ONTAP overview

DM Series OS overview

ONTAP – preinstalled software features

Scroll down for more information

Software feature	Function
All supported storage protocols included in configuration	Unified support for SAN and NAS workloads Ability to support both block and file workloads on one storage system
FlexGroup	Massively scalable, high-performance data containers
RAID DP	Double parity RAID protection protects against two concurrent failed drives or a single drive failure and an uncorrectable media error
RAID TEC	Triple parity RAID protection protects against three concurrent failed drives, two failed drives and an uncorrectable media error, or a single drive failure and uncorrectable media errors on two drives
SnapMirror	Asynchronous replication for backup and disaster recovery Array-based replication protects against failures that affect an entire data center and regional disasters over any distance
MetroCluster	Synchronous mirroring to enable business continuity Two ONTAP clusters synchronously replicating to each other, which protects against external power failures of power, cooling and network issues, and natural disasters

ONTAP – preinstalled software features

Scroll down for more information

Inline Deduplication	Allows users to store just one copy of each unique data object, greatly reducing capacity requirements
Inline Compression	Ability to store data in less space
FlexClone	Creates instant read/write virtual copies Allows for multiple copies of data which can be modified for various uses within the organization
Advanced Drive Partitioning	Increases usable storage capacity Enables capacity sharing of physical drives, HDDs, or SSDs between aggregates and controllers in an HA pair
NetApp Volume Encryption (NVE)	Software-based data-at-rest encryption on any type of disk **Note**: WW ships with encryption except for China and Russia
Storage Quality of Service (QoS)	Used to prevent workloads that consume more than their share of performance Can be used to dedicate performance for critical workloads
FlashCache	Accelerates hot data performance with read caching on flash media Ideal for multiple heterogeneous workloads, as storage latency for random reads is reduced

ONTAP – preinstalled software features

Scroll down for more information

Advanced Drive Partitioning	Enables capacity sharing of physical drives, HDDs, or SSDs between aggregates and controllers in an HA pair
NetApp Volume Encryption (NVE)	Software-based data-at-rest encryption on any type of disk **Note**: WW ships with encryption except for China and Russia
Storage Quality of Service (QoS)	Used to prevent workloads that consume more than their share of performance Can be used to dedicate performance for critical workloads
FlashCache	Accelerates hot data performance with read caching on flash media Ideal for multiple heterogeneous workloads, as storage latency for random reads is reduced
Flash Pool	Accelerates hot data performance with read and write caching on aggregate of HDDs and SSDs Ideal for database and transactional applications
SnapRestore	Recovers file systems or entire data bases in an instant
SnapCenter	Unified, scalable software for data protection and clone management
Snapshot	Incremental point-in-time copies

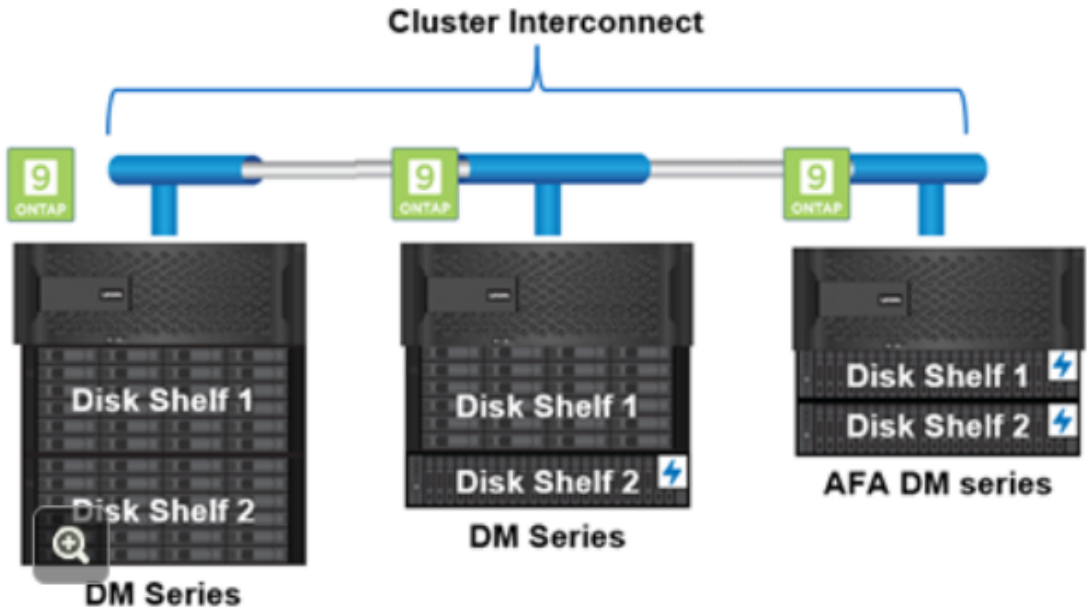
ONTAP – optional software features

Software features	Functions
FlexArray Virtualization	Storage virtualization of third-party arrays
FabricPool	Enables automated tiering of Lenovo Snapshot and secondary data to low-cost object storage tiers on premises or off premises
SnapLock	SnapLock is an alternative to the traditional optical "write once, read many" (WORM) data SnapLock is used for the storage of read-only WORM data, and it supports regulatory compliance and organizational data retention requirements
OnCommand Insight	Simplifies operational management for private/hybrid cloud and virtual IT environments

ONTAP cluster

Each DM Series controller that runs ONTAP software is called a node. Nodes are paired for high availability (HA). Together, these pairs (up to 12 nodes for SAN, and up to 24 nodes for NAS) comprise the cluster. Nodes communicate with each other over a private, dedicated cluster interconnect. The cluster interconnect makes the nodes appear as a single cluster.

Depending on the controller model, node storage consists of flash disks, capacity drives, or both. Network ports on the controller provide access to data. Physical storage and network connectivity resources are virtualized, visible to cluster administrators only, not to NAS clients or SAN hosts. Nodes in an HA pair must use the same storage array model; otherwise, you can use any supported combination of controllers. You can scale out for capacity by adding nodes like storage array models or for performance by adding nodes with higher-end storage arrays.



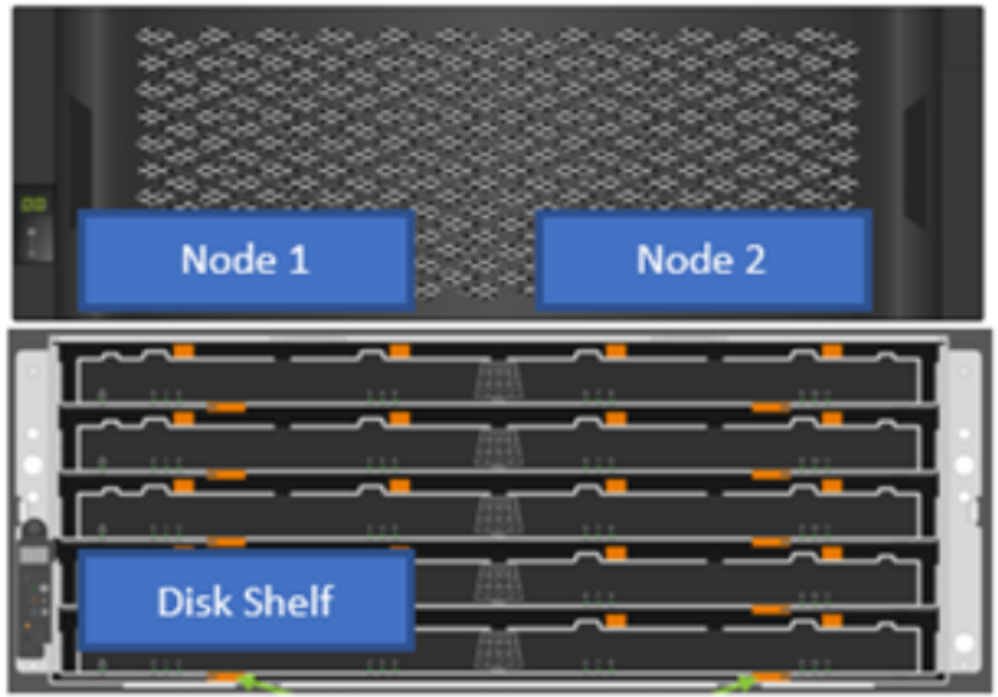
Click to enlarge the graphic

HA pair of DM controllers

Typically, DM Series storage systems are deployed in HA pairs. HA pairs provide hardware redundancy that is required for nondisruptive operations and fault tolerance. The partner node can take over storage services during maintenance events or if a controller hardware failure occurs.

