Power Your Digital Transformation with Cisco UCS Integrated Infrastructure for Big Data and Analytics with Couchbase

Deliver a high-performance NoSQL distributed database platform.

As the world migrates to a digital economy, businesses are adopting NoSQL databases as they discover that traditional relational databases can no longer handle the scale and speed at which data is being stored and served.

Highlights

Cisco Unified Computing System foundation

- The Cisco Unified Computing System™ (Cisco UCS®) M5 server platform offers complete integration of computing, networking, and storage resources with unified management while providing high performance, expandable storage, and scalability for big data systems.

- Cisco UCS uses a fabric-centric architecture designed for business acceleration, providing true on-demand infrastructure, with a system that grows gracefully and incrementally.

High performance with linear scalability

- Cisco UCS Integrated Infrastructure for Big Data is a simplified, intelligent infrastructure with high performance and scalability to meet growing business demands.

- Couchbase enables users to define the scalability characteristics of their systems based on their actual technical business requirements.

General-purpose database

- Couchbase Server is an open-source, distributed database engineered for scalability, performance, and availability. It is a general-purpose database that can be deployed as a document database, key-value store, or distributed cache.

- Couchbase also supports a robust collection of adapters for popular big data and analytics products.

Distributed SQL-based querying

- N1QL is a full-featured, SQL-based query language that extends SQL to JavaScript Object Notation (JSON) and enables low-latency queries on distributed data, regardless of scale.

ıllıllı
CISCO

Couchbase

Develop with Agility and Operate at Scale

The business world is undergoing massive change as enterprises are shifting toward a digital economy. At the core of every digital economy business are its web, mobile, and Internet of Things (IoT) applications. They are the primary means through which companies interact with customers today and run more and more of their business. The experiences that companies deliver through these applications largely determine customer satisfaction and retention.

Today's enterprise applications share one or more of the following characteristics. They need to:

- Support large numbers of concurrent users (tens of thousands, or perhaps even millions, of users)
- Deliver highly responsive experiences to a globally distributed base of users
- Be always available, no downtime
- Handle semistructured and unstructured data
- Rapidly adapt to changing requirements with frequent updates and new features

Building and running these applications has created a new set of technology requirements. The enterprise data platform must be more agile and scalable, and its approach to real-time data management must accommodate new levels of speed and data variability.

Cisco UCS® Integrated Infrastructure for Big Data and Analytics with Couchbase provides a low-latency data management platform for large-scale applications with a flexible data model and elastic scaling. Couchbase Server is extremely simple to deploy and manage, and it is built with a strong emphasis on reliability and high availability. It is designed to deliver an excellent experience with robust features and easy administration.

Cisco UCS Integrated Infrastructure for Big Data and Analytics

Organizations today must be sure that the underlying physical infrastructure can be deployed, scaled, and managed in a way that is agile enough to change as workloads and business requirements change. Cisco UCS Integrated Infrastructure for Big Data and Analytics has redefined the potential of the data center with its revolutionary approach to managing computing, network, and storage requirements to successfully address the business needs of IT innovation and acceleration. Cisco UCS Integrated Infrastructure for Big Data and Analytics provides an end-to-end architecture for processing high volumes of structured and unstructured data for both real-time processing and archival purposes.

Cisco UCS 6300 Series Fabric Interconnects

Cisco UCS 6300 Series Fabric Interconnects provide high-bandwidth, low-latency connectivity for servers, with Cisco UCS Manager providing integrated, unified management for all connected devices. The Cisco UCS 6300 Series Fabric Interconnects are a core part of the Cisco Unified Computing System™ (Cisco UCS), providing low-latency, lossless 40 Gigabit Ethernet, Fibre Channel over Ethernet (FCoE), and Fibre Channel functions.

Cisco® Fabric Interconnects offer the full active-active redundancy, performance, and exceptional scalability needed to support the large number of nodes that are typical in clusters serving big data applications. Cisco UCS Manager enables rapid and consistent server configuration using service profiles and automates ongoing system maintenance activities such as firmware updates across the entire cluster as a single operation. Cisco UCS Manager also offers advanced monitoring with options to raise alarms and send notifications about the health of the entire cluster.

