How SolidFire Data Efficiencies Work

SolidFire's always-on, global data efficiencies maximize system capacity without performance vs. capacity trade-offs.

SolidFire's storage system combines all-SSD performance with highly efficient data distribution and management. Embedded granular thin provisioning, multi-layer data compression and global data deduplication techniques that solve traditional storage deficiencies making flash at scale an economic reality while delivering superior performance.

Inline Data Reduction

SolidFire's scale-out architecture ensures network, cache, and compute resources grow in tandem with capacity as the cluster is expanded. Compared with a more conventional scale-up model, the SolidFire scale-out architecture delivers linear performance gains as customers increase capacity by adding nodes to the cluster.

Those efficiency gains, combined with the SolidFire's scale-out design, ensure there is plenty of CPU and RAM horsepower to reduce and rehydrate data, regardless of the performance and capacity demands being placed on the system.

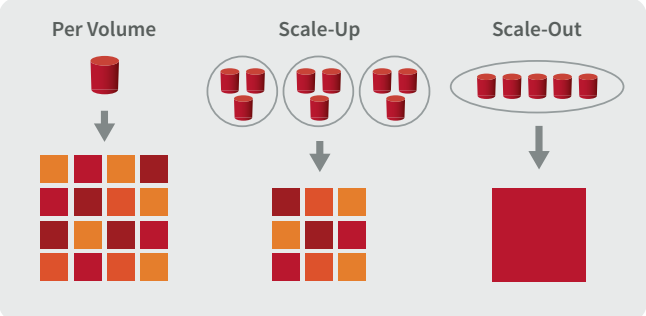
When coupled with with a shared-nothing architecture and a distributed replication-based configuration, the SolidFire system enables maximum resource utilization, reducing cost and accelerating innovation.

Global Deduplication:

SolidFire's Deduplication Block Service receives a WriteBlock message, which contains the unique BlockID and data. The BlockID is then computed based on a hash from the 4K data. As a result, the Block Service will automatically identify whether the BlockID has already been written.

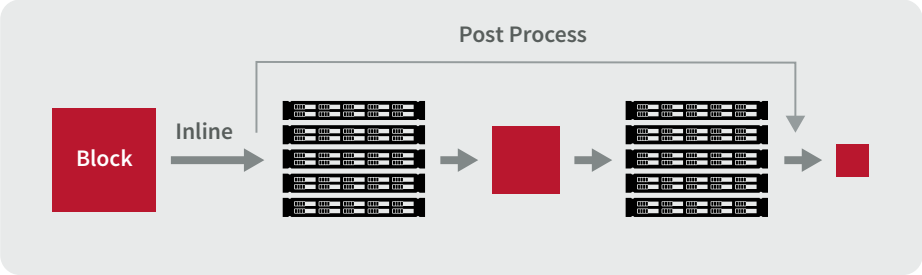
If the Block Service recognizes the BlockID, duplicate objects will never translate into physical data writes onto the block drives. This entire process is performed in-line with no performance impact to the system, and enables customers to:

- Reduce repetitive writes to media, increasing the life of the drives
- Increase system performance by minimizing system resources
- Evenly distribute capacity and performance loads across the system, eliminating hot spots



Multi-layer Compression:

One key to SolidFire's technology is its recognition that flash needs to be handled far differently from traditional spinning disk storage. SolidFire has architected its entire storage system accordingly, minimizing writes by compressing and deduplicating data before writing to its flash SSDs.



When a host writes data to a SolidFire storage node, that write is divided into 4KB data blocks. These blocks are immediately compressed and stored in the node's NVRAM write cache. Each compressed block is synchronously replicated to one or more additional storage nodes for data protection. An acknowledgement is returned to the host when — and only when — the data has been safely stored in the NVRAM of multiple storage nodes.