An HA pair includes an interconnect between the two nodes to mirror write operations to the partner's non-volatile RAM (NVRAM). When a write request is made to either of the two nodes in an HA pair, the write request is logged on both nodes before a response is sent back to the client or the host.

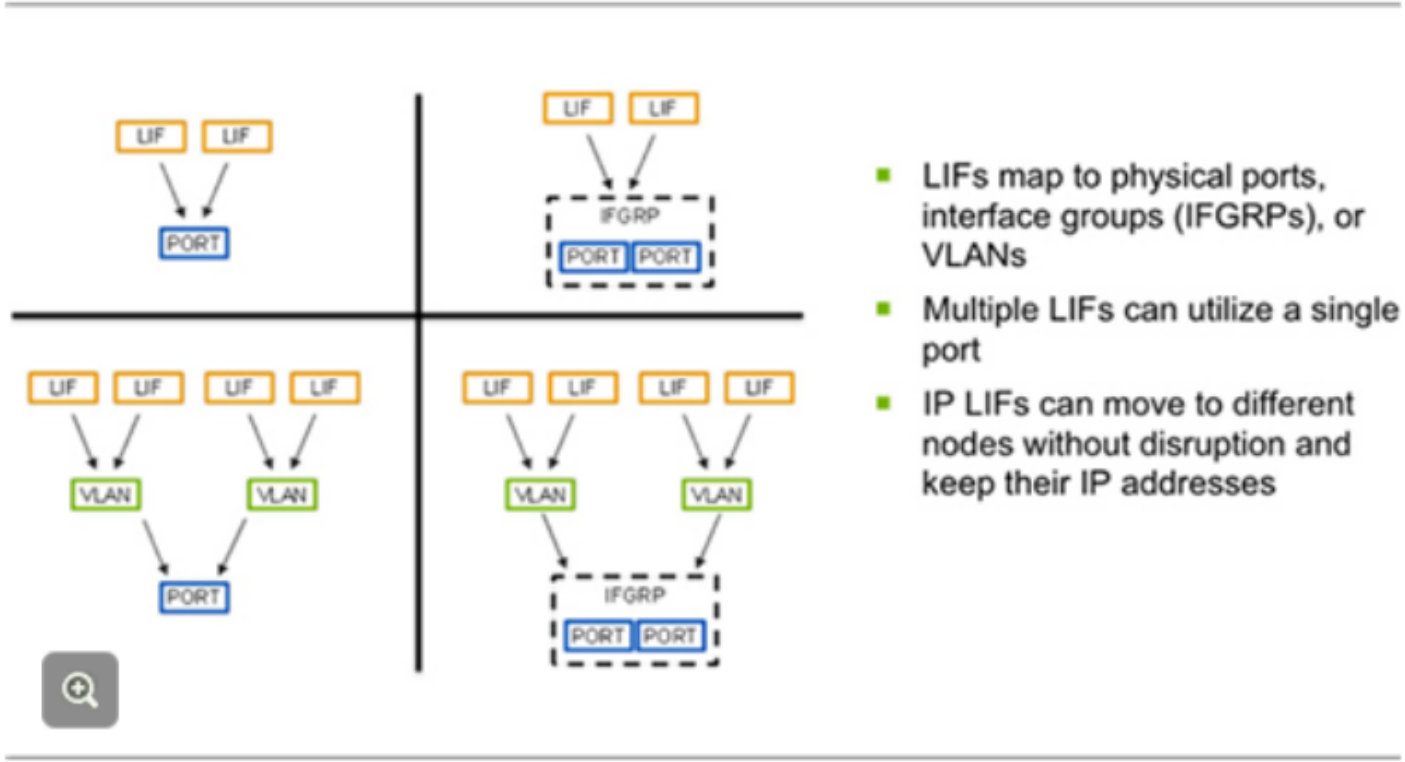
Each controller in an HA pair also has connections to the other controller's disk shelves. The connections provide continued access to the data if one of the two nodes needs to take over storage services for the other node. The controller failover function provides continuous data availability and preserves data integrity for client applications and users.



Ports and logical interfaces

Nodes have various physical ports that are available for cluster traffic, management traffic, and data traffic. The ports must be configured appropriately for the environment.
A logical interface (LIF) represents a data network interface for an IP address or World Wide Port Name (WWPN).
LIFs can be placed on physical ports, interface groups, or VLANS. The different types of LIF are as follows:

- Node Management LIF
- Cluster Management LIF
- Cluster LIF
- Data LIF
- Intercluster LIF

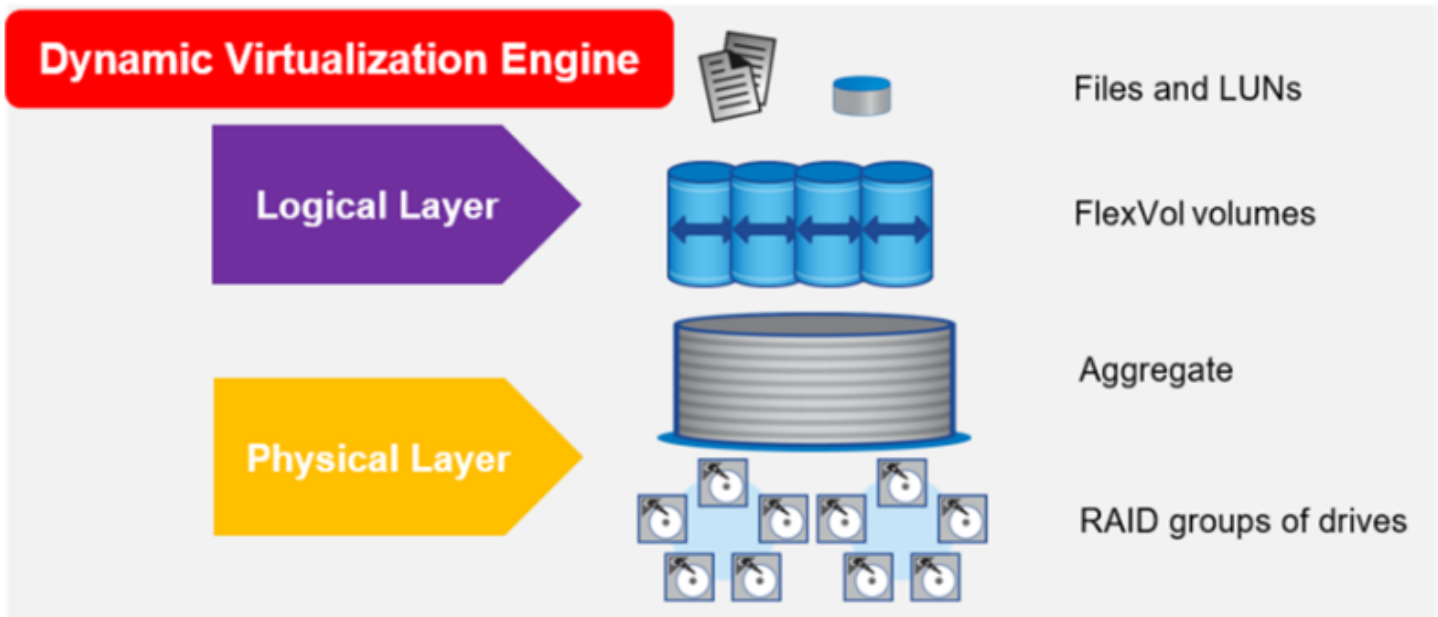


Ports are the underlying physical hardware, and there are different hierarchical stacks to serve LIFs.

