Gartner

Predicts 2021: Data Management Solutions — Operational Efficiency Rises to the Top

Published 1 December 2020 - ID G00735776 - 26 min read

By Analysts Adam Ronthal, Donald Feinberg, Ehtisham Zaidi, Mark Beyer, Guido De Simoni, Andrew White, Ted Friedman, Melody Chien

Initiatives: Data Management Solutions; Legal and Compliance Technology and Analytics

The diversity of data production, consumption and persistence is driving interest in new ways to deploy, integrate and govern data across the enterprise. Data and analytics leaders must consider these predictions and their implications as they plan for a radically changed data management landscape.

Overview

Key Findings

- Cloud DBMS revenues for 2019 were $17 billion, up 54% from 2018; 99% of that revenue is attributed to the top 13 vendors (cloud solution providers and the largest independent software vendors), leaving only 1% for the hundreds of smaller vendors in the market.
- Data integration, data preparation and cleansing, data ingestion, and data exploration are among the top eight technologies that organizations would like to automate within a year. [1]
- Organizations tend to architect and deploy solutions oriented around data sharing and governance on a piecemeal basis. This reflects historical organizational structures as well as a lack of awareness on the business side of the data's dependency on business outcomes: the same piece of data might be subject to multiple policies at different times.
- According to Gartner's 2020 Data Quality Solution Magic Quadrant Survey, poor data quality costs organizations an average of $12.9M per year. [2] This number is likely to rise as business environments become increasingly digitized and complex.

Recommendations

As data and analytics leaders responsible for data management solutions and seeking greater operational efficiencies in cloud operations, data governance and data quality initiatives, you should:

- Pay attention to the cloud database management system (DBMS) market and the changes taking place. You must be ready to change decisions around the database-platform-as-a-service (dbPaaS) offerings that are "standard" in the organization.

- Establish a data fabric strategy by evaluating your current tools for their ability to collect and share metadata and data (such as data catalogs, metadata management solutions, data integration, data virtualization, semantic technology and DBMSs) and plan a way forward to greater integration.
- Classify your data to be governed based on its value and importance to prioritized business outcomes. Orient your master data management (MDM), data quality and enterprise metadata Management program to this and they will be leaner, smaller, more adaptable and more likely to drive business benefit.
- Maximize the effectiveness of data quality initiatives by evaluating data quality vendors that are able to read metadata and business rules and policies from data catalogs and remediate quality issues in self-service approaches.

Strategic Planning Assumptions

By 2025, 50% of independent database management system (DBMS) vendors will cease operations, causing customers to adjust strategies and migrate back to their strategic DBMS suppliers.

By 2024, organizations that utilize active metadata to enrich and deliver a dynamic data fabric will reduce time to integrated data delivery by 50% and improve the productivity of data teams by 20%.

By 2024, 75% of organizations will have deployed multiple data hubs to drive mission-critical data and analytics sharing and governance.

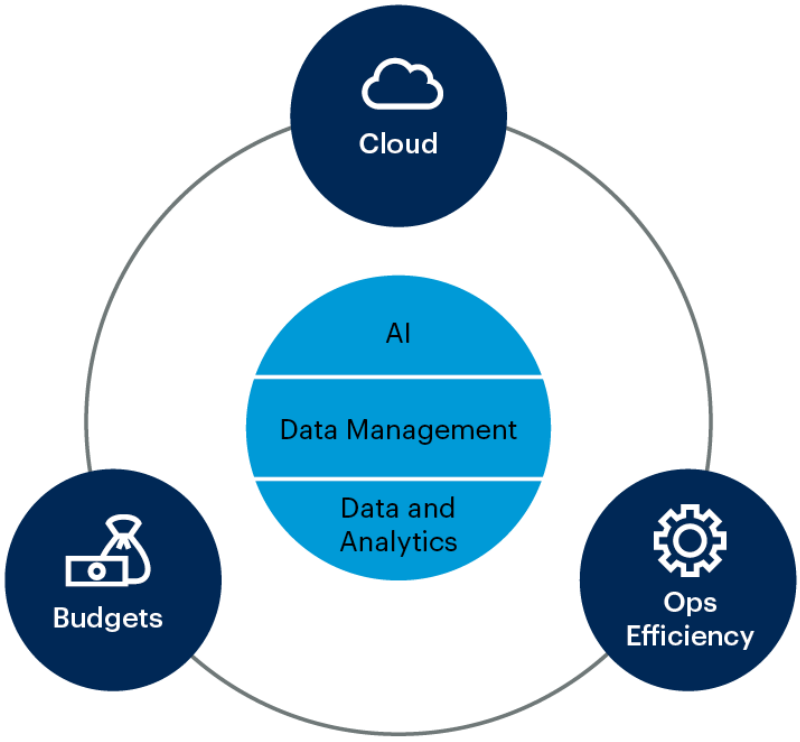
Through 2024, 50% of organizations will adopt modern data quality solutions to better support their digital business initiatives.