Cisco UCS C240 and C220 M5 Rack Servers

Cisco UCS M5 rack servers are dual-socket, 2-rack-unit (2RU) servers offering industry-leading performance and expandability for a wide range of storage and I/O-intensive infrastructure workloads, such as big data and analytics. These servers use the new Intel® Xeon® Processor Scalable Family CPUs with up to 28 cores per socket. The servers support up to 24 double-data-rate 4 (DDR4) dual in-line memory modules (DIMMs) for improved performance and lower power consumption. The DIMM slots are 3D XPoint ready, supporting next-generation nonvolatile memory technology.

Depending on the server type, Cisco UCS rack servers have a range of storage options. The Cisco UCS C240 M5 Rack Server supports up to 24 small-form-factor (SFF) 2.5-inch drives (with support for up to 10 Non-Volatile Memory Express [NVMe] Peripheral Component Interconnect Express [PCIe] solid-state drives [SSDs] on the NVMe-optimized chassis version), or 12 large-form-factor (LFF) 3.5-inch drives plus 2 rear hot-swappable SFF drives with a Cisco 12-Gbps SAS Module RAID Controller. The Cisco UCS C220 M5 Rack Server supports up to 10 SFF 2.5-inch drives (with support for up to 10 NVMe PCIe SSDs on the NVMe-optimized chassis version) Additionally, all the servers have two modular M.2 cards that you can use for boot. A modular LAN-on-motherboard (mLOM) slot supports dual 40 Gigabit Ethernet network connectivity with the Cisco UCS Virtual Interface Card (VIC) 1387.

Couchbase Server

Couchbase Server is a popular database system used to power the growing digital economy. Its core architecture is designed to simplify building modern applications with a flexible data model, a powerful SQL-based query language, and a secure core database platform that provides high availability, scalability, and performance.

Couchbase Server can be deployed in multiple ways:

- Document database: Access, index, and query documents while taking advantage of integrated caching for high-performance data access.
- Key-value store: Meet high-performance requirements for read and write data access while maintaining data durability and availability.
- Distributed cache: Provide scalable, low-latency access to large in-memory data sets, optionally through the memcache API and clients.

Couchbase enables applications to run with consistent high performance while supporting very high rates of throughput. Couchbase can also scale easily and efficiently: from a few nodes to many nodes and from one data center to multiple data centers, with push-button simplicity. A comprehensive, web-based console simplifies management and monitoring by performing complex tasks and providing more than 200 metrics covering everything from cluster health to individual node performance. Figure 1 shows the real-time flow of data with Couchbase.
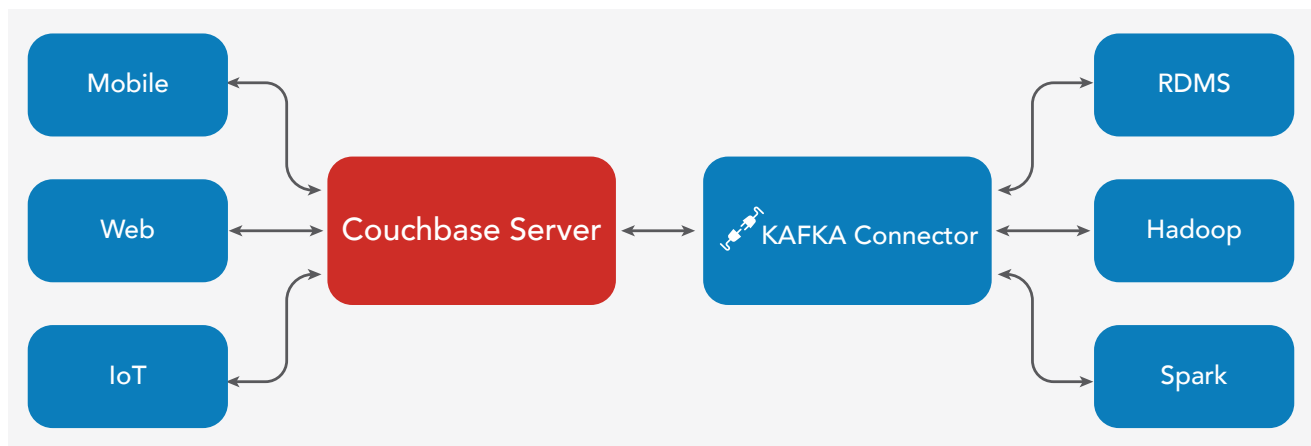

