

# The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

by Noel Yuhanna and Mike Gaultieri  
April 22, 2016

## Why Read This Report

“We want to do Hadoop in the cloud.” That’s what we increasingly hear from enterprises that have successfully run on-premises Hadoop proof-of-concepts. CIOs charge their enterprise architecture (EA) pros with understanding the strengths and weaknesses of Hadoop in the cloud and to make the case for it versus on-premises deployments. Forrester’s 37-criteria evaluation of eight leading big data Hadoop cloud solution vendors — Altiscale, Amazon Web Services (AWS), Google, IBM, Microsoft, Oracle, Qubole, and Rackspace — will help EA pros understand the available solutions and recommend the best one for their organization.

## Key Takeaways

### **Eight Viable Hadoop Cloud Solutions**

Among the eight vendors evaluated, we found four Leaders -- AWS, Google, IBM, and Microsoft -- and four Strong Performers -- Altiscale, Oracle, Qubole, and Rackspace. All offer competitive solutions that provide enterprise-scale Hadoop in the cloud.

### **EA Pros Look At Speed, Automation, And Cost**

This market is growing largely because EA pros see the cloud as a strategic platform to support their big data initiatives. When selecting a solution, enterprises should look for speed to provision, the level of automation, and cost as the key factors.

### **Broader Cloud Services Can Be A Deal-Maker**

While all of the vendors offer compelling value and features, many Hadoop cloud solutions also provide a broader range of cloud services that enterprises can use to build analytics into their applications.

# The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016

## Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

by [Noel Yuhanna](#) and [Mike Gualtieri](#)  
with [Gene Leganza](#) and Shreyas Warriar  
April 22, 2016

---

### Table Of Contents

- 2 **The Hadoop Cloud Market Gains Momentum Across Industries**
- 4 **Big Data Hadoop Cloud Solutions Evaluation Overview**
  - Evaluation Criteria: Current Offering, Strategy, And Market Presence
  - Hadoop Cloud Forrester Wave Evaluation Assessed The Capabilities Of Eight Vendor Offers
- 6 **Larger Providers Have An Edge With Broader Range Of Functionality**
- 9 **Vendor Profiles**
  - Leaders
  - Strong Performers
- 12 **Supplemental Material**

### Notes & Resources

Forrester conducted product evaluations and interviewed eight vendor companies: Altiscale, Amazon Web Services, Google, IBM, Microsoft, Oracle, Qubole, and Rackspace.

### Related Research Documents

- [Big Data Fabric Drives Innovation And Growth](#)
- [TechRadar™: Big Data, Q1 2016](#)
- [The Forrester Wave™: Big Data Hadoop Distributions, Q1 2016](#)

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

## The Hadoop Cloud Market Gains Momentum Across Industries

Big data initiatives are no longer just the domain of early adopters — robust data management capabilities are critical to every organization's success. And enterprises of all types and sizes are increasingly embracing the cloud as an important catapult for their business technology (BT) agenda.<sup>1</sup> Hadoop in the cloud offers a new opportunity to deliver next-generation insights quickly and economically — insights that were previously limited to enterprises with the deep pockets to fund an on-premises big data sandbox. Most enterprises see the opportunity in big data and are not standing idly by: 40% of technology decision-makers in our 2015 survey tell us they have added Hadoop to their mix of technologies supporting data warehousing and business intelligence (BI), with an additional 20% reporting they will do so by the end of 2016.<sup>2</sup> About half of our respondents are increasing their cloud deployments: 37% told us they plan to increase their investment in big data cloud by 5% to 10%, while another 14% plan to increase it by more than 10% (see Figure 1).

