



FAS6200 Running Data ONTAP 8 Cluster-Mode: Performance and Scalability



Chris Lemmons

Senior Manager
Workload Engineering



Bhavik Desai

Performance Engineer
NetApp

On November 1, 2011 NetApp set a record on the SPECsfs2008 NFS benchmark using NetApp® FAS6200 series systems running Data ONTAP® 8 in Cluster-Mode. NetApp achieved over 1.5 million SPECsfs2008_nfs.v3 ops/sec for a 35% performance increase over the previous record holder while using less than half the infrastructure.

This article explains how Data ONTAP delivers HPC-level performance in conjunction with the rich feature set that you expect from NetApp. We'll begin with a little background on Cluster-Mode and then discuss the benchmark work we did that demonstrates not only great top-line performance but also great scalability as you grow a cluster.

Cluster-Mode Fundamentals

As you might already know, Data ONTAP 8 merges the capabilities of Data ONTAP 7G and Data ONTAP GX into a single code base with two distinct operating modes: 7-Mode, which delivers capabilities equivalent to the Data ONTAP 7.3.x releases, and Cluster-Mode, which supports multicontroller configurations with a global namespace and clustered file system. As a result, Data ONTAP 8 allows you to scale up or scale out storage capacity and performance in whatever way makes the most sense for your business.

With Cluster-Mode the basic building blocks are the standard FAS or V-Series HA pairs with which you are already familiar (active-active configuration in which each controller is responsible for half the disks under normal operation and takes over the other controller's workload in the event of a failure). Each controller in an HA pair is referred to as a cluster "node"; multiple HA pairs are joined together in a cluster

Explore

NetApp Versus the Competition

Two other companies have posted recent SPECsfs2008 NFS benchmark results in excess of 1 million operations per second—one used an HPC-focused solution, the other a cache appliance in front of third-party storage. However, we believe that only NetApp can deliver the winning combination of unified storage with HPC-level performance, scalability, storage efficiency, and integrated data protection that enterprise environments require.

You can check out [NetApp TR-3990](#) for a more detailed comparison of the NetApp result versus one recent published result. [A recent blog post by Dimitris Krekoukias](#) of NetApp also provides a deep dive into SPECsfs results.

What's the Buzz?

Check out what people inside and outside NetApp had to say about the NetApp performance announcement:

- [Official Press Release](#)
- [Val Bercovicci Blogpost: HPC Comes to Enterprise IT](#)
- [Dimitris Krekoukias Blogpost: NetApp Posts World-Record SPECsfs NFS Benchmark result](#)
- [Storage Without Borders Blogpost: Breaking Records...Revisited](#)
- [Vaughn Stewart Blogpost: NetApp Takes SPECsfs2008 Benchmark to New Levels](#)

using a dedicated 10 Gigabit Ethernet (10GbE) cluster interconnect. This interconnect is redundant for reliability purposes and is used for both cluster communication and data movement.

The building blocks you use in Cluster-Mode need not be the same, and you can build a cluster using the FAS HA pairs you already own. When it's time to scale performance, you can add a new FAS system to your cluster, and you can nondisruptively "retire" older systems from the cluster should that become necessary.

To learn more about Data ONTAP 8 Cluster-Mode, check out this [recent NetApp technical report](#).

