A Custom Technology Adoption Profile Commissioned By Dell EMC And Hortonworks | September 2016

Optimize Your Enterprise Data Warehouse With Hadoop

GET STARTED ▶

Optimize Your Enterprise Data Warehouse With Hadoop

Enterprise data warehouse (EDW) technology has been a reliable resource for organizations for more than a decade. It can provide business units with timely and actionable insights that drive business decisions and outcomes. However, EDWs and the teams that manage them are strained by the rapid increases in data volume and the demand for more analytics. To meet the demand, organizations are implementing Hadoop as an effective way to augment and optimize the use of EDW resources.

In June 2016, Dell EMC and Hortonworks commissioned Forrester Consulting to evaluate the adoption of Hadoop to offload enterprise data management and analytics.

Demographics

109 US and UK IT decision-makers in infrastructure, storage operations, and app dev roles responsible for storage or big data decisions

Position

› C-level: 9%
› VP: 20%
› Director: 42%
› Manager: 29%

Level of responsibility for storage or big data

› 32% are the final decision-maker
› 47% are part of a team making decisions
› 21% influence decisions

Company size by employees

› 500 to 999: 9%
› 1,000 to 4,999: 46%
› 5,000 to 10,000: 22%
› 10,001 to 19,999: 15%
› 20,000 or more: 8%

Hadoop Is Mainstream

Hadoop has become data's darling. Survey results indicate that 78% of IT decision-makers work for companies that have already adopted, are in the process of adopting, or are expanding their adoption of Hadoop. Another 22% of respondents are planning to adopt Hadoop in the next 12 months.

Hadooponomics, the cost-effective scalability of Hadoop, drives adoption for a wide range of data management and analysis requirements. IT decision-makers look at Hadoop as a core component of a modern data architecture. They use it primarily to lower the overall cost of data and analytics, perform advanced analytics, and build a data lake to aggregate data from multiple sources.

Companies are adopting Hadoop primarily to lower data and analytics costs and perform more advanced analytics.

What drove/is driving your company to adopt Hadoop?
(Select all that apply)
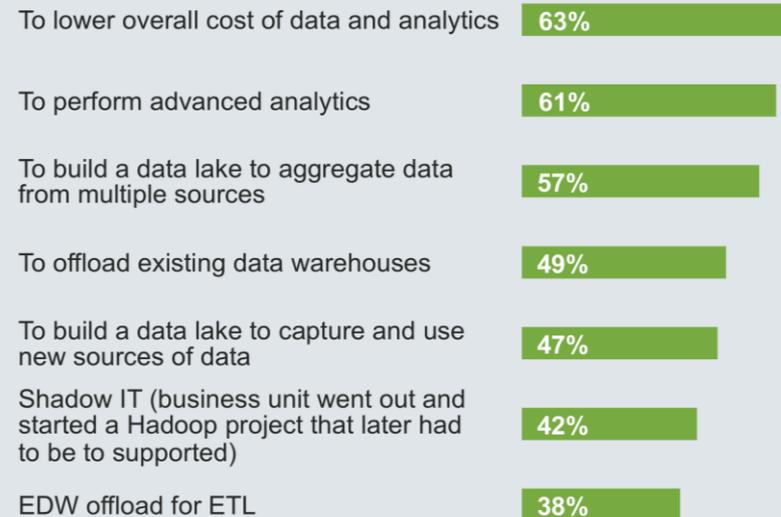
To lower overall cost of data and analytics	63%
To perform advanced analytics	61%
To build a data lake to aggregate data from multiple sources	57%
To offload existing data warehouses	49%
To build a data lake to capture and use new sources of data	47%
Shadow IT (business unit went out and started a Hadoop project that later had to be to supported)	42%
EDW offload for ETL	38%

Base: 109 IT decision-makers in infrastructure, storage operations, and application development roles in the US and the UK
Source: A commissioned study conducted by Forrester Consulting on behalf of Dell EMC and Hortonworks, July 2016

Optimize Your Enterprise Data Warehouse With Hadoop

1 **2**

Hadoop Data Lakes Coexist Nicely With Data Warehouses

While lowering costs and improving analytics are the primary drivers for adopting Hadoop, about half of IT decision-makers use it to offload existing data warehouse workloads. And these decision-makers populate their Hadoop data lake with a variety of data sources, including network attached storage (NAS) or EDWs.

