

(<http://www.gartner.com/home>)

LICENSED FOR  
DISTRIBUTION

# Magic Quadrant for Data Warehouse and Data Management Solutions for Analytics

25 February 2016 | ID:G00275472

**Analyst(s):** Roxane Edjlali, Mark A. Beyer

## Summary

Disruption is accelerating in this market, with more demand for broad solutions that address multiple types of data and offer distributed processing and repository options. Against this backdrop, this report will help data, analytics and IT leaders find the right vendor for their needs.

## Market Definition/Description

*This document was revised on 29 February 2016. The document you are viewing is the corrected version. For more information, see the Corrections ([http://www.gartner.com/technology/about/policies/current\\_corrections.jsp](http://www.gartner.com/technology/about/policies/current_corrections.jsp)) page on gartner.com.*

Organizations now require data management solutions for analytics that are capable of managing and processing internal and external data of diverse types in diverse formats, in combination with data from traditional internal sources. Data may even include interaction and observational data — from Internet of Things sensors, for example. This requirement is placing new demands on software in this market as customers are looking for features and functions that represent a significant augmentation of existing enterprise data warehouse strategies.

For this Magic Quadrant, a data warehouse or data management solution is defined as a complete software system that supports and manages data in one or many file management systems (most commonly a database or multiple databases) that can perform relational processing (even if data is not stored in a relational structure) and support access and data availability from independent analytical tools and interfaces.

Our definitions also state that:

The data warehouse (see Note 1) and data management solutions for analytics (DMSAs) are systems that perform the processing required to support analytics. They can be extended to support new structures and data types, such as XML, text, documents and geospatial data, and access to externally managed file systems. They must support data

availability to independent front-end application software, include mechanisms to isolate workload requirements, and control various parameters of end-user access within managed instances of data.

A data warehouse can comprise an entire DMSA, or be part of a larger system serving as a broader, more widely applied DMSA.

A DMSA is simply a system for storing, accessing and delivering data intended for a primary use case that supports analytics.

A DMSA is not a specific class or type of technology.

A DMSA may consist of many different technologies in combination. However, any offering or combination of offerings must, at its core, exhibit the capability of providing access to the files or tables under management by open-access tools.

A DMSA must support data availability to independent front-end application software, include mechanisms to isolate workload requirements (see Note 2) and control various parameters of end-user access within managed instances of data.

A DMSA must manage the storage and access of data in some form of storage medium, which may include (but is not limited to) hard-disk drives, flash memory, solid-state drives and even DRAM.

There are many different delivery models, such as stand-alone DBMS software, certified configurations, cloud offerings (public and private) and data warehouse appliances (see Note 3). These are evaluated together in the analysis of each vendor.

## Magic Quadrant

**Figure 1.** Magic Quadrant for Data Warehouse and Data Management Solutions for Analytics



HPE = Hewlett Packard Enterprise

Source: Gartner (February 2016)

## Vendor Strengths and Cautions

### 1010data

1010data (<http://www.1010data.com/>), which was acquired by Advance in August 2015, is a managed service data warehouse provider. Its integrated DBMS and business intelligence (BI) solution is aimed at the financial services, retail/consumer packaged goods, telecom, government and healthcare sectors.

**STRENGTHS**

1010data has continued to grow, and now has over 750 clients. The market has shown growing interest in cloud solutions and reduced concern around governance, which has benefited 1010data.

Reference clients praise 1010data's ease of use for interactive analysis, specifically its query performance, ease of data loading and analytical capabilities.

1010data's acquisition by Advance provides additional funding to enhance its offering at a time when competition is intensifying.

**CAUTIONS**

1010data has only a limited focus on the logical data warehouse (LDW) (see Note 4). It focuses mainly on supporting analytical needs for data managed in its own cloud.

1010data is still mainly U.S.-based. It has, however, recently opened a German data center for European operations, which is live and has customers.

Reference customers identify 1010data's overall integration with their ecosystem for BI and analytics, data access, and integration as requiring further enhancement.

**Action**

Action (<http://www.action.com/>) offers the Action Analytics Platform for data warehouse and data management solutions. The platform is composed of three products: Matrix, a massively parallel processing (MPP) DBMS engine; Vortex for analytics on Hadoop; and Vector, a symmetric multiprocessing (SMP) analytics DBMS.

**STRENGTHS**

Action offers an integrated solution supporting all four use cases (see Note 5) for data warehouses and DMSAs. In particular, it offers integration capabilities paired with analytical capabilities on top of Hadoop as part of Vortex.

Reference customers indicate that the technology is used for LDW and context-independent use cases combining data of diverse formats.

Overall, reference customers praise Action for its query performance and ability to support analytical capabilities.

**CAUTIONS**

While ParAccel (now Matrix) formed the basis of Amazon Redshift, both technologies are distinct and evolving separately. As a result, Action customers cannot expect to use both technologies in combination for a hybrid cloud and on-premises deployment. However, Action offers Matrix as a cloud service on third-party cloud service providers such as Microsoft (Azure) and Rackspace. Customers use Matrix deployed in the cloud for development and testing, combined with Matrix on-premises for production use.

Inquiries about Actian from end-user organizations received by Gartner in 2015 remained rare, which indicates that Actian has not increased its visibility in the market. However, Actian claims growth rates in excess of the market average.

The degree of completeness of Actian's solution was mentioned as an issue by reference customers. In particular, they identified its administration capabilities and distributed processing as posing challenges.

## Amazon Web Services

Amazon Web Services (<http://aws.amazon.com/>) (AWS) offers Amazon Redshift, a data warehouse service in the cloud, AWS Kinesis for streaming data, Amazon Simple Storage Service (S3) and Amazon Elastic MapReduce (EMR).

### STRENGTHS

AWS is often considered the leading cloud data warehouse platform-as-a-service provider. It continues to achieve strong adoption, driven by its broad acceptance of the cloud, flexibility, and agility from both a technical and a financial standpoint.

AWS supports a wide variety of use cases when its offerings are combined with other data management solutions. For example, our client interactions indicate adoption of S3 in support of data lakes, in combination with Redshift for analytics.

The vast majority of reference clients indicate that they plan to invest more in Redshift, which demonstrates continued satisfaction with this product. Strong scores for customer experience and rapid, significant market penetration are major contributors to AWS's position on the Ability to Execute axis.

### CAUTIONS

All the major vendors — IBM, Microsoft, Oracle, SAP and Teradata — are now actively competing with AWS in the cloud with varying degrees of support for true data warehouse platforms as a service. This growing competition on pricing and functional capabilities provides more cloud options for customers to choose from, but requires more careful scrutiny to truly compare offerings.

As AWS is a pure-play cloud vendor, Redshift lacks support for the hybrid cloud-and-on-premises data warehousing combinations that Gartner predicts will be the norm for most organizations by the end of 2018.

As AWS's reference clients mature in their use of Redshift, they are starting to report limitations in relation to their expectations for complex, mixed-workload management.

## Cloudera

Cloudera (<http://www.cloudera.com/>) provides a data storage and processing platform based on the Apache Hadoop ecosystem, as well as proprietary system and data management tools for design, deployment, operation and production management.

## STRENGTHS

Cloudera differentiates itself from other Hadoop distribution vendors by continuing to invest in specific capabilities, such as further improvements to Cloudera Navigator (which provides metadata management, lineage and auditing), while at the same time keeping up with the Hadoop open-source project.

Cloudera has successfully positioned its solution as a complement to the traditional data warehouse and made use of its relationships with traditional DBMS vendors, particularly Oracle.

Cloudera has continued to expand both geographically, with a growing number of European, Asian and Latin American customers, and through a strong network of partners for its full ecosystem.

## CAUTIONS

Although organizations have a growing interest in cloud deployments, Cloudera mainly addresses the cloud using an infrastructure-as-a-service approach that does not offer scalable, elastic and managed service support. However, Cloudera is addressing these needs with enhancements to Cloudera Director, to ease deployment of elastic clusters in the cloud.

Hadoop modularity enables new components to be added easily, and Cloudera continues to expand its set of components to meet new use cases and requirements. Although this approach enables Cloudera to deliver new capabilities without disrupting existing customers, it makes the overall landscape more challenging for clients to understand.

Although Cloudera has expanded into new geographies and added new clients, reference customers consider that the availability of support or professional service resources is becoming constrained. Cloudera has recognized this as an issue, and worked to address these points in 2015 by, for example, expanding its support team in Europe.

