

Magic Quadrant for Data Warehouse and Data Management Solutions for Analytics

Published: 12 February 2015

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

Entering 2015, the data warehouse has expanded to address multiple data types, processing engines and repositories. We now see a much wider separation in the Leaders quadrant. Enterprise architects and data warehouse managers must work with their CIOs to address the new demands and solution options.

Market Definition/Description

In 2015, organizations require solutions capable of managing and processing external data in combination with their traditional internal sources, and may even include data from the Internet of Things. This is creating new demands on the data warehouse market — for broader data management solutions for analytics, with features and functionality that represent a significant augmentation to existing enterprise data warehouse strategies.

In 2015, the data warehouse market continues its evolution of the past several years and now includes broader data management solutions for analytics. It now supports all types of data for analytics under a coordinated approach that will demand different types of integrated solutions and an interoperable services tier for managing and delivering data. Data warehouse managers, solution architects for analytics and CIOs establishing IT modernization strategies must take note of this change in direction, and prepare to meet it with hybrid technology platforms that expand the data warehouse beyond any of the current practices. This is especially important because the influence of the logical data warehouse (LDW, see Note 1) has created a situation in which multiple repository strategies are now expected, even from a single vendor. Strengths and cautions relating to a specific offering (or offerings), when noted by customers, are also made clear in the individual vendor sections.

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

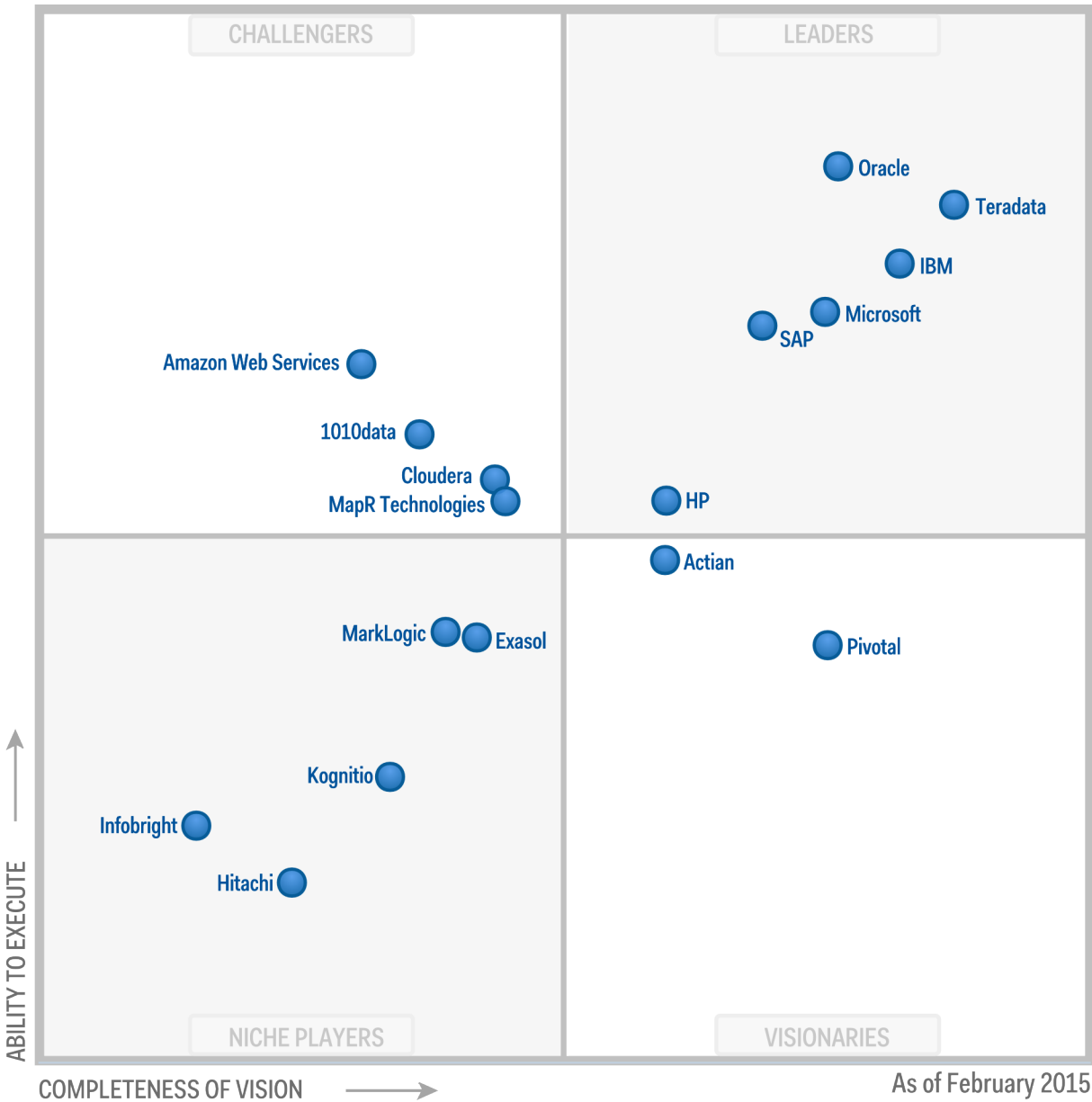
Furthermore, our definitions state that:

- The data warehouse (see Note 2) and data management solution for analytics (DMSA) are systems that can perform the processing required to support analytics and can be extended to support new structures and data types, such as XML, text, documents, geospatial 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 the data.
- A data warehouse can comprise an entire DMSA, or a data warehouse can 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. At the core, however, any vendor offering or combination of offerings must 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 3) and control various parameters of end-user access within managed instances of the data.
- A DMSA must manage the storing and access of data in some form of storage medium (which can include, but is not limited to, hard-disk drives, flash memory, and solid-state drives or even RAM).

There are many different delivery models, such as stand-alone DBMS software, certified configurations, cloud (public and private) offerings and data warehouse appliances (see Note 4). These are also evaluated together in the analysis of each vendor.

Magic Quadrant

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



Source: Gartner (February 2015)

Vendor Strengths and Cautions

1010data

1010data (www.1010data.com) was established 14 years ago as a managed service data warehouse provider with an integrated DBMS and business intelligence (BI) solution primarily for the financial sector, but also for the retail/consumer packaged goods, telecom, government and healthcare sectors.

Strengths

- 1010data has a long track record in specific vertical markets. As a result, it is effective in delivering to new and existing clients in its targeted vertical markets, with a high success rate and rapid deployment into production due to the nature of the offering. Potential customers should review 1010data's track record in their specific vertical or an adjacent or similar vertical.
- Clients report 1010data is a mature offering with fully all the features and functionality of a high-end analytics data warehouse and DMSA. This includes a nearly unanimous list of reference customers reporting high levels of satisfaction with speed/performance as well as a technically robust solution.
- 1010data also offers a certified configuration version, which can be an answer to penetrating markets outside of North America (95% of its customer base is located in North America) when governance of data becomes an issue.

