Optimized Infrastructure for Big Data Analytics on MongoDB from Hitachi Vantara

Reference Architecture Guide

By David Pascuzzi (Hitachi Vantara) and Abhinav Mehla (MongoDB)

September 2018
Feedback

Hitachi Vantara welcomes your feedback. Please share your thoughts by sending an email message to SolutionLab@HitachiVantara.com. To assist the routing of this message, use the paper number in the subject and the title of this white paper in the text.

Revision History

<table>
<thead>
<tr>
<th>Revision</th>
<th>Changes</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL-027-00</td>
<td>Initial release</td>
<td>February 15, 2018</td>
</tr>
<tr>
<td>MK-SL-027-01</td>
<td>Added support for the Hitachi Advanced Server DS220.</td>
<td>September 8, 2018</td>
</tr>
</tbody>
</table>
# Table of Contents

**Key Solution Elements**  
- Hardware Elements  
- Software Elements  

**Solution Design**  
- Hitachi Advanced Server DS120 Configuration  
- Mongod Nodes Configuration  
- MongoDB Config Server  
- Mongos Node  
- MongoDB OpsManager Node  
- Processing Node  
- Hardware Management Server  
- Network Architecture  
- Operating System  
- Sample Rack Configuration  

**Engineering Validation**  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Solution Elements</td>
<td>2</td>
</tr>
<tr>
<td>Hardware Elements</td>
<td>2</td>
</tr>
<tr>
<td>Software Elements</td>
<td>3</td>
</tr>
<tr>
<td>Solution Design</td>
<td>9</td>
</tr>
<tr>
<td>Hitachi Advanced Server DS120 Configuration</td>
<td>9</td>
</tr>
<tr>
<td>Mongod Nodes Configuration</td>
<td>10</td>
</tr>
<tr>
<td>MongoDB Config Server</td>
<td>13</td>
</tr>
<tr>
<td>Mongos Node</td>
<td>14</td>
</tr>
<tr>
<td>MongoDB OpsManager Node</td>
<td>15</td>
</tr>
<tr>
<td>Processing Node</td>
<td>15</td>
</tr>
<tr>
<td>Hardware Management Server</td>
<td>16</td>
</tr>
<tr>
<td>Network Architecture</td>
<td>16</td>
</tr>
<tr>
<td>Operating System</td>
<td>18</td>
</tr>
<tr>
<td>Sample Rack Configuration</td>
<td>18</td>
</tr>
<tr>
<td>Engineering Validation</td>
<td>21</td>
</tr>
</tbody>
</table>
Optimized Infrastructure for Big Data Analytics on MongoDB from Hitachi Vantara

Reference Architecture Guide

Accelerate your deployment by leveraging this reference architecture to help guide you in implementing Hitachi Vantara's optimized infrastructure for big data analytics on MongoDB. Reduce the risk of implementing an improper architecture with this guide.

Use this reference architecture guide to see how to configure our optimized infrastructure for big data analytics on MongoDB. This shows an example environment on how to deploy big data infrastructure software for advanced data analytics.

This integrated big data infrastructure uses rack-optimized servers. There are guidelines to create the best infrastructure for your big data analytics projects with these products:

- **Hitachi Advanced Server DS120** — This is a flexible 1U server designed for optimal performance across multiple applications.
- **Hitachi Advanced Server DS220** - This is a flexible 2U server designed for optimal performance across multiple application.
- **MongoDB Enterprise** — This is a leading modern, general purpose database platform.
- **Cisco Nexus 3048 switch** — This 48-port 1 GbE switch provides a management network as a leaf switch and a spine switch.
- **Cisco Nexus 93180YC-E/FX**— This 48-port switch provides 10 GbE connectivity for intra-rack networks. It is used as the leaf switch for the data network. Designed with Cisco Cloud Scale technology, it supports highly scalable cloud architectures.
- **Cisco Nexus 93180LC-EX**— This 24-port switch provides 40 GbE connectivity for inter-rack networks, it is used as the spine switch for the data network and supports flexible migration options. The switch is ideal for highly scalable cloud architectures and enterprise data center.

This reference architecture is certified by MongoDB.

