

FAS8000: Scale-Out Storage for the Enterprise



Steven Miller

Senior Technical Director and Platform Architect

The first NetApp® storage system, introduced in 1993, had a single Intel® i486 CPU and a maximum raw disk capacity of 14GB. (That's

Gigabytes.

Less total

storage than most smartphones have.) Modern FAS storage traces back to that original platform, but now scales to petabytes of capacity—an increase of over 300,000 fold—and capabilities scarcely dreamed of 20 years ago.

Which brings us to the next generation of NetApp storage. With the [FAS8000](#), we've refined and enhanced every aspect of the FAS platform, and in the process created a hybrid storage system that is uniquely well suited for the needs of today's enterprise—without leaving the past behind. The FAS8000 is designed to help you run your business operations faster while reducing management overhead, simplifying IT operations, and improving return on investment.

We've focused on flexibility, so that the FAS8000 adapts to your changing needs without ever requiring planned downtime or disruptive hardware changes. With FlexArray storage virtualization software (the subject of a [separate article in this issue of Tech OnTap](#)), the FAS8000 can also virtualize and manage existing storage arrays, extending the capabilities of the Data ONTAP® operating system to more of your storage infrastructure.

In this article, I describe the architecture of the FAS8000, focusing on the many enhancements we've made to processors, memory, flash, and I/O to deliver the performance, scalability, reliability, and flexibility you expect from FAS storage.



Explore

Learn More About the FAS8000

Ready to find out more about the FAS8000? Be sure and read the article on [FlexArray](#) in this issue, and check out the following additional resources as well:

- [FAS8000 Product Page](#) Be sure to click on the featured content from the Evaluator Group, *NetApp FlexArray Virtualization Software: Enabling Storage IT as a Service* in the upper right.
- [FAS8000 Data Sheet](#)
- [FAS8000 Press Release](#)
- [Video](#) George Kurian, EVP, Product Operations introduces the FAS8000 and FlexArray.
- [Blog: NetApp Unveils FAS8000](#)
- [Listen to the FAS8000 launch episode on the NetApp Communities Podcast](#)

What Makes FAS Hardware Different?

NetApp designs all the models in its FAS hardware line to deliver tremendous performance, scalability, and availability with superior manageability, ease of service, and ease of use. Find out how NetApp optimizes its hardware for I/O.

▶ [More](#)

Nondisruptive Operations Savings Calculator

Find out how much you can save with NetApp clustered Data ONTAP. Use this calculator to estimate the benefits of eliminating both planned and unplanned storage downtime.

Introducing the FAS8000

[More](#)

For many IT organizations, adding storage capacity can be disruptive. Most storage systems have pretty narrow limits and capabilities, so you end up with multiple storage systems, silos of storage, orphaned capacity, and greater management complexity. The FAS8000 solves this problem with unified scale-out storage that lets you scale your storage environment in the way that makes the most sense for your business needs. Flexible hybrid storage options let you provide an optimal level of acceleration for each workload.

Because the FAS8000 combines innovative hardware design with the proven capabilities of clustered Data ONTAP, industry-leading management capabilities, and unmatched support for well-known hypervisors, applications, and management and orchestration tools, it delivers the benefits of scale-out without sacrificing any of the capabilities that your operation depends on.

