The Forrester Wave™: Big Data Warehouse, Q2 2017

Adoption Grows As Enterprises Look To Revive Their EDW Strategy

by Noel Yuhanna
June 15, 2017 | Updated: June 21, 2017

Why Read This Report
In our 26-criteria evaluation of big data warehouse solutions, we identified the 15 most significant vendors — Amazon Web Services (AWS), Cazena, Cloudera, Hortonworks, Hewlett Packard Enterprise (HPE), IBM, MapR, MarkLogic, MemSQL, Microsoft, Oracle, Phemi, SAP, Snowflake, and Teradata — and researched, analyzed, and scored them. This report shows how each vendor measures up and helps enterprise architecture (EA) pros make the right choice for their data warehouse strategy.

Key Takeaways

Fifteen Big Data Warehouse Vendors Compete In This Hot Market
Among the commercial big data warehouse vendors Forrester evaluated, we found six Leaders, six Strong Performers, and three Contenders.

EA Pros Are Looking For Integration, Security, Automation, And Performance Features
The market is growing because more enterprise architects see that the big data warehouse is critical for their analytical strategy.

Machine Learning, Self-Service, Security, And Scalability Are Key Differentiators
The Leaders we identified offer large and complex deployments, support a broader set of use cases, and deliver a high degree of automation. The Strong Performers have turned up the heat as high as it will go on the incumbent Leaders and are ramping up to offer more integrated and lower-cost solutions.
The Forrester Wave™: Big Data Warehouse, Q2 2017
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

by Noel Yuhanna
with Gene Leganza and Jun Lee
June 15, 2017 | Updated: June 21, 2017

Table Of Contents

2 Big Data Warehouse Delivers A Modern EDW Platform
   Big Data Warehouse Combines The Best Of Old And New Approaches
   The Big Data Warehouse Market Includes Traditional EDW, Big Data, And Niche Vendors
   Big Data Warehouse Offers Use Cases That Go Beyond Traditional Ones

4 Big Data Warehouse Evaluation Overview
   Evaluated Vendors And Inclusion Criteria

7 Vendor Profiles
   Leaders
   Strong Performers
   Contenders

14 Supplemental Material

Related Research Documents

The Forrester Wave™: Big Data Fabric, Q4 2016
The Forrester Wave™: Big Data Hadoop Cloud Solutions, Q2 2016
The Forrester Wave™: Enterprise Data Warehouse, Q4 2015
Big Data Warehouse Delivers A Modern EDW Platform

For decades, enterprise architects have used data warehouse architectures to support various customer analytics, integrated analytics, and advanced and predictive analytics across the financial services, retail, manufacturing, oil and gas, telecommunications, and insurance industries. However, organizations are realizing that traditional data warehouses are failing to meet new business requirements, especially around streaming data, real-time analytics, large volumes of messy and complex data sets, and ad hoc requirements. On the other hand, big data technologies have gaps in maturity, especially for security and integration, and organizations lack both business competency (to deal with data that is messy, diverse, or high volume) and technology skills and legal issues (see Figure 1).

Big Data Warehouse Combines The Best Of Old And New Approaches

Big data warehouse (BDW) is a modern architecture that combines both of the best worlds — leveraging existing data warehouse and new big data technologies that enable organizations to support their growing analytical requirements. BDW minimizes the complexity of messy data and hides heterogeneity by embodying a trusted model and adapting to changing business requirements. It leverages unified metadata, distributed in-memory compute resources, and distributed data repositories to deliver scalable and integrated analytics. Adoption of BDW will accelerate as enterprises run into EDW challenges and look at using existing investments and technologies along with newer architectures to support innovation, boost growth, and minimize customer churn.

Forrester defines big data warehouse as:

A specialized, cohesive set of data repositories and platforms that supports a broad variety of analytics running on-premises, in the cloud, or in a hybrid environment. BDW leverages traditional and new big data technologies such as Hadoop, Spark, columnar and row-based data warehouses, ETL and streaming, and elastic in-memory and storage frameworks.

