

CREATING A FOUNDATION FOR BUSINESS VALUE

Building initial use cases to drive predictive
and prescriptive analytics

ABSTRACT

This white paper highlights three initial big data use cases that can help your organization create a foundation for predictive and prescriptive analytics: offloading data preparation work from the enterprise data warehouse, creating an online active archive and establishing a data lake to serve as a single repository for all types of data.

October 2016

Dell is a trademark of Dell Inc. and EMC is a trademark of EMC Corp. Dell Technologies is a trademark of Dell Inc.

Copyright © 2016 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 trademarks of their respective owners. Intel, Xeon and the Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

TABLE OF CONTENTS

TURNING DATA INTO INSIGHTS4

THE PATH TO ADVANCED ANALYTICS4

KEY STEPS ON THE PATH TO MODERN ANALYTICS5

 ETL offload 5

 Active archive 5

 Data Lake 6

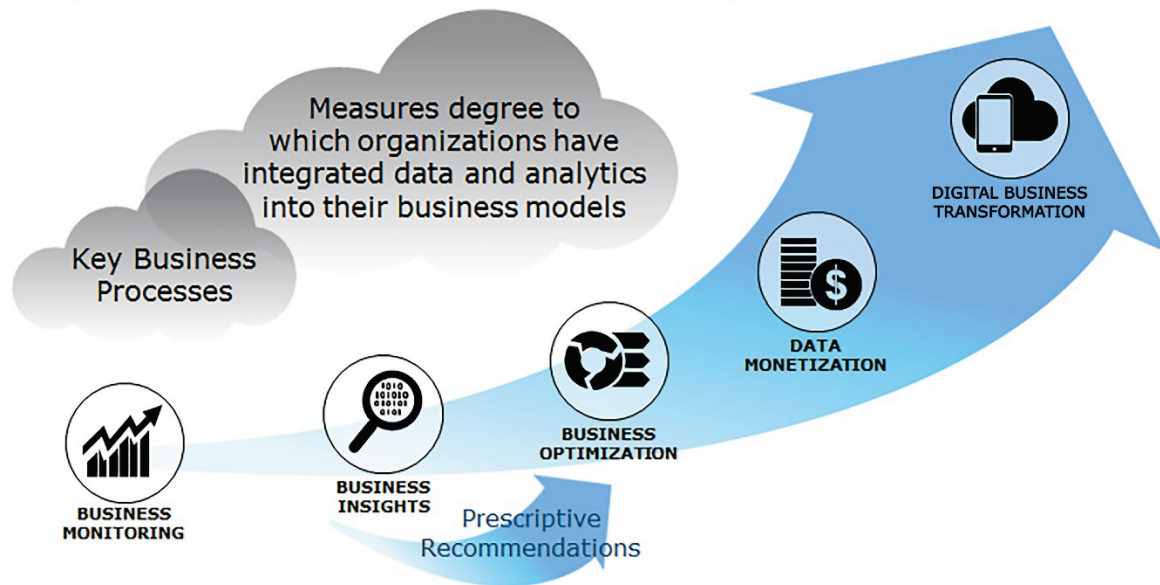
LET’S GET STARTED.....6

TURNING DATA INTO INSIGHTS

Organizations are realizing that they are sitting on a wealth of data. However, they lack a process for turning that data into superior customer, product, operational and market insights that can be used to power key business processes, uncover new monetization opportunities and drive a more compelling, more prescriptive customer engagement. A business use case-centric approach can ensure that organizations are making their data and analytics investments in areas that hold the highest business potential or return on investment (ROI).

At Dell EMC, we use the Big Data Business Model Maturity Index as a benchmark to help our clients understand where they sit today versus their competitors, and also as a roadmap for helping them to become more effective at leveraging data and analytics to power their business models.

BIG DATA BUSINESS MODEL MATURITY INDEX



THE PATH TO ADVANCED ANALYTICS

Data is now recognized as a critical asset for all types and sizes of organizations, but data alone will not enable organizations to compete. The value emerges through the strategic processes that organizations put in place to integrate and manage the data that can drive analytics use cases, which then allows them to use analytics tools that mine data to extract meaningful information and rich insights that help business units throughout the organization make better decisions. In recognition of this, many organizations are now embracing modern analytics tools that use sophisticated algorithms to predict what is likely to happen in the future and to prescribe courses of action.

These predictive and prescriptive analytics tools go far beyond the capabilities of traditional tools that mine static data to present a view of “what happened.” However, that doesn’t mean those modern analytics tools are the starting point for an organization that wants to capitalize more fully on its data. In fact, advanced analytics projects are prone to failure for organizations that don’t have the right big data foundation in place — a point borne out by the many stories of well-intention analytics projects that never came to fruition.