Analysis

What You Need to Know

Data is now more distributed and varied than ever before, driving new demands from users and an increasing diverse set of data producers. A typical enterprise has many data persistence stores, multiple sources of data that they are trying to integrate, all spread out across on-premises, hybrid cloud and multiple CSP environments. It is hardly a surprise that the top three topics of inquiry to Gartner's broad data and analytics team over the last six months were cloud, budgets and operational efficiency.

Cloud and Operational Efficiency Lead Inquiries to Gartner's Data and Analytics Teams



n = 8,225

Source: Gartner

Inquiries to KIs: Analytics, BI and Data Science Solutions; Artificial Intelligence; Data and Analytics Leaders; Data and Analytics Strategies; Data Management Solutions; Data Management Solutions for Technical Professionals, April – September 2020

735776_C

Gartner

The four predictions explored in this research align with these core areas of interest. The rise of cloud will have a broad impact on data management vendors in the market, which will certainly affect the available options to data and analytics leaders as they plan their long term cloud strategies, as well as how they budget for them.

Data hubs have emerged as a core approach to solving data integration, governance and sharing challenges, and if the predictions in this research come to pass, most organizations will be working with multiple data hubs. Data fabrics remain an emerging key area of interest, enabled by active metadata. The relationship between data hubs and data fabrics continues to evolve and may ultimately coalesce. Finally, data quality continues to be a challenge, as its absence throws all other data and analytics efforts into doubt. Data quality initiatives will necessarily be a key component of data hub and data fabric discussions.

Cloud, data hubs, data fabrics and data quality initiatives will be intertwined as data and analytics leaders strive for greater operational efficiency across their data management landscape.

Strategic Planning Assumptions

By 2025, 50% of independent database management system (DBMS) vendors will cease operations, causing customers to adjust strategies and migrate back to their strategic DBMS suppliers.

Analysis by: Donald Feinberg

Key Findings:

- Cloud DBMS revenues for 2019 were $17 billion USD, up 54% from 2018; 99% of that revenue is attributed to the top 13 vendors (CSPs and largest ISVs), leaving only 1% for the hundreds of smaller vendors in the market.
- Open source dbPaaS from the CSPs are growing in adoption, displacing many on-premises choices, even those with cloud services.
- CSPs are increasingly offering competition to the ISVs, displacing many smaller vendors. This also forces smaller vendors and new startups to offer dbPaaS products to compete with the larger ISVs and CSPs.

Market Implications:

Over the next five years, data and analytics leaders must pay attention to the cloud DBMS market, as there will be many shifts, mostly toward the CSPs and largest ISVs. The DBMS market today has hundreds of choices and the majority of them are small vendors. The top 13 vendors (by cloud revenue) account for greater than 99% of the cloud dbPaaS market. This includes the eight CSPs (Alibaba, Amazon Web Services (AWS), Google Cloud Platform (GCP), IBM, Huawei, Microsoft, Oracle and Tencent; accounting for 95.9% of the revenue) and the five largest ISVs (Cloudera, MongoDB, SAP, Snowflake and Teradata; accounting for 3% of the revenue). We do not believe that over the next five years, this high-percentage majority will vary much below the 99% of today.

