

POWER NEW POSSIBILITIES

Solutions for your data analytics journey

About this brochure

This brochure explains the capabilities and benefits of the Dell EMC options for starting on and maturing in your data analytics journey. It also outlines the overarching benefits of the many data platforms now available to your organization.

Unlock the value of your data

To realize the wealth available in data, your organization needs solutions, processes and tools that allow you to collect, integrate, process, manage, store and analyze data quickly, efficiently and cost-effectively. You will find all these capabilities, and more, in the Dell EMC™ portfolio of solutions for data analytics.

The Dell EMC portfolio provides a wide range of data analytics offerings, from infrastructure bundles, to strategic consulting, to data lakes to pre-engineered solutions, and in-house capabilities for advanced analytics and app-dev. The breadth of the Dell EMC portfolio makes it easy to find a solution that is right for where you are on your data analytics journey. This portfolio encompasses foundations, systems and solutions and spans the three phases in the customer journey: consolidation, analytics and analytics-based applications.

Together with partners, Dell EMC also offers a full complement of consulting, installation, implementation and support, and education services for data analytics. In short, the Dell EMC portfolio has everything your organization needs to unlock the value that exists within your data.

The digital transformation will cause disruption

Welcome to the fourth industrial revolution: digital transformation. This is a revolution that is fundamentally altering the way we live, work and relate to one another, and will transform the way every business in every industry operates. The growing accumulation of data, increasing diverse sources of data especially through the Internet of Things, and leveraging it all through data analytics is transforming the way every business operates. This ability to gather and make sense of this data transformation is dramatically reshaping industries and reinventing our future. Though considered an imperative, some companies are faring better than others in the race to become a data-driven organization. The revolution is already here.

Business leaders see a chaotic, uncertain future ahead. Consider these findings:

- 48 percent don't know what their industry will look like in three years
- 78 percent feel threatened by digital startups
- 45 percent fear they may become obsolete in three to five years¹

Digital transformation efforts are being inhibited by a wide range of challenges — two most often heard from organizations are the lack of budget and skilled resources. The lack of qualified resources points to a bigger issue: digital transformation is not a strategic focus for many organizations. Additionally, Gartner cited that early stage

¹"Embracing a Digital Future Study" research by Vanson Bourne and Dell Technologies, October, 2016

failures, up to 61 percent by 2017, are the direct result of the inability to define a use case and process to gain data insights.²

Despite these challenges, lines of business, competitors, customers and C-suite executives are recognizing the need to understand more about digital transformation to stay relevant. Change is patchy across organizations. At times, investments that are made are focused on parts of the organization rather than a holistic approach across the entire organization. The holistic and collaborative approach is required to bring IT and business together to support a strategic digital transformation approach. Leaders must agree that the following digital business attributes are imperatives for success:

- Predictively discover new opportunities
- Deliver unique and personalized experiences
- Innovate in agile ways
- Operate in real time

Successfully navigating in this new era requires a commitment from organizations to transform not only people and processes, but also modernizing the environment upon which a business is built — from the edge to the core to the cloud. Data analytics is at the core of these initiatives, and there are many opportunities for starting on the data journey. Data consolidation, whether adopting a data lake solution, self-service analytics or the Apache® Hadoop™ platform, will be a major prerequisite for organizations looking to capitalize on data using analytics.

The Value of Apache Hadoop

For organizations working to extract value from mountains of structured and unstructured data, the Hadoop data storage and processing system offers compelling benefits. It can store any kind of data from any source, cost effectively and at very large scale, and it can do sophisticated analysis of that data easily and quickly.

Hadoop is scalable, fault-tolerant and distributed. The open source software was originally developed by the world's largest Internet companies to capture and analyze the massive amounts of data they generate. Today we work with key strategic partners like Cloudera, Hortonworks and others to make this open source solution enterprise ready. Unlike earlier platforms, Hadoop can store any kind of data in its native format, and can be used to perform a wide variety of analyses and transformations on that data.

Both robust and reliable, Hadoop handles hardware or system failures automatically, without losing data or interrupting data analyses. Better still, Hadoop can run on clusters of highly optimized industry-standard servers. Each of those servers has local CPU and storage resources, and each has the flexibility to be configured with the proper balance of CPU, memory and drive capacity to meet specific performance needs.

