



# Data Deduplication with Veritas NetBackup Appliances.

# Contents

---

INTRODUCTION . . . . .	3
EXECUTIVE SUMMARY . . . . .	3
SCOPE . . . . .	3
TARGET AUDIENCE . . . . .	3
SOLUTION VALUE . . . . .	3
WHAT IS DEDUPLICATION? . . . . .	3
TYPES OF DEDUPLICATION . . . . .	4
Post-Process Deduplication . . . . .	4
Inline Deduplication . . . . .	4
Factors that Affect Deduplication . . . . .	5
DEDUPLICATION FOR BACKUP AND DISASTER RECOVERY . . . . .	6
DEDUPLICATION WITH VIRTUALIZATION ENVIRONMENTS . . . . .	6
SECURE, SCALABLE MSP WITH NETBACKUP APPLIANCES . . . . .	8
EASY TO UPGRADE A NETBACKUP APPLIANCE MEDIA SERVER . . . . .	9
MSP AVAILABILITY ARCHITECTURE FOR THE ENTERPRISE . . . . .	9
Software Architecture . . . . .	9
Veritas Critical System Protection . . . . .	10
Self-Healing Dedupe Architecture Safeguarded by NetBackup . . . . .	10
Veritas Appliance Autosupport and Call-Home Monitoring . . . . .	10
Veritas InfoScale Maximizes Storage Efficiency, Availability and Performance . . . . .	10
HARDWARE ARCHITECTURE: BUILT-IN REDUNDANCY AND HIGH AVAILABILITY OPTIONS . . . . .	11
Built-In Redundancy and Hot Swap . . . . .	11
High Availability . . . . .	11
Fast Ingest and Recovery Rates . . . . .	12
REPLICATING DEDUPLICATED DATA . . . . .	12
DEDUPLICATION TO THE CLOUD WITH CLOUD CATALYST . . . . .	13
MSP WITH VERITAS FLEX APPLIANCE . . . . .	14
MSP WITH NETBACKUP VIRTUAL APPLIANCES . . . . .	14
CONCLUSION . . . . .	14
REFERENCES . . . . .	14

## INTRODUCTION

### EXECUTIVE SUMMARY

Veritas Technologies has been a leader in backup and recovery solutions that focus on protection and management of companies' digital assets critical for their success and business continuity.

Veritas builds on this with advanced data deduplication technology built into every NetBackup Appliance media server. NetBackup Media Server Deduplication Pool (MSDP) is the result from years of engineering and over 80 patents specifically on data deduplication. MSDP enables deduplication to the cloud with compression and encryption for shorter backup times and fast data recovery at scale.

It should be no surprise that Gartner has once again named Veritas as the Market Share Leader in Backup and Recovery Software, Archiving Software and Management Software Defined Storage by Revenue.

With its innovative and industry leading products, Veritas has been able to maintain 15 years of Gartner *Magic Quadrant* leadership.

- Gartner: #1 Archive, Backup and Recovery Market Share
- Gartner: #1 Software Defined Storage Management Market Share
- IDC: #1 Integrated Backup Appliance Market Share

### SCOPE

The purpose of this document is to provide technical details and assist in understanding the value of the Veritas Deduplication technology that is present in NetBackup Appliances.

### TARGET AUDIENCE

This document was written for customers, partners, decision makers and Veritas field personnel interested in learning more about Veritas' Data Deduplication Appliance Solutions. It provides a technical overview of the technology and the unique values of its features.

### SOLUTION VALUE

As Datacenters have continued their transition into virtualized, cloud based and cost optimized footprints, storage has become one of the more challenging areas to optimize. Data is increasing at an accelerating pace and companies are constantly challenged to find the most cost-effective way to securely protect their data across multiple types of infrastructures -physical, virtual and in the cloud.

To address this challenge, Veritas has designed one of the most flexible and effective software defined deduplication engines on the market. It presents a workload agnostic, hardware independent deduplication solution with compression and encryption that can be run on a NetBackup Appliance, NetBackup Virtual Appliances on-premises or in the cloud as well as on Build Your Own ("BYO") hardware.

## WHAT IS DEDUPLICATION?

At the basic level, deduplication is the elimination of non-unique data segments within a data set. On one level, deduplication is not a whole lot different from compression. However, the real distinction of deduplication is that there is data reduction against historical data, allowing storage savings indexed against previously written data from a multitude of sources. Before deduplication, compression was the primary storage savings. When backing up to a tape or disk, a backup solution could compress the data stream as the data was written. However, the compression savings is only at that point in time. This means that the same data backed up at another point in time would compress, but the compressed data would take up the same amount of space as the previous data set. With deduplication, the data is segmented and checked against a matrix representing all previously written data. Only segment data that is found to be unique will be sent on for storage. Non-unique segment data will create a reference to the unique instance of the same segment data.

