IDC
ANALYZE THE FUTURE

Sponsored by: **MapR**

Authors:
Carl W. Olofson
Matthew Marden

January 2020

Taking a Short Path to Value: The Business Value of the MapR Data Platform

Business Value Highlights

567%
five-year ROI

11 months
to payback

20x
more queries per organization

33%
higher productivity for data analytics teams

30%
higher productivity for impacted developers

85%
less unplanned downtime

57%
lower five-year cost of operations for data analytics platform

IDC OPINION

For over 10 years now, many organizations of all sizes have derived considerable benefit from concentrating large amounts of data in data lakes and creating a variety of applications that can perform analytics on that data. Often, however, the community open source Hadoop and its constellation of Apache projects have been found to form a hard-to-manage and cumbersome analytics environment. Some analytics are very difficult to perform with reasonable performance and scalability on Hadoop. This is due to a fundamental limitation of the platform: HDFS is an append-only sequential file system. This forces all analytic processes to read every record in order to do the job.

Various software vendors have provided enhancements and workarounds to address this problem, but despite these efforts, many, if not most, users today collect and curate their data on Hadoop but move it to other environments, such as object storage clusters with Spark, to do their analytics. Real-time streaming data may need to be moved to specialized data engines that support time series analysis.

All this data movement, and the operational details required, makes for complicated manual processes that are fraught with the possibility of human error and that increase the time interval between acquiring and curating the data and getting real value from the data. What can simplify this picture, enabling far less operational complexity and affording much lower time to value, is a data platform that can provide support for a range of analytics across diverse data types without messy and time-consuming data movement and transformation.

The MapR Data Platform is an example of such a platform. It was built with its foundation in a new filesystem which is compatible with the HDFS API but offers more. MapR was acquired in August of 2019 by HPE, where the MapR Data Platform serves as an important element of the HPE Container Platform. MapR Data Platform offers an alternative approach to large-scale data ingestion, curation, and analysis that is designed to be more scalable and cost effective than

IDC
ANALYZE THE FUTURE

the classic Hadoop approach or, indeed, any alternative approach. To explore the business value of the MapR Data Platform, IDC undertook a study that involved interviewing a variety of MapR customers to gauge their experiences.

IDC interviewed organizations that have deployed the MapR Data Platform to run business-critical data analytics activities. These MapR customers described using data to support business operations more effectively and efficiently, which they have leveraged along with much improved platform performance to develop new business-focused use cases for data and better serve existing customers. Overall, IDC's research demonstrates that study participants have significantly increased the value they realize with data using the MapR Data Platform, which will result in an estimated five-year ROI of 567% by:

- **Opening new opportunities for data analytics to support their businesses** by enabling far more queries and completing queries in much less time
- **Enabling data analytics teams** with real-time insights, more robust data, and visibility into operational data
- **Increasing revenue** by developing new data-based services and differentiating existing services through embedded analytics and functionality
- **Reducing operational risk** by minimizing platform outages and supporting robust compliance and antifraud activities
- **Providing a more cost-effective data analytics platform** by requiring less hardware, allowing for retirement of legacy solutions, and taking less staff time to operate and support

SITUATION OVERVIEW

At the start of the 21st century, most enterprises concentrated their general business intelligence (BI) analytics in a big relational database into which internal and well-defined data was periodically shipped via a method called ETL (extract, transform, load). This approach was good enough to allow the use of historical data to perform business planning for the future.

But new data sources began to emerge, including machine-generated sensor and log data, data from mobile devices and other internet-enabled things, and customer experience data from ecommerce applications. At first, these enterprises attempted to incorporate this new data into the data warehouse, but this soon proved unworkable. The relational database

management systems (DBMSs) designed to handle the classic data warehouse were ill-suited for this new class of data, and the explosion in data volume made data warehouses massively expensive.

At the same time, a new style of data collection, curation, and analysis emerged. With the support of Google, Doug Cutting and Mike Cafarella introduced a technique called MapReduce for performing data operations on very large data collections by deploying a specialized system on clusters of commodity servers with internal storage. This system was called Hadoop. It drastically cut the cost of collecting and organizing very large amounts of data in comparison with the cost of conventional data warehouse and data integration software.

