storage meets the RTOs for your data types:

    ☐ Instant VM recoveries

    ☐ Large database recoveries (e.g. Oracle RMAN or SQL databases)

    ☐ Latest data version recoveries

☐ Using storage as a short-term host for data /VMs on the appliance

☐ Recovering recent versions of files

☐ Recovering older or archived versions of files

☐ Recovering from a site disaster from the offsite DR data. Assess the deduplicated data risk: do the recovery needs of the environment require a non-deduplicated data version on accessible disk landing zone or primary storage pool?

☐ Enable tiered storage – Choose a solution that will automatically enable long-term data retention at a lower cost, with slower recovery rate than more critical copies of data.

Data Backup, Backup Window and RPO Considerations

Among the primary considerations in assessing backup storage capabilities is how fast you need to complete your backups (backup window) and how frequently you need to backup your data to meet your Recovery Point Objectives (RPO).

Ingestion Rate
Standard disk solutions have ingestion rates that are well documented with data written via standard protocols. The majority of enterprises are running with multiple 10g Ethernet, eliminating network limitations. The backup performance bottleneck is typically linked to the ability of the backup storage to ingest and process the backups.

Deduplication Impact on Ingestion
As noted in the deduplication section, it is critical to choose a backup storage solution that meets both the short- and long-term performance needs, the data reduction needs, and the scalability needs of the environment.

Load Balancing
Load balancing is the distribution of ingestion and storage over resources to optimize the resource utilization. The ideal backup storage solution will enable load balancing across compute and network, enabling fast intake, fast restorations, and fully utilized back end storage.

Security Considerations: Encryption of Data at Rest and in Flight
All backup storage solutions should include fundamental security protection. All systems must encrypt data in flight across the WAN and at rest, protect against ransomware, and provide ongoing security updates. Data should also be continually checked using MD5 checksum technology.

Enterprises should consider using enterprise-class Self-Encrypting Drive (SED) technology that ensures all data on the disk drive is encrypted automatically without any action required by users. Note that unlike software-based encryption methods, SEDs typically have a better throughput rate, particularly during extensive read operations.

Backup Checklist

☐ No lengthening of backup time as data grows – Look for solutions that will add full appliances, for consistently fast backup performance and a fixed-length backup window as data increases.

☐ Choose a system that scales to meet the RPOs for your data types.

☐ Ensure processing the data backups will not put your backup window at risk.

☐ Understand whether you need to add backup server compute power to achieve your quoted performance numbers.

☐ Consider whether the system utilizes backup storage in a method that enables full use of capacity. Does the backup storage require manual intervention to handle an imperfect balance of data backup distribution?

☐ Review data at flight and at rest encryption protocols.

DR, Second Site, and Cloud Support

Regulatory requirements, and best practices, require that data is stored in more than one location. Organizations must determine if the data replication solution built into the backup application is adequate or if they should take advantage of backup storage replication which may reduce the overall replication time and bandwidth consumption. Cloud support should also be assessed within the environment and considered during the purchase process.

Deduplication features built into hyperconverged solutions only replicate changed data, dramatically reducing WAN network utilization and costs. Systems that use Adaptive Deduplication deduplicate backups and replicate the data as backups are coming into the system in parallel with the backup window. Parallel processing within the backup window typically provides the best RPO. Consider the following architectures:

Hub and Spoke
Hub-and-spoke configurations enable you to configure storage at multiple data centers so that one data center can be the disaster recovery site for another. This can be use in an active-active configuration where both local and remote storage solutions are in use for active backups. Look for a solution that allows for the second site to be at your own data center, a co-location facility, or a managed service provider.

Single Direction Replication
For one-way replication, you can configure one data center as an active site, and the second data center as a passive location. Some deduplication solutions require the same sized system onsite and offsite which increases overall cost. Consider backup storage solutions that allow you to configure "asymmetrical storage" in which the capacity of offsite appliances is allocated to 100% repository, giving you a disaster recovery site that is half the cost of the primary backup site. This is available in hyperconverged storage.

Public and Private Cloud Configuration
Many storage can replicate data to the cloud. This allows the disaster recovery and long term
retention data to be stored in locations such as a location owned and operated by a third-party
DR provider, AWS, or Microsoft Azure.

/

DR, Second Site, and Cloud Checklist

☐ Choose a system that scales to meet the RPOs for your data types.

☐ Support for deduplicated backup replication – Look for solutions that will optimize the
data transferred across the network.

☐ Parallel processing – Ensure that the offsite RPO can be met and data can be replicated in
the timeframe required. Adaptive deduplication solutions can parallel process backups
with deduplication and replication to provide the best RPO.

☐ Find a solution that supports your data center architecture – Determine whether your
needs are best met by a hub-and-spoke solution or an active-passive configuration.
Organizations without a secondary data center will need to consider cloud.

☐ Consider asymmetrical storage – Active-passive configurations, especially those with
hosted secondary storage locations, should consider solutions that do not require full-size
storage for the second site.

☐ Cloud considerations – Choose a solution that will meet both the short- and long-term
cloud vision.

