

Enterprise-Ready Scale-Out with Clustered Data ONTAP 8



Mike McNamara

Senior Manager, Product Marketing

This article, originally released in June, 2012, remains a great introduction to the features and capabilities of clustered Data ONTAP. Minor updates were made in June, 2013, to bring the article up to date with some changes that occurred in subsequent releases. After you read this article, you may want to see our article on [Clustered Data ONTAP 8.2](#) to learn about the latest features. - Editor

The idea behind scale-out storage is a simple one. Instead of growing a monolithic storage system to its limits and then adding another separate storage system, scale-out storage provides a cluster of storage nodes that operates as a single entity. Adding a node to the cluster adds a predictable amount of storage capacity and performance, allowing you to grow storage resources incrementally as demand increases.

Although scale-out storage products have been available from various vendors for some time, they have been deficient in terms of their ability to meet mainstream storage needs. They are typically focused on technical and engineering applications, are NAS only, are limited to a particular type or model of specialized storage controller, and offer limited or no support for the sorts of storage efficiency and data protection capabilities that most storage users need and expect.

NetApp® Data ONTAP® 8 is the first to offer a complete, unified scale-out solution that provides an adaptable, always-on storage infrastructure for highly virtualized environments. The NetApp approach to scale-out provides:

- **Unified storage** with the ability to support FC, FCoE, iSCSI, NFS, pNFS, and CIFS
- **Nondisruptive operations** that eliminate planned downtime, even during servicing and upgrades
- **Integrated data protection** with NetApp Snapshot™ technology and SnapMirror® replication
- **Heterogeneous cluster configurations** that incorporate different controller and disk types and the ability to provide storage tiering within a cluster
- **Innovative storage efficiency**, including deduplication, thin provisioning, and compression
- **Simplified management** for large-scale deployments
- **Secure multi-tenancy** using Storage Virtual Machines (VSMs, formerly referred to as Vservers) to provide secure isolation
- **Massive scalability and performance** capable of scaling to over 50PB of storage and more than 1.5 million SPECsfs2008_nfs.v3 ops/sec

In addition, the latest release, Data ONTAP 8.1.1, offers several new features that enhance performance, manageability, and supportability whether you're using clustering or 7-Mode.

This article focuses on the scale-out capabilities of clustered Data ONTAP 8.

Explore

More on Data ONTAP 8

Want to know more about Data ONTAP 8? Check out these recent Tech OnTap articles and other resources:

- [Get Future Ready with Data ONTAP 8](#)
 - [FAS6200 Running Data ONTAP 8 Cluster-Mode: Performance and Scalability](#)
 - [Data ONTAP 8.1 Cluster-Mode: An Introduction](#)
 - [ESG Data ONTAP 8.1 Lab Validation Report](#)
-

Understanding Clustered Data ONTAP

Clustered Data ONTAP is the enabler for NetApp scale-out storage configurations. The basic building blocks of a cluster are the familiar NetApp HA pairs in which two storage controllers are interconnected to the same set of disks. If one controller suffers a failure, the other takes over its storage and continues serving data.

In a Data ONTAP cluster, each storage controller is referred to as a cluster node, and nodes are allowed to be different models and sizes. For example, a FAS3250, FAS6290, and a [V-Series](#) open storage controller (acting as a front end for a third-party storage array such as EMC, HDS, HP, or IBM) could all be in the same cluster. Disks are grouped into aggregates, which are groups of disks of a particular type composed of one or more RAID groups protected using [NetApp RAID-DP®](#).

A dedicated, redundant 10 Gigabit Ethernet network is used for communication between cluster nodes and for moving data from one logical or physical storage device to another. A separate network is used for management functions.

Figure 1 depicts a heterogeneous cluster with a mix of controller types (including V-Series), storage protocols, and supported disk types to match the performance and cost of the storage to the requirements of the data and workloads. The systems on the left of the diagram represent high-performance systems, built using high-end controllers (for example, FAS6290) and fast SAS and/or SSD drives. The systems on the right side of the cluster represent midrange controllers (for example, FAS3250) and high-capacity SATA drives chosen to achieve the lowest cost per gigabyte of storage. If requirements change, data can be nondisruptively moved within the cluster to optimize for performance or capacity. For example, when a high-performance project such as a semiconductor design reaches completion, it can be moved to lower cost storage. When it is time to do the next revision of the chip, the project can be moved back into high-performance storage.