ONTAP storage architecture

The ONTAP storage architecture uses a dynamic virtualization engine, for which data volumes are dynamically mapped to physical space.

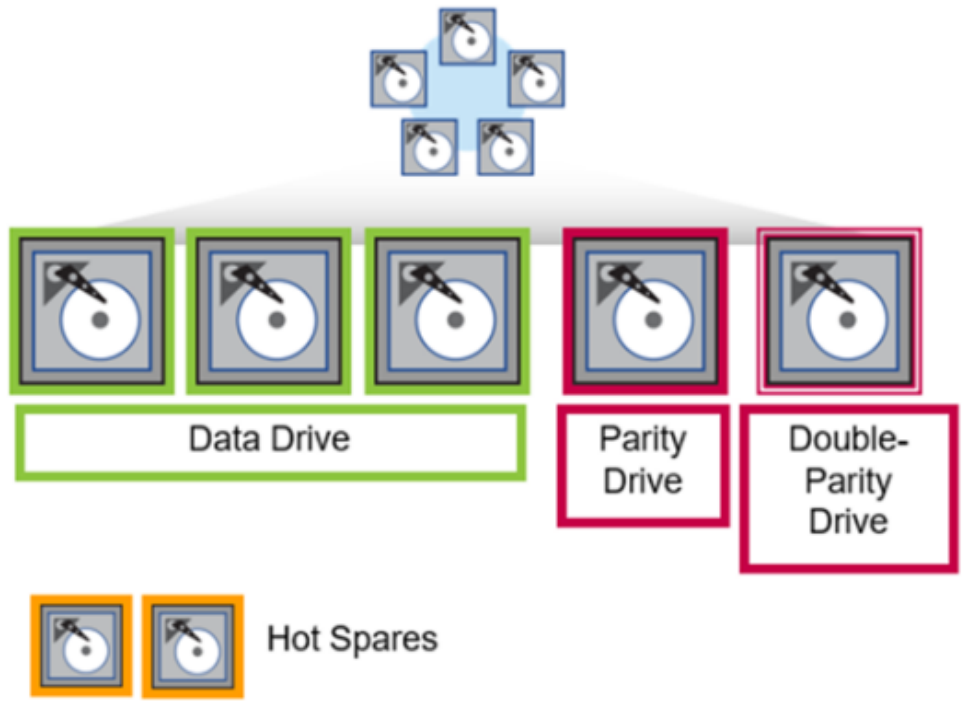
- Physical storage layer: Storage drives, RAID groups, and aggregates make up the physical storage layer. Drives are grouped into RAID groups. An aggregate is a collection of physical drive space that contains one or more RAID groups. Each aggregate has a RAID configuration and a set of assigned drives.
- Logical storage layer: FlexVol volumes, files, and LUNs make up the logical storage layer. Within each aggregate, you can create one or more FlexVol volumes. A FlexVol volume is an allocation of drive space that is a portion of the available space in the aggregate. A FlexVol volume can contain files or LUNs.



Physical layer – RAID groups

Drives are grouped into one or more RAID groups. Drives within a RAID group protect each other when a drive failure occurs. In a RAID group, users must use drives of the same type and the same speed. Users should always provide enough hot spares for each drive type. If a drive in the group fails and there are enough hot spares, the data can be reconstructed on a spare drive.
The RAID group contains the following drive types:

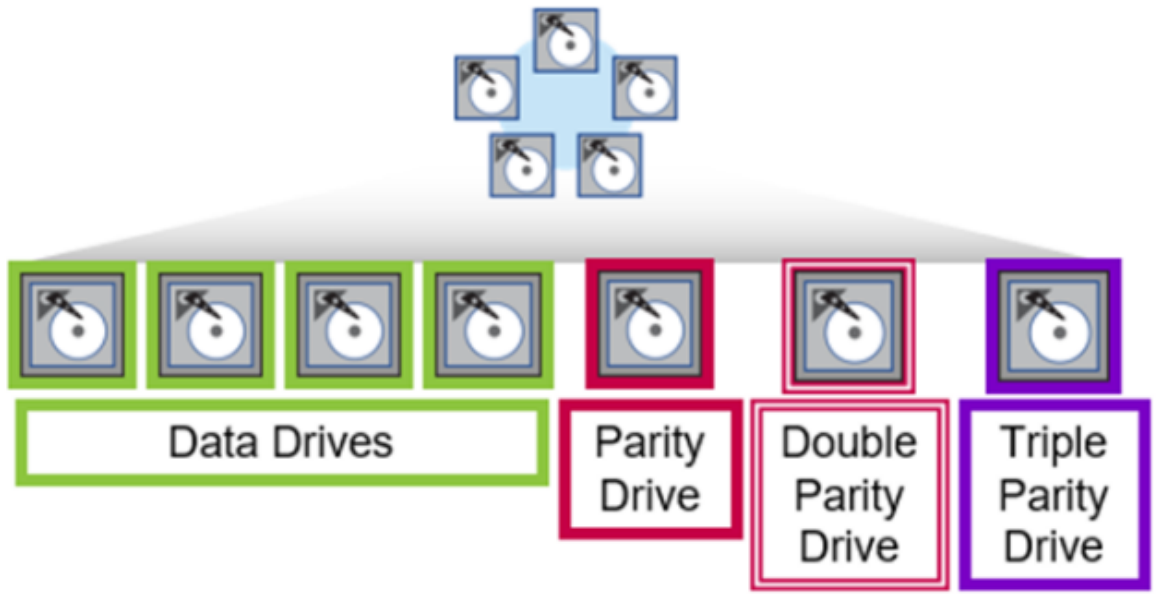
- Data drive: Holds data stored on behalf of clients within RAID groups and any data generated about the state of the storage system as a result of a malfunction
- Parity drive: Stores row parity information that is used for data reconstruction when a single disk drive fails within the RAID group
- Double parity (dParity) drive: If RAID-DP is enabled, stores diagonal parity information that is used for data reconstruction when two disk drives fail within the RAID group
- Spare drive: Any functioning disk that is not assigned to an aggregate but is assigned to a system functions as a hot spare disk – does not hold usable data, but it is available to be added to a RAID group in an aggregate