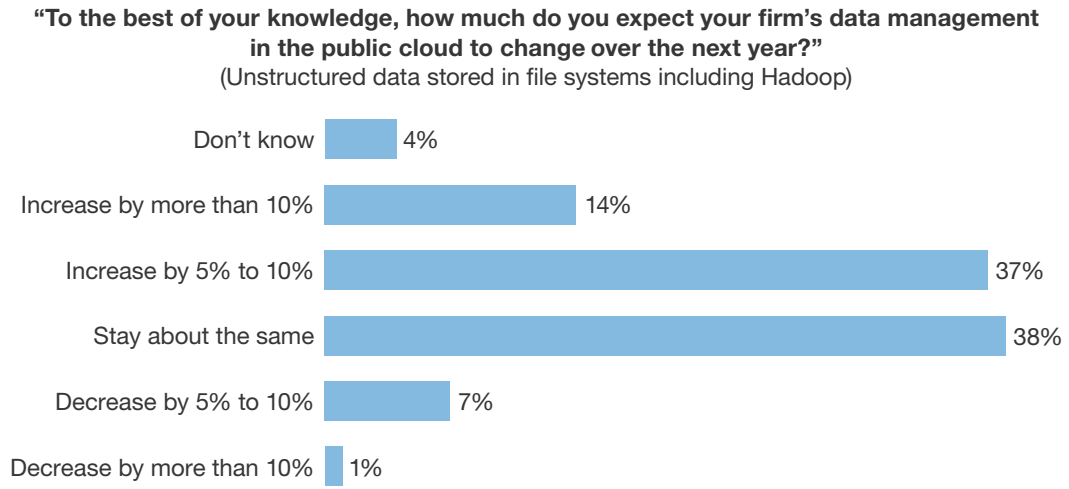
Hadoop deployments in the cloud offer five key benefits. They:

- › **Economize the analysis of more data using cheaper infrastructure.** Hadoop is an economical solution for both structured and unstructured data because it is based on open source and is designed to handle various workloads. Organizations deploying in the cloud report saving 20% to 60% over on-premises infrastructure cost, since most over-provision their servers and storage and then need to manage these environments.
- › **Enable faster advanced analytics.** Traditional BI is mostly about producing reports and populating dashboards with historical business performance analytics. That's enormously valuable, but to get the most out of big data, enterprises must also use advanced analytics such as machine learning to build predictive models and text analytics to analyze unstructured data. Forty percent of decision-makers told us that they plan to increase their investments by 5% to 10%, while another 34% plan to invest about the same (see Figure 2).
- › **Scale to unlimited resources in minutes.** Analyzing big data often means running short-lived workloads that process huge amounts of data for doing trend analysis or to support a marketing campaign. With Hadoop cloud, enterprise architects have the flexibility to scale to thousands of nodes rapidly to run workloads in minutes that would take hours or days on a few nodes. The flexibility, agility, and elasticity of cloud compute and storage makes it attractive for any organization to do any type of analysis.
- › **Provide a self-service analytics platform.** A past criticism of Hadoop was that only developers could use it to analyze data by coding custom MapReduce jobs. With today's solutions, BI professionals, data scientists, and less specialized business users can use Hadoop cloud through extract, transform, load (ETL), BI, and advanced analytics tools designed for a broad set of skill levels.
- › **Simplify big data deployment.** Deploying a Hadoop cluster can take months, including cross-training staff, procuring hardware, and testing and loading test beds. Big data cloud helps accelerate deployment through automation and simplification processes, provisioning a cluster

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**  
 Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In  
 The Cloud

in minutes. In addition, with storage and compute resource separation, organizations only pay for the resources they use, eliminating unwanted costs and enabling a stronger focus on business analytics and insights.