The Big Data Warehouse Market Includes Traditional EDW, Big Data, And Niche Vendors

The BDW market comprises traditional data warehouse vendors such as HPE, IBM, Microsoft, Oracle, SAP, and Teradata; big data vendors like Cloudera, Hortonworks, and MapR; and niche big data warehouse vendors that focus on analytics and big data, such as Cazena, MemSQL, Phemi, and Snowflake. Traditional data warehouse vendors are expanding their software to integrate with Apache’s Hadoop and Spark ecosystems, in-memory technologies, and employ improved automation and machine learning capabilities. Big data vendors focus on data lakes, streaming data, machine learning, real-time processing, and integration with traditional data sources. Niche big data warehouse vendors focus on a platform that’s easy to deploy and use; integrates with Hadoop, Spark, and traditional sources; leverages in-memory; and focuses on ad hoc and real-time use cases.
Big Data Warehouse Offers Use Cases That Go Beyond Traditional Ones

Enterprises are already leveraging BDW to support all kind of analytics and insights, including social analytics, customer 360, internet-of-things (IoT) analytics, pricing trends, healthcare patient monitoring, machine failure analysis, fraud detection, and other actionable insights. The top BDW use cases are:

› **Integrated analytics that deals with diverse data sources and types.** A key challenge in the traditional EDW approach was that if data didn’t exist in the warehouse, you couldn’t do any analytics. BDW performs integrated analytics across data warehouse and Hadoop clusters by leveraging Hadoop to store and process large sets of semistructured and unstructured data, log files, and streaming data with ease. For example, health research often requires looking at complex patient data and determining the likely efficacy of a treatment based on factors like age, sex, and health status. BDW enables EA pros to gather and store millions of data points in Hadoop and perform complex navigation and modeling using traditional data warehouse and in-memory technology.

› **IoT analytics that deals with high ingestion speed.** Traditional data warehouses don’t deal with IoT data. However, BDW offers the ability to efficiently store, process, and access large volumes of IoT data from sensors and devices in Hadoop repositories through automation and machine learning technologies. Manufacturers use highly sophisticated machinery to support their plants, whether they’re building a car, airplane, or tire or bottling wine or soda. Every minute of machine downtime can cost a manufacturer dearly. IoT analytics on BDW platforms enables manufacturers to predict machine failures based on sensor data, minimizing or eliminating production slowdown.

› **Right-time business analytics that deals with quicker and ad hoc requirements.** In traditional EDW architectures, preparing data for analytics was mostly done in batch mode, with ETL doing the heavy lifting of moving data from transactional systems to operational systems to data warehouses.³ As a result, by the time business and data analysts could leverage the data warehouse to do analytics, the data was already 12 to 48 hours old. BDWs enable right-time analytics by using streaming and replication with direct access to data sources, whether on-premises or in the cloud, bypassing traditional ETL approaches. The financial services industry has been an early adopter of BDW to support right-time analytics for portfolio management, fraud detection, and asset management.

› **Adaptive and self-service analytics that enables more data intelligence.** Most EDWs use predefined data sources to deliver predictive analytics, trends, and insights. The big data warehouse enables organizations to dynamically leverage new data sources quickly to deliver new insights. It enables self-service capabilities for business users to ask complex and new questions so they can make more accurate decisions. The BDW also adapts to the new sources and can help correlate data using machine learning and adaptive intelligence. For example, a major European bank recently built a BDW framework that business units use to support self-service for making better decisions on investments and risks. The platform represents a major shift from the static reports the bank used previously.
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

FIGURE 1 Big Data Creates New Business And Technology Challenges

“What are/were the biggest challenges in executing your vision for big data?”

- Lack of business competency to deal with data that is messy, diverse, or large: 20%
- Maturity of technology around security: 20%
- Lack of technology skills: 19%
- Legal and compliance issues: 19%
- Data silos created by management and organizational challenges: 18%
- Organizational business issues with data stewardship and governance: 17%
- Maturity of technology around governance and metadata management: 16%
- Maturity of technology around preparing messy, diverse, and large data: 16%

Base: 3,343 global data and analytics technology decision makers
Note: Multiple responses accepted; top eight responses shown.
Source: Forrester Data Global Business Technographics® Data And Analytics Survey, 2016

Big Data Warehouse Evaluation Overview

To assess the state of the big data warehouse market and see how the vendors stack up against each other, Forrester evaluated the strengths and weaknesses of top BDW vendors. After examining past research, user need assessments, and vendor and expert interviews, we developed a comprehensive set of 26 evaluation criteria, which we grouped into three high-level buckets:

- **Current offering.** To assess the breadth and depth of each vendor’s BDW offering, we evaluated each solution’s architectural and operational functionality.