## Exasol

Exasol (<http://www.exasol.com/>) offers an in-memory column-store DBMS, which is available as a free single-node edition, a clustered solution and a Dell appliance. It is also offered as a fully managed solution on EXACloud and on third-party cloud service providers such as AWS, Microsoft (Azure) and Rackspace.

## STRENGTHS

Exasol continues to report consistent growth, with over 100 customers to date. Although its customers are still mainly based in Europe, Exasol is seeing uptake in the U.S.

Exasol introduced virtual schema development (for external data sources) and the use of script language containers, along with existing parallel distribution. This combination enables customers or partners to develop, deploy and execute their analytics algorithms on Exasol in any language (for example, R, Scala, Java, Lua and Python).

Exasol's reference customers praise its technology for offering value for money. They particularly appreciate its performance.

## CAUTIONS

Exasol suffers from a lack of market visibility. This is likely to remain the case throughout 2016 as the company has opted to expand outside Europe mainly via partners. Exasol scaled back its U.S. operations in 2015, despite recent successes there. However, it retained U.S.-based expertise for sales and customer support and appears positioned to re-enter North America in 2016.

Exasol's reference customers report a lack of deployment and life cycle management capabilities, such as cluster downsizing and SQL client functionality.

Exasol's reference customers report that documentation is at times insufficient and that the limited availability of market skills hampers adoption.

## Hitachi

Hitachi (<http://www.hitachi.co.jp/>) entered the data warehouse and DMSA market in 2014 with the Hitachi Advanced Data Binder (HADB). It is offered in three configurations, including desktop, "entry" model and "standard" model — priced and delivered on the basis of expected capacity, number of processor cores and amount of memory.

## STRENGTHS

Hitachi's roadmap focuses on addressing Japanese market demands. HADB is a high-speed, traditional analytics solution for structured data analytics that combines structured data and sensor data with a focus on industry use cases.

Hitachi customers consider massive volumes of trade data, sensor data and even geological data to be structured data for analysis with HADB. Specifically, HADB users benefit from "out of order" execution, which bypasses traditional, synchronous operations to increase the degree of parallelization for I/O processes.

Reference customers rate highly the support that Hitachi provides. They also report that their self-reliance is enhanced by, for example, the ease with which trace logs can be accessed.

## CAUTIONS

Hitachi offers the Japanese market a mature, efficient, but sometimes basic solution for data warehousing. In more recent developments, Hitachi has been pursuing engagements in North America.

Hitachi achieved limited growth in 2015, and this resulted in only a small number of production references. As a result, Hitachi barely qualified for inclusion in this Magic Quadrant.

Hitachi's positioning of HADB focuses on high-performance analytics for large volumes of structured data and does not address, by itself, the LDW approach. However, Hitachi markets its Pentaho acquisition as a federation offering, as an alternative to the LDW approach that is now established in the market.

## Hortonworks

Hortonworks (<http://www.hortonworks.com/>) offers the Hortonworks Data Platform (HDP) on Linux and Windows. It also offers Hortonworks DataFlow (HDF) on Linux on an on-premises basis and through various cloud providers. Hortonworks partners with Microsoft (for its Azure HDInsight service) for hybrid on-premises-and-cloud deployments. A free, laptop-capable sandbox version of HDP is available. Hortonworks did not provide information specifically for this evaluation. Gartner's analysis is therefore based on other credible sources, such as vendor briefings, publicly shared financial results and discussions with users of these products.

### STRENGTHS

In December 2014, Hortonworks became the first Hadoop distribution vendor to go public – an aggressive move, with a dedicated posture designed to prove the viability and relevance of Hadoop to enterprises. Hortonworks has publicly disclosed significant increases in new customers since going public.

Hortonworks has gained market traction with an increased number of recognized partners, including traditional DBMS vendors. It avoids direct competition with them, which is precisely in keeping with the role we expect Hadoop distributions to play in expanding the data warehouse.

Hortonworks' commitment to the Open Data Platform initiative supports the growth of new Apache Foundation projects. Hortonworks differentiates itself from other distribution vendors by taking a public leadership role in the open-source community.

### CAUTIONS

It will be a challenge for Hortonworks' to maintain its differentiation based on deep partnerships and integration with the larger analytics ecosystem, as vendors such as Teradata and Microsoft are also partnering with other Hadoop distribution vendors.

Hortonworks' financial reports show potential challenges with market adoption, given the enterprise-ready demands of a well-established data management for analytics market. That said, Hortonworks' financial reports and comments indicate that, in terms of financial performance, it is progressing as was planned when it became a public company.

Gartner clients report that, even with demand shifting toward open-source solutions, they do not select providers or solutions on the basis of this aspect alone. Implementers are consistently increasing pressure for features that already exist in commercially licensed



solutions. Hortonworks must make a concerted effort to expand the market's skill base for its solution.

## HPE

Hewlett Packard Enterprise's (<http://www.hpe.com/>) (HPE's) portfolio for addressing data management solutions for analytics, HPE Vertica, is based on the core Vertica DBMS, a column store analytic DBMS. It is available as a cloud solution, as a software-only option and as an appliance. It offers integration with Hadoop with HPE Vertica for SQL on Hadoop.

### STRENGTHS

Customers rate HPE Vertica highly in terms of value for money. This differentiates HPE from other major vendors in this market.

Reference customers use HPE Vertica for a variety of use cases and types of data, which demonstrates adoption by leading-edge customers. Additional investment in polyglot capabilities will fuel this trend.

HPE Vertica caters to the major market trends, with support for cloud delivery, the LDW (with Vertica SQL on Hadoop) and rich in-database analytics capabilities.

### CAUTIONS

HPE has satisfied customers, but we have seen no rise in the number of users of Gartner's client inquiry service who ask about HPE Vertica. This indicates that the vendor still faces a challenge to raise the product's visibility in the market. However, formation of HPE could help address this issue.

Reference customers indicate challenges with the overall administration and management of HPE's DBMS, although they also point out that it is gradually improving.

HPE Haven OnDemand offers a promising set of cloud data management and analytical services, but is separate from the Vertica offering, and as result demonstrates a fragmented strategy across the two. This will affect customers' ability to use both in combination, particularly in cloud and hybrid cloud-and-on-premises deployments.

## IBM

IBM (<http://www.ibm.com/>) offers stand-alone DBMS solutions, data warehouse appliances, a z/OS solution, and a Hadoop distribution with BigInsights. Its appliances include the IBM PureData System for Analytics, the IBM PureData System for Operational Analytics, the IBM DB2 Analytics Accelerator (IDAA) and the IBM Smart Analytics System. IBM offers IBM DB2 with Blu Acceleration, as well as data warehouse managed services. It brought dashDB (a cloud data warehouse service) in October 2014.

### STRENGTHS

IBM has rolled out dashDB and DataWorks as cloud offerings. These give customers the opportunity to rapidly deploy analytic models and data in an elastic environment. They address the growing demand for cloud solutions.

In 2015, IBM introduced IBM Fluid Query, with connectors to relational and NoSQL sources and polyglot support. It enables access to, and processing across, a wide variety of environments.

IBM's commitment to the Apache Spark open-source project will bring value to IBM's products by enabling streaming, machine learning and advanced analytics. It may also help Spark mature faster in terms of technology and availability of skills.

## **CAUTIONS**

The level of adoption of dashDB is uncertain. IBM reports significant adoption, but Gartner has received few client inquiries about this offering. Cloud adoption is frequently preceded by experimentation and pilot efforts, but after five quarters, inquiries ought to be more numerous.

IBM's market share, growth and net revenue continue to fluctuate in year-over-year comparisons. This might be interpreted in many ways. However, we think it is currently the result of retrenchment as IBM repositions its offerings for the modern cloud and data management solution markets.

Reference customers rated IBM in the lower third of surveyed vendors for overall value for money. Customers also identified pricing as an issue — some, for example, stated that IBM's cloud products are not priced competitively.

## **Infobright**

Infobright (<http://www.infobright.com/>) is a global company that offers a column-vectorized, highly compressed DBMS under a MySQL- or PostgreSQL-based API layer. It markets the commercial Infobright Enterprise Edition (IEE), for which there is a trial download.

## **STRENGTHS**

Customers report speed of processing as a differentiator of Infobright. They attribute this to Infobright's approach, and specifically commended its compression, load rates and lack of a need for indexing.

During 2014, Infobright underwent a restructuring that resulted in positive cash-flow changes. Particularly noteworthy were decisions that included reducing commitments to products that were determined to be unsustainable in terms of cost.

