Technical white paper

DataStax Enterprise on HP Moonshot System with HP ProLiant m710 Server Cartridges

Right-sizing Big Data deployments

Table of contents

Executive summary

The crush of Big Data has affected every company in every industry. As a result, enterprise adoption of new data storage, processing, and analytics platforms has become more widespread than ever before. As the ecosystem continues to change, enterprises may find it tough to keep up with the changing hardware requirements and best practices for deploying each individual software ecosystem. Many customers use a single hardware platform to deploy their Big Data applications, following a "bigger is better" paradigm. This can lead to extreme under-utilization of hardware, more in upfront costs to acquire these machines, and more in recurring power and associated data center costs than is necessary.

The DataStax Enterprise on HP Moonshot System with HP ProLiant m710 Server Cartridges solution was designed to tackle this problem.

DataStax Enterprise, built on top of Apache Cassandra, delivers what Internet enterprises need to compete in today's high-speed, Always-On, data-driven economy. Cassandra has become the leading distributed database technology, driving many of today's modern online applications. It offers continuous availability, high scalability and performance, strong security, and operational simplicity.

The ability of DataStax Enterprise to effortlessly scale with exceptional performance has proven it to be the perfect match for the revolutionary HP Moonshot System utilizing HP ProLiant m710 Server Cartridges.

Target audience: HP Technical Presales, solution architects, customers, and partners. A working knowledge of server architecture, networking architecture, and storage design is recommended.

Document purpose: The purpose of this document is to describe a recommended architecture/solution, highlighting recognizable benefits to technical audiences. This white paper describes testing performed by HP in September 2014.

An ideal platform for DataStax Enterprise

Right-sized for data storage and processing applications

DataStax Enterprise is the leading distributed database choice for online applications that require fast performance with no downtime. DataStax Enterprise provides in-memory computing, built-in security, powerful integrated analytics, enterprise search, and visual management. Software and solution architects are frequently tasked with choosing the best hardware platform on which to deploy their data storage and processing applications. Cloud-based services are certainly convenient, but performance and control are sacrificed in these environments. Dedicated hardware provides the levels of performance and control that are critical in high-volume, low-latency solutions. But choosing a hardware platform can be a daunting task, especially for those unfamiliar with the landscape. Some may follow a "bigger is better" approach that is common to other Big Data platforms; however, this is not necessarily the best approach when deploying DataStax Enterprise.

By providing compute units that match the characteristics of the application being deployed, costs are reduced along with power and space requirements. That's what makes the HP Moonshot System with HP ProLiant m710 Server Cartridges an ideal platform for deploying DataStax Enterprise.

HP Moonshot System with HP ProLiant m710 Server Cartridges

HP Moonshot System (figure 1) represents the next generation of server technology. It delivers breakthrough efficiency and scale by aligning just the right amount of compute, memory, and storage for specific workloads, delivered in a dense, power- and energy-efficient package. Built just for Moonshot, the HP ProLiant m710 Server Cartridge (figure 2) is based on the lower-power Intel® Haswell processing architecture, providing the right balance between performance and operational efficiency. The complete system supports the massive scalability required by modern, data-driven, online applications. It also offers reductions in power, space, and complexity compared to traditional rack servers.

Figure 1. HP Moonshot 1500 System

Figure 2. HP ProLiant m710 Server Cartridge

This white paper details how HP Moonshot with ProLiant m710 Server Cartridges can be deployed to meet the needs of customers wishing to integrate DataStax Enterprise into their Big Data ecosystems. As the HP Moonshot System is built to scale out, providing just the right amount of compute resources for the target workload, it is a perfect match for this environment.

Solution components

HP Moonshot System

In the HP Moonshot 1500 System, a 4.3U chassis houses two downlink switches, two uplink modules, fans, power supplies, a chassis management (HP iLO) module, and 15, 30, or 45 server cartridges.

Networking

To connect the HP Moonshot 1500 Chassis to the main network, an HP 5900-series switch is recommended, as was used in this testing. For this testing, one 40GbE QSFP port from each uplink switch was connected to the HP 5900 switch, along with several machines used to drive load to the platform under test, connected via 10GbE.