IT professionals also continue to recognize the importance of EDWs in their data and analytics architecture. Almost 80% of respondents expressed interest in optimizing their EDW by offloading unknown-value, lower-value, or archive data storage and processing functions to a more cost-effective Hadoop data lake platform.

Almost 80% of respondents expressed interest in optimizing their EDW to a more cost-effective Hadoop data lake platform.

What are the data sources that populate (or plan to populate) your Hadoop cluster? (Select all that apply)
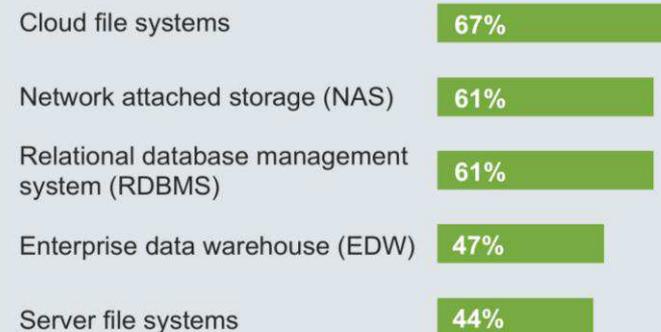
Data source	Percentage
Cloud file systems	67%
Network attached storage (NAS)	61%
Relational database management system (RDBMS)	61%
Enterprise data warehouse (EDW)	47%
Server file systems	44%

Base: 109 IT decision-makers in infrastructure, storage operations, and application development roles in the US and the UK
Source: A commissioned study conducted by Forrester Consulting on behalf of Dell EMC and Hortonworks, July 2016

Optimize Your Enterprise Data Warehouse With Hadoop

1 **2**

Yet Organizations Are Immature When Offloading EDW Data To Hadoop

Despite Hadoop's growing adoption, organizations struggle to decide how to balance data between their EDW and data lake. Over 60% of IT decision-makers stated that they face challenges with identifying exactly what data should remain within the EDW and what should move to Hadoop. Others indicated they struggle with managing costs and risk with custom solutions, especially when ensuring business continuity or compliance. This means IT professionals and their organizations need help with best practices and tools when it comes to identifying the data that can be moved.

Organizations struggle to decide how to balance data between their EDW and data lake.

What challenges is your organization experiencing with offloading EDW data to Hadoop? (Select all that apply)
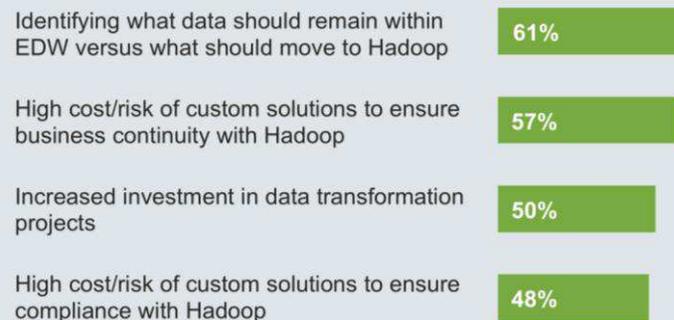
Identifying what data should remain within EDW versus what should move to Hadoop	61%
High cost/risk of custom solutions to ensure business continuity with Hadoop	57%
Increased investment in data transformation projects	50%
High cost/risk of custom solutions to ensure compliance with Hadoop	48%

Base: 109 IT decision-makers in infrastructure, storage operations, and application development roles in the US and the UK
Source: A commissioned study conducted by Forrester Consulting on behalf of Dell EMC and Hortonworks, July 2016

Optimize Your Enterprise Data Warehouse With Hadoop

1 2

Hadoop Alone Is Not Enough

To be successful, organizations require a solution that manages, governs, and processes data in a Hadoop data lake. Over 80% of IT decision-makers in this study value a solution that ensures business continuity, data protection, and meeting regulatory and enterprise policy requirements with a data lake. Around 60% of IT decision-makers want a solution that allows them to consolidate cross-enterprise data, archive cold data, and offload ETL (extract, transform, and load) and ELT processing from the EDW.

