

# SolidFire: All-Flash for the Next Generation Data Center



**Dave Wright**

SolidFire Founder, Vice President and GM

*In February 2016 NetApp completed its acquisition of SolidFire, a market leader in all-flash storage systems built for the next-generation data center. Tech OnTap is pleased to welcome SolidFire founder, Dave Wright, to introduce our readers to this technology. Dave started SolidFire in 2010—the third company he founded—to build a unique flash storage architecture that delivers the performance, automation, and scale to advance the way the world uses the cloud.*

Public-cloud architecture is making its way into enterprise and service provider data centers, creating a new set of challenges for administrators. To be successful with the cloud model—with dynamically allocated pools of compute, networking, and storage—infrastructure has to be extremely cost effective and deliver scalability, automation, and support for multi-tenancy and mixed workloads. These are the core principles of the SolidFire design.

**March / April 2016**

## **Additional SolidFire Resources**

### **Videos, Blogs, and Articles**

- [SolidFire and NetApp](#) (Dave Wright, video)
- [SolidFire + NetApp = Unstoppable Force](#)
- [SolidFire: The Perfect Addition to NetApp's Flash Portfolio](#) (Dave Hitz, blog)
- [NetApp and SolidFire: Flash to Meet All Your Data Management Needs](#) (Lee Caswell, interview)
- [The Beginning of the End of the All-flash Array Market](#) (Dave Wright, blog)

### **Podcasts**

The *Tech OnTap Podcast* team has been working hard to bring you the latest on SolidFire:

- [Episode 25: NetApp has Acquired SolidFire](#) (Andy Grimes)
- [Episode 26: SolidFire Deep Dive](#) (Aaron Delp, Amy Lewis, Jeremiah Dooley, Andy Banta)
- [Episode 27: Infrastructure Automation](#) (Josh Atwell)
- [Episode 28: SolidFire Virtualization](#) (Aaron Patten)

### **Storage Field Day Videos**

Dave Wright and the [SolidFire team](#) presented at Storage Field Day 5 including:

- [SolidFire Solution Overview](#)
- [Comparing Modern All-Flash Architectures](#)

## **Ultimate Software Chooses SolidFire**

Ultimate Software is a leading provider of human capital management (HCM) solutions delivered via Software as a Service (SaaS). More than 2,800 customers operating in 150 countries count on Ultimate's UltiPro® suite.

Ultimate has four main data centers in North America, built on a pod architecture with 200-500 customers per pod. Plagued by storage performance issues, Ultimate turned to SolidFire, originally purchasing five SF3010 nodes, and quickly adding 66 SF2405 nodes, with plans to purchase an additional 44 nodes.

In addition to performance, SolidFire was chosen for its security, availability, and integration with both OpenStack and VMware. Pods supported by SolidFire deliver 4x the performance, while QoS eliminates the noisy neighbor problems seen in other pods.

Read the [Ultimate Software case study](#).

# All-flash storage platform for the next generation data center



If you're a long-time NetApp user, you'll be curious to know more about SolidFire. This article explains some of the technology choices that make its design different from other all-flash arrays.

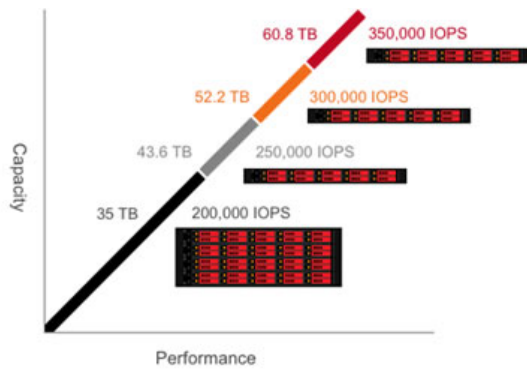
## Scale-Out, Shared Nothing Architecture

Most all-flash storage systems—including Pure and XtremIO—use a dual-controller design to protect against failure. Two controllers share access to a set of drives, and one controller takes over for the other in the event of a failure; data is protected with some type of RAID.

The SolidFire design takes a different approach, using a scale-out, shared nothing architecture. Each SolidFire node is a standard 1U x86 system with 10 internal MLC or TLC SSDs. Nodes are interconnected via 10GbE, and nothing is shared between nodes. Hosts access data via either iSCSI or Fibre Channel block protocols.