HP ProLiant m710 Server Cartridges

The HP ProLiant m710 Server Cartridge has the following specifications, as shown in table 1.

Table 1. HP ProLiant m710 Server Cartridge technical specifications

CPU	1 x 4-core (Haswell-H) Intel® Xeon® E3-1284L v3 processors
Memory	4 LV SODIMM (non-ECC) plus 8 ECC DRAM down; total 32 GB
Network	2 x integrated 10GbE NICs (one to switch A, one to switch B)
Storage	1 x solid-state storage module (120 GB, 240 GB, or 480 GB)

DataStax Enterprise and other software components

DataStax Enterprise, built on Apache Cassandra, is the distributed database technology and transactional backbone of choice for the world's most innovative companies. The "no single point of failure," peer-to-peer architecture lends itself very well to the HP Moonshot System. To benchmark this database platform, we chose to use Cassandra-stress. Cassandra-stress is packaged with all Cassandra distributions, both Apache and DataStax, and has been the *de facto* standard for benchmarking Cassandra. The Java-based tool sits outside of the Cassandra cluster, issuing reads and writes to the cluster, continuously reporting on the operation rate and latency that it is able to achieve.

Table 2 shows the versions of each software component in the solution. When deploying this solution in production, unless there are compelling reasons not to, one should typically use the newest generally available software component versions.

**Table 2.** Software component versions

Red Hat® Enterprise Linux® Operating System	6.5
DataStax Enterprise	4.0.3
Java	1.7
Cassandra-stress	2.1.0-rc1

Testing and results

Test hardware setup

Figure 3 shows the physical configuration of the hardware under test within the HP Lab. An HP Moonshot 1500 Chassis is fully loaded with 45 HP ProLiant m710 Server Cartridges. The 1500 Chassis is connected to a HP 5900AF-48XG-4QSFP+ Top of Rack (TOR) switch using a 40 Gb to 4 x 10 Gb DAC cable. Inside the HP Moonshot 1500 Chassis, we used 40 HP ProLiant m710 Server Cartridges to test the scale-out performance of DataStax Enterprise. This leaves five ProLiant m710 Server Cartridges free to use for other purposes, such as additional DataStax Enterprise nodes, supporting services such as HP Cluster Management Utility (CMU), application servers, etc. There are five HP ProLiant DL380p Gen8 Servers used as load generators running Cassandra-stress, connected to the TOR switch via 10 Gb DAC cable.