**FIGURE 1** Increased Investment Of Unstructured Data Stored In The Public Cloud

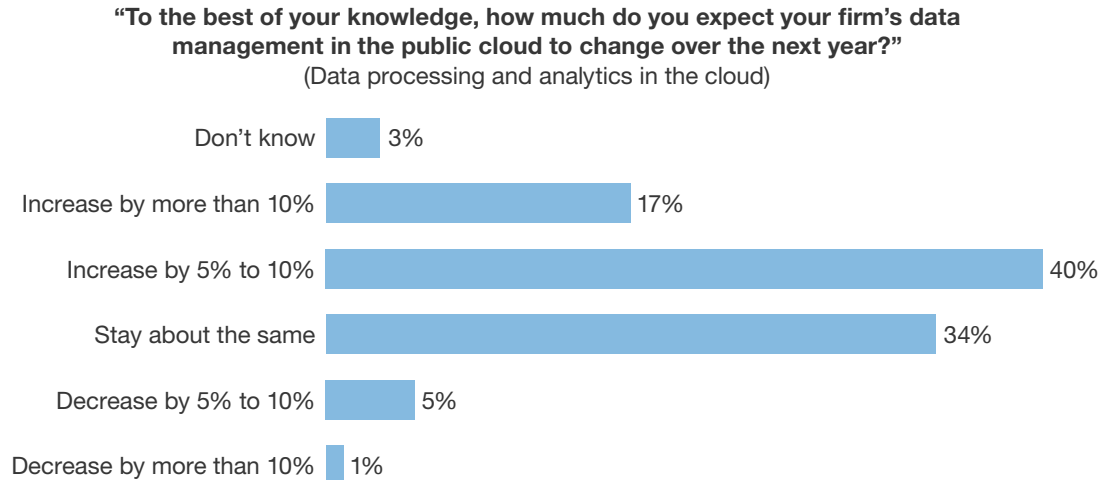


Base: 814 global data and analytics technology decision-makers who use or have purchase influence over data and analytics in the public cloud

Source: Forrester’s Global Business Technographics® Data And Analytics Survey, 2015

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

**FIGURE 2** Increased Investment Of Data Processing and Analytics In The Public Cloud

Base: 814 global data and analytics technology decision-makers who use or have purchase influence over data and analytics in the public cloud

Source: Forrester’s Global Business Technographics® Data And Analytics Survey, 2015

## Big Data Hadoop Cloud Solutions Evaluation Overview

To assess the state of the market and see how the vendors stack up against each other, Forrester evaluated the strengths and weaknesses of eight of the top commercial Hadoop cloud vendors: Altiscale, Amazon Web Services, Google, IBM, Microsoft, Oracle, Qubole, and Rackspace.

### Evaluation Criteria: Current Offering, Strategy, And Market Presence

After examining past research, user requirements, and vendor interviews, we developed a comprehensive set of evaluation criteria. We evaluated vendors against 37 criteria, which we grouped into three high-level buckets:

- › **Current offering.** We evaluated each product’s configuration options, scalability, performance, high availability, disaster recovery, administration, security, core data capabilities, development tools, and other features to establish the capabilities of the vendor’s current offering.
- › **Strategy.** We reviewed each vendor’s strategy to assess its ability to compete and grow in the commercial Hadoop cloud market. Key criteria include Forrester’s level of confidence in the vendor’s ability to execute on its stated strategy and to support current and future customers. Forrester also reviewed each vendor’s product road map to assess how it will affect the vendor’s competitive position compared with the others in this evaluation.

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

- › **Market presence.** To determine each vendor's market presence, we evaluated overall Hadoop revenue, the number of paying customers, the global distribution of paying customers, market awareness of the vendor's product, and partnerships with other technology and services firms.

**Hadoop Cloud Forrester Wave Evaluation Assessed The Capabilities Of Eight Vendor Offers**

Each of the eight vendors (Altiscale, Amazon Web Services, Google, IBM, Microsoft, Oracle, Qubole, and Rackspace) we included in this evaluation has (see Figure 3):

- › **A comprehensive Hadoop cloud offering.** The evaluated vendors provide a Hadoop cloud service that delivers a pay-as-you-go service offering that integrates all the necessary hardware, software, tools, and technologies. Their offerings are based on the Apache Hadoop open source codebase and include value-added features and add-on tools appropriate for use by enterprises, such as security, data management, and Hadoop management tools. The providers also needed to provide complete high availability, disaster recovery, provisioning, and administration services directly and/or through partners. Most offer Apache Spark as well.<sup>3</sup>
- › **A cloud solution.** We only included Hadoop cloud solutions that are not embedded into any particular application, BI, predictive analytics, ETL, or middleware stacks. They must be standalone Hadoop cloud solutions that can support various Hadoop use cases. Enterprises must be able to configure, add, remove, drop, and reconfigure compute and data nodes to support their dynamic Hadoop requirements.
- › **A referenceable install base.** Included vendors had to have 10 or more unique enterprise customers using the Hadoop cloud offering that span more than one major geographical region.
- › **Client inquiries or noteworthy technologies that put the vendor on Forrester's radar.** Forrester monitors which vendors and products clients discuss through inquiries. We included vendors that were mentioned often. Alternatively, the vendor may, in Forrester's judgment, warrant inclusion or exclusion because of technology trends and market presence.
- › **A publicly available offering.** The participating vendors must have actively marketed a Hadoop cloud service as of December 31, 2015.

