

Critical Capabilities for Data Management Solutions for Analytics

Published: 16 March 2017 **ID:** G00314186
Analyst(s): Rick Greenwald, Donald Feinberg, Adam M. Ronthal

Summary

Data management solutions for analytics are evolving rapidly, with new deployment scenarios, data sources and analytics requirements driving the market. Data and analytics leaders can use this research to guide evaluation and buying criteria for modern DMSA offerings.

Overview

Key Findings

- Traditional data management solutions for analytics (DMSA) vendors have integrated nonrelational access and functionality into their products and regained dominant positions in the market.
- Access to external data sources, such as data stored in the Hadoop Distributed File System (HDFS), has become "table stakes," indicating further acceptance and implementation of the core of the logical data warehouse infrastructure.
- Cloud-centric features, such as independently scalable storage and compute and subscription pricing, have made their way into private cloud and on-premises offerings.
- Hadoop- and cloud-based vendors are seeing a high percentage of surveyed customers that indicate plans to purchase more product in the next 12 months.

Recommendations

For data and analytics leaders tasked with modernizing their information infrastructure:

- Evaluate cloud-based solutions as options for new DMSA use cases. Existing DMSA use cases may also benefit from hybrid solutions, which can ease the migration path and deliver greater flexibility.
- Compare products within a given cohort for DMSA use cases. For instance, Hadoop-oriented products, such as those from Cloudera, Hortonworks, MapR, Transwarp and Huawei, will present more valid comparisons between themselves than with products in other cohorts, such as traditional relational DBMS (RDBMS) vendors like Oracle, IBM, Teradata and Microsoft.
- Plan for additional skill sets and an ongoing cost of data integration, which should be considered when evaluating multiple platform environments.

What You Need to Know

- Organizations now require data management solutions for analytics (DMSA) that can manage and process 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 — for example, from Internet of Things (IoT) sensors — as well as nonrelational data like text, images, sound and video. These requirements are placing new demands on DMSA offerings in this market as customers look for features and functions that represent a significant augmentation of existing enterprise data warehouse capabilities.
- Moreover, expectations are now turning to cloud as an alternative deployment option for its flexibility, agility and operational pricing models. As combining cloud and on-premises in a hybrid configuration is quickly becoming the norm, organizations expect vendors to support them in meeting such deployments.

Finally, the traditional data warehouse use case, while still the most common, is losing importance in the market. Among Gartner clients, traditional data warehouse inquiries are now fewer than logical data warehouse inquiries. This trend was first described in 2014 in "The Data Warehouse DBMS Market's 'Big' Shift" and is reflected in the change in name of the 2017 "Magic Quadrant for Data Management Solutions for Analytics." This change has also resulted in an expansion of the types of vendors included, and reflected in the inclusion criteria that have become more demanding to meet.

For this market, a data management solution for analytics 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 include specific optimization strategies designed for supporting analytical processing, including — but not limited to — relational processing, nonrelational processing (such as graph processing), and machine learning or programming languages such as Python or R. Data is not necessarily stored in a relational structure and can use multiple models (relational, document, key value, text, graph, geospatial, and others).

Our definitions also state that a data management solution for analytics:

- Is a system for storing, accessing, processing and delivering data intended for one or more of the four primary use cases that Gartner identifies as supporting analytics
- Is not a specific class or type of database management system (DBMS)
- May consist of many different data management technologies in combination; however, any offering or combination of offerings must, at its core, exhibit the capability of providing access to the data under management by open-access tools via APIs (e.g., via ODBC, JDBC, OLE DB)
- 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
- Must manage the storage and access of data residing in a type of storage medium, which may include (but is not limited to) hard-disk drives, flash memory, solid-state drives and DRAM

This research is aimed at data and analytics leaders. In this Critical Capabilities document, we have focused on the 10 most important functional capabilities that are required to support the major use cases identified. The research combines product capabilities and customer experiences to evaluate the support for these critical capabilities. User experience is evaluated based on the companion Magic Quadrant reference survey, Gartner inquiries, peer insights, in-depth reference calls and interactions with vendors. In addition to customer experience, capability ratings include Gartner analysis of differentiating product capabilities as described in the capability definitions.

Although this research shared survey results gathered for the 2017 "Magic Quadrant for Data Management Solutions for Analytics," the critical capabilities ratings are focused on how well a specific vendor product addresses one of four use cases, rather than the overall estimation of the vendor. In addition, the critical capabilities focus on a single offering from each vendor, unlike the Magic Quadrant, where all relevant offerings were considered.

Customers in the survey were initially identified by the vendors, and a subset of those customers completed the survey. This methodology was employed to provide a level playing field for vendors of different sizes and orientations. The actual composition of the customer base of a vendor may differ in its overall use of a particular solution or capability. The survey helps to determine if a vendor's user base has widely adopted a capability for production use.

This research does not include all criteria that should lead organizations to select a particular DMSA vendor. Many other criteria not included in our analysis will come into play, such as whether the offering is a stand-alone DBMS software package, appliance or cloud solution. Other requirements – pricing, vertical industry offerings, the availability of services, for example – are not included but would need to be part of a formal RFP process (see "Toolkit: RFP Template for Data Warehouse and Data Management Solutions for Analytics"). These aspects do factor in the evaluations for the Magic Quadrant for this market space.

Our research covers the ability of various vendors to provide certain capabilities that are critical to one or more of the use cases described. Be aware that the vendors represented in this Critical Capabilities report range from best-fit solutions to best-of-breed solutions.

The ability to deploy on-premises, in the cloud or in a hybrid model was not considered a critical capability in this document.

Additionally, capabilities are chosen based, in part, on their ability to differentiate between different solutions. As vendors approach parity on a capability, the range of scores for that capability narrow and the capability moves toward a status of a basic requirement, which may lead to it being dropped in future Critical Capabilities documents.

Readers should understand that our scores are meant to convey a vendor's standing in relationship to the market at the time of the report. As such, scores for any capability are not absolute from year to year, but relative and only relevant within the context of this specific yearly report.