So how do you get started with a foundational plan that will guide you down the path to more advanced analytics? This journey begins with basic big data capabilities and use cases, and progresses forward in clear, logical steps as shown here:

1. Data access and collection	Lay the groundwork for your data to thrive. Get the fundamentals in place to track and store your data, even if it is in a spreadsheet.
2. Information consolidation	Derive relationships from your data. Consolidate and connect your data in a logical and meaningful way.
3. Proactive reporting	Establish benchmarks. Measure success. Begin using a business intelligence platform to track your performance against key performance indicators (KPIs) that you set for yourself.
4. Predictive analytics	Study the past to forecast the future. Go beyond standardized reporting. Predict what the future holds based on the patterns observed on historical data.
5. Prescriptive analytics	Recommend what actions to take. Go beyond predicting what the future will look like to prescribe what specific actions your organization needs to take.

KEY STEPS ON THE PATH TO MODERN ANALYTICS

While there is no universal roadmap that defines the path to modern analytics, many organizations begin with a Hadoop-based data lake while leaving the traditional business intelligence (BI) tools and processes in place. This allows organizations to continue to meet the demands of an existing enterprise data warehouse while adopting technologies that support the path to a unified big data platform that accommodates all types of data.

Let's walk through a few scenarios that help organizations gain control over big data and put the foundational elements in place to support advanced analytics applications.

ETL OFFLOAD

This use case for big data technology focuses on driving down the costs of an enterprise data warehouse (EDW) environment by offloading the heavy lifting of data preparation — known as ETL, for “extract, transform and load” — to a lower-cost data processing environment, such as the open source Apache™ Hadoop™ platform on Intel-based systems, where all forms of data can be collected, managed, analyzed and stored. This can help organizations stem skyrocketing costs driven by a constant need to expand the EDW to keep pace with growth, as it enables organizations to manage and analyze both structured and unstructured data.

In another benefit, ETL offload eases the processing burden on the EDW, so its resources can be used for the things it was designed to do, such as running analytics queries and generating reports on the data from enterprise systems.

ACTIVE ARCHIVE

Another common initial use case for big data technology is active archiving. An active archive is a system that enables an organization to capture, retain, search and query data within a cost-effective online archiving environment. Unlike deep archiving solutions and offline archives, which tuck data away mainly for safekeeping, an active archive keeps data readily accessible to business users while storing it at a relatively low cost.

To support this use case, many organizations deploy the Hadoop platform on Intel-based Dell EMC™ PowerEdge™ servers to collect, manage, analyze and store any type of data, in any format, from any source, inexpensively and at very large scale, without requiring a great deal of data cleansing and parsing before archiving. And while providing cost-effective data archiving, a Hadoop environment can enable broad organizational access to varied data sets for ad-hoc analysis — without pulling the data into an EDW.

DATA LAKE

A data lake creates a single repository for all types of data — structured, semi-structured and unstructured. It consolidates storage silos into a scale-out platform with massive capacity, to form a single system to capture, store, process, analyze and protect an organization's data. Many organizations have successfully built data lakes based on the combination of Apache Hadoop, Dell EMC servers with Intel® Xeon® processors and the Dell EMC Isilon scale-out network-attached storage (NAS) platform, which has native integration of the Hadoop Distributed File System (HDFS).

A data lake goes beyond the capabilities of an EDW. While the typical EDW has tools for data analysis, it doesn't handle unstructured and semi-structured data, and it is very expensive to scale. Nevertheless, due to its history in an organization and through its ability to enable queries and reporting on the data from enterprise systems, the EDW will continue to play an important role in the big data environments of organizations that create data lakes to augment their existing EDWs.

LET'S GET STARTED

When it comes to capitalizing on big data, advanced analytics is a means to an end. The goal is to drive strategic business initiatives by using data and analytics to make better business decisions and uncover monetization opportunities.

If your organization is on this path to business value, Dell EMC is your ideal technology partner. Together with Intel and our software partners, we offer a broad portfolio of solutions and services that covers all of the layers of the analytics maturity model. We can help you evaluate the readiness of the systems you have in place today as well as the available building blocks for your big data foundation — from appliances to engineered systems, from modern infrastructure to data integration tools.

In short, Dell EMC is ready to help you put analytics solutions to work to unlock the value hidden in your data.

To learn more, visit Dell.com/AllData or EMC.com/BigData



IDC prediction

By 2020, 50% of all business analytics software will incorporate prescriptive analytics built on cognitive computing functionality.¹

¹ IDC FutureScape: Worldwide Big Data and Analytics 2016 Predictions. November 2015.