Cautions

- 1010data's strength is delivering a data warehouse solution for vertical industries in the cloud. However, this approach does not support all of the four use cases identified for data warehouse and DMSAs. In particular, 1010data does not address the needs of the LDW.
- 1010data users report frustration with integration to existing systems and, more notably, with analytic tools. While 1010data does provide its own visualization, the proprietary nature of the DBMS has emerged as an issue for customers. Customers also report that the interfacing and tooling is associated more with expert implementers.
- End users do not follow the standard market distribution of users — across casual, business analysts, data analysts and scientists (see Note 5). Instead, 1010data is used more often by the skills-centered data analysts and data scientists, but also boasts support for the entire spectrum of users, including casual and business users.

Actian

Actian (www.actian.com) offers its Actian Analytics Platform for data warehouse and data management solutions. The analytics platform was first released in July 2013 and is the result of the integration of Vectorwise, ParAccel and Pervasive Software.

Strengths

- Actian provides integrated and complementary components for analytics data management, data integration and embedded analytics, with a portfolio nearly as extensive as those of so-called megavendors (such as IBM, Oracle, Microsoft and others), but at a lower price. This complete offering allows Actian to cover all of the four use cases for data warehouse and DMSA.
- In 2013, the Matrix DBMS (formerly ParAccel) management capability for the hardware, network and storage environment was significantly enhanced to allow Matrix to operate on mixed, commodity hardware. The evolution of that approach has continued aggressively and customers report ease of deployment, speed on any hardware, and scale as principle strengths.
- Actian's Matrix DBMS was the basis for Amazon's Redshift. The two systems have a very similar data management tier, but utilize different file systems. However, there is close compatibility between Redshift's cloud data warehouse market and Matrix.

Cautions

- Actian's data warehouse and analytics numbers remain small, if uncertain. It claims growth rates in excess of the overall market; yet Gartner inquiries of over 8,000 end-user organizations indicate no increased presence of Actian in the market. Actian indicates both new customers and organic growth within existing customers.
- Actian offers online training and customers can contract for skills training; however, the small number of customers means a dearth of skills in the market. Customer distribution is split almost evenly between North America and EMEA. For a vendor with this small share of the market, its ability to deliver in both regions is notable. There is a very small customer base in Asia/Pacific.
- While portability and operations on commodity-class hardware are touted by Actian, reference customers rate the ease of implementation as below average — with references to dependency on hardware configuration or workload management.

Amazon Web Services

Amazon Web Services (AWS) (aws.amazon.com) offers Amazon Redshift, a data warehouse service in the cloud, AWS Data Pipeline (designed for orchestration with existing AWS data services) and Elastic MapReduce (EMR).

Strengths

- Amazon debuted on the Magic Quadrant in 2014, with the highest customer experience rating among all the vendors. It has maintained high ratings throughout 2014 regarding reliability, cost and ease of implementation.

- AWS customers have multiple applications deployed in the cloud that require a fundamental capability to integrate disparate data and analyze it — also in the cloud. Amazon provides both integration and analysis and eliminates the need to move data back to on-premises. Amazon has introduced machine data or "operational technology" by insuring S3 is available to Redshift and EMR in a "data lake" approach (lakes can be components of an overall analytics environment), which creates an entrée into LDW delivery.
- With a presence in Europe (Frankfurt) and data warehouses deployed in the hundreds of terabytes, and even into the petabyte scale, Amazon has demonstrated significant execution in multiple verticals with significant complexity (for example, healthcare/life sciences, financial services and others).

Cautions

- Best practices have not yet emerged for data warehouse platform as a service; the talent pool in this market is effectively an adapted skill set, and hard to find. There is also a need for technical maturity, as noted in the user experiences of Amazon customers (see below).
- Amazon "owned" the cloud infrastructure as a service (IaaS) and platform as a service (PaaS) for supporting data warehousing during 2013 and 2014. However, with other vendors introducing IaaS and PaaS solutions that also provide data warehousing, Amazon will face greater competition and will have to maintain price pressure on the market while enhancing functionality.
- Users report high levels of satisfaction with Amazon overall, but go on to cite an absence of user-defined functions, some SQL functionality, administration and management tools and aggregation/statistical functions. Amazon does, however, have plans for enhanced functionality.

Cloudera

Cloudera (www.cloudera.com) provides a data storage and processing platform based upon an Apache Hadoop Project, as well as proprietary system and data management tools for design, deployment, operation and production management.

Strengths

- Cloudera is tightly focused on embedding customer experiences in emerging product functionality such as with Impala, Cloudera Director or Cloudera Navigator — and a focus on intellectual property in solutions and software instead of consulting and professional services. This is an important shift from its early days and adds value to its offerings.
- Unlike many emerging solutions, Cloudera benefits from a large set of BI and data integration partners. This eases adoption for organizations with existing investments in these tools. It has also been able to partner with large traditional players in this market such as Teradata or Microsoft.

- Customer references reinforce Cloudera's execution strategy, reporting excellent support and services, broad interoperability with industry standard tools, and a high-speed performance. This is specific to analytics use-case customers.

Cautions

- The combination of a lack of metadata management, available skills in the market and difficulties in finding standard approaches to loading the data make Cloudera adoption slow and sometimes painful — according to its references.
- The uneven mix of its user base, sometimes for repetitive batch run processing and at other times heavily loaded with data analyst and data scientist users, makes it difficult for Cloudera's solution to find a "home base" in the market for DMSAs. However, the user base is maturing along with the tools and this is expected to be a short-lived issue.
- With only the beginnings of growth in EMEA and Asia/Pacific, Cloudera remains a vendor with largely North American experience. We anticipate further growth in both these other large regions during the next two years.

Exasol

Exasol (www.exasol.com) is a small DBMS vendor headquartered in Nuremberg, Germany, and has been in business since 2000. Its first in-memory column-store DBMS, EXASolution, became available in 2004. EXASolution is still used primarily as a data mart for analytic applications, but also occasionally in support of LDWs.

Strengths

- Exasol has grown its customer installed base by 90% in 2014 (from 42 customers in 2013). This has been possible through the expansion of Exasol in the U.S., the U.K. and Brazil in addition to its European base.
- Exasol customer references constantly praise Exasol for its performance, because it is an in-memory columnar DBMS, but also for its price. Its ability to manage hybrid disk and memory persistency allows customers to adjust price/performance and hardware needs.
- Exasol is being used in a variety of use cases, not only for traditional data warehouse. Its support for in-database execution of Java, Lua, Python and R makes it appealing to customers requiring advanced analytics.

Cautions

- Despite its success in sales execution, Exasol mainly competes with other smaller or niche vendors and is not invited to the table with other megavendors, although its technology could very well compete.

- Reference customers point out Exasol's lack of partnerships in BI and data integration, making it harder to implement in the overall ecosystem. The addition of Tableau is important.
- A focus on the emerging context-independent data warehouse is promising but ahead of overall market demand, which will continue to make Exasol a Niche Player.

Hitachi

Hitachi (www.hitachi.co.jp) entered the data warehouse and DMSA market during 2014 with the Hitachi Advanced Data Binder (HADB). It includes a combination of a data integration OEM solution, deployment strategies and a DBMS. It is offered in three configurations including a desktop-deployed offering, an "entry" model and a "standard" model — priced and delivered based upon expected capacity, processor cores and memory.