Internet of Things (IoT) data presents a clear opportunity for Infobright, and it plans various technological innovations in this area. It has augmented its Distributed Load Processor (DLP) to accommodate stream ingestion, which is critical to IoT adoption. Related areas such as telecommunications and networks also offer opportunities.

## CAUTIONS

According to some reference customers, some MySQL data types are not supported by Infobright's software, but this may reflect the use of different versions. The absence of certain features (such as vacuuming of tables) was noted, and the process of upgrading versions deemed cumbersome.

Traditionally, Infobright's customer base has focused on OEM business within the networking, telecommunications and advertising technology sectors. Infobright will continue to earn revenue in these sectors, but we expect its results to be mixed as new analytics demands have emerged over the past four years and the market continues to change.

Although Infobright's penetration of the IoT market is deepening, and it seems to be making the right adaptations, this market is also targeted by almost every other large data warehouse and DMSA vendor.

## Kognitio

Kognitio (<http://www.kognitio.com/>) offers the Kognitio Analytical Platform both as a software data warehouse DBMS engine and as an appliance. Kognitio delivers a cloud solution with Kognitio Analytical Services as a public or private cloud offering; it is also available on AWS's public cloud.

## STRENGTHS

Some of Kognitio's largest customers run traditional analytics solutions, like data marts and data warehouses, alongside Hadoop distributions (including Kognitio's analytical engine for data residing in Hadoop) in a single data management for analytics environment – the LDW.

With plans for coexistence on Hadoop nodes and integration with Apache Hadoop YARN and Kerberos, as well as JSON parsing, Kognitio is likely to enhance its appeal as the universal processing engine for distributed analytics processing.

Kognitio continues to focus strongly on technology. It is bringing to market capabilities tailored to the emerging demands of a few leading-edge customers, such as those requiring an analytical engine on Hadoop distributions. This includes external scripting to support R, Python, Java and C to enhance distributing processing approaches.

## CAUTIONS

Building market presence and sales remains a challenge for Kognitio. Skills in its software are hard to find, the community being very small. Reference clients suffer due to Kognitio's small size. They remarked, for example, that Kognitio's small installed base and limited vendor community make it challenging to build a support network and identify best practices.

Kognitio's reference customers identified a lack of integration with many third-party BI and data integration products.

Reference customers also identified database administration issues, such as with memory management. In addition, they observed that the administrative UI needs further improvement.

## MapR Technologies

MapR Technologies (<http://www.mapr.com/>) offers a Hadoop distribution with performance and storage optimizations, high-availability improvements, and administrative and management tools.

### STRENGTHS

Clients praise MapR for its enterprise-readiness, high availability and cluster management. MapR has continued to enhance these capabilities with the addition of authorization and auditing features.

MapR focuses on addressing a wide range of use cases. It supports streaming, operational and analytical use cases, all from the same platform, with multimodel support and SQL capabilities across all models.

MapR has been expanding further across the world. It has engaged in more partnerships and joint solutions with major players, such as AWS, Google, HPE, IBM, Microsoft, SAP, SAS and Teradata.

### CAUTIONS

Despite MapR's progress in this market, it still suffers from a lack of market visibility, as indicated by Gartner's receipt of fewer inquiries about MapR than about other Hadoop distribution vendors.

MapR reference customers report themselves to be generally advanced. Actual users appear to be advanced types who create analyses that are deployed as complete data products for use in production. Direct use of the technology remains the preserve of skilled adopters.

MapR's reference customers identified challenges with upgrades and installations. To address these challenges, MapR now offers GUI installers, playbooks, a patch installer and professional service upgrade packages.

## MarkLogic

MarkLogic (<http://www.marklogic.com/>) offers a NoSQL database that uses XML, JSON, text, Resource Description Framework (RDF) triples and binary storage, and provides a strong metadata-driven semantic entity management layer. The product includes indexes, tiered storage, Hadoop Distributed File System (HDFS) support, Amazon S3 support, mobile replication, full-text search, geospatial capabilities, and SQL/ODBC support.

## STRENGTHS

MarkLogic's growth has increased dramatically in the past three years, with increases in both revenue and customers. This is a key reason, along with product enhancements, for its improved position on the Ability to Execute axis. Its customers come from across the globe — North America, China, Singapore, South Korea, France and Eastern Europe, among other places.

MarkLogic's customers use its product when various forms of semistructured data assets (and more structured ones) are involved. Recent additions to optimization approaches — namely the introduction of statistics on use cases and how they relate to the data under management — have increased MarkLogic's administrative and design capabilities.

Customers use MarkLogic's software as an intervening data layer to support data delivery in Web services-centered architectures and deployments. They consider some of the functions offered — such as support for atomicity, consistency, isolation and durability (ACID) on NoSQL and integrated full-text indexing — to be unexpected bonuses.

## CAUTIONS

MarkLogic is a small vendor with some distinctive capabilities that are attracting increased demand, but it is competing in a large market in which much bigger vendors are starting to deploy, and seeking to improve, similar solutions.

MarkLogic's reference customers identified a lack of qualified expertise in the market, which represents a risk to growth in user implementations. They are frustrated by functionality gaps (a lack of declarative joins for optimization analysis) and a lack of mature management/administrative tools — but MarkLogic has plans to remedy these shortcomings.

Although its year-over-year results are much improved, MarkLogic remains dependent on implementers and word of mouth to spread awareness of its platform. MarkLogic is not mentioned by large numbers of clients during interactions with Gartner.

## MemSQL

MemSQL (<http://www.memsql.com/>) offers an in-memory DBMS for transactional use cases, combined with a disk-based column store for analytics.

## STRENGTHS

MemSQL's strategy focuses on supporting transactional and analytical use cases with low-latency requirements. Integration with Apache Spark enables stream ingestion, transaction processing and analytics.

MemSQL entered the DBMS market in 2013, addressing operational use cases. It is now entering the data warehouse and DMSA market by addressing operational analytics use cases involving data from the transactional application, as well as other sources.

Some of MemSQL's reference customers indicated that they use its technology for data warehousing, which is unusual for a vendor that originated in the operational DBMS space. This indicates that its technology is quite flexible.

## CAUTIONS

Reference customers identified MemSQL's pricing as an issue and ranked it in the bottom third of the surveyed vendors for overall value for money. However, MemSQL does offer a free community edition, and it does not charge for disk storage in its column store.

MemSQL's positioning, which caters to hybrid transactional/analytical processing (HTAP) use cases allowing both transactional and analytical processing, addresses only a nascent demand stemming from the operational DBMS market.

MemSQL's offering can be used for analytical-only use cases, such as the traditional data warehouse. But clients' expectations for analytical use cases differ from those for operational use cases. Attempting to address analytical-only use cases could prove challenging, as it could distract MemSQL from focusing on excelling in the operational DBMS market and dilute its product and marketing execution efforts.

## Microsoft

Microsoft (<http://www.microsoft.com/>) markets SQL Server, a reference architecture, the Microsoft Analytics Platform System (which combines SQL Server Parallel Data Warehouse and HDInsight), a cloud offering with Azure SQL Data Warehouse (in preview) and Azure HDInsight for Hadoop.

## STRENGTHS

Microsoft's cloud-based solutions will include an analytics data environment with Azure SQL Data Warehouse. Although Azure SQL Data Warehouse is still in preview, Microsoft's solution has already attracted strong customer interest. In addition, the broad choices for deployment and Microsoft's "complete solution" approach pose a challenge to traditional vendors with strengths in cloud and user-driven, self-service capabilities.

Microsoft Azure Federated Query, Azure Machine Learning and Azure Data Factory introduce a growing self-service analytics data management environment to Microsoft users. With the additional availability of the Azure Data Lake Store and Azure Data Lake Analytics, Microsoft's user community is being supported by this vendor's own brand of citizen scientist capability.

Microsoft's reference customers often commend its security, scale and familiarity. They also mention the T-SQL over PolyBase feature for structured and unstructured analytics, which includes embedded R and interoperability with open-source R. The Stretch Database feature in SQL Server 2016 supports on-premises deployments combined with the cloud.

## CAUTIONS

Reference customers identify occasional back-end and infrastructure issues. They also indicate that Microsoft's disaster recovery functionality, integration of SMP and MPP, and platform binding exhibit management and deployment issues at times.

Microsoft's market positioning for data management solutions for analytics in the cloud is unclear. Microsoft needs to clarify how it will simultaneously support the self-service audience and enterprise-class customers.

