VERITAS™

Data Deduplication with Veritas NetBackup Appliances

White Paper | October 2022

Contents

Introduction

Executive Summary

As data centers continue their transition into virtualized, cloud-based, cost-optimized footprints, storage has become more challenging to optimize. As data increases at a rapid pace, companies are challenged to find the most cost-effective way to secure their data across multiple types of infrastructures—physical, virtual, and in the cloud. Pair this with scenarios and regulations that have multi-year data retention requirements for backup and archiving, and the amount of data to store can be staggering.

Deduplication technology can combat this challenge and present significant savings when utilized in the backup infrastructure. Backups are an ideal candidate for deduplication due to the redundant nature of backup data. NetBackup provides deduplication that allows you to:

- Deduplicate data everywhere, as close to the source of data as you require
- Use encryption to protect your data during your remote client backups

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 run on NetBackup Flex and NetBackup Flex Scale appliances, in the cloud, or on build-your-own (BYO) hardware. With NetBackup Appliances, you can:

- Deploy NetBackup data protection in minutes
- Protect against cyberattacks with multi-layer security
- Reduce total cost of ownership (TCO) with optimized integration

Combined, you can reduce cost and increase backup and restore speed by:

- Minimizing the amount of data being stored, thereby decreasing your infrastructure and licensing costs.
- Reducing the amount of network backup bandwidth
- Decreasing your recovery point objective (RPO) with faster and more frequent backups

Veritas Technologies is a leader in backup and recovery solutions that focus on protection and management of companies' critical digital assets for their success and business continuity. Gartner has 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 maintained 17 years of Gartner Magic Quadrant leadership. NetBackup Media Server Deduplication Pool (MSDP) is the result of years of engineering and more than 80 patents specifically on data deduplication. MSDP also deduplicates data stored in the cloud. Because deduplication eliminates redundant data and compresses the data in flight and on disk, it reduces backup times, decreases bandwidth requirements, and enables fast data recovery at scale.

Scope

The purpose of this document is to provide technical details and assist in understanding the value of the deduplication technology that is present in NetBackup Flex and NetBackup Flex Scale appliances.

What is Deduplication?

At the basic level, deduplication is the elimination of non-unique data segments within a data set. In some ways, deduplication is not a 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 unique segments will be sent on for storage. Non-unique segments will simply reference previously stored segments containing the same data.

Fingerprinting | Redundancy Identification | Redundancy Elimination

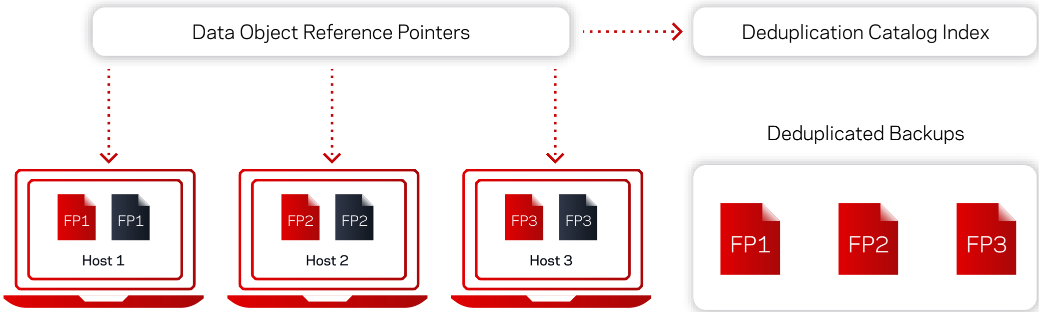

Figure 1. Data deduplication tasks

Deduplication Types and Options

Client-side deduplication or Client Direct is an easy way to improve the performance of your backups to an MSDP target. Part of the MSDP deduplication architecture is the use of a distributed plug-in-based fingerprinting service. Instead of moving all the data to the storage server before it is deduplicated, the fingerprinting, compression, and encryption can be performed on the source. This leads to ideal optimization and acceleration, with minimal network overhead.

With NetBackup, you can choose point in the backup process to perform deduplication.

There are three ways to use Deduplication Location options for MSDP:

- Always use the media server—All data is sent to the media server and the plug-in deduplication occurs on that server before the MSDP storage target is written to.
- Prefer to use client-side deduplication—At the beginning of a backup, a quick test is performed to verify that the client can successfully use client-side deduplication. If the test fails, the job falls back on the use of server-side deduplication.
- Always use client-side deduplication—The backup job explicitly uses client-side deduplication. If the functionality does not work, the job fails.

Whether deduplication is performed on the server side or the client side, the same plug-in library is loaded, and the savings are the same.

If your backup client has sufficient system resources, Veritas recommends using client-side deduplication or Client Direct as an easy way to optimize the performance of your backups to an MSDP target, and to minimize network overhead. Instead of moving all the data to the storage server before its deduplication, fingerprinting, compression, and encryption can all be performed right on the source.

	Media Server Deduplication—Server Side	NetBackup Client Direct Deduplication—Source Side
Deduplication Location	All data is streamed to the target device and then deduplicated as it is written to storage	Client agent identifies duplicate segments on data source and only sends unique segments to target device
Benefit	▪ Load balance capability ▪ Minimizes usage of client CPU	▪ Reduces network traffic ▪ Moves deduplication processing load from the storage server to client server

Figure 2. Comparison of client-side and source-side deduplication

Deduplication Methods

Different types of data get different levels of deduplication based on the data format. MSDP uses intelligent stream handlers that employ Veritas technology to determine the type of data format necessary 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. Since we understand the data stream and the header, we are able to make the fix length extremely efficient.