Strengths

- The combination of data integration, content or unstructured analytics — via the use of an underlying Hitachi Content Platform instance (HCP) — and database management as a "stack," presents a solution that provides for the traditional data warehouse and limited aspects of the context-independent use case.
- HADB is being used by customers utilizing a diverse number of BI platforms (such as Tibco Spotfire and OpenText). Hitachi has BI partners including SAP Business Objects, Qlikview and others; as well as OEMing IBM's InfoSphere DataStage for data integration and partnering with multiple other vendors.
- Customers report a slightly above-average experience with HADB. References indicate performance, ease of development and quality of technical support are rated highly.

Cautions

- Currently, HADB's entire customer base is in the Japanese market. The quality of technical support is therefore unproven for consistency in other markets, and the availability of skills is currently lacking outside of Japan.
- References report issues with product maturity and some issues with mixed workload or multiple (use) processing. With a small customer base, it is difficult to determine the extent of these issues and whether they originate with the product or the customer's utilization of the product.
- The context-independent use case is complex in the Hitachi environment and is its primary use case. HADB has the capability to read open fields in databases. HCP can utilize Hadoop as a file and processing source and be accessed via HADB.

HP

HP's (www.hp.com) portfolio for addressing data management solutions for analytics is based on HAVEn. The HAVEn concept combines many acquisitions in analytics under one banner. HP's

offering is anchored by Vertica, a column-store analytic DBMS, Autonomy and Hadoop. Vertica is delivered as software for standard platforms (excluding Windows), and as a Community Edition (free for up to 1TB of data and three nodes). HP Factory Express (a predefined certified configuration) is available for Vertica.

Strengths

- Vertica is a robust column-vectorized solution. It is a system that combines big data assets with traditional back-office and business transactional analytics, which meets the dual vision of traditional data warehouse enhancements and support for more advanced analytic needs. A final word from user references — "speed" — HP Vertica is fast on query response.
- Vertica customers report positive ROI experiences. A small number of reference customers report staff savings estimates at significantly fast rates of return and high value (based on query and analytics support as well as deployment practices), and when combined with anticipated business benefits in productivity at least one of these large references claims ROI in less than one year.
- HP references expand their use beyond structured data. Users report analysis of social media data, image and video data, machine data and document/text analytics. HP claims significant customer and revenue growth, year over year, but distinct revenue numbers are not officially reported.

Cautions

- Vertica references mention production issues. Users report issues with query optimization, workload distribution, system stability — plus incomplete administrative and production controls. However, HP Vertica's 7.1 release features advancements in dynamic mixed workload management, the management console and enhancements to query optimization and performance.
- Vertica has experienced uneven delivery in the market, but is starting to stabilize its growth curve. Customers indicate lowered levels of satisfaction with HP Vertica during 2014, compared with 2013, and an increase in issues reported that underscores the need to stabilize the customer experience.
- Vertica is offered as a stand-alone product, but also as an integral component of other HP products. When viewed as a stand-alone product, the growing scope of its synergy with other HP products is not easily discerned. Customer reports of production issues are concerning, because HP's "other" software is IT operations — suggesting that the Vertica product is not benefiting in a reciprocal fashion from the rest of HP.

IBM

IBM (www.ibm.com) offers stand-alone DBMS solutions, data warehouse appliances and a z/OS solution. IBM has also recently brought to market its cloud data warehouse service dashDB (BLU with PureData with Cloudant). Its various appliances include: IBM zEnterprise Analytics System,

PureData System for Analytics and Operational Analytics, IDAA, IBM Smart Analytics System and others. IBM offers data warehouse managed services.

Strengths

- IBM demonstrates a broad offering and integration across products that can support all four major data warehouse use cases. In addition, IBM's approach to the cloud (which includes a solution for data integration) and transformation in the cloud sets a new tone for analytics in the market.
- IBM has continued to invest in product innovation and meeting emerging customer and market demands such as the delivery of its data warehouse PaaS offering dashDB, including in-memory columnar capabilities and integrated with its Cloudant NoSQL database (also delivered as a cloud service), as well as PureData for in-database analytics.
- Customer references report higher levels of satisfaction than in the past regarding pricing and value for money, which demonstrates IBM's intention to compete more aggressively on price.

Cautions

- Overall, IBM suffers from complex marketing and branding that confuses the market. For example, the rebranding of products under Watson Foundations and "cognitive" causes some confusion because it is unclear what value Watson brings relative to customers' demands.
- By the end of 2013, IBM's worldwide market share ranking for database software dropped to No. 3. While the data warehouse market is extending to include DMSA, the relational DBMS portion of the market constitutes over 90% of the combined market. During 2014, Gartner inquiry clients expressed unease regarding IBM's commitment to traditional database delivery. Organizations need to carefully decide, when taking a best-fit engineering approach, what combination of technology aligns with IBM's offering.
- IBM customers are unclear how to move from traditional IBM offerings — that combine software licenses, appliances (PureData, formerly Netezza) and on-premises concepts — to modern infrastructure deployment (such as "data refinery"). IBM has been driving innovation, addressing new demands for data preparation, preserving governance and differentiating from data-lake approaches; however, it remains unclear how existing data warehouse DBMS investments and new information approaches live side by side, or how to correctly identify the appropriate solutions for specific use cases.

Infobright

Infobright (www.infobright.com) is a global company offering a column-vectorized, highly compressed DBMS. With open-source — Infobright Community Edition (ICE) and commercial Infobright Enterprise Edition (IEE) — versions, the company also offers an Infopliance database appliance.

Strengths

- Infobright is a low-cost solution that is familiar to MySQL users. Experience demonstrates rapid learning curves if you are already familiar with MySQL, so fast adaptation is a reasonable expectation.
- With the exception of some loading commands, Infobright demonstrates good performance when compared with other analytic solutions above open-source DBMSs.
- Infobright is making progress in the Internet of Things' space (such as embedded analytics in mobile advertising) because of its low cost. At first, some of the cautions regarding Infobright might give pause; however, this does not deter IT value-added reseller (VAR) providers from leveraging the technology.

Cautions

- Infobright customers report significant use by casual users with repetitive reporting and queries. However, the solution is actually designed for rapid deployment of ad hoc capabilities, and the ability to deploy them rapidly to many casual users often masks its capability.
- Users report excellent performance, but indicate that high availability and disaster recovery must be built around the solution. Users report minimal difficulty in doing so as long as they have moderate experience in managing database system availability.
- In 2014, Infobright customers report more issues than in previous years regarding missing functionality — for example, issues with scale out or SQL support. These issues result in complexity of implementation. Infobright's customer base also reports it is losing some of its perceived value and pricing differentiation.

Kognitio

Kognitio (www.kognitio.com) started out offering Whitecross in 1992 and a managed service in 1993. It now has customers using the Kognitio Analytical Platform either as an appliance, a data warehouse DBMS engine, data warehousing as a managed service (hosted on hardware located at Kognitio's sites or those of its partners), or as a data warehouse platform using Kognitio's private cloud or the AWS public cloud.