A SolidFire cluster starts with 4 nodes and can scale out incrementally to 100 nodes. Data is automatically distributed across all nodes in a cluster, so each additional node expands performance and capacity linearly. A storage volume is never constrained by the performance limits of a single controller.

**Figure 1)** SolidFire delivers linear scale-out of performance and capacity in a shared nothing design.



- 4 – 100 nodes
- 35TB – 3.4PB Usable Capacity
- 200k – 7.5M Controllable IOPS

Source: SolidFire 2016

This approach has significant advantages:

**Nondisruptive scale-out / scale-in.** Add or remove nodes without disrupting service or compromising Quality of Service (QoS). Data is automatically redistributed in the background across all nodes, maintaining balance as the system grows.

**Instant resource availability.** Newly added storage resources are instantly available to every volume within the system, eliminating the need to reallocate volumes over new drives.

**Ability to mix nodes.** Some scale-out systems require nodes to be identical. SolidFire gives you the ability to mix nodes of different types and generations. A selection of performance and capacity points lets you scale to match your needs.

**Simpler capacity planning.** Because scale-out occurs in 1U increments, performance and capacity can be added in a very granular fashion. This eliminates reliance on multiyear capacity and performance projections. It also eliminates upfront over-provisioning, allowing you to take advantage of price reductions over time.

**No forklift upgrades.** New-generation nodes can simply be added to an existing cluster. When the time comes, old nodes can be removed and retired or repurposed. Compatibility between storage nodes is guaranteed. Each time you add a node, you are able to add the most up-to-date technology.

## Data Assurance

To provide data redundancy, SolidFire maintains two copies of every data block on two separate nodes—a

technology called Helix, that is built into our Element OS operating software. This allows a cluster to sustain application performance after failures occur. It also takes away the need to have separate storage shelves with shared drive access, making the hardware less complex and less expensive. For instance, SolidFire uses single-attach SSDs instead of more expensive enterprise-grade, dual-attach SSDs.

The system self-heals quickly, reducing the risk that a second failure will occur before redundancy is restored. Because a cluster responds gracefully to nodes going offline, the feature also facilitates nondisruptive hardware and software upgrades.

## Self-Healing from Failures

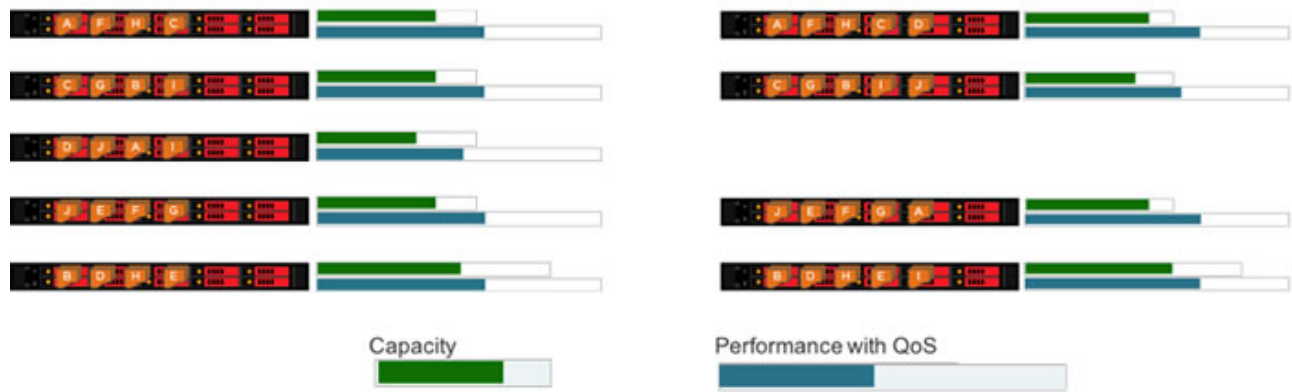
All resources in the system are always in the active pool; there is no need to have spare drives or spare nodes sitting idle in case of failure.

**Drive failure.** If a drive fails, the system automatically restores full redundancy by redistributing copies of data using a meshed rebuild process. There is no degraded mode operation and no performance penalty during a rebuild. The process typically completes in 5 minutes or less. Because of the speed with which full redundancy is restored, it provides a level of data protection that exceeds RAID-6 in a typical system.

**Node failure.** Because data copies are distributed on separate nodes, all data remains accessible if a node fails. Connections to the failed node are automatically redirected to other nodes. As with a drive failure, full redundancy is restored quickly and automatically by making sure there are two copies of each block.

