

Dell EMC Ready Bundles for Hadoop

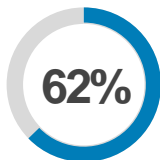
Leverage Hadoop solutions designed to address data analytics requirements, reduce costs and optimize performance

Customer Presentation



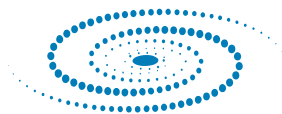
Data analytics are the core of digital transformation...

Newer, smarter competitors are here



62% of companies have seen new competitors enter the market as a result of digital technologies¹

And the **sea of insightful data** is immense



The “digital universe” is already immense – and is doubling in size every two years.²

Now is the time to **ACT**



Data analytics and data processing are amongst the top four technology priorities in 2017¹

...Yet tantalizing benefits of data analytics are elusive



Struggle to begin

Many organizations struggle to begin their data analytics journey or to make Hadoop projects successful once they've begun

Uncertainty and disruption

Digital transformation is causing churn, uncertainty and disruption for businesses as pressure increases from all directions

Lack of expertise

They are impeded by a lack of Hadoop expertise and spend resources on front-end work before getting to results

What is your business goal?

Apache® Hadoop® is a foundational component of the big data and analytics solution stack that enables analytic workloads to run on the same data concurrently at scale — on industry-standard hardware

Reduce operational TCO/Improve ROI

- Offload extract, transform, load (ETL) workloads
- Enable better data exploration and analytics

Secure your enterprise

- Prevent security breaches and threats
- Detect operational anomalies

Manage data storage and archiving

- Lower data storage costs while keeping access
- Streamline inquiry processes

MARKETING

Goal: Anticipating customer needs

Hadoop can help with:

- Customer “360 degree” insight
- New product/service launch

FINANCE

Goal: Reducing risk and detecting fraud

Hadoop can help with:

- Credit scoring
- Sarbanes-Oxley Act (SOX) compliance

HEALTHCARE

Goal: Improving patient care and reducing costs

Hadoop can help with:

- Patient safety
- Claims management

PHARMACEUTICAL

Goal: Ensuring regulatory compliance and validation

Hadoop can help with:

- Biomedical analytics
- Stability and shelf life

MANUFACTURING

Goal: Achieving continuous process improvement

Hadoop can help with:

- Product quality
- Demand forecasting

What if you could...



Leverage an optimized solution to deploy and manage Hadoop in a production environment?



Reduce costs by using optimized, industry-standard servers and storage?



Get Outstanding performance with solutions that have been engineered to work specifically with Hadoop?

Dell EMC Ready Bundles for Cloudera Hadoop

A fast, easy and secure modern data platform.

cloudera

Delivering the core elements of Apache® Hadoop®, including scalable storage and distributed computing, within a solution based on Cloudera® Enterprise software, Dell EMC hardware and services

Benefits

- ① **Fast for business** — Cloudera delivers the performance you need to unlock the potential of unlimited data.
- ② **Easy to manage** — Focus on results, not fighting fires. Keep mission-critical applications up and running — especially at scale.
- ③ **Secure without compromise** — Meet stringent data security and compliance needs without sacrificing business agility.

Dell EMC Ready Bundles for Hortonworks Hadoop

Open and stable with a growing ecosystem



Makes business intelligence on Hadoop a reality, combining a fast in-memory SQL engine to create data marts with an online analytical processing (OLAP) cube engine that enables huge data set queries in seconds

Benefits

- ① **Open** — Hortonworks is committed to a 100% open approach to software development that spurs innovation.
- ② **Interoperable** — 100% open source architecture enables Hortonworks to interoperate with a broad range of business intelligence and other applications.
- ③ **Enterprise ready** — Hortonworks is built for enterprises, providing consistent operations with centralized management and monitoring of clusters through a single pane of glass.

