



**CLOUDBERA**

# Cloudera DataFlow: IoT Data Management from Edge to Cloud

FEATURING RESEARCH FROM FORRESTER

Now Tech: IoT Data Management  
Solutions, Q2 2019

# Cloudera DataFlow: IoT Data Management from Edge to Cloud

**An enterprise's IoT initiative is only as good as its ability to harness the value of the data captured from its IoT devices in real-time. The enterprise will require an IoT data management solution to see true ROI with their IoT initiatives.**

## KEY DATA CHALLENGES WITH IOT INITIATIVES

The sheer sophistication and the complexity of today's IoT use cases are making enterprises struggle with data management. On one hand, the enormity of the data sets being generated by the IoT devices are hard to process and on the other hand, the speed at which the data must be ingested also makes it extremely challenging to process and analyze the data in real-time. A third dimension to the problem is the sheer diversity in the types of data that need to be processed. In today's use cases, a large set of unstructured data are part of the mix.

With 5G still in its infancy stages, network latency and bandwidth issues are seriously constraining IoT initiatives from transmitting all the data generated from the edge back to the cloud or the enterprise. So, there is a strong need for processing a lot of the data at the edge itself and transmitting what is relevant for subsequent analysis. And then, there are use cases in Oil & Gas or Aviation where the data cannot be transmitted instantaneously because the network is not available for longer periods. This warrants a need for local edge storage to accommodate store-and-forward scenarios.

Finally, two key data challenges that often get overlooked are security and governance. While enterprises are waking up to the fact the device security is of utmost importance in an IoT implementation, they have to understand that data security is equally important right from the edge to the cloud. In this age of GDPR and other such compliance regulations, it is also critical to understand your data better, establish a trust model, and track it from edge to cloud. Data lineage tracking is critical to understand where a piece of data came from, who touched it and what changes were made to it.

## IN THIS DOCUMENT

- 1 Cloudera DataFlow: IoT Data Management from Edge to Cloud
- 4 Research From Now Tech: IoT Data Management Solutions, Q2 2019
- 19 About Cloudera

## CLUDERA DATAFLOW

Cloudera DataFlow (CDF) is a comprehensive edge-to-cloud IoT data management platform. It addresses the key data management challenges with streaming and IoT data for all types of enterprises. CDF can capture data at the edge, stream it into the enterprise or the cloud, process the streams in real-time and be able to generate actionable intelligence from that data in real-time.

Cloudera DataFlow comprises of four key parts –

- **Cloudera Edge Management (CEM)** is made up of edge agents and an edge management hub. It manages, controls and monitors edge agents to collect data from edge devices and push intelligence back to the edge. CEM allows you to develop, deploy, run and monitor edge flow apps on thousands of edge devices. Edge Flow Manager (EFM) is an agent management hub that supports a graphical flow-based programming model to develop, deploy & monitor edge flows on thousands of MiNiFi agents. EFM can also push and execute ML models at the edge to do avoid costly roundtrips between the edge and the cloud.
- **Cloudera Flow Management (CFM)** is a no-code data ingestion and integration engine powered by Apache NiFi. CFM offers 300+ processors out-of-the-box to integrate with scores of data sources on-premises as well as on the cloud. CFM allows you to do data transformations, enrichment and setup content routing rules with drag-and-drop ease.
- **Cloudera Stream Processing (CSP)** enables stream processing capabilities with Apache Kafka to a scale of millions of IoT messages per second. CSP also supports advanced messaging, schema management and analytics capabilities with its support for Schema Registry and Kafka Streams as well.
- **Cloudera Streaming Analytics (CSA)** offers a choice of multiple streaming analytic solutions like Kafka Streams, Spark Structured Streaming, Apache Storm, Streaming Analytics Manager etc. These engines help with pattern matching, complex event processing, windowing etc to generate real-time insights from IoT streams.