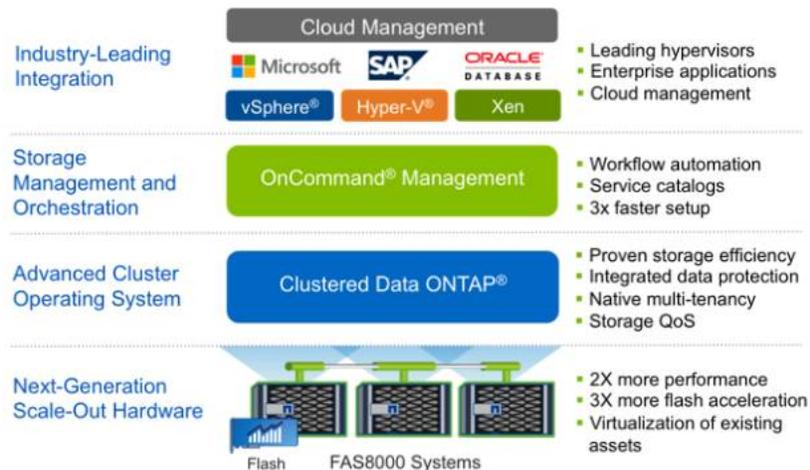


Figure 1) FAS8000 combines next-generation hardware with clustered Data ONTAP, OnCommand, and leading integration.

Scale Out and Scale Up

FAS is the only unified scale-out storage architecture. Like its predecessors, the FAS8000 supports both SAN and NAS protocols as needed, and the FAS8000 is the first FAS architecture designed specifically for clustered Data ONTAP. All FAS8000 models scale out to a maximum of 24 nodes.

The FAS8000 also offers significant scale-up capabilities. You can scale up FAS controllers as needed to meet your exact storage requirements by adding more capacity, by adding different types of media, or by installing Flash Cache™ intelligent caching or additional interface cards. You can upgrade from one model controller to another without disruption; NetApp customers have long appreciated these data-in-place "head" upgrades as a way to gain performance and capacity without disruptive data migration.

With [FlexArray](#), a FAS8000 system can also incorporate existing EMC, HDS, and NetApp E-Series storage arrays as part of your scale-out cluster without necessitating the purchase of additional hardware.

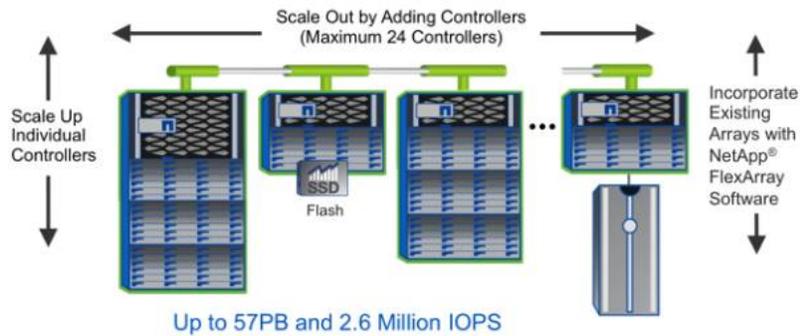


Figure 2) The FAS8000 is designed to scale out, scale up, and incorporate existing arrays.

Mix Generations and Eliminate Disruptive Tech Refresh

If you've got an existing FAS cluster, you can combine existing nodes with FAS8000 models. This means that you can continue to grow your cluster or make the transition to the latest controller technology, all without taking any downtime or disrupting important business operations. (When mixing nodes of different generations, cluster-node limits are dictated by the limits of the least capable node.)

FAS8000 Models

Although previous FAS generations were capable of scale out, the FAS8000 is the first generation designed from the ground up for that purpose. That change in mindset resulted in a product line that is simpler and more compact. You achieve the exact capabilities you need—performance, capacity, storage tiers, connectivity—by mixing and matching controllers (nodes), media types, and network connectivity in a cluster.

The FAS8000 series consists of three building blocks: the FAS8020, FAS8040, and FAS8060. The FAS8020 takes the place of the FAS3220, the FAS8040 takes the place of the FAS3250, and the FAS8060 takes the place of the FAS6220.

You can combine FAS8000 models in any combination up to a maximum of 24 nodes; clusters running SAN protocols can have up to 8 nodes. You can create a homogeneous scale-out infrastructure with identical nodes or a heterogeneous one in which nodes are of different types and/or configurations.



Figure 3) FAS8020 (3U) and FAS8040/8060 (6U) controllers.

Each FAS8000 enclosure can contain up to two controllers, providing single-enclosure HA and keeping the packaging extremely dense. As Table 1 shows, we've enhanced these platforms and added significantly more on-board I/O, minimizing the need for add-on I/O cards.