Business-driven results

Hadoop solves the hard scaling problems that come with consolidating and analyzing large amounts of complex data. As the amount of data in a Hadoop cluster grows, new servers with local storage can be added incrementally and cost-effectively to expand the processing power of the cluster. And to further accelerate time to insights, Hadoop can be paired with many new technologies like Splunk, HBase, Impala, Kudu as well as Apache Spark™, a fast engine for large-scale data processing that uses in-memory computing for interactive queries, iterative processing, graph analysis and data streaming.

² Gartner Report: Predicts 2017: Digital Initiatives Must Focus on Long-Term Transformation to Avoid Failure, November, 2016

The Hadoop edge

Hadoop delivers several key advantages:

Store everything. Hadoop stores data in its native format, exactly as it arrives at the cluster. This allows you to avoid the downside of a common alternative, translating data on arrival so that it fits into a fixed data warehouse schema, which puts information into context. Because Hadoop stores data without forcing that transformation, no information is lost. Downstream analyses run with no loss of fidelity.

Control costs. Hadoop is open source software that runs on highly optimized industry standard x86 servers with massive storage capacity. That combination means that the cost per terabyte, for both storage and processing, is much lower than on older proprietary systems. As your storage and analytic requirements evolve, your Hadoop installation can, too.

Use with confidence. The Hadoop community, including both developers of the platform and its users, is global, active, established and diverse. Companies across many industries participate, including social networking, media, financial services, telecommunications, retail, healthcare and others.

Scale with confidence. You may not have petabytes of data that you need to analyze today. Nevertheless, you can deploy Hadoop with confidence knowing that organizations around the world, including some of the giants of search and e-commerce, run very large high performance Hadoop instances that manage enormous amounts of data. Hadoop is ready to scale with your needs.

Build new analytics capabilities. With Hadoop, there is value in an organization's ability and flexibility to apply traditional SQL and next generation non-SQL analytics to drive competitive advantage.

Hadoop makes it possible to conduct the types of analyses that would be impossible or impractical using any other database or data warehouse. Along the way, Hadoop helps you reduce costs and extract more value from your data.

Diverse use cases

Hadoop is different from traditional database and data warehousing systems, and those differences can be confusing to IT professionals. What data belongs in a Hadoop cluster? What kinds of questions can the system answer?

Here are a few of many examples of common use cases for Hadoop, as well as the questions that the use cases can help your organization answer:

1. **Risk modeling:** How do we better understand our customers and markets?
2. **Customer churn analysis:** Why do we lose customers?
3. **Recommendation engine:** How do we predict customer preferences?
4. **Ad targeting:** How do we increase the efficiency of your ad campaigns?
5. **Point-of-sale transaction analysis:** How do we target retail promotions that are sure to make customers buy?
6. **Analyzing network data to predict failure:** How do we use machine-generated data to avoid IT outages?
7. **Threat analysis:** How do we detect threats and fraudulent activity?
8. **Trade surveillance:** How can we spot a rogue trader?
9. **Search quality:** What's in our customers' search?
10. **Data sandbox:** What can we do with new data?

One of the keys to getting started with Hadoop is to identify the specific use cases that are right for your organization, and then identify the workloads that will best leverage a

Data analytics solutions at a glance

Getting started with Hadoop

- Dell EMC Ready Bundle for Cloudera Hadoop
- Dell EMC Ready Bundle for Hortonworks Hadoop
- Dell EMC Customer Solution Centers

Building a data lake for analytics

- Dell EMC Isilon Scale-out NAS

Extending your analytics capabilities

- Dell EMC Ready Bundle for SAP HANA Scale-Out
- Dell EMC Ready Bundle for Microsoft Exchange

Enabling and accelerating your journey

- Dell EMC Reference Architectures
- Dell EMC Ready Bundles
- Dell EMC Ready Bundle for MS SQL
- Dell EMC Boomi Integration Tools
- Pivotal Cloud Foundry

Hadoop environment. A related key to getting started is to identify the metrics you will use to gauge the success of your Hadoop deployment.

The Dell EMC portfolio for data analytics

Every organization's data is unique and must be treated as such. Solutions that are perfect for one company may not address the needs of another.

With this thought in mind, Dell EMC offers a wide range of products, solutions, and services to address diverse data analytics challenges — from learning how to identify your best use cases for data analytics, to infrastructure, starter bundles and validated reference architectures, to integrated appliances and engineered solutions, even completely customized solutions for your specific environment.