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

**FIGURE 3** Evaluated Vendors: Product Information

Vendor	Product name
Altiscale	Altiscale Data Cloud 4.0
Amazon Web Services	Amazon Elastic MapReduce (EMR)
Google	Cloud Dataproc
IBM	IBM BigInsights on Cloud 4.1.02
Microsoft	Azure HDInsight
Oracle	Oracle Big Data Cloud Services
Qubole	Qubole Data Service (QDS)
Rackspace	Rackspace Private Cloud

**Vendor inclusion criteria**

**Comprehensive Hadoop cloud offering.** The vendors included in this evaluation provide a Hadoop cloud service that delivers a pay-as-you-go service offering that integrates all the necessary hardware, software, tools, and technologies. Their offerings are based on the Apache Hadoop open source codebase and include value-added features and add-on tools appropriate for use by enterprises, such as security, data management, and Hadoop management tools. The providers must also provide complete high availability, disaster recovery, provisioning, and administration services directly and/or through partners.

**Cloud solution.** We only included Hadoop cloud solutions that are not embedded into any particular application, business intelligence (BI), predictive analytics, extract, transform, load (ETL), or middleware stacks. They must be standalone Hadoop Cloud solutions that can support various Hadoop use cases. Enterprises should be able to configure, add, remove, drop, and reconfigure compute and data nodes to support their dynamic Hadoop requirements.

**Referenceable install base.** Included vendors had to have 10 or more unique enterprise customers using the Hadoop cloud offering that span more than one major geographical region.

**Publicly available.** The participating vendors must have actively marketed a Hadoop cloud service as of December 31, 2015.

## Larger Providers Have An Edge With Broader Range Of Functionality

Forrester's evaluation of big data Hadoop cloud providers uncovered a market with four Leaders and four Strong Performers (see Figure 4):

- › **Amazon Web Services, Google, IBM, and Microsoft are Leaders.** The larger Hadoop cloud providers have the most comprehensive, scalable, and integrated platforms. Each of the Leaders has a sweet spot strong enough to vigorously compete in the Hadoop cloud market (read the vendor profiles, below). AWS's Elastic MapReduce (EMR) solution is often shortlisted by many

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

enterprises because of its global and scalable data centers, engineering prowess, strong EMR offering, and integration with the portfolio of Amazon services. While Google is not widely known for Hadoop in the cloud, its strategic commitment and competitive pricing appeals to many customers. IBM's BigInsights is more than just Hadoop-as-a-service; it's an end-to-end analytical cloud platform. Microsoft's cloud-first strategy is paying off as the company ramps up its Azure HDInsight, Data Lake Analytics, and Data Lake Store products as managed services to support faster time-to-value implementation.