Table 1) Comparison of FAS8000 models.

	FAS8060	FAS8040	FAS8020
Maximum capacity	4800TB	2880TB	1920TB

Maximum disk drives	1,200	720	480
Controller form factor	Single-enclosure HA; 2 controllers in 6U	Single-enclosure HA; 2 controllers in 6U	Single-enclosure HA; 2 controllers in 3U
Memory	128GB	64GB	48GB
Max Flash Cache	8TB	4TB	3TB
Max Flash Pool	18TB	12TB	6TB
Max total VST flash	18TB	12TB	6TB
NVRAM	16GB	16GB	8GB
PCIe expansion slots	8	8	4
On-board I/O: 10GbE	8	8	4
On-board I/O: UTA2 (10GbE/16Gb FC)	8	8	4
On-board I/O: 6Gb SAS	8	8	4
OS version	Data ONTAP 8.2.1 RC2 or higher		



Figure 4) FAS8020 (3U) and FAS8040/8060 (6U) controllers, rear view.

FAS8000 Architecture

To drive our scale-out design to the utmost and address the need for front-end performance in parallel with back-end data management and storage efficiency functions, we took a close look at areas in the FAS architecture where clustered Data ONTAP could take advantage of additional hardware resources. Then we optimized and tuned those resources accordingly. We've significantly expanded almost every hardware metric—and increased overall reliability, availability, serviceability, and manageability (RASM).

Table 2) New FAS8000 models versus comparable previous-generation models.

	FAS8060	FAS8040	FAS8020
Takes the place of	FAS6220	FAS3250	FAS3220
Cores	2X	1X	1.5X
Memory	1.3X	1.6X	2X
Memory bandwidth	1.6X	4X	3X
NVRAM	2X	4X	2.5X
I/O bandwidth	2X	10X	10X
Flash (for VST)	1.5X	3X	3X+
Overall performance	1.7X	1.9X	2.3X

Multicore Sandy Bridge Processors

The FAS architecture has supported multicore operation for over a decade. We've steadily increased parallelization and tuned Data ONTAP to drive higher core counts as fast as Intel can deliver them.

The FAS8000 utilizes the Sandy Bridge architecture; each model offers more and/or faster cores than its predecessor. The net is a tremendous increase in the amount of compute in these platforms, which we utilize to drive simultaneous front-end and back-end work. We've worked closely with Intel on new instructions in Sandy Bridge for encryption and compression as well as Advanced Vector Extensions that allow a single instruction to be applied to multiple pieces of data.

Memory Enhancements

In addition to beefing up compute capabilities, we substantially enhanced the memory architecture, supporting more memory and increasing the memory bandwidth on each model by as much as four times over the previous model being replaced.

FAS8000 controllers deliver the maximum available memory bandwidth from each processor to provide maximum throughput for I/O requests. We've designed for a single DIMM per memory channel to deliver the maximum bandwidth to memory—as much as two times the bandwidth of some other Sandy Bridge implementations. The FAS8060 is able to deliver 85GB/sec bandwidth between memory and cores.

New NVRAM Design

At this point you're probably starting to detect a pattern. The new NVRAM9 architecture used in the FAS8000 delivers twice the bandwidth of the NVRAM8 and we've added significantly more capacity on all models. The NVRAM9 is integrated on each motherboard, so it doesn't consume any expansion slots. The contents of the NVRAM are automatically destaged to flash at the onset of a power failure so that data remains protected no matter how long an outage lasts.

The additional NVRAM on the FAS8000 has particular performance benefits for large sequential writes. Many random I/O workloads also have a sequential write component. For example, with OLTP, logging activity is sequential. Data ONTAP has been tuned to optimize performance for OLTP and similar workloads on the FAS8000. Other applications that generate significant sequential write activity include backup, media applications, satellite data acquisition, and seismic processing.

