

TERADATA ASTER DATABASE

TERADATA

DELIVER POWERFUL BIG DATA ANALYTICS

The explosion of multi-structured data types necessitates the application of new and powerful discovery techniques to unlock valuable insights. Leveraging these techniques in a cost-effective sense requires a solution that does not force new investments in hard-to-find programming skills and system administration capabilities. The Teradata Aster Database contains all the important features that accelerate time to insights with minimal resource outlays.

TERADATA ASTER DATABASE

The massively parallel Teradata Aster Database with its unique Applications-Within™ architecture runs analytic application logic inside the system and invokes its patented SQL-MapReduce®, SQL-Graph, and SQL engines to deliver powerful insights speedily at scale. It is continuously available, easy to manage, and linearly scalable to meet enterprise needs.

BREAKTHROUGH PERFORMANCE AND SCALABILITY

The Teradata Aster Database enables insights delivery on massive data scales and combines the analytics power of MapReduce with the familiarity of SQL. Its powerful suite of analytic functions delivers industry leading visualizations for SQL, MapReduce and Graph applications. The result: **10x-1000x better performance than traditional systems and scalability to hundreds of terabytes and petabytes of data.**

“Always-Parallel” Pervasive Parallelism

The Massively Parallel Processing (MPP) executes loads, queries, exports, backups, recovery, installs, and upgrades in parallel, thereby optimizing performance for speedy insights discovery.

Aster SQL-GR™ Engine

This native Graph processing engine enables powerful Graph analysis across complex data sets. Organizations can easily solve use cases such as social network/influencer analysis, fraud detection, supply chain management, security network analysis, threat detection, money laundering, and more. SQL-GR™ is based on the

OVERVIEW

The Teradata Aster Database is an integral component of the Teradata Aster Discovery Platform that embeds MapReduce analytic processing with data stores for big data analytics on multi-structured data sources and types. It delivers new analytic capabilities for a wide group of users with varying technical skills at breakthrough performance and scalability.

Highlights

- Massively Parallel Processing (MPP) architecture for performance and scalability.
- Patented SQL-MapReduce framework embeds and parallelizes analytic applications to deliver ultra-fast, deep analytics on massive data volumes.
- Innovative and industry first SNAP Framework™ that enables easy inclusion of new discovery engines.
- New SQL-Graph framework that facilitates insights discovery on new use cases that require the application of Graph analytic techniques: Network influencer analysis, disease progression, etc.
- Ability to intervene quickly when processing issues are seen on the intuitive Aster Management Console.
- Scalability to thousands of server cores and petabytes of active data.
- Dynamic mixed workload management ensures scalable performance even with large numbers of concurrent users and workloads.
- Fault-tolerant design and online management to minimize downtime.
- Hybrid row and column stores with unified computation and management layers.
- Next-generation loading architecture for leadership throughput to minimize time needed for loading data.
- Deployed on optimized Teradata hardware or Teradata certified commodity server hardware.