Several trends in the cloud DBMS market will accelerate this consolidation of vendors in the space:

- **Multicloud:** There are two perspectives here: CSP and ISV. First, we believe a growing number of enterprises support more than one CSP. In many cases, this is a result of application gravity, where the customer already uses one or more services from each CSP. Remaining with the services from these CSPs reduces financial governance complexity while offering even greater flexibility of choice. Second, the larger ISVs offer greater flexibility in supporting multiple CSP platforms (multicloud), allowing organizations greater flexibility of choice. However, for the ISVs, multicloud support comes at a cost that may be more than a smaller ISV can afford.
- **Open-source compatible alternatives:** The CSPs are introducing an increasing number of products with “wire compatibility” to an open source database management system (OSDBMS). The front-end

of the dbPaaS is API compatible with the OSDBMS, while the CSPs' infrastructure allows for greater performance, use of cloud object storage and stronger availability, adding a level of compatibility from one CSP to another. Although the majority of OSDBMS offerings today are PostgreSQL- and MySQL-based, an increasing number of competitors are arising, such as Cassandra and MongoDB.

- **Financial governance complexity:** Financial governance is becoming a major challenge in the cloud, with many different pricing models, elasticity and serverless offerings, in addition to the many disparate vendors. The greater the number of vendors, the greater the number of pricing models, which creates more complex financial governance. By consolidating to a smaller number of vendors, financial governance is more manageable and less complex.
- **"Good enough":** Many of the smaller DBMS vendors have products that are either unique or claim superior performance. As the CSPs release new products that equal this unique functionality, customers begin to realize the performance difference is not so great and in fact, the CSPs' offerings are getting even better. If the CSP has a product with the same functionality that is "good enough" for the project and application, the argument for using a single vendor (the CSP) becomes strong. Not only does this reduce application integration costs and complexities, it also simplifies financial governance.
- **Large ISV momentum and penetration:** There are several ISVs which are multicloud, large and have good market penetration. Additionally, they may have a large on-premises presence or a strong customer base. Examples are SAP with their application customers, Cloudera as the sole standing Apache Hadoop distribution vendor and Teradata with a loyal customer base and of 40-year history.

We believe these trends will drive customers to consolidate to a smaller number of vendors, thus reducing the overall vendor landscape in the cloud DBMS market.

Recommendations:

- Pay attention to the cloud DBMS market and the changes taking place. You must be ready to change decisions around the dbPaaS offerings that are "standard" in the organization.
- For the least risky strategy, choose a dbPaaS offering from the CSPs and the largest ISVs.
- Hesitate jumping on "shiny, new objects" when choosing a dbPaaS solution. Your standard CSP or ISV may already support or supply a choice that is "good enough."
- For financial governance, remember "the fewer the better," as this approach is far less complex than managing many disparate products and contracts.

Related Research:

Market Share: Enterprise Public Cloud Services, Worldwide, 2019

Choosing Between Multimodel DBMS and Multiple Specialized Engines

The Future of the DBMS Market Is Cloud

Understanding and Planning for Database Management Transformation to the Cloud

By 2024, organizations augmenting their data fabrics with active metadata will reduce the overall time to integrated data delivery by 50% and improve the productivity of their data teams by 20%.

Analysis by: Ehtisham Zaidi, Mark Beyer, Guido De Simoni

Key Findings:

- Data integration (44%), data preparation and cleansing (29%) and data ingestion (21%) are among the top 10 technologies that organizations would like to automate within the next two years.
- Cloud data management solutions are regularly crossing boundaries of old data management practices that historically separated data quality, master data management, data governance and data integration as point solutions.
- Data management teams are spending more time catching up with tactical data pipeline delivery through manual data preparation and engineering. Moreover, organizations are investing in fewer data engineers despite the rising demand for integrated data, adding to their supply-demand imbalance.
- Data management initiatives are most often prioritized by balancing business demand, cost, staff availability and planned time-to-deliver, as opposed to focusing on efficient delivery models to apply during the “next” project.

Market Implications:

As data and analytics leaders continue to struggle with the ever surmounting challenges of cataloging heterogenous data assets, lack of automation in integration design and delivery, complex (and seriously expensive) migration plans and a lack of experienced data engineers, modern data fabric architectures are now not only necessary investments, but urgent ones.

Data fabrics are not a technology that companies can buy and implement at this stage. Data fabrics as a design concept augment “intentional data management” and “governance expectations” with “observable experiences” for flexible data management, integration and orchestration.