## KEY BENEFITS

The Cloudera DataFlow platform is 100% open source, giving management teams the confidence of having no vendor lock-in. The no-code user interfaces of Apache NiFi and Edge Flow Manager boost the productivity of developers by being able to build sophisticated data orchestrations with drag-and-drop ease. Operations teams also stand to benefit from the ease of control, management and monitoring of thousands of edge agents from a single, central console. Data analysts benefit from edge intelligence and streaming analytics and the ability to keep refining ML models over time. Cloudera DataFlow supports a multi-cloud, hybrid architecture for your IoT needs but at the same time offers the flexibility to be deployed on-premises, at the edge or on the cloud. With hundreds of customers using it mission-critical use cases, read more about why CDF is included in this Forrester report among other IoT Data Management Solutions.

# Now Tech: IoT Data Management Solutions, Q2 2019

Forrester's Overview Of 22 IoT Data Management Solution Providers

by Noel Yuhanna and Michele Goetz

June 6, 2019

## Why Read This Report

You can use internet-of-things (IoT) data management solutions to deliver a comprehensive view of IoT data, become significantly more productive, and accelerate IoT initiatives. But to realize these benefits, you'll first have to select from a diverse set of vendors that vary by size, functionality, geography, and vertical market focus. Enterprise architecture (EA) professionals should use this report to understand the value they can expect from IoT data management solution providers and select vendors based on size and functionality.

## Key Takeaways

### **Improve IoT Insights With Integrated IoT Data Management**

Data management solutions foster business agility and innovation by allowing organizations to focus on business outcomes rather than deal with technology challenges.

### **Select Vendors Based On Size And Functionality**

Forrester has identified three key market segments: cloud services, on-premises, and systems-based data management. Leverage the platform based on your IoT strategy.

### **Use IoT Data Management Solutions To Accelerate IoT Initiatives**

IoT devices and sensors get all the press, but it's data and analytics that make IoT a reality for enterprises.

# Now Tech: IoT Data Management Solutions, Q2 2019

## Forrester's Overview Of 22 IoT Data Management Solution Providers

by [Noel Yuhanna](#) and [Michele Goetz](#)

with [Gene Leganza](#) and Elizabeth Hoberman

June 6, 2019

---

### Table Of Contents

[Improve IoT Insights With Integrated IoT Data Management](#)

[Select Vendors Based On Size And Functionality](#)

[IoT Data Management Solutions Market Presence Segments](#)

[IoT Data Management Solutions Functionality Segments](#)

[Align Individual Vendor Solutions To Your Organization's Needs](#)

---

[Recommendations](#)

[Use IoT Data Management Solutions To Accelerate IoT Initiatives](#)

---

[Supplemental Material](#)

### Related Research Documents

[The Forrester Wave™: Big Data Fabric, Q2 2018](#)

[The Forrester Wave™: Cloud Hadoop/Spark Platforms, Q1 2019](#)

[Predictions 2019: The Internet Of Things](#)



**Share reports with colleagues.**  
Enhance your membership with  
Research Share.

## Improve IoT Insights With Integrated IoT Data Management

IoT devices and solutions are becoming commonplace in businesses, homes, and cities, powered by affordable sensors, cloud and edge infrastructure, networks, platforms, and applications.<sup>1</sup> Most specialized IoT use cases focus on transforming operations and processes in specific sectors, such as remote patient monitoring in healthcare, proactively monitoring machines in manufacturing, and security monitoring across cities.<sup>2</sup> While collecting IoT data from sensors and devices is often straightforward, enterprises struggle to integrate, transform, wrangle, and aggregate it with other non-IoT data sources to deliver comprehensive and integrated analytics to support EA pros, data scientists, app developers, and others.<sup>3</sup> Solving this problem is why enterprises are adopting and implementing IoT data management.

Forrester defines IoT data management solutions as:

*Optimized and comprehensive end-to-end data management that focuses on automating the process of ingesting, storing, integrating, wrangling, securing, and transforming IoT and non-IoT data to support IoT analytics.*

By leveraging IoT data management solutions, EA pros enable the business to:

- › **Deliver a comprehensive view of IoT-related data.** These tools ingest, prepare, transform, and orchestrate data from multiple IoT and non-IoT data sources in real time to support advanced analytics, integrated dashboards, and actionable insights. They ensure end-to-end data lineage to deliver trusted insights and share data with other apps and frameworks.
- › **Become significantly more productive.** Finding and using insights from IoT data is an iterative and continuous process of data ingestion, preparation, and integration and applying various machine learning (ML) models. IoT data management tools have the breadth and depth EA pros need to rapidly work through messy processes and data for faster IoT insights.
- › **Accelerate IoT initiatives through ML and intelligence.** IoT data is voluminous, making it difficult to manage, integrate, or blend it with other sources. IoT data management solutions automate the process of discovering, integrating, and orchestrating data by leveraging AI and ML capabilities built into the platform.