Homogeneous clusters can also be built to maximize performance or capacity for various applications. The later section on scalability and performance includes examples of homogeneous clusters.

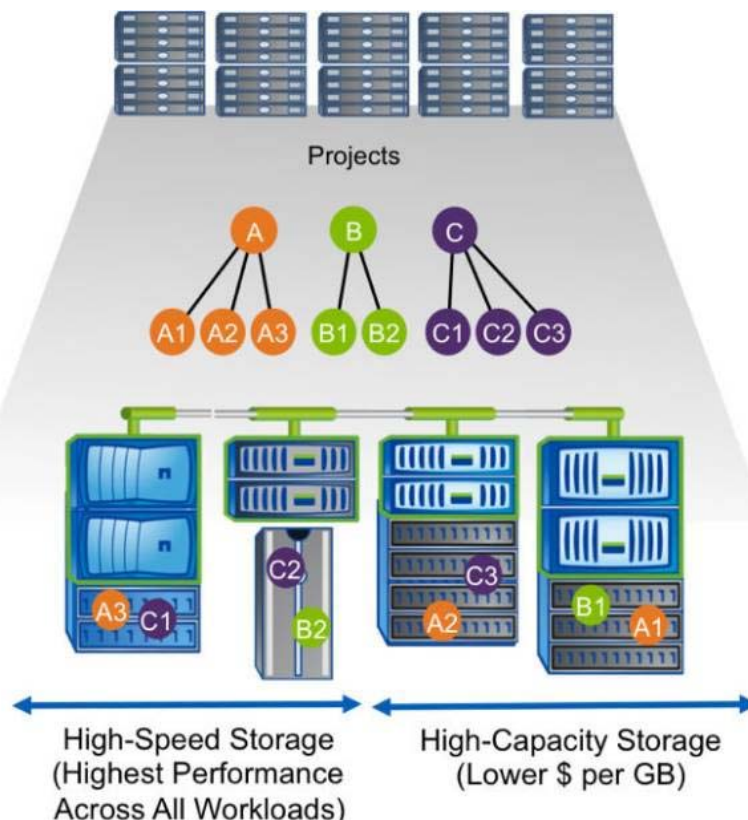


Figure 1) A single heterogeneous storage cluster can include both high-performance and high-capacity storage options to meet the requirements of different workloads. A single namespace can provide many classes of service.

Nondisruptive Operations

As business operations become more and more dependent on IT services, downtime—both unplanned and planned—becomes increasingly unacceptable. Downtime can result in lost business, poor customer satisfaction, and competitive weakness. Storage infrastructure must be always on and data always accessible. Nondisruptive operations are an integral part of clustered Data ONTAP, allowing your storage infrastructure to remain up and serving data during hardware and software maintenance and refresh operations.

When it is time to refresh your hardware, you can use aggregate relocate (new in clustered Data ONTAP 8.2) to upgrade your storage controllers without moving data, or you can move data from one HA pair to another within the storage cluster nondisruptively. Either method allows you to retire old hardware from the cluster without ever taking the data offline.

The ability to move individual data volumes, known as DataMotion™ for Volumes, allows data to be redistributed across a cluster at any time and for any reason. DataMotion is transparent and nondisruptive to NAS and SAN hosts. This rolling-refresh approach means that you can manage, upgrade, and service your storage infrastructure nondisruptively over the life of your data—even during business hours.

The specific hardware and software maintenance operations that can be performed nondisruptively in a clustered environment include upgrading from one version of Data ONTAP to another; upgrading firmware (system, disk, and switch); replacing a failed controller or component within a controller (for example, HBA, NIC); and replacing failed cables, drives, and I/O modules.

In addition, with Data ONTAP 8, adding storage controllers or shelves to a cluster, adding HBAs and Flash Cache, and upgrading components can be done nondisruptively. You can also redistribute data across controllers to improve performance, move data across controllers to rebalance capacity, and redistribute data across storage tiers within a cluster to optimize disk performance.

