

FIVE CRITICAL SUCCESS FACTORS FOR BIG DATA PROJECTS

Take a strategic approach to capitalizing on your data

ABSTRACT

This white paper explores five critical success factors for big data projects, from establishing your vision to executing your project. The paper notes that the path to project success begins not with a particular technology or solution but with a clear business use case and a strategic roadmap to the future.

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.

TABLE OF CONTENTS

THE ROAD TO SUCCESS WITH BIG DATA4

FIVE CRITICAL SUCCESS FACTORS FOR YOUR BIG DATA JOURNEY.....4

1. Establish your vision.....4

2. Align your business and IT use cases5

3. Assess your readiness5

4. Establish your key performance indicators6

5. Execute your project.....6

KEY TAKEAWAYS.....7

THE ROAD TO SUCCESS WITH BIG DATA

The combination of big data and powerful analytics tools is a key to success in the era of the digitally driven business. Organizations now have the raw material they need to make better decisions, build closer customer relationships, optimize their operations and meet countless other goals that would not be possible without the insights derived from big data. Yet they are still struggling to build end-to-end data architectures that enable deeper and faster business insights.

In this new era, all organizations capture and store enormous amounts of data, from enterprise business systems to new data sources like social media, log data and the Internet of Things. That is not only a good but a necessary direction. Yet as many organizations have found, huge datasets alone will not do anything for you unless you have the right solutions in place to turn raw data into valuable insights. Organizations must build a foundation where there is an operational model and process for how data is ingested, stored, processed, transformed, analyzed and then shared across the organization.

Here is where things get a bit more difficult. Finding the right big data solutions, focused on the right use cases, is a challenging proposition. That point is evident in headlines that warn of the potential for problems with big data initiatives. A few examples:

- 8 Reasons Big Data Projects Fail (Information Week)
- Study reveals that most companies are failing at big data (CIO.com)
- Why do so many analytics projects fail? (Analytics magazine)
- Big Botched Data (Gartner)

Analysts and other industry observers point to various reasons why some big data projects struggle, including a lack of business-IT alignment, poorly defined end goals, unrealistic use cases, cultural issues and skills shortages, just to name a few.

So how do you put your organization on the path to a successful big data initiative? In Dell EMC's experience with enterprises around the world, the process of capitalizing on data doesn't begin with the deployment of a Hadoop environment or the adoption of leading-edge analytics solutions. It begins with understanding your business use case(s) and then building out a strategic roadmap that will take you from where you are today to where you want to be tomorrow.

With that thought in mind, let's look at five key milestones on the road to success with big data.

FIVE CRITICAL SUCCESS FACTORS FOR YOUR BIG DATA JOURNEY

This roadmap to a successful big data initiative has five key milestones that mark your forward progress. Each of these milestones amounts to a critical success factor for a big data initiative.

1. ESTABLISH YOUR VISION

The first milestone is to gain a clear view of what you are trying to accomplish with big data. The fact that you capture terabytes of data on a daily basis is meaningless if you don't have a clear view with a plan of action as to what you want to accomplish with that data. Ask yourselves: What are the use cases? What business insights are we trying to drive? Do we have the correct tool set?

For example, do you want to become a data-driven organization, in which the analysis of information guides corporate strategies, business initiatives and customer relationships? If this is a top goal, then the vision for your big data initiative should include the ability to provide business executives with a steady stream of actionable information derived from big data.

Your vision can also include operational goals, such as putting new technologies to work to improve response times or to contain costs, such as those associated with runaway growth of an enterprise data warehouse (EDW).

2. ALIGN YOUR BUSINESS AND IT USE CASES

The second milestone on the road to a successful big data initiative is the alignment of IT with business goals and use cases. This is a critical success factor for any big data initiative. Without business and IT alignment, projects will likely fail because IT is not prepared to deploy or maintain big data technologies that require a new skill set.

Additionally, it's essential that business leaders feel they have an ownership interest in big data initiatives and believe that their investments in big data solutions will deliver measureable returns. One way to gain this buy-in is to discuss in detail and agree on the goals, objectives and use cases for big data projects.

Let's look at a few use cases that can help your organization get started with big data in a manner that meets the objectives of both business and IT.

ETL OFFLOAD

This use case focuses on containing the costs and productivity drains associated with explosive data growth. ETL offload can control the costs of an enterprise data warehouse environment by offloading the heavy lifting of data preparation work — known as ETL, for “extract, transform and load” — to a lower-cost data processing environment, such as the open source Apache™ Hadoop™ platform. This use case can help organizations stem skyrocketing costs driven by a constant need to expand the EDW to keep pace with growth.

ACTIVE ARCHIVING

Another common initial use case for big data 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, such as the Hadoop data storage and distributed processing platform. Unlike deep archiving solutions and offline archives, which tuck data away mainly for safekeeping, an active archive keeps data readily accessible to business users. At the same time, the use of an environment like Hadoop enables enterprises to archive data at a fraction of the cost of a traditional EDW.