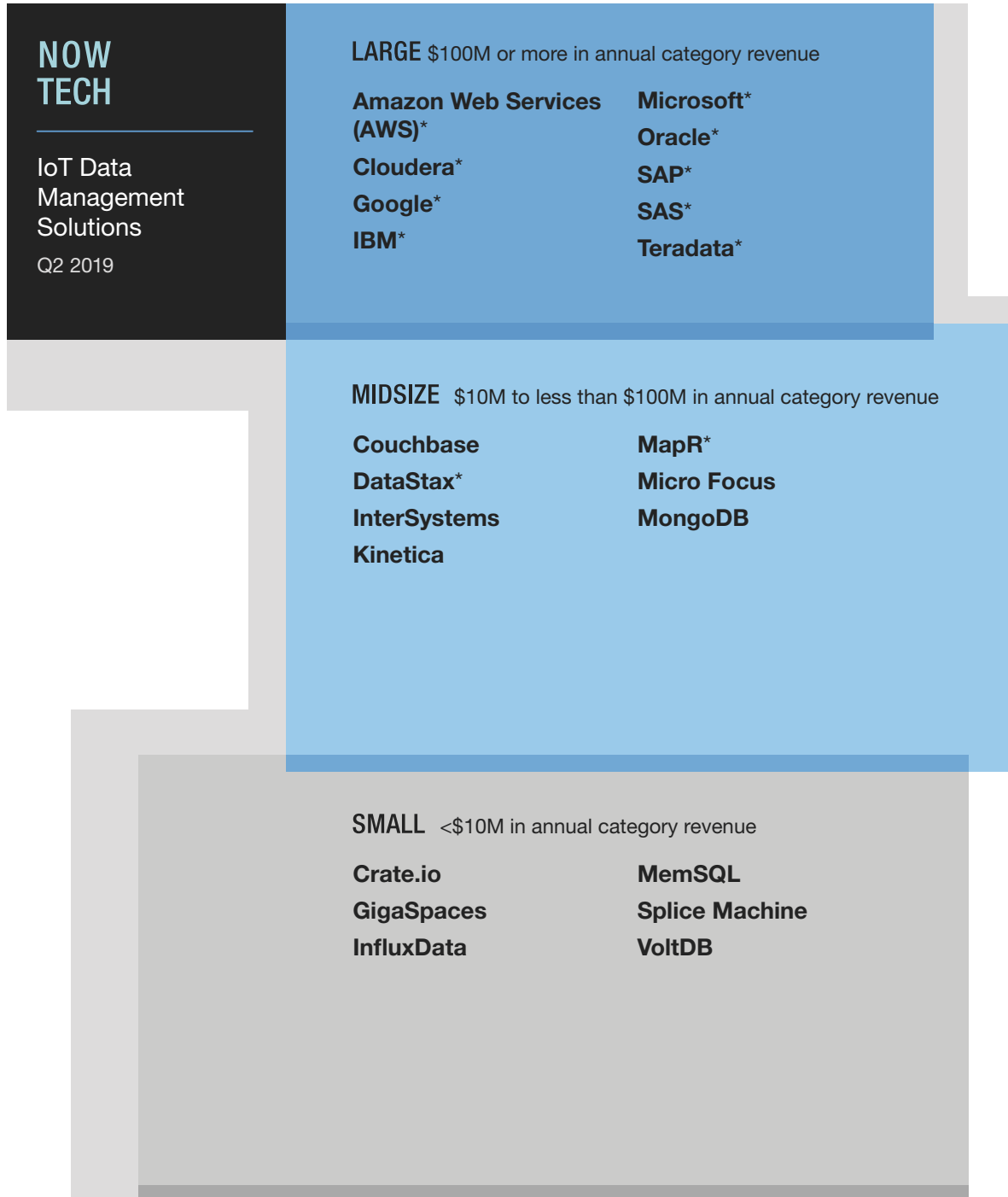
## Select Vendors Based On Size And Functionality

We've based our analysis of the IoT data management solutions market on two factors: market presence and functionality.