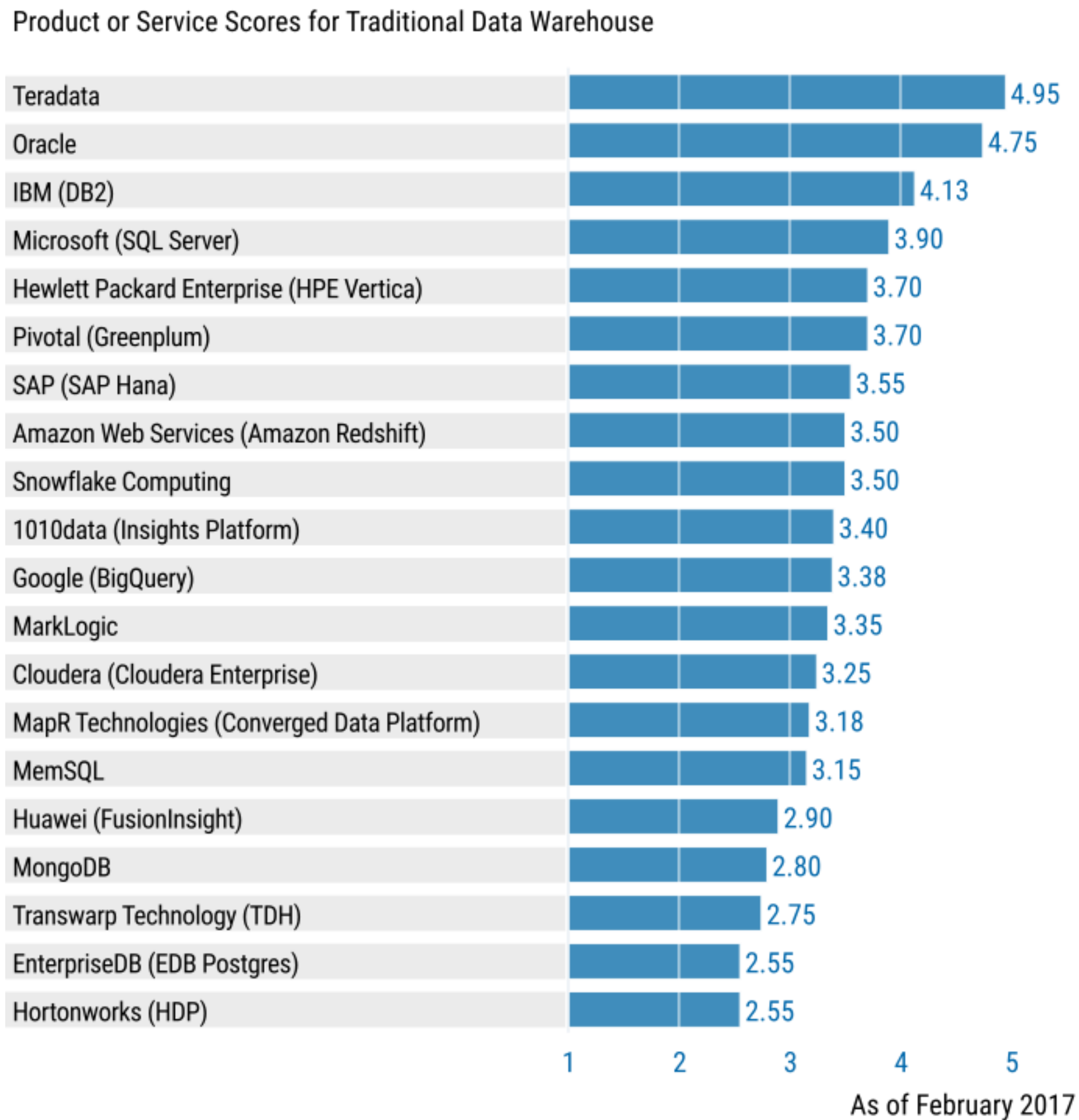
As detailed below, a score of 3.0 indicates that a product met the requirements for a particular use case. Although vendors are listed in the order of their relevant ranking (and alphabetically in the case of an equivalent score), be aware of the meaning of the individual ratings. Furthermore, Gartner does not recommend using any rating as the sole or primary basis for product selection, as there are many factors outside the scope of this research that can impact the suitability of a product.

In some cases, the overall range of these scores may shift from year to year. These changes are the result of both changing market conditions and refinements in the calculations used to evaluate these capabilities. In this year's research, there have been changes to criteria for evaluating support of external data sources and the replacement of a criterion for repeated queries with a criterion for query optimization.

Analysis

Critical Capabilities Use-Case Graphics

Figure 1. Vendors' Product Scores for the Traditional Data Warehouse Use Case



Source: Gartner (March 2017)