But HDFS was soon found to have severe limitations. For one thing, it lacked the kind of data management support that a full database management system affords, so using it required a great deal more human effort. Also, the MapReduce technique, which works very well when the problem at hand involves examining every single record, proved very inefficient in dealing with selective subsets of the data. So additional projects were spawned to address these issues, and several vendors emerged to provide commercial support for Hadoop-based environments, each with its own approach to overcoming these basic limitations.

At the same time, other approaches emerged to deal with the issue, including delegating management of different kinds, against data of various types, to specialized engines. The result was that what was to be a simple, cost-effective way of dealing with very large and diverse data collections became complex and labor intensive.

MAPR DATA PLATFORM

One vendor offering an alternative to the cumbersome "pure" Apache Hadoop approach to large-scale data collection, curation, and analysis is MapR. The MapR Data Platform is designed to provide a single platform environment for collecting, curating, and analyzing the full range of data needed to power a digitally transformed enterprise. It supports conventional structured analytics but also data science work, artificial intelligence (AI), and machine learning (ML).

At the heart of MapR's solution is a highly distributed, enterprise-grade set of capabilities for ingesting, managing, and using data. The platform is able to manage bits in the form of files, database tables, and event stream data. The file system is an alternative to HDFS that is called MapR-FS, which is API compatible with HDFS, but can also enable key value, document, or relational data management. It can do this because MapR-FS is an indexed file system that

supports random retrieval and update in place. MapR-FS underlies a facility designed to support a wide variety of data types and access methods, called the MapR Distributed File and Object Store (MapR XD). MapR-FS underlies a multimodel DBMS called the MapR Database. MapR also offers a specialized Event Store for Apache Kafka that mitigates the complexity in dealing with streaming data. The platform also supports easy access to other external data sources. It is designed to be deployed in the datacenter on the premises or in a public single cloud or multicloud configuration and supports IoT data from edge applications as well.

BUSINESS VALUE OF MAPR DATA PLATFORM

IDC's study demonstrates the strong value that interviewed organizations are achieving with the MapR Data Platform. IDC projects that study participants will realize benefits worth an average of $597,200 per 100TB ($21.98 million per organization) in their MapR environments over five years, which would result in a five-year ROI of 567%. Interviewed MapR customers spoke to the platform's significant impact on their data analytics and business operations:

- **Business enablement and increased analytical capabilities:** *"With MapR, we've been able to build our services faster, ship them faster, and then rebuild them as needed much more efficiently. We are providing capabilities that we would otherwise not be able to do."*
- **Analytics scalability and capabilities:** *"We've solved some scaling issues we had with older solutions with MapR. MapR is [also] opening up our ability to do large-scale historical data analysis, enabling our ability to do more real-time data captures of large-scale data sets."*
- **Providing business with single analytics platform:** *"The benefit [of MapR] is that we now have a platform for the entire business to use. We're going to do everything from custom analytics to mining analytics; we're going to use it across the board."*

Firmographics of Study Participants

IDC spoke with nine organizations using the MapR Data Platform about their experiences. Interviews covered topics related to the impact of MapR on data analytics performance, operations, and associated business outcomes. Overall, interviewed organizations had the profile of a large enterprise, with more than 35,000 employees and $14.5 billion in annual revenue on average (6,000 employees and $2.50 billion medians, respectively). Interviews

represented the experiences of organizations from North America, EMEA, and APAC and various industries — financial services (2), information services, IT security, manufacturing, pharmaceutical, software (2), and technology (see Table 1).

TABLE 1 Demographics of Interviewed Organizations

	Average	Median
Number of employees	35,121	6,000
Number of IT staff	4,825	900
Number of business applications	1,675	375
Revenue per year	$14.53 billion	$2.50 billion
Countries	United States (6), France, Germany, and Australia	
Industry	Financial services (2), information services, IT security, manufacturing, pharmaceutical, software (2), and technology	

n=9 Source: IDC, 2019

Use of MapR Data Platform

Interviewed organizations use MapR to run and support substantial data analytics efforts across their organizations. They moved to MapR from a mix of more traditional database solutions and manual process-based analytics efforts. They chose MapR for various reasons but commonly described needing to improve the performance and efficiency of analytics required to fuel business-critical applications that included:

- Fraud detection activities core to a financial services organization's business
- Predictive maintenance algorithms for a manufacturing company
- Embedded machine learning and analytics functionality in products for a software company
- Real-time market data forecasting for a financial technology services company
- Security intelligence functionality for customers of an IT security company

Study participants noted making substantial use of MapR's artificial intelligence and machine learning capabilities in terms of both supporting their own analytics activities and embedding these functionalities in customer-facing services. Several interviewed organizations also reported using MapR's containerization functionality. These MapR customers are using the

Source: IDC, 2018

platform for a data environment of 3,680TB on average (1,200TB median) that directly affects 853 data-focused employees, including data scientist, business intelligence, and analytics engineering teams (see Table 2).