### **IOT DATA MANAGEMENT SOLUTIONS MARKET PRESENCE SEGMENTS**

We segmented the vendors in this market into three categories, based on annual IoT data management solution revenue: large established players (\$100 million or more in category revenue), midsize players (\$10 million to less than \$100 million in revenue), and smaller players (less than \$10 million in revenue) (see Figure 1). We did not include vendors that we estimated to have less than \$1 million in annual IoT data management solution revenue.

**FIGURE 1** Now Tech Market Presence Segments: IoT Data Management Solutions, Q2 2019



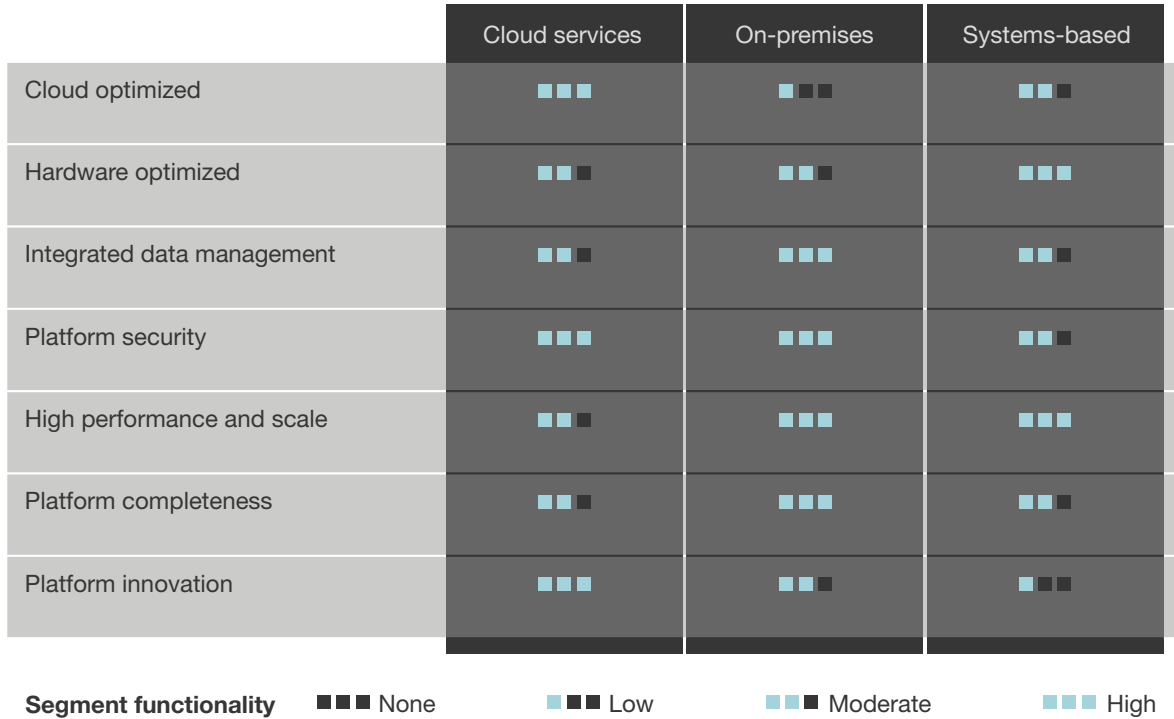
\*Forrester estimate



### IOT DATA MANAGEMENT SOLUTIONS FUNCTIONALITY SEGMENTS

To explore functionality at a deeper level, we broke the IoT data management solutions market into three segments, each with varying capabilities (see Figure 2):