No matter the failure mode—drive, node, backplane, network failure, software failure—the recovery process is the same. Because the recovery workload is distributed across all nodes in the cluster, redundancy is restored quickly, and no single node (or application workload) takes a performance hit. The more nodes in the cluster, the faster the activity occurs and the lower the overall impact.

**Figure 2)** After a node failure, data redundancy is restored by distributing new copies of all blocks from the failed node across the surviving nodes. Performance and capacity utilization increase evenly across all nodes.



Source: SolidFire 2016

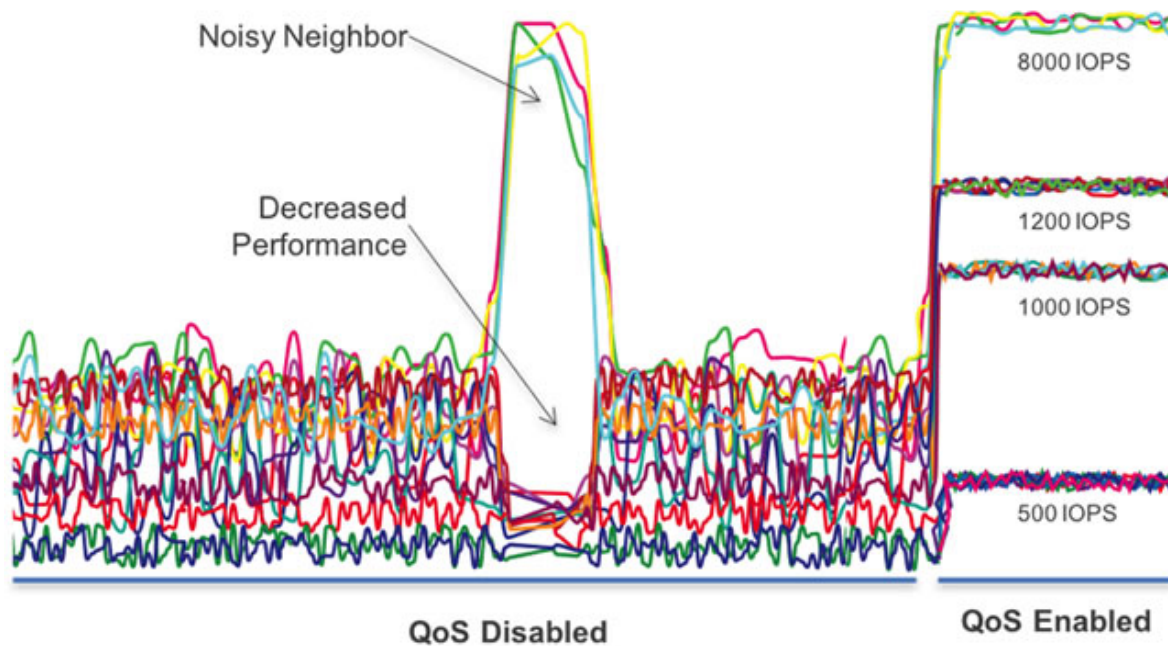
## Guaranteed Performance

To support mixed application workloads and multi-tenant environments on a single cluster, SolidFire provides guaranteed QoS. Unlike implementations that provide QoS on a best-effort basis, SolidFire is able to guarantee performance to each workload.

You can allocate performance and capacity independently for every volume in a system. When you create a volume, you simply set the desired size, and specify three QoS parameters: **Min**, **Max**, and **Burst**. If you change the settings on a volume, it will immediately start receiving service at the new levels.

The **Min** setting defines a minimum level of performance measured in IOPS (weighted by I/O size). The volume is guaranteed to deliver at least that level of performance under all circumstances. The **Max** setting defines how many IOPS a volume can consume. Because hard rate limits can create problems for applications—a transient VDI boot storm is a good example—there is also the **Burst** parameter. Applications build up credits when they run under their maximum limit, allowing them to burst for short periods when necessary.

**Figure 3) Guaranteed QoS.** The left side illustrates the effect of “noisy neighbors” with QoS disabled. A few poorly behaved workloads rob performance from everything else. On the right, the effect of enabling QoS with various settings for different workloads can be seen.



