DELL EMC QUICKSTART FOR CLOUDERA HADOOP

Accelerate Proofs of Concept

Dell EMC Hadoop expertise

In 2009, Dell started custom designing and building the first Hadoop server platforms for the largest Web 2.0 companies through the Dell Data Center Solutions Group (DCS). DCS was created in 2008 as a startup within Dell. With a history in data solutions that launched in 2008, when Dell started discussions with the leading innovators in big data, Dell turned that experience into big data expertise that enables Dell EMC to make Hadoop as simple as possible for organizations. Dell was the first server provider to build server hardware optimized to run demanding Hadoop workloads.

The Dell DCS FS12 server was the first purpose-built platform for Hadoop. As Hadoop workloads evolved, the Dell PowerEdge™ C2100 server was created. The PowerEdge C2100 platform would eventually lead to the development of the Dell PowerEdge™ R720xd server. In 2014, the next-generation Hadoop platform was introduced — the Dell PowerEdge™ R730xd server — to provide an ideal, flexible building block for Hadoop workloads. Now we add the PowerEdge™ FX2, a 2U hybrid rack-based computing platform that combines the density and efficiencies of blades with the simplicity and cost benefits of rack-based systems.

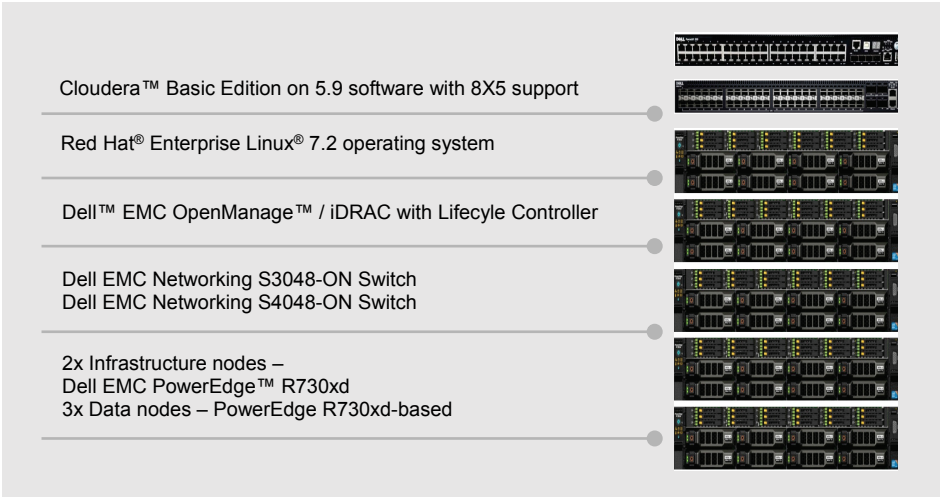
A QUICK-TO-DEPLOY, ALL-IN-ONE OFFERING WITH AN AGGRESSIVE PRICE POINT

At Dell EMC we recognize the importance of data as the new currency and as a key competitive differentiator. Data is being created and consumed at rates never before seen, and this is not unique to a single industry. Rapid data growth is affecting all verticals, all segments and all geographies. With this data explosion, organizations are recognizing the need for more than traditional, structured systems to take control of their data.

The Dell EMC QuickStart for Cloudera Hadoop offers an easy entry point for organizations to begin managing and analyzing data. An all-in-one system designed to reduce the complexity of deploying, configuring and managing Hadoop systems, Dell EMC QuickStart for Cloudera Hadoop includes the hardware, software and services needed to deliver a Hadoop cluster that will start your organization on a proof of concept to begin working with big data.

If your organization is on this path, Dell EMC, Cloudera and Intel are your ideal technology partners. Dell has worked with Hadoop and Cloudera, the leading provider of Hadoop-based software and services, and Intel, the world leader in computing innovation, to deliver end-to-end Hadoop solutions for more than six years. This combination of technology leaders delivers the benefits of proven architectures for Hadoop deployments, processors purpose-built for big data, powerful analytics engines, and technologies to enable enterprise-class security in Hadoop environments.

