

Stretch Your Storage Dollar: Optimizing HPC at the Value Tier

SUSE® Enterprise Storage, a Cost-Efficient Alternative for Value-Tier Storage in HPC

Storage systems come in a variety of sizes, shapes and prices, and experts understand that basic parameters such as speed and capacity are only part of the story. The HPC environment places extraordinary demands on the storage system, but economy still matters—any money you save on storage preserves budget that you can then use to improve and extend the rest of your system. SUSE® Enterprise Storage is a self-healing, fault-tolerant storage solution that will help you stretch your storage dollar.

The Problem

Storage serves several roles in the HPC environment. Along with low-latency data accessed directly from a running application, most HPC systems also include archives and other background files that do not have a strong effect on system performance.

Should all data be treated equally?

Although it is possible to outfit a high-performance cluster with only high-performance storage media, most HPC experts would consider this approach limiting. The biggest problem with high-performance storage is that it is quite expensive. Flash and NVMe storage systems are significantly more expensive than disk storage of equal capacity. In addition to the high acquisition cost, many high-end storage appliances are proprietary systems,

with significant maintenance and upgrade costs. Unnecessary spending on non-critical storage can quickly drive up the total cost of the project. For projects that are designed to a maximum overall budget, this overspending reduces the available budget for processors and other critical components, thus leading to a slower and lower-performing system.

In the Tiers

Experts agree that the key to HPC storage design is to provide fast storage for low-latency data that could potentially cause a bottleneck and to optimize other storage components for minimum expense and maximum utility. In other words, put speed where you need it and optimize other storage components for a well-balanced system that achieves performance targets with minimal cost.

A popular approach is to organize the HPC storage environment into two tiers:

- **Fast tier**—*ultra-fast, high-performing, low-latency storage with data accessed directly from within the application.*
- **Value tier**—*archives, long-term storage, infrequently accessed files, and other data that can reside on high-latency storage.*

This classic storage configuration is depicted in Figure 1.

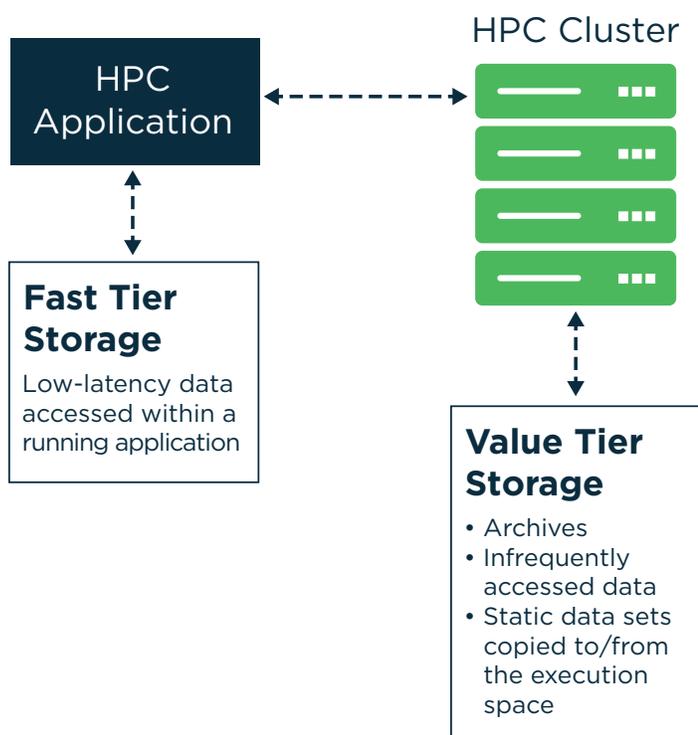


Figure 1. Optimizing HPC storage: collect the data that doesn't specifically require low latency into a separate "value tier."

Understanding the Value Tier

The fast storage tier is more expensive in cost-per-GB, but it is far simpler conceptually. Shopping for fast storage is a relatively simple exercise in studying specs and comparing prices. Value storage, on the other hand, requires careful analysis and consideration of several counter-balancing factors to achieve objectives and still minimize cost.

For instance, an optimum solution for the value tier should come with automation and self-management features to minimize the need for administrative intervention. It should be easy to deploy and, more importantly, easy and inexpensive to upgrade if project requirements change. Although performance demands at the value tier are not as strict as they are for fast storage, most HPC designers also keep an eye on performance.

The secret to optimizing at the value tier is to keep all these factors in focus while still seeking a solution that minimizes expense. SUSE Enterprise Storage is a Ceph-based storage alternative that is gaining popularity as a universal solution for the HPC value tier.