Source: SolidFire 2016

A study from Enterprise Strategy Group (ESG) estimates that SolidFire with guaranteed QoS can eliminate up to 93% of traditional storage-related issues—including problems caused by workload imbalances, monopolization of a fixed set of resources, insufficient resources in a pool, moving VMs, inefficient tiering, and controller bottlenecks. The study concluded that guaranteed QoS and automated load balancing allow an organization to consolidate a greater variety and volume of workloads on a single storage system. With traditional storage and no QoS, you would either spend more time addressing performance issues or over-provision storage arrays to minimize problems.

## Automated Management

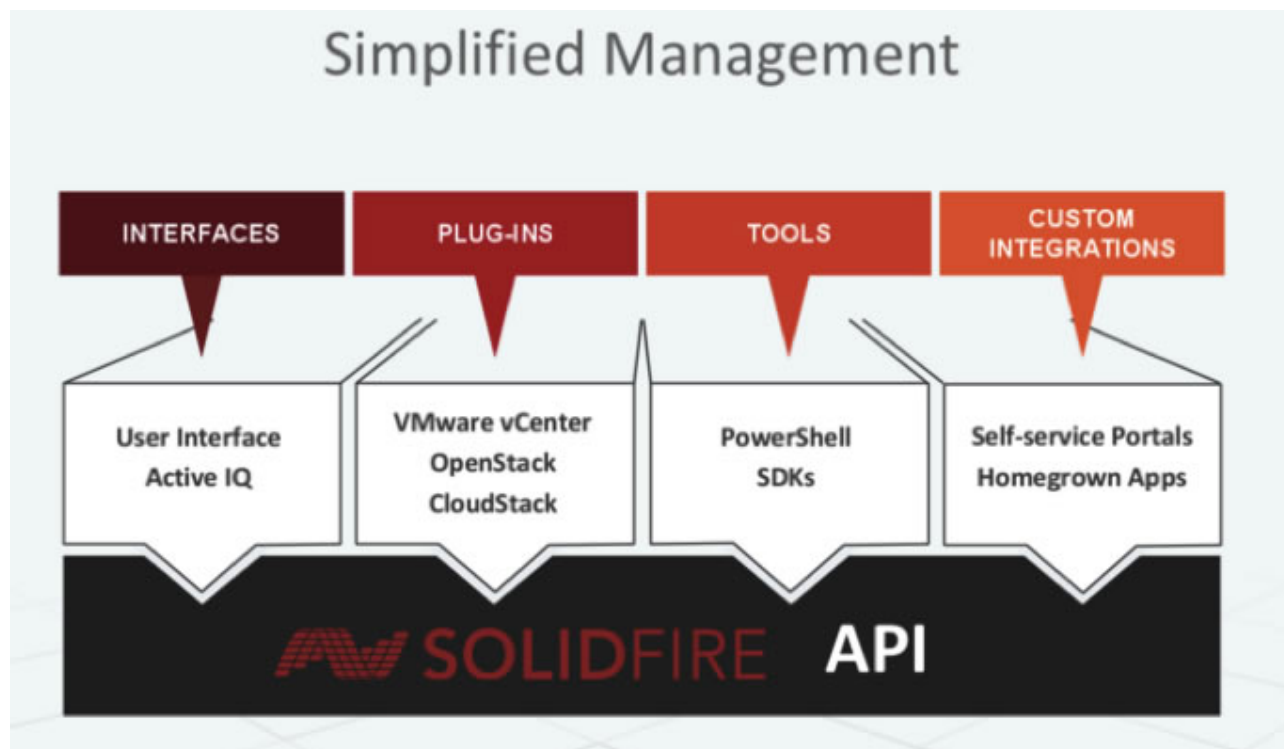
The SolidFire design eliminates much of the complexity that would otherwise complicate automation. Performance and capacity are global pools, and workloads are automatically distributed across the cluster. Provisioning is extremely simple, and many traditional storage tasks are eliminated such as:

- Performance tuning and load balancing
- Managing tiering, prioritization, or caching
- Short stroking or over-provisioning
- RAID group and spare drive management
- Generational upgrades or platform migrations

Everything SolidFire does is exposed through a comprehensive REST-based API. Automation reduces the risk of

human error associated with complex administrative tasks.

**Figure 4) SolidFire REST API.** The SolidFire REST API underpins all SolidFire management interfaces, plug-ins, and tools, and facilitates custom integrations.



Source: SolidFire 2016

The SolidFire API enables deep integration with management and orchestration platforms and supports the development of user-facing storage controls. It also enables rapid deployment of applications and services. All of SolidFire's tools and all third-party integrations including those for VMware, OpenStack, and others are built using the API.