## All deduplication products have three primary phases or tasks:

Fingerprinting | Redundancy Identification | Redundancy Elimination

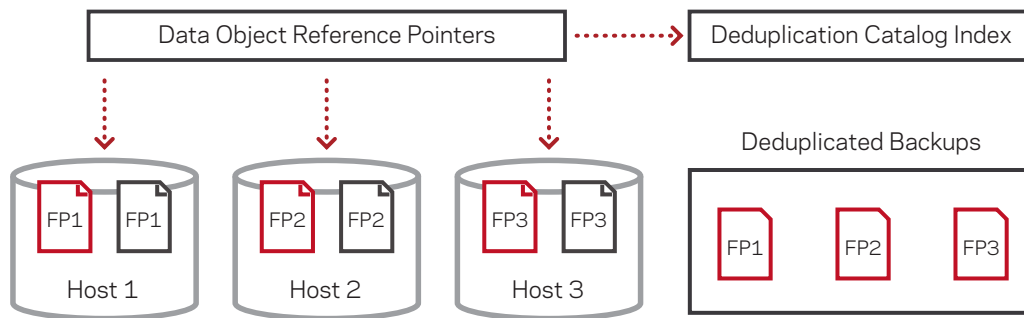


Figure 1: Data Deduplication Tasks

## TYPES OF DEDUPLICATION

Fundamentally, it is not hard to make a deduplication engine. However, it is very hard to make a deduplication solution that is performance optimized, fault tolerant and at the same time scalable. Where and how deduplication can be performed makes a big difference in the quality of service.

There are three major types of deduplication:

- Post-process
- Inline at target
- Inline at source

## POST-PROCESS DEDUPLICATION

Post-process deduplication is the least efficient form of deduplication. This method requires a large disk cache to temporarily store a full set of data plus another disk cache to store deduplicated data. The deduplication process is not applied until after successfully writing the data to the target where it is processed using a post-process deduplication method and then stored in a deduplication repository. This might help get the data off the source without having to worry about processing time but is an inefficient use of space and can be prone to data integrity issues. Veritas deduplication does not do post process deduplication. By using post processing deduplication, the available backup window time may need to be reduced to allow enough time for this post process to complete.

## INLINE DEDUPLICATION

Inline Deduplication means that the deduplication process applied to the data stream prior to writing to the storage. This ensures that only unique segments of data are written to storage. **Inline deduplication at the target level** means that all data is streamed to the target device and then deduplicated as it is written to storage. **Inline deduplication at the source level** means that the data is deduplicated against previously written data prior to being sent to the target device. Source side deduplication is the most efficient from a data transport perspective since there is a dramatic reduction of the amount of data that needs to be sent across the network. Veritas deduplication can perform both source and target inline deduplication as well as compression.

## FACTORS THAT AFFECT DEDUPLICATION

When planning for a deduplication deployment, care should be taken to consider how well the data being protected will deduplicate. There are different types of data that can get different levels of deduplication based on the makeup of the data. Image files, compressed data, virtual images, NDMP streams, and encrypted data do not deduplicate very well. In addition, databases with a high rate of change can require some extra effort to ensure that the data is presented in a manner that will give optimal deduplication results. Separate policies can be implemented within NetBackup for each data type based on how well the data can be deduplicated.

Veritas has two different methods that are designed to improve deduplication:

1. Fixed length segmentation with stream handlers
2. Adaptive Variable length segmentation

MSDP uses intelligent stream handlers. Stream handlers employ Veritas technology to use knowledge of the type of data format to optimize the stream for deduplication. Stream handlers are data aware and adaptive. They help improve backup performance and storage efficiency based on the type of data ingested. This turns the data stream into something that will get consistently good deduplication rates at high speed using fixed length segmentation. Stream handlers will get engaged in standard filesystem backups as well as VMware, Hyper-V, NetApp, EMC NDMP and other snapshot-based solutions, like FlashBackup.



Figure 2: Fixed Length with Stream Handler

Adaptive Variable Length (VLD) segmentation was introduced in NetBackup to deliver optimal deduplication results when a stream handler can't be employed. VLD uses a defined segment size range to find the optimal segmentation for the data being deduplicated. This will get the best deduplication of opaque data while utilizing the CPU a little more than fixed length segmentation. While testing has shown that VLD doesn't have a significant impact to the backup time, it still is better to use a stream handler when they are available.

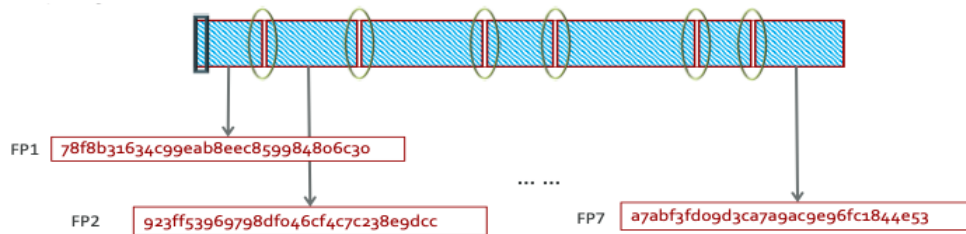


Figure 3: Variable Length Deduplication

NetBackup, NetBackup Appliances and NetBackup Virtual Appliances can create a deduplication pool that is not limited to shelf boundaries and does not restrict the disk shelf from being used for other storage. MSDP also provides the flexibility to choose between fixed-length, variable length or no deduplication on the same media server.

