PERFORMANCE BRIEF
Communications and Storage Infrastructure
Huawei OceanStor* 9000 Big Data Storage System
Intel® Xeon® Processor E5-2400 Product Family

Big Data Storage from Huawei and Intel: Linear Scalability to 288 Nodes and 60 Petabytes

The Huawei OceanStor* 9000 big data storage system, based on the Intel® Xeon® processor E5-2400 product family, scales linearly to 60 petabytes (PB) of data, under a single file system. In testing, the symmetric, distributed architecture completed queries at peak workloads in less than two seconds, in configurations of 72, 144, and 288 nodes.

The growing volume and diversity of today's enterprise data creates challenges for IT organizations as they strive to maintain storage infrastructure that can shift and scale rapidly to meet unpredictable requirements. Payloads may range from conventional relational databases to stores of multimedia files and unstructured social media, while workloads range from queries and batch operations to real-time streaming analytics. And it all happens at unprecedented scale, with tremendous ongoing growth, and against a background of flat (or even shrinking) IT budgets.

The industry is responding to this set of needs with standards-based solutions that are more powerful, flexible, interoperable, and cost-effective than the complex and costly proprietary solution architectures they replace. Powered by the latest data-handling innovations in Intel® architecture, accelerated solutions enable sophisticated and dynamic storage capabilities. The high performance and scalability of these systems supports the full range of emerging usage models for enterprise storage, with new efficiencies through server-storage convergence in areas such as data management, analytics, reporting, and archiving.

Innovation and Leadership from Huawei OceanStor 9000

To help meet the business and IT challenges of big data storage and analytics, the Huawei OceanStor 9000 storage system provides leadership in performance, scalability, and flexibility for structured or unstructured data. Based on the Intel® Xeon® processor E5-2400 product family, the platform is capable of more than five million network file system operations per second, a record-setting result as measured by the SPECsfs2008*_nfs.v3 benchmark.[1] This extraordinary system performance is enabled by a variety of factors, including high-speed, intelligent networking, as well as a global cache of up to 55 terabytes, pooled from non-volatile DIMMs (NV-DIMMs) that reside on nodes throughout the OceanStor 9000 storage system.

The platform achieves outstanding scalability by means of a shared-nothing symmetric distributed architecture that allocates data and metadata evenly to all nodes, as well as a global namespace supported by the Wushan distributed file system, which can be dynamically expanded to up to 60 PB of data. The OceanStor 9000's innovative "all-in-one" design enables data storage, archiving, and analysis within one system, across the entire data lifecycle. Unique zero-loading functionality enables data to flow among multiple servers, avoiding resource-consuming data loading and migration between devices and optimizing data-utilization efficiency.

The technical architecture allows for linear growth in performance as nodes are added for capacity, up to the maximum of 288 nodes.[2] Automatic deployment and configuration services allow one-click capacity expansion, with the ability to add a new node in under 60 seconds, without interruption of service.[2] The storage system provides more than 520 MB/s read-write bandwidth for each node, and more than 150 GB/s read-write bandwidth for the entire system,[2] providing sufficient I/O for highly demanding storage workloads. These factors position the Huawei OceanStor 9000 as the storage solution for a growing number of big data implementations in diverse fields such as broadcasting, satellite mapping, energy exploration, and scientific research.

Joint Verification of Capabilities by Huawei and Intel

Huawei and Intel performed comprehensive lab testing of the OceanStor 9000 storage system, using a rigorous set of requirements that were designed to simulate real-world scenarios. This work serves to both validate the system's capabilities and act as a reference solution for commercial implementations.

Testing Setup and Objectives

A collaborative team of engineers from Intel and Huawei built the test environment in a lab operated by Huawei. The OceanStor 9000 system under test consisted of nodes based on the Intel Xeon processor E5-2400 product family. Each node was enabled for high-speed, advanced networking, powered by the Intel® 82599 10 Gigabit Ethernet Controller, and the nodes were interconnected through a 10 Gigabit Ethernet (10GbE) switch. A 60 PB test data set was provided as the basis for read/write workloads to enable robust stress testing of the storage system.

Testing was conducted on clusters of 72 nodes, 144 nodes, and 288 nodes. As the scale of the environment increased, performance was measured in terms of throughput based on the speed of node-to-node data writes to disk (in MB/s). Other metrics of interest included query response time and the time required to deploy additional nodes. Test objectives also included verification that additional nodes could be deployed while the system was online, without service interruptions or downtime.

Test Results and Analysis[2]

Testing outcomes were analyzed with the explicit goal of verifying the OceanStor 9000's suitability for real-world big data implementations, within the following domains:

- **Performance**. Responses to user queries were obtained within two seconds, regardless of the number of nodes deployed, even with the large data sizes used during testing.
- **Scalability**. Linear increases in performance were observed as the environment scaled out from 72, to 144, to 288 nodes.
- **Availability**. Testing showed the ability to add nodes to expand the configuration while the storage system was online, without service interruptions or downtime.
- **Configurability**. A management GUI enabled rapid deployment of new nodes, supporting single-node expansion within 60 seconds and deployment of the entire 288-node system within 45 minutes.

In summary, testing demonstrated that the Huawei OceanStor 9000 big data storage system provides consistently high levels of performance, scalability, availability, and configurability under demanding conditions. The solution also proved to be stable, without any discernable hardware or software errors throughout the testing.