**Note** — Testing of this configuration was done in a lab environment. Many factors affect production environments that are beyond prediction or hard to duplication in a lab environment. Follow the recommended practice of conducting proof-of-concept testing for acceptable results in a non-production, isolated test environment that otherwise matches your production environment before you implement this solution in production.
Key Solution Elements

These are the key hardware and software components this integrated solution uses. Use this example environment to deploy an optimized big data infrastructure for advanced analytics on MongoDB. As these systems are complex to design, contact Hitachi Vantara and MongoDB prior to planning your deployment.

Hardware Elements

These key hardware elements power this big data solution. You can create a scale-out configuration to power your Hadoop needs.

*Hitachi Advanced Server DS120*

Optimized for performance, high density, and power efficiency in a dual-processor server, Hitachi Advanced Server DS120 delivers a balance of compute and storage capacity. This rack mounted server has the flexibility to power a wide range of solutions and applications.

The highly-scalable memory supports up to 3 TB RAM using 24 slots of 2666 MHz DDR4 RDMM. DS120 is powered by the Intel Xeon scalable processor family for complex and demanding workloads. There are flexible OCP and PCIe I/O expansion card options available. This server supports up to 12 storage devices with up to 4 NVMe.

Figure 1 shows the front and back of this server.

![Figure 1](image)

*Hitachi Advanced Server DS220*

Optimized for performance, high density, and power efficiency in a dual-processor server, Hitachi Advanced Server DS220 delivers a balance of compute and storage capacity. This rack mounted server has the flexibility to power a wide range of solutions and applications.

The highly-scalable memory supports up to 3 TB RAM using 24 slots of 2666 MHz DDR4 RDMM. DS220 is powered by the Intel Xeon scalable processor family for complex and demanding workloads. There are flexible OCP and PCIe I/O expansion card options available. This server supports up to 12 large form factor storage devices and an additional 2 small form factor storage devices.

This 2U solution using large form factor storage allows you to have a dense storage solution with lower power consumption. This is ideal for archival and backup nodes. This larger form factor also provides you with more expansion options.
Figure 2 shows front and back views of the DS220.

Figure 2

Cisco Nexus Data Center Switches

Cisco Nexus data center switches are built for scale, industry-leading automation, programmability, and real-time visibility.

This solution uses the following Cisco switches to provide Ethernet connectivity:

- Cisco Nexus 3048
- Cisco Nexus 93180YC-E/FX
- Cisco Nexus 93180LC-EX

This architecture uses a leaf-spine network configuration. You can use other configurations to best fit your environment.

Software Elements

These key software elements power this big data solution.

Operating Systems

The environment in the reference architecture uses RedHat Enterprise Linux 7.3.

MongoDB Enterprise Server

MongoDB Enterprise is a highly performant, scalable, resilient, and secure modern data management platform, designed to meet the challenge of today’s digital economy. It uses document model to store data in flexible, JSON-like documents. This reference architecture is verified using MongoDB Enterprise 3.4.xxx. When deploying it is recommended that the latest stable production release of MongoDB Enterprise is used. It can be downloaded from the MongoDB Download Center.

With distributed design and non-relational data model from MongoDB, architects can deliver applications that were never possible with traditional relational databases:

- **Fast, Iterative Development.** Scope creep and changing business requirements no longer stand between architecture teams and successful project delivery. A flexible data model coupled with dynamic schema and idiomatic, native language drivers make it fast for developers to build and evolve applications. Automated provisioning and management enable continuous integration and highly productive operations.
- **Flexible Data Model.** MongoDB’s document data model makes it easy for developers to store and combine data of any structure, without giving up sophisticated governance controls, data access and rich indexing functionality. The schema can be dynamically modified without application downtime. As a result, development teams spend less time preparing data for the database, and more time putting data to work.

- **Multi-Region Scalability.** MongoDB can be scaled within and across geographically distributed data centers and cloud environments, providing new levels of availability and scalability to meet the demands of customers, wherever they are. As deployments grow in data volume and throughput, MongoDB scales easily with no downtime, and without changing the application.

- **Multi-model Feature Set to Build Rich Operational and Analytical Apps.** Analytics and data visualization, text search, graph processing, geospatial, in-memory performance and global replication allow architects to deliver a wide variety of real-time applications on one technology, reliably and securely. No need to install multiple databases to meet the needs of different applications. No need to move data into expensive data warehouses, or complex and un governed data lakes, to extract insight and value from data.