Many applications today are using encryption at rest and the industry trends for security are driving this at a rapid pace. If using encryption at rest, deduplication rates are more like 2:1. With NetBackup, there is no requirement to dedicate storage shelves for deduplicated data storage; those workloads can be directed to a non-deduplicated storage pool on the Media Server, saving 100%-200% for the storage cost.

When comparing vendor data deduplication rates, something to consider is how the deduplication rates are calculated. Other solutions may report erroneously high deduplication rates by counting each backup as a full backup. This will exaggerate the amount of data deduplication received.

## DEDUPLICATION FOR BACKUP AND DISASTER RECOVERY

Deduplication technology can present significant savings when utilized in the backup infrastructure. Backup images taken over time will have a high chance of encountering duplicated data. Consider a series of clients where there are similar data sets or documents being worked on by multiple parties. Often there can be wholly or partially duplicated data sets across multiple systems. Pair this with scenarios and regulations that have multi-year data retention requirements and the amount of data to store can be staggering. Historically, this retention has been regulated to tape storage as the only cost-effective solution. The cost of keeping all of that data that may or may not be needed for years or decades on a disk storage location was a staggering problem. Tape certainly keeps costs lower than a massive set of arrays. At the same time, tape is not ideal. The media can take up a lot of physical space. There is datacenter footprint of specialized tape management hardware. Long-term data shipping and warehousing and logistical difficulties come into play to get the tapes to where the data is needed. This add time to an emergency restore situation that can have a critical impact on a businesses' operational ability and total cost of ownership.

To deliver a well-rounded data protection solution Veritas went beyond the orchestration and management of operations to delivering an integrated and powerful data deduplication storage engine. The integration between MSDP and NetBackup allows for a complete solution in a single converged application. The deduplicated data format now becomes highly portable with new possibilities. The deduplicated data format facilitates the replication of data across multiple location and a diverse number of targets. There is also a dramatic reduction of recovery time compared to recalling tape shipments or copying whole data back from a DR location. Other features also leverage the deduplication technology. Veritas was able to deliver Accelerator to optimize reads before deduplication for even better performance. Veritas also has recently delivered an Instant Recovery and Universal Shares, which also leverage MSDP.

NetBackup Universal Shares provide direct access to deduplicated storage via NFS and CIFS protocols. This makes it easy to protect open source databases that do not provide a backup API but do support dump or export operations.

NetBackup provides faster backup and recovery with direct data paths.

- Client Direct: Up to 50% faster and up to 98% reduction in network bandwidth
- Application Direct: Up to 50% faster and up to 98% reduction in network bandwidth
- Storage Direct: Up to 20% faster backup and 10% faster recovery
- Improved recovery performance with compression.

All NetBackup clients support client-side deduplication.

MSDP does not have limits on the number of streams coming in and unlike other deduplication solutions, will not refuse connections.

## DEDUPLICATION WITH VIRTUALIZATION ENVIRONMENTS

Virtualization solutions have opened new doors in the data center that have both simplified things and added a new type of complexity. A common core infrastructure is shared by any number of virtual entities. This can lead to VM sprawl; thousands of hosts that share a common template or data sets while having their own unique elements. Being able to protect all those points in time while maintaining the guest systems independence from each other could lead to having to store a massive amount of historical data. Again, deduplication is the key that makes protecting all this data possible. NetBackup MSDP delivers archival quality, historical protection of VM data while providing instant operational recovery as well as disaster recovery. With NetBackup appliances, customers can leverage NetBackup Universal Share along with MSDP to get Instant Access to individual files from VMs or complete secondary copies of virtual machines to leverage for replication, testing or several other uses. With the enablement of Accelerator technology, there are reduced backup windows. That means that there is faster snapshot reconciliation with lower VM quiesce time. NetBackup is the only backup product that gives backup administrators the option to exclude the data contained within the paging and swap files of the guest operating systems. With less data to backup, deduplicate and compress in the first place, there is less to deduplicate and compress.

## Space reclamation

Unlike simple storage targets, like a filesystem or tape, what the user sees as written to deduplication storage is a logical representation of what was sent to storage. The actual data is stored in unique segments of data in specialized container structures. The logical file information uses reference pointers to keep track of what unique segments of data would be needed to reassemble the actual file if needed. Since the result is multiple files written to storage, many of which share segment objects, there is a logistical challenge when it comes to deleting data.

## Writing is easier than removing

With many deduplication solutions, the immediate concern is with the process of deduplicating the data, but what about cleaning up data segments that aren't needed anymore? Freeing up space is a problematic endeavor. Many solutions struggle with identifying and releasing space. This is because they have no good way of understanding what the current state of the segment data is. This means that there are resource intensive and time-consuming periods of garbage collection or filesystem cleaning activity that are required to scan and find space that needs to be freed.

To remove bottlenecks that can be associated with the cleaning process used by other deduplication products, MSDP includes a patented process called rebase that runs for minutes every day. This efficient process makes it possible to deduplicate data in the cloud as well.