Supportability

There are several challenges to supporting a backup storage solution, including quick and
easy installation of the backup storage, the ability to easily monitor and resolve issues, and the
availability of competent vendor support when required. Some considerations for assessing
supportability include:

Solution must be fully manageable within a single, easy-to-use GUI
Implementation and long-term management of a backup storage solution should not increase
the burden on an enterprise IT department. The solution should employ a simple-to-use GUI
and present few reasons for ongoing active management.

**Email notification of service alerts to external support personnel for proactive support;
send alerts and alarms to leading SNMP consoles**
Most enterprise IT organizations have systems in place for data center monitoring and issue
management. Users should look for a solution that fits into their existing enterprise structure.

**Proactive monitoring (will you know when you need a new drive, vs automated /included
in support?)**
Some appliances and hyperconverged solutions also supply proactive monitoring with support.
Proactive monitoring will identify hardware issues prior to system disruption and supply replacement

parts for free. Look for support contracts that include shipment of failed hardware components via next business day air in most countries and within two days in countries with customs.

Redundant vs Resilient

Enterprises should determine the level of redundancy required within the backup storage solutions and the overall cost of implementation. Some backup storage solutions enable full node redundancy, an entire appliance can be lost and data can be recovered from other appliances. This adds to the overall cost of the infrastructure and may be necessary for enterprises without an offsite copy of backup data.

For most enterprises utilizing a secondary data storage location, backup storage solutions that employ RAID6 plus a hot spare will be sufficient. This implementation would protect against two simultaneous disk failures in order to ensure that your backup data is always available for restores. This method of data protection redundancy provides:

- System operation, even with a simultaneous disk drive failure
- The ability to automatically rebuild the RAID set to the included spare drive, if a drive failure occurs
- Hot swappable drives; if a drive fails, the system still runs; replace the drive at any time while the system is live in production

Vendor Considerations

Customer satisfaction ratings are an important evaluation criteria for your backup storage vendor. Before choosing a vendor, review the organization's Net Promoter Score® to gain a reasonable overview of the overall satisfaction of existing customers.

Enterprises with highly skilled IT employees should look for vendors that include a premier support model that assigns Level 2 support techs your account to provide high level support without the frustration of Level 1 triage.

Supportability Checklist

☐ Simple management – single GUI

☐ Alerts and integration with existing monitoring solutions

☐ Proactive monitoring with free hardware replacement

☐ Resiliency – can the hardware withstand basic failures?

☐ Is the vendor known for supporting their customers?

Calculating the Cost

Initial purchase price on backup storage is only a small part of the ownership cost.

The Cost of Storage

There are the fundamental storage cost factors that need to be determined to choose the right technology. Costs are driven by exactly how much storage will be required in the short term and

in the long run. This depends on the amount of data to be protected, the data deduplication ratios, and the processing resources that will be required to run the environment.

Past growth can sometimes be used as an indicator of future growth. Most environments are growing at a steady rate of 30% year-over-year which means their data will double every 2.5 years. At 30% yearly growth, 100TB of data today will be 400TB in just 5 years. Look for a solution that can keep up with your ingest rates on the backup storage by either purchasing a front-end controller sized for what your data volumes will be in three years, or by purchasing a hyperconverged scale-out system that will grow as needed to the total size required over time. Users should also look for deduplication with data reduction ranging from 10:1 to as high as 50:1.

Be wary of solutions that require forklift upgrades of the entire backup storage infrastructure if you exceed your growth projections.

Product Lifespan
Beyond storage requirements, enterprises should ensure the storage solution they choose will not be obsolete before the amortization period has ended. Solutions with fixed media servers and controllers need to be refreshed with the latest processor and memory. Find out whether the vendor has recently discontinued older models and replaced them with new ones. Hyperconverged backup storage solutions typically support the use of a mixture of older models with newer models and allow you to combine appliances of various capacities in a single scale-out system.

Cost of Maintenance and Support
Calculate the historical cost of maintenance for the solution you choose. Some vendors charge a reasonable initial support fee, but dramatically increase that fee – sometimes with a double-digit increase in that fee for ongoing maintenance and support on the discontinued model. Look for vendors with a history of keeping maintenance and support cost increases to a few percent per year.

Price Protection
Many vendors offer a low entry price for their product and/or support, and will dramatically raise their prices for subsequent purchases. Look for vendors who employ a price protection program that they put in writing. Some examples of this type of program include:
- The price you pay up front for the product is the price you'll pay for that product over the subsequent five years.
- Your initial maintenance and support cost will not increase more than a few percent per year for the subsequent five years.

Calculating the Cost Checklist

☐ Determine the cost of storage for the next three years – Forklift upgrades and controller challenges should be weighed against the cost of a scale-out hyperconverged system.

☐ Product lifespan – Estimated lifecycle of the backup storage solution. Will the solution besupported in the long-term and will that impact the cost of support and maintenance?

☐ Price protection – Ensure the cost of the solution and support is predictable over the next few years.

Checklist Summary

Storage Options Consideration

- ☐ Is the environment protecting more than 50TB of data?
- ☐ Does the environment need to store more than six versions of any backup?
- ☐ Are scalability, recovery performance, and backup performance important?