Veritas offers two types of stream handlers, designed to improve deduplication:

1. Fixed-length deduplication (FLD)
2. Adaptive variable-length deduplication (VLD)

Veritas provides the flexibility to choose between fixed-length, variable-length, or no deduplication on the same media server. By default, the variable-length deduplication is disabled on a NetBackup client.

NetBackup fixed-length deduplication method chunks data streams into fixed-length segments (128 KB) and then processes them for deduplication.

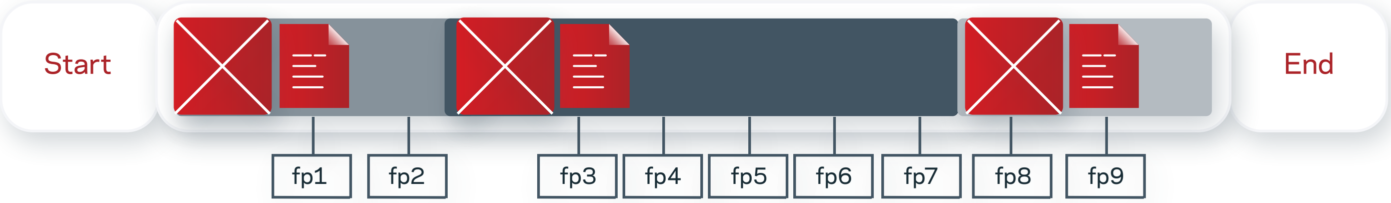

Figure 3. Fixed-length with stream handler

Adaptive variable-length deduplication 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 on the backup time, it still is better to use a stream handler when available.

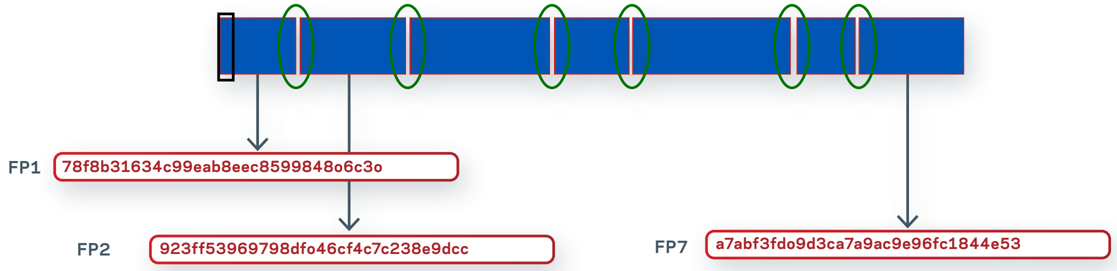
FP1
78f8b31634c99eab8eec8599848o6c3o
FP2
923ff53969798dfo46cf4c7c238e9dcc
FP7
a7abf3fdo9d3ca7a9ac9e96fc1844e53

Figure 4: Variable-length deduplication

	Fixed-Length Deduplication	Variable-Length Deduplication
Data Stream Chunk	The data streams are chunked into fixed-length segments (128 KB)	▪ Uses a defined segment size range to find the optimal segmentation for the data being deduplicated ▪ Every segment has a variable size with configurable size boundaries
Use Case	File backup	Optimal deduplication results when a stream handler can't be employed
Advantage(s)	▪ Swift method ▪ Consumes fewer computing resources	▪ Reduces backup storage ▪ Improves backup performance
Disadvantage(s)	Low deduplication ratios (in some cases)	▪ CPU performance may be affected ▪ Does not work with compression

Figure 5. Comparison of fixed-length deduplication and variable-length deduplication

* VLD can and should be used for workloads where an FLD stream handler does not exist.

MSDP Software Architecture

MSDP Storage Server has five main components: deduplication plug-in, multi-threaded agent, deduplication engine, deduplication manager, and media server deduplication pool.

- The deduplication plug-in is the data interface to the NetBackup deduplication engine on the storage server. It separates the file's metadata from the file's content, divides the client data files into data segments, calculates fingerprints for each data segment, and deduplicates them. The deduplication plug-in can run on the direct client, media server, or load balancing server.
- The multi-threaded agent works inline between the deduplication plug-in and the deduplication engine. It uses multiple threads for asynchronous network I/O and CPU core calculations to enhance deduplication performance for client-side deduplication and media server deduplication. The resources that are used are adjusted dynamically, based upon the hardware characteristics of the system it is running on. Like the deduplication plug-in, the multi-threaded agent runs on deduplication media servers, load balancing media servers, and NetBackup clients that deduplicate their own data.
- The deduplication engine stores and manages deduplicated file data.
- The deduplication manager maintains the configuration, controls internal processes, optimizes duplication, and manages security and event escalation. It also maintains a database that is referred to as the MSDP catalog.
- The media server deduplication pool defines a storage path—the directory in which NetBackup stores the raw deduplicated backup data as unique segment objects. The segment objects are stored in data containers, each having a unique data container ID (DCID). The Media Server deduplication pool also uses a database path—the directory that houses the reference database.