Figure 3. Lab hardware configuration
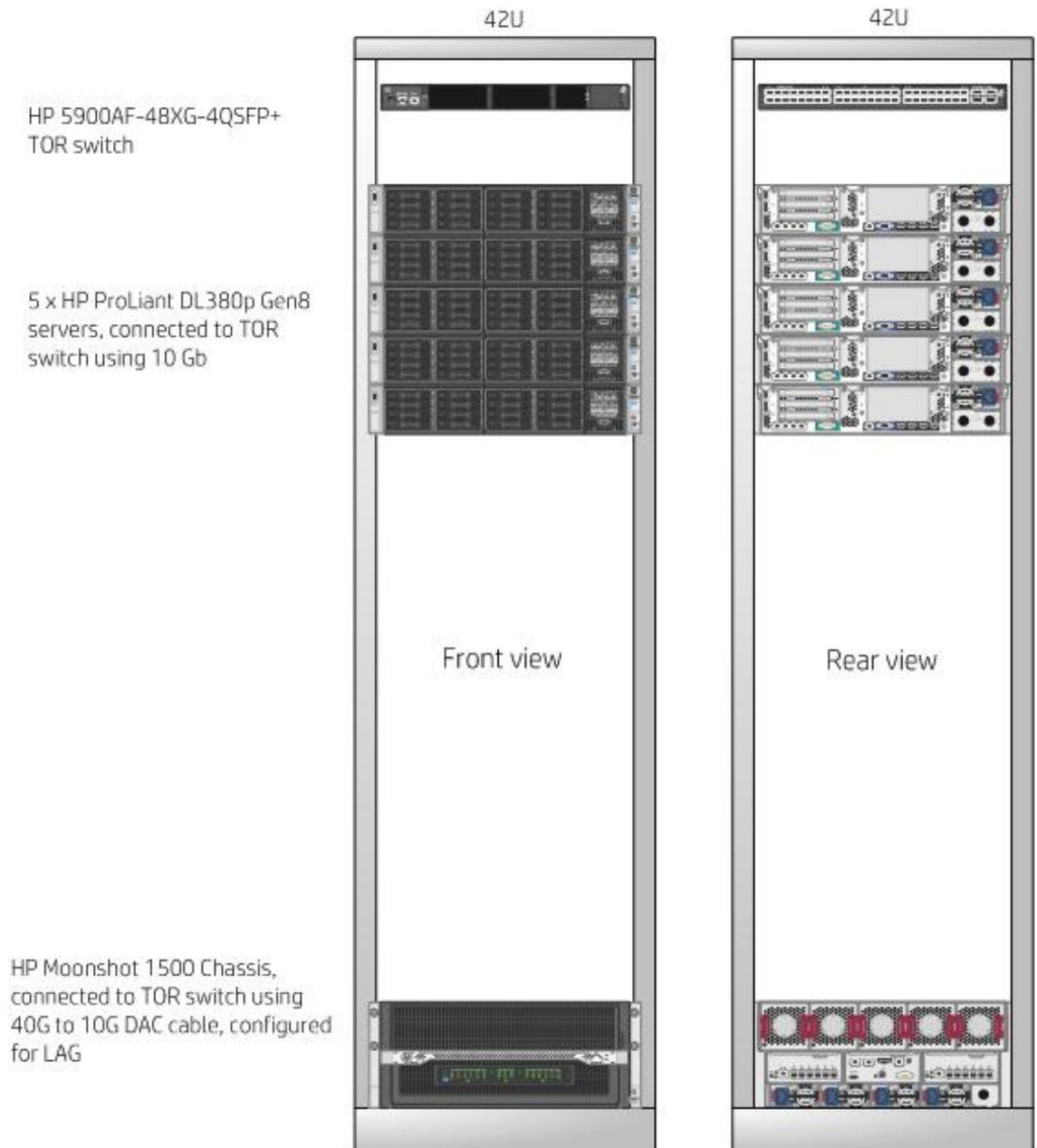
Workload description

For this test, the user first issues a write-only command to Cassandra-stress, which populates a Cassandra Keyspace. The user can then issue commands to begin testing various ratios of reads to writes, to simulate an actual production deployment. For this testing, we chose a combination of 10 distinct workloads that range from read-heavy, to balanced, to write-heavy. These are shown as follows, as a percentage of writes in the workload:

- 1 percent
- 3 percent
- 6 percent
- 10 percent
- 20 percent
- 50 percent
- 80 percent
- 94 percent
- 97 percent
- 99 percent

Results

Utilizing 40 HP ProLiant m710 Server Cartridges, with four machines used to drive client traffic with Cassandra-stress (500 threads per machine), we were able to achieve between 180,000 operations per second and 260,000 operations per second, as determined by the workload. Across workloads, the median operation rate was 226,820 operations per second. Figure 4 shows how each workload performed.

