IDC MarketScape: Asia/Pacific (Excluding Japan) Cloud Data Analytics Platform 2020 Vendor Assessment — Gen D Tribes Gather to Build the Intelligent Enterprise
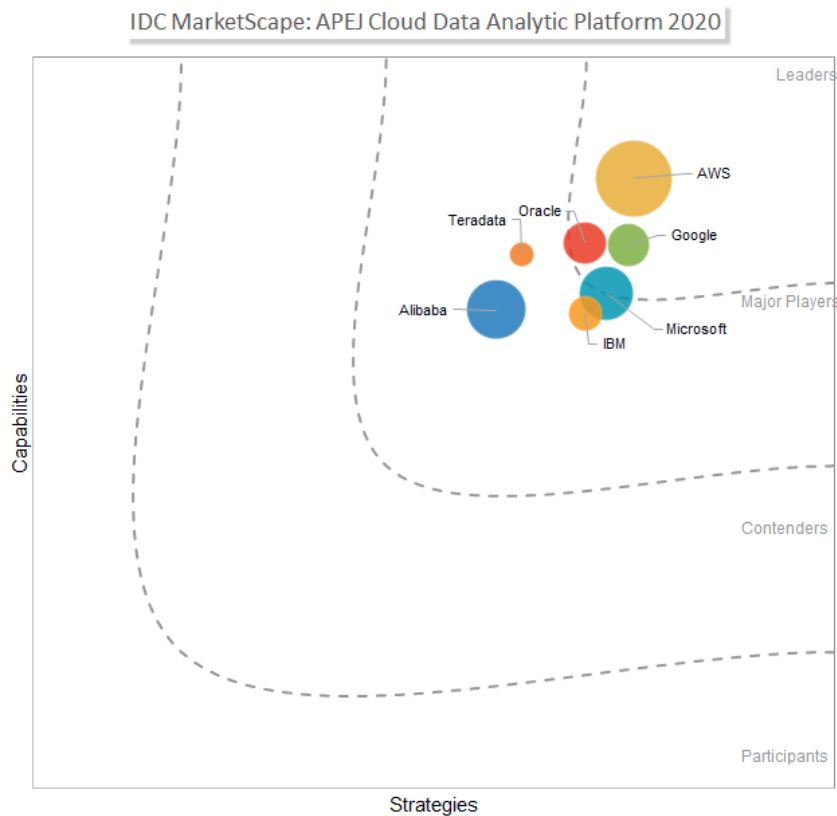
Christopher Lee Marshall   Jessie Danqing Cai   Dylan Yap

THIS IDC MARKETSCAPE EXCERPT FEATURES TERADATA

IDC MARKETSCAPE FIGURE

FIGURE 1

IDC MarketScape: Asia/Pacific (Excluding Japan) Cloud Data Analytics Platform Vendor Assessment

Source: IDC, 2020

Please see the Appendix for detailed methodology, market definition, and scoring criteria.

IN THIS EXCERPT

The content for this excerpt was taken directly from IDC MarketScape: Asia/Pacific (Excluding Japan) Cloud Data Analytics Platform 2020 Vendor Assessment – Gen D Tribes Gather to Build the Intelligent Enterprise (Doc # AP45938520). All or parts of the following sections are included in this excerpt: IDC Opinion, IDC MarketScape Vendor Inclusion Criteria, Essential Guidance, Vendor Summary Profile, Appendix and Learn More. Also included is Figure 1.

IDC OPINION

Within just three years, IDC expects the global economy to reach a tipping point with almost half (46%) of the world's gross domestic product (GDP) driven by digitally transformed products and services (*IDC FutureScape: Worldwide IT Industry 2019 Predictions*, IDC #US44403818, October 2018). Very soon, onboarding digital transformation (DX) will no longer be associated with a competitive advantage because it's already mainstream. Businesses need to think a few steps further ahead of the curve to move their target from becoming a digital-native enterprise to becoming an intelligent enterprise to having all organizational processes respond to changes in their operating environments by acquiring as much data as possible, extracting maximal insight from data, and enabling anyone in the enterprise to act promptly.