Backup Application Considerations

- ☐ Are more than one method of protecting data employed (heterogeneous vs homogeneous)?
- ☐ Will the environment support all backup methods simultaneously?
- ☐ Do you require integration with one of the vendor communication APIs, such as Veritas OST, Veeam SOBR, Veeam Data Mover, Oracle RMAN Channels?
- ☐ Are scalability and performance considerations?

Scale Considerations

- ☐ Understand the data growth rate in the protected environment.
- ☐ Look for solutions with seamless scalability that eliminate the need for forklift upgrades.
- ☐ Ensure the architecture will outlast the amortization term of the hardware purchase.
- ☐ Consider a scale-out implementation for low up-front cost and better performance over time.
- ☐ Ensure the environment meets the throughput requirements.
- ☐ Choose a backup storage solution that does not make previous model systems obsolete – Can the environment support models of different ages and capacities to ensure future scalability without worrying about unsupported hardware?

Deduplication Considerations

- ☐ Does the environment require deduplication or is the retention low?
- ☐ Will inline deduplication impact the backup performance?
- ☐ Do you require a landing zone or disk pool of undeduplicated data to meet RTOs?

Recovery

- ☐ Determine whether the storage meets the RTOs for your data types:
    - ☐ Instant VM recoveries
    - ☐ Large database recoveries (e.g., Oracle RMAN, SQL databases)
    - ☐ Latest data version recoveries
    - ☐ Using storage as a short-term host for data / VMs on the appliance
    - ☐ Recovering recent versions of files
    - ☐ Recovering older or archived versions of files
    - ☐ Recovering from a site disaster using the offsite DR data
- ☐ Assess the deduplicated data risk – Do the recovery needs of the environment require a non-deduplicated data version on accessible disk landing zone or primary storage pool?
- ☐ Enable tiered storage – Choose a solution that will automatically enable long-term data retention at a lower cost, with slower recovery rate than more critical copies of data.

Backup

- ☐ No lengthening of backup time as data grows – Look for solutions that will add full appliances, for consistently fast backup performance and a fixed-length backup window as data increases.
- ☐ Choose a system that scales to meet the RPOs for your data types.
- ☐ Ensure processing the data backups will not expand your backup window.
- ☐ Understand whether you need to add backup server compute power to achieve your quoted performance numbers.
- ☐ Consider whether the system utilizes backup storage in a method that enables full use of capacity – Does the backup storage require manual intervention to handle an imperfect balance of data backup distribution?
- ☐ Review data-in-flight and data-at-rest encryption protocols.

DR, Second Site, and Cloud

- ☐ Support for deduplicated backup replication – Look for solutions that will optimize the data transferred across the network.
- ☐ Parallel processing – Ensure that the offsite RPO can be met and data can be replicated in the timeframe required. Adaptive deduplication solutions can parallel process backups with deduplication and replication to provide the best RPO.
- ☐ Find a solution that supports your data center architecture – Determine whether your needs best fit a hub-and-spoke solution or an active-passive configuration. Organizations without a secondary data center will need to consider cloud.
- ☐ Consider asymmetrical storage – Active-passive configurations, especially those with hosted secondary storage locations, should consider solutions that do not require full-size storage for the second site.
- ☐ Cloud considerations – Choose a solution that will meet both the short- and long-term cloud vision.

Supportability

- ☐ Look for solutions that provide:
    - ☐ Simple management – single GUI
    - ☐ Alerts and alarms compatible with existing monitoring solutions
    - ☐ Proactive monitoring with free hardware replacement
- ☐ Resiliency – Can the hardware withstand basic failures?
- ☐ Is the vendor known for supporting their customers?

Calculating the Cost

- ☐ Determine the cost of storage for the next 3 years – Forklift upgrades and controller challenges should be weighed against the cost of a scale-out hyperconverged system.
- ☐ Product lifespan – Estimated lifecycle of the backup storage solution. Will the solution be supported in the long-term, and will that impact the cost of support and maintenance?
- ☐ Price protection – Ensure the cost of the solution and support is predictable over the next few years.

Conclusion

The method in which backup data is stored dramatically affects the cost of storing, replicating, and recovering it. Organizations need to evaluate a number of options, including standard disk backup, deduplication appliances, and intelligent hyperconverged storage in order to choose the most appropriate solution for their environment.

For organizations that need to meet stringent recovery time and recovery point objectives, hyperconverged solutions may offer the most cost-effective method of delivering fast, scalable backup, recovery, and replication.

For a demonstration of the power of **Intelligent Hyperconverged Storage for Backup**, contact ExaGrid now.

REQUEST A DEMO NOW
INTELLIGENT HYPERCONVERGED STORAGE FOR BACKUP

EXAGRID®

United States: 350 Campus Drive | Marlborough, MA 01752 | (800) 868-6985
United Kingdom: 200 Brook Drive | Green Park, Reading, Berkshire RG2 6UB | +44 (0) 1189 497 051
Singapore: 1 Raffles Place, #20-61 | One Raffles Place Tower 2 | 048616 | +65 6808 5574