- **Lower TCO.** Application development teams are more productive when they use MongoDB. Single click management means operations teams are as well. MongoDB runs on commodity hardware in your own data center, in the cloud, or as a fully-managed service, dramatically lowering costs. Finally, MongoDB offers affordable annual subscriptions, including 24x7x365 global support. Applications can be one tenth the cost to deliver compared to using traditional relational databases.

- MongoDB Enterprise features extensive security and governance capabilities to defend, detect, and control access to data. Please refer to this whitepaper on [MongoDB Security Architecture](#) to learn more. Before deploying a system in production verify that the system follows the current [MongoDB Production Notes](#) and [Operations Checklist](#) guidelines.

Below are some of the key concepts that should be understood to use and operate MongoDB deployments. To find out more details see the [MongoDB Manual](#).

- **Mongod** — This is a primary daemon that handles data requests, manages data access, and performs background operations.

- **Mongos** — A routing service that processes queries from the application layer, it determines the location of data in a sharded MongoDB cluster.

- **Mongo shell** — An interactive interface to MongoDB that you can use to query and update data, as well as perform administrative operations.

- **Replica set** — This is a group of mongod processes that maintain the same data set. Replica sets provide redundancy and high availability, and are the basis for all production grade deployments. It is recommended to maintain at least three replicas or copies of data for failover purposes.

- **Shard** — This is a single mongod instance or replica set that stores some portion of the total data set. It is used to scale out or partition a collection across multiple nodes.

- **Config Server** — This is a mongod instance that stores all the metadata associated with a sharded cluster. It contains information about data location across the shards. One set of config servers are required for each sharded cluster.

- **Collection** — A grouping of MongoDB documents that, broadly speaking, can be regarded as the equivalent of a RDBMS table.
A replica set provides resiliency. Figure 3 is a replica set with one primary and two secondary nodes. All writes go through the primary member. A minimum of a three-node replica set is required for a production system. For information on how MongoDB’s replicate sets work and their features include high availability, see the “Replication” in the online MongoDB user documentation. This document only covers primary members and secondary data-bearing members.

Figure 3

Use a shard to horizontally scale MongoDB. It allows you to split the data in a collection across multiple replica sets. Using shards introduces a router (mongos) into the solution along with config servers.
A config server keeps track of what data resides in each shard. Figure 4 shows a configuration with three shards. Each shard can have different number of nodes and a different hardware configuration of its members. For more information on replica sets and recommended deployment architectures, see high availability and automatic failover documentation from MongoDB.

**Figure 4**
**MongoDB Ops Manager**

*MongoDB Ops Manager* is the best way to run MongoDB within the data center. This makes it fast to deploy, monitor, back up and scale MongoDB. It incorporates best practices to help keep managed databases healthy and optimized. Ops Manager ensures operational continuity by converting complex manual tasks into reliable, automated procedures with the click of a button. Ops Manager assists with the following:

- **Deployment** — Deploy any topology, at any scale.
- **Management** — Deploy new clusters. Manage, monitor, and back up existing ones.
- **Upgrades** — Upgrade in minutes, with no downtime.
- **Scaling** — Add capacity, without taking the application offline.
- **Point-in-time, Scheduled Backups** — Restore to any point in time, because disasters are not predictable.
- **Performance Alerts** — Monitor over 100 system metrics, getting custom alerts before the system degrades.
- **Query Optimization** — Identify slow-running queries, get index suggestions, and automate index builds.

**MongoDB Cloud Manager**

*MongoDB Cloud Manager* is a cloud-hosted alternative to MongoDB Ops Manager. With automated provisioning, fine-grained monitoring, and continuous backups, you get a full management suite to reduced operational overhead, while maintaining full control over your MongoDB deployments.

**MongoDB Compass**

*MongoDB Compass* is a graphical user interface for MongoDB. Offering rich schema exploration and management, Compass allows you to modify documents, create validation rules, and efficiently optimize query performance by visualizing explain plans and index usage.

Build and execute sophisticated queries by selecting document elements from the user interface. View the results graphically and as a set of JSON documents. Accomplish these tasks from a point and click interface, without knowledge of MongoDB's query language.

MongoDB Compass is included with MongoDB Professional and MongoDB Enterprise Advanced subscriptions.