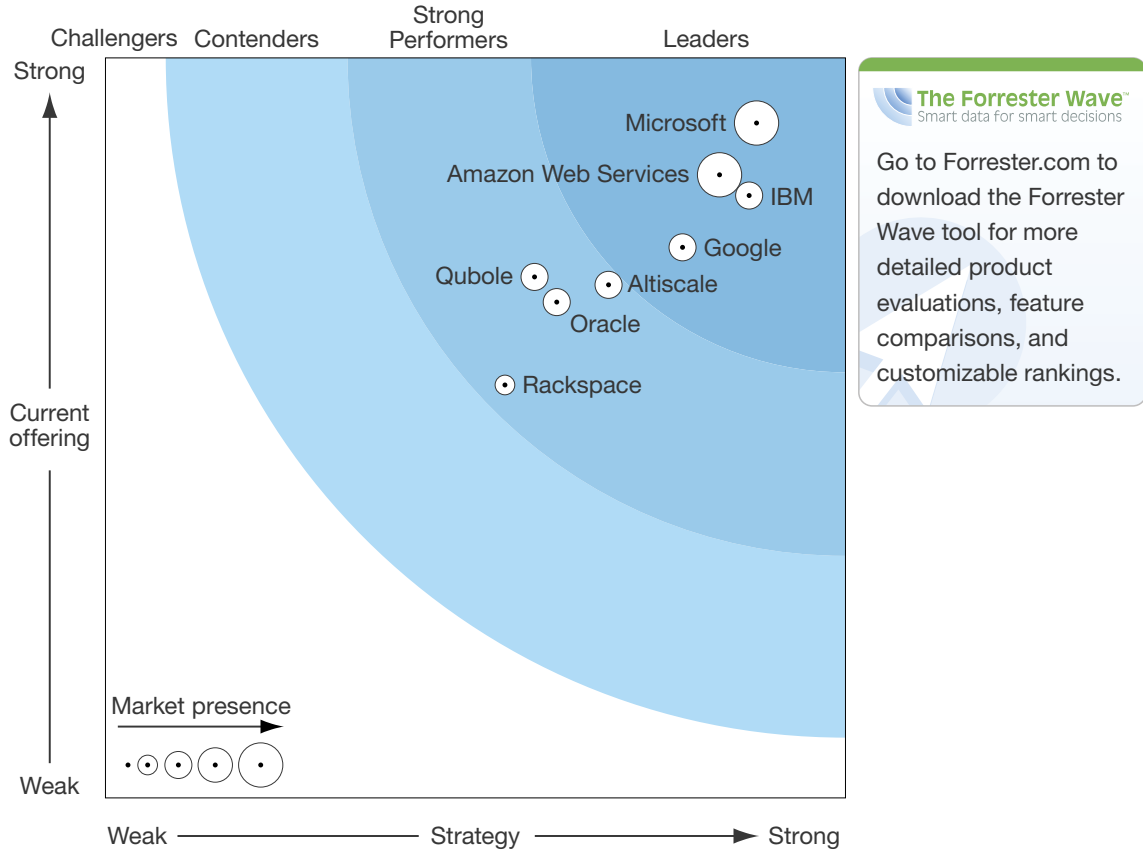
- › **Altiscale, Oracle, Qubole, and Rackspace are Strong Performers.** A Strong Performer among a Forrester Wave dominated by Leaders can still be a strong choice for an enterprise, especially if price/performance, broader big-data-as-a-service, integration-as-a-service, and appliance in the cloud are important considerations. Altiscale Data Cloud's fully managed big data platform delivers rapid access to Apache Hadoop and Apache Spark. Oracle is aggressively ramping up its Hadoop cloud offering by extending its portfolio of big data appliances, software, and solutions into the public cloud. Qubole has done well to go head-on against the larger vendors overall. Its high degree of automation, which focuses on simplifying the provisioning, management, and scaling of big data analytics workloads, is a big appeal to large customers. Rackspace's broad range of cloud offering entices customers who look for flexibility ranging from a multitenant cloud to fully dedicated Hadoop clusters.



**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

**FIGURE 4** Forrester Wave™: Big Data Hadoop Cloud, Q1 2016



**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

**FIGURE 4** Forrester Wave™: Big Data Hadoop Cloud, Q1 2016 (Cont.)

	Forrester's Weighting	AltiScale	Amazon Web Services	Google	IBM	Microsoft	Oracle	Qubole	Rackspace
Current offering	50%	3.44	4.21	3.72	4.07	4.56	3.35	3.52	2.79
Solution configuration	5%	4.40	3.80	3.20	4.40	4.40	4.40	3.20	3.20
Architecture	30%	3.60	4.20	3.30	3.80	4.20	2.90	3.80	2.90
Provisioning and administration	30%	3.50	5.00	5.00	3.50	5.00	4.00	5.00	3.50
Data security	15%	3.00	4.00	3.00	4.50	4.00	3.00	2.00	0.75
Data management	5%	3.80	4.20	3.40	4.60	4.60	4.20	3.40	3.00
Development	5%	3.00	3.00	3.00	5.00	5.00	2.00	3.00	3.00
Cloud platform integration	10%	3.00	3.00	3.00	5.00	5.00	3.00	1.00	3.00
Strategy	50%	3.40	4.15	3.90	4.35	4.40	3.05	2.90	2.70
Subscription	10%	5.00	4.00	5.00	3.00	5.00	3.00	4.00	4.00
Ability to execute	25%	3.00	5.00	4.00	5.00	4.00	4.00	3.00	3.00
Road map	35%	3.00	4.00	4.00	4.00	4.00	3.00	3.00	3.00
Professional services	10%	3.00	3.00	3.00	5.00	5.00	3.00	1.00	1.00
Support	10%	3.00	5.00	4.00	4.00	5.00	3.00	3.00	3.00
Fixes	10%	5.00	3.00	3.00	5.00	5.00	1.00	3.00	1.00
Market presence	0%	2.32	5.00	2.96	3.00	4.44	3.00	2.04	1.92
Product revenue	40%	2.00	5.00	3.00	2.00	4.00	3.00	2.00	2.00
Customer base	40%	2.80	5.00	3.40	3.00	4.60	3.00	2.60	1.80
Partnerships	20%	2.00	5.00	2.00	5.00	5.00	3.00	1.00	2.00

All scores are based on a scale of 0 (weak) to 5 (strong).