Strengths

- Kognitio was one of the first data warehouse vendors offering cloud, including private and public cloud options, as well as on-premises solutions. It can now also be used as the analytics platform sitting on top of other solutions such as Hadoop distributions. Offering the flexibility of deployments is one of the areas that Kognitio has been pioneering in.
- Kognitio had an early understanding of emerging trends. It delivers new capabilities enabling both the LDW as well as the context-independent use case. Kognitio continues to deliver innovative technology to address emerging trends.

- Kognitio is praised consistently by references for its performance. Performance is a key enabler of the LDW and context-independent data warehouse use cases.

Cautions

- Kognitio could focus on becoming the administrative and coordinating platform for multiple analytical solutions. This would ease adoption by adding onto existing solutions, but would also limit the breadth of the go-to-market use cases that Kognitio can address.
- Kognitio had to scale back its U.S. operations. It is now supporting customers in the U.S. through partners who are increasingly responsible for sales delivery there.
- Kognitio's recent refocus on product development and delivering value through tight co-development with marquee customers or partners can be of great value for organizations looking for such tailored solutions, but makes it less of a direct fit for organizations looking for standard off-the-shelf software.

MapR Technologies

Founded in 2009, MapR Technologies (www.mapr.com) offers a Hadoop distribution with performance and storage optimizations, high availability improvements and administrative and management tools. It offers training and education services.

Strengths

- Based on information provided by MapR, it is the Hadoop distribution vendor with the largest number of paying customers — which is an important indication of adoption for an emerging category of technology.
- MapR's strategy is to deliver a data platform that combines Hadoop and operational database technologies to support a wide range of workloads in a single deployment. To enable this strategy, the company has compensated for Hadoop deficiencies by creating alternatives to Apache components while including a number of open-source projects from other distributions. For example, it substitutes Hadoop Distributed File System (HDFS) with its Posix-compliant, standard Network File System (NFS) file system, and it also supports Impala. This inclusive strategy offers customers the greatest number of options.
- MapR is praised by references for its reliability, performance and scalability, making it a solution suitable for enterprise use.

Cautions

- MapR has a smaller partner ecosystem than the other Hadoop distribution vendors. For example, the number of DBMS, BI or data integration partners is modest. However, MapR is actively addressing this by recently adding partnerships with Teradata, SAS and HP Vertica (for example).

- Reference customers struggle to find enough skilled resources in the market. The growing interest in Hadoop will help to relieve some of this pressure, but this is a multiyear cycle rather than one measured in quarters or months.
- Reference customers indicate that it can take time for MapR to support the latest Hadoop capabilities, although it can support multiple versions of the same Hadoop project. To address this concern, MapR accelerated its ecosystem update process during March 2013, and now has monthly Hadoop ecosystem releases.

MarkLogic

MarkLogic (www.marklogic.com) was founded in 2001 and offers a NoSQL database that utilizes XML, JSON, text, Resource Description Framework (RDF) triple and binary storage and offers a strong metadata-driven semantic access management layer. The product includes indexes, tiered storage, HDFS support, backup (to Amazon S3), mobile replication, full text search, geospatial capabilities and Simple Protocol and RDF Query Language (SPARQL) access.

Strengths

- MarkLogic includes text analytics based on a profiling function when ingesting documents and text sources. This results in classifying metadata that is then clustered for retrieval in analytics. This is then combined with a declarative ontology input by analysts. By combining profiling at ingest with near-real-time unstructured classification with the ontology system, MarkLogic represents a somewhat unique vendor in this market.
- MarkLogic runs MapReduce instructions (hardly mentioned by Hadoop distributions but the basis for Hadoop analytics) natively or in Hadoop clusters via connectors and also uses HDFS for tiered storage. The result is that customers can use MarkLogic for significant statistical analysis on semistructured data as a native capability, instead of leveraging a Hadoop cluster — but can also do both.
- Customers report ease of use, vendor support and robust capability to remain in production as advantages gained through the solution. MarkLogic offered developer outreach, free training and a host of new APIs in 2014.

Cautions

- Customers report that obtaining skilled personnel is a major issue. Some specifically cite the issues of managing to size — midsize organizations find it difficult to resolve implementation issues by spending more on servers. While servers are cheaper every day, the solution needs better management tools in smaller environments — which the company is addressing in recent releases with a new monitoring UI and API, and training and certification programs.
- Reference customers report issues with support and availability of skilled resources. This situation will be further impacted by MarkLogic's exhibited excellent growth during 2014 — adding new customers and licenses. In addition, MarkLogic is not a new company and at this

point in its life it needs to increase its interaction with implementers by training them and providing a business incentive model that encourages delivery by those same partners.

- The most appropriate use case for this solution is for the analysis of complex mixed information assets (by a data scientist or data analyst). But the customer base uses MarkLogic for repeating mostly "canned" or repetitive and somewhat unsophisticated analysis — limiting its organic growth with existing customers.

Microsoft

Microsoft (www.microsoft.com) markets SQL Server, a reference architecture, and Microsoft Analytics Platform System (which combines SQL Server Parallel Data Warehouse and HDInsight), as well as Azure HDInsight for Hadoop.

Strengths

- Microsoft has continued to bring to market product innovation with SQL server 2014, allowing, for example, the ability to support both in-memory analytical and transactional capabilities (HTAP) as well as through the delivery of Microsoft Azure Machine Learning.
- Microsoft was also shortlisted by Gartner clients more commonly in 2014, alongside — and competing against — other major vendors, demonstrating a significant increase in market traction and adoption. Customers indicate that common skills, tools and integration with the overall Microsoft stack are hard to resist when considering their DMSA.
- Microsoft has gained momentum and continues to grow faster than the overall market. In 2014, Microsoft was No. 2 in the DBMS market share of Software Worldwide for the first time. While it is not possible to split the revenue between operational and analytical use cases, this momentum is applied to both.

Cautions

- Despite its strong cloud product offering with its SQL Server offering on Azure, Microsoft is not yet competing in data warehousing in the cloud. While this is likely to be addressed in the future, this is a gap in Microsoft's offering at a time when interest in cloud data warehousing is increasing.
- In 2014, overall pricing emerged as a concern for many of the SQL Server references. This is due to changes in the Microsoft pricing models and, while satisfied with the integration of the "stack," it is also affecting the perceived value of the solution.
- Microsoft has established a reputation for supporting data warehouses for small or midsize business customers as they grow into large companies. It is now time for Microsoft to further its reputation with what Gartner defines as "extra-large" organizations (that is, those with \$1.5 billion annual revenue and higher).

Oracle

Oracle (www.oracle.com) provides the Oracle DBMS in multiple offerings as well as a Hadoop appliance called Oracle Big Data Appliance. Customers can choose to build a custom warehouse, a certified configuration of Oracle products on Oracle-recommended hardware or on appliances.