TABLE 2 MapR Data Platform Use by Interviewed Organizations

	Average	Median
Number of datacenters	5	2
Number of business applications	16	4
Number of terabytes	3,680	1,200
Number of core users (data-focused employees)	853	240

n=9 Source: IDC, 2019

Quantifying the Value of MapR Data Platform

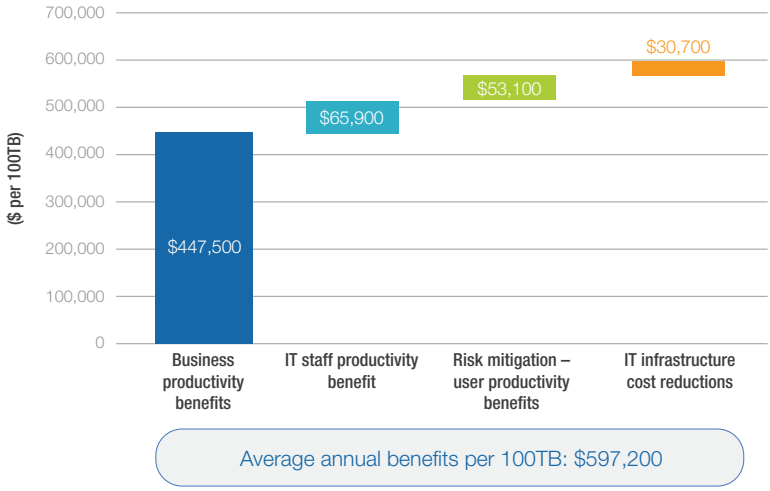
IDC's research demonstrates the significant value that interviewed MapR customers are achieving by increasing the performance and depth of use of analytics at their organizations. By increasing the functionality of their data analytics activities to a significant degree with MapR, they can use data to operate more efficiently and open up new avenues of monetizing data. Based on interviews with these organizations, IDC projects that they will realize benefits worth an annual average of $597,200 per 100TB ($21.98 million per organization) in the following areas (see Figure 1):

- **Business productivity benefits.** MapR handles far greater analytical activity while delivering relevant and timely real-time insights to business teams. Having a highly functional platform enables data analytics teams, allows for delivery of new data-driven services, and provides differentiation for existing products and services with real-time insights. IDC quantifies the value of higher analytics team productivity and revenue at an annual average of $447,500 per 100TB ($16.47 million per organization).
- **IT staff productivity benefits**. MapR provides more robust and timely data in support of development activities and requires less IT time for ongoing maintenance and support. IDC puts the value of higher developer productivity and IT team time savings at an annual average of $65,900 per 100TB ($2.42 million per organization).
- **Risk mitigation — user productivity benefits.** MapR incurs fewer and less impactful outages and provides compliance and antifraud teams with timely access to needed

insights. IDC estimates the value of resultant higher productivity to be worth an average of $53,100 per 100TB per year ($1.95 million per organization).

- **IT infrastructure cost reductions.** MapR allows interviewed organizations to retire legacy data analytics platforms that were often more expensive, in terms of both licensing costs and hardware requirements. IDC calculates that interviewed organizations will realize savings worth $30,700 per 100TB per year ($1.13 million per organization).

FIGURE 1 Average Annual Benefits per 100TB



n=9 Source: IDC, 2019

Improved Analytical Performance

Study participants have fundamentally changed how they leverage data in support of their businesses with the MapR Data Platform. For these organizations, rapid data growth has created both opportunities and challenges. Data creates opportunities when it is used in support of business operations but can also produce friction and missed opportunities when organizations are ill-equipped to handle it. Interviewed organizations reported transforming how they share, consume, and use data with the MapR Data Platform:

- **Turning data into actionable insights with ML functionality:** *"MapR has made it possible for us to convert data into actionable insights. [We can now] run machine learning on the volume of data and actually complete the job."*
- **Driving real-time streaming system with analytics:** *"We would not have been able to build and perform the analytics that we're doing now [with MapR] inside of a real-time streaming system."*

For interviewed organizations, this ability to create value through data and analytics comes back to being able to do more analytics work and deliver analytics insights in a timely way with MapR. Compared with their legacy platforms, MapR both handles far greater volumes of data analytics work and delivers outputs in much less time. On an average organizational basis, IDC's research shows that interviewed organizations run around 20 times more analytical queries more than two times faster (52% less time per query) on the MapR Data Platform, reflecting much increased capabilities (see Table 3).