DATA LAKE

A data lake acts as a comprehensive, unified repository for all of an organization's data. Users gain a consolidated view of data for analytics implementations that generate deeper insights than would be possible if data were held in multiple silos. The goal is to make all relevant data quickly and easily accessible, with no wasted motion in locating and identifying required data sets and running analytics queries.

LOG AGGREGATION

The log aggregation use case continually collects and analyzes log data from different sources in the IT environment, allowing IT operators to work proactively to maintain the health and performance of IT systems, rather than responding to potentially catastrophic outages.

CUSTOMER 360

Another common use case is to put big data and analytics tools to work to create a 360-degree view of an organization's customers. This all-encompassing view enables the business to interact with individual customers through more channels and on their own terms, which can be key in driving both improved customer experience and greater sales volumes.

3. ASSESS YOUR READINESS

The next key milestone is the assessment of the readiness of your IT environment and in-house skill sets to implement your big data project and empower members of your existing team as citizen data scientists throughout your organization to put the power of analytics to work to drive your business forward.



“By 2020, data monetization efforts will result in enterprises pursuing digital transformation initiatives increasing the marketplace's consumption of their own data by 100-fold or more.”

— IDC prediction¹

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

TECHNOLOGY

This assessment phase includes an evaluation of the systems you have in place, as well as the consideration of available solutions and engineered systems for big data environments that may be used to augment your current infrastructure. The key here is to look for offerings that integrate with your current environment and are based on proven technologies for your use cases.

PEOPLE

If your use cases call for technologies that are new to your organization, such as the Apache Hadoop platform to store and process large datasets, you have to decide how you will cultivate the required technical expertise. Will you rely on training existing personnel or create new positions that bring in expertise from outside the organization? Do you really need in-house deployment and integration expertise, or will you rely on professional services to deploy the solutions while you concentrate your efforts on developing ongoing operations capabilities? Similarly, you need to determine whether your organization needs to cultivate or hire personnel to fill new managerial positions, such as chief data officer.

PROCESSES

It's also essential to look at the processes you will need to have in place to operationalize your big data environment. For example: How will you eliminate today's data silos? Will all data be accessible via a data lake? What are your criteria for authorizing access to data based on users' roles within the organization? How will you enable authorized users to access the data they need? Where will you create the perimeter access controls? How will you enable and educate your teams? Who will provide technical support for your data integration tools and analytics applications?



"Our research also found that most organizations lack the required skills, technical capabilities and culture to truly gain the greatest advantage from their information. In fact, three in four businesses extract little or no advantage whatsoever."

— PricewaterhouseCoopers (PwC)²

4. ESTABLISH YOUR KEY PERFORMANCE INDICATORS

Key performance indicators (KPIs) are an essential component of any big data project. Your KPIs are like a gauge for your success, and they are not one-size-fits-all. They will be unique for every organization, for each use case, for each phase of a project and for each step in your big data journey.

Here are some examples of KPIs you might consider:

- By augmenting the EDW with Hadoop, our organization will reduce costs by X percent while giving business analyst Y percent faster access to data for queries or reporting.
- Our big data environment will allow us to reduce data storage costs by X percent while processing data XX faster.
- We will use data-driven analytics to increase our customer acquisition rates by X percent.
- We will use analytics to personalize and streamline the user experience when interacting with our e-commerce site, including reducing the number of required clicks to buy a product by X percent.

As a general rule, more specific KPIs will help you better gauge the success of your project.

5. EXECUTE YOUR PROJECT

The final milestone on the road to a successful big data initiative is the implementation of your first project. It's widely held that the best practice is to start with a narrowly targeted proof of concept (PoC), achieve success at that level, and then expand your PoC into a production environment.

² PricewaterhouseCoopers (PwC). "Seizing the information advantage: How organizations can unlock value and insight from the information they hold." A PwC report in conjunction with Iron Mountain. September 2015.

The goal is to achieve incremental successes as your team develops its skills and comfort level with the technologies available for big data and analytics projects, and then move on to new use cases. For example, you might get started with an ETL offload project, and then move on to applications that incorporate technologies for predictive and prescriptive analytics to drive better business decisions.

KEY TAKEAWAYS

All large organizations capture and store big data, yet many don't have the systems and processes in place to capitalize on all that data. This puts organizations at a disadvantage in today's digitally driven marketplaces.

The question isn't whether you need to invest in big data and analytics solutions. All organizations need to do that to ensure they provide the best customer experience or remain competitive. The real questions must be asked to determine the steps you need to take to get on the path to a successful big data project.

At Dell EMC, we know that a successful big data project doesn't begin with the deployment of a particular technology or solution. It begins with a business use case and a strategic roadmap that will take you from where you are today to where you want to be tomorrow.



"By 2020, organizations able to analyze all relevant data and deliver actionable information will achieve an extra \$430 billion in productivity benefits over their less analytically oriented peers."

—IDC prediction³

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

³ IDC FutureScape: Worldwide Big Data and Analytics 2016 Predictions. November 2015.