## How much free space is needed?

One important thing to note – with deduplication storage, there isn't a one to one reference for the segment data. The value of deduplication is that with multiple files referring to shared segments of data there is a storage savings. If new segment data comes in, then it is written independently. This means that most of the heavy lifting is done when writing new data. As a result, it is ideal to avoid having to write segment data as new if it has already been stored before. As a result, deduplication storage only frees up space it needs to use. Consider the example of a document that was backed up on day 1, was not there on day 2, but a slightly modified version of the document was backed up on day 3. If there was a 1:1 cleanup of space when the original document was backed up, the storage would have gone through the bother of freeing up that segment data only to have in ingest 98% of the same data on day 3. If all of this was going on while there was ample free storage, what is the point? Deduplication services benefit from persisting the segment data if there is a reasonable threshold of available space.

## What about when there isn't free space

Many deduplication solutions are designed for backups first and cleanup is another matter. In a critical scenario where a system needed to free space, this could result in hours and days of downtime depending on the size of the system. These deduplication solutions don't have the awareness to know what is going on with the stored segment data. They need to be able to scan the stored data on the system, often in multiple passes, before they can even begin to process the actual freeing of space. On some systems, this can take hours or even days to complete check and free space. Some vendors may tell a tale of efficiency gains in not spending time to free space in a timely manner with a weekly cleanup pass instead. At Veritas, efficiency is valued at all levels and it is recognized that sometimes space reclamation needs to be a priority without impacting overall system performance. To do this Veritas employs a reference tracking system that allows insight into the status of segments and the containers they reside in. This results in efficient space reclamation that runs in minutes to hours instead of hours to days. Veritas Deduplication will not remove space aggressively unless directed to. The reference architecture is normally "trued up" every 12 hours with what is referred to as queue processing. As the queue processing runs, it can optimize space reclamation by identifying entire containers that are no longer needed and removing them outright. Outside of the regularly scheduled queue processing, there is a compaction process. Compaction is a performance optimized space reclamation process that focuses on freeing segment data in containers with a mix of expired and non-expired data. Compaction is constantly waiting to run when the appropriate thresholds are hit. When active, compaction targets containers identified by the reference architecture that will deliver the best results with a minimal impact to system performance. The frequency, scale and impact of compaction can be configured to meet the desired behavior. Both queue processing and compaction are automatic processes, but they can easily be initiated at any time.

## Beyond “just” Deduplication

It is evident that deduplication in a data protection solution dramatically reduces costs. Veritas goes beyond “just” deduplication. The compatibility of the MSDP data format with other targets delivers significant solution flexibility. Customers are able to reduce footprints to disk only deduplicated storage across the entire data lifecycle. With no more disk/tape mixed environments, it is easier to distribute load across multiple targets. Traditional replication is costly, requiring solutions to move an entire data set between two points and constantly copy all new changes between locations. With MSDP and Veritas OpenStorage Technology (OST), in domain Optimized Duplication (OptDup) and Automated Image Replication (AIR) between domains uses duplication technology for the data transport. With this, inline deduplication methods are applied between the source and target systems. The source system handles fingerprinting on its end. This reduces load on the target systems and guarantees only the unique segment data needs to leave the source location. When replication is to a different domain with AIR, Additional technology allows for the target master server to have full image recovery capabilities without having to analyze the replicated image. This creates **highly available backup data**, meaning that customers can save money with data redundancy instead of managing expensive replication and failover solutions.

## Data Encryption

To keep data protected, data encryption can be used so that it is not readable without first being decrypted. NetBackup provides encryption for the deduplicated data. It is separate from and different than NetBackup policy-based encryption. With client-side deduplication, the data can be encrypted before it is sent to the media server so that is secure as it is sent across the network.

## Tape

While some vendors may offer deduplication to tape, Veritas provides deduplication to the cloud instead using Cloud Catalyst Appliances. The problem with deduplication to tape is that if one tape in the backup set is lost, it can render the entire backup set useless. Many customers are leveraging the cloud instead as an efficient replacement for tape.

## **SECURE, SCALABLE MSDP WITH NETBACKUP APPLIANCES**

NetBackup Appliances and Virtual Appliances enable organizations to deploy NetBackup MSDP services in a way that is secure, flexible, scalable and easy manage.

A single NetBackup Appliance can scale up 960TB of deduplicated data, the current maximum that has been tested. And a NetBackup Virtual Appliance can support up to 250TB of deduplicated data. Each NetBackup Media Server Appliance can host a mix of deduplicated and non deduplicated data.

Rather than running multiple virtual machines with different operating systems to provide a solution, NetBackup Appliances run a single, secure operating system. Running multiple virtual machines with multiple operating systems under a hypervisor is potentially less secure because it increases the potential attack surface for malicious activity.

NetBackup Appliances provide security and intrusion detection through Symantec Data Center Security (SDCS) and through Role Based Access Controls. NetBackup Appliances also include FIPS 140-2 validation that can be enabled at no additional cost. It is important to keep in mind that enabling FIPS-140-2 is a one-way choice.

For more information on the security that is built into every NetBackup Appliance, please see the Veritas white paper titled: [Ransomware Protection with NetBackup Appliances](#)

NetBackup Appliances provide the rapid recovery speeds that are required to restore at scale. NetBackup Appliances can support many concurrent recoveries and do so without limitations and without additional requirements such as SSD.



Not all data deduplication appliances are created equal. High Availability (HA) is available as an option on the NetBackup 5340 appliance. NetBackup Appliances use active-active clustering on a two node, fully redundant cluster running on tried and true Veritas InfoScale High Availability Cluster. With a HA, jobs are automatically load balanced providing up to 80% faster backups and faster deduplication fingerprinting. Existing appliances can be upgraded in place via a 30-minute guided setup.