- Designed data models, integration and usage patterns are enriched with additionally inferred metadata through continuous analytics to document actual experience as observable data utilization.
- New definitions, data models, data integration, business intelligence, analytics, data sharing and so forth are discovered.

A data fabric is augmented in nature, with artificial intelligence (AI)and graph technologies as key enablers. At its most mature state, a data fabric provides flexible and augmented data integration and orchestration (see Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration).

Modern data management use cases require a data fabric design that can converge, consolidate and deliver multiple data management technologies as services based on active metadata utilization and analysis. Data fabrics enriched with active metadata will liberate data management teams from having to manually plan, manage, implement and retire data pipelines based on the ever-changing requirements of the business. These data fabrics will reduce the time to integrated data delivery, not only by reusing existing mappings, but also by preempting upcoming transformations of and changes to the data management infrastructure, and by recommending the most optimal paths for data execution (including orchestration, persistence and execution engines).

Current data fabric designs (which do not accommodate active metadata analysis) can consolidate myriad different data management tools and technologies and provision integrated views of data based on the most optimal data delivery modes. This is already a big win for most organizations. However, these fabrics are not yet dynamic, in that they cannot predict or automate future changes to data integration demand. This is where organizations find it incredibly challenging to continue to invest in data fabrics to keep them up to date with current metadata for newer use cases or existing ones (when schema change). Data management teams that are already struggling to stay productive cannot further be tasked with keeping data fabrics current, which will only deteriorate their productivity. This is where active metadata is important to impart the much-needed automation to the most critical aspects of the data fabric, i.e., knowledge graph creation and updation.

Data fabrics can be increasingly automated (in terms of design, deployment, maintenance and evolution) if an enterprise metadata management approach is used to utilize active metadata as input to the data fabric design. The data fabric that uses active metadata is constantly evolving, is more dynamic and in some cases even self-driving. It uses performance metadata, location metadata, frequency of access metadata, and so on to autofix pipeline failures as a result of schema drift or performance challenges. In order to make the metadata broadly accessible and usable, it is important for organizations to have the right technology architecture to search for, connect to, collect and analyze metadata. These are all important steps toward being able to activate metadata. For a deeper explanation of active metadata analysis and enabling technologies, see Critical Capabilities for Metadata Management Solutions. The overall effects are to lower enterprise data management costs and present the most appropriate cost justification to date for metadata in the enterprise. Those organizations that are able to activate their existing passive metadata and use it as an input to enrich their data fabrics will continue to drive significantly lower spending in manual data integration design and delivery. They will be able to drive down costs significantly due to the automation of otherwise error-prone manual tasks arising from frequent requests by business teams. These tasks include data movement, schema assignment and other mundane tasks involved in the creation, management and operationalization of data pipelines.

Recommendations:

- When drafting a roadmap for data fabric enabling technologies, first evaluate your existing tools for their ability to collect, analyze and share metadata (such as data catalogs, metadata management solutions, data integration, data virtualization, semantic technology and DBMSs). Tools that are unable to collect and share metadata cannot be part of the data fabric design.
- When cataloging metadata, include all possible inputs including systems data, usage logs, query logs and more, and build an inventory of metadata. These inputs can support human analysts and machine learning (ML) in optimizing your data fabric design.
- Pilot metadata export/import solutions combined with federated approaches. Choose providers that are pursuing active metadata-driven ML-enabled initiatives.
- Prioritize providers that can not only inventory your existing metadata, but also enable a graph of your metadata along with its intricate relationships in the form of triples, and then assist with continuous graph analytics and ML over this connected metadata graph. This would enable insightful recommendations, ranging from basic tasks, such as repetitive data transformation automation, to complex data migration path recommendation.

Related Research:

Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration

How to Activate Metadata to Enable a Composable Data Fabric

Augmented Data Catalogs: Now an Enterprise Must-Have for Data and Analytics Leaders

Modernize Your Data Integration Architecture for Digital Business by Combining Data Delivery Styles

Cool Vendors in Data Management

By 2024, 75% of organizations will have deployed multiple data hubs to drive mission-critical data and analytics sharing and governance

Analysis by: Andrew White, Ted Friedman

Key Findings

- Organizations tend to architect and deploy solutions oriented around data sharing and governance on a piecemeal basis. This reflects historical organizational structures as well as a lack of awareness on the business side of data's dependency on business outcomes: the same piece of data might be subject to multiple policies at different times
- Due to the fragmented view described above, business and IT costs related to sharing and governing data are often higher than they should be, and the benefits lost or not recouped. Though everyone

agrees consistent sharing and governance is important, it remains a huge challenge.