Dell EMC Ready Bundle for Hadoop configurations

Configurations of Dell EMC Ready Bundles for Hadoop include:



- **Dell EMC Hadoop solution on PowerEdge R740xd or R640 Server**

Active archive/customer 360-degree

- **Dell EMC Hadoop solution on PowerEdge FC630 Server with Isilon option**

Active archive/customer 360-degree

- **Dell EMC ETL offload Hadoop solution on PowerEdge R740xd or R640 Server**

ETL offload

- **Dell EMC ETL offload Hadoop solution on PowerEdge FC630 Server with Isilon option**

ETL offload

- **Dell EMC QuickStart for Hadoop on PowerEdge R740xd Server**

Testing and proofs of concept (POCs)

Shared storage Hadoop vs distributed storage Hadoop

Hadoop supports alternative storage deployment models accounting for varying budget, performance, capacity and density requirements.



SHARED STORAGE HADOOP:

The file system data for Hadoop is stored on an Isilon NAS cluster

- Single copy of data for IT workloads and analytics
- Reduced datacenter footprint (storage density)
- Enterprise file management: data protection, security, storage tiering, etc.
- Independent scaling of storage and compute

DISTRIBUTED STORAGE HADOOP:

Data is spread across the local storage of the Hadoop nodes themselves

- Massive (100s+ of petabytes)
- Ability to use commodity servers
- Linear scaling
- Flexible replica model

Dell EMC Services customized for your Hadoop needs

Dell EMC Big Data Vision Workshop

focuses on business leaders to identify and agree on the analytics use case and path forward for using Hadoop

Dell EMC Hadoop Consulting

helps Hadoop customers achieve business value from data analytics

Dell EMC Jumpstart Services

offers best practice guidance, hands-on labs, roadmap planning and knowledge transfer of Hadoop

Dell EMC Cloudera Hadoop Health Check Service

makes recommendations for tools, testing and operational practices

Big Data Technology Assessment Services

assesses Isilon customers with Hadoop installations on the current state of their Hadoop environment

Big Data Technology Implementation Services

supports Isilon customers with Hadoop installations of one to three clusters and up to hundreds of nodes.

Dell EMC Hadoop on Isilon Implementation Starter Kit

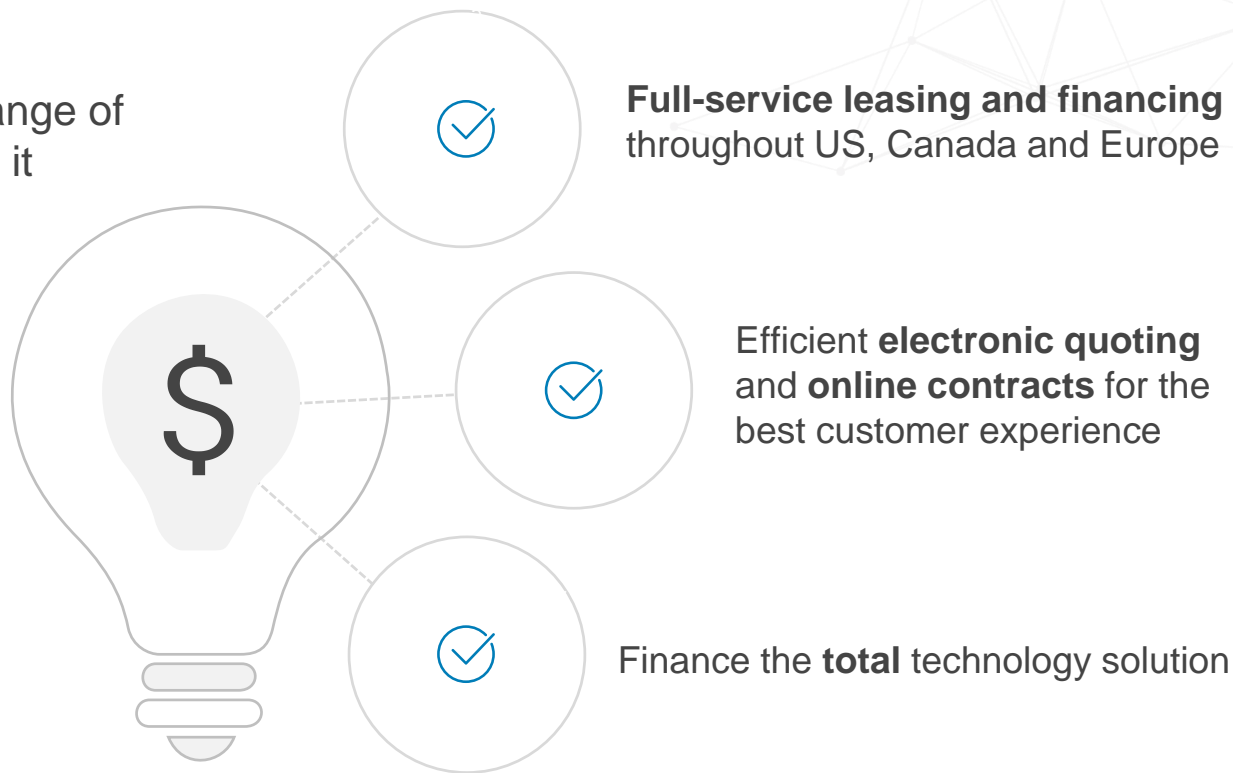
installs proofs-of-concept, or tests Hadoop clusters using best practices to leverage Isilon for HDFS storage