- **Strategy.** We reviewed each vendor’s strategy to assess how it plans to evolve its BDW services to meet emerging customer demands. We also evaluated each vendor’s go-to-market approach, commitment, and direction strategies.
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

Market presence. To determine each offering’s market presence, we evaluated the providers’ company financials, adoption, and partnerships.

Evaluated Vendors And Inclusion Criteria

Forrester included 15 vendors in the assessment: AWS, Cazena, Cloudera, Hortonworks, HPE, IBM, MapR, MarkLogic, MemSQL, Microsoft, Oracle, Phemi, SAP, Snowflake, and Teradata. Each of these vendors has (see Figure 2):

A comprehensive big data warehouse offering. Evaluated vendors must provide BDW functions as defined in the Forrester report “The Next-Generation EDW is The Big Data Warehouse.” A key component of the BDW architecture is the ability to leverage various specialized data repositories, such as traditional relational data warehouses, columnar data warehouses, and Hadoop. Unlike traditional data warehouses, BDW minimizes complexity and hides heterogeneity by embodying a trusted model; supports all types of data, including unstructured data; and adapts to changing business requirements more rapidly through a self-service platform. BDW centralizes the administration of distributed data repositories, in-memory compute resources, metadata, storage, access, and processing functions.

A standalone big data warehouse solution. Evaluated vendors provide a software solution that organizations can implement independent of analytical solutions. The solution should not be technologically tied or bundled to any particular application, product, or solution. The vendor must market the data warehouse or Hadoop or BDW-like product, solution, or service. The solution can run in cloud and/or on-premises platforms.

Big data use cases. The solution must be able to support big data use cases, such as customer churn, IoT, 360-degree view of customer and business, advanced analytics, real-time analytics, and others.

A referenceable install base. The vendor should have 10 or more unique enterprise paying customers using the BDW product, service, or solution that span more than one major geographical region. Each vendor provided at least two customer references for Forrester to interview.

A publicly available solution. The participating vendors must have actively marketed a BDW product, service, or solution as of December 1, 2016.

Customer interest. Forrester included only vendors that customers mentioned several times during Forrester inquiry calls in the past 12 months related to data warehouse, big data warehouse, and big-data-related topics.

Client inquiries and/or technologies that put the vendor on Forrester’s radar. Forrester clients often discuss the vendors and products through inquiries and interviews; alternatively, the vendor may, in Forrester’s judgment, warrant inclusion or exclusion in this evaluation because of technology trends and market presence.
FIGURE 2 Evaluated Vendors: Product Information And Selection Criteria

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Product name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon Web Services</td>
<td>Amazon Redshift</td>
</tr>
<tr>
<td>Cazena</td>
<td>Cazena</td>
</tr>
<tr>
<td>Cloudera</td>
<td>Cloudera Enterprise 5.9</td>
</tr>
<tr>
<td>Hewlett Packard Enterprise</td>
<td>HPE Vertica Analytics Platform 8.0.1</td>
</tr>
<tr>
<td>Hortonworks</td>
<td>Hortonworks Data Platform 2.5</td>
</tr>
<tr>
<td>IBM</td>
<td>IBM dashDB, IBM DB2, IBM BigInsights, IBM Information Server,</td>
</tr>
<tr>
<td></td>
<td>IBM Security Guardium, IBM InfoSphere Optim, IBM Fluid Query</td>
</tr>
<tr>
<td>MapR Technologies</td>
<td>MapR Converged Data Platform 5.2</td>
</tr>
<tr>
<td>MarkLogic</td>
<td>MarkLogic Server v8.0-6</td>
</tr>
<tr>
<td>MemSQL</td>
<td>MemSQL 5.5</td>
</tr>
<tr>
<td>Microsoft</td>
<td>SQL Server 2016, Analytics Platform System (APS), Azure SQL Data Warehouse,</td>
</tr>
<tr>
<td></td>
<td>HDInsight for Hadoop</td>
</tr>
<tr>
<td>Oracle</td>
<td>Oracle Database 12c, Oracle Exadata Database Machine X6, Oracle Big Data</td>
</tr>
<tr>
<td></td>
<td>Appliance X6, Oracle Big Data SQL</td>
</tr>
<tr>
<td>Phemi</td>
<td>PHEMI Central Big Data Warehouse</td>
</tr>
<tr>
<td>SAP</td>
<td>SAP HANA 2, SAP Vora 1, SAP BW/4HANA 1, SAP Data Services,</td>
</tr>
<tr>
<td></td>
<td>SAP Cloud Platform Big Data Services (formerly Altiscale)</td>
</tr>
<tr>
<td>Snowflake</td>
<td>Snowflake Elastic Data Warehouse</td>
</tr>
<tr>
<td>Teradata</td>
<td>Teradata Database 16.0, Teradata Unified Data Architecture,</td>
</tr>
<tr>
<td></td>
<td>Teradata Appliance for Hadoop 6</td>
</tr>
</tbody>
</table>