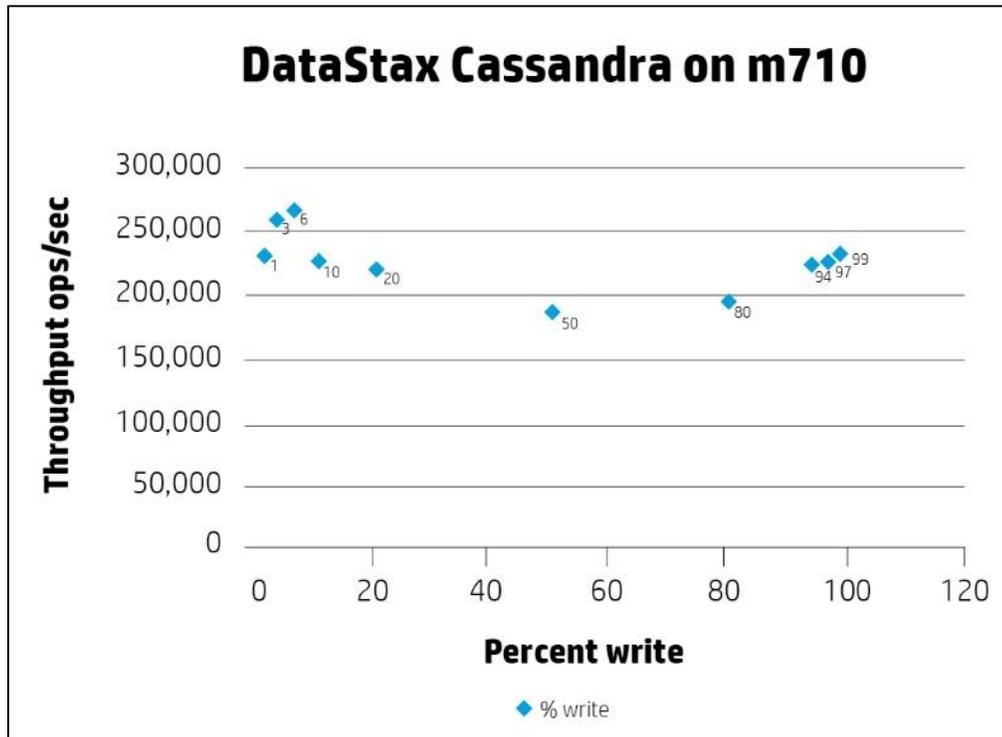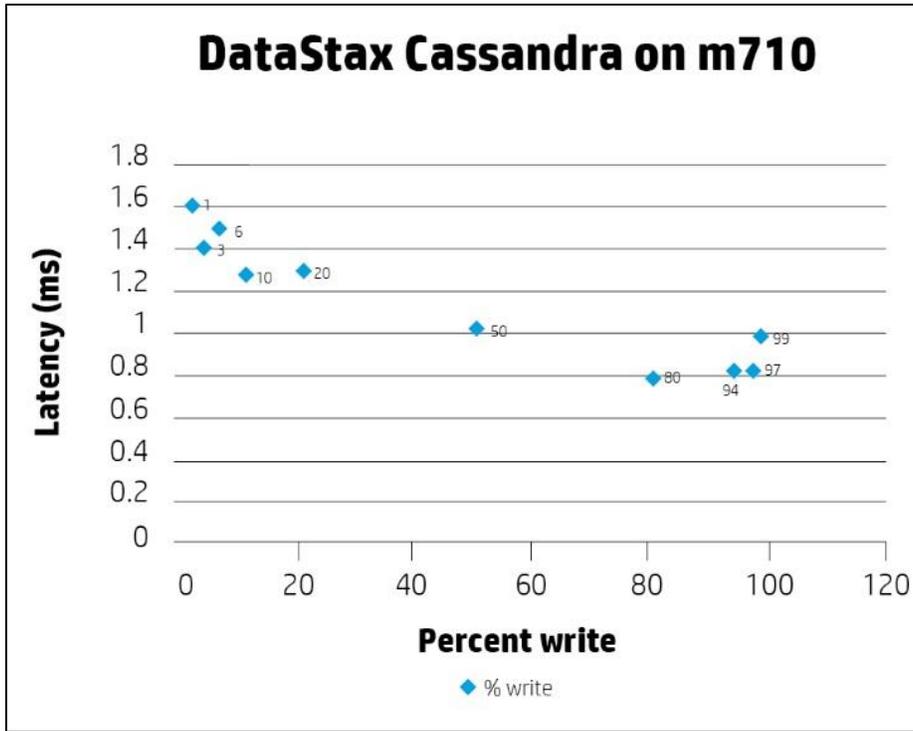**Figure 4.** Operations per second with various workloads

Figure 5 shows the median latency for each workload. We observed a cross-workload median latency of 1.15 ms.