To optimize deduplication performance, Veritas recommends that you use separate disks, volumes, partitions, or spindles for the storage path and the database path of the MSDP. This is automatically configured on NetBackup Appliances.

Fingerprinting Process (FP)

NetBackup uses a unique identifier for each file and file segment that is backed up. The deduplication plug-in reads the backup image and separates the image into files. The plug-in separates the files into segments. For each segment, the plug-in calculates the hash key (or fingerprint) that identifies each data segment. To create a hash, every byte of data in the segment is read and added to the hash. NetBackup uses SHA-2 (SHA256) for the hash algorithm. The chunk segment size unit used to compute fingerprints is by default a fixed-length of 128 KB, or configurable to a variable-length size based on the chunk boundary. You have an option to encrypt deduplicated data. Compression and encryption (if enabled) are performed after the fingerprint is calculated, prior to sending data to the target storage. NetBackup 9.1 added support for immutable object storage for the deduplicated data, to ensure backup data cannot be tampered with.

The fingerprinting is done by the deduplication plug-in, which can be on the media server (server side dedupe) or on the source side (direct deduplication client).

The storage server maintains an index cache of the fingerprints in RAM. For each backup, a client requests a list of the fingerprints from its last backup from the storage server. Initial lookups are performed against the local cache. If the fingerprint is not located in the local cache, then a second lookup is performed against the global FP cache held on the storage server.

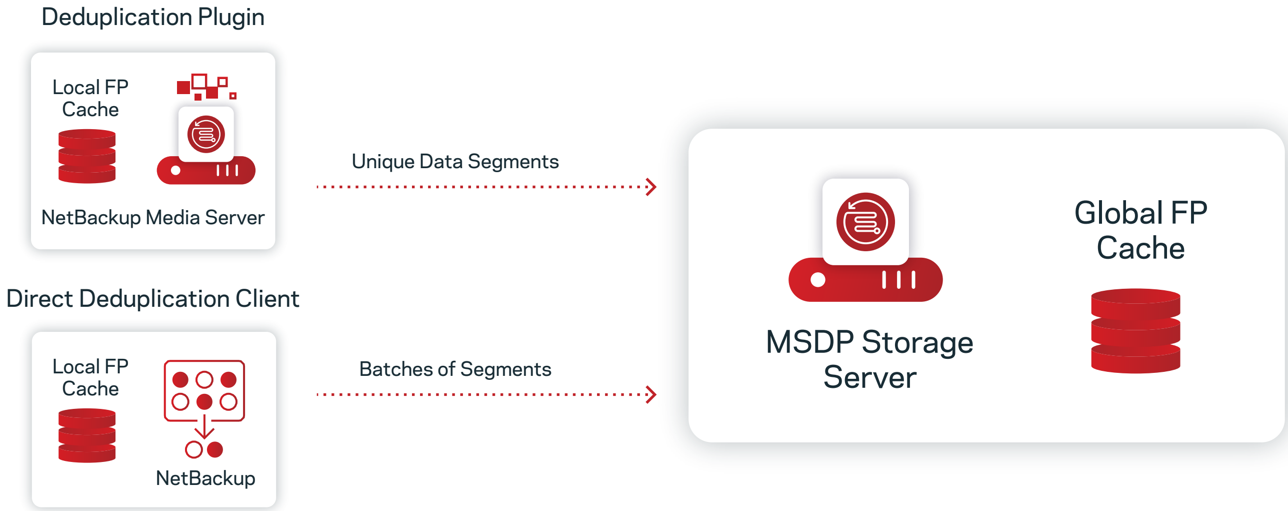

Figure 6. The fingerprinting deduplication plug-in

Positioning the backup stream on file boundaries also ensures that data can be deduplicated with files backed up from other clients as well. The deduplication plug-in does the following tasks:

- Identifies the metadata and file boundaries to separate the image into files
- For fixed-length deduplication, aligns the 128 KB segments for lookup
- Fingerprints each data segment by calculating hash keys based on the secure SHA-2 algorithm and look up in the FP cache
- Performs redundancy elimination