Ceph is:

- **Automated**—*the software seamlessly organizes and optimizes the storage environment.*
- **Self-healing**—*the cluster detects problems and rebalances workloads as needed, continuing operation without downtime or data loss.*
- **Fault-tolerant**—*all data saved to the cluster is protected using mirroring or erasure coding techniques.*
- **Flexible**—*the system is easy to adapt and interacts with the HPC environment in a variety of ways.*
- **Efficient**—*an object storage architecture reduces waste and maximizes storage efficiency.*
- **Inexpensive**—*Ceph is open source and is thus inexpensive to acquire and free of vendor lock-in.*

The Ceph project is under active development and is supported by a creative and vibrant community. The power and economy of the Ceph storage solution have made it a leading alternative for large and small HPC environments. The massive HPC system attached to the Large Hadron particle accelerator at CERN in Geneva, Switzerland, is an example of a critical HPC system built around Ceph.

SUSE Enterprise Storage: Frugal Solution for HPC

SUSE Enterprise Storage is an enterprise-grade storage solution built around Ceph. The SUSE developers have tuned and tailored SUSE Enterprise Storage to serve as a versatile storage solution for HPC environments. Along with the basic features found in all versions of Ceph, SUSE Enterprise Storage adds technical support, hardware certification and a powerful collection of management tools for efficient installation and easy administration.

SUSE Enterprise Storage delivers the benefits of Ceph with lowest-in-class subscription pricing. Most enterprise Ceph providers charge per-GB storage costs—the more data you save, the more you pay. In the data-intensive realm of HPC, per-GB charges drive up overall costs and reduce efficiency.

SUSE Enterprise Storage does not charge a per-GB storage premium. Subscriptions are assessed only on a per-server basis, so you won't pay more just because you save more data.

The automation features and management tools built into Ceph mean that a single storage admin can manage up to six times more data than an admin in an equivalent block storage environment (see Figure 2). The high-end feature set and low-end pricing of SUSE Enterprise Storage leads to a total cost of ownership that is often less than half the cost of other storage solutions.

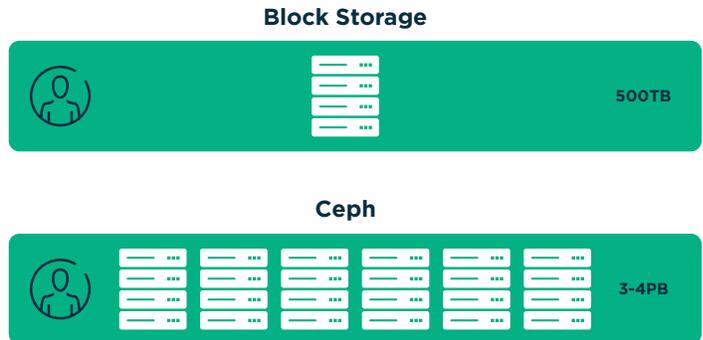


Figure 2. A storage admin in the SUSE Enterprise Storage environment can manage up to six times more data than an admin on an equivalent block storage system.

Versatility First

For many HPC customers, the flexibility of SUSE Enterprise Storage is the most compelling benefit. SUSE Enterprise Storage supports a number of powerful gateway and interface options (see Figure 3). Your storage cluster can appear to the network as a block device, object store, or network filesystem. Other interfaces support iSCSI network block storage and SMB.