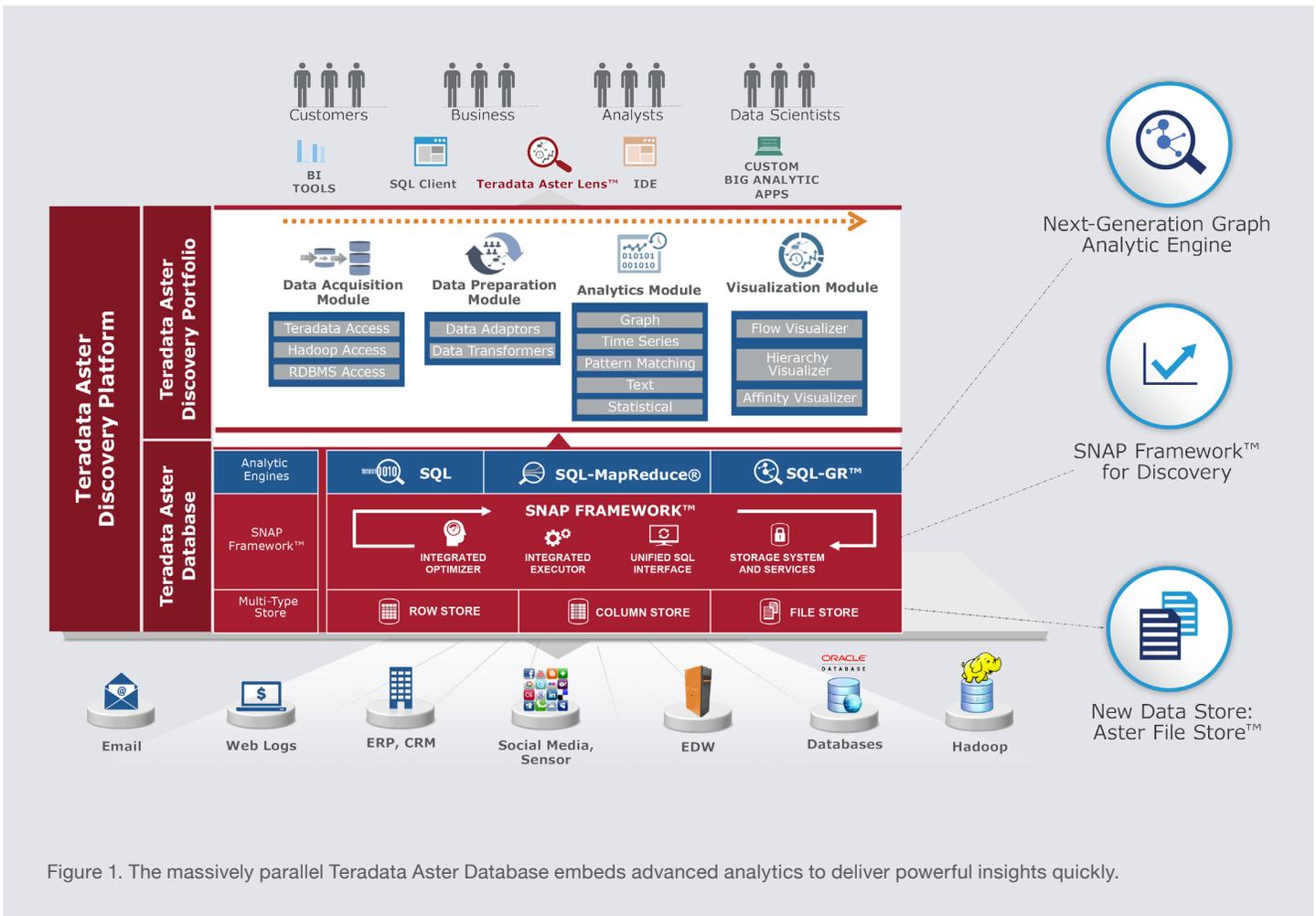


Figure 1. The massively parallel Teradata Aster Database embeds advanced analytics to deliver powerful insights quickly.

Bulk Synchronous Processing (BSP) model, not bound by memory limits, and Teradata Aster’s massively scalable parallel processing architecture distributes the Graph processing across multiple servers/nodes. The Graph engine is designed with very elegant spill-over structures wherein the engine will store the working data set in memory and will page-in and page-out depending Graph engine needs.

Seamless Network Analytic Processing (SNAP) Framework™

The SNAP Framework™ includes a new Optimizer and Executor, a unified SQL interface and storage system and support services. Users can snap together multiple analytic engines (e.g., SQL-MR, SQL-GR™, and SQL) and file stores based on the tailored needs of the organization.

The architecture is designed to be non-disruptive and additional workload-specific processing engines and data stores can be added to the discovery platform as they become available without sacrificing performance. As more analytics engines are added in the future, the SNAP Framework™ ensures that all the current workload management, query optimization, execution, and resource planning functions are automatically applied to the new engine. The SNAP Framework™ is designed to be standards-based, extensible and integrates with existing IT infrastructures. Teradata is the first organization to bring this to market.

Query Optimization

The new Integrated Optimizer is adaptive for fast iterative discovery across SQL, SQL-GR™ and SQL-MR

engines. This enables multiple analytic techniques to be simultaneously applied while still delivering high performance and scalability. The global collaborative planning feature adapts uniquely to the different processes, and computing resources are utilized most efficiently. The Optimizer deconstructs a multi-genre analytics query into sub-queries and executes each within the relevant engine. The Optimizer supports late-bind interfaces, and schema definition on-the-fly to support rapid processing execution. For example, if a query invokes a SQL-GR™ function and orders the results by some field (e.g., name) and if that SQL-GR™ function already has an order by clause with the “name” field, the Optimizer would recognize that the “order by” command is repeated and hence ignores it.