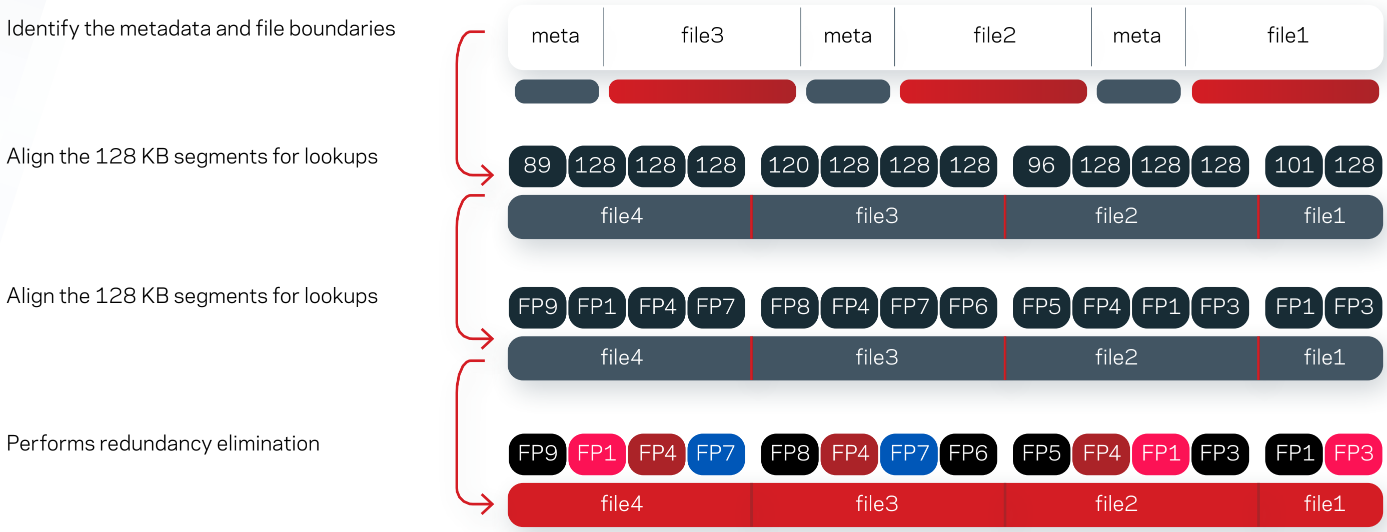

Figure 7. MSDP Deduplication Process

To increase overall performance, the NetBackup deduplication multi-threaded agent uses multiple threads for asynchronous network I/O and CPU core calculations. The agent runs on the storage server, load balancing servers, and clients that deduplicate their own data.

NetBackup uses a distributed, plug-in-based fingerprinting service to identify files and file segments.

MSDP Catalog

The MSDP catalog links NetBackup backup images with deduplicated client data in MSDP storage. The MSDP catalog is stored in a separate path from the MSDP backup data. NetBackup Flex Scale protects the catalog using triple mirroring, and can tolerate the loss of one full node and one disk failure. Flex Appliance uses RAID 6 to protect the catalog, and can tolerate two disk failures.

Beyond hardware protection such as mirroring, NetBackup provides additional levels of MSDP catalog protection to increase availability:

- Daily shadow—NetBackup automatically creates copies of the MSDP catalog
- Snapshot copies of the catalog—Veritas provides a utility that you can use to configure a snapshot schedule of the MSDP catalog
- Catalog backup policy—Veritas provides a utility that you can use to configure a NetBackup policy that backs up the MSDP catalog.
- Data on disk is self-describing and can be imported without the catalog

The ability to retain the metadata describing deduplication means that you always have a method for quickly retrieving your data in a its original form.

NetBackup Flex Scale automatically configures a catalog backup policy and snapshot copy policy as part of the initial deployment.

NetBackup Appliances include a catalog backup policy that is automatically configured on the appliance to protect the metadata.

Data Integrity

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

- All data records are periodically checked and verified in the background
- Multiple copies of metadata for extra redundancy and auto-rebuild when needed
- Network resiliency assures backup and replication
- Enables recovery from whatever is left, in the event of a complete system failure

Veritas Appliance data integrity features:

- Two-stage catalog and fingerprint database protection
    - 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
- 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 have an end-to-end checksum
- Online checking
    - Reference integrity checking
    - CRC checking
    - Storage leak checking
    - Reference DB automatic repair if detected as corrupted

Deduplication Space Management

MSDP was designed to store data quickly and efficiently with a high level of accuracy. In order to maintain performance, certain operations such as deletion may be deferred in order to reduce overhead. Unlike simple storage targets such as a file system or tape, what the user sees as written to deduplication storage is a logical representation of what was sent to storage. The actual data stored is the unique segments of data in specialized container structures. The logical file information uses reference pointers to keep track of what unique segments of data are needed to reassemble the actual file if necessary. 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? 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 file system cleaning activity that are required to scan and find space that needs to be freed.

Reclaim the Space with Deduplication

When you delete a file in deduplicated storage, the reference to the file gets removed but not the data within the MSDP storage pool. 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.

Note: Rebasing is not required on NetBackup Flex Scale, which uses block-based erasure coding that does not require garbage collection to delete blocks that no longer have pointers referencing it.

How Much Free Space is Needed?

NetBackup reserves four percent of the storage space for the deduplication database and transaction logs. Therefore, a storage full condition is triggered at a 96 percent threshold. If your storage requirements exceed the capacity of an MSDP, you can use more than one media server deduplication node. In order to maintain consistent peak performance, Veritas recommends not exceeding 85 percent of the total MSDP capacity.

The high water mark indicates that the volume is full. When the volume reaches the high water mark, NetBackup fails any backup jobs that are assigned to the storage unit. NetBackup also does not assign new jobs to a storage unit in which the deduplication pool is full.

The high water mark includes the space that is committed to other jobs but not already used.

The default is 98 percent.

The low water mark is a threshold at which NetBackup stops image cleanup. The low water mark setting cannot be greater than or equal to the high water mark setting. The default is 80 percent. Select this option to limit the number of read and write streams (jobs) for each volume in the disk pool.

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 needs to free space, this could result in hours or 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. 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 in which they reside. 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. 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. Queue processing and compaction are automatic processes, but they can easily be initiated at any time.