Use of the cloud does not necessarily lower overall costs over the life span of a data warehouse, though it does enable a different initial cost model and changes how costs scale. In general, the cloud presents a challenge to Microsoft's low-cost starting warehouse model, though Microsoft has begun to adapt with its cloud offering.

## MongoDB

MongoDB (<https://www.mongodb.com/>) offers an open-source document DBMS that provides a data model similar to JSON. The offering supports automatic sharding, failover, secondary indexes (including arrays), geospatial data and text search, as well as management tools. It is offered as a cloud service or as on-premises software.

### STRENGTHS

MongoDB is known for its strength in processing operational data. Its software is used frequently in an application-embedded manner. Analytic models are deployed into applications using data integrated from additional sources. As a result, MongoDB addresses HTAP use cases that provide significant benefit to those with operational analytics needs.

Reference customers underline MongoDB's strength in operational analytics. They report that flexibility — manifest in dynamic schema and schema evolution — is key. They also commend MongoDB's management interface, ease of tuning and easily surmounted learning curves.

MongoDB is embraced by application developers (who far outnumber data management practitioners).

### CAUTIONS

Although MongoDB has successful customer deployments for operational analytical needs, customers have difficulty identifying MongoDB as an enterprise analytics platform.

Although customers appreciate the design of MongoDB's offering and its ease of deployment, reference customers consider data management within MongoDB's DBMS to be a chore. They also reported issues with read/write lock management, authorization/authentication, replication and performance with network-attached storage.

MongoDB is embraced by application developers and architects, but it is difficult to find a data management expert willing to learn MongoDB's platform, primarily due to the lack of maturity of its data management functionality. MongoDB has, however, recently introduced schema discovery and validation, enhancements for query performance and indexing tools, more granular lock management and advanced security features.

## Oracle

Oracle (<http://www.oracle.com/>) provides the Oracle Database 12c, the Oracle Exadata Database Machine, the Oracle Big Data Appliance, the Oracle Database Exadata Cloud Service (including 19 Tier 4 data centers worldwide), the Oracle Big Data Management System, Oracle Big Data SQL and Oracle Big Data Connectors.

### STRENGTHS

Oracle customers report that the combination of hardware and software in a single solution eases deployment and management. Oracle customers are satisfied by the performance and stability of Oracle's engineered systems. In 2015, Oracle reported a surge in adoption of the Oracle Big Data Appliance, with significant growth across North America and EMEA, but also in Asia/Pacific.

The current version of Oracle Database includes SQL capabilities over JSON data and an in-memory columnar organized data store. In addition, Oracle introduced a software-on-silicon strategy in late 2015 to make use of specific instructions on processors by, for example, scanning highly compressed data at in-memory speeds.

Broad-based analytics offered by Oracle include the capability to use "any Hadoop" in combination with any Oracle Database 12c. Oracle's roadmap includes augmentation of its Hadoop capabilities with a global metadata directory that, combined with statistics, is intended to enable global query optimization and resource management. Additionally, Oracle's software can perform graph analytics (including for internal system optimization).

### CAUTIONS

Oracle reference customers and Gartner inquiry clients indicate they are concerned by the software's initial cost, maintenance cost and, therefore, overall value. More reference customers reported concerns about ease of use than was the case for other providers in this Magic Quadrant. Oracle does, however, serve a broad base of customers with a highly diverse range of skills.

Up until now, Oracle's customers have demonstrated a willingness to follow Oracle's cloud vision by combining on-premises data centers and using familiar tools with a cloud environment for cost control and deployment capability.

Gartner clients consistently report frustration with Oracle's contracts, negotiations and ongoing costs. However, Oracle customers span a wide spectrum of ongoing costs. At the same time as some say they intend to limit these costs, others are willingly increasing their



commitment to Oracle.

## Pivotal

Pivotal (<http://www.pivotal.io/>) is an independent software company owned by EMC, VMware and GE. It provides an open-source managed solution. Its products include Pivotal Greenplum and Pivotal HDB. It also offers services via Pivotal Labs. In addition, Pivotal combines and delivers these products through its Big Data Suite.

### STRENGTHS

Pivotal continues to challenge the market with the idea that distributed processing can take any form in any data management system. Its pursuit of balanced workload management across different data stores, and even the optimization of analytics processes across differing processing languages, may well be the future of data management for analytics.

Pivotal has the potential to capitalize on its position in relation to the new style of data management solutions (beyond the "repository only" style of data warehouse). It has potential in the LDW arena, but could also challenge the notion that a DBMS should be the central platform for analytics data management.

Pivotal's customers report a broad range of benefits, including speed, distributed processing capacity, scalability and high availability.

### CAUTIONS

Open-source DMSAs are changing from development platforms into management, governance and optimization functions. Open-source software-based data warehouse infrastructure may dominate the market after 2020, but Pivotal's open-source strategy is premature in 2016 (significantly, Pivotal reports that open-source inputs have already increased the pace of innovation in its offerings). IoT providers may ultimately absorb significant functionality from open-source providers into their own platforms.

Reference customers highlighted some important issues with Pivotal's software: high concurrency of access and query failure when a segment is down. Pivotal, however, provides references that indicate concurrency issues can be resolved with additional hardware resources. Pivotal indicates that complexity of analysis often demands greater availability of resources than mass concurrency.

Pivotal's open-source community is not ready to address a general lack of auditing tools, unexpected hardware-interfacing issues and weak documentation. Auditing is currently provided through a third-party solution (Datometry), but native tools are being improved. At present, the community requires more contributing experts, who are currently very rare for Pivotal.

## SAP

SAP (<http://www.sap.com/>) offers both SAP IQ and SAP Hana. SAP IQ, the first column-store DBMS, is available as a stand-alone DBMS. SAP Hana is an in-memory column store that supports operational and analytical use cases; it is also offered as an appliance, a cloud solution (SAP Hana Cloud Platform) and a reference architecture (SAP Hana tailored data center integration [TDI]). SAP also delivers SAP Business Warehouse (BW) on Hana.

## **STRENGTHS**

SAP has continued to succeed with data warehousing use cases. Adoption of SAP Business Warehouse (BW) on Hana is growing across its customer base.

SAP has continued to improve on the flexibility of its cloud deployment options and high-availability/disaster recovery capabilities.

SAP has continued to update its SAP Hana positioning by expanding on the role of Hadoop with SAP Hana Vora and other DBMSs, as part of an ecosystem that contributes to a more balanced LDW approach.

## **CAUTIONS**

Despite all of SAP's positioning efforts, SAP Hana is still adopted mainly by existing SAP customers. This presents SAP with a challenge to expand its penetration of the market for analytics data management. A lack of focus from SAP with regard to SAP IQ reduces its potential for growth outside the company's installed base.

Gartner clients continue to question the role of SAP BW as a true enterprise data warehouse. This is a frequent topic, specifically in light of SAP S/4HANA, as customers question both the need for a data warehouse solution and the role of SAP BW in their future application landscape.

Reference customers identified challenges with SAP Hana, such as functionality gaps, stability issues and a lack of available skills in the market.

## **Teradata**

Teradata (<http://www.teradata.com/>) 's offerings include a DBMS solution, data warehouse appliances and a cloud data warehouse solution. Teradata offers both traditional solutions and LDW solutions, which come under what Teradata calls the Unified Data Architecture (UDA). It offers a combination of tuned hardware and analytics-specific database software, which includes the Teradata Database (on various form factors of appliance), Aster Analytics and Hadoop via all three major distributions, as well as analytic consulting services.

## **STRENGTHS**

Teradata has continued to evolve its offerings to meet the demands of the changing market, with the Teradata Cloud, Teradata on AWS, and Aster Analytics on Hadoop, which is Aster Analytics running natively in Hadoop. This new approach offers more flexibility to clients and creates the opportunity to adopt Teradata products separately. This is important as organizations increasingly look for "best fit" engineering approaches.

With the introduction of Listener, Teradata is addressing the growing demand for streaming data, which will be increasingly important in the age of the IoT. This is also an important shift in terms of addressing market trends that expand analytical use cases, based on data at rest.

Reference customers continue to praise Teradata's software for its performance, scalability and workload management. They also value its feature-richness with QueryGrid and the UDA.

## CAUTIONS

Teradata's traditional appliance business is being challenged. After 35 years of dominance with a revolutionary appliance (one emulated by almost every other major vendor in the past 10 years), Teradata is shifting to software components and cloud computing at the same pace as its clients, but later than some of its competitors. Teradata has a reputation for eventually challenging all competitors, even when it is not the first to market.