IT decision-makers require a solution that manages, governs, and processes data in a Hadoop data lake.

How valuable are the following features and capabilities in a solution that manages, governs, and processes data in a Hadoop data lake?
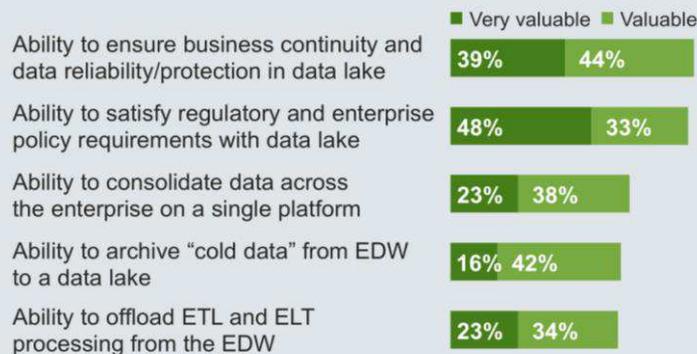
■ Very valuable   ■ Valuable

	Very valuable	Valuable
Ability to ensure business continuity and data reliability/protection in data lake	39%	44%
Ability to satisfy regulatory and enterprise policy requirements with data lake	48%	33%
Ability to consolidate data across the enterprise on a single platform	23%	38%
Ability to archive "cold data" from EDW to a data lake	16%	42%
Ability to offload ETL and ELT processing from the EDW	23%	34%

Base: 109 IT decision-makers in infrastructure, storage operations, and application development roles in the US and the UK
Source: A commissioned study conducted by Forrester Consulting on behalf of Dell EMC and Hortonworks, July 2016

1 2

Seamless Data Access To All Sources Overcomes That Data Deluge

Organizations expect to experience a variety of benefits as a result of implementing a data lake with Hadoop technology built over enterprise-grade storage. IT professionals want enterprise-class features and capabilities to manage and scale their Hadoop environment. IT professionals anticipate more comprehensive analytics on more data as the primary benefit. Lower costs of data and analytics is also an important benefit for 67% of IT decision-makers.

IT professionals anticipate more comprehensive analytics and lower costs as a result of implementing a data lake with Hadoop technology built over enterprise-grade storage.

What benefits does your organization expect to experience by implementing a data lake with Hadoop technology built over enterprise-grade storage? (Select all that apply)

Benefit	Percentage
More comprehensive analytics as a result of removed data silos	72%
Lower overall data and analytics costs while conforming to enterprise policy/regulatory requirements	67%
New products and services based on new data sources and analytics	62%
Ability to cope with unstructured and structured data	57%

Base: 109 IT decision-makers in infrastructure, storage operations, and application development roles in the US and the UK
Source: A commissioned study conducted by Forrester Consulting on behalf of Dell EMC and Hortonworks, July 2016

Hadoop Is The Epoch Of A Data And Analytics Renaissance

More data. More analytics. Faster time-to-insights without breaking the bank. That's what organizations strive for in building out a modern data and analytics architecture that includes both Hadoop and EDWs. Seamless and complete access to all enterprise application data is key.

Organizations aim to build a modern data and analytics architecture that gives them more data, more analytics, and lower costs.

METHODOLOGY

› This Technology Adoption Profile was commissioned by Dell EMC and Hortonworks.

› To create this profile, Forrester leveraged its existing research on data management.

› Forrester Consulting supplemented this research with custom survey questions asked of US and UK IT decision-makers responsible for storage or big data decisions.

› The custom survey began and was completed in July 2016. For more information on Forrester's data panel and Tech Industry Consulting services, visit forrester.com.

Project Director:

Andia Vokshi
Market Impact Consultant

Contributing Research:

Forrester's Application Development and Delivery research group