Efficient Replication

To improve data availability and accessibility, you need to replicate the same data to multiple locations in case of system failure, disaster, or to improve data access speed. Traditional replication is costly, requiring solutions to move an entire data set between two points, and constantly copying files with new changes between locations. With MSDP, replication of the deduplicated data is fast and efficient because only incremental changes are transferred to the remote MSDP server. Whether replicating between appliances, NetBackup BYO (build your wwn), to the cloud, or within or across NetBackup domains, only the incremental data changes are replicated. NetBackup has two replication options:

- Domain-optimized duplication (OptDup) for duplication within a NetBackup domain
- Auto image replication (AIR) for replication between domains

Both use the duplication technology for the data transport, and both use the source system to handle the fingerprinting, guaranteeing only the unique segment data is sent from the source location. Further, both reduce the amount of data sent across the network as well as the amount of data stored on the target.

Efficient, high-performance replication is extended through the NetBackup Appliance's networking options. Replication is available across IP networks and Fibre Channel on NetBackup Appliances. With NetBackup Flex Scale, an external media server is needed for Fibre Channel.

Domain-Optimized Duplication

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 offline storage table (OST) interface. This is known as optimized duplication or OptDup.

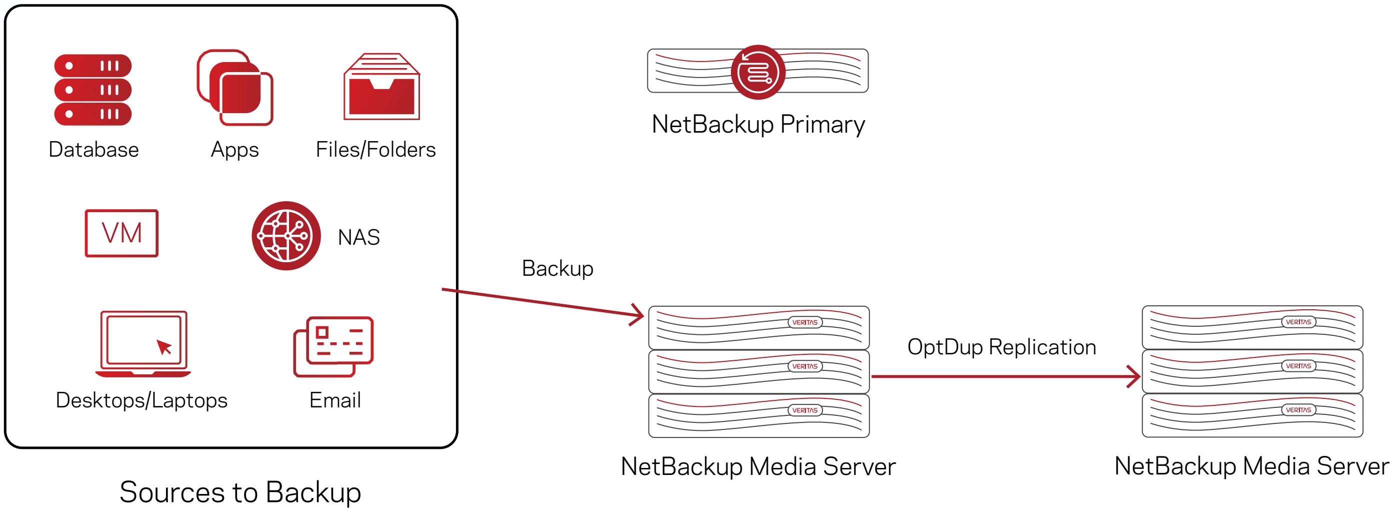

Figure 8. Replication with optimized duplication (OptDup)

NetBackup AIR is used when replicating data from one MSDP pool to another across NetBackup domains. In this scenario, deduplication is also maintained because only the unique data segments are replicated. Further, additional technology allows for the target primary server to have full image recovery capabilities without having to analyze the replicated image. This creates highly available backup data, saving customers time and money, versus managing expensive replication and failover solutions.

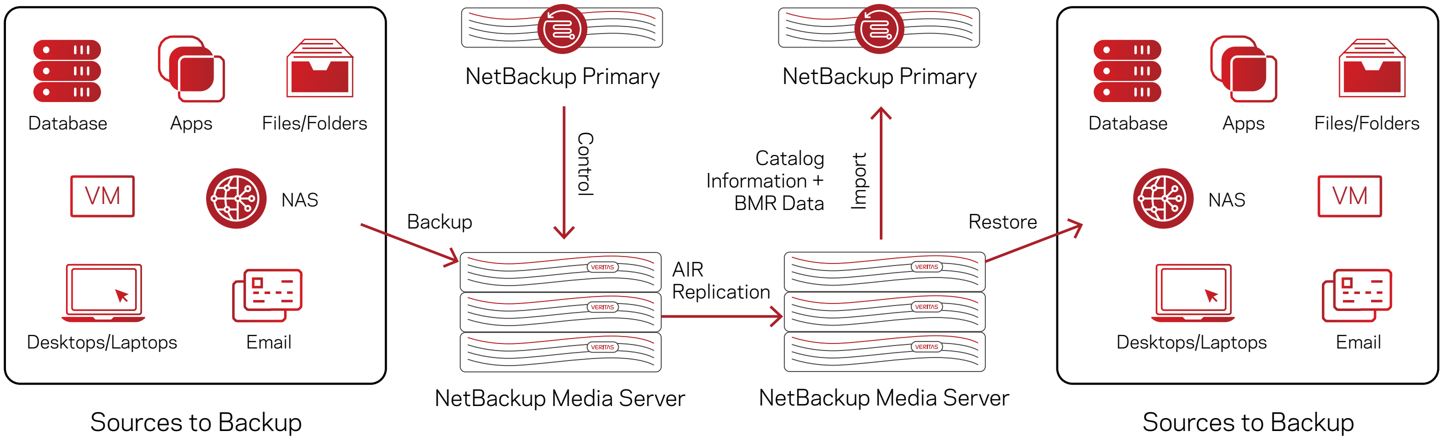

Figure 9. Replication with auto image replication (AIR)