Central to an intelligent enterprise is the management and leverage of enterprise data as a portfolio of trusted and operationalized information pipelines. We believe that this inevitably implies the use of platform as a service (PaaS) by the cloud-native architecture precisely because of its unmatched ability to scale in response to changes in enterprise demands. Each information pipeline requires multiple organizational capabilities to manage data, generate insights, empower workforce, and scale decision delivery across the enterprise (*Future of Intelligence Defined*, IDC #US45720619, January 2020). Many of these capabilities remain embodied in people in a variety of enterprise roles situated at different stages of these data pipelines, roles that depend on data to complete tasks and make critical enterprise decisions. IDC has defined these workers as "Generation Data" or "Gen D" for short (*Defining the Data-Native Worker: Gen D*, IDC #US46194920, April 2020).

Gen D workers often have distinct lenses concerning their work, and we can identify some standard architypes. Some are operational managers focusing on engineering or transforming data operationally, whereas others are data analysts who analyze data to create new models and insights. A few data governance executives worry about regulatory-, security-, and risk-related issues. Similarly, a handful of Gen D workers are strategists who look at how these information pipelines produce enterprise value typically in the form of DX initiatives. Gen D workers make deep financial, intellectual, and emotional investments in tools, technologies, and especially platforms. They often behave a lot similar to tribes operating across the enterprise. Tribal investments run deep and will inevitably color many of these workers' views about these tools and potential alternatives. All too often, the selection of a platform merely reflects the power of the Gen D tribe in the ascendency at that moment rather than the true strategic needs of the enterprise.

It is to add some light while removing some of the heat to these Gen D tribal arguments that we write this paper about the development, management, and enhancement of these enterprise information pipelines. This builds on a view of shared capabilities relevant to all Gen D tribes. For example, they must be able to trust the data they use. The platforms they use must be integrated across a variety of critical data sources owned by different functionalities. Their tools must be operationally efficient in

terms of deployment and orchestration. At the same time, they must be simplified for use by new entrants to the tribe. Finally, they must be accelerated in terms of lowered cost for higher performance and a faster time to value. Precisely because these themes are shared across Gen D workers to a greater or lesser extent, they provide a fair basis for comparison and analysis.

Our research evaluates the major vendors according to the IDC MarketScape model, incorporating vendors' market shares in Asia/Pacific (excluding Japan) (APEJ) as reflected in IDC's tracker products and using surveys and interviews conducted between April and August 2020.

All selected vendors provide cloud data analytics platform solutions. As discussed, we chose five general assessment dimensions to explain vendors' past successes and their anticipated future provisions: trusted, integrated, operationalized, accelerated, and simplified.

Our intention here is to help buyer organizations organize around the most critical success factors in building their technology foundations.

The rest of this document fills out the details, but some of the key findings of this report include:

- **Fewer technology differences**. Thanks to the continuous innovation of open source projects, the pervasive adoption of cloud-native architectures, and the constant competitive pressures to increase value-add, the solution disparity among the scoped vendors is inevitably decreasing. All the vendors in scope, for example, have revamped their solution stacks with cloud-native architectures to provide fast and flexible scaling options, embrace open formats to certain extent, and work to continuously improve query performance and user experience.

- **Move to hybrid and multicloud**. Given the relative increase in interoperability between cloud platforms, the desire to avoid undue dependence of a single vendor makes larger enterprises look to hybrid cloud and multicloud solutions as their preferred approach. Ironically, this is organizationally problematic as different tribes sit uneasily with a hybrid approach and will often try to parcel out different functional requirements for different vendor platforms (e.g., data warehouse for one, data science tools for another) as this retains tribal control over particular functional themes, such as data governance, analytics, operations, and DX.

- **Greater dependence on skills and integration services in a hybrid world**. Although all the reviewed solutions are provisioned on cloud with many administrative tasks done autonomously, readying all the necessary data pipelines and connections in a hybrid system landscape requires substantial time and effort by Gen D workers. This is exacerbated by the diverse system landscapes found in APEJ, with many executives on the lookout for both future-proof and past-compatible options. Access to skilled Gen D solution architects and engineers is very limited and highly valued. This calls for closer collaboration among partners, and the center piece in such collaborations is about a trusted working relationship to ensure governance and security, more than anything else.