Strengths

- Market share for data warehouse delivery is imprecise, but inquiry data, market share data and other data indicate that Oracle has the largest market share in the traditional data warehousing space. Further, the traditional data warehouse market is the largest market segment at the beginning of 2015. Additionally, reference survey data indicates that for the past five years approximately 70% of existing Oracle customers have simply selected Oracle for data warehouse by default.
- Oracle has built out an integrated offering that addresses multiple form factors — software only and appliance. It is also further expanding its offering to support these various form factors in the cloud.
- Oracle continues to deliver innovation to the market, such as adding in-memory columnar optimized data store or delivering big data SQL in support of the LDW. Its product execution combined with market share sets a high bar for vendors competing with Oracle.

Cautions

- Oracle continues to push its engineered systems perspective as a means of addressing all four of the data warehouse use cases. However, this may not meet the emerging demand for what Gartner refers to as the "best-fit engineering approach" — where organizations are now willing to implement widely varied software and hardware configurations based upon specific service-level expectations, rather than using a stack of offerings from a single vendor.
- Oracle's focus on the largest market segments allows it to pursue a more deliberate product road map that introduces innovations to address timely delivery of robust features to its customers. While beneficial to existing customers, this does create some frustration over when and even how Oracle enhancements are introduced to the product.
- Oracle reference customers report the cost of the solution as the major concern. This is a major issue for Oracle at a time when the market is transforming and when traditional data warehouse DBMS vendors are being challenged.

Pivotal

Pivotal (www.pivotal.io) is a wholly-owned subsidiary of EMC financed by EMC, VMware and General Electric (GE owns 10% of Pivotal). It carries the following products: Pivotal Labs, Greenplum DB, Pivotal HDGemFire and GemFire XD. It combines and delivers these various products also through its Big Data Suite.

Strengths

- Pivotal continues to demonstrate market vision, addressing both the needs of operational and transactional processing and taking into account that data will have different models and will need to be accessed and analyzed despite these different models.
- Pivotal continues to benefit from the backing, resources and the installed base of EMC. This has allowed Pivotal to dedicate a larger chunk of its resources to product development and innovation.
- Greenplum DB references quote database performance and scalability as major strengths.

Cautions

- Pivotal is moving away from the traditional data warehouse DBMS market to focus on the broader scope of the Pivotal Big Data Suite. Gartner clients rarely mention Greenplum DB in their data warehouse DBMS selection. Additionally, reference customers are concerned about the commitment to the Greenplum DBMS in particular.
- While the combined offering of EMC and VMware plays to Pivotal's advantage, other providers of DMSAs are also pursuing the new market direction. Pivotal's ability to integrate products to deliver on the vision will be key to its success.
- Gartner inquiries and Gartner search analytics show that market traction for Pivotal is limited. This applies both to the data warehouse Greenplum DBMS and to Pivotal HD.

SAP

SAP (www.sap.com) offers both SAP IQ and SAP Hana. SAP IQ was the first column-store DBMS. It is available as a stand-alone DBMS and on an OEM basis via system integrators. SAP Hana is an in-memory column store that supports operational and analytical use cases — also offered as an appliance and reference architecture (TDI).

Strengths

- SAP leverages its existing installed base very well. It continues to add features that enhance all four data warehouse use cases. Adoption and production implementations of SAP Hana, as well as of Sybase IQ, are driven by the SAP installed base. SAP has been able to sell both on strategy and the tactical benefits of Hana — such as SAP BW powered by Hana.
- SAP continues to deliver on capabilities extending SAP Hana and delivering on LDW and context-independent use cases, with the addition of smart data access for federation or in database processing for predictive or spatial analytics.
- Reference customers continue to praise SAP Hana for its performance. Continued adoption and extension of deployments further demonstrate the growing maturity of the solution.

Cautions

- While Hana represents a unified solution for transactions and analytics for SAP customers regarding SAP data, its impact on the traditional data warehouse is modest. However, SAP is working to address this with Hana SPS9 by adding new capabilities such as better workload management. The market is expressing concern and confusion regarding SAP IQ, because it is largely absent from SAP's market positioning.
- SAP has been leading with in-memory, but will now face greater competition — including in its own installed base — because all megavendors now provide in-memory alongside mixed tiers of data storage for analytics.
- Reference customers report a series of issues that highlight challenges when adopting SAP Hana: lack of third-party skills, inability to integrate easily with other solutions, even issues with the road map and plans for SAP's own applications and business warehouse (SAP BW). This is an issue for growing outside of the SAP customer base, and even within it.

Teradata

Teradata (www.teradata.com) has more than 30 years of history in the data warehouse market. Teradata offerings include: DBMS licenses, appliances or via the cloud. It offers traditional and LDW solutions, which 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 appliance form factors), the Aster Database and Hadoop.

Strengths

- Teradata generally receives high marks from references. In 2014, a small number of references reported improved experiences. More importantly, the much larger Gartner inquiry base showed evidence of "re-engaging" with Teradata — with more cost proposals for upgrades and confirmation of Teradata's ability to dominate specific environments and use cases.
- Teradata demonstrates continuous technology enhancements to meet production demands. Teradata customers report the capability to isolate new and untested workloads — as they are inserted into previously full-production mixed workloads — as critical. In addition, the query-grid approach and the "data fabric" optimize query performance using built-in statistical models that optimize query plans, transfer query plans and steps between systems, and even perform data redistribution when needed across disparate platforms.
- Teradata has broad user type support. Users report cases of more than 10,000 casual users mixed with hundreds of more advanced business and data analysts and dozens of data scientists working on the same system. In addition, enhancements are enabling new analytic workloads to be added to Teradata systems (JSON, geospatial, 3D geospatial and others).

Cautions

- Clients question the role of the Teradata DBMS when evolving their solution toward a best-fit engineering practice. Two long-standing workloads — historical data analysis and extraction, loading and transformation (ELT), as opposed to extraction, transformation and loading (ETL) — often constitute as much as 30% of the workload on Teradata systems. Some organizations are shifting pieces of the transformation portion of these workloads to other platforms including Teradata appliances; however, most Teradata customers actually broaden their use cases when they recover capacity on their system. We have seen some delay in purchasing Teradata while customers develop their new strategy, but Gartner believes this is short-term issue.
- Customers report complaints about pricing and cost — they indicate this is further complicated because Teradata is often a stand-alone budget line item. Teradata offers various system configurations (and associated flexibility in pricing), and with its recent rollout of a cloud-based offering Teradata can be obtained at cloud-rate pricing.
- Teradata presented a unique value proposition in the data warehouse appliance and dedicated platform era, with most other vendors evolving toward appliances as much as 20 years later. A new era has begun in which customers are seeking to support multiple use cases with a value proposition that is more difficult to understand. Teradata must leverage its new value proposition quickly if it is to eventually take a similar value leadership role in the DMSA era.

Vendors Added and Dropped

We review and adjust our inclusion criteria for Magic Quadrants and MarketScopes as markets change. As a result of these adjustments, the mix of vendors in any Magic Quadrant or MarketScope may change over time. A vendor's appearance in a Magic Quadrant or MarketScope 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

- MapR Technologies.
- Hitachi (with HADB).

Dropped

- InfiniDB (Calpont) has ceased operations as a private company, but the InfiniDB platform continues to be available as an open-source project on GitHub.