Market Implications

For thirty years the focus of data integration and governance was deep inside the centralized IT organization and focused on the ETL work associated with a data warehouse. This approach enjoyed mostly mediocre success in general, not least because business involvement — required to set policy and enforce it — was far removed and not motivated as much to help govern data in a data warehouse. The rise of the data lake in recent years has brought this situation into focus, and the COVID-19 pandemic has made what was a chronic issue into an acute challenge, since data is at the center of every business decision (see The State of Data and Analytics Governance Is Worse Than You Think). In addition, the point-to-point approaches to integration and sharing of data are now catching up to organizations as they are forced to deliver greater consistency and move faster. Current environments don't enable them to scale.

There have been many innovations to help address these well-known issues. MDM and improved data quality programs have sought to help, and have done so with some success. A new generation of data catalogs have been thrown at the problem, mostly by hasty vendors, in order to find data that might warrant governance. A new generation of ML-based graph analysis is making headway in discovering important data in organizations (see How Augmented Data Management Capabilities Are Impacting MDM and Data Governance). And APIs, data virtualization and event streaming technologies have sought to make data and systems more accessible and connected. All of these have tended to focus on a technological approach to help solve the challenges. What has been missing has been a more orchestrated or architectural approach that seeks to align and link all the discrete technologies and the important missing work that business leaders need to be involved in.

A data hub strategy emerged a few years ago, but it remains misunderstood, mainly due to the confusing use of the term "hub." In our definition, a data hub strategy combines the work involved in data and analytics governance (business leaders set policy; business users enforce policy) with the requirements for data sharing across the organization. To this is added an orchestration of all the technology-centric solutions. Implementing MDM leads to a data hub that connects to many business applications (one node of many), which is the focal point of the work of setting policy and identifying data sharing requirements. This is a good example of a data hub but there are others (see Data Hubs: Understanding the Types, Characteristics and Use Cases). For example, a data hub may help streamline and enforce integration transformations and data quality policies on data moving from applications to a data warehouse or data lake.

As such, organizations that have implemented MDM may well have already implemented their first data hub. An implementation of a reference data or application data hub for governance and sharing of said data are other forms of hubs. Long-term orchestration and consolidation of discrete approaches and technologies have led to simpler data management landscapes and a more focused effort to improve the governance and sharing of data in key locations across that landscape, leading to lower business and IT costs and an increase in the business value of that data.

Recommendations

- Classify your business data to be governed based on value and importance to prioritized business outcomes. Orient your MDM, data quality and enterprise metadata management program to this and they will be leaner, smaller, more adaptable and more likely to drive business benefit.
- Only seek to govern data that is aligned to the most important business outcomes. Use outcome priorities to scale your data and analytics governance and sharing program, one outcome at a time.
- Prioritize the least amount of work related to data governance to precede your data integration approach. It will yield the requirements you will need to determine where, what type and how many data hubs you will require over time.
- Begin to shift point-to-point integration patterns toward data hub architectures where sharing and governance challenges are most acute.

Related Research:

Data Hubs, Data Lakes and Data Warehouses: How They Are Different and Why They Are Better Together

Design an Effective Information Governance Strategy

Use a Data Hub Strategy to Meet Your Data and Analytics Governance and Sharing Requirements

Implementing the Data Hub: Architecture and Technology Choices

Infuse Your Data Hub Strategy with Data and Application Integration

Through 2024, 50% of organizations will adopt modern data quality solutions to better support their digital business initiatives.

Analysis by: Melody Chien

Key Findings

- According to Gartner's 2020 Data Quality Solution Magic Quadrant Survey, poor data quality costs organizations an average of $12.8M per year. This number is likely to rise as business environments become increasingly digitized and complex.
- The acceleration of digital initiatives and the dynamics of business change are creating demand for greater scale and shorter time to value in relation to data quality improvement.