Dell EMC, in partnership with Cloudera and Intel®, has helped many organizations begin their big data journeys and solve the skills gap by providing expert guidance and knowledge to streamline the architecture, design, planning and configuration of Hadoop environments.



DELL EMC QUICKSTART FOR CLOUDERA HADOOP

Dell EMC QuickStart for Cloudera Hadoop enables your organization to quickly engage in Hadoop testing, development and proof-of-concept work. Through the combination of Dell™ PowerEdge™ servers, Cloudera Enterprise Basic Edition and Dell EMC Networking and Services, your organization can quickly deploy Hadoop and enable your development and application teams to test business processes, data analysis methodologies and operational needs against a fully functioning Hadoop cluster.

Every QuickStart comes with deployment and software configuration services. In addition, our one-week Cloudera Accelerator includes training, knowledge transfer, interactive architectural design workshops, engagement, we provide a roadmap and best-practice deliverable with recommendations on how to achieve your goals with your QuickStart bundle.

REDUCE THE RISK

According to Gartner, through 2018, 70 percent of Hadoop deployments will fail to meet cost savings and revenue-generation objectives due to skills and integration challenges.[1] The high failure rate has left many organizations wary of adopting Hadoop, yet the demands of the business will continue to require big data technologies. With this knowledge, Dell EMC recognized the need to assist and enable organizations to lower the risk using a tested, validated and certified plan for implementing Hadoop.

In addition to the Dell EMC QuickStart for Cloudera Hadoop, Dell EMC has delivered 19 Hadoop DAS reference architectures and engineered solutions over the past six years and you can leverage that experience and expertise to fill the skills gap and build an architecture that meets the needs of your business, all while reducing the risks that come with technology projects.

DELL EMC POWEREDGE

Maximize server-based storage flexibility and performance with the Intel-based Dell EMC PowerEdge R730xd server, part of the 13th generation of Dell EMC PowerEdge servers. The PowerEdge R730xd server, based on latest Intel® Xeon® processors, offers an optimal balance of storage utilization, performance and cost with an optional in-server hybrid storage configuration that can support tiering and capacity for up to 28 drives in a 2S/2U system, including up to 18 1.8-inch SATA SSDs.

DELL EMC NETWORKING S-SERIES 10GBE SWITCHES

Deploy modern workloads and applications designed for the open networking era with an optimized data center top-of-rack (ToR) networking solution.

Among other benefit, this switch series:

- Delivers low latency, superb performance and high density with hardware and software redundancy
- Offers Active Fabric designs using S- or Z-Series core switches to create a two-tier, 1/10/40 GbE data center network architecture
- Provides an ideal solution for applications in high-performance data center and computing environments

DELL EMC DATA LAKE

Another option for beginning a big data journey is a Dell EMC Data Lake. At Dell EMC, we continue to keep customer use cases first as we offer robust Data Lake opportunities using Cloudera Enterprise together with Dell EMC scale-out NAS Isilon storage running the OneFS operating system. We provide a fully tested and complete solution encompassing enterprise-grade Apache Hadoop and related projects to form a Data Lake for maximizing the value of data.

The combination of Cloudera and Isilon shared storage helps your organization speed time to insights, improve storage utilization, eliminate islands or silos of storage, and lower storage management costs of migration, security and protection.

DELL SERVICES

After your exploration of the technologies, Dell EMC Services makes getting started easy. Options include custom solution design, hardware and software deployment, ongoing support and training. With Dell EMC, you have the assurance that your Cloudera solution is backed by expert hardware and software support that can be tailored to your specific needs.

CLOUDERA BASIC

The right technology is key for turning your data into real business value. Powered by Apache Hadoop, Cloudera Enterprise is a fast, easy and secure modern data platform. From analytics to data science, anyone can now get results from any data and across any environment — all within a single, scalable platform.

When you make Cloudera Enterprise the center of your business, you open up limitless possibilities with your data. Whether you're powering data engineering and data science workloads, building an operational or analytic database, or looking to bring them all together in an enterprise data hub, Cloudera has the right platform to fit your needs.

[1] Gartner. "Market Guide for Hadoop Distributions." January 6, 2015.

