



Cisco

Industry

Platform

Website

www.cisco.com

Company Overview

Cisco is the worldwide IT leader that enables businesses to seize the opportunities of tomorrow by proving amazing things happen when you connect the previously unconnected. Cisco brings together people, process, data, and things to transform how organizations meet next-generation demands.

About Cisco Unified Computing System (UCS)

Cisco UCS Integrated Infrastructure for big data architecture provides a secure and scalable infrastructure. Cisco leads the market by integrating technology with innovative and disruptive software companies. While the industry is bringing data to computing, Cisco is bringing the computing and analytics to the data to take advantage of the valuable insight that it reveals.

Cloudera and Cisco Unified Computing System (UCS) Implementing Big Data Analytics for Business Success

Big data is now a significant element in many industries and gaining meaningful traction across multiple industries. Organizations heavily using big data technologies include healthcare, education, and energy along with financial, utility, advertising, retail, public sector, and manufacturing. But how do you put all of this information to work for you?

Apache Hadoop, a technology developed to handle large volumes of data of any format, is much more efficient than traditional enterprise data warehouses. Because Hadoop is designed to run on industry standard hardware with infinite scalability, the cost savings are significant. Overall savings increase as organizations' data volumes grow.

However, while this powerful technology can help transform large volumes of data into actionable information, many organizations are also struggling to deploy effective and reliable Hadoop infrastructure that performs and scales and is appropriate for mission-critical applications in the enterprise. Many of the challenges arise from the friction between the rapid pace of change inherent in open-source software and the need for enterprise-class performance, security, reliability, and support.

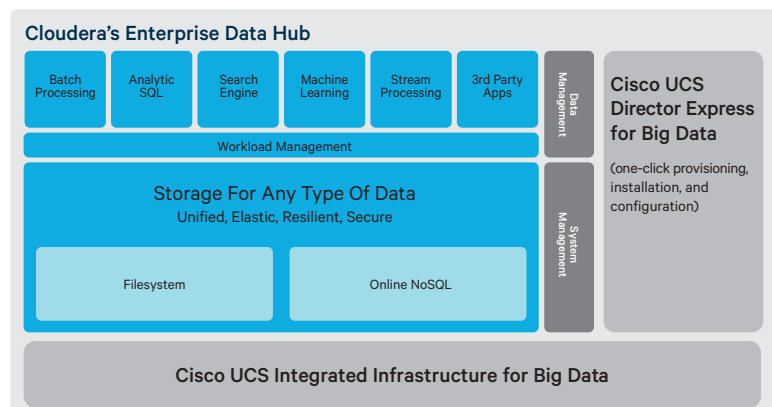
Cloudera and Cisco offer an industry-leading solution for enterprise Hadoop deployments.

Cisco UCS Integrated Infrastructure for Big Data and Cloudera Enterprise

Cloudera is the leading provider of enterprise-ready big data software and services. Cloudera Enterprise, a sophisticated administrative and management tool, is a native end-to-end data governance solution that provides comprehensive security tool sets and technical support.

Cisco UCS® Integrated Infrastructure for big data includes computing, storage, connectivity, and unified management capabilities to help companies manage the dramatic increase in data they must cope with today. It is built on Cisco UCS infrastructure using Cisco UCS 6200 Series Fabric Interconnects, Cisco Nexus® 2200 platform fabric extenders, and Cisco UCS C-Series Rack Servers. Installed in pairs, the fabric interconnects offer redundant, active-active connectivity and embedded management using Cisco UCS Manager.

Together, Cisco and Cloudera provide organizations with an enterprise-ready data management platform, as well as management integration with an enterprise application ecosystem. They seamlessly combine to provide a uniquely capable, industry-leading architectural platform for Hadoop-based applications. (Figure 1)



Solution Highlights

Ease of deployment

- Cisco UCS® Manager automates deployment and scaling processes, helping reduce the risk of configuration errors that can cause downtime.

Scalability for big data workloads

- Cisco UCS Integrated Infrastructure for big data solutions offer linear scalability and help simplify essential operations for single-rack and multiple-rack deployments.

Comprehensive integrated infrastructure

- Cisco UCS Integrated Infrastructure for big data solutions include computing, storage, connectivity, and unified management.

Complete security and data protection

- Cisco UCS Integrated Infrastructure for big data with Cloudera offers complete data protection with built-in support for Kerberos authentication, enterprise-class access controls, and data encryption.