Dell EMC ProSupport

provides unlimited access to 24x7 chat, email and phone support services with how-to assistance and disaster recovery

Dell EMC Financial Services support the solution

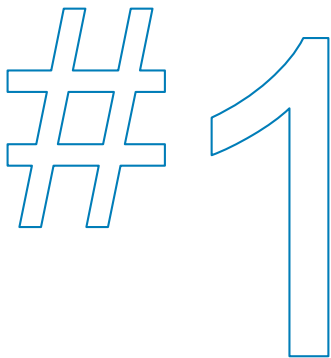
Dell EMC offers a wide range of payment options to make it easier than ever to meet financing needs.



Why Dell EMC for Hadoop?



The combination of Dell and EMC brings together two industry-leading companies with strong reputations for value and innovation. Dell EMC holds leadership positions in some of the biggest and largest growth categories in the IT infrastructure business, and that means you can confidently source all your IT needs from one provider — Dell EMC



- converged infrastructure¹
- in traditional and all-flash storage²
- virtualized data center infrastructure³
- cloud IT infrastructure⁴
- server virtualization and cloud systems management software (VMware)⁵
- in data protection⁶
- in software-defined storage⁷

1 IDC WW Quarterly Converged Systems Tracker, June 2016, Vendor Revenue — EMC FY 2015; 2 IDC WW Quarterly Enterprise Storage Systems Tracker, June 2016, Vendor Revenue — EMC CY 2015; 3 Dell EMC Annual Report, 2015; 4 IDC WW Quarterly Cloud IT Infrastructure Tracker, Q1 June 2016, Vendor Revenue — EMC FY 2015; 5 IDC WW Virtual Machine and Cloud System Market Shares 2015, July 2016; 6 Dell EMC Pulse, Gartner Recognizes EMC as a Leader in the 2016 Data Center Backup and Recovery Software Magic Quadrant, June 2016; 7 IDC white paper, "Software Defined Storage: A Pervasive Approach to IT Transformation Driven by the 3rd Platform," November 2015

Measuring what's in your wallet with Hadoop



Business need

Turning 2.2B global cards
160M transactions/hour, 52B
transactions / year into
intelligence with 1.9M rules
applied to those transactions.

Solution

MasterCard worked with Dell EMC and Cloudera to build a secure, PCI-certified Hadoop cluster. MasterCard's Enterprise Data Hub fully conforms to the PCI-DSS V 2.0 security standards so it can host PCI datasets and integrate with other systems.

Benefits

- Understand how, when and where customers buy leading to predictive analytics
- Increased security and fraud protection via analyzing customer purchasing patterns, affinities and rhythms using machine learning
- Expand business with anonymized data re: share of wallet compared to competition, average ticket and frequency to identify marketing opportunities and measure return on investment

“Data privacy and protection is a top priority for MasterCard. As we maximize the most advanced technologies from partners and vendors, they must meet the rigorous security standards we’ve set.”