NetBackup Appliance Engineering has a fully staffed team of performance experts whose sole job is to test and validate the performance of all NetBackup Appliance versions. No Appliance is released without this performance validation.

### **EASY TO UPGRADE A NETBACKUP APPLIANCE MEDIA SERVER**

The benefits of using MSDP on a NetBackup Appliance include making upgrading the media servers easy. An appliance upgrade includes built in safeguards to ensure success, including a set of pre-flight checks as well as an option to roll back to the previous version in case of an unsuccessful upgrade. When upgrading earlier versions of NetBackup that did not support variable deduplication are upgraded to a newer version that does support variable rate deduplication, existing data goes through a one-time data conversion.

When it comes time to upgrade the appliance hardware itself, the NetBackup licenses are fully transferable to another NetBackup Appliance. There is no need to purchase additional licensing when upgrading the NetBackup Appliance hardware.

### **MSDP AVAILABILITY ARCHITECTURE FOR THE ENTERPRISE**

For years, Veritas has been architecting invulnerability features into the NetBackup software to make it the enterprise-class solution that it is today. This architecture has been extended even further with NetBackup Appliances to create a family of robust backup appliances that include industry-leading data deduplication technology. NetBackup Appliances are designed with a software and hardware availability architecture that provide resiliency for both the systems and the data that they contain.

### **SOFTWARE ARCHITECTURE**

The fingerprint database is an integral part of a backup's data deduplication metadata and without it, recovery will not be possible. For this reason, it is critical that it be fully protected. For fingerprinting, Veritas uses SHA-2 hashing. The SHA-2 family of algorithms are FIPS-compliant; SHA-1 is not.

With the MSDP Availability Architecture, the MSDP catalog is fully protected using off-host backups through NetBackup. NetBackup Appliances include a catalog backup policy that is automatically configured on the appliance to protect the fingerprint database.

With some other deduplication solutions, a second appliance is needed to protect the fingerprint database. Having off-host backups means that backup and disaster recovery are simpler and that a second appliance at significant extra cost is not needed to serve as a replication target.

The fingerprint database is also more easily recovered in the event of a disaster. Some solutions can require a complete wipe and restore of the database to recover, which can take significantly longer to do.

During backup operations, NetBackup checks the integrity of the MSDP catalog and constantly checks the integrity of the data being backed up at many points during the backup and recovery processes.

The built-in data integrity features are as follows:

- Two Stage Catalog and Fingerprint Database Protection is built into every NetBackup Appliance.
  - Provides real-time catalog shadow copies of the MSDP catalog
  - Two-stage commit to ensure that the MSDP catalog is transactionally-consistent
  - In case of MSDP catalog corruption, failover to a shadow copy is automatic and with zero down time.
- Crash-consistent Off Host backup of catalog
  - Can be enabled on NetBackup Appliances
  - Policies are automatically configured (and deactivated by default)
  - Automatic MSDP catalog backup as part of initial appliance deployment
- Container check after every backup
- CRC checks of all segments are performed during a restore
- Metadata is protected with multiple copies for recovery
- All records are check summed from end to end
- Online Checking
  - Reference Integrity Checking
  - CRC Checking
  - Storage Leak Checking
  - Automatically Repairs Problems in the Background

The more backups that are performed, the more consistent the data becomes. As the system scales and the more data that it contains, it becomes more consistent and more reliable.

#### **VERITAS CRITICAL SYSTEM PROTECTION**

- Monitors changes and behaviors to reduce operational downtimes and meet critical security compliance.
- Protects against malicious behaviors, file system changes and known and unknown attacks by utilizing host intrusion software (SDCS) based monitoring, notification and auditing.

#### **SELF-HEALING DEDUPE ARCHITECTURE SAFEGUARDED BY NETBACKUP**

- Self-healing deduplication architecture originally introduced in NetBackup 7.6 and has been enhanced and extended in subsequent releases.
- End-to-end verification of backup data with compromised disk isolation and fault avoidance technology. (The more backups that are created, the more verification is performed)

#### **VERITAS APPLIANCE AUTOSUPPORT AND CALL-HOME MONITORING**

- Faster problem resolution leads to higher availability
- Less unplanned downtime leads to higher availability and better performance overall

#### **VERITAS INFOSCALE MAXIMIZES STORAGE EFFICIENCY, AVAILABILITY AND PERFORMANCE**

- Dynamic Multi-Path technology (DMP) provides storage path failure protection from the controller to the primary storage.
- Storage pools are distributed across multiple RAID sets to limit overall system degradation in the event of disk failure.
- Enterprise-proven technology built to improve storage Quality of Service (QoS) and application availability

## **HARDWARE ARCHITECTURE: BUILT-IN REDUNDANCY AND HIGH AVAILABILITY OPTIONS**

Veritas combines industry leading NetBackup data protection and storage software together with field-proven, highly reliable hardware to create data protection and storage appliances that bring high levels of performance and five nines (99.999%) of uptime to the enterprise.