- › **Cloud IoT data management is elastic.** The new generation of cloud-based IoT data management identifies vendors that provide IoT platforms in one or more public clouds, such as Amazon Web Services (AWS), Google Cloud Platform, and Microsoft Azure. These are vendors that offer IoT data management services by leveraging elastic compute and storage infrastructure. Many of the cloud-based vendors also integrate with other big data and analytical services and, of course, provide a pay-per-use pricing model.
- › **On-premises software-based IoT data management offers flexibility.** The on-premises IoT data management vendors are, in most cases, the founders of the IoT data management category. They focus solely on bringing innovation from multiple open source projects and often commercial technology for enterprises. They also make significant contributions to the open source community. Some on-premises vendors also support cloud implementations and work with infrastructure vendors to deliver prebuilt IoT-optimized systems.
- › **Systems-based IoT data management focuses on optimized performance.** An IoT data management solution ultimately must run on some hardware infrastructure, whether on-premises or in the cloud. Several hardware vendors have designed systems that are optimized for or can be configured to support various IoT workloads. Although the hardware system can be deployed on-premises, it is also used by cloud vendors to deliver optimized IoT analytics.

**FIGURE 2** Now Tech Functionality Segments: IoT Data Management Solutions, Q2 2019


## Align Individual Vendor Solutions To Your Organization's Needs

The following tables provide an overview of vendors with details on functionality category, geography, and vertical market focus (see Figure 3, see Figure 4, and see Figure 5).

**FIGURE 3** Now Tech Large Vendors: IoT Data Management Solutions, Q2 2019

**LARGE** \$100M or more in annual category revenue

	Primary functionality segments	Geographic presence (by revenue %)	Vertical market focus	Sample customers
<b>Amazon Web Services (AWS)</b>	Cloud services	NA 60%; LATAM 10%; EMEA 20%; AP 10%	Telecom; energy, oil, and gas; manufacturing; automotive/transportation	Bayer Crop Science; iRobot; Modjoul; Pentair; Rachio; Vantage Power
<b>Cloudera</b>	Cloud services; on-premises; systems-based	NA 60%; EMEA 30%; AP 10%*	Industrial/manufacturing; automotive/transportation; energy, oil, and gas	American Water; Clearsense; Komatsu; Navistar; Octo; Renault
<b>Google</b>	Cloud services	NA 60%; EMEA 30%; AP 10%*	Manufacturing; energy, oil, and gas; automotive/transportation; telecom*	Aker BP; Flex Digital Health; Oden Technologies; Smart Parking; Swiss Steel; Vagabond
<b>IBM</b>	Cloud services; on-premises; systems-based	NA 46%; EMEA 32%; AP 21%; other 1%*	Manufacturing; energy; transportation; building management; retail	L'Oréal; Sandvik; Sugar Creek Brewing
<b>Microsoft</b>	Cloud services; on-premises	NA 60%; LATAM 10%; EMEA 20%; AP 10%	Manufacturing; supply chain management; transportation/logistics	Ecolab; Johnson Controls; Rockwell Automation; Schneider Electric; Tetra Pak; thyssenkrupp
<b>Oracle</b>	Cloud services; on-premises; systems-based	NA 30%; LATAM 15%; EMEA 35%; AP 20%	Industrial/manufacturing; transportation/logistics; utilities	Hitachi Consulting; Noble Plastics; Nucor

\*The vendor did not provide information for this cell; this is Forrester's estimate.

**FIGURE 3** Now Tech Large Vendors: IoT Data Management Solutions, Q2 2019 (Cont.)**LARGE** \$100M or more in annual category revenue

	<b>Primary functionality segments</b>	<b>Geographic presence (by revenue %)</b>	<b>Vertical market focus</b>	<b>Sample customers</b>
<b>SAP</b>	Cloud services; on-premises; systems-based	NA 60%; LATAM 10%; EMEA 20%; AP 10%*	Industrial/ manufacturing; healthcare; telecom; transportation/ logistics; retail*	City of Antibes; dormakaba; Hilti; Kaiserwetter Energy Asset Management; Severstal
<b>SAS</b>	Cloud services; on-premises	NA 60%; LATAM 10%; EMEA 20%; AP 10%*	Manufacturing; energy (oil and gas, utilities); transportation	Honeywell; Lockheed Martin; Siemens; Volvo
<b>Teradata</b>	Cloud services; on-premises; systems-based	NA 50%; EMEA 40%; AP 10%	Manufacturing; utilities; transportation; automotive; communications; gaming; agriculture; government; retail	Blizzard Entertainment; Boeing; Caterpillar; General Motors; Maersk; Siemens; Union Pacific; Volvo

\*The vendor did not provide information for this cell; this is Forrester's estimate.