## Vendor Profiles

This evaluation of the big data Hadoop cloud market is intended to be a starting point only. We encourage clients to view detailed product evaluations and adapt criteria weightings to fit their individual needs through the Forrester Wave Excel-based vendor comparison tool. Clients can also schedule an inquiry with the authors of this report to have a conversation about the market trends and specific vendor products. In addition, be sure to read our sister report on big data Hadoop solutions, which compares the evaluation of Hadoop distributions from Cloudera, Hortonworks, IBM, MapR, and Pivotal.<sup>4</sup>

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

**Leaders**

- › **Microsoft provides a data lake with all the trimmings.** Microsoft's cloud Hadoop offering includes Azure Marketplace, which runs Hortonworks Data Platform (HDP), Cloudera Enterprise, and MapR in a virtual machine, and Azure Data Lake, which includes Azure HDInsight, Data Lake Analytics, and Data Lake Store as managed services. Azure HDInsight uses the HDP Hadoop distribution, which is designed for the Microsoft Azure cloud. It also includes Spark, HBase, and Storm besides Apache Hadoop, and enterprise architects can use C#, Java, and .NET to create, configure, submit, and monitor Hadoop jobs, in addition to a fully user-focused user interface. PolyBase allows SQL Server customers to execute queries against data stored in Hadoop. Microsoft has significant engineering efforts on other open source Hadoop subprojects. Its strong presence in the database, data warehouse, cloud, spreadsheet, collaboration, BI, OLAP, and development tools markets delivers a growing Hadoop stack to Microsoft customers.
- › **Amazon Web Services reigns among the Leaders for a reason.** Offering Hadoop in the cloud is among the many AWS cloud computing firsts. Amazon EMR is available across 12 regions worldwide. AWS offers versions of Hadoop, Spark, and Presto that can work off data stored in Amazon S3. With Amazon EMR, you can leverage multiple data stores, including Amazon S3 using the EMR File System (EMRFS), the Hadoop Distributed File System on local storage or Amazon EBS, Amazon RDS, Amazon DynamoDB, and Amazon Kinesis. Amazon EMR automatically configures Amazon EC2 firewall settings that control network access to instances, and customers can launch clusters in a virtual private cloud. Amazon EMR compliance certifications include SOC 1/2/3 and PCI-DSS Level 1, and it is HIPAA-eligible under the AWS BAA.<sup>5</sup> AWS offers one of the broader ranges of partners for Hadoop data access/query, modeling and development, data integration, cluster management, and business applications.
- › **Google spins up fast and bills by the minute.** Cloud Dataproc is Google's managed Hadoop (and Spark) service. With Cloud Dataproc, the company recognizes that many cloud customers want to use Spark and Hadoop ecosystems on the Google Cloud Platform alongside other tools such as Google BigQuery and Google Bigtable. That's a wise step forward — Cloud Dataproc will make enterprise customers more attracted to the Google Cloud, and after they benefit from Google's extremely competitive pricing, they may just try the other Google Cloud services.
- › **IBM differentiates BigInsights with end-to-end advanced analytics.** IBM BigInsights runs atop IBM's SoftLayer cloud infrastructure and can be deployed on any of 17 global data centers. IBM's client relationships require it to be flexible in how it offers Hadoop in the cloud and offer highly customized configurations. IBM is making significant investments in Spark, offering data science notebooks that run with the platform. Enterprises using IBM's data management stack will find BigInsights a natural extension to their existing data platform. The company has also launched an ambitious open source project, Apache SystemML, from its newly minted Spark Technology Center. IBM's customers value the maturity and depth of its Hadoop extensions, such as BigSQL,

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

which is one of the fastest and most SQL-compliant of all the SQL-for-Hadoop engines. In addition, BigQuality, BigIntegrate, and IBM InfoSphere Big Match provide a mature and feature-rich set of tools that run natively with YARN to handle the toughest Hadoop use cases.