Physical layer – RAID types

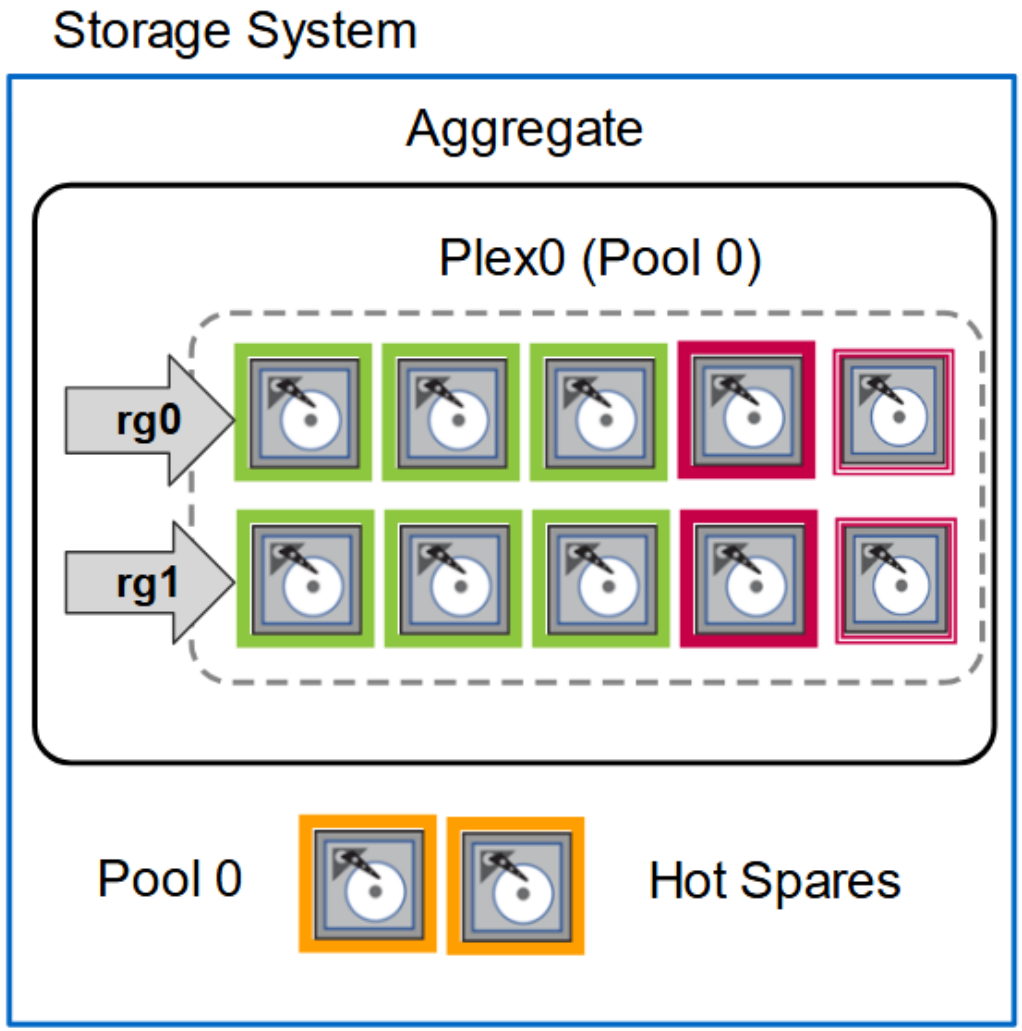
ONTAP uses three primary RAID types:

- RAID 4:
  - Provides a parity drive that protects the data if a single-drive failure occurs
  - Uses data aggregates that require at least three drives
- RAID-DP technology:
  - Provides two parity drives that protect the data if a double-drive failure occurs
  - Uses data aggregates that require at least five drives
- RAID-TEC technology:
  - Provides three parity drives that protect the data if a triple-drive failure occurs
  - Uses data aggregates that require a minimum of seven drives



Physical layer – Aggregates

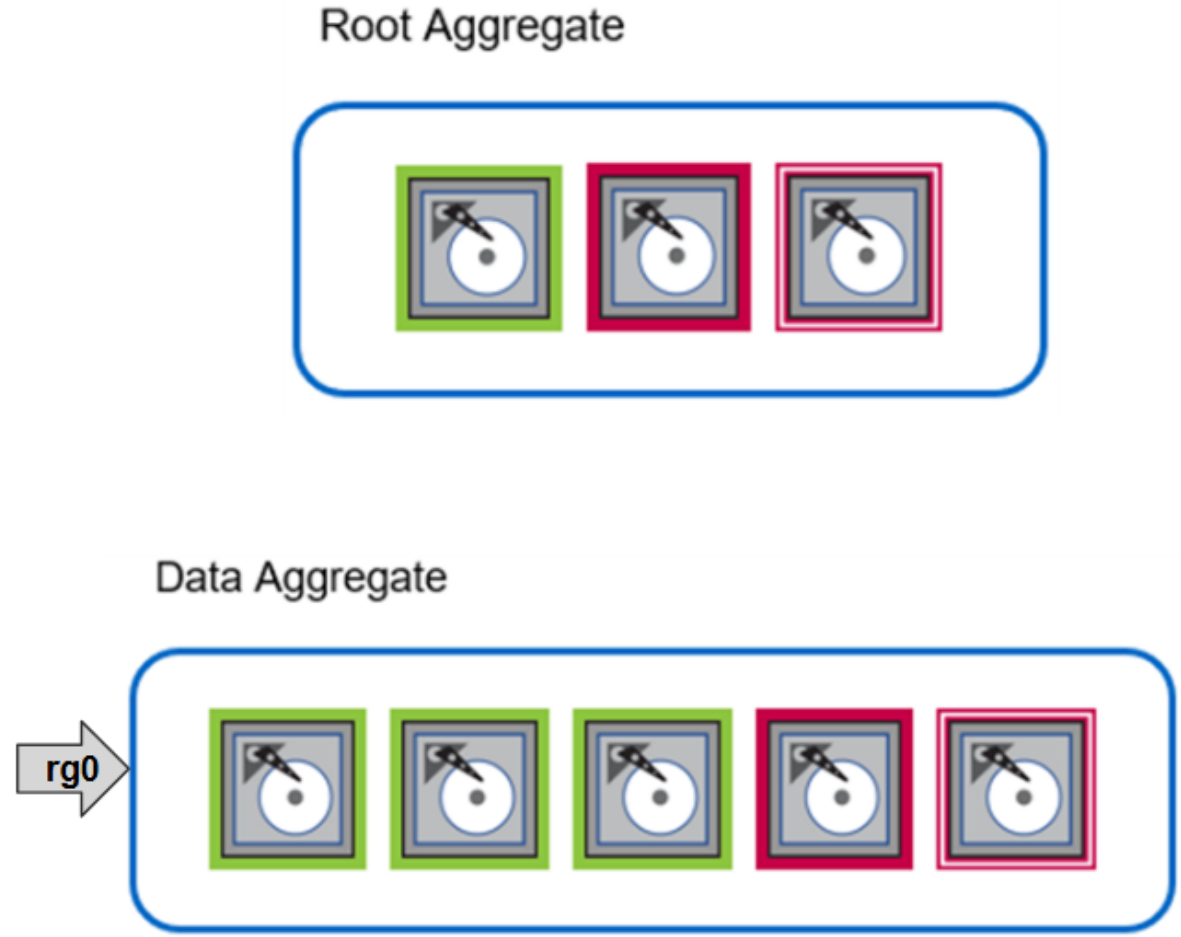
Aggregates are composed of RAID groups (rg) that contain drives or array LUNs.
All RAID groups must be of the same RAID type.
Drives in the aggregate must be of the same drive type.
Aggregates have a single copy of data, which is called a plex.
A plex contains all the RAID groups that belong to the aggregate.
Mirrored aggregates have two plexes.
A pool of hot spare disks is assigned to each plex.