Multi-Tenancy

It is important for separate customers in public clouds and different business units in private clouds to be securely isolated at the compute, network, and storage layers. Data ONTAP 8 clustering provides multi-tenancy at the storage layer by isolating storage entities such as logical interfaces, LUNs, and volumes within an SVM.

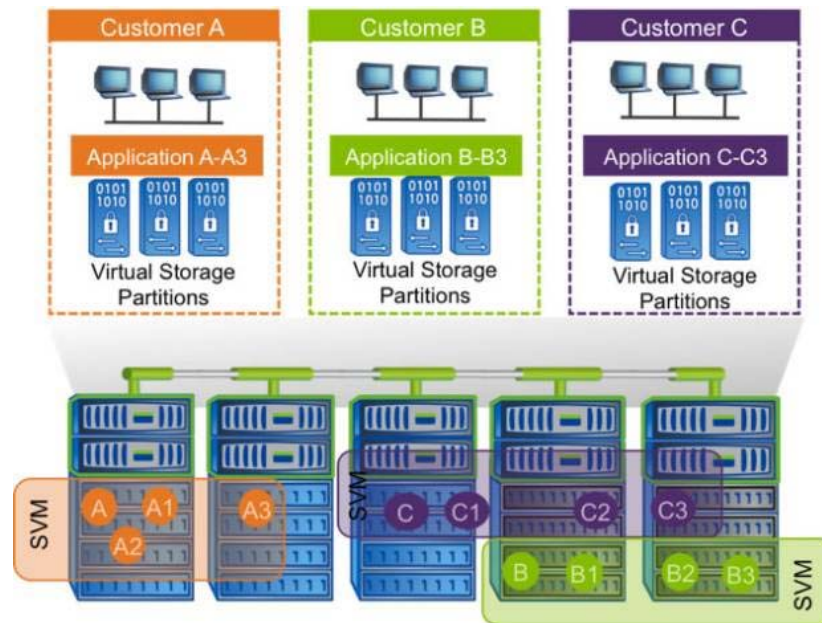


Figure 2) Clusted Data ONTAP uses Storage Virtual Machines to provide multi-tenancy.

An SVM is a secure, virtualized storage container that includes its own administration security, IP addresses, and namespace. An SVM can include volumes residing on any node in the cluster. Data ONTAP clustering supports from one to hundreds of SVMs in a single cluster.

Each SVM enables one or more SAN (FC, FCoE, iSCSI) and NAS (NFS, pNFS, CIFS) access protocols and contains at least one volume and at least one logical interface, or LIF. LIFs present either an IP address (which will be used by NAS clients and iSCSI hosts) or a WWN (for FC and FCoE access). Each LIF is mapped to a home port on a NIC or HBA. LIFs virtualize NIC and HBA ports. Each SVM requires its own dedicated set of LIFs, and up to 128 LIFs can be defined on any cluster node. Because each SVM operates in its own namespace, each customer mapped to an SVM is completely isolated. The administration of an SVM can also be delegated if desired, so that separate administrators can be responsible for provisioning volumes and other SVM-specific operations, a capability that is particularly valuable in multi-tenant environments.

You can use clustered Data ONTAP quality of service (QoS), new in clustered Data ONTAP 8.2, to set performance limits on SVMs, volumes, LUNs or files. QoS lets you limit the bandwidth or IOPS that can be consumed by a particular SVM or by an application workload.

Scalability and Performance

Currently as of clustered Data ONTAP 8.2, NAS configurations can scale up to 24 nodes, while SAN configurations scale up to 8 nodes. NetApp has published both NAS and SAN (file and block) performance numbers based on industry-standard benchmarks.

NAS Performance

NetApp evaluated clustered Data ONTAP file services performance using the SPECsfs benchmark. A cluster of FAS6240 nodes demonstrated linear scaling as nodes are added with a maximum performance of over 1.5 million SPECsfs2008_nfs.v3 ops/sec using 24 nodes. This performance shows that clustered Data ONTAP has the scalability and performance to accommodate any workload. More details of these results have been described in a [previous Tech OnTap® article](#).

