

Teradata Aster Analytics



DATA ANALYTICS

Multi-Genre Advanced Analytics at Scale

Teradata® Aster® Analytics is the industry's first multi-genre advanced analytics solution that provides powerful, high-impact insights on all data types of any volume. Multi-genre capabilities refer to the seamless application of different analytics techniques to address any use case within one solution. Aster Analytics does precisely this by allowing users to execute a wide variety of advanced techniques (e.g., Path, Graph, Text and Sentiment, Visualizations, R packages) to ingest, prep, and analyze data in addition to visualizing and operationalizing insights all within a single interface and at speed and scale. Aster Analytics enables organizations to attain unmatched competitive advantage and drive pervasive adoption of big data analytics by every user based on their skills and preferences.

Next-Generation Analytics for New Insights

While most big data solutions require highly specialized and expensive data science skills to design, build and maintain their analytics solutions, Aster Analytics provides interfaces to enable a broad range of users to discover insights from their data at speed and iterate. Technical users such as data scientists can code directly to develop their own polymorphic next generation analytics that can be applied on any data store. Business analysts with SQL skills can execute over 100 prebuilt functions (Aster Analytics Portfolio) in any combination to run sophisticated analytics against their data. Data Scientist, Citizen Data Scientist and Business Analyst can leverage KNIME, a popular open source analytic workbench with the Aster Analytics extension for KNIME