Teradata is generally perceived as a stack vendor at a time when many organizations are looking to expand their technology portfolio piece by piece. Teradata offers components that can participate individually in analytics infrastructure. It also offers a comprehensive stack option through the UDA, both in the cloud and on-premises.

Teradata continues to face challenges in terms of sales execution, even though it is addressing market demands with new offerings. On the other hand, feedback from existing Teradata customers, gathered during interactions with users of Gartner's inquiry services and from Gartner surveys, indicates an above-average customer experience, especially for the delivery of professional services.

## Transwarp

Transwarp (<http://transwarp.io/?lang=en>) offers the Transwarp Data Hub (TDH), a full suite of Hadoop distribution components, which is supplemented by its SQL engine, machine learning, NoSQL search engine and stream processing.

## STRENGTHS

Although a young vendor, Transwarp has gained traction in the Chinese market. It has won 200 clients in less than 18 months.

Transwarp has a unique set of capabilities, such as its Inceptor SQL component based on Apache Spark, with Oracle SQL and PL/SQL compatibility supporting create, read, update, delete (CRUD) and ACID operations. This component is particularly praised by reference customers.

Reference customers indicated that they are very satisfied with Transwarp's product, as well as with the support and training that the company offers.

## CAUTIONS

So far, Transwarp operates in China only. That said, the size of the Chinese market, and its specific requirements, offers plenty of scope for Transwarp to expand.

Transwarp has yet to offer a cloud solution, although it indicates that the cloud is on its roadmap.

Reference customers pointed to some missing functionality, particularly with regard to administration and management, and highlighted a lack of skills in the market. However, across the whole spectrum of customer experience, Transwarp's customers awarded scores equal to the average for this market.

## Vendors Added and Dropped

We review and adjust our inclusion criteria for Magic Quadrants as markets change. As a result of these adjustments, the mix of vendors in any Magic Quadrant may change over time. A vendor's appearance in a Magic Quadrant one year and not the next does not necessarily indicate that we have changed our opinion of that vendor. It may be a reflection of a change in the market and, therefore, changed evaluation criteria, or of a change of focus by that vendor.

### Added

Hortonworks

MemSQL

MongoDB

Transwarp

### Dropped

None

## Inclusion and Exclusion Criteria

To be included in this Magic Quadrant, vendors had to meet the following criteria:

Vendors had to have DBMS software generally available for licensing or supported download for approximately one year (since 10 December 2014). We did not consider beta releases.

We used the most recent release of the software to evaluate each vendor's current technical capabilities. For existing data warehouses, and direct vendor customer references and reference survey responses, all versions currently used in production were considered. For older versions, we considered whether later releases may have addressed reported issues, but also the rate at which customers refuse to move to newer versions.

Product evaluations included technical capabilities, features and functionality present in the product or supported for download through 8:00 p.m. U.S. Eastern Daylight Time on 1 December 2015. Capabilities, product features or functionality released after this date could be included at Gartner's discretion and in a manner Gartner deemed appropriate to ensure the quality of our research product on behalf of our nonvendor clients. We also considered how such later releases might reasonably impact the end-user experience.

Vendors had to provide 10 verifiable production implementations that will exhibit generated revenue from 10 distinct organizations with data management solutions for analytics or data warehouses indicating they are in production. Revenue can be from licenses, support and/or maintenance. Gartner may include at its discretion additional vendors in cases of known use for classified but unspecified cases.

Customers in production must have deployed data warehouses that integrate data from at least two operational source systems for more than one end-user community (such as separate business lines or differing levels of analytics).

To be included, any acquired DBMS product must have been acquired and offered by the acquiring vendor as of 30 June 2015. Acquisitions after 30 June 2015 are considered legacy offerings and represented by a separate dot until publication of the following year's Magic Quadrant.

Support for the included data warehouse and data management for analytics product(s) had to be available from the vendor. We also considered products from vendors that control or participate in the engineering of open-source DBMSs and their support.

We included in our assessments the capability of vendors to coordinate data management and processing from additional sources beyond the DBMS, but continued to require a DBMS that met Gartner's definition — particularly regarding support for at least one of the four major use cases .

Vendors participating in the data warehouse and DMSA market had to demonstrate their ability to deliver the necessary services to support a data warehouse through the establishment and delivery of support processes, professional services and/or committed resources and budget.

Products that exclusively support an integrated front-end tool that reads only from the paired data management system did not qualify for assessment in this Magic Quadrant.

We also considered the following factors when deciding whether data warehouses and DMSAs were eligible for inclusion:

- Relational data management.

- Nonrelational data management.

No specific rating advantage was given with regard to the type of data store used (for example, relational DBMS, HDFS, key value, document, row and column).

Multiple solutions used in combination to form a DMSA were considered valid, but each solution must demonstrate maturity and customer adoption.

Cloud solutions (such as platform as a service [PaaS] offerings) were considered viable alternatives to on-premises warehouses. The ability to manage hybrid of on-premises and cloud solutions was considered advantageous for inclusion.

DMSAs were expected to coordinate data virtualization strategies for accessing data outside the DBMS, as well as distributed file and/or processing approaches.

Gartner analysts were the sole arbiters of which vendors and products were included in this Magic Quadrant.

For details of our research methodology, see Note 6.

## Evaluation Criteria

### Ability to Execute

Ability to Execute is primarily concerned with the ability and maturity of a vendor and its product(s). Criteria under this heading also consider a product's portability, its ability to run and scale in different operating environments (giving the customer a range of options), and the plurality of viable offerings answering diverse market demands. Ability to Execute criteria are critical to customers' satisfaction and success with a product, so customer references are weighted heavily throughout.

**Product/Service:** This criterion relates to increasingly divergent market demands — for traditional logical data warehousing, operational data warehousing and context-independent data management for analytics (for definitions, see "Critical Capabilities for Data Warehouse and Data Management Solutions for Analytics"). The largest and most traditional portion of the analytics and data warehouse market is still dominated by the demand to support relational analytical queries over normalized and dimensional models (including simple trend lines through complex dimensional models). Data management for analytics solutions are increasingly expected to include repositories, semantic data access (such as federation/virtualization) and distributed processing in combination — what is referred to in the market as LDWs. All the traditional demands of the data warehouse remain. Operational data warehouse use cases also exhibit traditional requirements, plus loading of streaming data, real-time data loading and real-time analytics support. Users expect solutions to become self-tuning, to reduce the number of staff required to optimize the data warehouse, especially as mixed workloads increase. Context-independent warehouses (CIWs) do not necessarily support mixed workloads, nor do they require the same level of mission-critical support. CIWs serve more as data discovery support or "sandboxes."

**Overall Viability:** This criterion includes corporate aspects, such as the skills of the vendor's personnel, its financial stability, its R&D investment, the overall management of its organization, and the likely persistence of a technology during merger and acquisition activity. It also covers a vendor's ability to weather market difficulties, which is crucial for long-term survival. Vendors are further evaluated on their capability to establish dominance in meeting one or many discrete market demands.

**Sales Execution/Pricing:** This criterion examines the price/performance and pricing models of the DBMS, and the ability of the vendor's sales force to manage accounts (judged by the feedback from our clients and feedback collected from a survey of reference customers). It also considers the market share of DBMS software. Also included is the diversity and innovative nature of packaging and pricing models, including the ability to promote, sell and support the product within target markets and around the world. Aspects such as vertical-market sales teams and specific vertical-market data models are considered.

**Market Responsiveness/Record:** Market demands change over time and track records are established over the lifetime of a provider. The availability of new products, services or licensing in response to recent market demands and the ability to recognize meaningful trends early in the adoption cycle is particularly important for this criterion. Diversity of delivery models, as demanded by the market, is also important for this criterion (for example, a vendor's ability to offer appliances, software solutions, data warehouse "as a service" offerings or certified configurations).

**Marketing Execution:** This criterion assesses a vendor's ability to generate and develop leads, channel development through Internet-enabled trial-software delivery, and partnering agreements (including co-seller, co-marketing and co-lead management arrangements). Also considered are a vendor's coordination and delivery of education and marketing events throughout the world and across vertical markets, and whether its participation in competitive situations is increasing or decreasing. In addition, this year, a vendor's events and educational activities are considered.

**Customer Experience:** Our assessment for this criterion was based on surveys of reference customers and discussions with users of Gartner's inquiry service during the previous six quarters. Also considered are a vendor's track record on proofs of concept, customers' perceptions of its product(s), and customers' loyalty to the vendor (this reflects their tolerance of its practices and can indicate their level of satisfaction). This criterion is sensitive to year-over-year fluctuations, based on customer experience surveys. Additionally, customer input regarding the application of products to limited use cases can be significant, depending on the success or failure of a vendor's approach in the market.