TABLE 3 Impact of MapR Data Platform on Analytical Queries

	Previous Environment	With MapR Data Platform	Increased Value with MapR	Benefit with MapR (Calculated)	Benefit with MapR (Average by Organization)
Number of queries run per week	6,800	124,200	117,400	18x more	20x more
Time to execute per query (minutes)	195	4	191	98%	52%

n=9 Source: IDC, 2019

Higher Analytics Team Productivity

Interviewed organizations maintain sizable teams whose primary responsibilities include using and transforming data into actionable insights that support their businesses. Increasingly, their business operations depend on these teams; without their support, study participants would struggle to identify trends and business opportunities and may fail to meet customer demand and expectations.

These teams, which include data scientists, business intelligence staff, and analytics engineers, rely on the timely flow of robust and high-quality data and insights to do their jobs. As such, they were hindered by legacy platforms that provided neither the requisite volume nor velocity to transform raw data into actionable insights. Thus the ability of interviewed organizations to use MapR to service many times more queries and significantly reduce the time required per query has enabled these data-driven teams to work more effectively and deliver more value to their organizations (refer back to Table 3). Study participants described the impact of MapR on these teams:

- *"Compared with a traditional architecture, our data scientist team is 50% more productive because they can work on a greater volume of data in a shorter amount of time."*
- *"Our data teams are more productive with MapR because they have data in one place versus having to go hunt it down in multiple silos. We are also able to run large processes in hours rather than days, which is a time saver for modeling teams."*

- *"We are in a self-service BI mode with MapR and it's made a big difference for our 70-person BI team because they used to make requests to IT first and now they do it themselves."*
- *"With MapR, because security is embedded, we give access to data scientists who can work straight on the information. So there's no copying of the data, and the data scientists save time and work more efficiently."*

Table 4 shows the substantial value for interviewed organizations of higher productivity for these analytics teams through use of the MapR Data Platform. Interviewed MapR customers reported significant productivity gains — 33% on average — for analytics teams that include hundreds of team members. For study participants, this higher productivity represents the extent to which MapR has enabled these teams to support and deliver additional value for their business operations.

TABLE 4 Impact of MapR Data Platform on Analytics Teams Productivity

	Per Organization	Per 100TB
Data scientists		
Number of users impacted	80	2
Average productivity gain	38%	38%
Net value of higher productivity per year	$1.86 million	$50,500
Business intelligence staff		
Number of users impacted	129	4
Average productivity gain	36%	36%
Net value of higher productivity per year	$2.88 million	$78,200
Analytics engineers		
Number of users impacted	148	4
Average productivity gain	38%	38%
Net value of higher productivity per year	$3.47 million	$94,300
Total (Data-focused teams)		
Number of users impacted	357	10
Average productivity gain	33%	33%
Net value of higher productivity per year	$8.21 million	$223,000

n=9 Source: IDC, 2019

Better Business Results

Study participants linked improvements to analytics performance and productivity levels of analytics teams with the MapR Data Platform to improved business results. They explained that MapR has allowed them to create differentiation for their businesses with data, in terms of both creating new services and improving the functionality of existing services. Several interviewed organizations spoke to the benefit of instilling their services with machine learning features with MapR, noting that this enhanced functionality helped maintain differentiation for their services:

- *"With MapR, we're able to provide an ML feature in our product because we can write the algorithm and run with MapR, so it's a key feature of our product."*
- *"We're using MapR to embed our services with ML principles and analytics."*

Several interviewed organizations have translated improved analytics with MapR into new types of services. One interviewed organization explained: *"We've enabled a data science consultancy with MapR …. We also have a software-as-a-service solution that we're now offering. These services are 100% because of MapR, which is millions of dollars of additional revenue."*

Overall, study participants have leveraged MapR to win significant amounts of new business and protect their existing businesses. On average, IDC calculates that they will gain $54.11 million of additional revenue per organization per year ($1.47 million per 100TB) (see Table 5).