Figure 2. Vendors' Product Scores for the Operational Data Warehouse Use Case

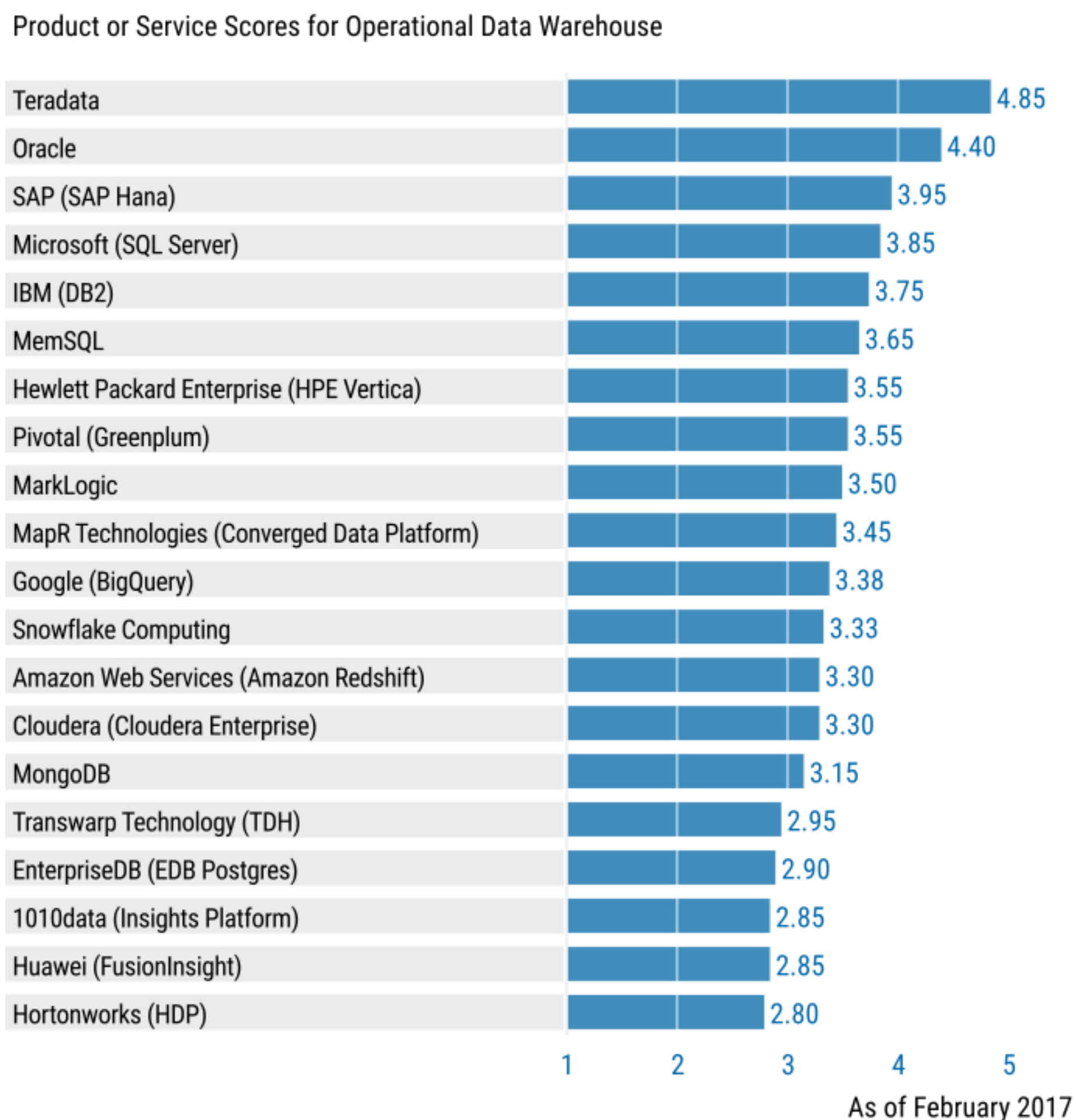


Figure 3. Vendors' Product Scores for the Logical Data Warehouse Use Case

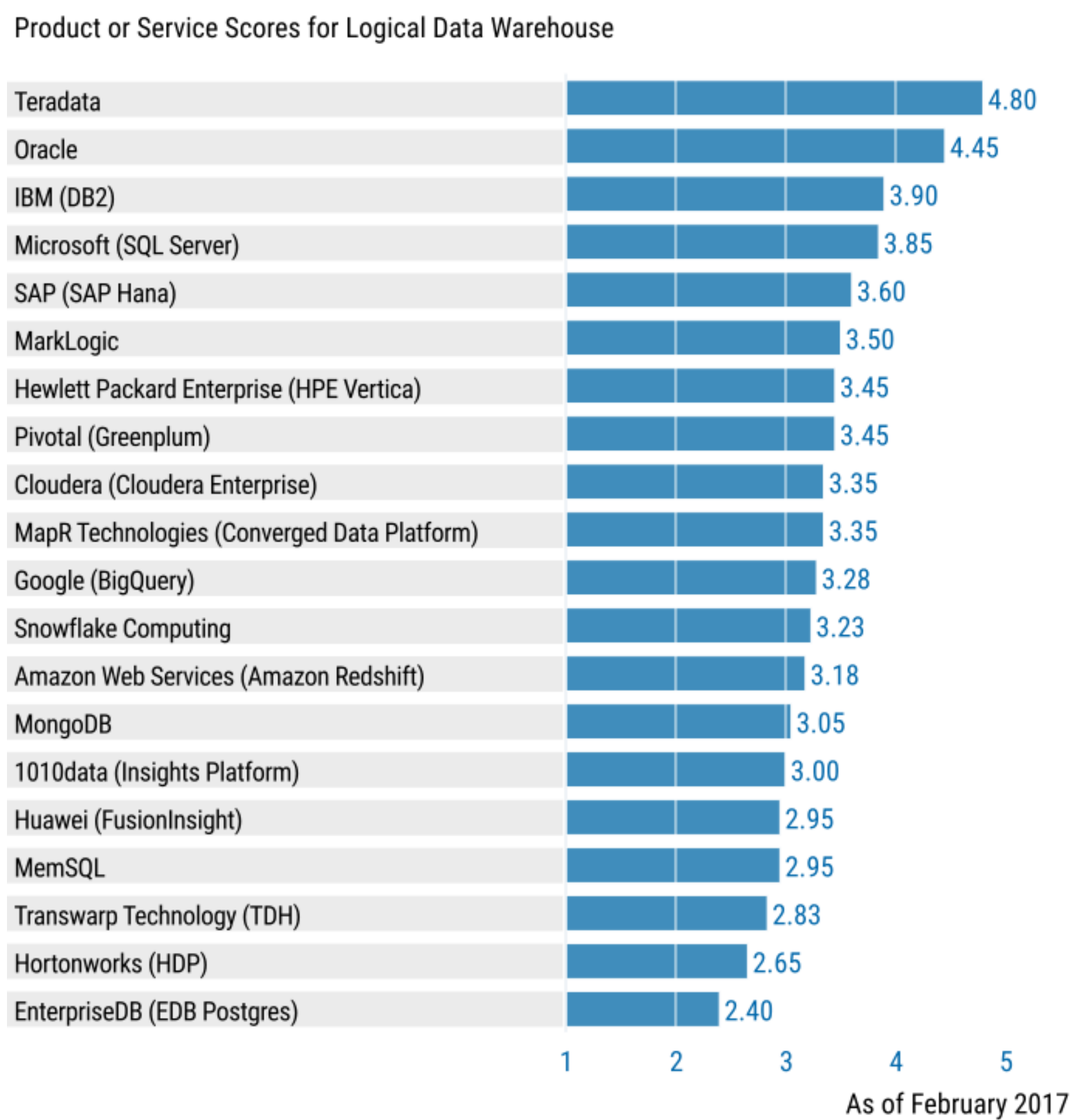
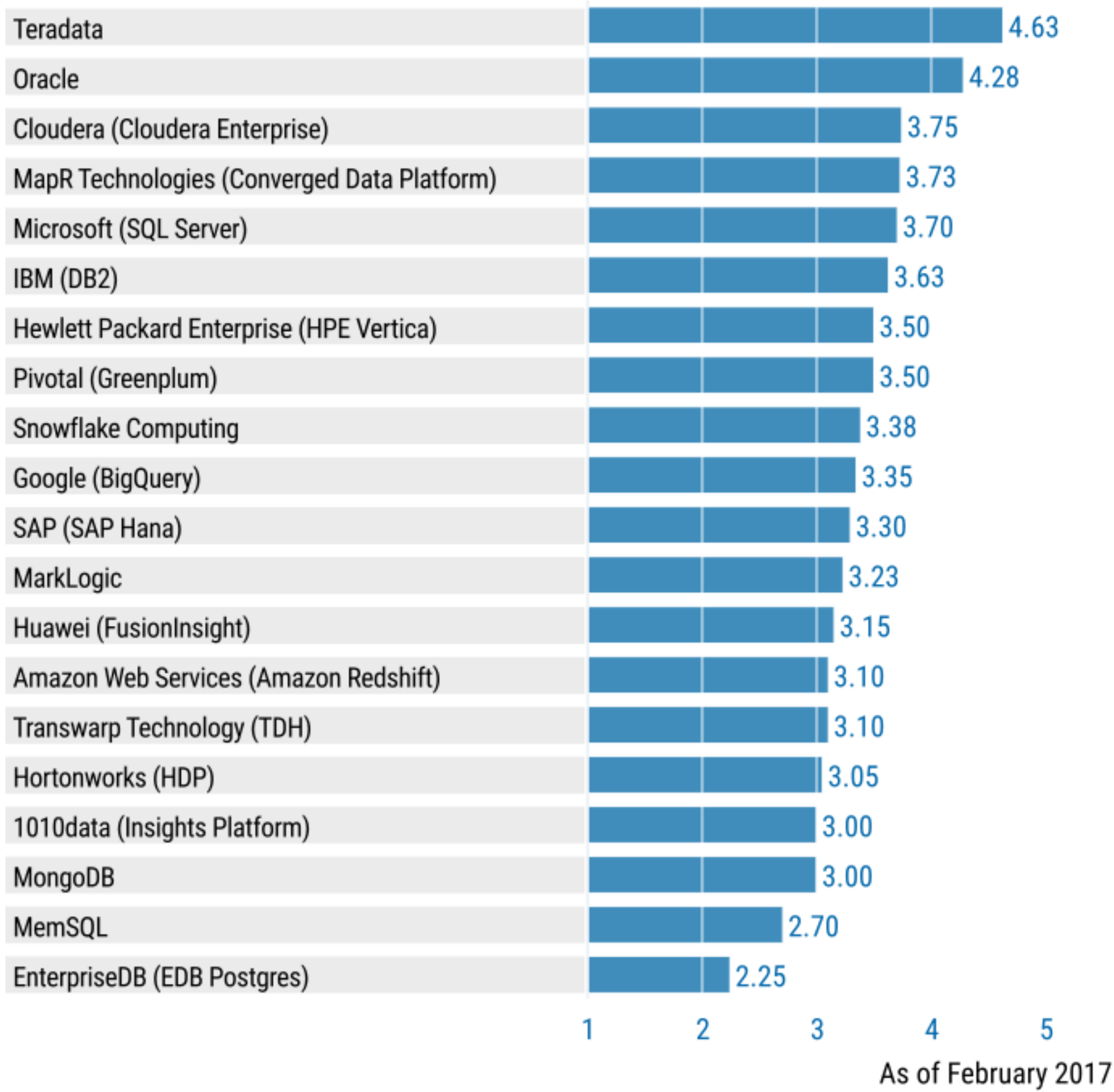


Figure 4. Vendors' Product Scores for the Context-Independent Data Warehouse Use Case

Product or Service Scores for Context-Independent Data Warehouse



Source: Gartner (March 2017)

Vendors

1010data (Insights Platform)

1010data's Insights Platform (<https://1010data.com/products/insights-platform/>) is primarily a managed service DMSA offering, which consists of an integrated DBMS and business intelligence (BI) solution. Most of its customers are in the financial services, retail/consumer packaged goods, telecom, government and healthcare sectors.

Although 1010data ranks comparatively low in three use cases, it must be noted that the vendor scores well above 3.0 (meets requirements) for the traditional data warehouse use case. This is the primary focus of 1010data, as shown by the largest number of its survey responses reporting this use case. That said, 1010data also demonstrates suitability for two of the other three use cases, scoring 3.0 for each.

Compared to 2016, 1010data ranks higher in all four use cases in relation to the criteria for meeting requirements, even though the requirements are more stringent this year. Survey results show most queries supported by 1010data are repetitive, prebuilt queries run by casual users, with little or no data from outside the 1010data database. This contributes to the lower scores for the nontraditional data warehouse use cases where continuous loading, external data sources and large volumes of data for data discovery are important criteria.