**MongoDB Connector for BI**

SQL-based business intelligence and analytics tools expect to see the data in tabular format. *MongoDB Connector for BI* lets you use SQL-based tools without sacrificing MongoDB’s flexible data model to store data in rich, multi-dimensional documents and quickly build new functionalities. MongoDB Connector for BI acts as a layer that passes queries and data between a MongoDB instance and your reporting tool. It stores no data, only bridging your MongoDB server with business intelligence tools.
An increasing number of businesses are leveraging Apache Spark with MongoDB to perform advanced analytics. MongoDB Connector for Apache Spark makes it easy and efficient. The connector, Figure 5, exposes all of Spark's libraries, including Scala, Java, Python and R. MongoDB data is materialized as DataFrames and Datasets for analysis with machine learning, graph, streaming, and SQL APIs.

Figure 5
Solution Design

This is the detailed design for an integrated infrastructure from Hitachi to implement big data and business analytics using MongoDB.

- "Hitachi Advanced Server DS120 Configuration" on page 9
- "Mongod Nodes Configuration" on page 10
- "MongoDB Config Server" on page 13
- "Mongos Node" on page 14
- "MongoDB OpsManager Node" on page 15
- "Processing Node" on page 15
- "Hardware Management Server" on page 16
- "Network Architecture" on page 16
- "Operating System ," starting on page 18
- "Sample Rack Configuration " on page 18

Hitachi Advanced Server DS120 Configuration

This solution uses multiple Hitachi Advanced Server DS120 servers placed in 42U racks. After reserving space for switches, this solution can accommodate up to 36 nodes per rack.

There are different types of nodes and configuration's serving different purposes. This reference architecture focuses on the nodes running the Mongod service and the config server.

Table 1 lists the hardware components this reference architecture uses for the various mongod nodes. For help in sizing your system, contact Hitachi Vantara Sales or MongoDB Sales.

TABLE 1. HARDWARE OPTIONS

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>Hitachi Advanced Server DS120</td>
</tr>
<tr>
<td>CPU</td>
<td>Use one of the following options when deploying the server for this solution:</td>
</tr>
<tr>
<td></td>
<td>- 2 Intel 4110 8 core processor, 85 W; 2.1 GHz</td>
</tr>
<tr>
<td></td>
<td>- 2 Intel 6128, 6-core processor, 3.4 GHz</td>
</tr>
<tr>
<td></td>
<td>- 2 Intel 6140, 18-core processor, 2.3 GHz</td>
</tr>
</tbody>
</table>
## Mongod Nodes Configuration

Depending on your needs, multiple hardware configurations are available.

In a single replica set configuration, there is a primary and at least two secondary nodes. As the secondary nodes provide high availability, they always use the same hardware configuration as the primary nodes. In a sharded cluster, there will be one or more replica sets of at least one mongos node, and three or more config servers.

The standard configuration listed in Table 2, “Standard Configuration for Mongod Nodes,” on page 11 is a good starting place for mongod nodes to meet the needs of most deployments.

### Memory Options

Use one of the following memory options when deploying the server for this solution:

- 64 GB — 2 × 32 GB DIMMs
- 128 GB — 4 × 32 GB DIMMs
- 256 GB — 8 × 32 GB DIMMs
- 512 GB — 16 × 32 GB DIMMs

### Network Connections

Use the following ports:

- 1 Intel XXV710 dual port 10/25 GbE port
- (Optional additional) 1 Intel XXV710 dual port 10/25 GbE port
- 1 GbE LOM management port

### Disk Controllers

1 LSI 3516 RAID controller

### Operating System

2 × 128 GB SATADOM, configured as RAID-1

### Disk

Use only one of the following options:

- 1 to 12 SSD, from 480 GB to 3.84 TB
- 1 to 4 NVMe, sized 1 TB to 4 TB

### Rack

42U Minkel Rack

### Switches

Use the following switches in this solution:

- Cisco Nexus 93180YC-E/FX
- Cisco Nexus 93180LC-EX
- Cisco Nexus 3048

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory Options</td>
<td>Use one of the following memory options when deploying the server for this solution:</td>
</tr>
<tr>
<td>Network Connections</td>
<td>Use the following ports:</td>
</tr>
<tr>
<td>Disk Controllers</td>
<td>1 LSI 3516 RAID controller</td>
</tr>
<tr>
<td>Operating System Storage</td>
<td>2 × 128 GB SATADOM, configured as RAID-1</td>
</tr>
<tr>
<td>Disk</td>
<td>Use only one of the following options:</td>
</tr>
<tr>
<td>Rack</td>
<td>42U Minkel Rack</td>
</tr>
<tr>
<td>Switches</td>
<td>Use the following switches in this solution:</td>
</tr>
</tbody>
</table>