**Operations:** This criterion relates to the alignment of a vendor's operations, and how this enhances its ability to deliver. Aspects considered include field delivery of appliances, manufacturing (including identification of diverse geographic cost advantages), internationalization of product(s) (in light of both technical and legal requirements) and

adequate staffing. This criterion also considers a vendor's ability to support clients throughout the world, around the clock, and in many languages. Anticipation of regional and global economic conditions is also evaluated.

**Table 1.** Ability to Execute Evaluation Criteria

| Evaluation Criteria          | Weighting |
|------------------------------|-----------|
| Product or Service           | High      |
| Overall Viability            | Low       |
| Sales Execution/Pricing      | High      |
| Market Responsiveness/Record | Medium    |
| Marketing Execution          | Medium    |
| Customer Experience          | High      |
| Operations                   | Low       |

Source: Gartner (February 2016)

## Completeness of Vision

Completeness of Vision encompasses a vendor's ability to understand the functions needed to develop a product strategy that meets the market's requirements, comprehend overall market trends, and influence or lead the market, when necessary. A visionary role is necessary for the long-term viability of both a product and a company. A vendor's vision is enhanced by its willingness to extend its influence throughout the market by working with independent third-party application software vendors that deliver data-warehouse-driven solutions (for BI, for example). A successful vendor will be able not only to understand the competitive landscape of data warehouses, but also to shape the future of this field with the appropriate focus of its resources on future product development.

**Market Understanding:** This criterion assesses a vendor's ability to understand the market and shape its growth and vision. In addition to examining a vendor's core competencies in this market, we consider its awareness of new trends, such as the increased demand from end users for mixed data management and access strategies, the growth in data volumes (see Note 7), and the changing concept of the data warehouse and analytics data management. This criterion also takes account of a vendor's position regarding emerging



terminology such as "data lakes" and "multimodel" and "polyglot" data management. Understanding the different audiences for traditional data warehousing and the new approaches is crucial, as is a demonstrable track record of altering strategy and tactical delivery in response to both opportunistic segments of the market and broader market trends.

**Marketing Strategy:** This criterion relates to a vendor's marketing messages, product focus, and ability to choose appropriate target markets and third-party software vendor partnerships to enhance the marketability of its products. This criterion includes the vendor's responses to the market trends identified above and any offers of alternative solutions in its marketing materials and plans. Investor relations is becoming an important part of marketing strategy — not investor sentiment (which can run contrary to vendors' fiscal health), but vendors' management and responses to that sentiment (see "The Data Warehouse DBMS Market's 'Big' Shift" ).

**Sales Strategy:** This criterion encompasses all plans to develop or expand channels and partnerships that assist with selling. It is especially important for younger organizations as it can enable them to greatly increase their market presence while maintaining lower sales costs (for example, through co-selling or joint advertising). This criterion also covers a vendor's ability to communicate its vision to its field organization and, therefore, to clients and prospective customers. Also included are pricing innovations and strategies, such as new licensing arrangements, and the availability of freeware and trial software.

**Offering (Product) Strategy:** When viewed from a vision perspective, this criterion is clearly distinguished from product execution. We evaluate the roadmap for enhancing traditional data warehouse capabilities (including plans to address currently missing execution components). Also considered are expected functionality and any timetable for meeting new market demands. We looked for, among other things, roadmaps and development plans for the following: a semantic design tier; system and solution auditing and health management to ensure use case SLA compliance; static and dynamic cost-based optimization, with the potential to span processing environments and data structures; management and orchestration of multiple processing engines; and elastic workload management and process distribution. Our assessment also bore in mind that end-user organizations are taking a "best fit" engineering approach that requires vendors to allow their technology to be easily combined with that of other vendors.

**Business Model:** This criterion assesses how a vendor's model of a target market combines with its products and pricing, and whether the vendor can generate profits with this model — judging by its packaging and offerings. We consider reviews of publicly announced earnings and forward-looking statements relating to an intended market focus. For private companies, and to augment publicly available information, we use proxies for earnings and new customer growth — such as the number of Gartner clients indicating interest in, or awareness of, a vendor's products during calls to our inquiry service.

**Vertical/Industry Strategy:** This criterion reflects a vendor's ability to understand its clients. A measurable level of influence within end-user communities is important, as is certification by industry standards bodies. We regard a specific product or solution roadmap to support a targeted industry as evidence of focus.

**Innovation:** Vendors demonstrate innovation by developing new functionality, allocating funds to R&D and leading the market in new directions. This criterion also covers a vendor's ability to innovate and develop new functionality for accomplishing data management for analytics. Also addressed is the maturation of alternative delivery methods, such as infrastructure as a service (IaaS) and cloud infrastructures, as well as solutions for hybrid on-premises-and-cloud and cloud-to-cloud data management support. A vendor's awareness of new methodologies and delivery trends is also considered. Organizations are increasingly demanding data storage strategies that balance cost with performance optimization, so solutions that address the aging and temperature of data will become increasingly important.

**Geographic Strategy:** This criterion considers a vendor's ability to address customer demands in different world regions using direct/internal resources or in combination with subsidiaries and partners. It also evaluates a vendor's global reach and roadmap for addressing specific geographic regulatory requirements, particularly for cloud deployments. We consider a specific product or solution roadmap to support a targeted geographic region as evidence of focus.

**Table 2.** Completeness of Vision Evaluation Criteria

| Evaluation Criteria         | Weighting |
|-----------------------------|-----------|
| Market Understanding        | High      |
| Marketing Strategy          | Medium    |
| Sales Strategy              | Medium    |
| Offering (Product) Strategy | High      |
| Business Model              | Low       |
| Vertical/Industry Strategy  | Low       |
| Innovation                  | High      |
| Geographic Strategy         | Medium    |

Source: Gartner (February 2016)

## Quadrant Descriptions

### Leaders

The Leaders quadrant includes traditional, large vendors that have had to adapt to this highly changing market. A notable thing about these vendors is that all have decided to pursue all four use cases for data warehousing with at least an average maturity of execution and vision. The span of ratings for Leaders' Completeness of Vision in this Magic Quadrant is quite narrow, with very similar scores for all five vendors — although each has a distinctively different vision.

Leaders are responding to market disruption caused by Challengers and Visionaries over the past three years by adapting their offerings and bringing innovations to market. The largest are developing alternatives by extending their processing logic into the operating systems of remote processors or clusters or rationalizing workloads using metadata engines guided by utilization, capacity and service-level policy controls. All now have a cloud strategy that they are executing with varying degrees of success.

The data warehouse is possibly the largest data management system in most organizations. The market for them is therefore large in terms of revenue, trained professionals and a variety of data management solutions, ranging from simple to complex. A priority for Leaders is to retain traditional customers and help them grow their existing warehouses, while expanding their engagement by introducing multiple data management and analytic processing platforms. At the same time, they must ensure that their marketing strategy and messages inspire confidence at both the established and the emerging ends of the market. This dual requirement results in a wide separation in the Magic Quadrant in terms of Ability to Execute.

Leaders also face a challenge in that appliances are no longer differentiators. Although 45% to 55% of the market's customers are late or middle-era adopters who will still benefit from appliances, Leaders will have to accept shrinking profit margins.

### Challengers

Challengers have made some very careful decisions to secure their market positions. Some have focused on a particular deployment strategy, such as to offer the dominant cloud offering. Others have focused on specifically large data volumes in a pre-cloud, pre-big data environment with fully optimized solutions that dominate vertical delivery in regional areas.

There are two keys to success for Challengers in 2016. First, the robust Challengers that command a significant market in a particular industry or region will have to determine how much they want to grow, before designing a strategy to enable that growth. Second, they must avoid the temptation to expand their reach into other use cases before they are mature enough to satisfy their existing customers' support needs.

### Visionaries

Visionaries have enormous potential over the next three years. They are challenging the more mature Leaders with new concepts. Vendors from the operational DBMS market have entered the data warehouse and DMSA market, indicating that operational analytical use cases, as well as the performance of analysis over operations, are gaining market traction. However, their customers often report specific tuning and deployment issues, a lack of practices and skills in the market, and patchy success, with intermittent difficulties. The amount of white space in the Magic Quadrant between the Visionaries and the Leaders reflects the Visionaries' difficulties with execution. This is a wide market, with high expectations in terms of technology maturity and geographical, sales and marketing execution. It is therefore challenging for new entrants.