TABLE 5 Business Productivity Benefits — Increased Revenue

	Per Organization	Per 100TB
Additional revenue per year	$54.11 million	$1.47 million
Recognized revenue per year — IDC model*	$8.12 million	$220,500

n=9 Source: IDC, 2019

**IDC model assumes a 15% operating margin for all additional revenue.*

More Effective Development

Study participants also discussed how they have leveraged more robust analytics with MapR to support data-driven development activities. These developers, like other data-focused employees, rely on the timely flow of robust data to support delivery of new applications and features. MapR has supported their efforts by helping interviewed organizations complete more analytics activities (e.g., queries) while speeding up delivery of actionable insights. This provides development teams with higher-quality data and insights to use and fewer delays.

For some organizations, the impact is considerable across large development teams. One interviewed organization explained: *"We have a couple of hundred developers, and they are more productive by about 25% on MapR Data Platform."* As shown in Table 6, the increased efficacy of impacted development teams is reflected in terms of higher productivity (30% on average), delivery of more new applications (14%) and features (8%), and faster development life cycles (8% for new applications and 38% for new features).

TABLE 6 Application Development KPIs

	Previous Environment	With MapR Data Platform	Increased Value with MapR	Benefit with MapR (%)
Number of new applications/features developed				
Number of new applications per year	1.3	1.5	0.2	14
Number of new features per year	75.1	81.3	6.3	8
Development life cycle (weeks)				
New applications	31.7	25	6.3	8
New features	5.1	3.1	2	38
Productivity of impacted developers				
Developer productivity per organization (FTEs)	44.9	58.5	13.6	30
Value of impacted developers per organization per year ($M)	4.49	5.85	1.36	30

n=9 Source: IDC, 2019

Reduced Operational Risk

Study participants reported reducing data-related risk with the MapR Data Platform. As previously noted, rapid growth in data volume and use has brought potential risks and rewards. Data-related risks can be both operational and financial in nature and are related to areas that include application performance, data loss, and regulatory compliance and auditing activities. Even from a more general perspective, one study participant noted its perception that the MapR Data Platform provided a more secure foundation for analytics than other platforms: *"Security was always an afterthought with Big Data Analytics, but security is a core part of the platform with MapR."*

Interviewed organizations also described specifically how MapR has helped them handle risk more efficiently and reduce their overall risk exposure:

- **Reduced impact of unplanned downtime and data loss.** With a more reliable analytics platform and functionality such as improved failover, study participants reported losing an average of 85% less staff time due to unplanned outages with MapR (see Table 7). Another organization commented on eliminating data loss: *"We've never lost any data with MapR."*
- **More efficient and effective regulatory compliance and antifraud activities**. *"All of our fraud identification is now done through ML and AI on the MapR Data Platform …. We are avoiding financial penalties from regulators."* On average, these organizations have improved the productivity levels of their regulatory compliance and antifraud teams by 18% with MapR.

TABLE 7 Impact of MapR Data Platform on Risk: Unplanned Downtime and Regulatory Compliance/Fraud

	Previous Environment	With MapR Data Platform	Increased Value with MapR	Benefit with MapR (%)
Unplanned downtime				
Lost user productivity per year (FTEs)	18.6	2.8	15.8	85
Value of lost productive time per year per organization ($M)	1.30	195,800	1.11	85
Productivity of impacted developers				
Productivity of impacted regulatory compliance and antifraud teams (FTEs per organization)	56.3	68.4	12.1	18
Value of impacted regulatory compliance and antifraud teams per organization per year ($M)	3.94	4.79	0.85	18