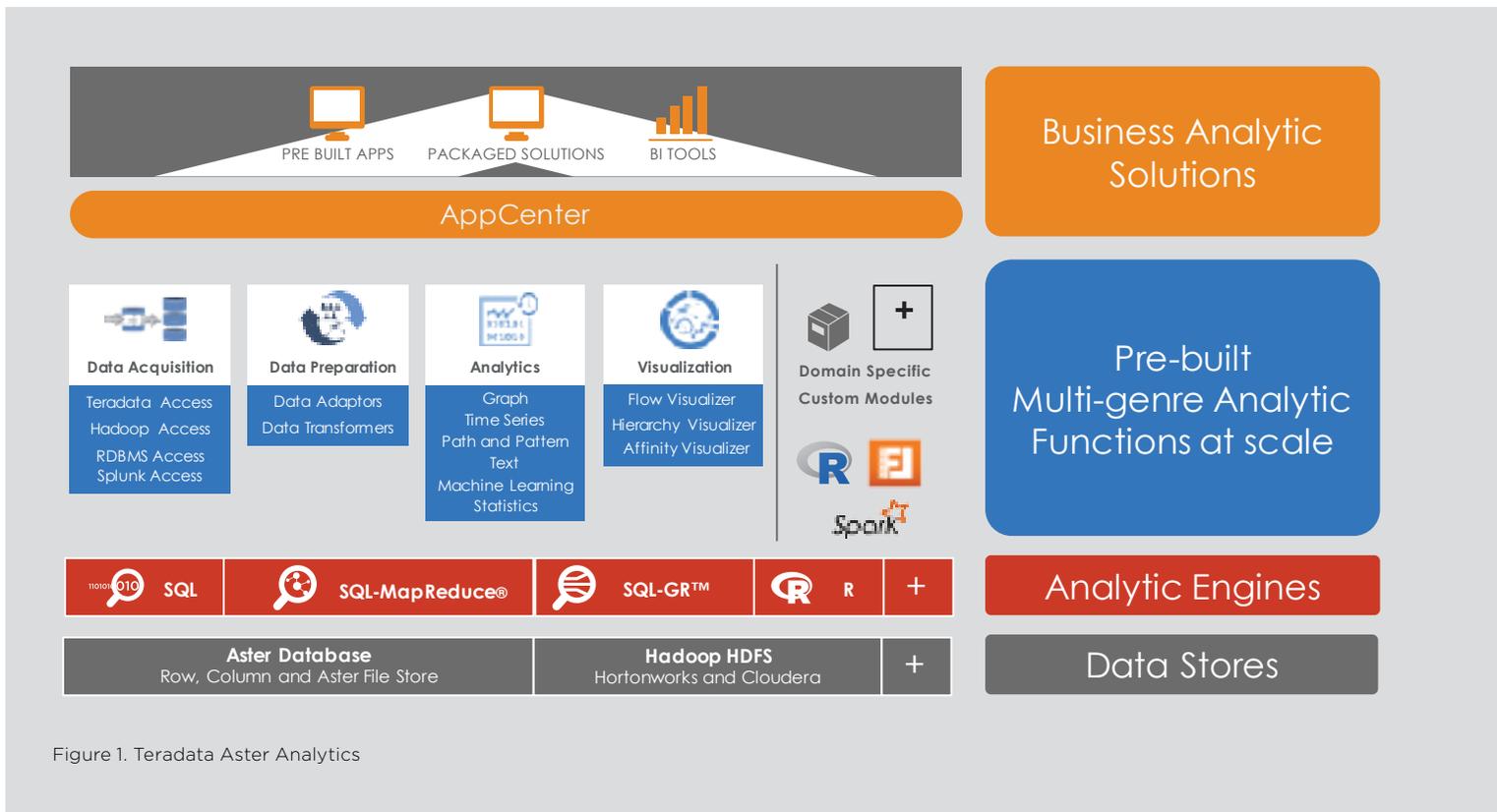


Figure 1. Teradata Aster Analytics

to simplify and accelerate analytic development. Furthermore, analysts who prefer R can build R scripts to implement multi-genre advanced analytics by combining the flexibility of the R language with Aster's advanced analytics capabilities. The results of these analytics can be easily deployed and operationalized via Teradata Aster AppCenter, an interactive application for the business, by a wide breadth of enterprise users. Aster Analytics can effortlessly scale analytic processing on massive volumes of data by leveraging Massive Parallel Processing (MPP) architecture where tasks are run across multiple nodes for parallel processing. The Aster database includes 3 analytic engines (SQL, MapReduce and Graph) designed to provide optimal processing of the analytic tasks across massive volumes of data. In addition, the analytics are performed directly in the database to eliminate data movement and to leverage the MPP, parallel processing. The Aster database can be easily extended to support other analytic engines such as R or Python. Teradata Aster Snap Framework™ for Advanced Analytics

Aster Analytics features the unique Teradata Aster Seamless Network Analytics Processing (SNAP) Framework™ for advanced analytics that is designed to be standards-based, extensible and seamlessly integrated with existing IT infrastructures. Teradata's revolutionary SNAP Framework™ enables users to snap together multiple analytic engines and file stores with ease, providing them with unmatched power and speed to delve deeply into data.

The SNAP Framework™ includes the following:

- Using the unified SQL interface, analysts and data scientists can invoke multi-genre advanced analytics (graph, path/pattern, text, SQL and statistical predictive analysis) in a single SQL statement.
- Integrated Optimizer takes the query with multi-genre advanced analytics and parses it into multiple sub-queries to be executed by workload-specific execution engines like Graph and MapReduce. The Integrated Optimizer features powerful advanced optimizations like global collaborative planning and adaptive optimization that result in order of magnitude gains in performance, such as eliminating redundant operations. Integrated Executor orchestrates the

An example of an application of multi-genre advanced analytics could be illustrated in the context of addressing the issue of customer churn. First, we can apply a path analytics function to identify the key drivers of churn (e.g., customer complaint, formal account review initiation, legal document download). Then we can determine a customer's influence level by applying graph analytics functions. Finally, once we have determined the top 20 percentile of influential customers who are likely to churn, we can determine the nature of their sentiments by applying text and sentiment analytics functions. The ultimate goal of multi-genre advanced analysis is to provide the business with a quick and powerful insight that allows them to focus on the riskiest high value (influence) customers and reduce churn likelihoods.