With any of these paths forward, Dell EMC offerings include enterprise-grade Hadoop solutions that are optimized for key workloads, and powered by Intel-based Dell EMC servers along with Dell EMC networking and storage technologies. You can now adopt Hadoop more rapidly and with confidence as key capabilities for security, performance, and management are jointly addressed by Dell EMC and its technology partners.

Getting started with Data Analytics

Deploying an end-to-end Hadoop environment

Dell EMC Ready Bundle for Cloudera Hadoop

For organizations ready to put a Hadoop environment into production, the Dell EMC Ready Bundle for Cloudera Hadoop provides a proven, end-to-end solution based on Dell EMC hardware and the latest Cloudera distribution of Hadoop.

These solutions enable organizations to streamline Hadoop deployments, from bare-metal server configuration to network setup to running CDH on a certified architecture based on the Dell EMC PowerEdge R730xd server, FX2 chassis, FC630 server nodes with FD332 storage modules, and Networking switches S4048-ON and S3048-ON. The end result is a solution that takes the guesswork out of building a Hadoop cluster.

Better still, 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.

Also included in this Ready Bundle is a QuickStart bundle to enable organizations to quickly and cost effectively deploy Hadoop for development and application teams to test business processes, data analysis methodologies, and operational needs against an end-to-end functioning Hadoop cluster. Additionally, this ready bundle includes a validated use case-driven solution for ETL offload. As traditional Enterprise Data Warehouses (EDWs) hit their limits, many forward-looking organizations are deploying the open source Hadoop platform as a complement to their existing EDWs and then offloading data warehouse processing functions to Hadoop. Working with Cloudera, Intel and Syncsort, Dell EMC delivers this solution to allow organizations to capitalize on the unique technical and cost advantages of the Hadoop platform while making better use of existing EDW investments.

Recently Dell EMC and Intel commissioned Forrester Consulting to conduct a Total Economic Impact™ (TEI) study to examine the potential return on investment (ROI) enterprises may realize by deploying Dell EMC Ready Bundle for Cloudera Hadoop, accelerated by Intel. The purpose of the study is to provide readers with a framework to evaluate the potential financial impact of the Hadoop solution on their organizations.

Following the Forrester study, the financial summary determined the three-year risk-adjusted results that include:

- A 97 percent ROI that equates to \$1.97 of benefit for every \$1.00 spent
- A 6 month payback period
- A net present value of \$5.6M

Key benefits

The Dell EMC Cloudera Hadoop Solution:

- Accelerates time-to-value with a validated solution
- Allows your team to focus on using Hadoop to gain deeper insights instead of designing a solution
- Maximizes server-based storage flexibility and performance with the Dell EMC PowerEdge R730xd server with Intel® Xeon® processors
- Provides access to Hadoop expertise to help reduce risk and fill the skills gap

Dell EMC Ready Bundle for Hortonworks Hadoop

The goal of faster time-to-value is at the heart of the Dell EMC Ready Bundle for Hortonworks Hadoop. It simplifies the architecture, design, configuration and deployment of Hadoop environments. Dell EMC engineers have validated and certified Hortonworks Data Platform (HDP) on Dell EMC PowerEdge™ R730xd servers and Dell EMC Networking switches S4048-ON and S3048-ON, allowing your organization to build a Hadoop cluster without the guesswork.

When you leverage this reference architecture, you have the assurance that Dell EMC engineers have certified the server and network configuration running Hortonworks HDP. Dell EMC develops an architecture document that provides guidance and know-how to help you successfully build a Hadoop cluster from bare-metal hardware.

With Dell EMC, you have the confidence that comes with experience. You can leverage our expertise to help fill the skills gap and build an architecture that will meet the needs of the business — while reducing risk.

Key benefits

The Dell EMC Hortonworks Hadoop Reference Architecture:

- Provides a secure, enterprise-ready, open source Hadoop distribution based on a centralized architecture (YARN)
- Addresses the complete needs of data-at-rest via HDP
- Delivers robust analytics that accelerate decision making and innovation
- Maximizes server-based storage flexibility and performance with the Dell EMC PowerEdge R730xd server with Intel Xeon processors
- Supports modern workloads and applications designed for the open networking era with an optimized data center top-of-rack (ToR) networking solution

Building a data lake for analytics

Dell EMC Isilon is an ideal foundation for a data lake, providing scale-out network-attached storage (NAS) to store, analyze, protect and manage your data. Everything is faster without the need to move and copy data across silos. With Dell EMC Isilon scale-out NAS, you have massive room for growth — up to 100 petabytes of capacity per cluster. You can scale both capacity and performance independently in about a minute to meet your specific business needs — all without additional IT burdens.