**Vendor inclusion criteria**

Each vendor included in this evaluation has the following: 1) a comprehensive big data warehouse offering; 2) a standalone big data warehouse solution; 3) big data use cases; 4) a referenceable install base; 5) a publicly available offering; 6) customer interest; and 7) Forrester client inquiries and/or technologies that put the vendor on Forrester's radar.
Vendor Profiles

This evaluation of the big data warehouse 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 (see Figure 3).
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

**FIGURE 3 Forrester Wave™: Big Data Warehouse, Q2 ’17 (Cont.)**

<table>
<thead>
<tr>
<th>Current Offering</th>
<th>Forrester’s weighting</th>
<th>Development</th>
<th>Deployment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50%</td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>3.78 1.94 3.28 3.14 3.52 3.90 3.10 3.14 2.34 3.56 3.80 2.10 4.08 2.48 4.28</td>
<td>3.50 2.30 2.80 2.90 2.80 3.90 2.70 2.70 2.30 4.00 4.00 2.70 4.60 2.80 4.20</td>
<td>4.20 1.40 4.00 3.50 4.60 3.90 3.70 3.80 2.40 2.90 3.50 1.20 3.30 2.00 4.40</td>
</tr>
</tbody>
</table>

| Strategy | 50% | 4.60 2.60 3.00 3.60 3.80 3.40 2.60 3.20 2.60 2.60 4.20 2.00 4.00 3.40 3.20 |
| Ability to execute | 40% | 5.00 3.00 3.00 3.00 3.00 3.00 3.00 3.00 1.00 5.00 3.00 3.00 3.00 3.00 |
| Road map | 30% | 5.00 3.00 3.00 3.00 5.00 3.00 3.00 3.00 3.00 1.00 5.00 5.00 3.00 |
| Professional services | 20% | 3.00 1.00 3.00 5.00 3.00 5.00 1.00 3.00 1.00 5.00 5.00 1.00 5.00 1.00 3.00 |
| Support | 10% | 5.00 3.00 3.00 5.00 5.00 3.00 5.00 3.00 3.00 3.00 3.00 5.00 5.00 |

| Market Presence | 0% | 4.30 1.00 3.20 3.80 3.20 4.50 1.70 2.50 1.20 4.30 4.50 1.00 4.30 1.50 4.20 |
| Product revenue | 35% | 3.00 1.00 3.00 3.00 3.00 5.00 1.00 3.00 1.00 5.00 5.00 1.00 5.00 1.00 5.00 |
| Customer base | 30% | 5.00 1.00 3.00 5.00 3.00 5.00 1.00 3.00 1.00 5.00 5.00 1.00 5.00 1.00 3.00 |
| Market awareness | 25% | 5.00 1.00 3.00 3.00 3.00 3.00 3.00 1.00 1.00 3.00 3.00 1.00 3.00 3.00 3.00 |
| Partners | 10% | 5.00 1.00 5.00 5.00 5.00 5.00 3.00 3.00 3.00 3.00 5.00 1.00 3.00 1.00 3.00 |

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

**Leaders**

› **Hortonworks delivers a viable open source BDW platform.** Hortonworks delivers actionable intelligence from all kinds of data-in-motion and data-at-rest. Through its open source strategy, Hortonworks continually evolves its offering by working closely with partners across the EDW ecosystem of tools and vendors. The vendor provides a cost-effective, nimble, and scalable architecture to implement big data warehouses, whether on-premises or in the cloud. All of the technology built into the Hortonworks Data Platform is an Apache open source project. Enterprises
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

like Hortonworks’ storage and compute processing, broad data ingestion, data governance, and open source support when deploying BDW, but they claim it lags in data transformation and data modeling capabilities.