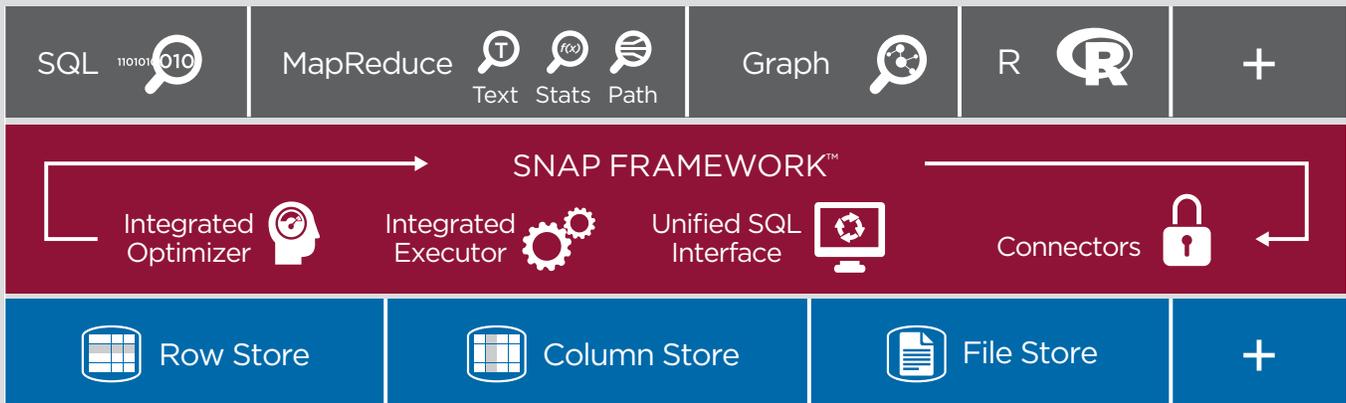


Figure 2. Aster Analytics SNAP Framework

execution of query and iterative phases across the multiple engines. It passes parameters and results across engines and performs common workload management and monitoring across engines.

- Connectors provide unified access to a number of data sources such as Hadoop, Teradata, Splunk and Spark with a plug-and-play system designed for data access and analytic processing.

Aster Analytics Portfolio

Teradata® Aster® Analytics Portfolio provides a suite of ready-to-use SQL, MapReduce® and Graph functions quickly and easily uncover non-intuitive business insights from big data. The Analytics Portfolio is part of the Teradata Aster Analytics solution. The Analytics Portfolio includes pre-built functions that enable data ingestion and preparation, advanced analysis, and visualization. The Analytics Portfolio provides new high-value insights for a wide variety of business use cases such as customer churn, social network, path to purchase, marketing attribution, product affinity, fraud, manufacturing optimization and more. All this is possible on a single solution and through a single query that integrates all the advanced analytics steps. It is meant for all user types and eliminates the need to maintain disparate tools, hire people with niche and expensive skill sets, or maintain silos of metadata.

Aster Analytics Portfolio—Key Capabilities

Data Acquisition

This module includes functions that enable access to multi-structured data stored in Apache™ Hadoop,™ Teradata Data Warehouse, Social Media sources, Call Center Records, other third party applications, and other RDBMS systems. Furthermore, as the number of data sources increase in disparity and volume, Teradata Querygrid and Aster’s connectors can link Aster analytics to just about any data source from social media, non-Teradata RDBMS’ to IoT data. This ensures that all available data are used to deliver crucial insights.

Aster Analytics also provides a number of connectors that allow users to push and pull data from Spark, Splunk and any data source. Security teams can apply Aster parsers and analytic functions on Splunk log data to detect insider threats.

Data Preparation

Ready-to-use data adaptors and transformers that enable interpretation and preparation of data such as Weblogs, XML, emails, and machine logs for analyses. A sample set of functions included provide capabilities such as: *Outlier filters* to remove outliers from data, *Apache Log Parser* to support custom log formats from Apache Web servers as defined by the user, *XML Parser* and *JSON Parser* to parse and prepare XML logs generated