SAN Performance

NetApp also submitted a 6-node FAS6240 unified storage cluster for Storage Performance Council SPC-1 block performance testing running Data ONTAP 8.1.1. The results for this SPC-1 benchmark were ~250,000 SPC-1 I/O operations per second (IOPS), \$6.69 \$ per I/O operations per second (\$/IOPS), and a least response time (LRT) of 0.99 milliseconds. These results demonstrate that this modular scale-out model provides a foundation for continued growth as both controller performance and node count increase over time.

In terms of relative performance, the 6-node FAS6240 is in the top 10% of submitted configurations as measured by LRT. The IOPS result represents a 267% increase in performance and 12% reduction in cost relative to our published FAS3270 SPC-1 result. Further, NetApp uses list pricing in its SPC-1 submissions, while almost all other vendors use discounted pricing, resulting in a \$/IOPS that is both more conservative and verifiable than that of the competition.

Infinite Volume Technology for Enterprise Content Repositories

As organizations generate petabytes of data, it becomes increasingly challenging to store, manage, and retrieve content in a flexible and efficient manner. Additionally, many organizations are faced with the need to keep data for long periods of time, often measured in decades, while retaining the ability to actually find data, regardless of where it is stored. NetApp created Infinite Volume specifically to address the scalability needs of enterprise content repositories; it eliminates the complexities of structuring data in multiple small containers.

NetApp Infinite Volume is a dedicated cluster running Data ONTAP 8.1.1 or later. It provides a single NFSv3 mount point that can scale up to 20PB or 2 billion files, all contained in a single SVM.

An Infinite Volume is a compound volume in which data is distributed across multiple constituent volumes (which we refer to as constituents) spread across the nodes of the cluster. The namespace hierarchy is stored in a single active namespace constituent volume for the entire content repository. NFS clients see the content of this volume. All metadata lookups (directory scans, file opens, get attributes, and so on) are performed on the namespace constituent volume. Subsequent reads and writes go directly through the node that "owns" the data constituent that contains the file being accessed. Data is automatically load balanced across Infinite Volume at ingest.

Infinite Volume provides simplified management using OnCommand® System Manager 2.1. Snapshot copies for data protection and replication purposes are created at the Infinite Volume level and are coordinated across all constituents in the repository to provide data consistency.

Infinite Volume provides all the storage resiliency and high-availability features of a Data ONTAP cluster, including nondisruptive operations and advanced storage efficiency features.

Simpler data management using a single mount for multi-PB data store

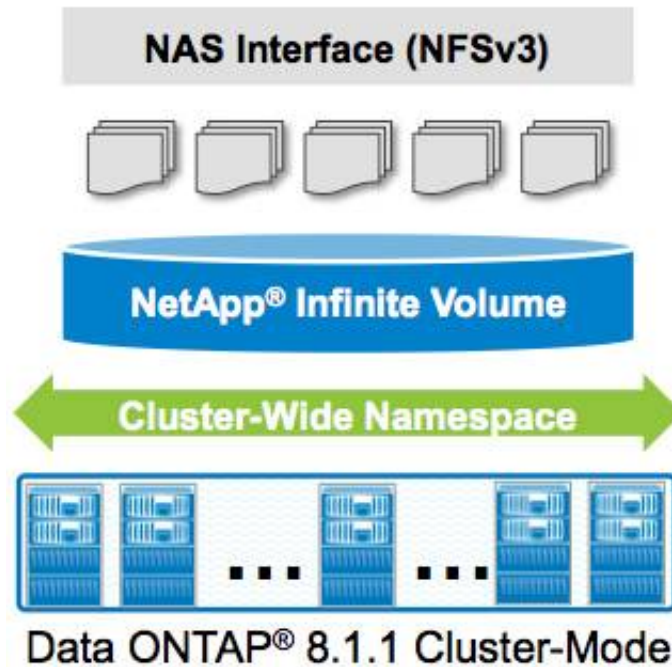


Figure 3) NetApp Infinite Volume is an enabling feature for an enterprise content repository.

Flash Pool Technology

With the release of Data ONTAP 8.1.1, NetApp has added a new Flash Pool technology to Data ONTAP to further boost the scalability and performance of both 7-Mode and clustered configurations. Flash Pool is supported on all NetApp storage systems, including entry-level systems. No other vendor offers this type of functionality for the entry-level storage market.