› **IBM’s BDW strategy evolves its focus on the cloud, big data, and real-time analytics.** IBM’s significant DB2 and data management solution install base and large professional services organization give it a competitive advantage. IBM’s hybrid data warehouse strategy includes IBM dashDB, IBM’s Fluid Query technology, Big SQL, IBM InfoSphere Information Server, IBM BigInsights, and IBM DB2 (including BLU acceleration). IBM is also deeply committed to open standards and interoperability and is an active contributor to various Hadoop, Spark, and other projects. The vendor offers advanced data compression, in-database analytics, real-time streaming, automated resource management, preconfigured vertical data models, virtualization governance, scalable appliances, and native integration with Spark and Hadoop platforms. Enterprises like IBM’s performance, security, data transformation, governance, and integration capabilities. However, some claim that its high-end scale and self-service capabilities are lagging.

› **Oracle expands its data warehouse to support cloud, real-time, and big data.** Oracle’s dominant position in the database market and its growing focus on big data platforms and appliances give it a competitive edge. Oracle offers enterprise architects the ability to deploy BDW on several form factors, including on an integrated Exadata appliance, the cloud, virtualized offerings, and commodity servers. Oracle’s Big Data SQL focuses on querying against Oracle, Hadoop, and other data repositories. The vendor continues to enhance its EDW solution by investing in storage optimization, data virtualization, compression, in-memory, cloud, hardware optimization, and automation. Enterprises like Oracle’s data warehouse performance, Big Data SQL, storage and compute processing, and data ingestion capabilities. However, they claim it lags in agile integration with big data sources such as Hadoop and NoSQL, but that’s changing. Organizations use Oracle’s solution for big data analytics, customer analytics, and self-service insights.

› **SAP’s BDW focuses on real-time analytics, scalability, and integration.** Enterprises use SAP HANA for in-memory data marts and SAP BW/4HANA implementations that integrate with other data warehouses, including SAP IQ. Enterprises also use SAP Cloud Platform Big Data Services (formerly Altiscale) for cloud-based data lakes. SAP’s key differentiators are its shared-nothing, distributed in-memory data platform for real-time analytics; optimized data streaming and query processing; integrated data services layer; advanced compression; and security. In addition, SAP Vora extends the Apache Spark execution framework to provide enriched interactive analytics. The combination of SAP HANA and SAP Vora delivers a powerful BDW capability that brings together in-memory, Hadoop, data warehouses, integration, streaming, and scalability to support large-scale, real-time analytical requirements. Enterprises like SAP’s performance for small to midsized BDW deployments as well as its data integration, data modeling, and storage processing capabilities, but they claim that it lags in high-end scale and simplified administration for large deployments.
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

- **Teradata expands its EDW to support cloud, in-memory and big data.** Teradata continues to expand its reach with Teradata Everywhere, providing the same software across all deployment options to give clients choice and flexibility. QueryGrid offers the ability to push query processing and data to any platform in the analytical ecosystem. Enterprises leverage Teradata with Amazon EMR and S3 storage together with Spark and Presto. Teradata offers advanced capabilities, including in-database analytics, distributed query processing, self-service, automation, workload management, and broader security. Enterprises like its ease of use and deployment, performance, scalability, storage, and compute processing capabilities, but they cite data governance and data modeling as a concern. Teradata recently changed its licensing strategy to provide subscription-based pricing and license portability through a simplified model. The vendor remains a prominent choice, especially for hybrid deployments where scalability and availability are critical.

- **Amazon Redshift has the largest adoption of BDW in the cloud.** With more than 5,000 deployments, Amazon Redshift has the largest data warehouse deployments in the cloud — some over 10 petabytes in size. Amazon Redshift is an MPP architecture that distributes nodes in a scale-out configuration. AWS’s key strengths lie in its dynamic scale, automated administration, flexibility of database offerings, good security, and high availability (HA) capabilities, which make it a preferred choice for customers. Enterprises like Amazon Redshift’s ability to quickly resize and size warehouse for large petabyte-sized data warehouse deployment, low subscription cost, and integration with other AWS services to support elastic search, interactive query, real-time streaming, backup, recovery, and data ingestion. However, organizations claim it lags in query and workload scalability, data modeling, and data integration services. AWS’s road map focuses on further improvements in automation, performance, and security.