Figure 5. Latency (in ms) with various workloads

DataStax Enterprise on the ProLiant m710 shows exceptional performance with respect to scalability, as measured separately both by adding nodes and by increasing the number of client threads. Figure 6 shows the scalability from 15 to 40 nodes, as represented by the throughput per node, running the 20 percent write workload.

We see a slight dip at 30 nodes, but this is attributed to the cluster being overly stressed during the 15 node test, as illustrated in figure 7, which shows the median latency decreasing substantially at larger node sizes.

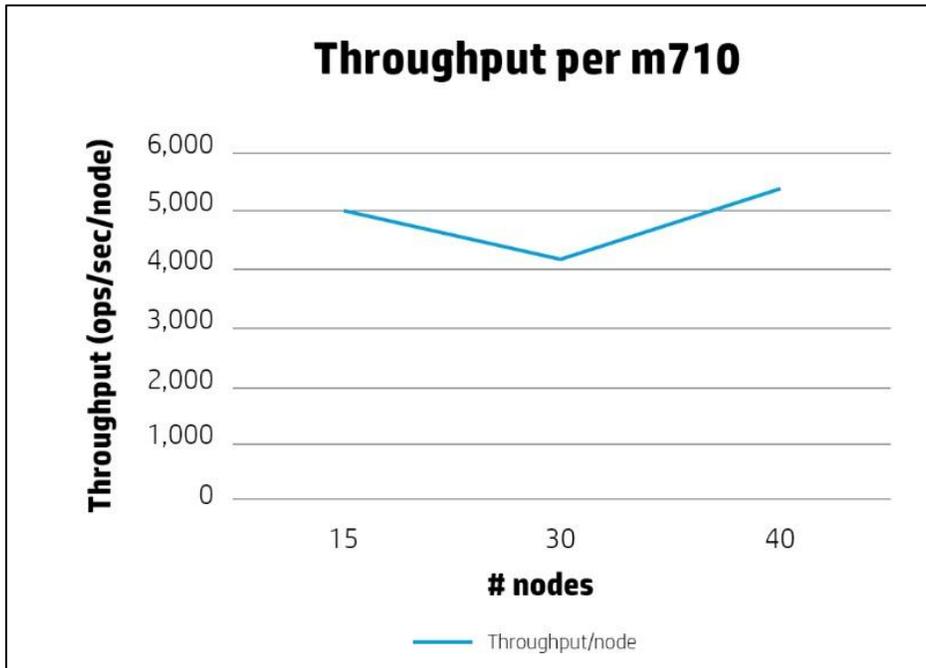**Figure 6.** Throughput when scaling number of nodes