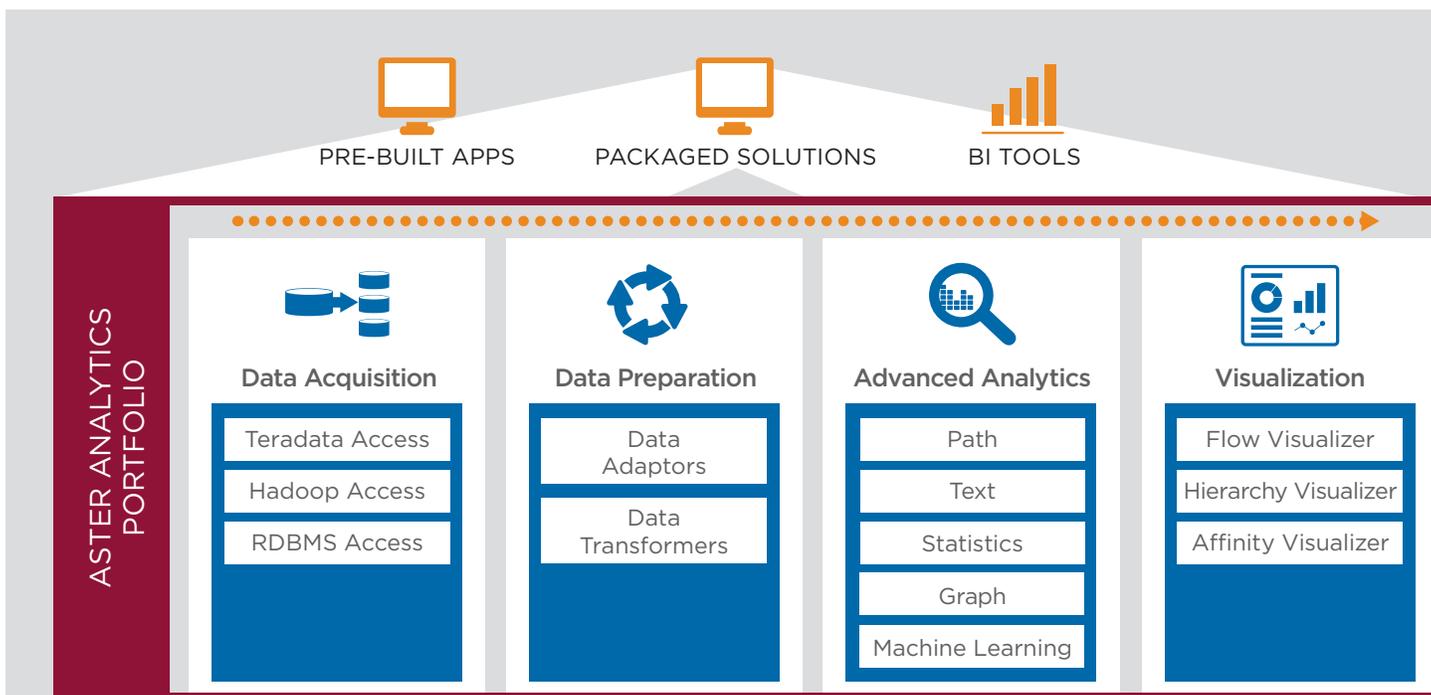


Figure 3. Teradata Aster Analytics Portfolio

by applications like Web logs and POS logs at retail stores and outlets, and *Sampling* to execute various sampling techniques for robust statistical analyses. Data transformation functions include *sessionization* and *unpack* to transform complex, unstructured data into meaningful formats suited for analytics.

Advanced Analytics

Vast array of ready-to-use functions for time series, statistical, text, graph, and SQL analysis in addition to custom SQL, MapReduce and Graph analytic functions. A sample set of functions include:

- **PageRank** – assigns an importance or influence value to nodes in a network based on which the value, importance, or influence of that node can be determined
- **nPath** – to identify the most common consumer paths to defined events such as buying a product, abandoning a shopping basket, or closing their account
- **Shapley Value** – a cooperative game theory construct where a specific distribution of values are generated across a set of participants who operate in a coalition
- **SAX** – machine data analysis, such as analysis of sensor data in manufacturing. In addition to identifying anomalies in manufacturing processes
- **Shapelets** – time series subsequences which are representative of an underlying data class
- **Confusion Matrix** – machine learning for quantifying the performance of an algorithm to improve predictive models
- **Single Decision Tree** – (build and apply a single decision tree for classification and Identify important variables that are critical to decision-making
- **Distribution Matching** – hypothesis testing of data origins from a specific distribution and parametric estimation
- **Collaborative Filter** – used in recommendation engines to create classes of entities for further action, eg., purchase, based on previous actions
- **Spark MLlib** – used by data scientist to access a growing library of open source analytic functions. For example, the Spark connector allows analyst to quickly access Spark deep learning algorithms without having to know Scala

Visualization

Visual MapReduce functions that are massively parallel, in-database, in-process and out-of-the-box, providing novel visualizations to make it faster and easier to discover

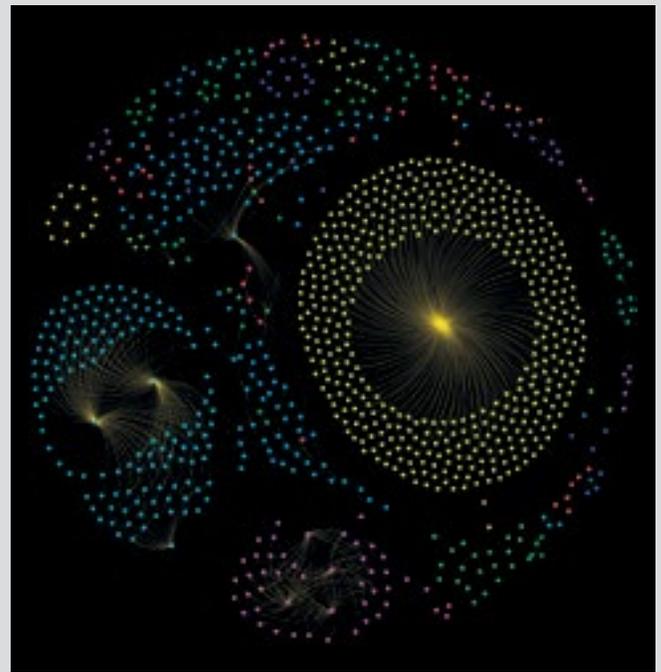


Figure 4. Aster Analytic Sigma visual illustrates the relationship between Aster Analytics Visualization via AppCenter.