**Strong Performers**

- **Microsoft’s BDW adoption grows with a strong focus on the public cloud.** Microsoft’s big data warehouse platform includes SQL Server 2016, Analytics Platform System (APS), Azure SQL Data Warehouse, and HDInsight for Hadoop. Although Azure SQL Data Warehouse can scale to petabyte-sized deployment, most deployments are typically around tens to hundreds of terabytes. Enterprises can deploy Microsoft SQL Server flexibly in diverse EDW and business intelligence (BI) topologies as well as integrate the database management system with the vendor’s diverse application platform, middleware, BI, performance management, and desktop software solutions. Microsoft offers sophisticated caching, compression, partitioning, indexing, cost-based query optimization, and workload management functionality for both cloud and on-premises platforms. Enterprises like its low cost, automation, and support for hybrid (on-premises and cloud) platforms. However, some have concerns around security, data integration, and high-end scale and performance.

- **Cloudera’s BDW strategy is evolving rapidly to challenge established EDW players.** The first commercial Hadoop startup, founded in 2008, Cloudera started the SQL-for-Hadoop craze with Impala, and its pace of innovation has been steady ever since. Cloudera delivers a data
management and analytics platform built on Hadoop and the latest open source technologies, available both on-premises and in cloud environments. Enterprises capture, store, process, and analyze vast amounts of data, empowering them to use advanced analytics quickly, flexibly, and at lower cost. Enterprises have successfully deployed a wide range of use cases, and they like Cloudera’s scalability, performance, technical support, HA, and storage and compute processing capabilities when deploying BDW. But they claim it lags in data modeling, security, and governance. Cloudera leverages Apache Hadoop, Spark, Impala, HBase, Kudu, and Kafka to support low-latency, high-concurrency interactions that deliver real-time insights from big data.

› Hewlett Packard Enterprise gains momentum in the BDW market. HPE Vertica Analytics Platform is an MPP, shared-nothing data warehouse that is scalable to support large workloads. Featuring an analyze-in-place unified architecture, HPE offers good high-end scale, performance, advanced compression, advanced indexing, automatic design and tuning, workload management, distributed query optimization, and automation. HPE’s BDW includes the HPE Vertica analytics database integration with Hadoop, Kafka, and Spark as well as connectors to various data sources. HPE continues to expand its services features in dedicated and multitenant cloud BDW deployments on AWS and Microsoft Azure. Enterprises like the vendor’s performance, especially in join operations, technical support, HA features, and high-end scale capabilities. However, some claim that it lags in integration with new APIs in heterogeneous environments, high-speed ingestion, and data governance.

› MarkLogic’s data warehouse offering is expanding to support broader use cases. MarkLogic has been powering large enterprises’ information management strategies for more than a decade. Its data warehouse solution enables enterprises to store, process, and access any kind of data for analytics, insights, and operational reporting. MarkLogic can run on-premises or in the cloud on AWS, Azure, and Google Cloud Platform. Enterprises use the solution for various use cases, including big data analytics, real-time analytics, fraud detection, operational analytics, and integrated analytics. Enterprises like MarkLogic’s data integration, performance, query processing, HA, low cost, technical support, and data security capabilities. However, some cited concerns around the depth of third-party vendor tools to support ETL, data governance, and data modeling.

› Snowflake offers a viable, low-cost, high-performance, and scalable BDW solution. Snowflake delivers an elastic BDW-as-a-service in the cloud, offering a highly automated platform to simplify ingestion, transformation, processing, and administration. Snowflake scales up and down to deliver computing resources as needed, depending on the query or workload requirements. The vendor combines the flexibility of big data platforms, cloud elasticity, data warehouse features, and workload optimization to deliver a viable BDW solution for agile, ad hoc, and real-time analytics. Enterprises like Snowflake’s low cost, separation of storage from compute, scalability, administration, and performance, but they claim that its tooling, data modeling, and data governance are lagging.
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

MapR delivers a high-performance and scalable BDW platform. MapR delivers a converged data platform that enables customers to leverage big data by combining real-time analytics with operational applications to support actionable insights. The MapR platform processes distributed files, database tables, and event streams in one unified layer that lets organizations support both operational and analytical apps in one cluster. MapR's road map focuses on improved automation, self-service, and further support for real-time analytics and cloud platforms. Enterprises use MapR for various use cases, including big data analytics, operational insights, text analytics, customer insights, and real-time insights. Enterprises like the vendor's capabilities around storage and compute processing, performance, and data ingestion as well as its ease to manage and scalability. However, some claim that MapR lags in data modeling, complex query processing, and data transformation.