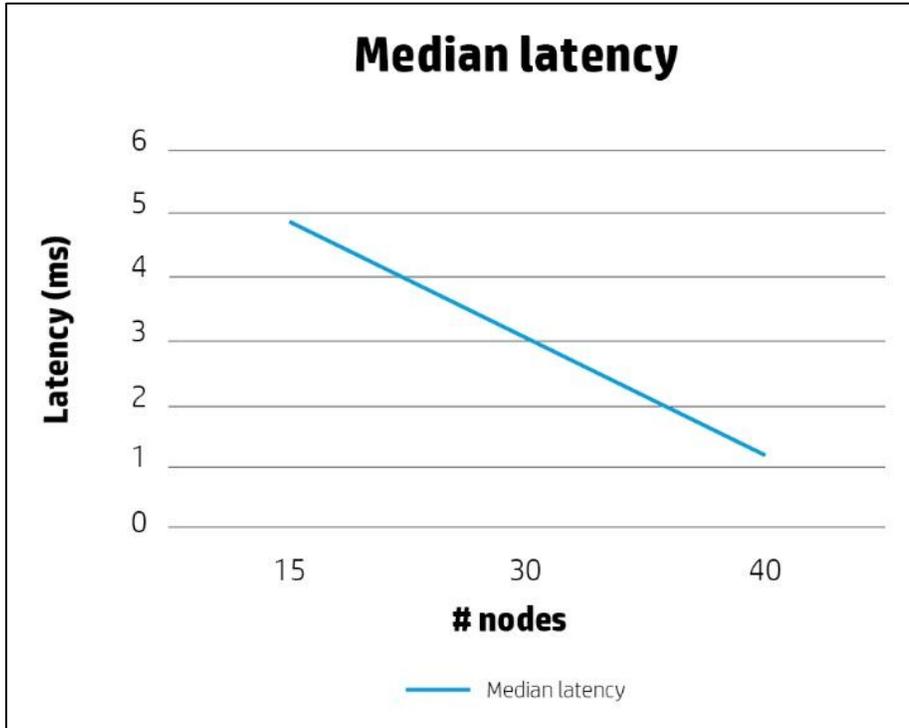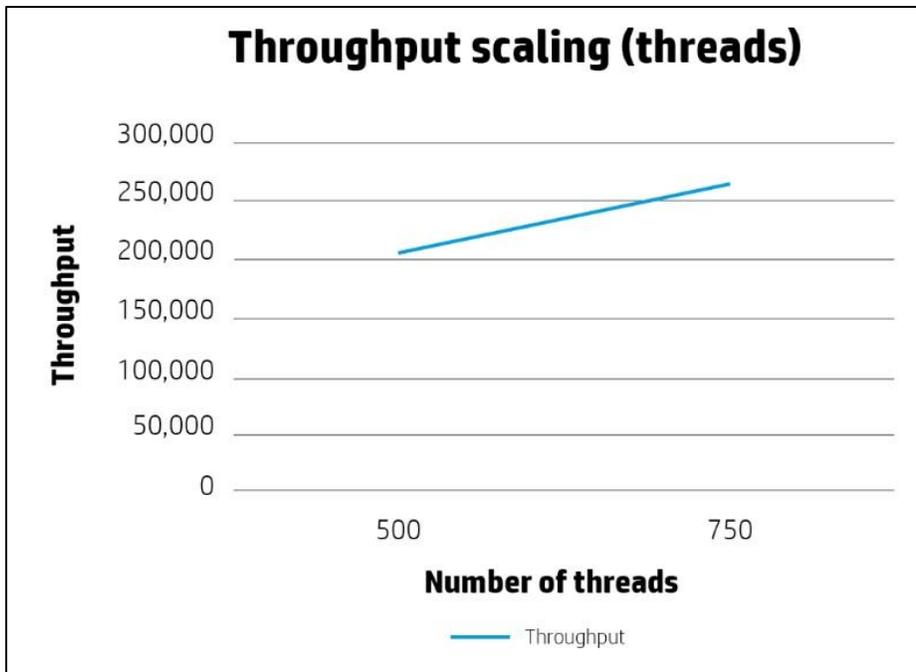**Figure 7.** Median latency when scaling number of nodes

Figure 8 represents the scalability of 40 nodes running the 10 percent write workload, when increasing the number of threads by 50 percent to simulate a spike in user activity.

Figure 8. Throughput with number of threads

Power draw under load is approximately 48 watts per cartridge, with an overhead of about 500 watts for the chassis. At idle, cartridges draw around 21 watts, with the same chassis overhead. This compares to a single HP ProLiant DL380p Server drawing approximately 300 watts under the load of the same software stack and 200 watts at idle. With a national average of 12.2 cents per kWh, this power savings will add up quickly, providing for tremendous power savings over traditional hardware platforms.

Analysis and recommendations

With a cross-workload median throughput of around 5,000 ops/sec/node, at a cross-workload median latency of 1.15 ms, DataStax Enterprise on the HP Moonshot System with ProLiant m710 Server Cartridges is a very viable alternative to traditional enterprise-class server hardware. With the ProLiant m710, in many workloads, a larger percentage of the compute resources are used, leading to a lower idle percent CPU utilization, meaning that customers will not pay for resources that their application will rarely, if ever, utilize.

The HP ProLiant m710 provides the perfect balance of compute, storage, memory, and network bandwidth for DataStax Enterprise. We recommend that customers choose this solution to reduce wasted compute resources and increase savings in power and rack space.