Direct Access to Deduplicated Storage with NetBackup Universal Share

NetBackup Universal Shares provide direct access to deduplicated storage via network file system (NFS) and common Internet file system (CIFS) protocols. Any data that is stored in a universal share is inline deduplicated against any other data previously ingested into the MSDP storage unit. Since a typical MSDP location stores data across a broad scope of data types, the universal share offers significant deduplication efficiency. The protection point feature lets you create a point in time copy of the data that exists in the specified universal share. Once a protection point is created, NetBackup automatically catalogs the data as a specific point in time copy of that data, and manages it like any other data that is ingested into NetBackup. The protection point only catalogs the universal share data that already resides in the MSDP, no data movement occurs. Therefore, the process of creating a protection point can be extremely fast.

NetBackup provides faster backup and recovery with direct data paths.

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

Deduplication with Virtualization Environments

With NetBackup Appliances, customers can leverage NetBackup Universal Share along with MSDP to get instant access to individual files from virtual machines (VMs); or complete secondary copies of VMs to leverage for replication, testing, or 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.

Resilient, Secure, Scalable MSDP with NetBackup Appliances

The integration of NetBackup Appliances with NetBackup software maximizes your ability to manage, protect, and control enterprise data. With NetBackup Appliances, you can create new secure NetBackup deployments in minutes, tailored to meet specific business unit needs.

Increase Operational Efficiency and Return on Investment

Deploy NetBackup Data Protection in Minutes

You can consolidate NetBackup primary servers and media servers on a single appliance to increase efficiency and TCO. With a few clicks, you can deploy or upgrade NetBackup software to quickly adapt to a changing business environment.

Fast Ingest and Recovery Rates

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

Appliance		Distributed Client Deduplication	Server-Side Deduplication
NetBackup 5350 Flex Appliance	Backup	197 TB/Hr	37 TB/Hr Single Node 58 TB/Hr HA
	Restore	13.2 TB/Hr	13.2 TB/Hr
NetBackup 5350 Flex Scale	Backup	263 TB/Hr - 4 Node cluster 957 TB/Hr - 16 Node Cluster	80 TB/Hr - 4 Node Cluster 263 TB/Hr - 16 Node Cluster
	Restore	17.1 TB/Hr	17.1 TB/Hr

Figure 10. Maximum backup throughput with 98 percent deduplication rate

Maximum Recommended Concurrent Instant Access Streams:

- NetBackup Flex 5350 Appliance: 50
- NetBackup Flex Scale: 800 (50/node)

End-to-End Resilience

NetBackup Appliances include multiple layers of resilience built in to provide the highest levels of data protection. Exact implementations vary slightly for NetBackup Flex Appliance and NetBackup Flex Scale, but overall include:

- Storage resiliency features such as RAID, erasure coding, and mirroring
- Hardware and service monitoring
- AutoSupport and Call Home functionalities
- Hardware resiliency
- Container isolation and management

NetBackup Media Server Deduplication Pool Cloud Tier

If you're like most companies, you have a vast amount of data types laid out across various data centers and the cloud. Now you are facing the challenge of figuring out how to protect the data silos in different locations, infrastructures, and management systems.

NetBackup Appliances with NetBackup software is the best one-stop-shop solution, allowing you to store data from more than 800 different data sources and more than 60 cloud storage targets, all with a single platform. NetBackup Media Server Deduplication Pool Cloud Tier (MSDP-C) enables you to back up to and restore data from cloud storage as a service (STaaS) vendors.

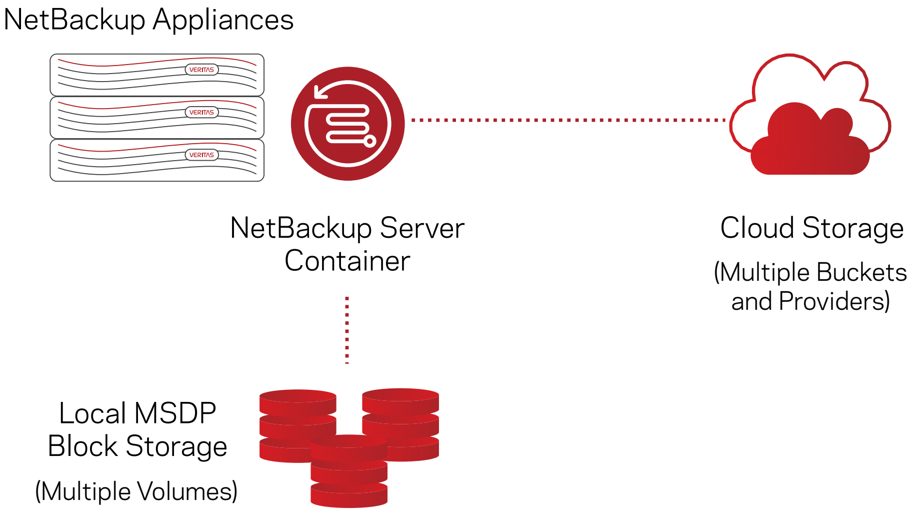

Figure 11. Data deduplication to the cloud with MSDP-C

Storage Efficiency

One of the most important factors organizations consider when choosing a long-term retention storage solution is the cost.