Simplified management

- Faster and easier big data infrastructure deployment: Cisco UCS Director Express for big data offers one-click provisioning, installation, and configuration.
- Cloudera Manager is a holistic interface that provides end-to-end system management and granular visibility into and control over every part of an enterprise data hub.

Scale to thousands of servers with Cisco® Application Centric Infrastructure (Cisco ACI™)

- With Cisco ACI, you can easily scale the cluster to thousands of nodes. Cisco ACI treats the network as a single entity rather than a collection of switches.

The Cisco Unified Computing System™ (Cisco UCS) solution for Cloudera is based on Cisco UCS Integrated Infrastructure for big data, a highly scalable architecture that includes computing, storage, connectivity, and unified management capabilities and is designed to meet a variety of scale-out application demands.

The current version of the Cisco UCS Integrated Infrastructure for big data offers the configurations summarized in Table 1. The configuration used depends on the computing and storage requirements of Hadoop.

Performance Optimized	Capacity Optimized
Connectivity: <ul style="list-style-type: none"> • 2 Cisco UCS 6296UP 96-Port Fabric Interconnects 	Connectivity: <ul style="list-style-type: none"> • 2 Cisco UCS 6296UP 96-Port Fabric Interconnects
Connectivity: <ul style="list-style-type: none"> • Up to 80 servers per domain • Up to 160 servers per domain with Cisco Nexus 2232PP 10GE Fabric Extender 	Connectivity: <ul style="list-style-type: none"> • Up to 80 servers per domain • Up to 160 servers per domain with Cisco Nexus 2232PP 10GE Fabric Extender
16 Cisco UCS C240 M4 Rack Servers (SFF), each with: <ul style="list-style-type: none"> • 2 Intel Xeon processor E5-2680 v3 CPUs • 256 GB of memory • Cisco 12-Gbps SAS Modular RAID Controller with 2 GB of flash-based write cache (FBWC) • 24 x 1.2-TB 10,000-rpm SFF SAS drives (460 TB total) • 2 x 120-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS VIC 1227 (with 2 x 10 Gigabit Ethernet SFP+ ports) 	16 Cisco UCS C240 M4 Rack Servers (LFF), each with: <ul style="list-style-type: none"> • 2 Intel Xeon processor E5-2620 v3 CPUs • 128 GB of memory • Cisco 12-Gbps SAS Modular RAID Controller with 2 GB of FBWC • 12 x 4-TB 7,200-rpm LFF SAS drives (768 TB total) • 2 x 120-GB 6-Gbps 2.5-inch Enterprise Value SATA SSDs for boot • Cisco UCS VIC 1227 (with 2 x 10 Gigabit Ethernet SFP+ ports)
Scale to tens of thousands of servers with Cisco ACI	

Table 1: Cisco UCS Integrated Infrastructure for big data with Cloudera

Delivering Fast, Simple, and Secure Big Data Platform Fast Performance

From analytics to data science and everything in between, only Cloudera delivers the performance you need to unlock the potential of unlimited data through a few data access and management components:

- Apache Spark provides in-memory data processing for developers and data scientists. Its easy development, flexibility, and faster performance have caused Spark to be the most popular Apache project, and the successor to MapReduce as the standard execution engine for Hadoop. Cloudera has the most experience with Spark as part of Hadoop for ensured success.
- Apache Impala (incubating) provides massive parallel processing (MPP) analytic SQL that opens up interactive BI on Hadoop for business analysts. Impala not only exceeds the speed and concurrency requirements necessary for building an analytic database, it is natively integrated across Hadoop and with the leading BI tools for a complete low-cost platform.
- Cloudera Search (powered by Apache Solr) provides full-text search that opens up data to the entire business. With its real-time exploration capabilities and flexible indexing, multiple users can discover new insights—faster. As the only natively integrated search solution, it delivers streamlined value as part of “find then do” work flows.
- Cloudera Navigator Optimizer gives you the insights and risk assessments you need to build out a comprehensive strategy for Hadoop success. Simply upload your existing SQL workloads to get started, and Navigator Optimizer will identify relative risks and development costs for off-loading these to Hadoop based on compatibility and complexity. Navigator Optimizer is the plan you and your business can trust to always use Hadoop with confidence.

Simple Management

Cloudera Manager is the industry's first and most sophisticated management application for Hadoop. Cloudera Manager sets the standard for enterprise deployment by delivering granular visibility into and control over every part of the data hub—empowering operators to improve performance, enhance quality of service, increase compliance, and reduce administrative costs.