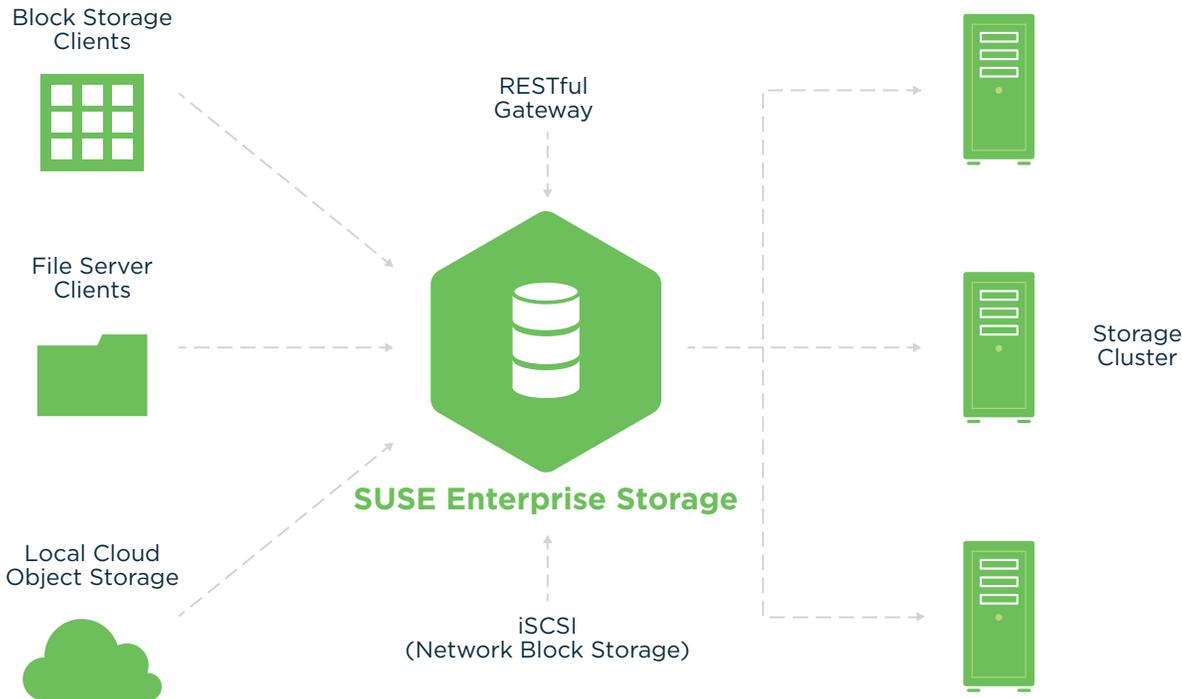


Figure 3. A collection of gateway and interface options provides a versatile system for the storage cluster to interact with the HPC environment.

The gateway and interface options available with SUSE Enterprise Storage enable you to deploy your cluster in a variety of configurations, tailoring the solution to the needs of your environment. Many HPC customers prefer to access Ceph using the CephFS filesystem. CephFS is a POSIX-compliant filesystem that integrates easily with existing applications and components that require a conventional filesystem.

Another important feature of SUSE Enterprise Storage is the ability to layer and organize storage components within the cluster (see Figure 4). It is possible, for instance, to configure your cluster so that data requiring low-latency storage is saved to a high-performing storage device such as an SSD or NVMe drive, and data that doesn't specifically require high performance is routed to inexpensive disk resources.

Single-Source Solution

For some HPC workloads, the versatility of SUSE Enterprise Storage leads to the option of encompassing the two-tiered storage structure depicted in Figure 1 entirely within the Ceph cluster. In other words, Ceph can act as a single-source storage solution, offering both fast tier storage and value tier storage using the organizational capabilities described in the preceding section. A single-source storage solution maximizes storage economy by extending Ceph's minimal TCO to all data within the cluster.

The single-source storage option is not optimized for all workloads. SUSE estimates that approximately 20% of HPC installations could benefit from a solution that manages all data from within SUSE Enterprise Storage. This configuration works best for HPC clusters with fewer than 250 nodes and for scenarios that do not require extremely low latency. Examples of scenarios that could benefit from this solution include scientific studies and some AI workloads, as well as test and development environments.

SUSE Enterprise Storage

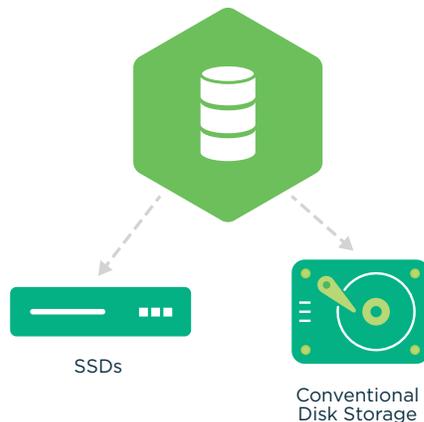


Figure 4. SUSE Enterprise Storage lets you organize storage resources within the cluster so that low-latency data is saved to a low-latency device, such as an SSD or NVMe drive.

Stretching Your HPC Dollar

Smart HPC storage design optimizes performance but still keeps an eye on the budget to preserve investment for other parts of the system. HPC experts understand that the best solution is often two storage tiers: a fast tier for low-latency storage accessed within the application, and a value tier for data with greater potential for optimization, such as archives and less frequently accessed files.

SUSE Enterprise Storage is a self-managing, self-healing, fault-tolerant, clustered storage solution that serves as a cost-efficient alternative for value-tier storage in HPC. For some workloads, SUSE Enterprise Storage also works well as a single tier, with fast storage integrated directly into the cluster through SSD drives or other low-latency hardware.

The experts at SUSE can help you build an HPC system that stretches your storage dollar by optimizing at the value tier.

Additional contact information and office locations:
www.suse.com

www.suse.com