- Data quality initiatives are increasingly driven by business teams (instead of IT), which creates demand-driven trends for intelligent capabilities in self-service approaches to help business users with greater insights and simpler tasks.
- Organizations are increasingly embracing software-as-a-service (SaaS) or platform-as-a-service (PaaS) deployment options for data quality solutions for cost optimization purposes, as well as to improve agility and flexibility.

Market Implications

According to Gartner's 2020 Data Quality Solution Magic Quadrant Survey, poor data quality costs an organization an average of $12.8M per year, making data errors extremely costly for business. Not only does poor data quality cause financial loss, it also negatively impacts an enterprise's efficiency, productivity, credibility and confidence in data. It has far-reaching effects and consequences. Some of these effects are more subtle, such as damage to employee morale, or increased organizational mistrust. If issues related to poor data quality are not properly addressed with appropriate tools and programs, they can easily become top inhibitors to the success of an enterprise's digital initiatives.

The acceleration of digital initiatives and the dynamics of business change are creating demand for greater scale and shorter time to value in relation to data quality improvement. Trusted, high-quality data is vital to an enterprise's digital initiatives. Especially when an enterprise is embracing a distributed infrastructure to decentralize data environments across new and existing platforms, it is difficult to achieve optimal levels of data quality. In addition, growing regulatory requirements from governments and industries, such as the EU's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA), put more pressure on organizations to manage personal data properly in order to reduce operational risks and costs. As data quality is seen as a mandatory aspect of every business, data quality solutions are in greater demand and are often embedded in the critical business operations. Adoption of modern data quality solutions represents an optimal option for tackling data quality efficiently and effectively.

Modern data quality solutions provide several critical capabilities (see Critical Capabilities for Data Quality Solutions) to address typical and emerging data quality use cases, ranging from conventional enterprise operation requirements to emerging AI development support. The solutions are aimed to help organizations streamline the data quality process, build data quality firewalls for critical data capture and gain more insights into their data. Some vendors also offer augmented data quality to further extend their capabilities, such as improving discovery of data characteristics, suggesting next best actions and automating data quality processes by using metadata, application logs, users' actions, common best practices in industries and AI algorithms (see Augmented Data Quality Represents a New Option for Upscaling Data Quality Capabilities). The discovery of certain data patterns or outliers can inspire corrective actions. Automation helps organizations to meet complex data quality requirements quickly and effectively. It leads to higher productivity, greater accuracy and quicker ROI.

While enterprises are quickly adopting data quality solutions to take advantage of out-of-the-box functionalities in support of their business requirements, we also notice a shift in the deployment model. According to the most recent Magic Quadrant survey, 35% of organizations use cloud or hybrid-cloud deployment models, which is an increase of 50% from the previous year. This deployment shift corresponds to the moving-to-cloud initiatives that are common among many enterprises. Data quality solutions are deployed to where the data is. Also, the consequences of cloud-based adoption show in cost reduction, based on the same survey. An organization spends an average of $250,000 annually on data quality solution software for either on-premises or cloud deployment, which is a slight drop from the $261,000 indicated in the previous survey. This cost reduction is the result of more competitive license models for SaaS- or PaaS-based deployment. In addition to cost optimization, flexibility and agility are also improved by the cloud-based adoption of data quality solutions. Some mainstream data quality vendors are investing in cloud-native data quality solutions to provide self-service approaches for cleansing, standardizing, and enriching data. We expect to see a fast-moving trend in the adoption of cloud-based solutions.

Recommendations

- If your organization is not currently using a data quality solution, evaluate current mainstream data quality solutions that address pain points you are encountering, and conduct a proof of concept (POC) with a solution that best fits your scenarios and business requirements.
- Start with a shortlist of specific business problems that are currently tackled manually and are either time-intensive or prone to exceptions.
- Maximize the effectiveness of data quality initiatives by evaluating data quality vendors that are able to read metadata and business rules and policies from data catalogs and remediate quality issues in self-service approaches.
- Exploit vendors' deployment models, pricing and licensing structures. Take advantage of cloud capabilities for initial deployments in the cloud or SaaS options.