I/O: Rock-Solid PCIe Gen3

The I/O architecture is an aspect of the FAS8000 that I'm particularly proud of. All FAS8000 models use PCIe Gen3, enhancing the available I/O bandwidth. The FAS8060 offers 80 Gen3 lanes (versus 72 Gen2 lanes in the 6220) for twice the bandwidth. The FAS8040 and FAS8020 each provide 40 Gen3 lanes and up to 10 times the internal I/O bandwidth of their predecessors.

What you may not know is that PCIe Gen3 had a bit of a rocky start. The earliest adopters had a lot of difficulties, which sometimes resulted in system hangs that necessitated power-cycling. NetApp engineering spent two years working on Gen3 to find root causes and fixes for all issues to make it completely solid.

The FAS8000 uses PCIe Gen3 throughout, with expansion slots for additional I/O cards and Flash Cache. The FAS8020 offers two slots per controller (four per HA pair) while the FAS8040 and FAS8060 offer four slots per controller (eight per HA pair).

On-Board Ports Deliver Greater Flexibility

A critical aspect of the I/O architecture of the FAS8000 is the number and variety of on-board ports. This gives you much greater user configurability right out of the box while preserving PCIe slots. Past FAS models had more

rigid, fixed configurations.

Each FAS8000 controller includes four major types of on-board ports: GbE, SAS, 10GbE, and Unified Target Adapter 2 (UTA2). The FAS8020 includes two of each port type per controller (four of each per HA pair) while the FAS8040 and FAS8060 include four of each port type per controller (eight per HA pair).

Industry-first, flexible UTA2 ports. Our new UTA2 ports are an industry first. By changing the optics and setting a software switch, these ports can deliver 16Gb/sec Fibre Channel—the first FAS platform to do so—or 10GbE. We think this gives you greater flexibility now and in the future.

You can connect to existing FC SANs with the highest available FC bandwidth; 16Gb FC is also optimal for connecting to and virtualizing existing storage with [FlexArray](#). If you don't have a current requirement for FC, you can configure the ports for 10GbE to support FCoE and iSCSI SAN connectivity and NFS and CIFS/SMB client connectivity, so nothing is wasted.

New UTA2 cards. If the flexibility of these UTA2 ports intrigues you, you'll be happy to know that NetApp is also releasing a new two-port PCIe Gen3 adapter with the same functionality. You can add these to your FAS8000 for even more connectivity, and they are also supported on the FAS3220, FAS3250, and all FAS6200 models.

On-board 10GbE. On-board 10GbE ports are used for the cluster interconnect or for client and host network connections. The FAS8020 uses dual cluster interconnects while the FAS8040 and FAS8060 support dual or quad interconnects for even greater cluster bandwidth.

When combined with our on-board UTA2 ports, the FAS8000 is capable of delivering a tremendous number of 10GbE ports and driving those ports at full bandwidth, delivering the performance and flexibility you need for mission-critical applications and cloud connectivity.

New 10GBase-T cards. If you need even more 10GbE connectivity, you can add UTA2 adapters as mentioned above or one of our existing 10GbE NICs. We've also added a new 10GBase-T NIC to our portfolio so you can connect your FAS systems using CAT 6A (or higher) cabling. These ports autorange between 10GbE and GbE speeds.

Flash Support

NetApp has made tremendous enhancements to our flash offerings in recent years; the majority of new FAS systems in the FAS8000 class ship in hybrid storage configurations, combining flash for performance and HDD for capacity. This means that, as with everything else, we both needed and wanted to enhance the flash capabilities of the FAS8000.

Hybrid storage. As you probably already know, our NetApp virtual storage tier (VST) portfolio includes several options that let you use flash as a high-speed cache for hot data in combination with traditional HDDs. This includes Flash Cache accelerator cards that can be installed in each FAS8000 controller, and Flash Pool configurations that combine HDDs and SSDs in the same NetApp aggregate.

We've significantly increased the amount of Flash Cache and Flash Pool we support in FAS8000 models. Up to three times more flash allows you to accelerate more workloads and further optimize your storage configuration for specific requirements.