**FIGURE 4** Now Tech Midsize Vendors: IoT Data Management Solutions, Q2 2019

**MIDSIZE** \$10M to less than \$100M in annual category revenue

	<b>Primary functionality segments</b>	<b>Geographic presence (by revenue %)</b>	<b>Vertical market focus</b>	<b>Sample customers</b>
<b>Couchbase</b>	Cloud services; on-premises	NA 63%; EMEA 37%	Travel and hospitality; telecom; manufacturing/logistics; retail; media; financial services; healthcare	BD; Coyote; GE; Royal Caribbean; Verizon
<b>DataStax</b>	Cloud services; on-premises	NA 70%; EMEA 20%; AP 10%*	Telecom; transportation; retail; media; healthcare*	Equinix; Traxens; Walmart
<b>InterSystems</b>	Cloud services; on-premises	NA 71%; LATAM 3%; EMEA 22%; AP 4%	Healthcare; manufacturing; logistics; government; energy; financial services	BTC; Fujifilm; Navantia; Partners Healthcare; STANLEY Healthcare
<b>Kinetica</b>	Cloud services; on-premises; systems-based	NA 50%; LATAM 10%; EMEA 20%; AP 20%	Telecom; financial services; government	Anadarko Petroleum; GlaxoSmithKline; OVO/Lippo Group; SoftBank; Telkomsel; US Postal Service
<b>MapR</b>	Cloud services; on-premises	NA 70%; EMEA 20%; AP 10%*	Automotive; financial services; healthcare; manufacturing; retail; energy; telecom	Anadarko Petroleum; Audi; Cisco; Ericsson; Whiting Petroleum
<b>Micro Focus</b>	Cloud services; on-premises; systems-based	NA 60%; EMEA 25%; AP 15%	Industrial/manufacturing; telecom; healthcare	Anritsu; Climate (part of Monsanto); Nimble Storage; Optimal Plus; Philips; Suunto
<b>MongoDB</b>	Cloud services; on-premises; systems-based	NA 65%; EMEA 30%; AP 5%	Manufacturing; healthcare; technology; energy	AXA; Bosch; City of Chicago; Florida Power & Light; Humana

\*The vendor did not provide information for this cell; this is Forrester's estimate.

**FIGURE 5** Now Tech Small Vendors: IoT Data Management Solutions, Q2 2019

**SMALL** <\$10M in annual category revenue

	Primary functionality segments	Geographic presence (by revenue %)	Vertical market focus	Sample customers
<b>Crate.io</b>	Cloud services; on-premises	NA 45%; EMEA 55%	Manufacturing; smart city/smart building infrastructure; software	ALPLA
<b>GigaSpaces</b>	Cloud services; on-premises	NA 50%; EMEA 35%; AP 15%	Transportation; retail; manufacturing; insurance; telecom	Avaya; CSX; Folksam; Magic Software Enterprises; Schneider
<b>InfluxData</b>	Cloud services; on-premises	NA 30%; EMEA 70%	Energy and utilities; manufacturing; consumer devices	BBOX; Equinor; Siemens; tado; Tesla
<b>MemSQL</b>	Cloud services; on-premises	NA 76%; EMEA 9%; AP 15%	Energy and utilities; finance; high-tech; media and communications	Akamai Technologies; Comcast; Uber Technologies; Verizon
<b>Splice Machine</b>	Cloud services; on-premises; systems-based	NA 100%	Financial services; healthcare; industrial	Anthem Insurance; Cetera Financial Group; ClearSense; Kroger; Wells Fargo
<b>VoltDB</b>	Cloud services; on-premises	EMEA 50%; AP 50%	Utilities; networks	CGI; Mitsubishi Electric; Openet; Sakura Network (Japan Foundation)