n=9 Source: IDC, 2019

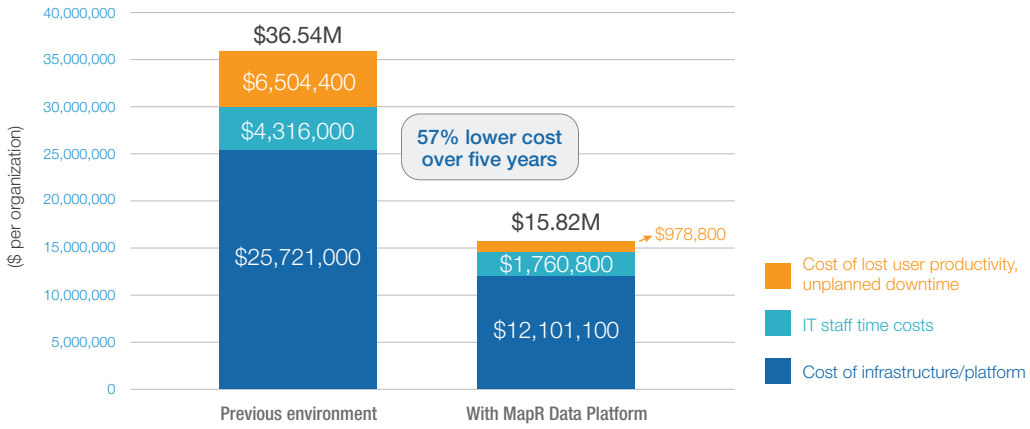
Lower Cost of Operations

In addition to the impact on business operations of improved analytics capabilities and performance, study participants reported benefiting from using MapR in terms of the cost of running data analytics activities. On average, IDC calculates that this sample of MapR customers will lower their costs over five years by 57% in terms of hardware and platform costs, IT staff time costs, and the cost of lost user productivity due to unplanned downtime (see Figure 2).

Interviewed organizations reported that MapR will cost 53% less over five years in terms of platform (including licensing) and infrastructure-related costs (cost of hardware, power, facilities, etc.). Several study participants attributed significant cost savings to moving off of legacy data analytics platforms that were too costly, especially as they had to scale to meet increasing organizational demand for analytics. Other interviewees noted that they can run equivalent analytics workloads on MapR with significantly fewer servers, providing further capital and operational cost savings.

Meanwhile, interviewed MapR customers also need less staff time to manage and support the platform on an ongoing basis. On average, they attributed a 47% efficiency to MapR, thereby helping their organizations extend their analytics capabilities without bringing in additional staff. One interviewed organization discussed the basis for significant IT staff efficiencies: *"MapR is very integrated into our security system, which makes administration very efficient. We'd probably need three times more staff without [MapR]."* Several organizations also noted that they have leveraged containerization functionality with MapR to make managing their environments more efficient.

FIGURE 2 Five-Year Cost of Operations



Source: IDC, 2019

ROI Analysis

IDC based its ROI analysis on interviews with organizations that are using the MapR Data Platform as an analytics platform to support their businesses. Based on these interviews, IDC has calculated the benefits and costs to these organizations of deploying and using the MapR Data Platform compared with their previous environments (see Appendix for details of IDC's business value methodology).

IDC's analysis of the benefits and costs related to use of the MapR Data Platform by interviewed organizations is presented in Table 8. On average, IDC projects that these MapR customers will achieve five-year discounted benefits worth $76.71 million per organization ($2.08 million per 100TB) in higher employee productivity, increased revenue, and reduced costs. This compares with total five-year discounted investment costs that include MapR licensing and hardware costs, staff time costs for deploying and managing MapR, and third-party expenses of $11.49 million per organization ($0.31 million per 100TB). These levels of benefits and investment costs

would result in an average five-year ROI of 567%, with breakeven on investment occurring in 11 months on average.

TABLE 8 ROI Analysis

	Five-Year Average per Organization	Five-Year Average per 100TB
Benefit (discounted)	$76.71 million	$2.08 million
Investment (discounted)	$11.49 million	$0.31 million
Net present value (NPV)	$65.22 million	$1.77 million
Return on investment (ROI) (%)	567%	567%
Payback period	11 months	11 months
Discount rate (%)	12%	12%

n=9 Source: IDC, 2019

CHALLENGES/OPPORTUNITIES

MapR's heritage is in Apache Hadoop, Spark and its associated projects. This is a strength because it means that developers with skills in these technologies can leverage those skills in a highly productive way working with the MapR Data Platform. It is a problem also, however, because of the market perception that Hadoop may be on the way out. All available evidence suggests, however, that Hadoop is not on the way out, but its use is morphing as new, complementary technologies for machine learning and AI arise that can complete the cost, usability, and performance picture for data analytics at scale. MapR clearly offers support for such technologies beyond Hadoop and has optimized them in a platform that is enterprise ready and, as this study indicates, shows distinct business benefits as well as advantages over other approaches that involve distributed scale-out technologies. This is chiefly realized through the MapR file system and the various database management facilities that MapR provides in conjunction with it, as well as a full suite of management tools.