### Strong Performers

- › **Oracle's Big Data Cloud is starting to gain momentum.** Although Oracle has only recently become aggressive in the cloud platform market, it offers a broad portfolio of products and services to support big data analytics. Oracle Big Data Cloud Services (BDCS) is composed of Cloudera CDH, Apache Spark, Oracle Big Data Connectors, Oracle Data Integrator, Oracle Big Data Spatial and Graph, Oracle R distribution, and Cloud Service Integration software. It focuses on automation, ease-of-use through single command patching, upgrading, and comprehensive dashboards. In addition, Oracle's Big Data SQL provides data federation with Hadoop and non-Hadoop sources such as databases and data warehouses. BDCS includes support for automated on-demand provisioning and integrated security with LDAP and Kerberos. Oracle's key differentiator lies in the ability to provide an end-to-end big data platform that can run on-premises and in the cloud with common tooling, integration, security, and delivery.
- › **Qubole is cloud-agnostic.** Qubole, launched in 2013, offers a portable, open source Hadoop service that runs on AWS, Google Cloud Platform, and Microsoft Azure. Qubole offers the ability to automate, simplify, and integrate data from sources and deliver advanced analytics. Once technology management sets policies, any number of data analysts can be set free to collaboratively "click to query" Hive, Spark, Presto, and many others in a growing list of data processing engines. Qubole will elastically scale Hadoop resources up or down according to what's needed to run the job, thus improving the cost efficiency of the clusters. It focuses on simplifying the provisioning, management, scale, and security of big data analytics workloads, using data stored in AWS, Google Cloud Storage, or Microsoft Azure infrastructure.
- › **Altiscale offers white-glove service for Hadoop at a self-service price.** Altiscale, founded in 2012, delivers a Hadoop- and Spark-based platform in the cloud that also includes operational services. The vendor handles automation, scaling, security, availability, and performance tuning. It manages the infrastructure and operational concerns on behalf of customers so that customers can focus on what really matters: insights. Altiscale experts monitor everything, including customer jobs. The company started as a service designed and run by Hadoop experts for IT customers, but it has recently branched out to offer tools like Alation to provide self-service BI to business analysts and other professionals. The Altiscale Data Cloud features Apache Hadoop, Apache Spark, Apache Hive, Apache Pig, and support for third-party applications such as Alation, AtScale, Datameer, H2O, and Zaloni.

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

- › **Rackspace is ramping up its Hadoop services.** Rackspace offers enterprises a broad range of cloud services, including fully managed Hadoop-as-a-service. Its offering gives customers choices — whether that means a flexible multitenant cloud offering or a fully dedicated and isolated Hadoop cluster. Rackspace provides support for the Hortonworks Data Platform stack, including Apache Spark and HBase. In addition, it supports Ambari, Hue, Tachyon, and Zeppelin in predefined stacks, including an optimized Spark stack with Tachyon and Zeppelin as well as a custom stack option where customers can build à la carte solutions based on their workload. Although Rackspace still has some way to go before becoming a threat to Amazon, IBM, and Microsoft in the Hadoop cloud space, its commitment to expanding key partnerships to support a fully managed big data platform is likely to interest customers, especially if the services come with an attractive price tag.

## Engage With An Analyst

Gain greater confidence in your decisions by working with Forrester thought leaders to apply our research to your specific business and technology initiatives.

### Analyst Inquiry

Ask a question related to our research; a Forrester analyst will help you put it into practice and take the next step. Schedule a 30-minute phone session with the analyst or opt for a response via email.

[Learn more about inquiry](#), including tips for getting the most out of your discussion.

### Analyst Advisory

Put research into practice with in-depth analysis of your specific business and technology challenges. Engagements include custom advisory calls, strategy days, workshops, speeches, and webinars.

[Learn about interactive advisory sessions](#) and how we can support your initiatives.

## Supplemental Material

### Survey Methodology

Forrester's Global Business Technographics® Data And Analytics Survey, 2015 was fielded online in January through March 2015 of 3,005 business and technology decision-makers located in Australia, Brazil, Canada, China, France, Germany, India, New Zealand, the UK, and the US from companies with 100 or more employees.

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

Forrester's Business Technographics provides demand-side insight into the priorities, investments, and customer journeys of business and technology decision-makers and the workforce across the globe. Forrester collects data insights from qualified respondents in 10 countries spanning the Americas, Europe, and Asia. Business Technographics uses only superior data sources and advanced data-cleaning techniques to ensure the highest data quality.

**Online Resource**

The online version of Figure 4 is an Excel-based vendor comparison tool that provides detailed product evaluations and customizable rankings.