Query Execution

The new Integrated Executor drives the distributed parallel execution among the SQL, Map-Reduce and Graph execution phases, as queries are executed. SQL is a data flow process. MapReduce is the process to first map and then reduce as a batch function. Graph executes in an iterative way. The Executor manages the various execution phases and the translations between them as a single expression. Teradata is the first in the market to deliver this new automated process that minimizes the need for data scientists to understand different languages or attempt data optimization by hand.

Aster File Store™ (AFS)

The Teradata Aster Database comes with new storage architecture, Aster File Store™ (AFS). AFS is based on a common storage system and services that leverage the discovery platform’s parallel processing to quickly and easily ingest multi-structured data, such as Web logs,

sensor data, machine log data etc., without any loss of fidelity or the need for upfront schema definition effort. AFS is purpose-built to capture and preprocess data in its native form. For example, images and large documents can be preprocessed more efficiently in AFS. AFS is 100% compatible with HDFS. Applications that are built for HDFS can easily run on data stored in Aster File Store™ without the need for data recompilation. Analysts can easily change the IP address from HDFS to AFS, making it easy to use AFS as a loading foundation and use the parallel processing of all the Aster workers to get the data into the system dynamically.

“Always-On” Fault Tolerance

Massive scale fault tolerance through replication, automatic failover, NIC bonding, failure heuristics, and clustered backup prevent unplanned downtime due to hardware or software failures.

Hybrid Row and Column Stores

Support for storing data in a row or column format for the highest levels of performance across diverse query workloads including ad hoc and interactive, investigative queries.

Dynamic Mixed Workload Management

Advanced workload management allows for granular rule and policy-based prioritization allowing administrators to adapt to changing priorities in real time.

SQL-MapReduce (SQL-MR)

Aster’s patented SQL-MR framework makes it easy to leverage the power of MapReduce within the familiarity of SQL. SQL-MR allows developers to write powerful and

TECHNICAL SPECIFICATIONS

- **Server Hardware** – certified x86-based commodity server hardware from Teradata partners, or fully-supported Teradata hardware for appliance delivery.
- **Operating Systems** – certified Linux-based operating system.
- **Drivers and APIs** – SQL, OLE DB, ADO.NET, ODBC, JDBC, Psycopg (Python).
- **SQL Standards** – ANSI SQL-92 compliant with SQL-99 and SQL-03 extensions.
- **Business Intelligence and Data Integration** – compatible with leading BI and data integration tools.
- **Ecosystem Adapters** – Aster-Teradata and Aster-Hadoop™ Adapters for rapid and parallel data transfer between Aster and Teradata Databases or Hadoop™.
- **Security** – compatible with Quest Authentication Services.

highly expressive functions in languages including Java, C, C++, C#, Python, and R and push them into the analytic platform where they can be called using standard SQL.

SEAMLESS MANAGEMENT AND CONNECTIVITY TO IT ECOSYSTEMS

The Teradata Aster Database provides intuitive tools for centralized administration and streamlines the management of data and applications even as the system scales to 10s and 100s of servers.

Powerful Applications Management Console

The Aster Management Console makes it easy to configure, manage and monitor data, applications, and infrastructure. An intuitive graphical interface enables easy monitoring with summary dashboards, graphical views of query and process execution. It also provides single-click scaling and point-and-click access to workload management policies.

“Always-On” Online Maintenance

The Teradata Aster Database enables simultaneous load and export during queries, online backup and recovery, online restoration, and online scaling to avoid scheduled downtime.

Extensibility Framework

Aster Management Console (AMC) allows administrators to extend the capabilities accessible through the AMC with custom scripts and programs for common tasks.

Enterprise Adapters

Aster-Teradata and Aster-Hadoop™ Adapters go beyond the capabilities of standard connectors to provide rapid and parallel data transfer between Aster and Teradata Databases or Hadoop™.

ABOUT TERADATA ASTER

The Teradata Aster MapReduce Platform is the market-leading big data analytics solution. This analytic platform embeds MapReduce analytic processing for deeper insights on new data sources and multistructured data types to deliver analytic capabilities with breakthrough performance and scalability. Teradata Aster’s solution utilizes Aster’s patented SQL-MapReduce, and SQL-Graph to parallelize the processing of data and applications and deliver rich analytic insights at scale.

ABOUT TERADATA

Teradata is a global leader in analytic data platforms, marketing and analytic applications, and consulting services. Teradata helps organizations collect, integrate, and analyze all of their data so they can know more about their customers and business and do more of what’s really important. Visit Teradata.com for details.

