Western Digital.

SOLUTION BRIEF

Gain Control of Data Overload in High Performance Computing

IBM Spectrum Scale

Highlights

- Increases accessibility by pooling disparate storage into a single pool under a global namespace.
- Improves storage efficiency by automatically moving infrequently accessed data to efficient cloud storage
- Automated management that puts right data in the right place throughout its life cycle
- Protects valuable data with extreme data durability and integrity even at petabyte scale
- Ensures all data is accessible and can be recalled rapidly when needed for long periods of time cost effectively
- Increases ability to react quickly to provision storage resources in response to new business requirements.

Challenge

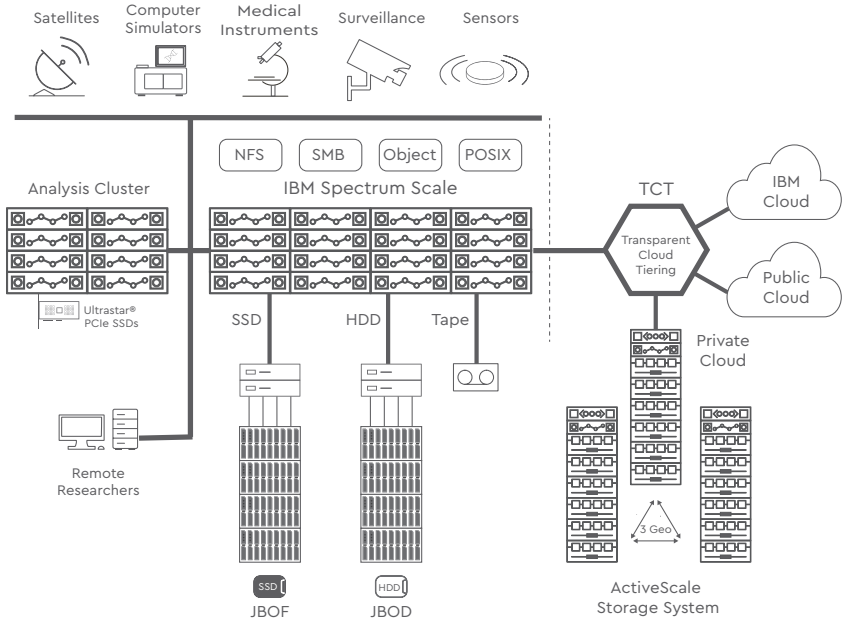
- Scaling storage capacity to meet increasing data growth related to capturing more IOT data
- Protecting large amounts of critical data throughout the information life cycle to meet regulatory requirements and to support ongoing research and development
- Supporting a complex mix of file sizes and access patterns, that require both high throughput and low latency
- Matching data to the right storage technology as it moves through its life cycle.
- Providing access to large amounts of data globally, enabling researchers to turn information into knowledge.

Rapid advances in the ability to collect a greater quantity of information from data generated by sensors, satellites, medical instruments, high definition video devices and computer simulators is increasing the size of data repositories to capacities never seen before. The ability to move seamlessly from raw data to analysis to actionable insights is making it possible for researchers and engineers to transform these vast amounts of data into insights and to deliver better, faster and more cost-effective results with their high performance computing (HPC) environments.

However, for many HPC environments, data intensive workloads can overwhelm data storage systems, creating severe computing bottlenecks and major performance issues. In addition, managing and monitoring these complex storage systems significantly adds to the burden on storage administrators and researchers. Organizations must have the ability to store, process and analyze large volumes of data without impacting the efficiency of their high performance computing environments.

As the requirements around data storage performance and capacity keep increasing, researchers, scientists, engineers and storage administrators are increasingly empowered to consider new approaches for deploying high-performance computing (HPC) and storage infrastructure.

IBM Spectrum Scale with TCT + ActiveScale for HPC Architecture



Life Sciences architecture example using an ActiveScale object storage system

Solution

IBM Spectrum Scale™ with Transparent Cloud Tiering and Western Digital ActiveScale™ object storage system combine to provide automated policy-based storage tiering from flash through disk to tape. This extremely scalability solution allows clients to seamlessly access all data from a shared global namespace and enables users to easily create and manage a private or hybrid cloud using the ActiveScale system to help reduce storage costs. Data that does not require high performance can be transparently moved to efficient ActiveScale object storage and data that is frequently accessed can be promoted to high-speed flash, ensuring the right data is stored in the right place at the right time.

Legacy storage architectures lack the flexibility, scalability and accessibility needed for modern workloads at petabyte scale. Data is often silo'd by architectural limitations even though total system capacity is sufficient. This often leads to data redundancy, driving up costs and increasing management complexity.

Next generation storage architectures based on modern technologies and storage constructs can reduce the cost and complexity of storing large amounts of data.

Breakthrough Storage Efficiency

IBM Spectrum Scale with Transparent Cloud Tiering is software-defined storage for applications that demand high performance and shared access to a common set of local or remote data. It provides extreme scalability with support for flash accelerated performance and automatic policy-based tiering across flash, spinning disk and tape media. Native protocol access allows clients to seamlessly access all data via a shared global namespace. A highly available cluster can range from two nodes to thousands in order to meet the most demanding HPC workload requirements.

Private and Hybrid Cloud Storage Made Easy

Transparent Cloud Tiering enables users to integrate Western Digital's ActiveScale object storage to create a private or hybrid cloud to help reduce storage costs. Data that does not require high performance storage can be transparently moved to ActiveScale automatically based on life cycle policies. This is ideal for data that needs to remain on-premises or has a higher recall frequency that would make off-premises cloud storage too expensive.

IBM Spectrum Scale combined with ActiveScale object storage allows High Performance Computing environments to gain control of data growth with:

- **Unified storage** by pooling redundant isolated storage resources under a single global namespace that supports a diverse set of HPC workloads where performance, reliability and availability of data are essential to the business.
- **Storage efficiency** by freeing up expensive storage capacity by transparently moving infrequently accessed data to more cost effective ActiveScale object storage and promoting data that is frequently accessed to high-speed flash.
- **Improved access and collaboration** provided by the global namespace and native protocol support for NFS, SMB, Object and POSIX, making it easier for teams to collaborate boosting overall productivity.
- **Extreme data durability and integrity at petabyte scale** helps ensure valuable data is protected and always available. ActiveScale delivers up to 19 nines durability and site-level fault tolerance in a multi-site configuration.
- **Modern approach to long-term retention** ActiveScale offers fast access to cold data, enabling organizations to eliminate antiquated tape technology.
- **Easy to install and manage** – simply add power and network connections and it is ready to go. The system self-protects and heals requiring significantly less IT intervention compared to traditional storage systems.

Conclusion

The enormous size of today's data repositories presents a daunting challenge to High Performance Computing environments. Huge amounts of data needs to be quickly stored, processed, analyzed and kept for long periods of time. The software defined storage solution from IBM and Western Digital delivers the necessary performance, scale, and efficiency needed to address these challenges with fewer resources and lower cost than traditional approaches.

To learn more about ActiveScale visit: www.wdc.com/dc-systems

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
US (Toll-Free): 800.801.4618
International: 408.717.6000

www.westerndigital.com