Other Vendors to Consider

Gartner's Magic Quadrant process includes 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 because they failed to meet one or more of the inclusion criteria. Unless otherwise noted, the information provided on these vendors derives from responses to Gartner's initial request for information for this document or from reference survey respondents.

This list is not intended to be comprehensive:

- **Hortonworks.** Located in Palo Alto, California, U.S., and founded in 2011, Hortonworks markets the Hortonworks Data Platform (HDP), derived entirely from the open-source Apache Hadoop stack. The company was a leading contributor to Hive for SQL interfacing. HDP includes services to support security, data governance and operations. Hortonworks participates in the "Stinger" Initiative to advance Apache Hive for interactive query capabilities and claims HDP enables interactive query operations at the petabyte scale.
- **MongoDB.** MongoDB offers two products: MongoDB, an open-source NoSQL document-style database, and MongoDB Enterprise Advanced (available in various service tiers). The database provides for mixed-type datasets targeted at broad scalability on commodity hardware and supports indexing (including geospatial and text search), a MapReduce engine and strong integration with Hadoop — using memory-mapped files to enhance read performance, which operates on commodity-class hardware. MongoDB Enterprise Advanced is available through several cloud providers, as well as on-premises. It has a large partner ecosystem, with partners such as AWS, Capgemini, Cloudera, Cognizant, EMC, Google, Hortonworks, IBM, Map Technologies, Microsoft, Pure Storage, Rackspace and Red Hat. Data integration tools such as Informatica, Pentaho and Talend offer native integration support. Customers use MongoDB to power a variety of data warehouse and analytics applications. MongoDB is dual-headquartered in New York City, New York and Palo Alto, California, U.S., with 15 offices worldwide.
- **Objectivity.** This vendor, which has a lineage dating back to 1989 when it first began to offer an object-oriented database, now offers InfiniteGraph and Objectivity/DB (version 10.2.1). Objectivity reports a global sales presence with thousands of direct-licensed customers as well as thousands of embedded licenses worldwide. Customers give it a mixed review: praising capabilities such as multiple versions of objects and high-level SQL capabilities; but also indicating that nonspecific SQL functionality that should be present is lacking, resulting in a demand for specialized, product-specific skills. In 2014, Objectivity focused on releasing additional product versions to address customer-requested features, functionality and performance. Objectivity is based in Sunnyvale, California, U.S.
- **ParStream.** ParStream is a columnar, in-memory database offering a high-performance compression index on a massively parallel processing 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, and a recently established partnership with QlikView, Datawatch and other front-end tools provides visual data analysis for large datasets in combination with time series data. The company is based in Cologne, Germany and Cupertino, California, U.S.
- **RainStor.** RainStor 6.0 was released in June 2014, with its first general availability release in June 2008. The product can be deployed on-premises or in the cloud and most of its more than 100 customers report use cases as a near-line, fully integratable data archive. It is integrated

with the Teradata data warehouse and capable of moving data bidirectionally between the two environments. However, RainStor also provides full DBMS capability in a highly compressed file format that can hold multiple data types. Customer solutions in production include analytical and compliance archives running non-Hadoop and Hadoop distributions from Cloudera, Hortonworks, IBM and MapReduce; certified configurations via Dell, EMC and other hardware platforms; and it is also part of an offering for data retention and analytics from HP. RainStor's technology appears viable to support the LDW as a primary component. It is based in San Francisco, California, U.S. Rainstor was acquired by Teradata during 4Q14.

- **XtremeData.** This privately owned company targets organizations that need a massively scalable DBMS solution for mixed read and write workloads in the cloud — both public and private. Their dbX database is a multithreaded solution that scales to efficiently use available processor capacity. Organizations benefit from low entry costs, fast time to market, elastic scalability and a pay-for-use billing model. XtremeData is available on AWS, Rackspace, CenturyLink, Microsoft Azure and other clouds. Our information on this vendor was obtained from reference customers, from previous direct communications with XtremeData and the ongoing research of Gartner analysts. XtremeData is based in Schaumburg, Illinois, U.S.

Inclusion and Exclusion Criteria

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

- Vendors must have DBMS software that has been generally available for licensing or supported download for approximately one year (since 10 December 2013). We do not consider beta releases.
 - We use the most recent release of the software to evaluate each vendor's current technical capabilities (we do not consider beta releases). For existing data warehouses, and direct vendor customer references and reference survey responses, all versions currently used in production are considered. For older versions, we consider whether later releases may have addressed reported issues, but also the rate at which customers refuse to move to newer versions.
 - Product evaluations include 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 2014. Capabilities, product features or functionality released after this date can be included at Gartner's discretion and in a manner Gartner deems appropriate to ensure the quality of our research product on behalf of our nonvendor clients. We also consider how such later releases can reasonably impact the end-user experience.
- Vendors must have generated revenue from at least 10 verifiable and distinct organizations with data warehouse DBMSs in production that responded to Gartner's approved reference survey questionnaire. Revenue can be from licenses, support and/or maintenance. Gartner may include additional vendors based on undisclosed references in cases of known use for classified but unspecified use cases. For this year's Magic Quadrant, the approved questionnaire was produced in English. Gartner exercises its option to provide for other languages as deemed appropriate only in the case of an extreme exception.

- 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 2014. Acquisitions after 30 June 2014 are considered a legacy offering and will be 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) must be available from the vendor. We also consider products from vendors that control or participate in the engineering of open-source DBMSs and their support.
- We also include the capability of vendors to coordinate data management and processing from additional sources beyond the DBMS, but continue to require that a DBMS meets Gartner's definition — in particular regarding support of at least one of the four major use cases (traditional data warehouse, operational data warehouse, LDW or context-independent data warehouse, see Note 6).
- Vendors participating in the data warehouse DBMS market must demonstrate their ability to deliver the necessary services to support a data warehouse via 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 do not qualify for this Magic Quadrant.

Data warehouses and DMSAs were also determined for inclusion or exclusion based upon the following:

- Relational data management.
- Nonrelational data management.
- No specific rating advantage is given regarding the type of data store used (for example, RDBMS, HDFS, key-value, document; row, column and so on).
- Multiple solutions in combination to form a DMSA are considered valid (although one approach is adequate for inclusion), but each solution must demonstrate maturity and customer adoption.
- Cloud solutions (such as PaaS) are considered viable alternatives to on-premises warehouses; and ability to manage hybrids between premises and the cloud are considered advantageous.
- DMSAs are expected to coordinate data virtualization strategies for accessing data outside of the DBMS, as well as distributed file and/or processing approaches.

For details of our research methodology, see Note 7.

Evaluation Criteria

Ability to Execute

Ability to Execute is primarily concerned with the ability and maturity of the product and the vendor. Criteria under this heading also consider the 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, therefore customer references are weighted heavily throughout.