### **BUILT-IN REDUNDANCY AND HOT SWAP**

Each NetBackup Appliance has the following field serviceable and hot swappable components:

- Raid Controllers
- Hard Disk Drives
- Power Supply Units
- Cooling Modules
- Expansion I/O modules

Redundancy is built in to each appliance with the following components:

- Dual Redundant FC Cards
  - Provides 4 FC connections to external RAID controllers
  - Spreads IO across the 4 redundant data paths to improve IO performance and availability
- Dual Redundant External hot-swap RAID Controllers
  - Dual load sharing RAID controllers
  - NVRAM protected write cache
- Field-Proven Storage Architecture
  - Dual path SAS to risk from RAID controllers
  - Intelligent storage layout to address all single point of failure scenarios

### **HIGH AVAILABILITY**

InfoScale is built into every NetBackup Appliance and makes HA (High Availability) availability as an option with the NetBackup 5340 Appliance. When used, HA provides up to 80% increased backup throughput and a 10x higher operational uptime for backup and recovery operations at a significantly lower operational cost. With the increase in storage capacity and performance of a high-end NetBackup Appliance, downtime on one such appliance will have a bigger impact on the overall backup and recovery operations. While taking care of the availability of the operations in case of failures, NetBackup HA cluster consisting of two nodes with shared storage provides significant benefits in planned downtime events as well as software upgrades via features like rolling upgrade.

When HA is enabled on a NetBackup Appliance, the following benefits are provided as an “active-active” cluster:

- Two active NetBackup compute nodes configured as A/A (Active/Active) in a HA configuration
- Storage workload is shared efficiently between two nodes
- All stored data is available and accessible through either node
- All backup jobs can be run on one node to facilitate upgrades or repairs on the partner node
- Automatic failover and job retry for a single node failure

## FAST INGEST AND RECOVERY RATES

With fast ingest rates, NetBackup Appliances are able protect and deduplicate data at scale. In the event of disaster, NetBackup Appliances can be recovered relatively quickly. Any inconsistencies are resolved, and the reference database is rebuilt in the background.

### Maximum Through with 98% Deduplication Rate:

- Distributed Client Deduplication: 197TB/Hr
- Server-Side Deduplication (single node): 37TB/Hr
- Server-Side Deduplication (HA): 58TB/Hr

### NetBackup 5340 Recovery:

- Up to 50 Concurrent Instant Access restores

## REPLICATING DEDUPLICATED DATA

Replicating deduplicated data is fast and efficient with NetBackup Appliances. Whether replicating to other NetBackup Appliances, NetBackup BYO (Build Your Own) domains or to the cloud, only the incremental data changes are replicated. This reduces the amount of data sent across the network as well as the amount of data stored on the target. NetBackup Appliances include different network options for data replication, which further extend their flexibility.

When replicating data from one MSDP pool to another within the same NetBackup domain, deduplication is maintained because only the delta changes are replicated through the OST interface. This is known as Optimized Duplication or *OptDup*.

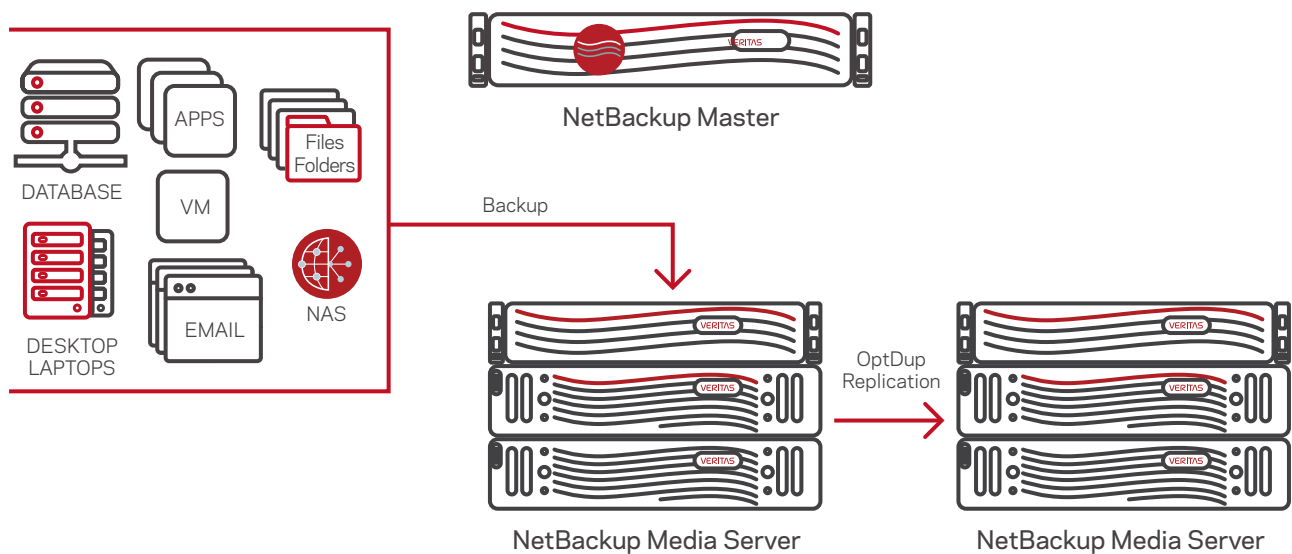


Figure 4: Replication with Optimized Duplication (OptDup)

When replicating data from one MSDP pool to another across NetBackup domains, deduplication is also maintained because only the delta changes are replicated in this scenario as well. This is done using NetBackup AIR (Auto Image Replication).