**TABLE 1. HARDWARE OPTIONS (CONTINUED)**

Component | Description
--- | ---
Memory Options | Use one of the following memory options when deploying the server for this solution:
- 64 GB — 2 × 32 GB DIMMs
- 128 GB — 4 × 32 GB DIMMs
- 256 GB — 8 × 32 GB DIMMs
- 512 GB — 16 × 32 GB DIMMs
Network Connections | Use the following ports:
- 1 Intel XXV710 dual port 10/25 GbE port
- (Optional additional) 1 Intel XXV710 dual port 10/25 GbE port
- 1 GbE LOM management port
Disk Controllers | 1 LSI 3516 RAID controller
Operating System Storage | 2 × 128 GB SATADOM, configured as RAID-1
Disk | Use only one of the following options:
- 1 to 12 SSD, from 480 GB to 3.84 TB
- 1 to 4 NVMe, sized 1 TB to 4 TB
Rack | 42U Minkel Rack
Switches | Use the following switches in this solution:
- Cisco Nexus 93180YC-E/FX
- Cisco Nexus 93180LC-EX
- Cisco Nexus 3048
Configure each node running any MongoDB component (that is, mongod, mongos, MongoDB Ops Manager) according to the MongoDB Production Notes which outline the latest best practices for operating system and MongoDB process configuration. Click to [read the latest production notes](#).

Never run any production MongoDB system without validating that all recommendations in the production notes are satisfied. Additionally, consider engaging MongoDB Global Consulting Services to conduct a comprehensive review of your MongoDB deployments. Contact your account manager and see MongoDB Global Consulting Services for additional information.

**TABLE 2. STANDARD CONFIGURATION FOR MONGOD NODES**

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>Hitachi Advanced Server DS120</td>
</tr>
<tr>
<td>CPU</td>
<td>2 Intel 4110 8 core processor, 85 W, 2.1 GHz</td>
</tr>
<tr>
<td>Memory Options</td>
<td>128 GB — 4 × 32 GB DIMMs</td>
</tr>
<tr>
<td>Network Connections</td>
<td>Use these ports:</td>
</tr>
<tr>
<td></td>
<td>1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>(Optional additional) 1 Intel XXV710 dual-port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>1 x GbE LOM management port</td>
</tr>
<tr>
<td>Disk Controllers</td>
<td>1 LSI 3516 RAID controller</td>
</tr>
<tr>
<td>Operating System</td>
<td>2 128 GB SATADOM, configured as RAID-1</td>
</tr>
<tr>
<td>Storage</td>
<td>4 SSD, 960 GB, configured as RAID-10</td>
</tr>
</tbody>
</table>

Change this configuration based on the following:

- “Memory Sizing Options for Mongod” on page 11
- “Storage Options for the Mongod Nodes ” on page 12
- “CPU Options for the Mongod Node” on page 13

**Memory Sizing Options for Mongod**

Different configurations can be used to support a system running MongoDB. Understanding your working set—data that is actively be used—is needed for high performance systems. Use the following recommendations for sizing memory:

- For best performance, the working set should be less than the total memory of the node.
- Index size is 15-30% of total data size.
- Data in the working set size depends on the application.
Deciding on memory sizes has many factors. In most cases, 128 GB is a good starting point. These are some of the factors that can influence your memory size:

- Working set size
- Performance requirements
- Concurrent users
- Throughput
- Data set size
- Application access pattern
- Number of nodes in a replica set
- Shard key

**Storage Options for the Mongod Nodes**

The storage can use either SSD or NVMe. In this reference architecture, the standard configuration uses SSDs. This provides high performance for a lower cost.

Perform storage high availability and redundancy through normal MongoDB replica sets. Configure storage as RAID-10 for an extra level of availability.

The following is the file system configuration:

- XFS
- Storage layout
  - /data — Primary mount point all for storage
  - /data/journal — Secondary mount point for journal
  - /data/log — Tertiary mount point for log
- Standard RAID configurations
  - 1 storage device — No RAID, mounted /data
  - 2 storage devices — No RAID disk, mounted /data and /data/journal
  - 3 storage devices — No RAID, mounted /data, /data/journal, /data/log
  - 4 or more storage devices — RAID-10, mounted /data
- Other storage configurations available
  - RAID-0
  - NVMe devices
  - Multiple RAID-10 devices
**CPU Options for the Mongod Node**

For the mongod nodes, three options are provided for CPUs:

- 2 Intel 4110 8-core processors is the standard option. This option can process 10s to 100s of operations per second.
- 2 Intel 6128 6-core processors. This option is used when there are relatively few number of concurrent transactions with the speed of a single transaction is foremost priority.
- 2 Intel 6140, 18-core processors. This option is used when there is a very high number of concurrent transactions.