Product or Service. Represents increasingly divergent market demands — ongoing traditional, logical data warehousing, operational data warehousing and context-independent data management for analytics (for definitions see "Critical Capabilities for Data Warehouse Database Management Systems"). 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 — referred to in the market as LDWs. All traditional demands of the data warehouse remain. Operational data warehouse use cases also exhibit traditional requirements plus loading streaming data, real-time data loading and real-time analytics support. Users expect solutions to become self-tuning, to reduce staffing required to optimize the data warehouse, especially as mixed workloads increase. Context-independent warehouses (CIWs) do not necessarily support mixed workloads (but can), nor do they require the same level of mission-critical support. CIWs serve more in the role of data discovery support or "sandboxes."

Overall Viability (Business Unit, Financial, Strategy, Organization), Financial. Includes corporate aspects, such as the skills of the personnel, financial stability, R&D investment, the overall management of an organization and the expected persistence of a technology during merger and acquisition activity. It also covers the company's ability to survive market difficulties (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. We examine the price/performance and pricing models of the DBMS, and the ability of the sales force to manage accounts (judged by the feedback from our clients and feedback collected through the reference survey). We also consider 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 for this criterion.

Market Responsiveness/Record. Based upon the concept that 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 more recent market demands and the ability to recognize meaningful trends early in the adoption cycle are particularly important. The diversity of delivery models as demanded by the market is also considered an important part of this criterion (for

example, its ability to offer appliances, software solutions, data warehouse "as a service" offerings or certified configurations).

Marketing Execution. Includes the 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 the vendor's coordination and delivery of education and marketing events throughout the world and across vertical markets, as well as increasing or decreasing participation in competitive situations. This year, events and education are part of marketing execution.

Customer Experience. Based on customer reference surveys and discussions with users of Gartner's inquiry service during the previous six quarters. Also considered are the vendor's track record on proofs of concept, customers' perceptions of the product, 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 the vendor's approach in the market.

Operations. Alignment of the vendor's operations, as well as whether and how this enhances its ability to deliver. Aspects considered include field delivery of appliances, manufacturing (including the identification of diverse geographic cost advantages), internationalization of the product (in light of both technical and legal requirements) and adequate staffing. This criterion 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 considered.

Table 1. Ability to Execute Evaluation Criteria

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

Source: Gartner (February 2015)

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, comprehends overall market trends, and influences or leads the market when necessary. A visionary leadership role is necessary for the long-term viability of both product and 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 for future product development.

Market Understanding. The 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 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 8), and the changing concept of the data warehouse and analytics data management; or, the value and position regarding early emerging terminology such as "polyglot" data management. Understanding the different audiences for traditional data warehousing and new approaches is crucial as is a demonstrable track record for altering strategy and tactical delivery in response to both opportunistic segments in the market and the broader market trends.

Marketing Strategy. Refers 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 vendor fiscal health), but vendor management and response to that sentiment (see, "The Data Warehouse DBMS Market's 'Big' Shift").

Sales Strategy. Encompasses all plans to develop or expand channels and partnerships that assist with selling, and 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 will evaluate the road map for enhancing traditional data warehouse capabilities (including, but not limited to, addressing currently missing execution components). This also includes expected functionality and a timetable for introducing new market demands that will specifically include, but are not limited to, road maps and development plans for: a semantic design tier; system and solution auditing and health management to assure 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.

Business Model. This is 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 is the vendor's ability to understand its clients. A measurable level of influence within end-user communities and certification by vertical industry standards bodies are of importance here.

Innovation. Vendors demonstrate this in developing new functionality, allocating R&D spending 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 here is the maturation of alternative delivery methods such as IaaS and cloud infrastructures, as well as solutions for hybrid premises-cloud and cloud-to-cloud data management support. Vendors' awareness of new data warehousing methodologies and delivery trends is also considered. Organizations are increasingly demanding data storage strategies that balance cost with performance optimization, so solutions that address aging and temperature of data will become increasingly important.

Geographic Strategy. This considers the vendor's ability to address customer demands in different global regions using direct/internal resources or in combination with subsidiaries and partners. We also evaluate a vendor's global reach.

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 2015)

Quadrant Descriptions

Leaders

The Leaders quadrant is the most heavily populated portion of this market, which makes the market unusual in both its maturity and the fierce competition it experiences. The most notable aspect of these vendors is that all of them have determined to pursue all four market use cases for data warehousing (warehouses for traditional, operational, context-independent and LDW deployments) with at least an average degree of mature execution and vision. The span of ratings is also the widest among any of the quadrants.

The Challengers and Visionaries have completed their disruption cycle; the Leaders are responding. The largest vendors are developing their own alternatives that range from extending their processing logic into the operating systems of remote processors or clusters, to rationalizing workloads using metadata engines guided by utilization, capacity and service-level policy controls. Whether a system takes control of the data management and processing or simply leverages remote clusters efficiently, the result is the same for these Leaders — they adapt to the disruption, incorporate it and then prepare for the next disruption. Data management is the bedrock of all IT; security is deployed to protect it; applications are deployed to capture it and use it; metadata was developed to describe and audit it; volumetrics for everything are based upon it; and now, there are attempts to place quantifiable value on it.

The data warehouse, which constitutes the largest data management system in most organizations, is therefore a large market in terms of revenue, trained professionals and a variety of data management solutions ranging from simple to complex. The high-stakes decisions that are being made by all of the Leaders regarding road maps and market delivery will be critical during the next two years. Buy, build or invent will be on every product manager's agenda every single day for the next 24 to 36 months. The emphasis for Leaders is to retain existing 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 messaging creates confidence for both the traditional and emergent ends of the market.

Almost all of these vendors appear to have moved "backward" in 2015, but that is an illusion of the Magic Quadrant in a mature market. All mature markets demand more vision and more execution every single year. That means the Magic Quadrant moves up and to the right every year and simply maintaining a position is difficult.

Challengers

In 2014 and heading into 2015, the Challengers have made some very careful decisions to secure their markets. Some focused on a deployment strategy, such as focusing on becoming the dominant cloud offering. Others focused on specifically large data volumes in a pre-cloud, pre-big data environment with fully optimized solutions that now dominate vertical delivery in regional areas. New-era vendors in this quadrant focused on growing products that capitalized on an eager market

that was willing to develop new skills first and accept improved tools later. All of them succeeded quite well in their individual pursuits.

Their success has created a barrier to enhancing their vision overall — with revenue in place and a customer base that demands greater functionality from products that are more timely and more specifically aligned with product road maps as announced. They are fast becoming incumbents that will find their support models being challenged by their user community.

The key to the Challengers in 2015 will lie in two directions. For the robust vendors with significant market captured in a vertical or region, they will have to first determine how much they will grow, then design the strategy to pursue that plan. On the other hand, they must avoid the temptation to expand their reach into other use cases before they are mature enough to satisfy existing customers' support needs. Focus, concentration of resources, speed of cloud product patching, fixing and enhancements will become the order of the day.

Visionaries

To qualify as Visionaries, vendors must demonstrate that they have customers in production — in order to prove the value of their functionality and/or architecture and not simply be experimental implementations "of interest."

In 2015, our Visionaries have enormous potential for the next three years. The Visionaries are challenging the more mature Leaders with new concepts. However, their customers often report specific tuning and deployment issues, lack of practices and skills in the market, and spotty success with intermittent difficulties. With a weaker execution, Visionaries are vulnerable for a variety of reasons — from revenue to customer count to reported issues during implementation and even into production. There is a cluster of vendors right around the central cross hairs of the Magic Quadrant, and any of them could enhance their vision into new analytic use cases, new features and functions or via hardware exploitation.