All-flash aggregates and all-flash configurations. In addition to VST flash acceleration, FAS8000 systems also support all-flash aggregates composed entirely of SSDs. All-flash aggregates can be deployed along with disk-only aggregates and Flash Pool aggregates. FAS8000 systems can also be deployed in all-flash configurations. This gives you tremendous flexibility to provide multiple tiers of storage from a single storage system or from a scale-out cluster.

All-flash options deliver optimal performance for those applications that require the absolute lowest latency for every transaction.

Full range of SSD options. Whether you're deploying Flash Pool or all-flash options, NetApp offers a broad portfolio of SSDs, including 200GB, 400GB, 800GB, and 1.6TB capacities. This makes it easier to configure the right amount of flash for your specific needs. Our new automated workload analyzer for Data ONTAP 8.2.1 helps you figure out exactly how much cache you need before deploying Flash Pool, making it simpler to choose the optimal SSD configuration to support your workloads.

Reliability, Availability, Serviceability, and Manageability (RASM)

All FAS8000 models are designed to deliver 99.999% or better availability. As you would expect, the FAS8000 builds on the RASM features of previous generations, including alternate control path for disk shelves and NVRAM destaging as described above. Each controller has an integrated service processor that is isolated from the main controller logic and remains operational even when the rest of the board is down.

Full processor diagnostics. The service processor can be accessed via Ethernet or the serial console for troubleshooting if a problem does occur. For instance, we can extract the complete state of a processor, even if it's locked up. Sandy Bridge provides even more information than previous generations of Intel processors, giving us greater forensic capabilities than ever before, both from processors and the various I/O devices embedded on the motherboard or installed in PCIe slots.

FAS8000 Performance

Now that you know how we've enhanced the hardware in each FAS8000 model, the obvious question is, how do they perform? We've released several benchmark results for the FAS8020 and FAS8040, and of course we've done an extensive characterization of FAS8000 performance internally.

I want to digress slightly to talk about benchmarks for a moment. With the recent popularity of all-flash arrays, IOPS numbers are being thrown around quite freely, often without qualifying the workload used to achieve a given result. The eye-popping numbers you sometimes hear typically correspond to "easy" workloads, consisting entirely of 512B to 4KB reads; in many if not most cases, the actual workload isn't specified. This makes it difficult to make meaningful comparisons when talking about performance.

By comparison, [SPECsfs](#) and [SPC-1](#) are independent, closely monitored benchmarks based on workloads designed to replicate real-world applications. A significant portion of each benchmark is made up of writes, and numbers in excess of 1M IOPS are rare, even with all-flash configurations. The official [NetApp result for the 24-node FAS6240](#) of over 1.5M SPECsfs2008_nfs.v3 ops/sec released in late 2011 remains a leader among the major storage vendors on the SPECsfs NFS benchmark.

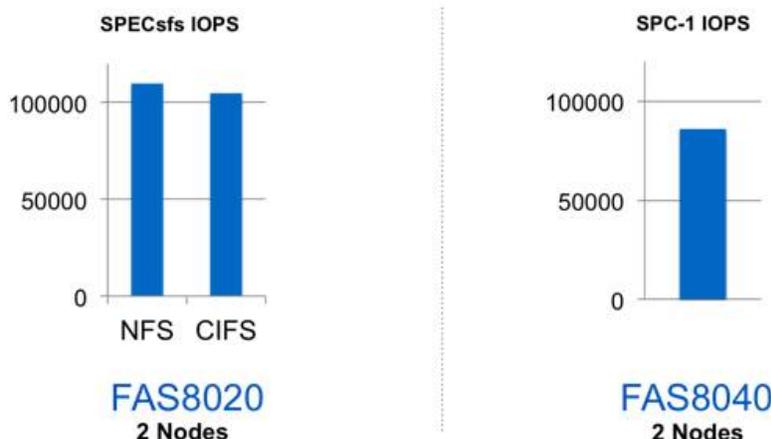


Figure 5) FAS8000 SPECsfs and SPC-1 performance results.