To become Leaders, Visionaries must demonstrate that they do more than disrupt the market: they must deliver true value to organizations. This requires them to have a growing number of customers in production — in order to prove that their functionality or architecture is of real value and not simply of interest for experimental implementations.

## Niche Players

In 2015, the Niche Players saw new entrants challenge and, in some cases, move swiftly past them. The overall broadening of vision in the market, with competing approaches, is creating openings for new vendors.

Niche Players generally deliver a highly specialized product with limited market appeal. A Niche Player might, for example, provide an exceptional data warehouse DBMS product, but one that is limited to a specific end-user community, region or industry. Although Niche Players' solutions' may have no such limitations, adoption is invariably limited.

## Context

The Leaders quadrant is populated mainly by large vendors that are relatively close to each other in Figure 1. The market has, however, continued to attract new vendors, which appear in the Visionaries and Niche Players quadrants. These emerging vendors will challenge the positions of the established vendors.

This market remains, therefore, in a state of significant flux. Disruption to this market is likely to continue throughout 2016 and into 2017, which will erode the installed bases of the large traditional vendors. When such disruption occurs, the entire market usually moves away from a single mature trajectory and splits in two in terms of vision and execution. The result is that over the next two years, until the end of 2017, this market is likely to be much more volatile, with changes in leadership a possibility.

This Magic Quadrant has a lot of white space in upper-right corner, indicating that the market continues to demand more innovation and better execution to address the needs of combined cloud and on-premises deployments, as well as cloud and big data combinations. Cloud data management scaling and workload issues have begun to emerge.

The next area of competition will be the race to deliver cloud solutions and hybrid cloud-and-on-premises offerings. Specialized cloud vendors and all the traditional vendors have already started working to deliver these offerings. In parallel, we are seeing appliances, which have been reliable sellers for many vendors, losing some of their appeal and becoming niche offerings.

Vendors are continuing to innovate and support new the use cases demanded by the IoT and digital business. For these, timeliness of access and analysis of data becomes more important than the volume and variety aspects of big data.

## Market Overview

The data warehouse and data management solution for analytics market continues to evolve. Customers now expect solutions that support all types of data for analytics and that take a coordinated approach. This demands different types of integrated solution and an interoperable services tier for managing and delivering data. Data lakes and the ability to manage streaming data are now being pursued by a growing number of organizations.

Data and analytics leaders must be aware of the market's evolution, and prepare hybrid technology platforms that expand the data warehouse beyond any current practice. This is especially important because the influence of the LDW has created a situation in which multiple repository strategies are now expected, even from a single vendor. Additionally, interest is growing in cloud solutions as alternatives to on-premises solutions, although we expect hybrid cloud-and-on-premises situations will become the norm.

Gartner notes the following key trends in this market.

**The definition of the data warehouse is expanding.** The term "data warehouse" does not mean "relational, integrated repository." The data warehouse is what we built to do that, but the new SLAs indicate that sometimes data should be integrated, and sometimes not. This new market demands a much broader data management solution for analytics. This is best explained by comparing two guiding architectural approaches (see below and "The Data Warehouse DBMS Market's 'Big' Shift" ).

**Enterprise data warehouse (EDW):** An integrated, subject-oriented, time-variant and physically centralized data management system mounted on hardware optimized for mixed workload management and large-query processing.

**Logical data warehouse (LDW):** An optimized combination of software and hardware that delivers a logically consistent, subject-oriented integration of time-variant data accessed via a centralized data management infrastructure. It uses repositories, virtualization and distributed processes in combination. The LDW is part of a larger movement to establish a wider market for DMSAs.

The  
eme  
pop  
the  
a

**The p**  
as m  
flexi  
for  
wh  
en  
a  
-

onloading of his

fit engineering preference from clients has supported the entrance of new vendors coming from the operational DBMS space (MongoDB and MemSQL), selected for their unique capabilities to perform analytics over operational data collected from diverse applications.

## Appendix

### Other Vendors to Consider

Gartner's Magic Quadrant process involves research on a wider range of vendors than appears in the published document. In addition to the vendors featured in this Magic Quadrant, Gartner clients sometimes consider the following vendors when their specific capabilities match the deployment needs (this list also includes recent market entrants with relevant capabilities). These vendors were not included in the Magic Quadrant either because they failed to meet the required number of production customers or because they entered the market too recently, but they can be valid alternatives to featured vendors. Unless otherwise noted, the information provided on these vendors derives from responses to Gartner's initial RFI for this document or from reference survey respondents. The following list is not intended to be comprehensive:

**LexisNexis.** The High-Performance Computing Cluster (HPCC) Systems is an open-source computing platform for big data processing and analytics that uses the data-centric ECL language to simultaneously query data and support that query with integrated/embedded analytics. Owing to ECL's capabilities, HPCC Systems can scale across very large datasets and grid computing clusters. The resulting two-stage processes scale out as MPP environments across commodity-class hardware/servers. HPCC Systems is offered primarily as a data management analytics solution as a service. The company is based in Alpharetta, Georgia, U.S. and is owned by the RELX Group (of London, Amsterdam and New York).

LexisNexis is not included in this edition of the Magic Quadrant as it did not have the required number of production, on-premises analytics customers.

**Snowflake.** Snowflake offers a data warehouse as a service via the cloud. It has a multicloud, logically integrated data architecture. The architecture uses an approach best described as separating the three "components" of a data warehouse — storage, processing and cloud services for interfacing and management of the warehouse. It supports Java Database Connectivity (JDBC), Open Database Connectivity (ODBC) and a Web user interface for access. The result is an ability to scale out three optimization regions separately, based on analytical demands. Snowflake is based in San Mateo, California, U.S.

Snowflake is not included in this edition of the Magic Quadrant as it entered the market too recently.

**ParStream.** ParStream was acquired by Cisco on 3 November 2015. Cisco uses ParStream's technology to support IoT use cases. ParStream offers a columnar, in-memory database that features a high-performance compression index on an MPP architecture. Version 3.0, released in 2014, provided broad support for SQL Joins and highly distributed query processing. ParStream achieved \$15.6 million in funding. Recently established partnerships with QlikView, Datawatch and other front-end tool suppliers provides it with visual data analysis capabilities for large datasets in combination with time series data. ParStream is based in Cologne, Germany, and has an office in Cupertino, California, U.S.

ParStream is not included in this Magic Quadrant as it did not have the required number of production customers.

## Acronym Key and Glossary Terms

|      |  |
|------|--|
| AWS  | Amazon Web Services                    |
| BI   | business intelligence                  |
| DMSA | data management solution for analytics |
| ELT  | extraction, loading and transformation |
| HDFS | Hadoop Distributed File System         |
| IaaS | infrastructure as a service            |
| IP   | intellectual property                  |
| LDW  | logical data warehouse                 |
| MPP  | massively parallel processing          |
| PaaS | platform as a service                  |
| SMP  | symmetric multiprocessing              |
| SSD  | source-system extracted data           |

## Note 1

### Gartner's Definition of a Data Warehouse

A data warehouse is a collection of data in which two or more disparate data sources can be brought together in an integrated, time-variant information management strategy. Its logical design includes the flexibility to introduce additional disparate data without significant modification of any existing entity's design.



A data warehouse can be much larger than the volume of data stored in the DBMS, especially in cases of distributed data management. Gartner clients report that 100TB warehouses often hold less than 30TB of actual data — that is, source-system extracted data (SSED).

## Note 2

### Expected Workloads

For the purposes of this evaluation, the workloads we expect to be managed by a data warehouse include batch/bulk loading, structured query support for reporting, views/cubes/dimensional-model maintenance to support online analytical processing, real-time or continuous data loading, data mining, significant numbers of concurrent access instances (primarily to support application-based queries at the rate of hundreds, if not thousands, per minute), and management of externally distributed processes.

## Note 3

### Gartner's Definition of a Data Warehouse Appliance

A data warehouse appliance consists of a DBMS mounted on specified server hardware with an included storage subsystem specifically configured for analytics use case performance characteristics. In addition, a single point of contact for support of the appliance is available from the vendor, and the pricing of the appliance does not include separate prices for the hardware, storage and DBMS components.