This product family can address your analytics and storage challenges with the Dell EMC Isilon Solution for Hadoop Analytics. Dell EMC Isilon is the first and only scale-out NAS

platform with native integration of the Hadoop Distributed File System (HDFS). When you use Hadoop with Dell EMC Isilon NAS, there is no need for data ingestion. You can run big data analytics in place, without moving data to a dedicated Hadoop infrastructure. The Isilon platform is also an ideal foundation for a data lake that requires the ability to scale out while supporting traditional and next-generation applications and workloads.

Additionally, Dell EMC can provide DAS-based data lakes featuring PowerEdge servers as well as data lakes that include the third-generation object platform from Dell EMC, Elastic Cloud Storage (ECS). This platform is designed for next-generation applications and traditional workloads with unmatched storage efficiency, resiliency and simplicity.

Key benefits

A Dell EMC data lake enables you to:

- Deliver faster times to insights
- Meet storage capacity demands easily
- Expand to a public or private cloud
- Eliminate silos of data within your organization

Extending your analytics capabilities

Dell EMC Analytic Insights Module

Analytic Insights Module from Dell EMC delivers a self-service analytics experience for rapidly transforming data into actionable insights with high business value. This helps organizations gather, analyze and act on data insights to monetize new digital business opportunities in weeks rather than months.

The Analytic Insights Module is delivered on the foundation of Native Hybrid Cloud, a turnkey developer platform based on Pivotal Cloud Foundry. Analytic Insights Module combines self-service data analytics with cloud-native application development into a single-cloud platform, eliminating the months it takes to build your own solution. This helps organizations gather, analyze and act on data insights to monetize new digital business opportunities and unlock the value of their data.

The Analytic Insights Module delivers flexible, deep data awareness, and the fastest time to realizing the value of your data through a unique Data Curator, Data Governor, Data and Analytics Catalog, and integrated data lake by providing a single view of all your discovered, indexed and curated data sets both within the data lake (Isilon) and selected external sources. Data analytics teams are provided security and resource quotas by IT for role-based, self-service access to easily stand up independent workspaces with Hadoop clusters in minutes without disrupting the enterprise infrastructure. Analytic Insights Module makes it easy to share data and analytics services with the application development framework for collaborating with developers on data-driven applications, services and business processes.

Analytic Insights Module is an open platform, allowing data analytics teams to choose from any of the analytics and visualization tools offered through Pivotal Cloud Foundry or from the expansive ecosystem beyond Analytic Insights Module. For example, Analytic Insights Module supports Cloudera and Hortonworks Hadoop distributions and a whole host of visualization and analytics tools. It is the perfect platform to support business use cases ranging from electronic fraud detection and customer experience optimization to predictive machine maintenance.

Key benefits

Analytic Insights Module provides the tools to:

- Gather the right data with deep awareness to provide a single view of all your data
- Analyze your data through a self-service experience
- Act on data insights for monetizing new opportunities through cloud native applications

Enabling and accelerating your journey

Adding to the Dell EMC Hadoop solutions, Dell EMC enables and accelerates the data analytics journey with the reference architectures, and tested and validated solutions that complete the Dell EMC big data and analytics portfolio.

Reference architectures

Dell EMC offers tested and certified reference architectures for small, medium and large data analytics deployments, drawing on technologies from Dell EMC and strategic partner industry leaders. These reference architectures include all the hardware and software needed to contain costs and increase the scalability of a solution for a complete big data environment.

Tested and validated solutions

The following solutions from Dell EMC bring together the components of complete data analytics environments in turnkey appliances that are easy to order and quick to deploy.

Dell EMC Ready Bundle for Microsoft

Dell EMC delivers a solution for real-time management of relational (SQL) and non-relational (Hadoop) data with Microsoft acting as your single point of contact for all patches, upgrades and issues. This complete data warehousing solution spans desktop systems, data marts, and the enterprise data warehouse — all while operating seamlessly with Microsoft business intelligence tools.