Make your journey with Dell EMC, Cloudera and Intel

Around the world, organizations large and small are seeking ways to uncover and act on opportunities hidden in data, regardless of the type of data and where it's located, inside or outside the organization.

Dell EMC enables your organization to explore opportunities to begin a big data journey using a bundled open source Hadoop software platform via Dell EMC QuickStart for Cloudera Hadoop or the Dell EMC Data Lake. Either gives your organization the ability to begin your journey to store and analyze data more affordably than ever before. With its power and flexibility, Hadoop offers a natural complement to existing data warehousing infrastructure.

Together, Dell EMC, Cloudera and Intel are ideally positioned to deliver end-to-end Hadoop solutions that enable your organization to manage and mine growing amounts of data to gain valuable insights and competitive advantages.

	Cloudera Express	**Cloudera Enterprise** Basic Edition Data Engineering Operational Analytic Enterprise Data Edition Database Edition Database Edition Hub Edition				
Term	Unlimited/ Free	Annual Subscription (per node) / Elastic Cloud Pricing Also Available				
Open Source Apache Hadoop Distribution						
CDH (100% open source data platform, including Apache Hadoop)	•	•	•	•	•	•
Automated Cluster Management (Cloudera Manager)						
Core Features: Multi-Cluster Deployment & management; Service & Configuration Management (including HA, cluster templates); Service, Host, and Job Monitoring (including health tests, health history, & charting); Kerberos (including Active Directory); Diagnostic Tools & Alerting; & a Comprehensive API	•	•	•	•	•	•
Advanced Features: Operational Reporting; Multi-Tenant Quota Management; Cluster Utilization Reporting; Configuration History & Rollbacks; Rolling Updates & Service Restarts; SNMP Support; Support Integration (including scheduled diagnostics & proactive maintenance); External Authentication (LDAP/SAML); Fine-Grained User Roles; & Automated Backup & Disaster Recovery		•	•	•	•	•
Hybrid Deployment & Management						
Cloudera Director: Flexible deployment across cloud environments; On-demand cluster creation/termination; Elastic cluster sizing; Cluster templates & cloning; Kerberos authentication & HA workflows; Rollback support; Multi-environment cluster management & monitoring; Comprehensive API & clients; & Customization	•	•	•	•	•	•
Components Covered by Cloudera Support						
Basic Hadoop (HDFS, YARN, MapReduce, Hive, Pig, HUE, Sentry, Flume, Sqoop, Kafka + Cloudera Manager, Cloudera Director)		•	•	•	•	•
Flexible data processing (Apache Spark including Spark Streaming, MLlib, and Spark SQL)			•	•	• (Hive-on-Spark only)	•
Analytic SQL (Apache Impala (incubating))					•	•
Cloudera Search (Apache Solr)			•	•		•
Online NoSQL (Apache HBase)				•		•
Active Data Optimization (Cloudera Navigator Optimizer)					•	•

Governance & Data Management (Cloudera Navigator including auditing, lineage, discovery, & policy lifecycle management)			•	•	•	•
Encryption & Key Management (Cloudera Navigator Encrypt & Key Trustee)			•	•	•	•
Support Features						
Only with Cloudera: Dedicated global support team, proactive technical guidance, predictive issue analysis, comprehensive knowledge base, production solution guides, open source community advocacy		•	•	•	•	•
Commercial License Warranty		•	•	•	•	•
Indemnification: Get protection from litigation stemming from use of open source technology			•	•	•	•
8x5 or 24x7 Expert Support: Get direct access to Cloudera's dedicated team of experts to help you resolve issues quickly and optimize your environment with the latest best practices, straight from the source		•	•	•	•	•
Premium Support: 15-minute time-to-first response for critical issues available as additional purchase option			•	•	•	•

Cloudera CDH 5.9	
Accumulo	1.7.2
Impala	2.7
Kafka	2.0
Navigator	2.8
Sentry	Cloudera Manager 5.X & Higher
Spark	1.6

To learn more, visit: Dell.com/Hadoop | Dell.com/BigData | EMC.com/BigData

For questions, please contact us at: Hadoop@Dell.com

November 2016 | Rev 1.0