Cloudera Manager is designed to make administration of your enterprise data hub simple and straightforward, at any scale. With Cloudera Manager, you can easily deploy and centrally operate the complete big data stack. The application automates the installation process, reducing deployment time from weeks to minutes; gives you a cluster-wide, real-time view of nodes and services running; provides a single, central console to enact configuration changes across your cluster; and incorporates a full range of reporting and diagnostic tools to help you optimize performance and utilization.

Robust Security

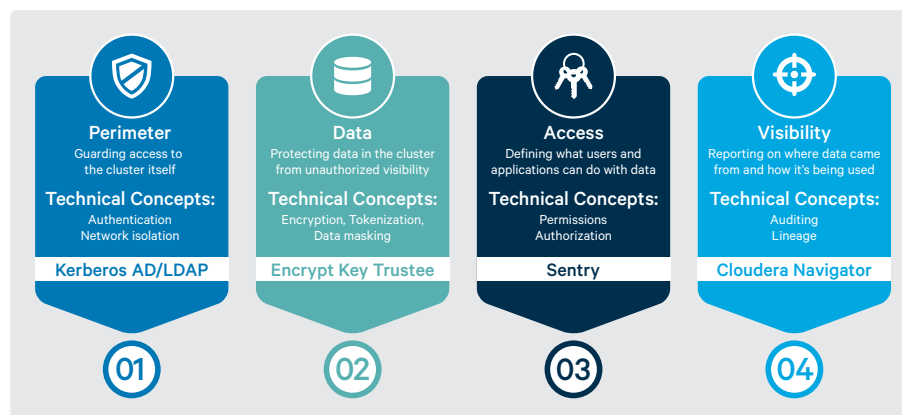
Cloudera Navigator is the only native end-to-end governance solution for Hadoop-based systems. Through a single user interface, it provides visibility for administrators, data managers, data scientists, and analysts to secure, govern, and explore the large amounts of diverse data that land in Hadoop. Cloudera Navigator is part of Cloudera Enterprise's comprehensive data security and governance offering and is a key part to meeting compliance and regulatory requirements.

Features include:

- **Guard the perimeter:** Cloudera Manager makes it easy to secure your cluster using industry-standard Kerberos, open Lightweight Directory Access Protocol (LDAP) and Microsoft Active Directory (AD), and Security Assertion Markup Language (SAML).
- **Protect data at rest and data in motion:** Cloudera can protect both data at rest and data in motion through encryption and powerful key management capabilities that are integrated into Cloudera Navigator.
- **Control access with Sentry:** Sentry integrates with the Hadoop platform and allows storage of sensitive data while meeting regulatory requirements.
- **Gain visibility with Cloudera Navigator:** Cloudera Navigator is the only native end-to-end governance solution for Hadoop. It provides full visibility into the source and use of the data to verify authenticity and easily comply with regulatory requirements.

Benefits of Cloudera

- **Powerful** – Store, process, and analyze all your data to drive competitive advantage
- **Efficient** – Hadoop unifies compute and data to improve operational efficiency
- **Open** – 100% open source: CDH is the world's most popular open source distribution powered by Apache Hadoop
- **Simple** – Easy to deploy and operate with centralized administration
- **Compatible** – Leverage your existing investments for rapid adoption and lower TCO
- **Economical** – Rethink the economics of data management with an open source platform on industry standard hardware—up to 90% more cost-effective than traditional solutions
- **Enterprise Ready** – Equipped with critical capabilities to support mission-critical operations





About Cisco

Cisco is the worldwide IT leader that enables businesses to seize the opportunities of tomorrow by proving amazing things happen when you connect the previously unconnected. Cisco brings together people, process, data, and things to transform how organizations meet next-generation demands.

About Cloudera

Cloudera delivers the modern platform for data management and analytics. The world's leading organizations trust Cloudera to help solve their most challenging business problems with Cloudera Enterprise, the fastest, easiest, and most secure data platform built on Apache Hadoop. Our customers can efficiently capture, store, process, and analyze vast amounts of data, empowering them to use advanced analytics to drive business decisions quickly, flexibly, and at lower cost than has been possible before. To ensure our customers are successful, we offer comprehensive support, training, and professional services. Learn more at cloudera.com.

cloudera.com

1-888-789-1488 or 1-650-362-0488

Cloudera, Inc. 1001 Page Mill Road, Palo Alto, CA 94304, USA

© 2016 Cloudera, Inc. All rights reserved. Cloudera and the Cloudera logo are trademarks or registered trademarks of Cloudera Inc. in the USA and other countries. All other trademarks are the property of their respective companies. Information is subject to change without notice.