Figure 1. Real-Time Data Flow with Couchbase

Typical use cases for Couchbase include the following:

- User profile management
- Operational dashboards
- Personalization
- 360-degree customer view
- Real-time data processing
- High-availability caching
- Metadata management

Reference Architecture

The Cisco UCS Integrated Infrastructure for Big Data and Analytics for Couchbase includes eight or more C240 M5 (or C220 M5) servers, each with dual Intel Xeon Processor Scalable Family 6132 CPUs (2 x 14 cores at 2.6 GHz), 384 GB of RAM, dual 40-Gbps network connectivity, and 8 (or 16) solid-state disks (SSDs). These servers are connected to Cisco UCS 6332 Fabric Interconnects (Figure 2).

Performance Tests and Results

Note: The tests discussed here were conducted using Cisco UCS C240 M4 Rack Servers. They demonstrate the linear scalability of the solution. The Cisco UCS C240 M5 Rack Servers are expected to provide even higher performance.

Several tests were performed to demonstrate the performance of Couchbase on Cisco UCS Integrated Infrastructure at a large scale. The tests focused on the most critical components for real-time analytics: ingestion and retrieval. Couchbase Server delivered high performance for both as the system scaled out from four nodes to eight nodes. These tests validate the linear scalability and high performance of Couchbase on Cisco UCS Integrated Infrastructure.

Test Configuration

The test configuration consisted of four to eight Cisco UCS C240 M4 Rack Servers, each with two Intel Xeon processors E5-2680 v4, 256 GB of RAM, eight 1.6-TB SATA SSDs with Couchbase Server 4.6.1 Enterprise Edition software, and the Red Hat Enterprise Linux (RHEL) 7.2 operating system.

Initial set of tests were run on a 4-node cluster that was later scaled out to an 8-node cluster. Multiple tests were performed using 50:50 and 95:5 read-write ratios:

- With 100% data present in memory
- With 50% data present in memory

The Couchbase bucket for storing data was created using the default setting for data compaction and a replica setting of 1, which means that for each data record, Couchbase Server automatically created a second copy to support high availability.

All tests were performed using cbc-pillowfight, a load-generation tool created by Couchbase. (The source code for this tool is publicly available at GitHub.)
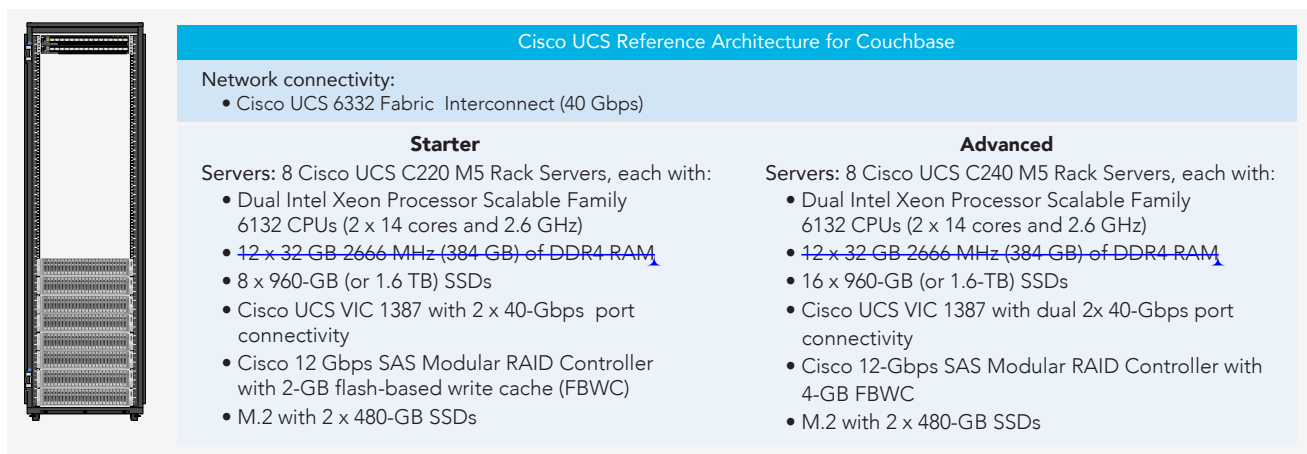
Cisco UCS Reference Architecture for Couchbase	
Network connectivity: • Cisco UCS 6332 Fabric Interconnect (40 Gbps)	
Starter Servers: 8 Cisco UCS C220 M5 Rack Servers, each with: • Dual Intel Xeon Processor Scalable Family 6132 CPUs (2 x 14 cores and 2.6 GHz) • 12 x 32 GB 2666 MHz (384 GB) of DDR4 RAM • 8 x 960-GB (or 1.6 TB) SSDs • Cisco UCS VIC 1387 with 2 x 40-Gbps port connectivity • Cisco 12 Gbps SAS Modular RAID Controller with 2-GB flash-based write cache (FBWC) • M.2 with 2 x 480-GB SSDs	**Advanced** Servers: 8 Cisco UCS C240 M5 Rack Servers, each with: • Dual Intel Xeon Processor Scalable Family 6132 CPUs (2 x 14 cores and 2.6 GHz) • 12 x 32 GB 2666 MHz (384 GB) of DDR4 RAM • 16 x 960-GB (or 1.6-TB) SSDs • Cisco UCS VIC 1387 with dual 2x 40-Gbps port connectivity • Cisco 12-Gbps SAS Modular RAID Controller with 4-GB FBWC • M.2 with 2 x 480-GB SSDs