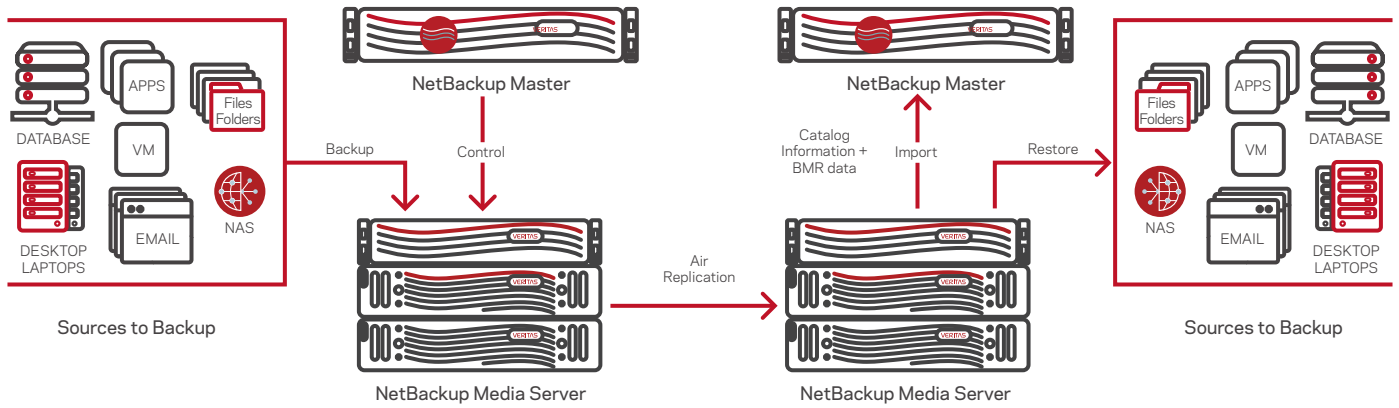


Figure 5: Replication with Auto Image Replication (AIR)

Efficient, high performance replication is extended through the NetBackup Appliance’s networking options. Replication is available across both 10GbE and Fibre Channel.

### DEDUPLICATION TO THE CLOUD WITH CLOUD CATALYST

The best way to reduce storage costs in the cloud is to reduce the amount of data that is sent to the cloud. This can be a challenge, especially for enterprises and large enterprises. The ideal approach is to use deduplication technologies that reduce the amount of data sent to the cloud in the first place. NetBackup Cloud Catalyst is a software defined option that enables NetBackup Media Servers to target the cloud directly with deduplicated, and not consume space on the media servers.