Intel® Architecture Big Data Building Blocks

Storage innovation in Huawei solutions benefits from a co-engineering relationship with Intel that draws on the expertise of both companies, creating technology synergies that drive robust results and business value. As part of that effort, the OceanStor 9000 storage system is optimized to take advantage of a broad range of hardware and software technologies from Intel, including the key elements that are described below.

Intel® Xeon® Processor E5-2400 Product Family

At the core of each of the server nodes that comprise the OceanStor 9000 storage system is the Intel Xeon processor E5-2400 product family. This processing engine incorporates a range of I/O and memory features and capabilities that contribute to robust storage infrastructure.

- **Accelerated I/O moves data quickly throughout the storage system**. 24 lanes of PCI Express* 3.0 interconnect support up to 8.0 gigatransfers per second, with the I/O controller integrated directly onto the processor die to reduce latency.
- **Enhanced memory subsystem supplies storage data efficiently to processors within nodes**. Three independent 64-bit DDR3 memory channels provide high-speed memory access, and integration of the memory controller onto the processor die further reduces latency.

These features of the Intel Xeon processor E5-2400 product family help turn the dramatic expansion of enterprise data from a challenge into an opportunity, improving data center efficiency and preparing for emerging usage models.

Intel® Data Direct I/O Technology

To increase the efficiency of system-level I/O data flows, Intel® Data Direct I/O Technology (Intel® DDIO) allows Intel® Ethernet Controllers and adapters to communicate directly with the processor cache, which acts as the primary destination and source of I/O instead of main memory. This innovation re-architects the flow of I/O data in and out of the processor, delivering increased bandwidth, lower latency, and reduced power consumption within the nodes of the Huawei OceanStor 9000 storage system.

Taking advantage of the large last-level cache provided by the Intel Xeon processor E5-2400 product family, Intel DDIO reduces the number of memory accesses required within nodes. By avoiding multiple reads from and writes to memory, Intel DDIO enables more efficient movement of storage data at the node level and reduces processing overhead, helping drive high overall performance of the OceanStor 9000 storage system.

Intel Ethernet Optimizations for Multi-Core Processors

The large data sets employed by the Huawei storage solution create the need for efficient processing of network data, and Intel Ethernet provides a number of features designed to take advantage of multi-core processors. Rather than handling data requests serially as in earlier systems, the Intel 82599 10 Gigabit Ethernet Controller handles data flows in parallel, taking better advantage of processing resources.

- **Multiple descriptor queues** allow network traffic within the OceanStor 9000 storage system to be distributed into queues that can each be associated with a specific processor core.
- **Receive-side scaling (RSS)** directs incoming traffic flows into specific queues, without needing to reorder packets, which helps to ensure the efficient handling of network processing among processor cores.
- **Virtual Machine Device Queues (VMDq)** offloads network-traffic management to the network silicon, avoiding software-based processes to reduce packet-processing delays and free up processor resources for other work.
- **Extended message-signaled interrupts (MSI-X)** improves the efficiency of communication between queues and processor cores, for efficient packet management and fine-tuned load balancing among cores.

These features work together to enhance network performance of the Huawei OceanStor 9000 storage system under heavy loads, helping ensure high application responsiveness for end users.

Intel® Intelligent Storage Acceleration Library (Intel® ISA-L)

The OceanStor 9000 takes advantage of Intel ISA-L, a library of algorithms that are optimized for Intel architecture to provide enhancements in storage efficiency, data integrity, data availability, and security/encryption. By using Intel ISA-L functions, the Huawei OceanStor 9000 enables efficient utilization of the multiple execution cores in Intel® processors, supporting high performance for the storage system as a whole, as well as for related applications.

Conclusion

The Huawei OceanStor 9000 big data storage system, based on the Intel Xeon processor E5-2400 product family, offers compelling performance, scalability, availability, and configurability for the enterprise. Taking advantage of hardware and software building blocks from Intel, this platform provides consistently robust results in configurations up to 288 nodes, with data sets up to 60 PB.

Learn more about Intel® Storage Technologies:
www.intel.com/storage

Learn more about Huawei storage products:
http://e.huawei.com/en/products/cloud-computing-dc/storage

[1] All SPEC SPECsfs2008_nfs.v3 Results Published by SPEC. http://www.spec.org/sfs2008/results/sfs2008nfs.html

[2] Testing by Nuawei and Intel in test laboratory facilities operated by Huawei.

Performance nodes (OPS density usage, 144 nodes): Intel® Xeon® processors E5-2402; memory: 48 GB, scale to 192 GB; disk configuration: 4 x 2.5-inch SSD/SAS hard disk drives and 21 x 2.5-inch 600 GB SAS hard disk drives; network adapters: Intel® Ethernet 1Gb/10Gb/40Gb and InfiniBand*.

Capacity nodes (big capacity and high bandwidth usage, 288 nodes): Intel® Xeon® processors E5-2402; memory: 48 GB, scale to 192 GB; disk configuration: 1 x 3.5-inch SSD/SAS hard disk drive and 35 x 3.5-inch 600 GB SAS hard disk drives; network adapters: Intel® Ethernet 1Gb/10Gb/40Gb and InfiniBand*.

Mini-capacity nodes (small capacity usage, 72 nodes): Intel® Xeon® processors E5-2402; memory: 32 GB, scale to 192 GB; disk configuration: 1 x 3.5-inch SSD/SAS hard disk drive and 11 x 3.5-inch 600 GB SAS hard disk drives; network adapters: Intel® Ethernet 1Gb/10Gb/40Gb and InfiniBand*.