Flash Pool is a persistent, aggregate-level read and write cache. It lets you add RAID groups consisting of SSDs to an aggregate containing HDDs with the goal of delivering performance comparable to that of an SSD-only aggregate while keeping cost closer to that of an HDD-only aggregate. A relatively small number of SSDs in an aggregate is used as a persistent cache to accelerate both random reads and writes.

Flash Pool is part of the NetApp [virtual storage tier \(VST\)](#) and operates in a manner that is in many respects similar to NetApp [Flash Cache](#). Flash Pool shares the same 4KB granularity, works in real time, is fully automatic, and works in conjunction with NetApp storage efficiency and data protection technologies. In addition, Flash Pool adds the ability to cache randomly written data and also provides consistent performance during failover and takeover operations because the aggregate-level data cache remains accessible and available during these events. Flash Pool and Flash Cache can coexist on the same system, and existing aggregates can be converted nondisruptively to utilize Flash Pool.

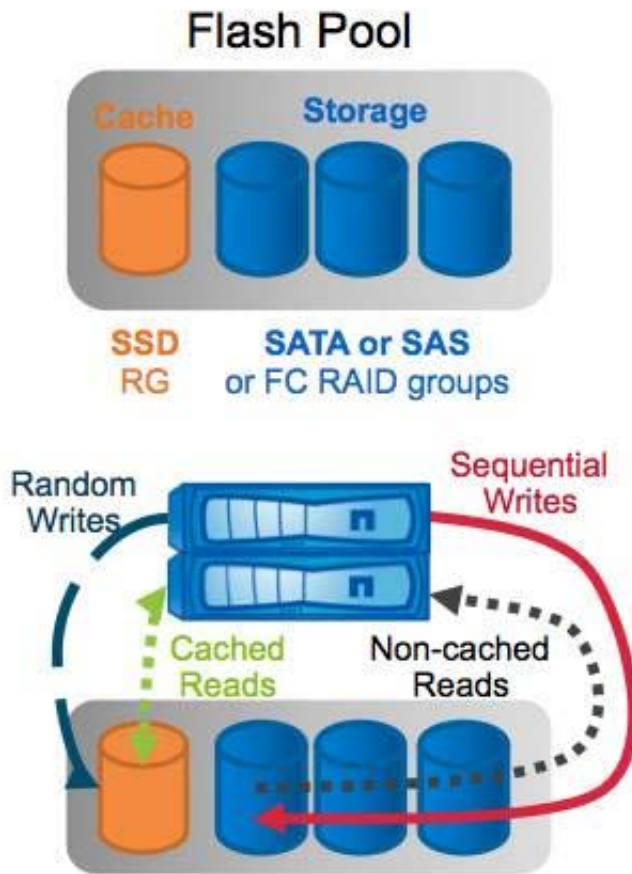


Figure 4) NetApp Flash Pool technology.

Conclusion


In addition to high performance and low latency, clustered Data ONTAP 8 delivers the full NetApp storage efficiency and data protection portfolio, nondisruptive operations, and complete support for secure multi-tenant and cloud environments.

Whether you choose FAS systems or V-Series, clustered Data ONTAP 8 gives you "always on" reliability for nondisruptive operations, tremendous flexibility to get ahead of market changes, and the operational efficiency you need to grow your business. Broad partner ecosystem integration helps to further enable your success.

Check out the recent article, ["What's New in Clustered Data ONTAP 8.2?"](#) to learn about enhancements that have been made to clustered Data ONTAP since this article was published. -Editor

› **Got opinions about Cluster-Mode?**

Ask questions, exchange ideas, and share your thoughts online in NetApp Communities.



Mike has over 23 years of computer industry marketing experience, 17 years of which have been specifically focused on storage. He worked at Adaptec, EMC, and Hewlett Packard before joining NetApp more than six years ago. Mike has been the marketing chairperson for the Fibre Channel Industry Association (FCIA), a member of the Ethernet Alliance, and a frequent speaker at industry events.

Quick Links

- › [Tech OnTap Community](#)
 - › [Archive](#)
 - › [User Groups](#)
 - › [PDF](#)
-