- **Preference for cloud in running business-critical applications**. Cloud is increasingly viewed as a utility. Many of the buying organizations we reached out to in the course of this study deemed their data and analytics platform to be business-critical, particularly for digital-native and digitally transformed companies. There is little dispute on cloud being the preferred platform to become agile and stay relevant. Cloud data analytics platforms are effectively becoming a utility layer running in between the infrastructure layer and business application layer to facilitate data as the new fuel to continuously power the intelligent enterprise.

- **Gen D's shifting preferences in consuming data**. The work culture and changing skill sets of Gen D knowledge workers are making the consumption of downstream insights more bite-

sized, mobile-friendly, actionable, and artificial intelligence (AI)–augmented or AI-centric. Furthermore, Gen D preferences are also pushing platform change in supporting common themes, such as regulation, privacy, analytics, AI, and operational efficiencies. These shifts not only increase demand for a trusted and operationalized data platform to embed on-demand real-time insights in products and business processes but also encourage organizations to adapt the modern scalable cloud infrastructure to support high-performance real-time consumptions.

With this report being completed in 2020, in the midst of the COVID-19 pandemic, it is important to underline how this has highlighted the benefits of cloud scalability and accelerated organizations' efforts to review decision bureaucracies and distribute their decision making to make the business more resilient. A consensus is emerging that success in developing and managing information pipelines is the basis of enterprise resilience and, ultimately, enterprise intelligence.

IDC MARKETSCAPE VENDOR INCLUSION CRITERIA

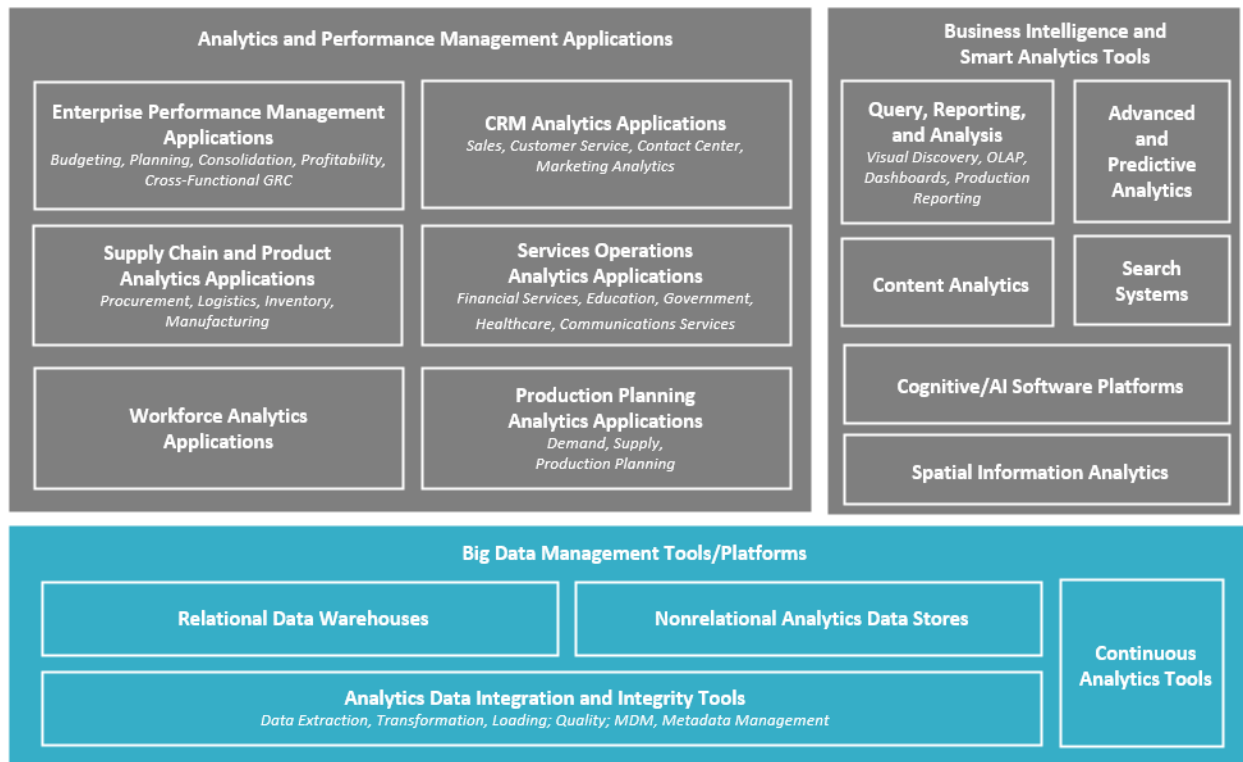Vendor inclusion in the study is based on the following criteria:

- Operates in more than one location in APEJ
- Provides at least one type of analytics data management products: relational data warehouse, nonrelational data store, or a unified product of both
- Offers at least one of the following two competitive markets: continuous analytics tools and/or data integration and integrity tools.
- Provides the solution in scope on public cloud for at least one year in APEJ
- Reports a regional annual revenue of not less than US$10 million for the solution in scope

To the best of our knowledge, vendors not mentioned in this report are yet to meet all the aforementioned inclusion criteria. Although we cannot claim the vendor list is exhaustive, we believe that it represents the most significant players by annual revenue in the cloud analytics space. Furthermore, as economies of scale are so critical in cloud, the vendors listed here are almost certainly going to dominate the landscape for cloud analytics for years to come.

Solutions satisfying these functional criteria are described as "big data management tools and platforms," shortened as "the cloud data platforms" or "the platforms" in the rest of the report, as highlighted in blue in Figure 2 according to *IDC's Worldwide Big Data and Analytics Software Taxonomy, 2019* (IDC #US44517318, December 2018). Note that computing frameworks that support the processing of unstructured data in data lake solutions, Apache Spark, and equivalents are considered a subcategory of nonrelational analytics data stores.

FIGURE 2

Functional Markets in Scope

Analytics and Performance Management Applications

Enterprise Performance Management Applications
Budgeting, Planning, Consolidation, Profitability, Cross-Functional GRC

CRM Analytics Applications
Sales, Customer Service, Contact Center, Marketing Analytics

Supply Chain and Product Analytics Applications
Procurement, Logistics, Inventory, Manufacturing

Services Operations Analytics Applications
Financial Services, Education, Government, Healthcare, Communications Services

Workforce Analytics Applications

Production Planning Analytics Applications
Demand, Supply, Production Planning

Business Intelligence and Smart Analytics Tools

Query, Reporting, and Analysis
Visual Discovery, OLAP, Dashboards, Production Reporting

Advanced and Predictive Analytics

Content Analytics

Search Systems

Cognitive/AI Software Platforms

Spatial Information Analytics

Big Data Management Tools/Platforms

Relational Data Warehouses

Nonrelational Analytics Data Stores

Analytics Data Integration and Integrity Tools
Data Extraction, Transformation, Loading; Quality; MDM, Metadata Management

Continuous Analytics Tools

Source: IDC, 2019

In this review, we focus on the platforms and not explicitly discuss two other related functional markets, "analytics and performance management applications" (shortened as "the applications") and "business intelligence and smart analytics tools" (shortened as "the tools"). The tools and applications are important in their own right as they support the two downstream steps of the data processing cycle — insight generation and embedded consumption, respectively — but they are only discussed as adjacent topics. One way to understand these differences is in terms of targeted members of the Gen D tribe: the platforms are used primarily by data engineers, data stewards, and data architects, the tools are used more by business analysts and data scientists, and the applications are used by business and operational roles.

ADVICE FOR TECHNOLOGY BUYERS

It is almost an impossible task to explore the twists and turns within the big data space in recent years, but a summary of some recent trends is appropriate:

- The widespread use of in-memory technology accelerating queries and embedded business analytics
- The rise of nonrelational big data stores running distributed computing for unstructured schemaless data on commodity hardware

- The use of virtualization tools for queries and analytics to run without requiring a native data store to minimize data movement and increase data portability
- The resurrection of high-performance data warehouses as public cloud brought down the cost of massive parallel processing
- The increased adoption of data catalog and lineage tools to capture data changes and truly manage data as products
- The availability of finer-grained data access control partially driven by compliance requirement to protect sensitive data
- The assured performance of autonomous services and the reduced cost for performance enabled by serverless
- The unified data query across SQL and Spark that has largely redefined data extraction, transformation, and load (ETL)