new insights from big data. They complement existing business intelligence and visualization tools by providing purpose-built visualization capabilities best suited to represent the in-depth insights offered by Teradata's patented MapReduce framework. A sample set of capabilities include:

- **Flow Visualizer** – understand the path taken leading to an outcome such as purchase or downloads
- **Affinity Visualizer** – understand how two sets of seemingly different products or services actually have a close connection and hence can be bundled
- **Hierarchy Visualizer** – organize all discrete entities and interactions into hierarchies to better understand and comprehend the relationships and behaviors. Teradata also provides the flexibility to integrate your favorite BI and visualization tools (e.g., Tableau, Microstrategy, Tibco Spotfire, IBM, Cognos) to easily operationalize your analytics. In addition, Teradata provides tools such Guided UI and Tango to simplify access and accelerate the analytic process.

The graph analytics visualization (see Figure 4) highlights the solution's ability to be rapidly iterative while quickly zeroing in on the most relevant signals while doing advanced analyses.

R Interface for Aster Analytics

R, the open source language for statistical analysis and data mining, has grown in popularity where over 70 percent of data miners and scientist report using R. And with thousands of packages contributed to open source, it provides multi-genre advanced analytics in a single language. But as powerful, flexible and extensible R is, it carries distinct memory, data and processing limitations. That is why Teradata provides an R interface to Aster Analytics, a solution that provides scalable R for business analytics. Integrating open source R engine delivers a single powerful analytic environment for the analytics community—including R programmers, business analyst and data scientist. The Aster R packages provides an R interface to Aster Analytics Portfolio functions that can be combined with open source R packages and Aster R Runners to successfully address critical business problems.

Flexible Deployment Options

Aster Analytics is available in a few options: as an Appliance, on the Cloud, software only, and on Hadoop.

- The Aster Analytics Appliance includes Dell Servers and can be configured to include Hadoop nodes as well.
- The software only solution is available to customers who want to install Aster Analytics on their own hardware.
- Aster Analytics as a Service is hosted in the Teradata Cloud, offering customers fast and flexible provisioning option and scalability as needed. Cloud services are available in monthly, annually or three year pricing agreement.

- Aster Analytics on AWS is offered on Amazon Marketplace and is hosted on Amazon Web Services (AWS) with all the advantages of Aster Analytics and AWS—flexible, robust, self-service analytics. Businesses can easily and quickly provision a multi-node instance of Aster on AWS to support multiple use cases from analytic sandbox or production analytic environment.
- Aster Analytics on Hadoop is the latest offering in which Aster Analytics will run as a native, first class citizen in Hadoop leveraging hdfs as the file system. Aster analytics can run directly on data loaded into Aster and also data loaded on hdfs via other formats such as Hive.

As much as these options are available to ensure maximum compatibility with the preferred infrastructure of the customer, all of the Aster analytics capabilities are consistent across these options. Furthermore, analytic activity conducted on one option can easily be ported on other options.

For More Information

To find out how the Teradata Aster Analytics solution can help you take advantage of big data volumes in a fast, efficient and cost effective way while you improve your decision-making capabilities and grow a stronger, more productive business please contact your local Teradata representative or visit Teradata.com.

Join the Aster Community at Aster-Community.Teradata.com to learn about Aster Analytics.

10000 Innovation Drive, Dayton, OH 45342 Teradata.com

SNAP Framework and SQL-GR are trademarks and Teradata, the Teradata Logo and MapReduce are registered trademarks of Teradata Corporation and/or its affiliates in the U.S. and world-wide. Apache and Hadoop are trademarks of the Apache Software Foundation. Teradata continually improves products as new technologies and components become available. Teradata, therefore, reserves the right to change specifications without prior notice. All features, functions, and operations described herein may not be marketed in all parts of the world.

Copyright © 2017 by Teradata Corporation All Rights Reserved. Produced in U.S.A.

01.17 EB7573



TERADATA.