The deduplication engine in the NetBackup MSDP stores the data in a storage-optimized, portable format, and is designed to transport to any compatible target on any infrastructure. It also features storage efficiency technologies such as deduplication and compression to reduce egress and ingress costs to and from the cloud. Customers have seen space savings up to 95 percent, which translates to a lower storage bill at the end of the month. NetBackup can write this deduplicated data directly to local storage as well as—or in addition to—writing it in the cloud, resulting in significant cost savings for your local and cloud storage.

Simplicity

Not only can the NetBackup deduplication engine store data efficiently locally, but also directly to multiple cloud storage targets.

For NetBackup Flex Appliance, you can choose whether to run the storage server, media server, or primary server containers on the same physical server or not. After you configure the MSDP storage server, you can add a cloud storage target to that storage server, and then MSDP can store data directly to the cloud target.

For NetBackup Flex Scale, your storage, media, and primary server instances as well as MSDP storage are automatically configured for you during initial deployment and during scaling operations; you never have to manually configure them. You can simply choose to add a cloud storage target and then MSDP can send deduplicated data directly to the cloud target.

No matter where your data is, NetBackup Appliances can help you protect and control valuable digital assets.

Security By Default

Veritas provides a unified, multi-layered, hardened, and secure appliance platform that optimizes operational efficiency and seamlessly integrates comprehensive protection into an industry-leading backup and recovery solution. As an industry leader in data protection, Veritas provides the technological depth and experience to safeguard your business-critical data across physical, virtual, and cloud environments.

Zero Trust

A zero trust architecture uses the least privileges necessary to complete a particular task based on roles and permissions, combined with robust user authentication, authorization, and policy-based data protection. Using a zero trust architecture, NetBackup Appliances provide a unified, multi-layered platform approach to seamlessly integrate intelligent protection, comprehensive detection, and industry-leading backup and recovery. NetBackup Flex Appliances offer multi-domain isolation, network segregation, and container separation. They feature write once, read many (WORM) storage, STIG-compliant operating system (OS) hardening, FIPS 140-2–compliant data encryption, and comprehensive security access controls. Flex Appliances provide a complete immutable and indelible storage solution to defend an organization's backup data and enable recovery in software and hardware.

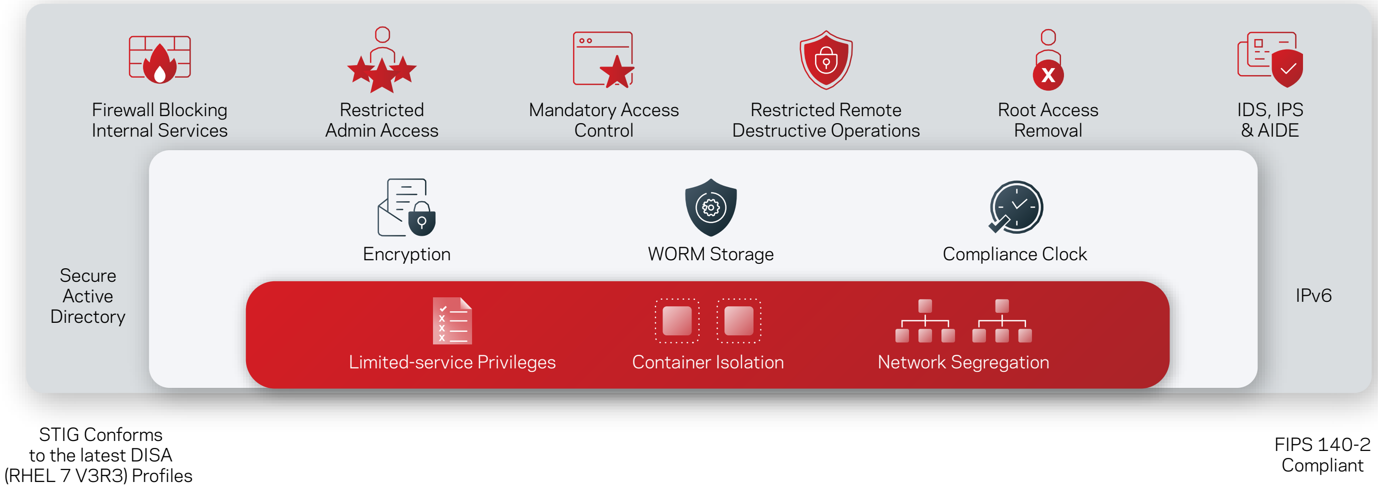

Figure 12. Veritas NetBackup Appliances Zero Trust Architecture

Veritas data protection appliances include native ransomware recovery for business-critical data—at any scale—with near-zero RPO and RTO. Key benefits include:

- Simplified IT management with immutable storage
- A secure-by-default architecture
- Integrated and highly available system configurations

Data Encryption

NetBackup security and encryption provide protection for all parts of NetBackup operations on NetBackup primary servers, media servers, storage servers, and attached clients. The backup data is protected through encryption processes. NetBackup data that is sent over the network is protected by dedicated and secure network ports.