SPECsfs

For the FAS8020, we ran both the NFS and CIFS versions of the SPECsfs benchmark. The configuration was the same in both cases: a 2-node FAS8020 cluster with 144 disks and a 512GB Flash Cache 2 card in each controller.

These are extremely strong results for the FAS8020 and indicate how the FAS8000 line will excel in terms of both scale-up and scale-out, delivering great throughput and extremely low latency.

For NFS, the FAS8020 achieved 110,281 SPECsfs2008_nfs.v3 ops/sec with an overall response time (ORT) or average latency of 1.18 milliseconds.

- This 2-node result matches the ops/sec of a 14-node configuration (seven times more controllers) from a scale-out competitor, while the competitor's ORT was almost three times higher.
- This configuration delivers submillisecond latency up to 65K ops/sec and stays under 2ms up to almost 100K ops/sec, demonstrating latencies approaching those of all-flash storage.

For CIFS, the FAS8020 achieved 105,050 SPECsfs2008_cifs ops/sec at 1.42ms ORT:

- This 2-node result matches the ops/sec of a 7-node configuration from a scale-out competitor, while the competitor's ORT was more than two times higher.
- This configuration delivers submillisecond latency up to 50K ops/sec and stays under 2ms up to almost 85K ops/sec.

SPC-1

We tested a 2-node FAS8040 on the SPC-1 benchmark. That configuration included 192 disks in total with a 512GB Flash Cache 2 card in each controller and achieved 86,072 SPC-1 IOPS.

- A FAS8040 competitor recently achieved 109,986 SPC-1 IOPS, but the configuration required more than four times the spindles (896 disks) to achieve only 30% higher performance. (http://www.storageperformance.org/results/benchmark_results_spc1#a00141. Data current as of February 19, 2014.)
- Many recent SPC-1 results have focused solely on performance using SSD-only configurations, but in many cases provide just a few TB of capacity.

These FAS8040 results clearly demonstrate the great performance and low latency possible with FAS8000 hybrid storage configurations, proving that if you need performance and capacity, hybrid storage is likely still the best option for a wide range of real-world applications.

Linear Scaling to Millions of IOPS

Based on our internal testing, we believe that the FAS8000 will demonstrate performance scaling that is extremely linear. That is, if a 2-node configuration delivers X, 4 nodes will deliver 2X, 6 nodes will deliver 3X, and so on. In addition, we've designed the FAS8020, FAS8040, and FAS8060 so they are evenly spaced in terms of performance. The FAS8060 delivers roughly twice the performance of the FAS8020, with the FAS8040 falling in the middle. Internal testing suggests that a FAS8060 cluster will deliver in excess of 2.6 million IOPS using a typical NFS workload.

With linear scalability, you can choose the best building blocks for a scale-out cluster today and be confident that you can add capacity and performance in the right increments for your needs and budget as your requirements grow.

Storage for Today—and Tomorrow

NetApp is excited by the capabilities of the FAS8000. These flash-accelerated, cloud-integrated enterprise storage systems deliver tremendous I/O to applications and end users while simultaneously executing the data management tasks critical to the success of your IT operations and your business. We've focused on delivering great flexibility in a more compact form factor to create storage building blocks uniquely well suited to today's changing IT environment and that are able to grow and adapt to meet your future needs without disrupting your business.



By Steven Miller, Senior Technical Director and Platform Architect

Steven has been the platform architect for NetApp for over six years and was responsible for the FAS3100, FAS3200, FAS6200, FAS2240, FAS2220, and FAS8000 as well as the Performance Acceleration Module (PAM) and Flash Cache (PAM II). He is also the NetApp engineering liaison to the National Security Agency, National Geospatial-Intelligence Agency, and Central Intelligence Agency. Steven is currently involved with several IEEE and industry groups. He is credited with 38 issued patents and 16 pending applications in the areas of storage and high-performance computing.

Quick Links

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