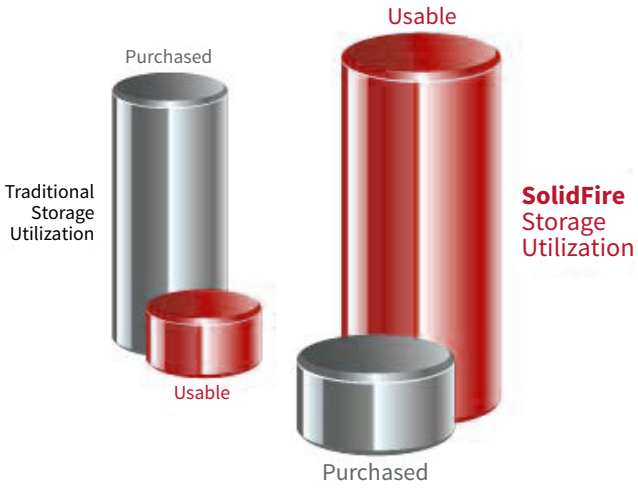
Data contained in the compressed data chunk is then hashed. The system looks for that hash value in its index of stored data, which is distributed across the entire cluster. If the data is already present, the SolidFire operating system updates its metadata to indicate that the previously stored block should be delivered when the host reads the data being written, and the newly written block is discarded without ever being written to flash.

If the compressed data block has unique data, it is stored in the system's block pool. The block pool is organized only by the content hash value, rather than by when data was written or from where it originated. To maximize the efficiency of storing compressed data blocks, the block storage pool doesn't allocate space in fixed-size chunks; instead, the system tightly packs blocks that vary in size, as determined by the compressibility of the data.

4K Granular Thin Provisioning:

Systems that use more traditional RAID sets of dedicated drives have typically had thin provisioning implemented after the fact, frequently by automating the LUN expansion process. This often entails adding another slice of a RAID set to a thinly provisioned volume when it reaches some preset threshold. Such automated extension of volumes can, on some systems, cause a performance stutter for a few seconds as space is allocated. In addition, using large allocation extents can reduce the efficiency gains of thin provisioning.

To minimize the effect of granularity on provisioning, SolidFire allocates data 4KB at a time. This increases efficiency and reduces overhead by utilizing the smallest allocation possible while maintaining alignment with the native 4KB allocation format used by many operating systems, applications, and modern disk drives.



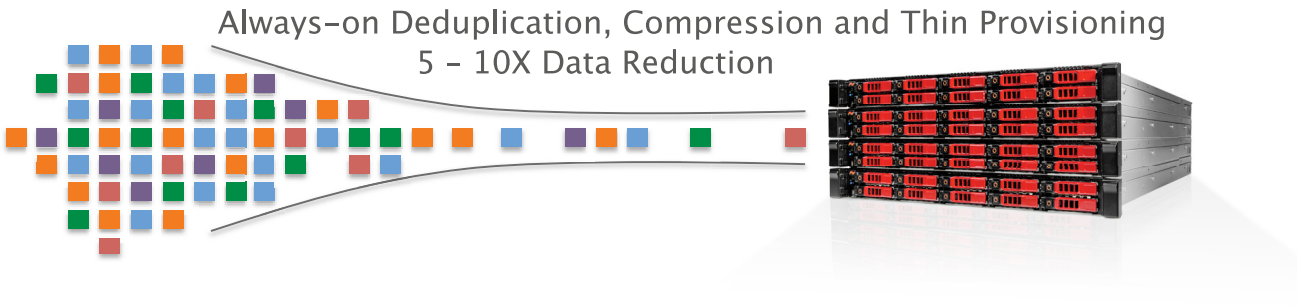
Benefits of Inline Data Reduction Techniques

Performance enhancement - Improved performance is a key benefit of deduplication and compression techniques being performed inline as there is no performance tax that results from their usage within a primary storage infrastructure.

Guaranteed systemwide efficiency - the SolidFire system tags and stores data, ensuring the data stored always resides in its most optimal form, compressed and deduplicated across both the NVRAM and SSD tiers.

Increased effective capacity - SolidFire increases effective capacity of the system by reducing the data footprint by five to 10 times inline, before it's written to flash. This reduces the cost per gigabyte, forging the path for flash in the data center.

Extended media endurance - SolidFire extends the life of SSDs by wear-leveling write data across all the flash capacity in all the SSDs in the system.



SolidFire customers see the following combined compression and deduplication gains, on average

Databases	VDI	Server Virtualization
3 to 4x	*4 to 10x*	*3 to 5x*