Configuration guidance

Uplink module

When configuring the solution, the HP Moonshot chassis uplink module should be connected directly to the TOR switch using a 40 Gb QSFP port. If your TOR switch does not have 40 Gb ports, you may use a splitter cable that brings 4 x 10 Gb SFP ports into one 40 Gb QSFP port, and configure the switch for a Link Aggregation Group (LAG). In a multiple-chassis configuration, each uplink module should be connected to the TOR switch in a manner similar to above. Alternatively, chassis uplink modules may be stacked together. Please refer to the Moonshot Networking Guide for more information.

Operating system

There are quite a few ways to deploy the operating system (OS) to HP Moonshot. We recommend using familiar configuration management utilities, following the best practices of the organization. For the work done in this paper, HP CMU was used for OS deployment. HP CMU makes it easy to backup and deploy complete drive images over the network using PXE. First, we created a "golden image" on one node containing the OS-level modifications (detailed in appendix A) and packages to install, then backed up that node, and deployed it to the entire chassis. This ensures that all settings are the same across nodes.

DataStax Enterprise

DataStax Enterprise can be deployed several ways: via tarball, via OS-specific packages (rpm, deb), or via DataStax OpsCenter. DataStax OpsCenter is a visual management console, simplifying the control of complex workloads (figure 9). OpsCenter makes it easy to manage Apache Cassandra and DataStax Enterprise clusters by giving administrators, architects, and developers an at-a-glance view of the system from a centralized dashboard. OpsCenter installs seamlessly and gives system operators the flexibility to monitor and manage even the most complex workloads with ease from any Web browser. We highly recommend using this tool to manage your DataStax environment.

For our testing, we installed OpsCenter, and used it to deploy the Cassandra cluster. To deploy another way, please follow the instructions here: datastax.com/download.

Figure 9. DataStax OpsCenter
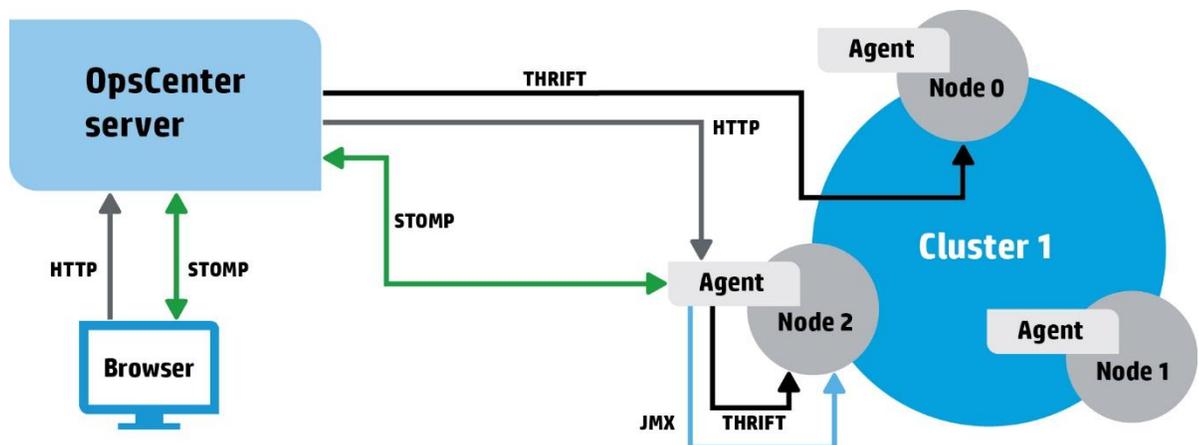
To install DataStax OpsCenter and deploy DataStax Enterprise:

1. Visit datastax.com/download to register for DataStax Enterprise

2. Make sure your OS package manager is configured, and you have set any necessary parameters for Proxy and Internet gateway

3. Configure your package manager to access the DataStax repository applicable to your OS, then install and start OpsCenter:

   A. Red Hat Enterprise Linux: datastax.com/documentation/opscenter/5.0/opsc/install/opscInstallRHEL_t.html

   B. Debian: datastax.com/documentation/opscenter/5.0/opsc/install/opscInstallDeb_t.html

   C. Other Linux: datastax.com/documentation/opscenter/5.0/opsc/install/opscInstallTar_t.html

4. After starting OpsCenter, navigate to the OpsCenter Web GUI: opscenter-host:8888/(replace opscenter-host with the appropriate IP address/hostname)