Contenders

MemSQL focuses on real-time analytics with its big data warehouse solution. MemSQL is a distributed, in-memory data warehouse that delivers full ACID compliance and extreme performance to support transactional, operational, and analytical workloads. It uses SQL and a horizontally scalable, distributed architecture that runs on-premises and in the public cloud. With the combined power of database, data warehouse, and streaming workloads in one system, MemSQL turns insights into actions and supports self-service predictive analytics. Enterprises like its performance, data ingestion, HA, and transformation capabilities when deploying BDW, but they claim it falls short in comprehensive tooling, data integration, data security, and data modeling. MemSQL's road map includes improvements with Apache Spark; enabling further integration with cloud platforms; and improving performance, scale, and administration for larger in-memory deployments.

Phemi offers a viable warehouse solution, especially when security is critical. Phemi, founded in 2013, is a BDW company that transforms data into analytics-ready digital assets for researchers, analysts, and administrators in healthcare and public sector organizations. Phemi Central provides the ability to collect, curate, and consume any kind of data, and it offers high-speed data ingestion and processing to support real-time operational and BI applications. The vendor has a strong focus on privacy, security, and governance, with deep knowledge in the healthcare space. Customers like Phemi’s data ingestion and data governance for BDW deployments, but some claim it lags in data modeling, broad third-party tooling support and data integration, and simplified administration for complex deployments. Phemi’s road map includes performance optimization, automation, simplified metadata management, and more self-service capabilities.

Cazena focuses on making BDW simpler to use through intelligence and automation. Cazena is a big-data-as-a-service solution in the public cloud that focuses on workload intelligence, end-to-end automation, and security and governance. The vendor provisions and optimizes cloud infrastructure and data technologies such as Hadoop, MPP SQL, and Spark to support various analytics and insights. Cazena’s road map includes improved scale, compliance certifications,
Adoption Grows As Enterprises Look To Revive Their EDW Strategy

security, and an even higher degree of automation. It offers intelligent provisioning and continuous optimization to ensure queries and insights are delivered with quickly with minimum infrastructure resources. Cazena supports AWS and Microsoft Azure and is likely to support other cloud platforms in the future. Enterprises like the vendor’s automation, simplified analytical data management process, and support, but some are concerned about its lack of high-end scale and broad third-party vendor tooling to support the platform.

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 iPhone® and iPad®

Stay ahead of your competition no matter where you are.
Supplemental Material

Online Resource

The online version of Figure 3 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 three data sources to assess the strengths and weaknesses of each solution. We evaluated the vendors participating in this Forrester Wave, in part, using materials that they provided to us by December 21, 2016.

› **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 demos.** We asked vendors to conduct demonstrations of their products’ functionality. We used findings from these product demos to validate details of each vendor’s product capabilities.

› **Customer reference online survey.** To validate product and vendor qualifications, Forrester also fielded an online survey with at least two of each vendor’s current customers.

The Forrester Wave Methodology

We conduct primary research to develop a list of vendors that meet our criteria for evaluation in this market. From that initial pool of vendors, we 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 evaluation — and then score the vendors based on a clearly defined scale. We intend these default weightings to serve only as a starting point and encourage readers to adapt the weightings to fit their individual needs through the Excel-based tool. The final scores 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

We conduct all our research, including Forrester Wave evaluations, in accordance with our Integrity Policy. For more information, go to http://www.forrester.com/marketing/policies/integrity-policy.html.

Endnotes

1 See the Forrester report “The Next-Generation EDW Is The Big Data Warehouse.”
2 EDW: enterprise data warehouse.
3 ETL: extract, transform, load.
4 See the Forrester report “The Next-Generation EDW Is The Big Data Warehouse.”
5 MPP: massively parallel processing.
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. We offer quantity discounts and special pricing for academic and nonprofit institutions.