**Data Sources Used In This Forrester Wave**

Forrester used a combination of 32 data sources to assess the strengths and weaknesses of each solution:

- › **Vendor surveys.** Forrester surveyed vendors on their capabilities as they relate to the evaluation criteria. Once we analyzed the completed vendor surveys, we conducted vendor calls where necessary to gather details of vendor qualifications.
- › **Product briefings and demos.** We asked vendors to conduct briefings and demonstrations of their product's functionality. We used findings from these product briefings and demos to validate details of each vendor's product capabilities.
- › **Customer reference calls.** To validate product and vendor qualifications, Forrester also conducted reference calls or conducted surveys with at least one of each vendor's current customers.

**The Forrester Wave Methodology**

We conduct primary research to develop a list of vendors that meet our criteria to be evaluated in this market. From that initial pool of vendors, we then narrow our final list. We choose these vendors based on: 1) product fit; 2) customer success; and 3) Forrester client demand. We eliminate vendors that have limited customer references and products that don't fit the scope of our evaluation.

After examining past research, user need assessments, and vendor and expert interviews, we develop the initial evaluation criteria. To evaluate the vendors and their products against our set of criteria, we gather details of product qualifications through a combination of lab evaluations, questionnaires, demos, and/or discussions with client references. We send evaluations to the vendors for their review, and we adjust the evaluations to provide the most accurate view of vendor offerings and strategies.

We set default weightings to reflect our analysis of the needs of large user companies — and/or other scenarios as outlined in the Forrester Wave report — and then score the vendors based on a clearly defined scale. These default weightings are intended only as a starting point, and we encourage readers to adapt the weightings to fit their individual needs through the Excel-based tool. The final scores

**The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016**

Elasticity, Automation, And Pay-As-You-Go Compel Enterprise Adoption Of Hadoop In The Cloud

generate the graphical depiction of the market based on current offering, strategy, and market presence. Forrester intends to update vendor evaluations regularly as product capabilities and vendor strategies evolve. For more information on the methodology that every Forrester Wave follows, go to <http://www.forrester.com/marketing/policies/forrester-wave-methodology.html>.

### Integrity Policy

All of Forrester's research, including Forrester Wave evaluations, is conducted according to our Integrity Policy. For more information, go to <http://www.forrester.com/marketing/policies/integrity-policy.html>.

### Endnotes

<sup>1</sup> In 2012, 52% of enterprises said developing a comprehensive cloud strategy was a high or critical priority; in 2014, this percentage climbed to 64%. See the "[Benchmark Your Enterprise Cloud Adoption](#)" Forrester report.

<sup>2</sup> Source: Forrester's Global Business Technographics Data And Analytics Survey, 2015.

<sup>3</sup> For more information, see the "[Apache Spark Is Powerful And Promising](#)" Forrester report and see the "[Brief: Apache Spark Ignites The Big Data Landscape](#)" Forrester report.

<sup>4</sup> Choosing the best Hadoop distribution for your organization is not easy. Each vendor's solutions have nuanced sweet spots. For an evaluation of the vendors, see the "[The Forrester Wave™: Big Data Hadoop Distributions, Q1 2016](#)" Forrester report.

<sup>5</sup> BAA stands for business associate agreement.

We work with business and technology leaders to develop customer-obsessed strategies that drive growth.

#### PRODUCTS AND SERVICES

- › Core research and tools
- › Data and analytics
- › Peer collaboration
- › Analyst engagement
- › Consulting
- › Events

---

Forrester's research and insights are tailored to your role and critical business initiatives.

#### ROLES WE SERVE

##### **Marketing & Strategy Professionals**

CMO  
B2B Marketing  
B2C Marketing  
Customer Experience  
Customer Insights  
eBusiness & Channel Strategy

##### **Technology Management Professionals**

CIO  
Application Development & Delivery  
› **Enterprise Architecture**  
Infrastructure & Operations  
Security & Risk  
Sourcing & Vendor Management

##### **Technology Industry Professionals**

Analyst Relations

---

#### CLIENT SUPPORT

For information on hard-copy or electronic reprints, please contact Client Support at +1 866-367-7378, +1 617-613-5730, or [clientsupport@forrester.com](mailto:clientsupport@forrester.com). We offer quantity discounts and special pricing for academic and nonprofit institutions.