Figure 2. Cisco UCS Integrated Infrastructure Reference Architecture for Couchbase

Performance results

Figure 3 shows the performance results for the benchmark with 100 percent of data in memory.

Operations per second with 100% data in memory

Throughput in millions

- 4 Nodes
- 8 Nodes

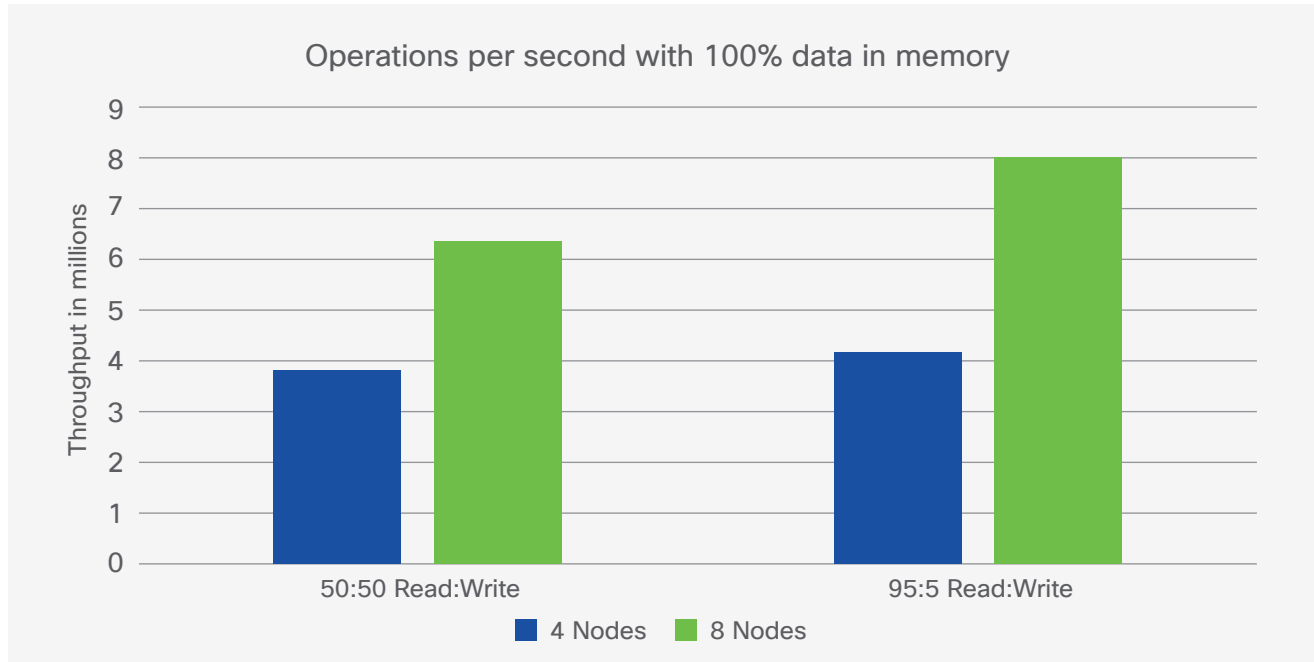50:50 Read:Write   95:5 Read:Write

Figure 3. Results with 100 Percent of Data in Memory

Figure 4 shows the performance results for the benchmark with 50 percent of the data in memory and 50 percent of the data in the SSDs.
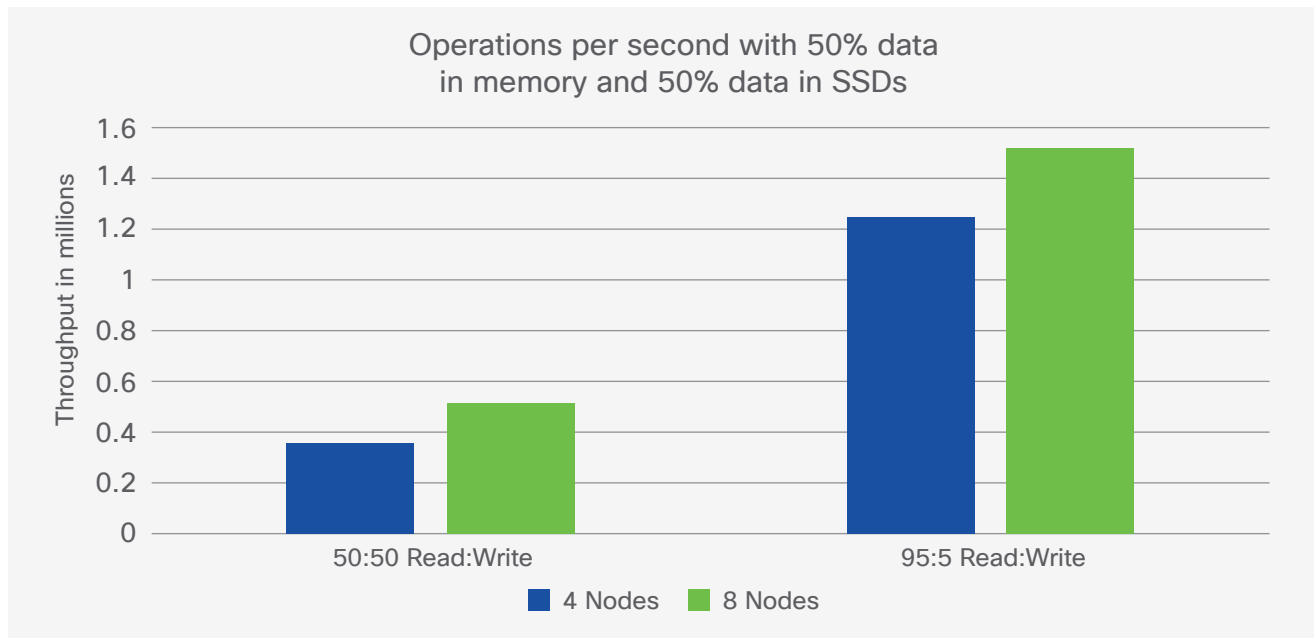

Operations per second with 50% data in memory and 50% data in SSDs

Throughput in millions

- 4 Nodes
- 8 Nodes

50:50 Read:Write   95:5 Read:Write

Figure 4. Results with 50 Percent of Data in Memory and 50 Percent in SSDs

Conclusion

The fifth generation of Cisco UCS Integrated Infrastructure for Big Data and Analytics is the next-generation platform, with new processors, faster memory, and more storage options. It is designed, tested, and validated for enterprises to lower the cost of ownership and to scale from small to very large environments as applications demand. This easily deployable server infrastructure offers linear scalability with high throughput to meet the requirements of enterprise users.

For more information

- For more information about Cisco UCS big data solutions, see https://www.cisco.com/go/bigdata.
- For more information about Couchbase, see http://www.couchbase.com/.