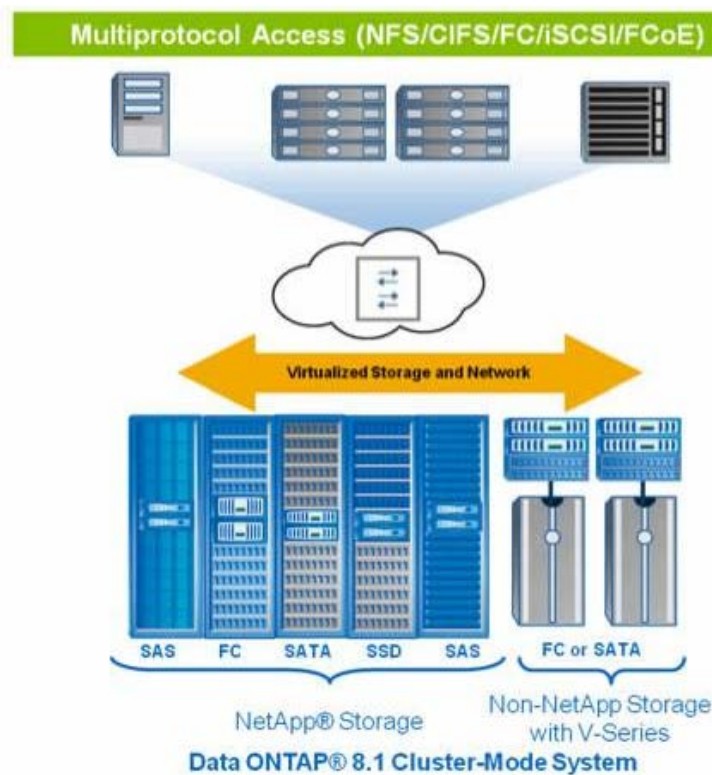


Figure 1) A Cluster-Mode system is composed of homogeneous or heterogeneous (as shown) FAS and/or V-Series storage systems and is capable of supporting all storage access protocols simultaneously.

Leading SPECsfs Results

Now that you understand the basics of Cluster-Mode, we'll explain the configuration that achieved our SPECsfs benchmark results. If you've spent any time looking at benchmarks, you know that many records are set using systems that are unlikely to ever see the light of day outside the testing laboratory. For this testing, we wanted to use a configuration that could be easily configured and operated under real-world conditions in a data center.

Our test storage system consisted of:

- 24 FAS6240 storage controllers (12 HA pairs)
- Each with a 512GB Flash Cache module
- A total of 1,728 disk drives (72 disks per controller) configured in RAID-DP® RAID groups

We chose the FAS6240—the middle platform of the FAS6200 enterprise storage series—because it represents a typical choice. The top-of-the-line FAS6280 would be expected to deliver higher performance under the same test conditions.

A single FAS6240 controller can support up to 3TB of Flash Cache, so limiting each cluster node to a single 512GB card is if anything stingy compared to what might be deployed in a typical data center configuration. It's also modest compared to the large numbers of solid-state disk drives (SSDs) deployed in some recent SPECsfs test configurations.

Similarly, the use of only 72 drives per controller again errs to the low side since a single FAS6240 controller supports up to

1,440 drives. The disks used were 450GB 15K SAS drives. Note that we had RAID-DP, the NetApp double-parity RAID 6 implementation, enabled for all testing. This is consistent with standard customer practices (RAID-DP is the NetApp default), but varies significantly from much of the SPECsfs testing done by others, which uses mirroring to avoid RAID overhead.

A graph of the SPECsfs2008 result is shown in Figure 2. Maximum throughput at the endpoint was 1,512,784 SPECsfs2008_nfs.v3 ops/sec with an overall response time (ORT) or average latency of 1.53 milliseconds (dotted line). You can see the [full NetApp SPECsfs2008 results at the SPEC Web site](#). If you are familiar with SPECsfs, you can see that the cluster demonstrates a profile very similar to what you would see for a standalone storage system. At lower throughput the cluster shows a faster response time (the time needed to satisfy a request), and the response time gradually rises as load is increased.

Fast response times have been a trademark for NetApp over the years, and this result continues that heritage, despite the fact that we made no attempt to optimize data access. In fact, our test environment was quite possibly the least optimal in this regard. Test clients requested data from any cluster node regardless of where the data actually resided. This means that on average 23 out of 24 of the requests received by any single node are for off-node data. What this means in the real world is that latency-sensitive workloads such as Microsoft® Exchange, database applications, and virtual machines should run effectively on a NetApp cluster.

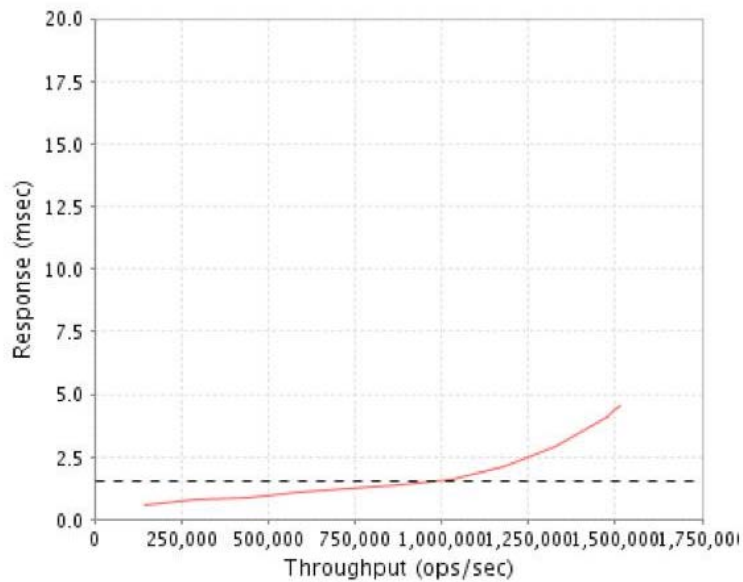


Figure 2) Data ONTAP 8.1 Cluster-Mode SPECsfs result showing throughput in SPECsfs ops/sec versus client response time. Maximum throughput was 1,512,784 ops/sec. The dashed line shows the overall response time (ORT) of 1.53 milliseconds.