## Recommendations

### Use IoT Data Management Solutions To Accelerate IoT Initiatives

IoT devices and sensors get all the press, but it's data and analytics that make IoT a reality for enterprises. IoT data management solutions help enterprises overcome IoT-related data challenges and simplify and accelerate IoT initiatives. To select the right solution, EA pros must understand the types of options that exist, ranging from on-premises software to cloud-based and optimized hardware systems offerings. As you consider your selection of IoT data management, you must:

- › **Shortlist IoT data management vendors that suit your requirements.** With more than two dozen vendors offering IoT data management solutions, shortlisting can be challenging. If you're choosing a cloud strategy, also look at system-based solutions that can run in the public and private cloud. Most of the innovation in IoT data management is happening in the cloud, so look at cloud offerings before on-premises unless you have a mandate to stay on-premises.
- › **Invest in IoT data management with an eye on your future needs.** Most enterprises typically make IoT data management solution choices based on their current needs, but as IoT becomes more strategic, requirements will likely change. Review your company's IoT road map and identify capabilities you'll need over the next three to four years. This exercise will help you shortlist vendors to include only those that can help you reach these forward-looking goals.
- › **Involve key stakeholders across the enterprise early on.** To align priorities and bolster buy-in to the final decision, you must collaborate with various stakeholders involved in the IoT data management evaluation process. Key stakeholders will include EA pros, data scientists, app developers, information security, and others who may use the platform.
- › **Secure all data for IoT from the start.** Enterprises struggle with disparate and sometimes immature security offerings that fail to properly secure IoT deployments.<sup>4</sup> Enterprises should take the utmost care to ensure that the IoT data management is accessed only by users, applications, and APIs that have gone through proper authentication, authorization, and access control measures, especially when dealing with sensitive IoT data.

## 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

To help you put research into practice, connect with an analyst to discuss your questions in a 30-minute phone session — or opt for a response via email.

[Learn more.](#)

### Analyst Advisory

Translate research into action by working with an analyst on a specific engagement in the form of custom strategy sessions, workshops, or speeches.

[Learn more.](#)

### Webinar

Join our online sessions on the latest research affecting your business. Each call includes analyst Q&A and slides and is available on-demand.

[Learn more.](#)



**Forrester's research apps for iOS and Android.**

Stay ahead of your competition no matter where you are.

## Supplemental Material

### MARKET PRESENCE METHODOLOGY

We defined market presence in Figure 1 based on the revenue the vendor derives from IoT data management solutions.

To complete our review, Forrester requested information from vendors. If vendors did not share this information with us, we made estimates based on available secondary information. We've marked companies with an asterisk if we estimated revenues or information related to geography or industries. Forrester fact-checked this report with vendors before publishing.



## Endnotes

- <sup>1</sup> IoT-enabled products, assets, and solutions power digital transformation initiatives to create sustainable customer relationships, enhance operations, and differentiate products in both B2C and B2B environments. See the Forrester report "[Predictions 2019: The Internet Of Things.](#)"
- <sup>2</sup> IoT use cases like building management, security, and surveillance apply to many firms; other use cases enhance operational processes; and still others address specialized needs in industries such as healthcare or industrial manufacturing. See the Forrester report "[Internet-Of-Things Heat Map 2018.](#)"
- <sup>3</sup> Forty-four percent of global data and analytics decision makers whose firms use IoT services say that integrating enterprise and IoT data is their firm's biggest challenge with data, insights, and analytics for IoT. Source: Forrester Analytics Global Business Technographics® Data And Analytics Survey, 2018.
- <sup>4</sup> While business leaders are excited about the business insights they will garner from connected devices, security and risk pros struggle with disparate and sometimes immature security offerings that fail to properly secure deployments, leading to increased risk of data loss and physical damage. See the Forrester report "[The State Of IoT Security 2018.](#)"

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.



# CLUDERA

At Cludera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cludera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cludera advances digital transformation for the world's largest enterprises. Learn more at [Cludera.com](https://Cludera.com).