To keep data protected, FIPS-compliant AES-256 encryption can be used so that it is not readable without first being decrypted. NetBackup provides encryption for the deduplicated data. It is separate and different from NetBackup policy-based encryption. With client-side deduplication, the data can be encrypted before it is sent to the media server, so that it is secure as it is sent across the network.

To ensure optimal security, NetBackup includes encryption features for data in rest and data in motion. You can encrypt your data before you send it to the cloud.

- For backups, the deduplication plug-in encrypts the data after it is deduplicated. The data remains encrypted during transfer from the plug-in to the NetBackup deduplication engine on the storage server. The deduplication engine writes the encrypted data to storage. For restore jobs, the process functions in reverse.
- For duplication and replication, the deduplication plug-in on MSDP servers encrypts the data for transfer. The data is encrypted during transfer from the plug-in to the NetBackup deduplication engine on the target storage server and remains encrypted on the target storage.

Key Management Server

NetBackup uses a primary server-based symmetric key management service (KMS) to manage the keys for the data encryption for cloud disk storage.. The service runs on the NetBackup primary server.

NetBackup incorporates KMS with MSDP. NetBackup uses security certificates to authenticate NetBackup hosts. The security certificates conform to the X.509 Public Key Infrastructure (PKI) standard. A primary server acts as the certificate authority (CA) and issues digital certificates to hosts. NetBackup can also retrieve keys from an external KMS server to encrypt MSDP backup data.

MSDP encryption carries out segment-level encryption and assigns a unique encryption key for every data segment. A customer key is retrieved from NetBackup KMS to encrypt the segment key. Key creation and activation must be done manually (or using scripts) by the user.

> You can configure the KMS service from the NetBackup administration console or the NetBackup command line during storage server configuration.

Note: You cannot disable the MSDP KMS service once you enable it.

Enhance Security and Data Efficiency with Immutable WORM Storage

NetBackup Flex Appliances are designed with security at the forefront and provide a complete immutable and indelible storage solution to ensure your system and data are recoverable.

The NetBackup primary server communicates with the storage unit to gather the immutability and indelibility capability and WORM retention period (min/max) settings. The primary server sets up immutability controls on the storage unit and applies the WORM retention period policy. NetBackup provides backup image management with visual representation of the immutable lock, image deletion after the WORM retention period (via the command line interface [CLI]), and honors legal hold on the catalog.

Isolated Recovery Environment (IRE)

For enhanced ransomware resiliency, it is important to not only secure your backup data on immutable storage, but to also maintain an isolated copy of your backup data. This is often referred to as an air-gapped copy. An IRE enables air-gapped backup copies by disabling network connectivity to a secure copy of your critical data, providing administrators a clean set of files on demand to neutralize the impact from a ransomware attack. The NetBackup IRE solution:

- Ensures data is immutable and indelible, minimizing threats from ransomware and rogue users
- Detects ransomware infections within the protected data to prevent reinfection when restoring data
- Enables recovery operations at scale so business services can meet service-level objectives
- Enables predictable recovery processes that can be rehearsed to on-premises or cloud infrastructures

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.

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 0.50 TB to 250 TB of deduplicated, on-premises storage.

Virtual Appliance	Use Case	Capacity
Primary + Media Server	Remote Office / Back Ooffice (ROBO)	0.50 TB to 16 TB
Primary Sserver	Data Center	N/A
Media Server	Data Center	0.50 TB to 250 TB

Figure 13.

MSDP-C is a bundle feature with MSDP. Once MSDP is created, MSDP-C will be available and can be deployed in the data center (same use case as the media server), its local cache capacity shares the available MSDP capacity.

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

Conclusion

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

Additional Resources

Product Documentation:

- NetBackup Deduplication Guide
- For installation, configuration, and administration of each of the products discussed in this white paper, see the appropriate Veritas product documentation

Flex Appliances:

- Benefits, features, use cases, and architectural details: Flex Appliance Design Guide
- Details on security features built into Flex Appliances: NetBackup Flex Appliance Security Guide

- Air gap deployment: NetBackup Isolated Recovery Environment Solution
- Best practices and sizing recommendations: NetBackup Flex Appliance Best Practices
- Integration and API guide: NetBackup Flex API Guide

NetBackup Flex Scale:

- Benefits, features, use cases, and architectural details: NetBackup Flex Scale Technical Overview

Details on security features built into NetBackup Flex Scale: NetBackup Flex Scale - Secure-By-Default

About Veritas

Veritas Technologies is a leader in multi-cloud data management. Over 80,000 customers—including 95 percent of the Fortune 100—rely on Veritas to help ensure the protection, recoverability, and compliance of their data. Veritas has a reputation for reliability at scale, which delivers the resilience its customers need against the disruptions threatened by cyberattacks, like ransomware. No other vendor is able to match the ability of Veritas to execute, with support for 800+ data sources, 100+ operating systems, 1,400+ storage targets, and 60+ clouds through a single, unified approach. Powered by Cloud Scale Technology, Veritas is delivering today on its strategy for Autonomous Data Management that reduces operational overhead while delivering greater value. Learn more at veritas.com. Follow us on Twitter at @veritastechllc.

VERITAS™

2625 Augustine Drive
Santa Clara, CA 95054
+1 (866) 837 4827
veritas.com

For global contact
information visit:
veritas.com/company/contact