Linear Scalability

Let's face it: it's unlikely that many of you have a need for a storage cluster capable of 1.5 million SPECsfs ops right now. You might, however, have a need for more performance than you can get from any single storage system. Given the current trend toward massive virtualization, cloud computing, and big data, your performance needs are probably set to grow more rapidly than you planned. With that in mind, we also wanted to test how well a NetApp cluster will scale.

In addition to the 24-node SPECsfs2008 result discussed earlier, we've also published five intermediate results using [4-node](#), [8-node](#), [12-node](#), [16-node](#), and [20-node](#) clusters of FAS6240 systems. In all cases the per-node configuration was the same as for the 24-node result shown earlier: 512MB of Flash Cache per controller node and 72 disks per controller node.

As you can see in Figure 3, performance scaling as nodes are added is extremely linear, and ORT holds steady for all configurations tested.

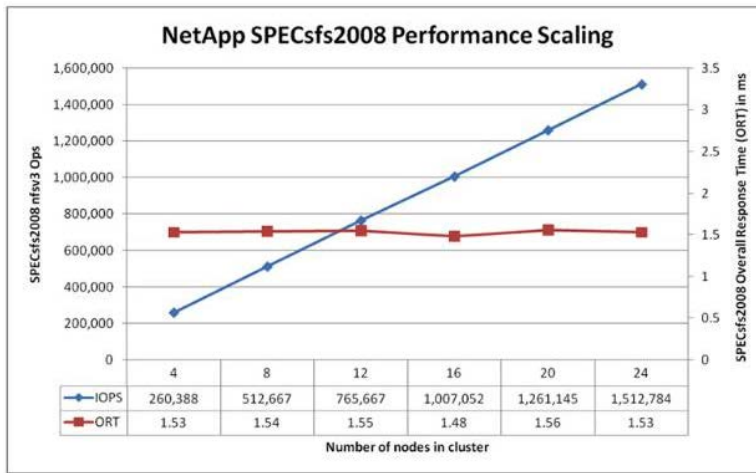


Figure 3) Cluster size versus SPECsfs2008 throughput and overall response time.

This means that you can expect great scaling as you build out a cluster over time with predictable improvements in performance and without unacceptable increases in response time that would negatively affect latency-sensitive applications.

Conclusion

The recent NetApp SPECsfs2008 NFS benchmark results demonstrate both leading performance and linear scalability. FAS6200 systems running Data ONTAP 8 Cluster-Mode bring HPC-level storage performance to mainstream IT. In addition to high performance and low latency, Cluster-Mode offers the full NetApp storage efficiency and data protection portfolio, nondisruptive operations (NDO), broad partner ecosystem integration, and complete support for secure multi-tenant and cloud environments.

Whether you choose FAS systems or V-Series, Data ONTAP 8 Cluster-Mode gives you “always on” reliability for nonstop operations, tremendous flexibility to get ahead of market changes, and the operational efficiency you need to grow your business.

➤ **Got opinions about the FAS6200?**

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



By Chris Lemmons, Senior Manager, Workload Engineering and Bhavik Desai, Performance Engineer, NetApp

Chris has devoted much of his 28 years in the IT industry to understanding the nuances of computer and storage performance, starting at Data General and continuing on at Ziff-Davis and Veritest. Since he joined NetApp 7 years ago he has focused almost exclusively on storage performance, with a particular emphasis on enterprise application environments including Oracle®, Exchange, and VMware®.

Bhavik started at NetApp as a new college grad after completing a master's degree in computer science at Clemson. In his 3

years at NetApp he has focused on the performance of a variety of technologies with NetApp storage, including Exchange, VMware, and Hyper-V™. Recent efforts have been concentrated on industry-standard benchmarks such as SPECsfs2008 and SPC-1.



Quick Links

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