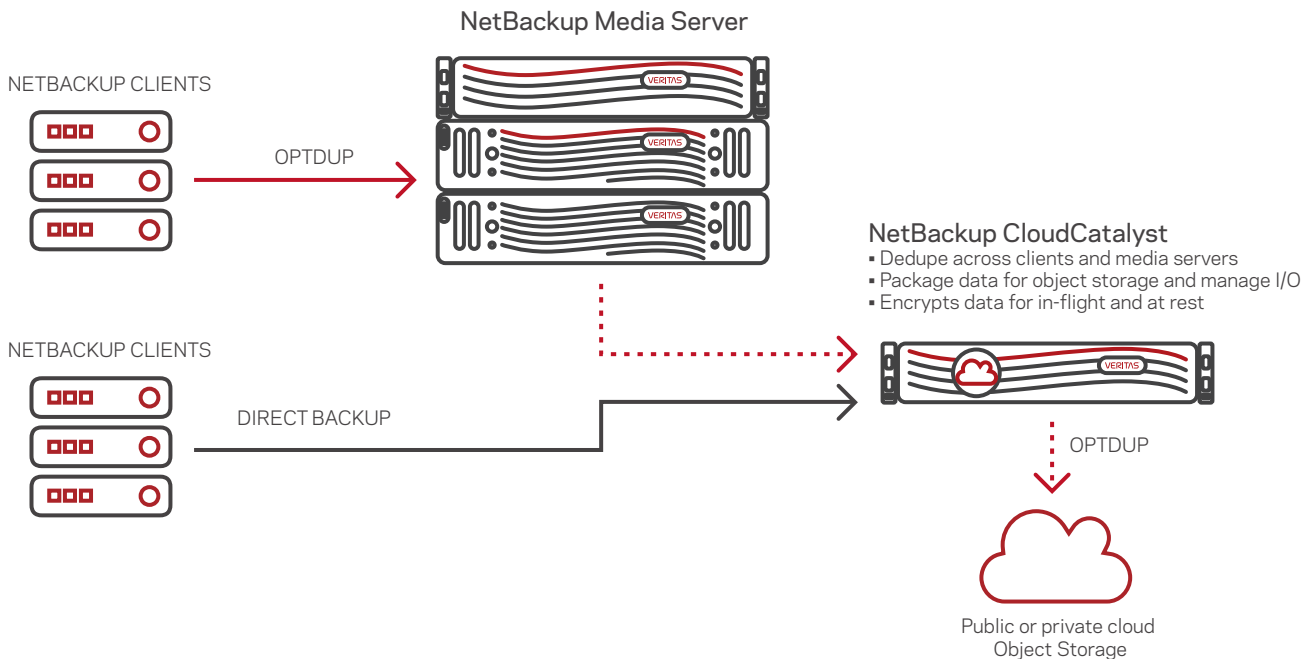


Figure 6: Data Deduplication to the Cloud with Cloud Catalyst

Both disk space and network bandwidth are saved by sending deduplicated data directly to the cloud in deduplicated format without going through a hydration-dehydration process. Once data is stored in the format, it becomes highly portable and is optimized for transport to any compatible target on any infrastructure. Cloud Catalyst includes support for over 60 cloud providers and does not require a storage shelf. Cloud Catalyst can be deployed on a NetBackup Appliance, Veritas Flex Appliance, NetBackup Virtual Appliance or built on your own hardware.

## MSDP WITH VERITAS FLEX APPLIANCE

The Veritas Flex Appliance also runs on a single, secure operating system. With a Flex Appliance, multiple NBU domains with discrete MSDP pools can be created and provided to multiple customers or tenants.

For further information, please see Veritas white paper: [Veritas Flex Appliance: Using Flex to Enable NetBackup Data Protection Services](#).

## MSDP WITH NETBACKUP VIRTUAL APPLIANCES

NetBackup Virtual Appliances deliver unified data protection and seamless hypervisor integration into existing virtual environments. Built on industry leading NetBackup software, these appliances provide deduplication with reduced network bandwidth utilization like their physical counterparts.

NetBackup Virtual Appliances are designed to be easily integrated into existing virtual environments and are supported with several popular versions of VMware.

- vSphere 6.5
- vSphere 6.0
- vSphere 5.5

There are four versions of the NetBackup Virtual Appliance available. These VMs can be easily deployed into existing virtual environments and the media servers can scale from 0.50TB to 250TB of deduplicated, on-premises storage.

Virtual Appliance	Use Case	Capacity
Master + Media Server	Remote Office / Back office (ROBO)	0.50TB to 16TB
Master Server	Data Center	N/A
Media Server	Data Center	0.50TB to 250TB
CloudCatalyst	Data Center	4TB local cache + 1PB cloud

For high availability, NetBackup Virtual Appliances can also be run on fully built VMware ESX clusters.

## CONCLUSION

Veritas NetBackup Appliances combine industry-leading data protection and deduplication technology with data encryption and compression in a single data protection appliance that is performant, secure and scalable. NetBackup Appliances with MSDP enable significant savings through a minimized data and backup footprints as well as optimized data transfer rates. NetBackup physical and virtual Appliances extend MSDP services to virtual environments as well as the cloud.

## REFERENCES

High Performance and Availability Solution for NetBackup Appliances :  
[https://www.veritas.com/content/support/en\\_US/doc/125171886-128639504-1](https://www.veritas.com/content/support/en_US/doc/125171886-128639504-1)

Ransomware Protection with NetBackup Appliances:  
[https://www.veritas.com/content/dam/Veritas/docs/white-papers/V0827\\_GA\\_ENT\\_WP\\_Ransomware-Protection-with-NBU\\_EN.pdf](https://www.veritas.com/content/dam/Veritas/docs/white-papers/V0827_GA_ENT_WP_Ransomware-Protection-with-NBU_EN.pdf)

Veritas Flex Appliance: Using Flex to enable Data Protection Services:  
[https://www.veritas.com/content/dam/Veritas/docs/white-papers/flex\\_appliance\\_technical\\_wp.pdf](https://www.veritas.com/content/dam/Veritas/docs/white-papers/flex_appliance_technical_wp.pdf)

## DISCLAIMER

THIS PUBLICATION IS PROVIDED “AS IS” AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID. VERITAS TECHNOLOGIES LLC SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS PUBLICATION. THE INFORMATION CONTAINED HEREIN IS SUBJECT TO CHANGE WITHOUT NOTICE.

No part of the contents of this book may be reproduced or transmitted in any form or by any means without the written permission of the publisher.

---

## ABOUT VERITAS

Veritas Technologies is a global leader in data protection and availability. Over 50,000 enterprises—including 87 percent of the Fortune Global 500—rely on us to abstract IT complexity and simplify data management. The Veritas Enterprise Data Services Platform automates the protection and orchestrates the recovery of data everywhere it lives, ensures 24/7 availability of business-critical applications, and provides enterprises with the insights they need to comply with evolving data regulations. With a reputation for reliability at scale and a deployment model to fit any need, Veritas Enterprise Data Services Platform supports more than 800 different data sources, over 100 different operating systems, more than 1,400 storage targets, and more than 60 different cloud platforms. Learn more at [www.veritas.com](http://www.veritas.com). Follow us on Twitter at [@veritastechllc](https://twitter.com/veritastechllc).

---

2625 Augustine Drive, Santa Clara, CA 95054  
+1 (866) 837 4827  
[www.veritas.com](http://www.veritas.com)

For specific country offices and contact numbers, please visit our website.  
[www.veritas.com/company/contact](http://www.veritas.com/company/contact)

**VERITAS™**