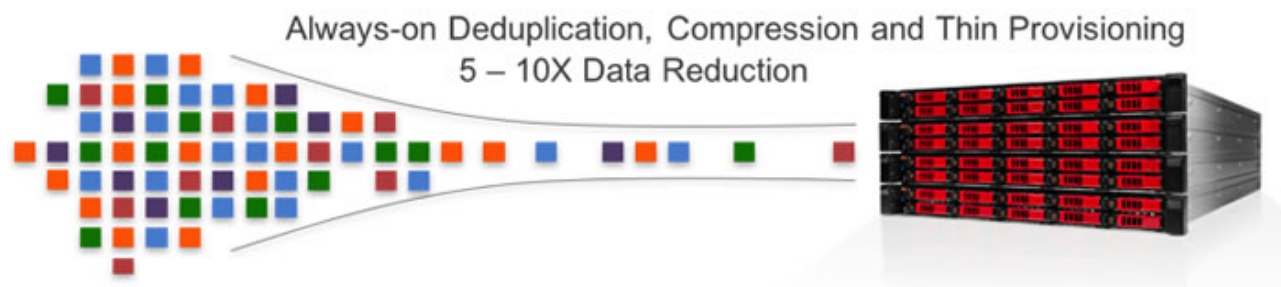
ESG makes the benefit of SolidFire automation clear, concluding that [SolidFire helps administrators spin up virtual machines up to 81% faster and lowers operating expenses up to 67% versus traditional storage.](#)

### Inline Data Efficiency

SolidFire offers a variety of storage efficiency technologies, including global thin provisioning and space-efficient snapshots and clones. These are similar in principle to NetApp technologies you are probably already familiar with.



Combined with **multilayer compression** and **global inline deduplication**, these technologies increase the effective storage capacity of a SolidFire cluster.



Each SolidFire node includes a PCIe NVRAM card that serves as a write cache. When a host writes data, the write is divided into 4KB blocks that are immediately compressed and stored in NVRAM. Each compressed block is synchronously replicated to an additional storage node. An acknowledgement is returned after data has been stored in NVRAM on both nodes, so writes are extremely fast and performance is predictable.

Each compressed block is hashed using a secure crypto hash algorithm. The resulting value serves as a BlockID. The BlockID determines block placement, resulting in a content-addressed storage system similar to those used in leading object stores. The hash algorithm distributes blocks across all nodes in a random fashion that ensures an even distribution of load.

Based on the BlockID, the SolidFire Deduplication Block Service identifies blocks that have previously been written. If a block already exists, metadata is updated accordingly and the duplicate is discarded.

The deduplication process is in-line and global—deduplication happens across the entire cluster, not per volume or per node.

The combination of inline compression and global deduplication has substantial advantages:

- **Reduces drive wear:** Repetitive writes are eliminated, increasing SSD life.

- **Increases system performance:** System resource consumption is minimized.
- **Eliminates hot spots:** Workloads are evenly distributed across entire clusters.

The inline compression algorithm was chosen based on speed. SolidFire also uses post-process compression in the background. A more computationally intense compression algorithm further optimizes storage capacity without impacting performance.

## Complete Storage Capabilities

This article provides a foundation to help you begin to understand the unique aspects of SolidFire's all-flash scale-out storage platform. While many of the key points have been touched on, you should know that SolidFire offers a comprehensive set of storage services:

- Replication (synchronous / asynchronous)
- Integrated cloud backup
- Snapshots and clones
- 256-bit encryption-at-rest
- Comprehensive logging
- Cloud-based monitoring
- Secure multi-tenancy
- Simultaneous multiprotocol support (FC / iSCSI)
- Deep integrations: VMware, OpenStack, CloudStack

To find out more about SolidFire, check out the resource list in the sidebar included with this article or visit [solidfire.com](http://solidfire.com).

*Dave Wright left Stanford in 1998 to help start GameSpy Industries, where he led a team that created a backend infrastructure powering thousands of games and millions of gamers. He later served as Chief Architect for IGN after it acquired GameSpy.*

*In 2007 Dave founded Jungle Disk, a pioneer in cloud-based storage and backup. Rackspace acquired jungle Disk in 2008, and Dave worked closely with the Rackspace Cloud division to build a cloud platform that supported tens of thousands of customers. In December 2009, Dave left Rackspace to start SolidFire.*