**MongoDB Config Server**

Use the config server to manage queries across a sharded cluster. It stores the metadata that reflects the state and organization of the data in the system, used with distributed lock management.

As shown in Table 3, every sharded cluster must have its set of config servers. A config server is for one shared cluster. For redundancy and consistency, configure the config servers in at least a three-node replica set. Table 3 list the configuration of a config server.

**TABLE 3. CONFIG SERVER CONFIGURATION**

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>Hitachi Advanced Server DS120</td>
</tr>
<tr>
<td>CPU</td>
<td>2 Intel 4110 8 core processor, 85 W, 2.1 GHz</td>
</tr>
<tr>
<td>Memory Options</td>
<td>64 GB — 2 × 32 GB DIMMs</td>
</tr>
<tr>
<td>Network Connections</td>
<td>Use these ports:</td>
</tr>
<tr>
<td></td>
<td>1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>(Optional additional) 1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>1 GbE LOM management port</td>
</tr>
<tr>
<td>Disk Controllers</td>
<td>LSI 3516 RAID controller</td>
</tr>
<tr>
<td>Operating System Storage</td>
<td>2 × 128 GB SATADOM, configured as RAID-1</td>
</tr>
<tr>
<td>Storage</td>
<td>1 × 480 GB SSD</td>
</tr>
</tbody>
</table>
Mongos Node
Even though mongos do not need to have its own nodes, running on the same node as application servers, the recommendation is to have separate servers for mongos. When providing a separate mongos node, the standard configuration is in Table 4.

TABLE 4. MONGOS CONFIGURATION

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>Hitachi Advanced Server DS120</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel 4110 8 core processor, 85 W, 2.1 GHz</td>
</tr>
<tr>
<td>Memory Options</td>
<td>64 GB — 2 × 32 GB DIMMs</td>
</tr>
<tr>
<td>Network Connections</td>
<td>Use these ports:</td>
</tr>
<tr>
<td></td>
<td>▪ 1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>▪ (Optional additional) 1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>▪ 1 GbE LOM management port</td>
</tr>
<tr>
<td>Disk Controllers</td>
<td>LSI 3516 RAID controller</td>
</tr>
<tr>
<td>Operating System Storage</td>
<td>2 × 128 GB SATADOM, configured as RAID-1</td>
</tr>
<tr>
<td>Storage</td>
<td>No separate storage is needed</td>
</tr>
</tbody>
</table>
**MongoDB OpsManager Node**

If you are deploying MongoDB OpsManger at your site, the recommended configuration is shown in Table 5. In a production environment, use at least three nodes running OpsManager.

---

**Note** — This configuration does not provide for using the OpsManager node for backup. Contact your Hitachi Vantara and MongoDB account managers for backup storage options.

---

**TABLE 5. MONGO OPSMANAGER NODE CONFIGURATION**

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>Hitachi Advanced Server DS120</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel 4110 8 core processor, 85 W, 2.1 GHz</td>
</tr>
<tr>
<td>Memory Options</td>
<td>64 GB — 2 × 32 GB DIMMs</td>
</tr>
<tr>
<td>Network Connections</td>
<td>Use these ports:</td>
</tr>
<tr>
<td></td>
<td>▪ 1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>▪ (Optional additional) 1 Intel XXV710 dual port 10/25 GbE port</td>
</tr>
<tr>
<td></td>
<td>▪ 1 GbE LOM management port</td>
</tr>
<tr>
<td>Disk Controllers</td>
<td>LSI 3516 RAID controller</td>
</tr>
<tr>
<td>Operating System</td>
<td>2 × 128 GB SATADOM, configured as RAID-1</td>
</tr>
<tr>
<td>Storage</td>
<td>1 × 480 GB SSD</td>
</tr>
</tbody>
</table>

---

**Processing Node**

Use these nodes to run software that interfaces MongoDB. Some examples are the following:

- Web servers
- Pentaho
- Apache Spark
- Business intelligence tools

The different processing nodes have diverse needs and requirements. While they are mentioned in this document, detailed description of each architecture is not presented.
Hardware Management Server

You can include an optional hardware management server in your solution. This server allows access to the out-of-band management network. Table 6 lists the hardware used for this server.