Gary VonderHaar,
Chief Technology Officer, Architecture, MasterCard



How Dell EMC Does IT: Data analytics



Business need

The Dell EMC Marketing Analytics team needed to give analysts easier, faster access to customer data to determine the best ways to interact with customers.

Solution

- Dell EMC | Cloudera® ETL Offload Reference Architecture
- Dell EMC PowerEdge R730 and R730xd rack servers powered by Intel® Xeon® processors
- Cloudera Hadoop

Benefits

- Uses predictive modeling helping sales professionals target the right customer at the right time
- Enables data analysis that is 4X faster than before
- Reduces time-to-insight from several days to less than two hours
- Realizes up to \$4 million in OPEX savings

“We can answer business questions faster with the Dell EMC | Cloudera solution. If I get an ad hoc query request from the business, I can get an answer for them within a few hours to a day. Before, it took me weeks to try to synthesize all the data. That makes our stakeholders very happy.”

Jeff Weidner, Director,
Information Management, Dell



Siemens PLM Software, improves customer response times with big data analytics



Business need

To enable global-brand owners to manage product performance and the customer experience, Siemens PLM Software needed to collect, manage, search and analyze vast amounts of diverse data types, and it sought the right software and hardware infrastructure to support this effort.

Solution

The organization worked with Dell and Cloudera to build a software solution on top of the Cloudera® Distribution of Hadoop® (CDH) platform running on a cluster of Dell PowerEdge C8220 servers, giving customers total product data visibility throughout their entire supply chains.

Benefits

- Provides a 360-degree view of supply chain, product and customer data
- Searches billions of data records in seconds
- Scales to support 1 billion new records every month
- Helps pinpoint emerging issues within the supply chain
- Saves millions of dollars and boosts productivity
- Improves product performance and the customer experience

“We are able to help customers search billions of records in seconds with Dell infrastructure and support, Cloudera’s Hadoop solution, and our knowledge of supply chain and product performance issues. That’s a powerful combination for our customers.”

Dave Mitchell, CTO, Cloud Services,
Siemens PLM Software, United States



National Center for Supercomputing Applications (NCSA)



Business need

The private sector program within NCSA (working with industry and other production-oriented organizations) wanted help with compute and data challenges. NCSA sought to become more competitive through HPC, utilizing big data and real-time analytics to speed innovation.

Solution

- Implement big data Hadoop® solutions – such as MapReduce and Spark – for managing and analyzing data
- iForge®
- Dell PowerEdge Servers powered by Intel® Xeon® processors

Benefits

- Foster innovation with more platform convergence between HPC and big data analytics
- More competitive through HPC, along with business value from having advanced real-time analytic capabilities



“The uncompressed size of the data set that we ingest every night is about a terabyte and a half and needs to be processed promptly and iForge is an excellent platform for us to do our work.”

Don Petravac, NCSA Senior Project Manager

Texas Advanced Computing Center (TACC) inspires innovation with next-gen analytics



Business need

Traditional high-performance computing (HPC) is compute-intensive. But with research problems increasingly requiring analytics of larger data sets – both structured and unstructured data, often in real time – HPC needs extreme infrastructure performance. That's why TACC sought to design and build an HPC platform optimized for the computing and the storage demands of big data and analytics.

Solution

- Big data
 - Cloudera Hadoop
 - Data analytics
- Cloud solutions
- High-performance computing

Benefits

- Brings innovation to market faster
- Increases cycles of learning
- Improves engineering efficiency
- Expands research options and inspires innovation with next-gen analytics

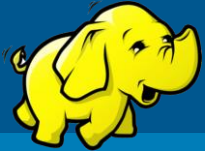


“We designed Wrangler around storage, whether researchers need to tap into 10 petabytes of long-term storage to preserve their data or access half a petabyte of DSSD flash memory to compute against data without having the boundaries of standard disk based file systems getting in the way...”

Niall Gaffney, Director, Data Intensive Computing,
Texas Advanced Computing Center

Find out more today...

Don't wait to harness the benefits of Hadoop on an optimized solution designed from the ground up to address data analytics requirements, reduce costs and optimize performance for deep data mining and analytics.



Contact your Dell EMC representative!

D~~E~~LL EMC