Altogether, these rapid innovations make it increasingly convenient and affordable to manage all types of data, in motion and at rest, from incoming pipelines to outgoing products. However, inevitably, this onslaught of changes and new options makes solution selection a rather difficult decision and creates arguments among Gen D users adhering to different schools of thoughts. Therefore, we recommend that technology buyers:

- **Think about your information pipelines and expected business outcomes.** Scope and tabulate the must-have data sources, use cases, workload types, and, most importantly, the expected outcomes of these use cases across all five assessment dimensions. It is quite pointless to talk about vendor selection without at least approximating the workloads in scope and expected performance level. When analytics/data-driven decisioning and digital product innovations remain as an afterthought of DX, retrofitting can be too complex to execute and inevitably delay the business impacts.

- **Explicitly acknowledge and manage your Gen D tribes.** Different Gen D tribes exist in organizations, whether executives are aware of them or not. It is important to explicitly identify these tribes in your organization, understand and acknowledge their beliefs and expectation when evaluating different platforms, and make sure the different tribes work in cohesion. Get your tribes involved in platform decisions. If this is not done, decisions viewed as purely strategic will soon come undone when they are implemented organizationally. Equating the move of all data into one technology platform as the entirety of change required is oversimplified and fundamentally the reason of delayed value realization.

- **Catalog and manage data operations as owned by the respective Gen D workers in connection with the business processes, decision owners, and the transactional data sources to cover**. Identify their skills and data/insight consumption preferences to reduce the hurdle of culture change. Business processes establishment or changes can be required to ensure cross-tribe data operations. This also calls for data leaders (chief data officers [CDOs], chief analytic officers [CAOs], CIOs), whichever title they hold, to interact proactively with executives across business functions and communicate requirement for data flows in business terms.

- **Understand vendor capabilities across all the five assessment dimensions compared with the scope using this document as a basis to look for deeper-dive information from the short-listed vendors**. All vendors included in this report provide service options, such as seminars, workshops, and product trials, to guide customers during the onboarding process. Consider also leveraging specialized data skills through their partners' network.

- **Factor in more cost elements and options.** Other than compute and storage resources for ETL queries, there are also the considerable costs of checking for data consistency, quality, compliance, and protection. Some advanced yet necessary features may have an additional price tag to be enabled, and of course, some compute resources can be pushed to the on-premises part of the hybrid infrastructure for better cost for performance.
- **Keep in mind the future enterprise intelligence capabilities to be enabled by the selected platform solution**. The platform is underutilized unless the applications, reports, and insight models running on top of it are mission-critical, in near real time, and machine learning–powered. Monitor the progress of making business and operational decisions data-driven or AI-augmented and proactively work with internal stakeholders, Gen D tribes, and vendors to close any gaps.

VENDOR SUMMARY PROFILE

This section briefly explains IDC's key observations resulting in a vendor's position in the IDC MarketScape. Although every vendor is evaluated against each of the criteria outlined in the Appendix, the description here provides a summary of the vendor's strengths and opportunities, ordered alphabetically.

Teradata

Teradata is positioned in the major players category in this APEJ 2020 IDC MarketScape for cloud data analytics platform.

Teradata provides Vantage as the revamped and unified cloud-native solution for relational data warehouse and graph processing for nonrelational analytics data store, data integration and integrity tools, and machine learning analytics, in general. Vantage, when delivered as a service, is accessible from AWS, Azure, GCP, and Teradata's own cloud infrastructure. In addition, Teradata has Listener as the standalone continuous analytics tools and Viewpoint as the analytics data integration and integrity tools.

At the adjacent Business Intelligence and smart analytics tools layer (the tools), Teradata provides Native Object Store and QueryGrid for native query users. Vantage also includes R and Python integration for data science notebook users and visualizations for business analytics users.

Strengths

Teradata cloud data and analytics platform (the platforms) solution, as detailed earlier, is highly accelerated for complex queries that prioritize performance optimization in a hybrid environment. The company has been in the accelerated data platform market for decades. Since the launch of Vantage in 2018, the company has seen a high renewal rate and expanded the scope of use among its customer base. Customer references in this region include Medibank and Tabcorp Holdings, among others.