Related Research:

Augmented Data Quality Represents a New Option for Upscaling Data Quality Capabilities

5 Steps to Build a Business Case for Continuous Data Quality Assurance

Build a Data Quality Operating Model to Drive Data Quality Assurance

Magic Quadrant for Data Quality Solutions

Critical Capabilities for Data Quality Solutions

Apply These Techniques to Maximize Your Return on Data Quality Technology Investments

A Look Back

In response to your requests, we are taking a look back at some key predictions from previous years. We have intentionally selected predictions from opposite ends of the scale — one where we were wholly or largely on target, as well as one we missed.

On Target: 2019 Prediction — By 2019, 90% of cloud DBMS architectures will support separation of compute and storage, rendering those that do not as irrelevant in the overall market.

The separation of storage and compute has become a basic requirement for any native cloud DBMS. Such separation allows for greater financial flexibility by allowing the separate provisioning of the storage and compute components. These match the needs of target workloads, leading to more efficient use and consumption of cloud resources. Additionally, the use of cloud object storage as the primary storage service has become the common data layer for all cloud data management service offerings.

Today, almost every vendor offering dbPaaS provides for this separation of architectures. In 2019, one of the last holdouts of direct attached storage — Amazon Redshift — introduced new RA3 node types, which explicitly enabled separation of storage and compute as part of a modernization effort. Almost all modern dbPaaS offerings provide this capability.

Missed: 2021 Prediction — By 2021, enterprises using a cohesive strategy incorporating data hubs, lakes and warehouses will support 30% more use cases than competitors.

Data and analytics leaders continue inquiring about data hubs, lakes and warehouses, but many of these conversations haven't shifted toward cohesive implementation plans. Many organizations remain confused, lacking clear ideas and concepts on these topics. At the same time, they are inundated with vendors pushing their own architectures, integration styles and performance characteristics. Often, these vendor-specific offerings address one type of data consumer persona, like data scientists, but fail to account for other personas within the enterprise, such as business analysts, data engineers or product managers. Gartner recommends data and analytics leaders and their teams continue to evolve a holistic plan for their infrastructure, involving a combination of data hub, data lake and data warehouse capabilities (see Data Hubs, Data Lakes and Data Warehouses: How They Are Different and Why They Are Better Together). Through such an approach, they can best position their organizations to support a wide range of use cases and requirements.

Evidence

[1] Survey of reference customers for the 2020 Magic Quadrant for Data Quality Solutions: The survey of vendors' reference customers was conducted in 1Q20 and formed part of a data-gathering effort to help Gartner build on its knowledge of vendors in the market for data quality solutions. At the start of the Magic Quadrant research process, all contacted vendors were asked to identify reference customers that generally represented the inclusion criteria. Vendors provided contact information that was used to invite their reference customers to complete a 35- to 40-minute online survey. A total of 154 reference

customers from 20 vendors completed the survey. Note: vendors' reference customer data is different from primary research and is not a representative knowledge base for the data quality solutions market.

[2] Gartner's Data and Analytics Adoption Survey, 2019: This study was conducted to learn how organizations use data and analytics.

The research was conducted online during November and December 2019 among 272 respondents from North America, Western Europe and APAC regions. Companies from different industries were screened for having annual revenues less than $100M.

Respondents were required to be at manager level or above and to have a primary involvement and responsibility for their organization's data and analytics solutions, including purchase and investments.

The study was developed collaboratively by Gartner analysts and the Primary Research Team. who follow data and analytics management

Disclaimer: the results of this study do not represent global findings or the market as a whole but reflect the sentiment of the respondents and companies surveyed.

Document Revision History

Predicts 2020: Data Management Solutions - 5 December 2019

Predicts 2019: Data Management Solutions - 6 December 2018

Predicts 2018: Data Management Strategies Continue to Shift Toward Distributed - 31 October 2017

Recommended by the Authors

Choosing Between Multimodel DBMS and Multiple Specialized Engines

The Future of the DBMS Market Is Cloud

Understanding and Planning for Database Management Transformation to the Cloud

Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration

How to Activate Metadata to Enable a Composable Data Fabric

Augmented Data Catalogs: Now an Enterprise Must-Have for Data and Analytics Leaders

Design an Effective Information Governance Strategy

Augmented Data Quality Represents a New Option for Upscaling Data Quality Capabilities

5 Steps to Build a Business Case for Continuous Data Quality Assurance