Amazon Web Services (Amazon Redshift)

Amazon Web Services (AWS) offers Amazon Redshift (https://aws.amazon.com/redshift/?nc2=h_m1) , a data warehouse service in the cloud. AWS is a wholly owned subsidiary of Amazon. This document focuses on a single product from each vendor, while Amazon believes in a best-fit approach, with customers using multiple services together.

Amazon Redshift's relative ranking for the traditional data warehouse use case and the operational data warehouse use case rose this year, based in part on high scores for continuous data loading and traditional data warehouse use, and an improved score for managing large amounts of data. Redshift ranked at the top for system availability, based on the robustness of the AWS Cloud platform.

AWS adheres to a best-fit philosophy, where different services are provided for different use cases. Because of the evaluation of a single product for this document, Redshift may not have scored as well in a use case such as the context-independent data warehouse, which might typically use a different AWS service (such as EMR).

In addition, Redshift does not have good access to data in external systems, which hurt the vendor in the logical data warehouse category.

Like the Hadoop cohort and Google, a high level of Redshift customers reported plans to purchase more product in the next 12 months.

Cloudera (Cloudera Enterprise)

Cloudera offers Cloudera Enterprise Data Hub (<https://www.cloudera.com/products/enterprise-data-hub.html>) , an Apache Hadoop distribution that combines components of Apache Hadoop and Spark but also delivers its own components such as Cloudera Navigator for data governance, Cloudera Director for cloud provisioning and Cloudera Manager for cluster administration, Impala for SQL access and Kudu for analytics on transactional data. Cloudera's platform is available both on-premises and across multiple cloud environments (such as AWS, Microsoft Azure or Google Cloud Platform), including with cloud-native support for object stores.

In 2016, customers rated Cloudera highly on supporting large volumes of data and discovery use. These positives led to the vendor's highest ranking for the context-independent use case. Cloudera also ranked well for continuous data loading, one of the factors that raised its rating for all four use cases over the level of 3.0 (meeting requirements).

Cloudera did not provide the same level of access to external data sources as many of the other vendors, which affected its suitability for the logical data warehouse use case. Although almost all vendors did well in terms of survey data for availability and management, Cloudera respondents included some negative comments, and their ease of management was relatively low within the vendor range. Some of these comments may have coincided with Cloudera progressing toward more complex use cases.

Cloudera customers indicated a high level of intent to purchase more of the product in the coming 12 months.

EnterpriseDB (EDB Postgres)

EnterpriseDB ships EDB Postgres (<https://www.enterprisedb.com/products/edb-postgres-platform>) Standard Edition and EDB Postgres Advanced Server, based on the PostgreSQL open-source DBMS.

EnterpriseDB and the PostgreSQL DBMS are new to the DMSA market. Until recently, the open-source RDBMS products have been primarily used for operational and web-based applications. The capabilities and functionality needed to meet the requirements for the DMSA use cases are not traditional strengths of this vendor, hence the scores in the 2 to 3 range for all use cases.

More than 50% of the surveyed customers have database sizes of less than 20TB, and more than 90% of them load data several times per day or less often (over 50% daily or less often). Surveyed customers report more sophisticated users, especially business analysts, followed by expert users. This correlates with the results showing a high number of automated and free-form queries, explaining a stronger score in the operational data warehouse use case. Few surveyed customers reported any external data sources used for queries. This explains the stronger position for both the traditional and operational data warehouse use cases, with few survey customers reporting use for other than these two use cases.

Google (BigQuery)

Google offers BigQuery (<https://cloud.google.com/bigquery/>) on its Google Cloud Platform as a managed, in-memory query execution engine, along with its adjacent products (Cloud Dataproc, Cloud Dataflow) to provide a cloud-based data management solution for analytics.

Google's BigQuery product is specifically designed to address the needs of the DMSA market. BigQuery met the requirements for all four use cases. BigQuery received higher scores for managing large volumes of data and administration and management, and a particularly high rating for system availability — a testimony to the robustness of the Google Cloud Platform infrastructure.

BigQuery is not up to market standards for accessing data residing in external sources, and there is no on-premises version of the product. As such, BigQuery is best suited for use cases where all the data will reside on the Google Cloud Platform.

Absent or weak functionality was cited by a high number of customers as a problem. Google's developer focus and favorable pricing metrics may help to increase adoption of BigQuery, and a large percentage of customers surveyed indicated that they would purchase more of the product in the next 12 months.

Hewlett Packard Enterprise (HPE Vertica)

Hewlett Packard Enterprise offers the HPE Vertica Analytics Platform (<https://www.vertica.com/overview/>) , a columnar relational analytical DBMS. This platform is delivered as a software-only solution, certified configuration (HPE ConvergedSystem reference architectures), SQL on Hadoop, and clouds (HPE Vertica Machine Images for Clouds) that are optimized and tested machine images for AWS and Microsoft Azure with a bring-your-own-license (BYOL) model.

HPE Vertica shows consistent strength in all four of our defined use cases, meeting requirements for all of them and rated in the top half of all vendors evaluated. HPE customers praise performance and scalability, though they cite in-depth knowledge of the platform as a requirement to truly leverage it effectively. Most HPE Vertica customers use it in support of the traditional and operational data warehousing use cases, but there is also a significant showing on the other two use cases — context-independent and logical data warehouses. Nearly half report loading data continuously or several times per day. Among survey respondents, HPE Vertica is overwhelmingly a standard for DMSA in those organizations that have deployed it.

Hortonworks (HDP)

Hortonworks offers the Hortonworks Data Platform (HDP) (<https://hortonworks.com/products/data-center/hdp/>) on Linux and Windows. It also offers Hortonworks DataFlow (HDF) for streaming analytics on an on-premises basis and through various cloud providers. Hortonworks partners with Microsoft (for its Azure HDInsight service) for hybrid on-premises plus cloud deployments, and offers a version of HDP on AWS. A free, laptop-capable sandbox version of HDP is also available.

HDP scores well for managing large volumes of data, continuous data loading and discovery use, and meets requirements for the context-independent data warehouse use case.

HDP allows for data to be stored in AWS S3, which provides data persistence outside of the persistence of a cluster node and therefore for more flexible pricing and implementations in the AWS Cloud.

Hortonworks' philosophy is to deliver an offering that sticks to the open-source code base of Hadoop and associated products. This philosophy is backed by the contributions Hortonworks makes to open source, such as improvements in Hive. The adherence to open source is seen as an advantage to Hortonworks' installed base.

Huawei (FusionInsight)

Huawei offers FusionInsight (<http://e.huawei.com/uk/products/cloud-computing-dc/cloud-computing/bigdata/fusioninsight>) , a data management platform combining components of Apache Hadoop, Spark and Storm, with a proprietary massively parallel processing database (MPPDB). Huawei's MPPDB provides support for the ANSI SQL standard based on PostgreSQL. Its MPPDB can be deployed either as a stand-alone product or within the Hadoop system as an enhancement component. Huawei is new to this Critical Capabilities document this year.

Huawei has added industry-specific domain models and proprietary extensions to the Hadoop platform for event stream processing, graph and machine-learning capabilities. Additional enhancements have been made to the Hadoop scheduler and to the HDFS file system. Huawei also provides cloud services based on its FusionInsight product through the Data Warehouse Service (DWS) and MapReduce Service (MRS) through Huawei's own cloud service in China, and through partnerships with Deutsche Telekom, Orange and Telefónica in Europe.

Like most Hadoop vendor offerings, Huawei FusionInsight scored in the bottom half for the traditional and operational data warehouse use cases, with a stronger showing in the context-independent use case — the realm of data exploration and data science initiatives. FusionInsight scored near the top for number of queries run on a daily basis in support of forecasting and predictive modeling, with only Teradata and Oracle doing better. Respondents to our reference survey praise Huawei's technical support, though the high percentage of references citing support as a strength also raises questions on the overall maturity of the product, which may cause more use of support.