- **Accelerated data**. Vantage packages off-the-shelf vertical data models and ETL sets for banking, insurance, retail, telecommunications, utilities, and transportations. Vantage is designed for high-concurrency high-complexity SQL workloads and provides hardware optimization when running as a service, which can be provisioned on public cloud, hybrid cloud, and multicloud environments; Native Object Store and QueryGrid can decompose complex queries and push-down processing to the right element of the analytical infrastructure

for higher parallelism, hence better performance. In addition, Vantage offers pricing flexibility (consumption-based and capacity-based) to help customer scale with cost control;

- **Integrated data**. Native Object Store and QueryGrid connect Vantage with databases, data warehouses, and data lake products of major vendors, such as AWS, Azure, GCP, IBM, Oracle, and Cloudera, to process data where it resides and minimize data movement. Teradata's identity registry solution can be opted in to pair identifiers across different data tables and build an identity graph for customer experience related use cases.

- **Trusted data**. Each Vantage-as-a-service subscription is provisioned as single tenant for every customer for greater control and protection; Vantage provides row-level security by user profile. Column-level encryption can be opted in via third-party software, such as Protegrity and Micro Focus. Vantage supports real-time data quality check and exception captures.

- **Operationalized data**. Viewpoint can monitor and manage workloads in the Vantage ecosystem, on-premises or in the cloud. It offers a wide variety of portlets for different roles, including systems management, monitoring, query submission, system health, and query progress.

- **Simplified data**. Vantage offers an interactive and role-specific user interface for different types of data workers.

- **Leverage of machine learning**. Vantage includes a built-in machine learning engine for data transformation, integration, and preparation. Teradata's Adaptive Optimizer continuously learns and adjusts to changes in customer workloads, data demographics, and changes in host platform, thereby providing fully automatic, hands-off performance. Vantage uses advanced AI/machine learning to "fingerprint" the system configuration and match it against known issues for event detection and patch distribution.

- **Cloud platform**. Vantage runs on AWS, Microsoft Azure, and GCP. Vantage can also be integrated into AWS/Azure/GCP storage and compute services, such as S3/Blob, SageMaker, Glue, and Dataproc. Vantage is built for hybrid cloud/multicloud; Teradata also offers frictionless license portability. There is no charge for customers moving between cloud platforms.

    Channel and community. Teradata has a growing network of over 100 partners, including resellers, systems integrators, and solution partners, such as MicroStrategy, IBM Cognos, Looker, Qlik, Tableau, SAS and TIBCO Spotfire.

Challenges

Teradata actively looks to transfer its deep knowledge in enterprise data warehouse and use case–specific business know-hows to new accounts through its consulting arm and expanded partner network to help organizations reap data-driven business outcomes. The continuous cost-down of public cloud resources will make the highly performant and highly reliable Vantage platform increasingly affordable for companies of all sizes. Teradata also aims to enable more dynamic pricing for customers who require both sporadic and fixed-capacity workloads.

Consider Teradata When

Teradata, and its cloud data analytics platform, is a good match for companies that need hybrid cloud flexibility and prioritize security and performance for their analytics workloads. Gen D workers in larger enterprises concerned with operations and governance often look to Teradata solutions.

Reading an IDC MarketScape Graph

For the purposes of this analysis, IDC divided potential key measures for success into two primary categories: capabilities and strategies.

Positioning on the y-axis reflects the vendor's current capabilities and menu of services, and how well aligned the vendor is to customer needs. The capabilities category focuses on the capabilities of the company and product today, here and now. Under this category, IDC analysts will look at how well a vendor is building/delivering capabilities that enable it to execute its chosen strategy in the market.

Positioning on the x-axis or strategies axis indicates how well the vendor's future strategy aligns with what customers will require in three to five years. The strategies category focuses on high-level decisions and underlying assumptions about offerings, customer segments, and business and go-to-market plans for the next three to five years.

The size of the individual vendor markers in the IDC MarketScape represents the market share of each individual vendor within the specific market segment being assessed. For this IDC MarketScape, vendor size was determined by IDC's 2019 Software Tracker, which represents an estimate of each vendor's 2019 software revenue. For additional information regarding more vendors in this scope and their market shares, see *New Orders and Newcomers in the Market of Big Data Management Tools & Platforms (APEJ, 2019)* (IDC #AP45220619, November 2019).