In addition to the challenge of dealing with the market perceptions of Hadoop-related technologies, MapR is likely to be challenged by new and competing technologies in this highly dynamic environment. These technologies also expose, in the minds of C-level executives, data needs that they had not considered. MapR, as its technologies continue to evolve, should be able to leverage this new awareness as an opportunity to demonstrate the value of the MapR Data Platform and expand its usage footprint in the enterprise.

CONCLUSION

There is a wide variety of distributed, scale-out data management technologies out there today, but when one examines them, one finds that in most cases while the scaling and performance may be there, the manageability, and therefore practical utility in an enterprise setting, is lacking. As the results of this study demonstrate, interviewed organizations have found that MapR's offering of proven elements of the open source array of scalable data management technologies, combined with enhanced database capabilities and the kinds of functionality required by major enterprises in areas such as manageability, data security, and continuous availability, have enabled them to capture strong value through better data analytics performance, opening up new revenue streams, higher reliability, lower operational costs, and reduced business risk.

The value of MapR Data Platform is clearly seen in the resulting strong ROI and relatively short payback period for interviewed organizations. The architectural design center for the platform, including its ability to provide an enterprise-grade foundation for a broad range of capability to manage a variety of data types enabling the use of a variety of compute engines and the ability to evolve with an changing ecosystem have resulted in very positive results for these MapR customers. The subjects of this study have confirmed that MapR's platform has delivered strong price/performance capabilities and fully supports the critical role of managing mission-critical data for their business operations.

APPENDIX

Methodology

IDC used the following three-step method for conducting the ROI and business value analysis informing this study's results and conclusions:

- **Gathered quantitative benefit information during the interviews using a before-and-after assessment for interviewed organizations of using the MapR Data Platform compared with their previous data analytics platforms and/or environments.** In this study, the benefits of using MapR Data Platform included improved analytics performance, higher productivity for analytics teams, increased revenue, higher productivity for other users, more efficient regulatory and auditing activities, IT staff efficiencies, and data platform –related cost reductions.

- **Created a complete investment (five-year total cost analysis) profile based on the interviews.** Investments go beyond the initial and annual costs of deploying and using the MapR Data Platform and can include additional costs related to migrations, planning, consulting, and staff or user training.
- **Calculated the ROI and payback period.** IDC conducted a depreciated cash flow analysis of the benefits and investments for the organizations' use of the MapR Data Platform over five years. ROI is the ratio of the net present value (NPV) and the discounted investment. The payback period is the point at which cumulative benefits equal the initial investment.

IDC's standard ROI methodology was utilized for this white paper. This methodology is based on gathering data from current users of the MapR Data Platform. Based on interviews with nine organizations, IDC performed a three-step process to calculate the ROI and payback period:

- Measure the benefits from use of MapR in terms of higher productivity for analytics teams, increased revenue, higher productivity for other users, more efficient regulatory and auditing activities, IT staff efficiencies, and data platform –related cost reductions.
- Ascertain the investment made in deploying the MapR Data Platform and associated migration, training, and support costs.
- Project the costs and savings over a five-year period and calculate the ROI and payback for use of the MapR Data Platform.

IDC bases the payback period and ROI calculations on assumptions that are summarized as follows:

- Time values are multiplied by burdened salary (salary + 28% for benefits and overhead) to quantify efficiency and productivity savings. IDC assumes a fully burdened salary of $100,000 per year for IT staff, including developers, and $70,000 for other employees, with an assumption of 1,880 hours worked per year.
- Downtime values are a product of the number of hours of downtime multiplied by the number of users affected.
- The impact of unplanned downtime is quantified in terms of impaired end-user productivity and lost revenue.
- Lost productivity is a product of downtime multiplied by burdened salary.
- The net present value of the five-year savings is calculated by subtracting the amount

that would have been realized by investing the original sum in an instrument yielding a 12% return to allow for the missed opportunity cost. This accounts for both the assumed cost of money and the assumed rate of return.

- Because every hour of downtime does not equate to a lost hour of productivity or revenue generation, IDC attributes only a fraction of the result to savings. As part of our assessment, we asked each company what fraction of downtime hours to use in calculating productivity savings and the reduction in lost revenue. IDC then taxes the revenue at that rate.
- Further, because IT solutions require a deployment period, the full benefits of the solution are not available during deployment. To capture this reality, IDC prorates the benefits on a monthly basis and then subtracts the deployment time from the first-year savings.

Note: All numbers in this document may not be exact due to rounding.

IDC Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC
ANALYZE THE FUTURE