FusionInsight is most frequently deployed in batch-oriented initiatives, with nearly 60% of surveyed customers loading data on a daily basis, but less than 20% loading multiple times per day or continuously.

IBM (DB2)

IBM's current offerings include the traditional solutions of stand-alone DBMS, appliances (IBM PureData System for Analytics, IBM PureData System for Operational Analytics, and IBM DB2 Analytics Accelerator for z/OS), Hadoop through IBM BigInsights, and managed data warehouse services. The database platform as a service offering, IBM dashDB, has added a private cloud capability, while The DataFirst Method supports data lake deployment and management. Our Critical Capabilities evaluation is based on DB2.

IBM DB2 (<https://www.ibm.com/analytics/us/en/technology/db2/>) ranks in the top six for all use cases. It has strengths in the traditional data warehouse and has been a leader in the logical data warehouse, ranking third in each of these use cases. The majority of surveyed customers have database sizes of 50TB or less; however, one reported greater than 1,000TB. The customer survey indicated that more than 80% of customers load data daily or less often, explaining the lower ranking for the operational use case.

IBM DB2 survey results also report a very high number of queries to external data, contributing to the strong ranking for logical data warehouse use.

MapR Technologies (Converged Data Platform)

MapR Technologies offers its Converged Data Platform (<https://www.mapr.com/products/mapr-converged-data-platform>) based on a Hadoop distribution (including both open-source and commercial software) with performance and storage optimizations (use of NFS), streaming, high-availability improvements, and administrative and management tools.

MapR has maintained its strong ranking for the context-independent use case, largely due to its use for discovery and large database support. Survey responses showed the majority of users with database sizes greater than 150TB and almost 40% greater than 1,000TB.

In the other three use cases, MapR scored greater than 3.0 (meets requirements) and ranked in the top 10 for all except traditional data warehousing. Query optimization influenced the MapR scores for both the traditional and logical data warehouse use cases, a general weakness of the Apache Hadoop products. MapR met requirements for all four use cases.

MarkLogic

MarkLogic (<http://www.marklogic.com/>) offers an ACID-compliant NoSQL document store DBMS in Essential Enterprise, Global Enterprise and Mobile editions, and a free, fully featured developer version. It can be deployed via leading cloud and virtualization platforms, including those of AWS, Microsoft Azure and VMware.

MarkLogic focuses on a unified source of metadata to bring data together, whether stored in its repository or accessed through virtualization. The vendor uses this metadata to create logical models and physical implementations of features automatically, like materialized views or optimization of queries. MarkLogic has been promoting its revised positioning in the past year, and has seen a significant improvement in relative ranking for the logical data warehouse and context-independent database use cases. In the more traditional use case areas, its relative ranking dropped, but this was more due to other players improving in these areas and overtaking MarkLogic.

MarkLogic had the highest rating in the ability to access external data sources, as would be expected reflecting its new focus.

Survey data showed that MarkLogic had a relatively high number of customers who had recently implemented solutions with the product. The overall survey also showed that MarkLogic lost in evaluations less frequently than any other vendor, displaying a high level of efficiency in finding appropriate use case opportunities.

MemSQL

MemSQL (<http://www.memsql.com/>) offers a distributed, scale-out, in-memory DBMS combined with a memory and disk-based column store for analytical and operational use cases. MemSQL extends the DBMS platform to include integration with Apache Spark and Apache Kafka for real-time analytics. Multimodel capabilities for relational, key value, document store and geospatial round out the DBMS capabilities. MemSQL is available in on-premises deployments, as a managed service, and on AWS via a CloudFormation template for clusters, or via the AWS Marketplace.

MemSQL's strongest showing is in the operational data warehouse use case, a segment in which it rated No. 1 last year. The strength here is not surprising given MemSQL's focus on the operational DBMS market. The focus on ingesting streaming data is a core differentiator, with more than half of surveyed customers loading data either continuously, or multiple times per day. MemSQL scored third in frequency of data ingest (continuous plus multiple times per day) in our survey. MemSQL's scores for the other three use cases are generally lower. It did, however, meet requirements for the traditional data warehouse use case, making it a strong candidate for when real-time analytics or converged infrastructure for transactional and analytics use cases is important.

Microsoft (SQL Server)

Microsoft offers SQL Server (<https://www.microsoft.com/en-gb/sql-server/sql-server-2016>) as an on-premises solution, Azure SQL Data Warehouse, Azure HDInsight (Hadoop distribution based on Hortonworks) and Azure Data Lake as cloud services. This document evaluates the critical capabilities of SQL Server.

Microsoft improved its relative ranking across all four use cases this year, placing fourth in each except for the context-independent data warehouse use case, in which it placed fifth. Survey results indicated that the most common reason for Microsoft to not be selected in a competitive evaluation was missing or limited functionality.

In each use case, SQL Server rated significantly higher than the rating for "meets requirements" (3.0). The addition of PolyBase, which allows access to external data sources in SQL Server 2016, helped Microsoft achieve a high score in this capability. A significant percentage of survey respondents had been using Microsoft in production for more than 10 years.

Although this Critical Capabilities document only evaluates one product per vendor, Microsoft has the broadest range of cloud and noncloud offerings of any of the established vendors. Survey respondents also indicated a high level of planned purchases in the next 12 months.

MongoDB

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

MongoDB is primarily an operational DBMS used for transaction processing, so it is no surprise that its highest score is for the operational data warehouse use case, where the surveyed customers also reported the highest amount of usage. Due to evolution of criteria and the market, MongoDB's absolute scores dropped this year, although its relative ranking in the use cases rose. MongoDB has all its scores above or near 3.0 (meets requirements).

The majority of survey responses have database sizes of less than 20TB, although one response had a database greater than 150TB in size. This usage is in the lower part of reports on managing large amounts of data and, coupled with poor scores for scalability, does influence all use cases.

MongoDB has the highest reported use by expert users (two to 20 times higher than all other products evaluated) and a high number of queries reported to be repetitive and automated with embedded analytics. Since the majority of surveyed customers have used MongoDB for two to five years, this suggests the level of necessary expertise for DMSA use cases is very high, making the use of MongoDB more difficult.

Oracle

Oracle (<https://www.oracle.com/index.html>) provides the Oracle Database 12c, the Oracle Exadata Database Cloud Machine, the Oracle Big Data Appliance, the Oracle Big Data Management System, Oracle Big Data SQL and Oracle Big Data Connectors. Oracle Cloud provides Oracle Database as a Service, Oracle Exadata Cloud Service, and Oracle Big Data Cloud Service.

Oracle placed second in all four use cases. The gap between Oracle and the leader in each use case was no larger than the gap between Oracle and the third place vendor in all cases. Oracle received the highest score for 6 of 10 capabilities, and only received a score below 3.0 (meets requirements) in the area of discovery use.

Oracle showed improved capabilities in accessing external data this year with a fully realized Big Data SQL, which allows not only access to Hadoop data but also pushes down operations like predicate analysis to the Hadoop platform.

Customers surveyed gave Oracle high marks for system availability. The survey also showed that Oracle's submitted customers had been in production for the second longest amount of time on average, and that the ease of use for the Oracle Database was above average.

Pivotal (Greenplum)

Pivotal offers Greenplum (<https://pivotal.io/pivotal-greenplum>) , a commercial version of the open-source massively parallel processing (MPP) database based on PostgreSQL. It also offers Pivotal HDB based on the open-source Hawq project for SQL processing on top of Hadoop. For this Critical Capabilities document, the focus is on Greenplum.

Pivotal Greenplum scored well in all four of our defined use cases. Greenplum demonstrated strengths in the context-independent data warehouse use case, reflecting the effort that Pivotal has put into developing robust in-database analytics capabilities. In the traditional data warehouse use case category, Greenplum scored near the top. Most customers report using the product in support of traditional data warehouse use cases, with less than 10% doing continuous real-time data loading.