IDC MarketScape Methodology

IDC MarketScape criteria selection, weightings, and vendor scores represent well-researched IDC judgment about the market and specific vendors. IDC analysts tailor the range of standard characteristics by which vendors are measured through structured discussions, surveys, and interviews with market leaders, participants, and end users. Market weightings are based on user interviews, buyer surveys, and the input of IDC experts in each market. IDC analysts base individual vendor scores, and, ultimately, vendor positions on the IDC MarketScape, on detailed surveys and interviews with the vendors, publicly available information, and end-user experiences in an effort to provide an accurate and consistent assessment of each vendor's characteristics, behavior and capability.

Market Definition

IDC describes the big data and analytics software market as a collection of three functional markets: the applications, the tools, and the platforms for data extraction, integration, governance, movement, curation, analysis, and visualization deployed to support or automate a broad range of strategic, operational, and tactical decision making. The platforms, which is the scope of this paper, consists of the following technologies:

- **Nonrelational analytics data stores**. The nonrelational analytics data stores market segment is primarily derived from the dynamic data management systems. Dynamic data management systems share some characteristics with database management systems (DBMS). They support the common storage and retrieval of data optimized in a managed environment for quick saving and retrieval or query or both. The key difference, and what makes dynamic data management systems dynamic, is that they have no schema but depend on program code to define their contents.

- **Relational data warehouses**. Relational data warehouse software market includes relational DBMS (RDBMS) software used to manage and process data in support of ad hoc queries and report generation.

- **Continuous analytics tools**. Continuous analytics tools are used for real-time and near-real-time decision support and decision automation. Continuous analytics is in-motion technology that continuously receives and transforms data in real time and in micro batches. It is software made up of two primary segments: streaming integration and streaming analytics.

- **Analytics data integration and integrity tools**. This market was formerly called "data warehouse generation tools." This software enables the access, blending, movement, and integrity of data among multiple data sources. The purpose of data integration is to ensure the consistency of information in which there is a logical overlap of the information content of two or more discrete systems. Analytics data integration software is used to capture, prepare, and curate data for analytics. It is also the conduit through which new data types, structures, and content transformation can occur in modern IT environments that are inclusive of relational and nonrelational data repositories.

The platforms can be provisioned on-premises or on-cloud. This paper addresses only those provisioned on-cloud, especially public cloud. An extended read about the market definition can be found in *IDC's Worldwide Big Data and Analytics Spending Guide Taxonomy, 1H19* (IDC #US46089419, March 2020).

Related Research

- *Defining the Data-Native Worker: Gen D* (IDC #US46194920, April 2020)
- *IDC's Worldwide Big Data and Analytics Spending Guide Taxonomy, 1H19* (IDC #US46089419, March 2020)
- *New Orders and Newcomers in the Market of Big Data Management Tools & Platforms (APEJ, 2019)* (IDC #AP45220619, November 2019)
- *IDC's Worldwide Big Data and Analytics Software Taxonomy, 2019* (IDC #US44517318, December 2018)
- *IDC FutureScape: Worldwide IT Industry 2019 Predictions* (IDC #US44403818, October 2018)

Synopsis

This IDC study provides an assessment of cloud data analytics platform solutions and presents the criteria most important for companies to consider when selecting such a platform in preparation for becoming an intelligent enterprise. This assessment discusses both quantitative and qualitative characteristics that explain success in cloud data analytics capabilities in organizational data initiatives. The evaluation is based on a comprehensive and rigorous framework that assesses vendors relative to the criteria and one another. The study highlights the success factors expected to be the most influential for now and in the future for companies considering, exploring, scrutinizing, or reviewing their investment on data platforms with an emphasis on the importance of managing different tribes of Generation Data (Gen D) workers.

"Data strategy is no longer reared by technology alone. Cloud data platforms need to acknowledge the competing expectations of the business as well as the skills and attitudes of enterprise data workers who make it happen," says Dr. Chris Marshall, Associate VP, Data, Analytics, and Artificial Intelligence, IDC Asia/Pacific. "Technology convergence and open source have increasingly leveled the playing field, but attitudes and skills remain slow to adapt to the new technical reality," he adds.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC Asia/Pacific Headquarters (Singapore)

80 Anson Road, #38-00
Singapore 079907
65.6226.0330
Twitter: @IDC
idc-community.com
www.idc.com