Physical layer – Aggregate types

There are two types of aggregate:

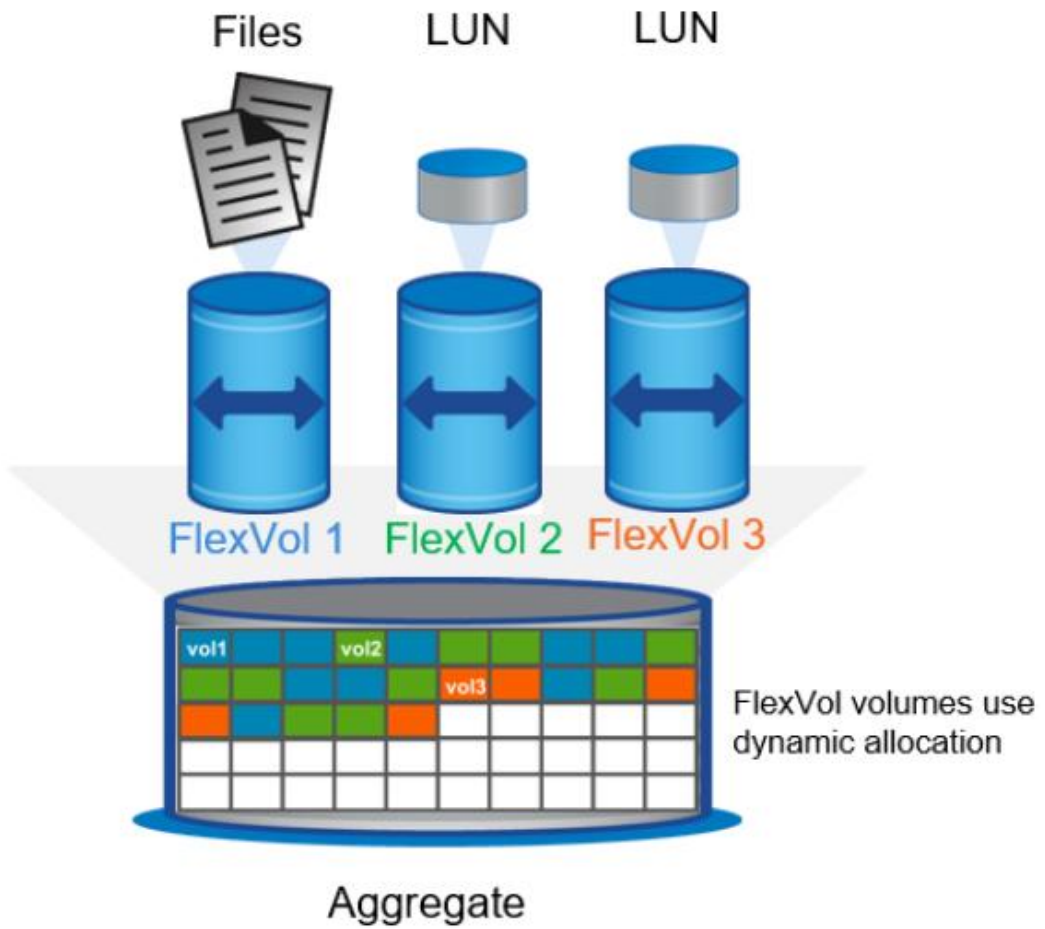
- Root aggregate (aggr0)
  - Each node of an HA pair requires three drives to be used for a RAID-DP root aggregate (aggr0), which is created when the system is first initialized. The root aggregate contains the node's root volume, named vol0, which contains configuration information and log files. ONTAP prevents the creation of other volumes in the root aggregate.
- Data aggregates
  - Aggregates for user data are called non-root aggregates or data aggregates. Data aggregates must be created before any data storage virtual machines (SVMs) or FlexVol volumes. When data aggregates are created, the default is a RAID-DP configuration with a minimum of five drives for most drive types. The aggregate can contain hard drives, SSDs, or array LUNs.



Logical layer – FlexVol volumes

A FlexVol volume is an allocation of drive space that is a portion of the available space in the aggregate.

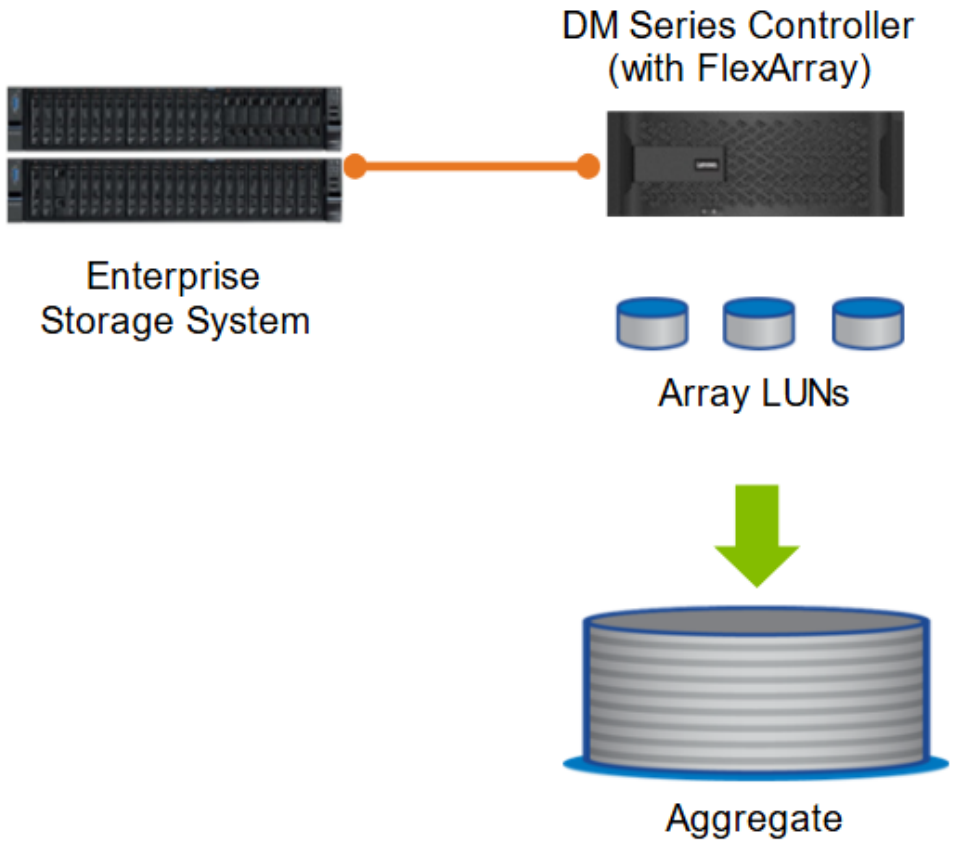
- FlexVol volume size is flexible and can be increased or decreased in size as needed.
  - FlexVol volumes use dynamic allocation.
- Aggregates can contain multiple FlexVol volumes.
- They can be used to contain files in a NAS environment or LUNs in a SAN environment.
- Files: refers to any data that is exported or shared to NAS clients (including text file, spreadsheet, and database)
- LUN: represents a logical disk that is addressed by a SCSI protocol (FC or iSCSI).
  - Block level
  - Data can be accessed only by a properly mapped SCSI host



Logical layer – Array LUNs

Like drives, array LUNs can be used to create an aggregate. With the FlexArray storage virtualization software licenses, an enterprise storage array can present an array LUN to ONTAP. An array LUN uses an FC connection type.

- FlexArray creates array LUNs on third-party storage systems:
  - Array LUNs are presented to ONTAP.
  - Array LUNs can function as hot spares or be assigned to aggregates.
- Array LUNs in an aggregate:
  - Aggregates use RAID 0.
  - Aggregates can contain only array LUNs.



Storage virtual machines

Storage virtual machines (SVMs) contain data volumes and one or more LIFs through which they serve data to the clients.
SVMs securely isolate the shared virtualized data storage and network, and each SVM appears as a single dedicated server to the clients. Each SVM has a separate administrator authentication domain and can be managed independently by its SVM administrator.
In a cluster, SVMs facilitate data access. A cluster must have at least one SVM to serve data. SVMs use the storage and network resources of the cluster. However, the volumes and LIFs are exclusive to the SVM. Multiple SVMs can coexist in a single cluster without being bound to any node in a cluster. However, they are bound to the physical cluster on which they exist.

SVM and FlexVol volumes

FlexVol volumes are the unit of data management in SVMs.

- FlexVol volumes can be moved, mirrored, backed up, or copied by using Snapshot copies in the SVM.
- FlexVol volumes can be mirrored (not moved) to other SVMs.