SAP (SAP Hana)

SAP has multiple offerings in the DMSA market, with SAP Hana (<https://www.sap.com/uk/product/technology-platform/hana.html>) chosen for this research. SAP Hana is an in-memory column-store DBMS that supports operational and analytical use cases. It is also offered as an appliance (from more than a dozen hardware vendors), a cloud solution (SAP Cloud Platform, of which SAP Hana is one component) and a reference architecture (SAP Hana Tailored Datacenter Integration [TDI]). SAP Hana can also be deployed on either AWS or Microsoft Azure.

Although SAP Hana has the potential to address the logical data warehouse and context-independent data warehouse use cases, surveyed customers reported almost no queries to external data. The majority of survey responses report using SAP Hana for both DMSA and transactional use, supported by very high numbers for queries supporting embedded analytics, allowing SAP Hana to jump to third position for the operational data warehouse use case. Its position for the logical data warehouse use case is due to the increase in scores for administration and management, support of analytics and system availability.

Snowflake Computing

Snowflake Computing (<https://www.snowflake.net/>) offers a fully managed data warehouse as a service on Amazon Web Services (AWS) infrastructure. Snowflake supports ACID-compliant processing on relational data and document store formats like JSON, Avro, XML and Parquet format files. A native bidirectional Spark connector, R integration, Python integration, support for user-defined functions (UDFs), dynamic elasticity, and temporal support round out the core capabilities. Snowflake is currently only available in the AWS Cloud.

Snowflake scored above 3.0 (meeting requirements) in all four defined use cases for DMSA, with its strongest showing in the traditional data warehouse use case. Snowflake's support for UDFs, along with native multimodel support for document formats like JSON and Avro, see it also placed in the top half for the context-independent data warehouse use case.

Snowflake reference customers largely deploy the solution in support of traditional data warehousing use cases, with only one-third doing real-time or multiple loads per day.

Teradata

Teradata (<http://www.teradata.co.uk/?LangType=2057&LangSelect=true>) 's offerings include a DBMS solution, data warehouse appliances and a cloud data warehouse solution — both on its own managed cloud and on public cloud provider infrastructures like AWS and Microsoft Azure. Teradata offers both traditional solutions and logical data warehouse solutions, which come under what Teradata calls the Unified Data Architecture (UDA). Teradata QueryGrid (part of the UDA) provides multisystem query support via Teradata's own software, as well as open-source Presto.

Teradata offers a combination of tuned hardware and analytics-specific database software, which includes the Teradata Database (on various form factors of appliance, software only, and cloud), Aster Analytics and Hadoop via all three major distributions, as well as analytic consulting services. The core Teradata DBMS was used in this Critical Capabilities research.

Teradata received the top score in all four of our defined use cases, demonstrating a mature product with market-leading depth and breadth of functionality. Every single reference customer reported using Teradata in support of the traditional data warehouse use case, with two-thirds also reporting its use in support of operational and logical data warehouse use cases. Teradata received very high scores for performance and shows strong customer loyalty.

Transwarp Technology (TDH)

Transwarp offers the Transwarp Data Hub (TDH) (<http://www.transwarp.cn/product/tdh?lang=en>) , a full suite of Hadoop distribution components, which is supplemented by its SQL engine, machine learning, NoSQL search engine and stream processing. TDH is available on Microsoft Azure (China). The vendor also offers Hadoop-as-a-service and the hyper-converged infrastructure (HCI) appliance for transactional data.

Unsurprisingly, TDH showed its greatest strength in the context-independent data warehouse use case. The core strengths of Hadoop-based product offerings naturally align to data exploration and data-science-related use cases. However, respondents to our reference survey reported using TDH for all four use cases, with close to 90% using it in support of traditional data warehousing. This indicates that Transwarp's investment in SQL engines on Hadoop with compatibility with major SQL dialects is resonating with its customer base. PL/SQL compatibility and the ease of implementation that it enables were frequently cited as core strengths.

Context

Well over 80% of the market still adheres to traditional approaches; however, we are seeing a growing number of leading organizations evolving toward the logical data warehouse use case to address the volume, variety and velocity needs driven by big data. This trend has accelerated in the past year, with all traditional DMSA vendors adding the capability of accessing data in big data stores to their offerings, sometimes with advanced features such as pushing down some operations to the target systems. Products had to support this type of access to external systems to receive a rating of "meets requirements" for the "queries to multiple data sources" capability. At the same time, mature organizations are turning their data warehouses into active contributors to their operational needs (for example, by embedding analytics in their applications), hence adding to their mission-critical nature.

As organizations start using their data warehousing and management solutions beyond descriptive requirements, and evolve toward predictive and prescriptive analytics, additional user profiles are emerging (data scientist) or growing (data analysts). These new types of users are being added to business analysts and casual users, and are creating new types of demands and interactions with the DMSA, supported by context-independent data warehouses. These demands include native support for new data types and in-database programming languages such as R, Lua or Python, all of which have an impact on workloads and query types.

This year saw the mainstream traditional megavendors add logical data warehouse capabilities to their products, allowing them to access Hadoop-based data stores. Survey results indicated that shared instances of traditional data warehouses (74%) and operational data warehouses (68%) were being widely used, but that logical data warehouses (48%) and context-independent data warehouses (45%) were also being used in a significant number of organizations. Using a single product for multiple use cases is moving toward the norm.

We have noted different groups of vendors within this market space, such as Hadoop-based vendors, other nonrelational vendors, and traditional megavendors. The Hadoop-based group suffered in terms of both execution and vision, while the megavendors' vision improved. Cloud-native vendors are seeing increasing significance, even when they are limited to cloud-only solutions. Cloud-based solutions have also impacted the popularity of appliances, since the ease-of-use factor for appliances is matched and exceeded by cloud solutions.

There is an increasing international presence in the overall market, and although some of these vendors are mainly active within a region (China for Transwarp and Huawei), they will exert increasing pressure on other vendors in the global market moving forward.

Product/Service Class Definition

The various capabilities identified below address the major needs identified above.

Critical Capabilities Definition

Managing Large Volumes of Data

This capability reflects if the volume of data managed by customers is large. This applies to data of multiple data structures and formats.

It plays a role in all use cases, but to various degrees, as it may not be equally important for all. In this context, we have defined "small" as being below 10TB and "large" as being over 150TB, with consideration given to those vendors whose survey respondents reported data stores of 1PB or larger.

In addition to customer experience, this capability takes into consideration the ability of the vendor to address management of query workloads and the availability of price performance optimization options, as well as strategies for query optimization in isolation.

Loading Data Continuously

This capability represents the prevalence of data being loaded continuously by customers. Some use cases — more than others — require data to be loaded from the operational sources in near-real time.

This capability does not assume all data is loaded continuously. Bulk and batch loading remains the most common loading process.

In addition to customer experience, this capability takes into consideration the vendor's ability to support ingestion of streaming data, and the ability to perform continuous updates for read optimization.

Query Optimization

This capability reflects the prevalence of the use of repetitive queries by customers across all types of queries and the capabilities of the vendor product.

The prevalence of repetitive queries in workloads is indicative of data warehouses supporting a wide set of casual users that only consume data that has been prepared for them. Repetitive queries can also extend to supporting dashboards that offer limited and prebuilt interactions.

In addition to customer experience, this capability takes into consideration the ability to perform automation of optimization and workload management, as well as the features and proven capabilities of the underlying optimizer.

Queries Support Analytics

This capability reflects the prevalence of advance analytics queries by customers across all types of queries.

Advanced analytics queries support forecasting, predictive modeling, in-database execution of R (an open-source statistical programming application), user-defined functions or other mining styles, as well as free-form analytical queries based on data marts, views, cubes or a semantic-enabled modeling interface. This criterion includes consideration of the integration of analytic languages such as R into the native capabilities of an offering.

