Solution Brief

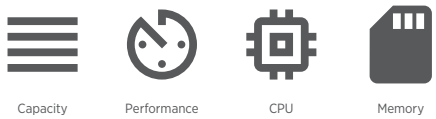
Scale-Out Storage for Cloud Infrastructure

Balance performance and capacity independently with SolidFire scale-out systems

Key Features

In the modern data center, a constant struggle exists between applications and storage. Applications are very sensitive and require adequate storage performance and capacity to operate. Any imbalance can negatively impact an application's ability to perform, causing a ripple effect to the business.

Flash solves many performance problems, but due to the significant performance available in an all-flash array, most systems run out of capacity before performance. With a scale-up storage system, the design of the storage controller dictates the system's performance. Regardless of how much capacity you have left, if you need to add performance, you are stuck buying another controller pair. Achieving an ideal balance of capacity and performance is nearly impossible.



The Scale-Out Process

To provide an all-flash storage system capable of scaling seamlessly and nondisruptively, SolidFire uses solid-state storage and provides hosts with separate pools of capacity and performance. Traditional storage architecture marries the two and cannot separate performance from capacity, resulting in stranded capacity and inconsistent performance. The solid-state pool of capacity is separated into metadata and block data capacity.

Metadata capacity consumes about a tenth of the node's storage capacity; it is simply the list of block identifiers used by each volume on the system. SolidFire block data capacity, which is the remaining nine-tenths of storage capacity on each node, stores two copies of each unique 4k block. This architecture was designed to ensure effortless in-line data reduction and guaranteed consistent performance.

Mix and Match

It's critical for a storage system to help avoid the lengthy downtime and business impact of a "forklift" upgrade. Many storage systems have specific shelves of disks that are compatible only with certain controllers, limiting expansion or requiring downtime for upgrades. The SolidFire scale-out architecture makes it possible to add the latest node model to any previously released model and operate them as a single cluster. When a node of the same or different configuration is added to an active cluster, the block and metadata services are responsible for rebalancing data in the cluster across all nodes.

All active volumes running on the cluster remain the same, and no remaps or rescans are necessary. Regardless of the SSD size, NAND type, or vendor/model, SolidFire systems are built to simply consume it as additional capacity and performance. This software-based approach to storage is what allows the SolidFire infrastructure to be so flexible, not forcing customers into a specific configuration path.

Capacity Management

As new nodes are introduced to a cluster, their performance and capacity (block and metadata) are added to the collective pool of the SolidFire cluster. When block drives are added to the system, SolidFire evenly distributes the existing block data onto the newly added block capacity. This results in the newly added node having an equal share of block data compared to every other node in the system.

NetApp®

Similar to the block service, the metadata service evenly distributes volumes to the newly added metadata service, ensuring constant redundancy across all nodes in the cluster. During the scale-out process, the cluster temporarily contains three copies of each block of data instead of its standard two copies. This extra copy of data ensures that the newly created copy is stored on the newly added capacity before the metadata service removes the copy that it was migrated from. This process ensures that while scaling is in progress, the maximum level of data assurance and protection is used. In addition, SolidFire calculates which volumes need to move to achieve a perfectly balanced I/O and capacity combination for the cluster

Planned (Sunny Day) Capacity Removal (Scale-In)

The SolidFire architecture allows storage to be shaped and shifted based on individual needs. For a planned node removal ("sunny day" — manual removal of nodes), an administrator marks a node and its drives for removal, which initiates the system to balance capacity off of the node being removed. For the block service, the data that exists on each drive is rebalanced to a node that is remaining in the cluster. Metadata follows the same process, but instead of blocks of data being moved, primary and secondary copies of metadata are gracefully migrated off the node to be removed. Once metadata and block data have both been successfully reprotected, the system marks the node as safe for removal.

Unplanned (Rainy Day) Capacity Removal

If data needs to be rebalanced due to an unplanned failure, the SF system self-heals and automatically rebalances any data that was on the component that failed. When a component has failed, and the system needs to rebalance and reprotect the data, there is a 5-minute waiting period. This wait time is important to prevent triggering an unnecessary cluster rebalance for a simple reboot or temporary power loss. Each block of data present on the lost node is automatically reprotected and rebalanced on one of the nodes remaining in the cluster.

Metadata and block data rebalancing and reprotection occur independently, and are triggered separately or together if needed. For example, in the case of complete power loss for a node, both block and metadata sync. If only the metadata drive fails, then only the metadata is rebalanced and reprotected.

Preventing Business Disruption

In every cluster, a cluster master service is responsible for maintaining a full copy of all cluster operations to ensure that operations such as adding, removing, or the failure of a node does not disrupt active operations on the cluster. This is important because the master service keeps track of the service profiles, capacity, authentication, snapshot schedule, and data protection strategy for every volume in the system.

Network Connectivity

All hosts connect to a SolidFire cluster through the storage virtual cluster IP address. Once the connection is established, a redirector service passes the connection to the node that contains the primary metadata copy for the volume for which the host is requesting access. During the node removal process, great care is taken to prevent a host from losing connection to its volume. If there is an active connection from the host to a node that is being removed, the secondary copy is promoted to primary and a new secondary is created. A seamless iSCSI redirect then occurs, and the application remains connected to its volume. A similar process occurs when adding a node to a cluster. The newly added node takes on its portion of volumes to help evenly distribute load and capacity on the system.

As volumes are rebalanced to provide optimal placement, a seamless iSCSI redirect occurs if the active connection is moved to one of the new nodes added to the system.

Benefits of SolidFire Scale-Out Storage

In a cloud, you are never going to need to add an exactly equal amount of performance or capacity, and it's never guaranteed that you need to add the same level of capacity and performance that you have been purchasing. That's why it's essential to choose an architecture that allows you to pick a ratio of performance and capacity that best fits your business needs.

- **Nondisruptive scale-out/scale-in.** Add nodes to or remove them from a SolidFire cluster without disrupting service or compromising volume-level quality of service (QoS) settings. Data is automatically redistributed in the background across all nodes in the cluster, maintaining perfect balance as the system grows.
- **Instant resource availability.** Newly added storage capacity and performance resources are instantly available to every volume in the system, eliminating the need to reallocate volumes over new drives.
- **Simplify capacity planning.** Initial implementations begin with a simple 4-node/4U cluster configuration and scale out easily in 1U node increments, allowing performance and capacity resources to be added as needed. Eliminate multiyear capacity and performance projections and scale on demand.
- **Seamless generational upgrades.** New nodes with more capacity and performance are simply added to the established cluster, while old nodes are removed, retired, or repurposed. No rebalancing, restriping, or volume reallocation required. And all QoS settings remain enforced.

About NetApp

NetApp is the data authority for hybrid cloud. We provide a full range of hybrid cloud data services that simplify management of applications and data across cloud and on-premises environments to accelerate digital transformation. Together with our partners, we empower global organizations to unleash the full potential of their data to expand customer touchpoints, foster greater innovation and optimize their operations. For more information, visit **www.netapp.com**. #DataDriven