FlexVol volume types:

- System (or node root):
  - Typically named vol0
  - Should only contain configuration and logs
  - Should never contain user data
- SVM root volume:
  - Top level of the namespace
  - Should not contain user data, only junction paths
- Data volume:
  - Contains user data
  - If it is NAS: contains file systems for user data
  - If it is SAN: contains LUNs

FlexGroup

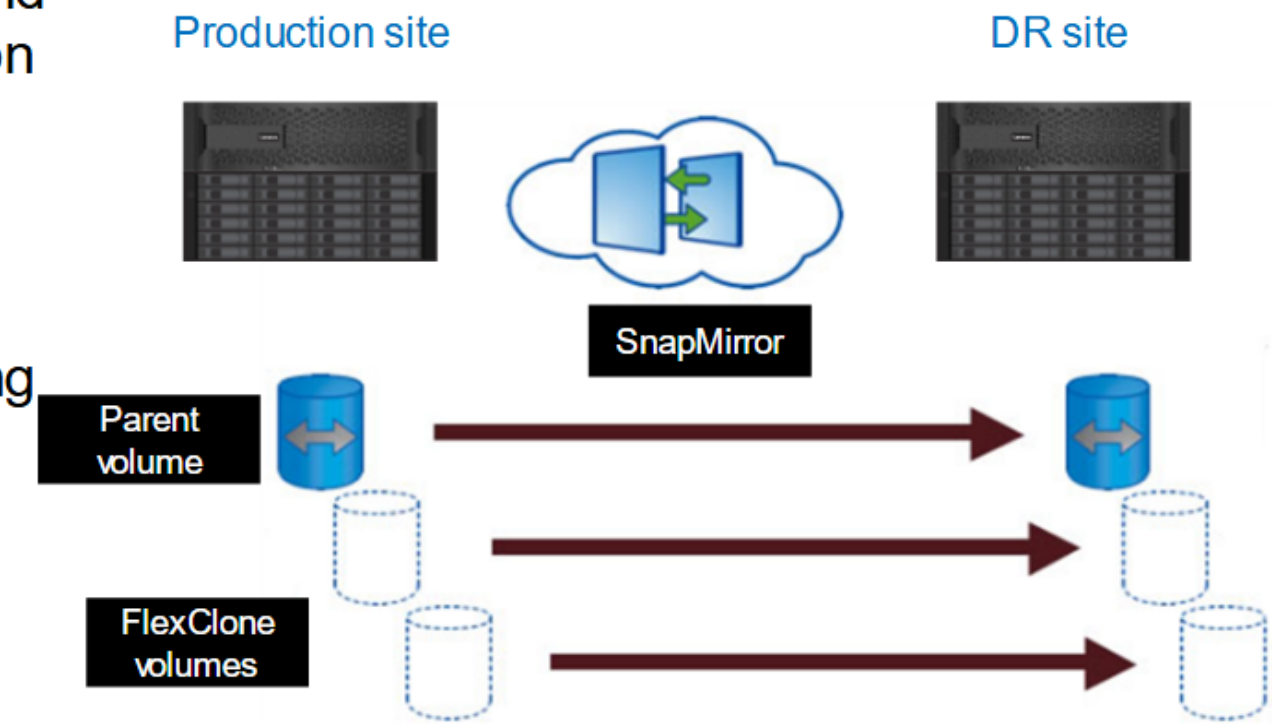
FlexGroup can support a massively scalable and high-performance data container. It supports compute-intensive workloads and data repositories that require a massive NAS container while maintaining consistent high performance and resiliency.

- Linear scale for performance and capacity
  - Scales to 20 PB and 400 billion files
- Operational simplicity
  - Single mount point with automated load and space distribution
- Consistent high performance
  - Predictable, consistent low latency
  - All-flash containers
- Leading resiliency
  - ONTAP® nondisruptive operations

SnapMirror

SnapMirror is an array-based replication and it provides an asynchronous, thin-replication solution based on ONTAP Snapshot technology.

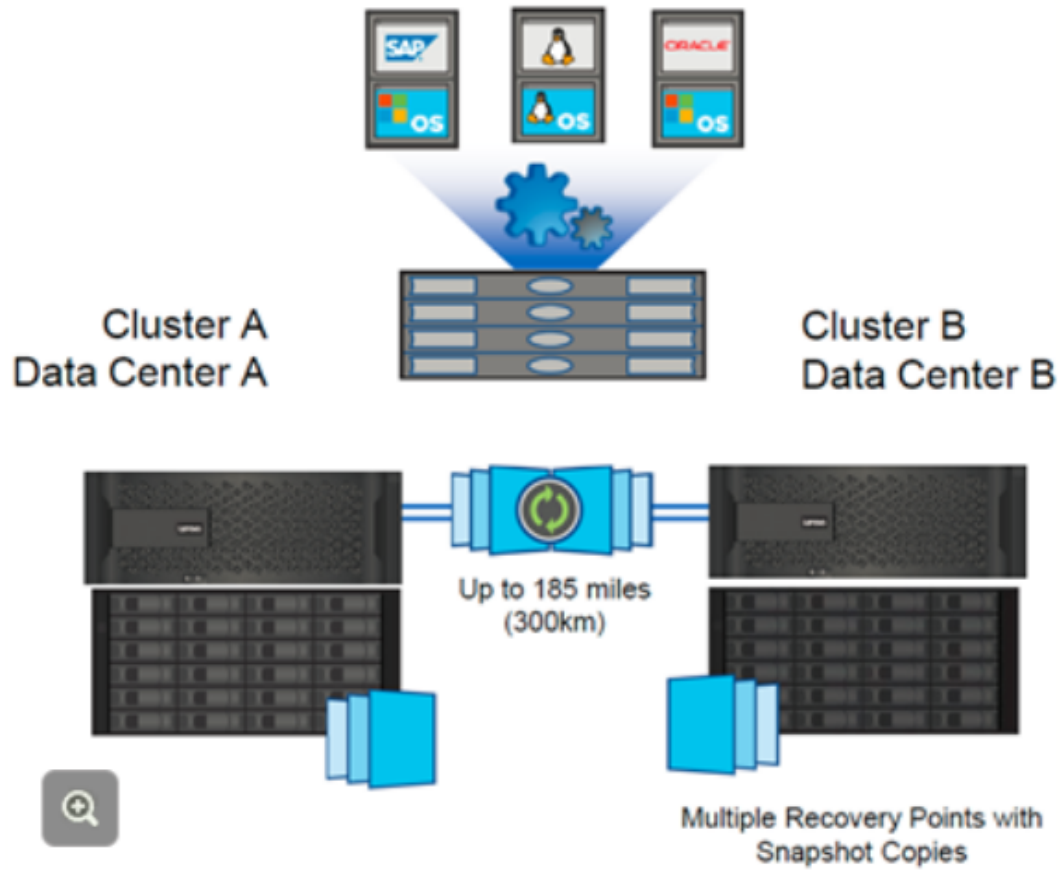
- Offers unified replication for disaster recovery, business intelligence, data distribution, and development and testing
- Leverages ONTAP software for nondisruptive operations and upgrades
- Lets users tune their Recovery Point Objective (RPO) to meet the business requirements
- Works equally well in virtual and traditional environments
- Provides volume- and SVM-level granularity



MetroCluster

MetroCluster means two ONTAP Clusters synchronously replicating to each other, which protects against external power failures of power, cooling and network issues, and natural disasters. It has the following features:

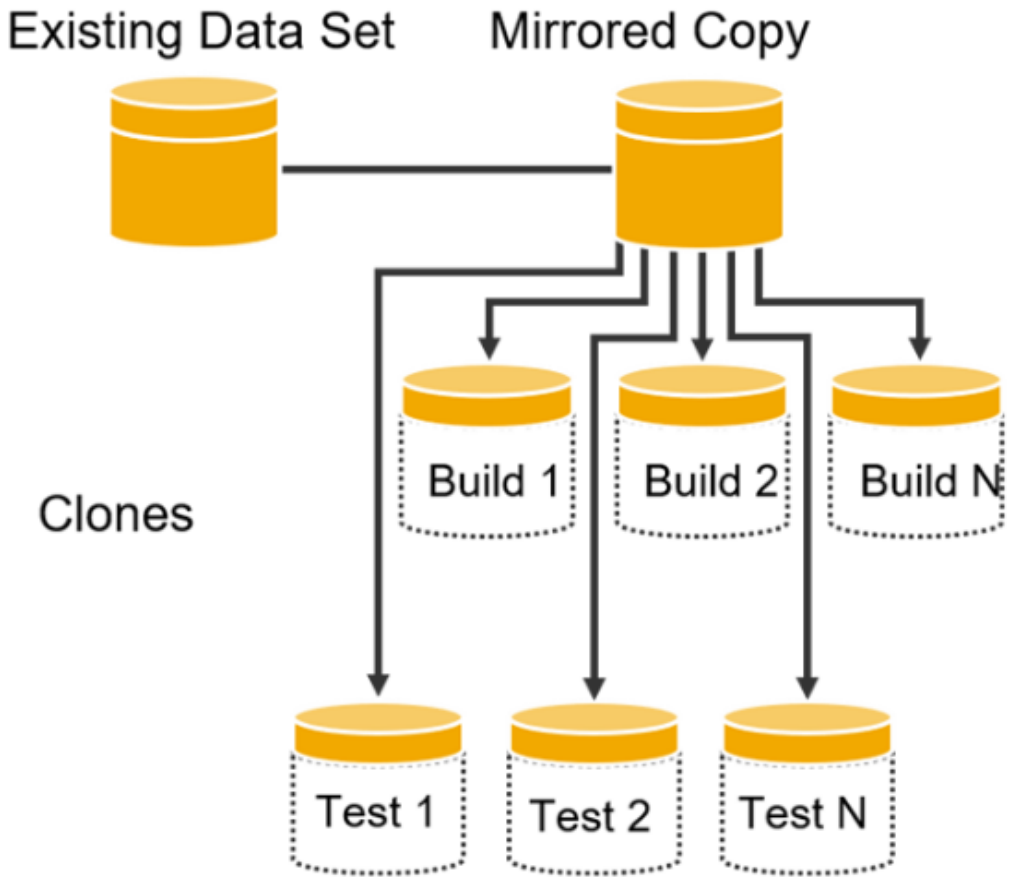
- Nondisruptive operations leading to zero data loss
  - With synchronous replication the recovery point objective (RPO) is equal to 0, so you have zero data loss
- Set-it-once simplicity with zero change management
  - Configuration changes are automatically passed to the remote node, so the administrator does not have to manage the nodes at each site
- Seamless integration with storage efficiency, SnapMirror, NDO, and virtualized storage
- Unified solution: supports SAN and NAS
- Synchronous replication over Fibre Channel or IP networks
  - Available with ONTAP 9.4
  - IP network support is currently only available with a four-node configuration
- No incremental licensing costs – it's built into ONTAP
  - MetroCluster interoperates with all of the features of ONTAP: deduplication, compression, flash, and SnapMirror



FlexClone

FlexClone is used when users need to have multiple copies of data that can be modified for various uses in an organization. It is particularly well suited for:

- Software application development and testing where multiple developers need individual copies of production-level data to work with
- Multiple simulations with data that closely resembles the production environment
- Provisioning for virtual infrastructure where many copies of similar but not identical data set images are needed



Volume encryption

The DM Series offers software-based, data-at-rest encryption onboard the system. The volume encryption feature keeps the system in compliance with the strictest regulations and comes standard with FIPS140-2. Users can also purchase SnapLock to be in compliance with standards, such as HIPAA, SEC 17a-4 through SLC certification. Volume encryption has the following features:

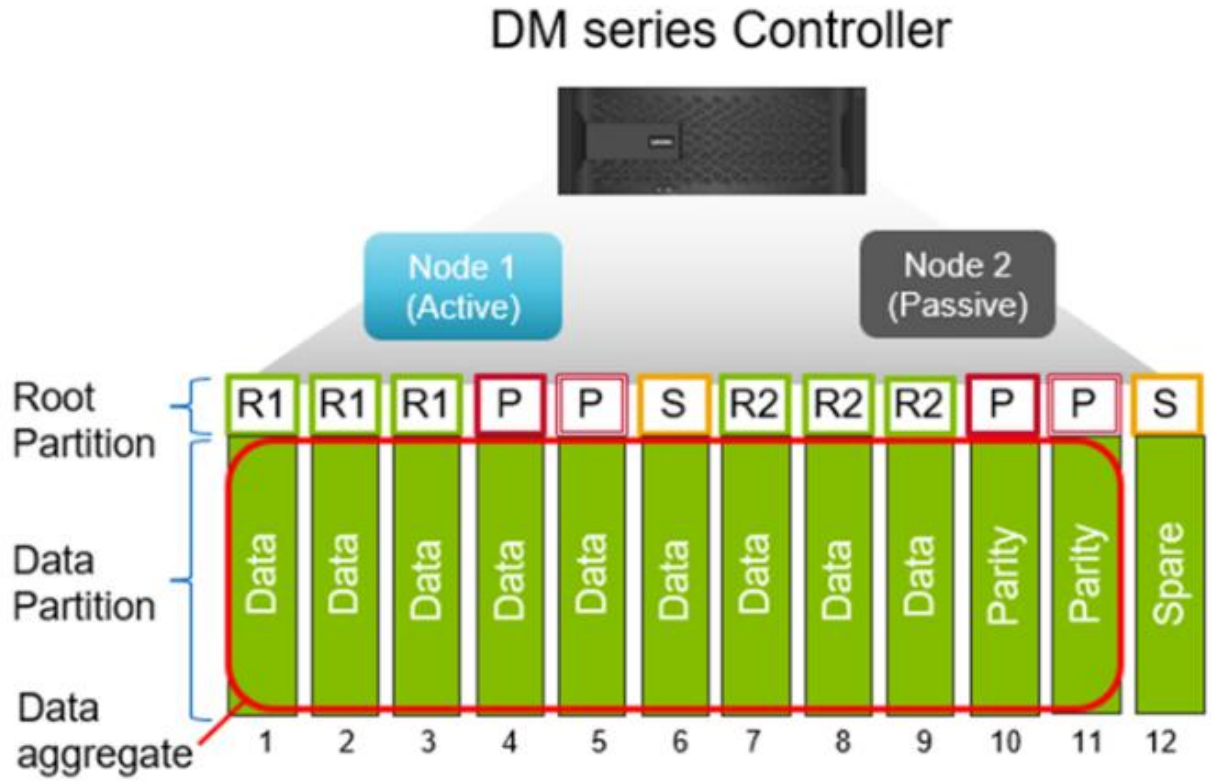
- Comes standard with Multifactor authentication (MFA), which strengthens administer accounts with an additional security layer
- Software encrypts any volume and any disk
  - Lower cost: no need for hardware self-encrypting drives
  - More granular: volume level
  - Onboard key manager
- Leverage storage efficiency features
  - Deduplication, compression, compaction
- Futureproof encryption solution
  - Software updates will keep the algorithms up to date

Advanced Drive Partitioning

To maximize performance, Advanced Drive Partitioning (ADP) can be configured to enable the controllers to share drives to increase efficiency.