In addition to customer experience, this capability takes into consideration the ability to support complex ad hoc queries, and to handle mixed workloads and data temperatures.

Queries to Multiple Data Sources

This capability reflects the prevalence of queries across multiple data types and sources by customers across all types of queries.

Queries across multiple data types and sources access data in other sources beyond the database management system, such as other RDBMSs or Hadoop distributions.

This capability is also rated on the functionality implemented in access to external data sources, such as whether some processing, like predicate evaluation, is passed to the external data source for implementation locally.

Operational BI Use

This capability reflects the prevalence of queries for operational reporting or embedded analysis in applications as part of the overall set of queries among customers, as well as features for operational BI usage.

While operational BI queries are primarily used in support of the operational data warehouse use case, other use cases also utilize these type of queries. Three of the four use cases attempt to support operational BI in some way, but do not have the same level of direct contribution.

In addition to customer experience, this capability takes into consideration the ability to perform operational queries with high service levels, as well as support for running in-database prescriptive analysis.

System Availability

This capability outlines the ability of the vendor to support general production availability, high availability and disaster recovery as expressed by customers.

In addition to customer experience, this capability takes into consideration the vendor's ability to provide high availability and disaster recovery, as well as the ease of deployment and management of these features.

Administration and Management

This capability demonstrates the vendor's or product's ease of implementation, upgrade and ease of use as expressed by customers.

It covers overall ease of administration and management, not only during implementation but also during ongoing use and upgrade phases. Scoring is also affected by the complexity of deployment and by vendor history. Some vendors have recent offerings for which upgrades may not yet have been released.

In addition to customer experience, this capability takes into consideration the completeness of vendor administration capabilities, such as role-based activities, advisors, utilization and capacity planning, resource allocation features and user interface, as well as complexity of deployment and management.

Traditional Data Warehouse Use

In this year's Critical Capabilities calculations, business analysts and casual users were classified as traditional data warehouse users, while data scientists and data miners were classified as discovery users.

The criteria for traditional data warehouse use were based, in large part, on the relative percentage of users classified as business analysts or casual users. These skill sets were defined as:

Business analyst — Utilizes online analytical processing and dimensional tools to create new objects, some faculty with computer languages and computer processing techniques.

Casual user — Regularly uses portals and prebuilt interfaces, minimally capable of designing dimensional analytics (if at all).

We also took into consideration some survey results and product evaluations relating to traditional data warehouse usage.

Discovery Use

In this year's Critical Capabilities calculations, data scientists and data miners were classified as discovery users, while business analysts and casual users were classified as traditional data warehouse users.

The criteria for discovery use were based, in large part, on the relative percentage of users classified as data scientists or data miners. These skill sets were defined as:

Data scientist — Expert in statistics, abstract mathematics, programming, business processes, communications and leadership.

Data miner — Expert in data, statistical software, statistical models; fully aware of computer processing "traps" or errors.

We also took into consideration some survey results and product evaluations relating to discovery use.

Use Cases

Traditional Data Warehouse

This use case involves managing historical data coming from various structured sources. Data is mainly loaded through bulk and batch loading.

The traditional data warehouse use case can manage large volumes of data, and is primarily used for standard reporting and dashboarding. To a lesser extent, it is used for free-form querying and mining, or operational queries. It requires high capabilities for system availability and administration and management, given the mixed workload capabilities for queries and user skills breakdown. Query optimization plays a role in this use case, as many of the BI-style queries are used repetitively, multiplying the effects of the optimization.

Operational Data Warehouse

This use case manages structured data that is loaded continuously in support of embedded analytics in applications, real-time data warehousing and operational data stores.

The operational data warehouse use case primarily supports reporting and automated queries to support operational needs, and will require high-availability and disaster recovery capabilities to meet operational needs. Managing different types of users or workloads, such as ad hoc querying and mining, will be of less importance as the major driver is to meet operational excellence. Query optimization plays a role in this use case, as many of the BI style queries are used repetitively, multiplying the effects of the optimization.

Logical Data Warehouse

This use case manages data variety and volume of data for both structured and other content data types.

Besides structured data coming from transactional applications, this use case includes other content data types such as machine data, text documents, images and videos. Because additional content types can drive large data volumes, managing large volumes is an important criterion. The logical data warehouse is also required to meet diverse query capabilities and support diverse user skills.

This use case supports queries reaching into other sources than the data warehouse DBMS alone. The logical data warehouse encompasses the features of the other use cases, combined with the ability to include a variety of data sources in analytic operations.

Context-Independent Data Warehouse

This use case concerns new data values, variants of data form and new relationships. It supports search, graph and other capabilities for discovering new information models.

It is primarily used for free-form queries to support forecasting, predictive modeling or other mining styles, as well as queries supporting multiple data types and sources. It has no operational requirements and favors advanced users such as data scientists or business analysts, resulting in free-form queries across potentially multiple data types. The context-independent data warehouse use case is more likely to use siloed data that has not been integrated with other data or subjected to full data governance practices.

Vendors Added and Dropped

Added

EnterpriseDB

Google

Hortonworks

Huawei

Snowflake Computing

Dropped

Actian is no longer actively engaging in data management solutions for the analytics market.

Exasol did not meet the revenue criterion.

Kognitio did not meet the revenue criterion.

Infobright did not meet the revenue criterion.

Inclusion Criteria

To be included in this Critical Capabilities research, vendors had to meet the following criteria, which is the same criteria as for our "Magic Quadrant for Data Management Solutions for Analytics."

To be included in this Critical Capabilities document, vendors had to meet the following criteria:

Vendors had to have DMSA software generally available for licensing or supported for download for approximately one year (since 1 December 2015). 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 solutions, 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 refused to move to newer versions.

Product evaluations included technical capabilities, features and functionality present in the product or supported for download on 1 December 2016. 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 20 verifiable production implementations that exhibited generated revenue from 20 distinct organizations with data management solutions for analytics, indicating they were in production and were satisfying either of the below conditions:

- Had a minimum of \$10 million in revenue with a 50% growth rate year over year
- Had more than \$40 million in revenue (revenue could be from licenses, support and/or maintenance)

The production customer base had to include customers from three or more vertical industries (see list in Note 1).

Customers in production needed to deploy DMSA 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).

Vendors had to demonstrate production customers from at least two distinct geographic regions (see list Note 2).

Any acquired product had to have been acquired and offered by the acquiring vendor as of 30 June 2016.

Support for the included data management for analytics product(s) had to be available from the vendor. We also considered products from vendors that control or contribute specific technology components to 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 evaluated DMSA. However, vendors in this Critical Capabilities report needed to offer the ability to manage physical persistence of the data.

Vendors had to provide support for at least one of the four major use cases.

Vendors had to at least provide relational processing. Depth of processing capabilities and variety of analytical processing options were considered as advantageous in the evaluation criteria.

Vendors participating in the 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 Critical Capabilities report.

We also considered the following factors when deciding whether products were eligible for inclusion:

RDBMS.

Nonrelational DBMS

Hadoop distributions.

No specific rating advantage was given with regard to the type of data store used (for example, RDBMS, graph DBMS, HDFS, key-value DBMS, document DBMS, wide-column DBMS).

Multiple solutions used in combination to form a single data management solution for analytics were considered valid, but each solution must demonstrate maturity and customer adoption.

Cloud solutions were considered viable alternatives to on-premises solutions. The ability to manage hybrid on-premises and cloud solutions is considered advantageous for inclusion.

Open-source solutions.

Gartner may include at its discretion additional vendors in cases of known use for classified but unspecified cases.

The following technology categories are excluded:

BI and analytical solutions that only offer a DMSA that is embedded, or that embed a DMSA from another provider.

BI and analytical solutions that only offer a DMSA that is limited specifically to the vendor's own BI and analytical solution, or whose customers exhibit only using the solution within the same vendor stack.