TABLE 6. HARDWARE MANAGEMENT SERVER CONFIGURATION

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chassis</td>
<td>Hitachi Advanced Server DS120</td>
</tr>
<tr>
<td>CPU</td>
<td>Intel 4110 8 core processor, 85 W, 2.1 GHz</td>
</tr>
<tr>
<td>Memory Options</td>
<td>64 GB — 2 × 32 GB DIMMs</td>
</tr>
<tr>
<td>Network Connections</td>
<td>Use these ports:</td>
</tr>
<tr>
<td></td>
<td>• 1 Intel XXV710 dual port 10/25 GbE ports</td>
</tr>
<tr>
<td></td>
<td>• 1 copper NIC</td>
</tr>
<tr>
<td></td>
<td>• 1 GbE LOM management port</td>
</tr>
<tr>
<td>Disk Controllers</td>
<td>Intel RTEs on the mother board</td>
</tr>
<tr>
<td>Operating System</td>
<td></td>
</tr>
<tr>
<td>Storage</td>
<td>2 × 128 GB SATADOM, configured as RAID-1</td>
</tr>
</tbody>
</table>

Network Architecture

This solution uses two networks.

- **Data Network** — Communication between nodes
- **Management Network** — Out of Band management of hardware

Figure 6 on page 17 shows the network configuration. The network supports 25 GbE. The default configuration uses 10 GbE.
The network architecture has these components:

- “Switches” on page 18
- “Data Network” on page 18
- “Management Network” on page 18
Switches

This solution requires the following three types of switches:

- **Spine Data Switches — Cisco Nexus 93180LC-EX**
  
  These spine data switches interconnect leaf switches from different racks.

  Connect the two switches using a peer-keepalive link. This lets both switches act together as a single logical switch. If one switch fails, there still is a path to the racks.

- **Leaf Data Switches — Cisco Nexus 93180YC-E/FX**
  
  These leaf data switches connect all the nodes in a rack together. Then, uplink the leaf switches to the spine data switches.

  Connect the two switches using a peer-keepalive link. This lets both switches act together as a single logical switch. If one switch fails, there still is a path to the hosts.

- **Leaf and Spine Management Switches — Cisco Nexus 3048**
  
  These leaf and spine switches connect the management ports of the hardware to the management server. When there is more than one rack, use a spine switch to connect all the management leaf switches together.

Data Network

Use the data network for communications between the nodes.

This is redundant network. Each node has one connection to each switch. To provide better throughput, configure these connections to use **active-active** network-bonding.

Connect the data network leaf switches to the data network spine switches using a 100 GbE Fibre connection.

Management Network

The management network allows for access to the nodes using the 1 GbE LAN on motherboard (LOM) interface. This network provides out of band monitoring and management of the servers and switches.

Connect leaf switches to the spine switches using a 1 GbE connection.

Operating System

This reference architecture was tested with Red Hat Enterprise Linux 7.3.

- The boot storage is 2 x 128 GB SATADOM drives configured to use disk mirroring.
- NUMA is disabled.
- Transparent huge pages are disabled.
- NICs are configured **active-active**.

Sample Rack Configuration

Install the master nodes and edge nodes in the same rack. Also, you can use this rack for spine switches. You can install worker nodes in the empty units.

Figure 7 on page 19 shows a sample rack with the following:

- Three config servers
- One processing node
- One sharded cluster
  - Scaled out across three shards
  - Three nodes per shard, to provide high availability
- Two data leaf switches
- One data management switch

Figure 7
Figure 8 shows a multiple rack configuration with the following:

- Three config servers
- Three processing nodes, running applications and mongos
- One shared cluster
  - Scaled out across 19 replica sets
  - Three mongo nodes per replica set to provide high availability
- Two leaf data switches per rack
- One leaf management switch per rack
- Two spine data switches
- One spine management switch

Figure 8
Engineering Validation

MongoDB defined the testing criteria and then evaluated this architecture. This environment was evaluated and tested by engineers at Hitachi Vantara in their lab environment. This is certified by MongoDB.