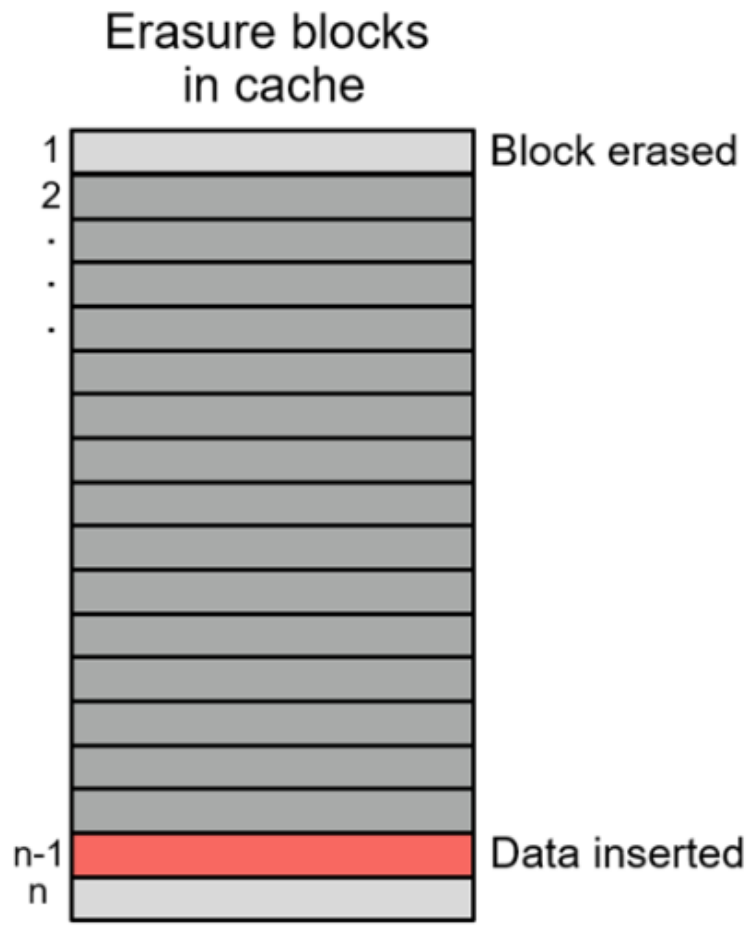
- Drives are shared, so resources are used more efficiently.
  - ADP reserves a small slice from each drive to create the root partition, which can be used for the root aggregates and hot spares. The remaining larger slices are configured as data partitions that can be used for data aggregates and hot spares.
- Fewer drives are required.
- ADP is the default configuration for some systems:
  - HA configurations for entry-level DM systems
  - All Flash DM systems

The graphic shows the active-passive configuration for a DM Series controller with 12 drives, which creates a single data aggregate.



Flash Cache

- Flash Cache is a First-In, First-Out (FIFO) read cache
  - When the cache is full, the oldest data is evicted even if it is still “hot”
- Repeat random read workloads benefit most from caching
  - Write-back policies enable pre-populating cache with new writes
- Controller-level cache
  - All volumes can use the cache
  - Each volume can implement a different caching policy
- A cache map is kept in controller memory
  - Cache is completely rewarmed after unplanned takeovers or shutdowns
  - Cached data is accessible after planned takeovers or shutdowns
- Works with deduplication, cloning, and volume encryption



Flash Pool aggregates

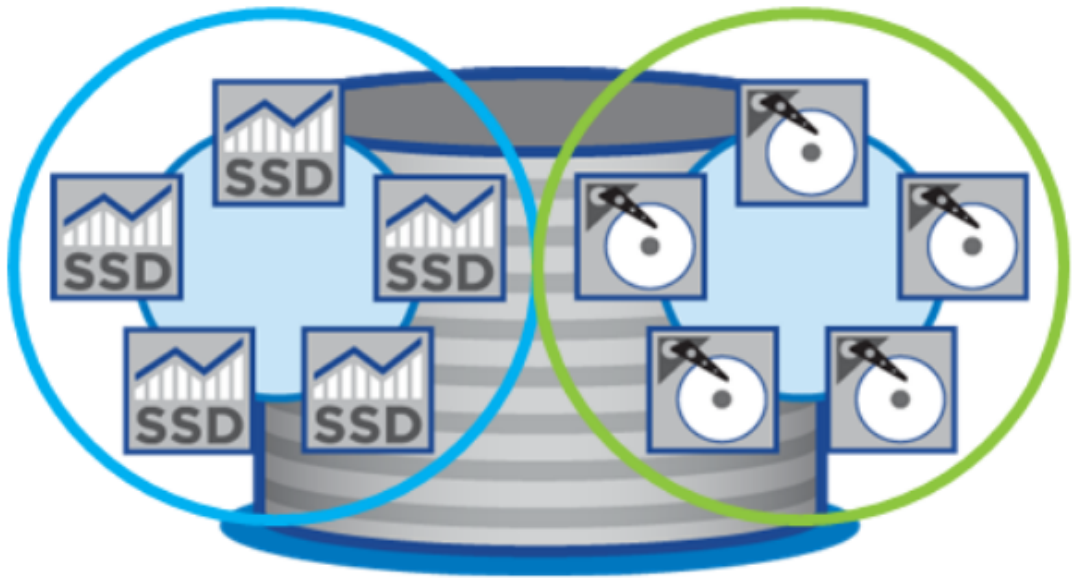
A Flash Pool aggregate is a special type of hybrid data aggregate. A Flash Pool aggregate combines SAS or SATA disks and SSDs to provide high performance at lower cost than a full SSD aggregate. Flash Pool aggregate include:

- SAS or SATA disks for user data
- SSDs for high-performance caching

Performance advantages of Flash Pool aggregates:

- Offloading of random-read operations
- Offloading of repetitive random-write operations

Flash Pool can improve workloads that use online transactional processing (OLTP), such as database application data. Flash Pool does not improve the performance of predominantly sequential workloads.



Flash Cache versus Flash Pool

Flash Cache

- **Controller-level** cache
- First-In, First-Out (FIFO) read cache
- Flash Cache modules in the expansion slots of a node
- Improved response time for repeated, random reads
- Simple use, without additional administration
- Cache for all volumes that are on the controller

Flash Pool

- **Storage-level** cache
- Hybrid aggregates of hard disks and SSDs
- Improved response time for repeated, random reads and overwrites
- Consistent performance across storage failover events
- Cache for all volumes that are on the aggregate

Data protection software tools

ONTAP provides tools that enable you to back up, replicate, and restore data. These tools help you to recover your data if a disaster strikes at the primary data storage site.

- Snapshot copies:
  - Create, schedule, and maintain multiple backups of data on the primary storage volume
  - Quickly and easily restore data that is accidentally modified or deleted
  - Create clones of FlexVol volumes
  - A SnapRestore license is required to perform fast Snapshot recovery from Snapshot copies
- SnapVault backup solution:
  - Efficiently retains backup copies
  - Backs up data and configures data retention policies that meet service-level agreements (SLAs)
- SnapMirror:
  - Replicates data from one FlexVol volume to another within the same cluster or a different cluster
  - Retains those Snapshot copies
  - Uses the destination volume for efficient and fast restore operations of data after a disaster

SnapManager

SnapManager software is a suite of intelligent tools that allows applications and storage to coordinate activities to make backup fast and space efficient, speeds up the restore process, and simplifies common data management tasks. SnapManager provides the following tools:

- SnapManager for Microsoft Exchange Server
- SnapManager for Oracle
- SnapManager for Virtual Infrastructure
- SnapManager for Hyper-V
- SnapManager for Microsoft SQL Server
- SnapManager for Microsoft Office SharePoint Server

SnapManager software runs on the server (or in a virtual machine) and interfaces directly with the application. Each SnapManager tool has its own user interface that lets you create and run backup and replication schedules and perform restores.

ONTAP supplement document

For detailed ONTAP How-To instructions and videos, refer to the supplement document in the course main page.