In-memory data grids.

Prerelational DBMS.

Object-oriented DBMS.

Gartner analysts are the sole arbiters of which vendors and products are included in this Critical Capabilities document.

Table 1. Weighting for Critical Capabilities by Use Case

Critical Capabilities	Traditional Data Warehouse	Operational Data Warehouse	Logical Data Warehouse	Context-Independent Data Warehouse
Managing Large Volumes of Data	15%	10%	10%	25%
Loading Data Continuously	5%	15%	5%	0%
Query Optimization	20%	15%	10%	0%
Queries Support Analytics	10%	10%	10%	20%
Queries to Multiple Data Sources	5%	0%	15%	10%
Operational BI Use	5%	20%	5%	0%
System Availability	10%	15%	10%	5%
Administration and Management	10%	10%	15%	10%
Traditional Data Warehouse Use	20%	5%	10%	5%
Discovery Use	0%	0%	10%	25%

Total	100%	100%	100%	100%
As of February 2017				

Source: Gartner (March 2017)

This methodology requires analysts to identify the critical capabilities for a class of products/services. Each capability is then weighed in terms of its relative importance for specific product/service use cases.

Critical Capabilities Rating

Each of the products/services has been evaluated on the critical capabilities on a scale of 1 to 5; a score of 1 = Poor (most or all defined requirements are not achieved), while 5 = Outstanding (significantly exceeds requirements).

Table 2. Product/Service Rating on Critical Capabilities											
Critical Capabilities	1010data (Insights Platform)	Amazon Web Services (Amazon Redshift)	Cloudera (Cloudera Enterprise)	IBM	EnterpriseDB (EDB Postgres)	Google (BigQuery)	Pivotal (Greenplum)	SAP (SAP Hana)	Hortonworks (HDP)	Huawei (FusionInsight)	MapR Technology (Converged Data Platform)
Managing Large Volumes of Data	3.0	4.0	4.0	3.5	2.0	4.0	4.0	2.0	4.0	3.0	4.5
Loading Data Continuously	1.0	3.0	5.0	3.0	2.0	3.5	4.0	4.0	5.0	4.0	5.0
Query Optimization	3.0	3.0	3.0	5.0	3.0	3.0	4.0	4.0	2.0	3.0	2.0
Queries Support Analytics	4.0	3.0	3.0	4.0	3.0	3.0	4.0	5.0	2.0	3.0	2.0
Queries to Multiple Data Sources	2.0	2.0	2.0	5.0	1.0	2.0	3.0	4.0	2.0	3.0	2.0
Operational BI Use	2.0	2.0	2.0	4.0	4.0	2.0	3.0	5.0	2.0	2.0	3.0
System Availability	4.0	5.0	3.0	3.0	3.0	5.0	3.0	4.0	3.0	3.0	4.0
Administration and Management	3.0	3.5	4.0	3.0	3.0	4.0	3.0	3.0	2.0	2.0	4.0
Traditional Data Warehouse Use	5.0	4.0	3.0	5.0	2.0	3.0	4.0	3.0	2.0	3.0	3.0
Discovery Use	2.0	2.0	5.0	3.0	2.0	3.0	3.0	3.0	4.0	4.0	5.0

Source: Gartner (March 2017)

Table 3 shows the product/service scores for each use case. The scores, which are generated by multiplying the use-case weightings by the product/service ratings, summarize how well the critical capabilities are met for each use case.

Table 3. Product Scores by Use Case											
Use Cases	1010data (Insights Platform)	Amazon Web Services	Cloudera (Cloudera Enterprise)	IBM	EnterpriseDB (EDB Postgres)	Google (BigQuery)	Pivotal (Greenplum)	SAP (SAP Hana)	Hortonworks (HDP)	Huawei (FusionInsight)	MapR Technology (Converged Data Platform)

		(Amazon Redshift)									Data Platform
Traditional Data Warehouse	3.40	3.50	3.25	4.13	2.55	3.38	3.70	3.55	2.55	2.90	3.18
Operational Data Warehouse	2.85	3.30	3.30	3.75	2.90	3.38	3.55	3.95	2.80	2.85	3.45
Logical Data Warehouse	3.00	3.18	3.35	3.90	2.40	3.28	3.45	3.60	2.65	2.95	3.35
Context-Independent Data Warehouse	3.00	3.10	3.75	3.63	2.25	3.35	3.50	3.30	3.05	3.15	3.73

Source: Gartner (March 2017)

To determine an overall score for each product/service in the use cases, multiply the ratings in Table 2 by the weightings shown in Table 1.

Evidence

This research is based on:

Gartner inquiry data on the topics of data warehousing, data integration and metadata management.

In-depth interviews with reference customers provided by vendors.

Gartner strategic advisory service, full-day sessions with end-user organizations in which Gartner is asked to review client implementation plans and designs.

Gartner's respondent survey (301 respondents) for "Magic Quadrant for Data Management Solutions for Analytics" for year 2016 (survey completed in 4Q15).

Note 1 Vertical Industry Sectors

Accommodation and food services

Administrative and support and waste management and remediation services

Agriculture, forestry, fishing and hunting

Arts, entertainment and recreation

Construction

Educational services

Finance and insurance

Healthcare and social assistance

Information

Management of companies and enterprises

Manufacturing

Mining

Professional, scientific and technical services

Public administration

Real estate rental and leasing

Retail trade

Transportation and warehousing

Utilities

Wholesale trade

Note 2

Geographic Regions

- North America (Canada and the U.S.)
- Latin America (including Mexico)
- Europe (Western and Eastern Europe)
- The Middle East and Africa (including North Africa)
- Asia/Pacific (including Japan)

Critical Capabilities Methodology

This methodology requires analysts to identify the critical capabilities for a class of products or services. Each capability is then weighted in terms of its relative importance for specific product or service use cases. Next, products/services are rated in terms of how well they achieve each of the critical capabilities. A score that summarizes how well they meet the critical capabilities for each use case is then calculated for each product/service.

"Critical capabilities" are attributes that differentiate products/services in a class in terms of their quality and performance. Gartner recommends that users consider the set of critical capabilities as some of the most important criteria for acquisition decisions.

In defining the product/service category for evaluation, the analyst first identifies the leading uses for the products/services in this market. What needs are end-users looking to fulfill, when considering products/services in this market? Use cases should match common client deployment scenarios. These distinct client scenarios define the Use Cases.

The analyst then identifies the critical capabilities. These capabilities are generalized groups of features commonly required by this class of products/services. Each capability is assigned a level of importance in fulfilling that particular need; some sets of features are more important than others, depending on the use case being evaluated.

Each vendor’s product or service is evaluated in terms of how well it delivers each capability, on a five-point scale. These ratings are displayed side-by-side for all vendors, allowing easy comparisons between the different sets of features.

Ratings and summary scores range from 1.0 to 5.0:

- 1 = Poor: most or all defined requirements not achieved
- 2 = Fair: some requirements not achieved
- 3 = Good: meets requirements
- 4 = Excellent: meets or exceeds some requirements
- 5 = Outstanding: significantly exceeds requirements

To determine an overall score for each product in the use cases, the product ratings are multiplied by the weightings to come up with the product score in use cases.

The critical capabilities Gartner has selected do not represent all capabilities for any product; therefore, may not represent those most important for a specific use situation or business objective. Clients should use a critical capabilities analysis as one of several sources of input about a product before making a product/service decision.



(https://www.gartner.com/technology/contact/become-a-client.jsp?cm_sp=bac_-_reprint_-banner)

© 2017 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) 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. 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)
 - Privacy (<https://www.gartner.com/privacy>)

[Site Index \(http://www.gartner.com/technology/site-index.jsp\)](http://www.gartner.com/technology/site-index.jsp)

[IT Glossary \(http://www.gartner.com/it-glossary/\)](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)