Dell EMC SAP HANA solutions

Dell EMC delivers an in-memory analytics appliance engineered by Dell EMC and certified by SAP to ensure non-disruptive, modular scalability from 2 to 18 nodes, without requiring replacement or reconfiguration. Dell EMC is currently the only SAP-certified partner to offer a truly affordable solution for in-memory, real-time analytics. The Dell EMC Ready Bundle for SAP HANA Scale-Out, the Dell EMC SAP HANA Scale-Up Appliance, and the Dell EMC SAP HANA Edge Appliance bring the analytical power and speed of SAP HANA to the masses.

Dell EMC Boomi Integration Tools

Dell Boomi API Management is a single scalable platform that supports the complete API lifecycle, enabling organizations to create, publish and centrally manage APIs on-premises or in the cloud. Its unified capabilities simplify the creation, distribution and governance of APIs, which accelerates business-to-consumer (B2C), business-to-business (B2B) and mobile interactions.

Dell EMC solution installation and implementation

Dell EMC offers a broad menu of installation and implementation services for data analytics, ranging from the initial steps into Hadoop to services that include onsite hardware and software installation, optional rack integration at a Dell EMC facility and validation of the installed solution. In all the Dell EMC data analytics solutions, Dell EMC takes care of the complete project management, from order drop to customer acceptance.

Dell EMC consulting services for Hadoop and data analytics

Dell EMC Hadoop Consulting

Dell EMC Hadoop Consulting is a best-in-class service delivered by certified Apache Hadoop experts to help organizations fully exploit the business value of big data using Hadoop. The services may include a big data assessment, workshop, proof-of-concept or production implementation. These Hadoop experts help determine where Hadoop is a good fit for an organization. They also help Dell EMC customers build their own teams of Hadoop experts through knowledge transfer at each step.

Dell EMC Services for data analytics

Dell EMC Services offerings start with strategic consulting and deployment services to help business leaders prioritize where and how to apply data and analytics to drive their business initiatives, and to help technology leaders understand how to build big data and analytics capabilities into their IT architectures and processes.

The Big Data Vision Workshop employs a unique methodology to identify and prioritize a single analytics use case with the best combination of implementation feasibility and business value. It's a three-week engagement that applies research, interviews, data science expertise and techniques to your organization — culminating in a one-day workshop to identify and agree on the best data analytics use case and path forward to solving a business problem. From here, customers move on to a Proof of Value service, to prove out the potential ROI of a proposed analytics solution.

Technology leaders who seek to expand analytics capabilities to meet growing demands for better/faster data and analytics, can start with the Big Data Technology Advisory service, a three-week engagement to identify strategic capabilities for data, determine future-state architecture, perform a gap analysis, and build a deployment roadmap. They can then opt for a Proof of Technology service to pilot their solution and confirm the desired capabilities before moving into production.

At this point, the customer can choose to deploy a use case as a custom solution that is designed to support that one use case, or deploy a custom data lake-based platform that can support additional use cases. This is the Big Data Technical Implementation. Alternatively, the customer can choose to implement an analytics use case by deploying a packaged, engineered solution, such as the Analytic Insights Module, with included One-Call support. Dell EMC Services offers comprehensive consulting engagements to develop and implement an analytics use case into production, including solution architecture, data ingestion, data science services, and application integration.

Dell EMC ProSupport

Dell EMC ProSupport offers a single point of accountability from experts with solution-specific training, along with premium hardware and software support available 24x7x365. ProSupport also includes collaborative support for Cloudera Enterprise software. Additionally, ProSupport includes next-business-day onsite service with four- and eight-hour parts and labor response options, and escalation management with customer-set severity level options. Start your data analytics journey with Dell EMC.

Accelerating your digital transformation with Dell EMC

Dell EMC has what it takes to help you gain hands-on experience across many solutions, from initial briefings, through a proof of concept, and into a full production environment. The Dell EMC approach is guided by proven reference architectures and Ready Bundles, enabled by solutions along the life cycle of data, from data lake to self-service analytics, enabled by validated solutions and fully supported by the Dell EMC Global Services organization.

Data analytics is a journey. Beginning with a partner who encourages collaboration, has a depth of understanding in data analytics, a wide breadth of solutions, and who drives to success through secure and supported growth can lead organizations to success at every step of the journey.

To learn more, visit DellEMC.com/BigData

Copyright © 2017 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners.

Published in the USA 03/17. Brochure

Dell EMC believes the information in this document is accurate as of its publication date. The information is subject to change without notice.