## Note 4

### Logical Data Warehouse Definition

The LDW is a new data management architecture for analytics combining the strengths of traditional repository warehouses with alternative data management and access strategies. It has seven major components:

- Repository management

- Data virtualization

- Distributed processes

- SLA management

- Auditing statistics and performance evaluation services

- Taxonomy and ontology resolution

- Metadata management

## Note 5

### Use Cases

1. *Traditional data warehouse* : Manages historical data coming from various structured sources. Data is mainly loaded through bulk and batch loading. It requires high capabilities for system availability, and administration and management, given the mixed workload capabilities for queries and user skills breakdown.
2. *Operational data warehouse* : Manages structured data that is loaded continuously in support of embedded analytics in applications, real-time data warehouses and operational data stores. Primarily supports reporting and automated queries, to support operational needs.
3. *Logical data warehouse* : Requires management of data variety and data volume for both structured and other content data types. Also includes other content data types, such as machine data, text documents, images and videos. Supports queries using data from sources other than the data warehouse DBMS alone.
4. *Context-independent data warehouse* : Has the capability to establish "schema on read" approaches for new and even existing data values, variant data forms and new relationships. Also supports search, graph and other advanced capabilities for discovering new information models. There are no specific performance requirements. This option is favored by advanced users such as data scientists and data miners, resulting in free-form queries across (potentially) multiple data types.

## Note 6

### Research Methodology for Magic Quadrant

Gartner uses multiple inputs to establish the positions and scoring of vendors in our Magic Quadrants. These are adjusted to account for maturity in a given market, market size and other factors. For this update of the Magic Quadrant, the following sources of information were used:

Original Gartner-published research, often utilizing our market share forecasts to establish the breadth and size of a market.

Publicly available data, such as earnings statements, partnership announcements, product announcements and published customer cases.

Gartner inquiry data collected from more than 16,000 inquiries conducted by the authors and the wider analyst community within Gartner during the previous 20 months. These inquiries provides inputs on, for example, use cases, issues encountered, license and support pricing, and implementation plans.

RFI surveys issued to vendors, in which they were asked to provide specific details about versions, release dates, customer counts and distribution of customers worldwide, among other things. Vendors could refuse to provide any information in this survey, at their discretion.

Surveys of reference customers (with almost 300 new responses added this year to four prior years of survey data). Vendors were asked to identify a minimum number of reference customers. Gartner augmented the vendor-provided population by adding Gartner inquiry client contacts as potential respondents. Responses were voluntary for all participants. These surveys included questions to confirm that customers were current license holders. Additionally (especially in the case of open-source utilization), customers provided information that confirmed the size and scope of their implementations. In addition, customers were asked to provide information about issues and software bugs, overall and specific sentiments about their experience of a vendor, the use of other software tools in the environment, the types of data involved, and the rate of data refresh or load. They were also asked about their deployment plans. Historical survey responses were used to identify trends only. Current-year survey responses were used for all the commentary in this Magic Quadrant.

Gartner customer engagements, in which we provided specific support, were aggregated and anonymized to add perspective to the other, more expansive research approaches.

It is important to note that this was qualitative research that formed a cumulative base on which to form the opinions expressed in this Magic Quadrant.

## Note 7

### Data Warehouse Data Volumes

The data warehouse volume managed in the DBMS can be of any size. For the purpose of measuring the size of a data warehouse database, we define data volume as source-system extracted data (SSED), excluding all data-warehouse-design-specific structures (such as indexes, cubes, stars and summary tables). SSED is the actual row/byte count of data extracted from all sources.

The sizing definitions of traditional warehouses are:

Small data warehouse — less than 5TB

Midsized data warehouse — 5TB to 40TB

Large data warehouse — more than 40TB

## Note 8

### User Categories

*Data scientist:* Expertise in statistics, abstract mathematics, programming, business processes, communications, leadership.

*Data miner:* Expertise in subject areas of data; uses statistical software and statistical models; is fully aware of computer processing "traps" or errors.

*Business analyst:* Uses OLAP and dimensional tools to create new objects; has some difficulty with computer languages and computer-processing techniques.

*Casual user:* Regularly uses portals and prebuilt interfaces; has minimal, if any, ability to design dimensional analytics.

## Evaluation Criteria Definitions

### Ability to Execute

**Product/Service:** Core goods and services offered by the vendor for the defined market. This includes current product/service capabilities, quality, feature sets, skills and so on, whether offered natively or through OEM agreements/partnerships as defined in the market definition and detailed in the subcriteria.

**Overall Viability:** Viability includes an assessment of the overall organization's financial health, the financial and practical success of the business unit, and the likelihood that the individual business unit will continue investing in the product, will continue offering the product and will advance the state of the art within the organization's portfolio of products.

**Sales Execution/Pricing:** The vendor's capabilities in all presales activities and the structure that supports them. This includes deal management, pricing and negotiation, presales support, and the overall effectiveness of the sales channel.

**Market Responsiveness/Record:** Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve and market dynamics change. This criterion also considers the vendor's history of responsiveness.

**Marketing Execution:** The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message to influence the market, promote the brand and business, increase awareness of the products, and establish a positive identification with the product/brand and organization in the minds of buyers. This "mind share" can be driven by a combination of publicity, promotional initiatives, thought leadership, word of mouth and sales activities.

**Customer Experience:** Relationships, products and services/programs that enable clients to be successful with the products evaluated. Specifically, this includes the ways customers receive technical support or account support. This can also include ancillary tools, customer support programs (and the quality thereof), availability of user groups, service-level agreements and so on.

**Operations:** The ability of the organization to meet its goals and commitments. Factors include the quality of the organizational structure, including skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and efficiently on an ongoing basis.

## Completeness of Vision

**Market Understanding:** Ability of the vendor to understand buyers' wants and needs and to translate those into products and services. Vendors that show the highest degree of vision listen to and understand buyers' wants and needs, and can shape or enhance those with their added vision.

**Marketing Strategy:** A clear, differentiated set of messages consistently communicated throughout the organization and externalized through the website, advertising, customer programs and positioning statements.

**Sales Strategy:** The strategy for selling products that uses the appropriate network of direct and indirect sales, marketing, service, and communication affiliates that extend the scope and depth of market reach, skills, expertise, technologies, services and the customer base.

**Offering (Product) Strategy:** The vendor's approach to product development and delivery that emphasizes differentiation, functionality, methodology and feature sets as they map to current and future requirements.

**Business Model:** The soundness and logic of the vendor's underlying business proposition.

**Vertical/Industry Strategy:** The vendor's strategy to direct resources, skills and offerings to meet the specific needs of individual market segments, including vertical markets.

**Innovation:** Direct, related, complementary and synergistic layouts of resources, expertise or capital for investment, consolidation, defensive or pre-emptive purposes.

**Geographic Strategy:** The vendor's strategy to direct resources, skills and offerings to meet the specific needs of geographies outside the "home" or native geography, either directly or through partners, channels and subsidiaries as appropriate for that geography and market.



(<http://gtnr.it/1KsfgQX>)

© 2016 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. or its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. If you are authorized to access this publication, your use of it is subject to the Usage Guidelines for Gartner Services ([/technology/about/policies/usage\\_guidelines.jsp](/technology/about/policies/usage_guidelines.jsp)) posted on [gartner.com](http://gartner.com). The information contained in this publication has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information and shall have no liability for errors, omissions or inadequacies in such information. This

publication consists of the opinions of Gartner's research organization and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice. Gartner provides information technology research and advisory services to a wide range of technology consumers, manufacturers and sellers, and may have client relationships with, and derive revenues from, companies discussed herein. Although Gartner research may include a discussion of related legal issues, Gartner does not provide legal advice or services and its research should not be construed or used as such. Gartner is a public company, and its shareholders may include firms and funds that have financial interests in entities covered in Gartner research. Gartner's Board of Directors may include senior managers of these firms or funds. Gartner research is produced independently by its research organization without input or influence from these firms, funds or their managers. For further information on the independence and integrity of Gartner research, see "Guiding Principles on Independence and Objectivity. (/technology/about/ombudsman/omb\_guide2.jsp)"

---

About (<http://www.gartner.com/technology/about.jsp>)

Careers (<http://www.gartner.com/technology/careers/>)

Newsroom (<http://www.gartner.com/newsroom/>)

Policies ([http://www.gartner.com/technology/about/policies/guidelines\\_ov.jsp](http://www.gartner.com/technology/about/policies/guidelines_ov.jsp))

Privacy (<http://www.gartner.com/privacy>)

Site Index (<http://www.gartner.com/technology/site-index.jsp>)

IT Glossary (<http://www.gartner.com/it-glossary/>)

Contact Gartner ([http://www.gartner.com/technology/contact/contact\\_gartner.jsp](http://www.gartner.com/technology/contact/contact_gartner.jsp))