

Hitachi Hyper Scale-Out Platform (HSP) combines server computing, server virtualization and storage in a hyper-converged appliance. HSP delivers high-performance NFS ingest integrated with Hadoop, to provide faster insight from mixed analytics workloads.



Hitachi Hyper Scale-Out Platform

Enterprise-Class Distributed File System for Hadoop

With Hitachi Hyper Scale-Out Platform, Hitachi Data Systems provides an enterprise-class file system and analytics environment, fully compatible with Hadoop. HSP is able to:

Provide higher availability, performance and enterprise robustness required by large enterprises by removing any single point of failure within the Hadoop standard architecture. The Hitachi solution uses distributed management of metadata on each node of the cluster, increasing the overall solution availability and performance.

Significantly enhance the flexibility of the solution by moving applications towards the data. This approach provides a unique big data environment in which Hadoop

clusters can be dynamically deployed close to the data. By sharing a common data platform and enabling computing nodes to host different VMs for different big data analytics needs, HSP enables increased consolidation. It helps to reduce the network traffic required to move data from different sources and to optimize resource utilization and cost.



A few important differences set the architecture of HSP apart from a typical Hadoop implementation by enabling data in-place analytics and faster time to results.

Hyper-Converged, Scale-Out Platform Is Optimized for Hadoop

HSP uses an enterprise-class scale-out file system that is optimized for Hadoop (supports the Hadoop Distributed File System or HDFS API). It also provides a portable operating system interface (POSIX)-compliant file system that allows other third-party data analytics applications to directly access the data without having to move the data out (export) to a separate system.

The scale-out solution provides compute power, memory and a virtual infrastructure, a combination that allows the deployment of many virtual instances, as needed, directly on HSP. Running additional analytics workloads is as simple as spinning a new VM instance.

The HSP approach brings the applications to the data, both for Hadoop analytics as well as for other applications' analytics, to avoid data copy and movement across the network. It provides a single, highly integrated and scalable platform that manages any type of workload and performs

different analytics, without moving data around between compute and storage resources. The approach allows much faster and more flexible deployments, minimizes data traffic, allows massive scalability and ensures ease of management. Connectors to archive systems allow archiving of data when it is no longer needed on the analytics platform.

HSP represents a scale-out architecture that is designed to provide efficient and reliable shared distributed storage and compute capabilities. The capacity and performance scales as nodes are added, on demand.

While Hitachi uses many of the same principals of Hadoop (64MB data chunk size, three-copy policy, data location API), we are able to avoid the NameNode bottleneck by distributing metadata, which enables every node to serve data without centralized metadata management. HSP accesses the data via standard POSIX semantics through our efficiently distributed file system. HSP is able to work in a Hadoop environment because we use the standard interface that Hadoop analytic jobs reference. Therefore, HSP has the ability to:

- Scale without NameNode bottlenecks.
- Fit into analytics workflow.
- Access the data in place, without loading the data into HDFS format.

The Hyper Scale-Out Platform Value Proposition

Hitachi Hyper Scale-Out Platform provides the following benefits.

- **Converged compute, storage and virtualization platform:** Helps organizations to gain faster business insight and save on capital expenditures.
 - **In-place data access** (data are local to the analytic jobs) both for Hadoop workload and other analytics applications; leverages the Hitachi enterprise scalable file system and the POSIX-compliant file system.
 - **Virtualization management integration** for flexible analytics; allows the deployment of mixed workloads running on the same physical node.
- **Availability:** No single point of failure or contention. Data is automatically distributed across the nodes in the system. The way this is done ensures no single point of failure; because of the distributed lock and metadata management, mechanisms avoid contention.
- **Recoverability:** Data protection and availability. Data is protected while written into the system. Automatic copies are made of the data to allow automatic protection and recovery.
- **Massive scalability (scale-out paradigm):** Designed to start relatively small (a few hundred terabytes) and to scale to multiple petabytes.
- **Performance:** Scale while the solution grows. As each node adds capacity, as well as CPU and memory, the solution scales in each direction.
- **Balancing and rebalancing:** Automatic data distribution. Distributed link and metadata management allow automatic distribution of large data sets across the nodes.

SPECIFICATIONS TABLE

Product Model: Hitachi Hyper Scale-Out Platform 400 (per node)	Product SKU: SSG-SOP400-SN0001A
Form factor (height)	2U (rack units)
Power	Redundant 920W platinum power supplies
Processor	Two Intel Xeon E5-2620 2.00GHz 15M cache processors
Memory	192GB and battery-backed NVRAM
Hard disk drive or HDD (data)	Twelve 3.5" 6TB SAS 7,200rpm HDD and all flash configurations
HDD (application or operating system)	Two 2.5" 300GB SAS 10,000rpm HDD
Network (back-end)	40GbE dual-port QSFP NIC
Network (front-end)	10GbE dual-port SFP + NIC
Width	17.2" (437mm)
Height	3.5" (89mm)
Depth	25.5" (638mm)
Weight	95.75 lbs. (with HDDs installed)
Operating temperature range	10 – 35 degrees C
Humidity range	8 – 90% noncondensing
Brocade ICX 7750-26Q Switch	
Width	21.9" (556.26mm)
Height	6.1" (154.94mm)
Depth	22.1" (561.34mm)
Weight	7.005 lbs.
Minkels 42U Rack Cabinet	
Width	23.62" (600mm)
Height	79.13" (2010mm)
Depth	47.24" (1200mm)
Weight	275 lbs.

- **Self-configuring, self-managing and self-repairing (self-healing):** Automatic configuration is applied throughout the system. There is a RESTful API and the support of Hadoop management tools. Errors are automatically identified and repaired.
- **Cloud ready and OpenStack compatible:** HSP provides APIs for OpenStack Glance (image services), Nova (cloud computing fabric controller) and Swift (object storage) projects to enable users to deploy infrastructure as a service (IaaS) functionality.

Hitachi Data Systems

Corporate Headquarters
 2845 Lafayette Street
 Santa Clara, CA 95050-2639 USA
www.HDS.com community.HDS.com

Regional Contact Information
Americas: +1 866 374 5822 or info@hds.com
Europe, Middle East and Africa: +44 (0) 1753 618000 or info.emea@hds.com
Asia Pacific: +852 3189 7900 or hds.marketing.apac@hds.com