Niche Players

The Niche Players in 2014 saw new entrants challenge them and, in some cases, move swiftly past them. However, the overall broadening of vision in the market, with competing approaches, is actually creating openings for new vendors. Niche Players generally deliver a highly specialized product with limited market appeal. Frequently, a Niche Player provides an exceptional data warehouse DBMS product, but is isolated or limited to a specific end-user community, region or industry. Although (sometimes) the solution itself may have no limitations, its adoption is limited.

Context

In 2015, the Leaders quadrant is widely spread out and heavily populated, that means the market is in a state of significant fluctuation. Challengers on one edge and Visionaries near the other edge have the potential to disrupt the market leaders — especially those near the edges of the Leaders quadrant themselves. When such disruption occurs, it usually means that the entire market moves

away from a single trajectory of maturity and splits in the two directions of vision and execution. The net result is that over the next two years, until the end of 2016, this market will exhibit a much higher level of volatility and could see disruptions and changes in leadership.

Gartner has previously referred to the coming "title fight" over data management for analytics (see "Magic Quadrant for Data Warehouse Database Management Systems"). We indicated the "fight" would start at the end of 2013 — and so it began in early 2014, and continues now. We originally anticipated this fight would last about three years and believe this is still the case. By the end of 2016, all of the traditional vendors will have a clear strategy to support all four data warehouse use cases and offer pricing, licensing and deployment options to meet small, medium, large and extra-large business profiles across multiple verticals.

At the same time, the winners and losers from vendors of the new era will begin to sort themselves out. Acquisitions will begin in early 2016 (and even at the end of 2015) — providing investor exit strategies. These new vendors are already scrambling to introduce high availability, disaster recovery, security, administrative consoles, programming language interfaces, semantic data management, hardware virtualization and workload distribution management tools.

New contenders will emerge that include newly adapted traditional vendors that have acquired or deployed new capabilities on their mature architectures; maturing new era vendors (few or possibly only one or two of which will actually survive intact past 2016); and wide-area network providers (such as Cisco) that will pioneer new approaches for efficient data management and processing management over wide, geographically distributed information assets.

Market Overview

End-user organizations in 2014 and beyond require data warehouse platforms that are capable of managing and processing data internal to their native repositories in combination with external data sources (including data external to their organization). The result is the emergence of a new demand in the market for DMSAs with features and functionality that represent a significant augmentation to the existing enterprise data warehouse strategy. The data warehouse market now includes management of all types of data for analytics under an integrated approach. Data warehouse managers, solution architects for analytics and CIOs establishing IT modernization strategies must take note of this change in direction and prepare to meet it with hybrid technology platforms that expand the data warehouse beyond any of its current practices. It will become the DMSA market — and traditional data warehousing will become a critical, but equal player in the larger DMSA market.

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 preintegrated, 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 also see "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.
- **Big data's role in the data warehouse.** Big data was an approach that served as a catalyst for change in the data warehouse environment. Implementers have identified three highly useful use cases for big data in analytics: data exploration/data science "sandboxes," offloading history from the warehouse, moving transformation support back off the data warehouse platform in ELT. Successful organizations pursuing the use of big data in advanced analytics are following a best-of-breed (BOB) approach to support the solution space, because no single product is a complete solution. They are even employing multiple products from a single vendor manifest as the BOB within their stack. However, they are now seeking approaches to integrate with these new, very large data management and analytic silos.
- **The emergence of best-fit engineering.** The BOB deployment model includes a combination of different software (proprietary license and open-source license), file management systems, communication and semantic middleware, and variable hardware/network components. Generally, BOB meant acquiring leading technologies in several areas and then hiring expert implementers to accomplish the deployment. However, BOB is being replaced by a concept Gartner calls "best-fit engineering" (see "The Data Warehouse DBMS Market's 'Big' Shift"). The difference between the two is that under BOB, implementers will select the best solution for part of their architecture and then reuse it in secondary roles that emerge — even if such secondary functionality is less than optimal. In best-fit engineering, the least required technology for each function is considered first. For example, it is possible to use a DBMS to facilitate access to external files and tables under BOB, but by adding a different technology to the stack that is specifically focused on data virtualization (and therefore has different, and possibly superior, optimization capabilities), it becomes best-fit engineering. Each technology is used for its most appropriate purpose and is therefore much more likely to exhibit a low cost for a precise need.

The concept of the LDW emerged as the first practical architecture for the newly emerging analytic data management requirements. The LDW will continue to grow in popularity during the next five years. The language around the LDW will become the de facto vocabulary for describing how to evolve a traditional data warehouse into a broader DMSA.

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
PaaS	platform as a service
SSED	source-system extracted data

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"Market Share: All Software Markets, Worldwide, 2013"

"Survey Analysis: Big Data Adoption in 2013 Shows Substance Behind the Hype"

"How Markets and Vendors Are Evaluated in Gartner Magic Quadrants"

Note 1 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 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

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 4 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 for the appliance does not include separate prices for the hardware, storage and DBMS components.

Note 5 User Categories

- Data Scientist (expert statistics skills, numerology, abstract mathematics, programming skills, business process skills, communication skills, leadership skills).
- Data Miner (expertise in subject areas of data, utilizes statistical software, statistical models, fully aware of computer processing "traps" or errors).
- Business Analyst (utilizes OLAP and dimensional tools to create new objects, some difficulty with computer languages and computer processing techniques).
- Casual User (regularly uses portals and prebuilt interfaces, minimally capable of designing dimensional analytics if at all).

Note 6 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 managing data variety and volume of data 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, variants of data form and new relationships. Also supports search, graph and other advanced capabilities for discovering new information models. There are no specific performance requirements and this option is favored by advanced users such as data scientists or data miners, resulting in free-form queries across (potentially) multiple data types.

Note 7 Research Methodology for This Update

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: inputs such as use cases, issues encountered, license and support pricing, and implementation plans.
- RFI surveys issued to the vendors, in which they were asked to provide specifics about versions, release dates, customer counts, distribution of customers worldwide and other data points. Vendors could refuse to provide any information in this survey, at their discretion.
- Customer reference surveys (with almost 300 new responses added this year to four prior years of survey data). Vendors were asked to identify a minimum number of references. 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 validate customers as current license holders. Additionally (especially in the case of open-source utilization), customers provided information that validated 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 the 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 trending only. Current-year survey responses were used for all commentary.

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

It is important to note that this is qualitative research and, as such, forms a cumulative base on which to form the opinions expressed in this Magic Quadrant.

Note 8 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 SSSED, excluding all data-warehouse-design-specific structures (such as indexes, cubes, stars and summary tables). SSSED 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
- Midsize data warehouse — 5TB to 40TB
- Large data warehouse — more than 40TB

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.

GARTNER HEADQUARTERS

Corporate Headquarters

56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters

AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit <http://www.gartner.com/technology/about.jsp>

© 2015 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](#) posted on 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. 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](#)."