5. Select "Create Brand New Cluster," and follow the prompts, entering your DSE credentials, the IP address details of your systems, and your system credentials

6. Make sure to check the "Advanced" tab for any specific settings applicable to your environment, such as data directory location, and enable Native Transport if your application requires it

7. Select "Build Cluster"

8. Packages will be installed via the Internet, and the cluster will be functional within a few minutes

Bill of materials

Note

Part numbers are subject to change. The bill of materials does not include complete support options or other rack and power requirements. If you have questions regarding ordering, please consult your HP sales representative or authorized reseller for more details. hp.com/large/contact/enterprise/index.html

Table 3. Bill of materials

Quantity	Part number	Description
1	755371-B21	HP Moonshot 1500 Chassis
3	AF556A	HP 1.83m 10A C13-UL US Power Cord
3	684532-B21	HP 1500W Hot Plug Power Supply Kit
1*	704652-B21	HP Moonshot-4QSFP+ Uplink Module Kit
1*	704654-B21	HP Moonshot-45XGc Switch Kit
1	681254-B21	HP 4.3U Rail Kit
1	681260-B21	HP 0.66U Spacer Blank Kit
45	755860-B21	HP ProLiant m710 Server Cartridge
45	765483-B21	HP Moonshot 480G SATA VE M.2 2280FIO Kit

* If switch redundancy is desired, add one more of both 704652-B21 and 704654-B21.

Appendix A: Tuning parameters used

In rc.local:

echo deadline > /sys/block/sda/queue/scheduler

blockdev --setra 0 /dev/sda

echo 0 > /sys/block/sda/queue/rotational

In /etc/security/limits.conf (if not already in /etc/security/limits.d/cassandra.conf)

* - memlock unlimited

* - nofile 100000

* - nproc 32768

* - as unlimited

In /etc/security/limits.d/90-nproc.conf

* - nproc 32768

In /etc/sysctl.conf

vm.max_map_count = 131072

sudo swapoff --all

For testing using Cassandra-stress, we chose a replication factor of three, and we used the "LZ4Compressor."

Enabling compression:

cqlsh

use "Keyspace1";

ALTER TABLE "Standard1"

WITH compression =

{ 'sstable_compression' : 'LZ4Compressor'};

Fstab settings:

noatime,nodiratime,discard,errors=remount-ro 1 2

List of up-to-date DataStax Enterprise tuning parameters

datastax.com/documentation/cassandra/2.0/cassandra/install/installRecommendSettings.html

support.datastax.com/entries/37991806-Some-tips-on-getting-the-most-IO-performance-out-of-SSDs

Appendix B: Glossary

CMU   Cluster Management Utility

DSE   DataStax Enterprise

LAG   Link Aggregation Group

PXE   Preboot eXecution environment

QSFP   Quad Small Form Factor

TOR   Top of Rack

Summary

Not all Big Data applications are created equal—it requires deep analysis to understand the characteristics of the software, and how those characteristics translate into server hardware requirements. Many customers may be tempted to use a single hardware platform on which to deploy their Big Data applications, but as we have seen, this is a less-than-optimal approach.

As detailed in this paper, the HP Moonshot System with HP ProLiant m710 Server Cartridges provides the perfect balance of compute, storage, memory, and network bandwidth for DataStax Enterprise, leading to a solution that increases efficiencies in power and physical space, with a reduction in complexity. These efficiencies provide for a lower overall total cost of ownership and are achieved with minimal performance impact over traditional servers.

Implementing a proof-of-concept

As a matter of best practice for all deployments, HP recommends implementing a proof-of-concept using a test environment that matches the planned production environment as closely as possible. In this way, appropriate performance and scalability characterizations can be obtained. For help with a proof-of-concept, contact an HP services representative (hp.com/large/contact/enterprise/index.html) or your HP partner.

Learn more at
hp.com/go/moonshot
datastax.com

Sign up for updates
hp.com/go/getupdated

Share with colleagues

Rate this document