

Aster Data® Analytic Platform

Deliver Powerful Big Data Analytics

The massive growth in data from new data types and sources is creating exploding volumes of multi-structured big data that are not typically stored and processed in traditional systems. At the same time, the need has emerged for new analytics that analyze this data quickly and efficiently, enabling iterative exploration that incorporates these massive amounts of new data to find the high value data and deep insights locked inside it. Taking full advantage of the value of this data requires a solution that simplifies and accelerates the development of rich analytic applications without requiring new skills and administrative resources.

The Aster Data Solution

The Aster Data analytic platform, Aster Data nCluster, delivers a massively parallel software solution that embeds MapReduce analytic processing with data stores for big data analytics that incorporate new multi-structured data sources and types to deliver new analytic capabilities with breakthrough performance and scalability. Its unique Applications-Within® architecture runs analytic application logic inside the system, leveraging its massively parallel processing architecture and patented SQL-MapReduce® to fully parallelize processing for deep and ultra-fast analysis of massive data sets. The Aster Data analytic platform delivers easy management of data and analytic applications, continuous availability, linear scalability and the cost advantages of commodity hardware. Combined with Aster Data's visual development environment—Aster Data Developer Express — and suite of optimized analytic modules—Aster Data Analytic Foundation —the Aster Data analytic platform makes deep and rich analysis of big data that was previously difficult or impossible not only possible but also fast and easy.

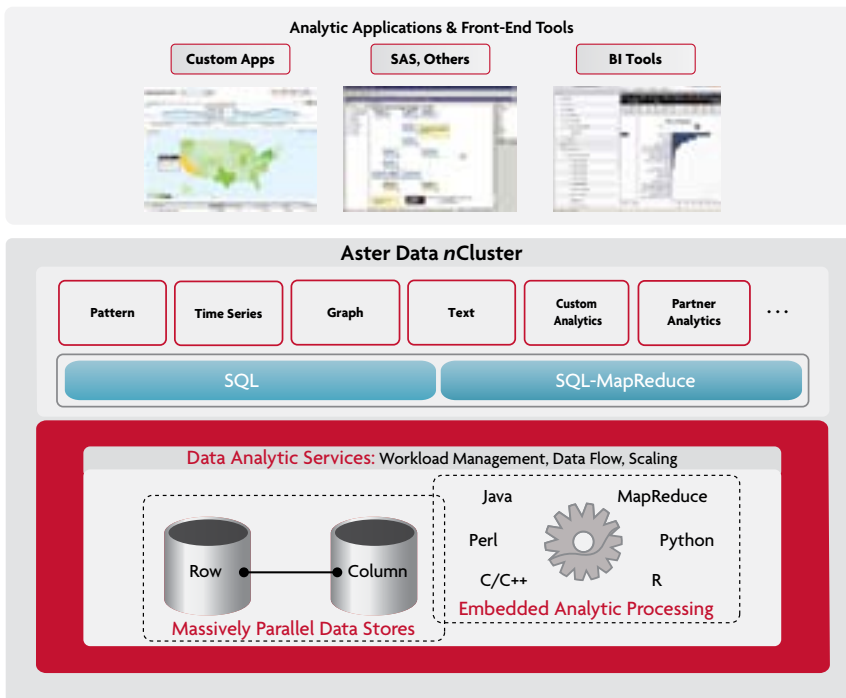


Figure 1: The massively parallel Aster Data analytic platform enables analytic applications to be embedded with data for ultra-fast analysis.

Breakthrough Performance and Scalability

The Aster Data analytic platform delivers leading performance and scalability for big data analytics. The result: 10x-1000x better performance than traditional systems and scalability to terabytes and petabytes of data.

- **“Always-Parallel” Pervasive Parallelism** – Aster Data nCluster’s massively parallel processing (MPP) architecture and embedded MapReduce analytics engine enable end-to-end parallelism of data and analytic processing. It executes loads, queries, exports, backups, recoveries, installs and upgrades in parallel to take full advantage of all resources, optimizing performance for analytic processing.
- **Embedded MapReduce** – Aster Data nCluster deeply embeds an implementation of the MapReduce processing framework, bringing the analytic power of MapReduce to developers of advanced analytics so that they can easily deliver ultra-fast results on large data sets.

Overview

The Aster Data Analytic Platform, Aster Data nCluster, is a massively parallel software solution that embeds MapReduce analytic processing with data stores for big data analytics that incorporate new multi-structured data sources and types. It delivers new analytic capabilities with 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
- Suite of powerful building blocks for simplifying and accelerating development of analytic applications—Aster Data Analytic Foundation
- First visual integrated development environment for SQL-MapReduce analytic applications—Aster Data Developer Express
- Rich monitoring and easy management of data and analytic processing with the intuitive Aster Data 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
- Runs on certified commodity server hardware from Aster Data partners

- **Unlimited Scalability** – Unique Online Precision Scaling™ provides linear scalability across loads, queries, and backups independently or in unison. Granular splitting and load balancing of virtual partitions ensures maximum parallelism for massive “no limits” scalability.
- **“Always-On” Fault Tolerance** – Massive-scale fault tolerance through replication, automatic failover, NIC bonding, failure heuristics, and clustered backup to 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 queries.
- **Dynamic Mixed Workload Management** – Advanced workload management that runs on commodity hardware. Granular rule- and policy-based prioritization and dynamic resource allocation balance processing and compute resources for in-progress transactions, allowing administrators to adapt to changing priorities in real time.

Next-Generation Analytic Applications

The Aster Data analytic platform makes it easy to create advanced analytic applications that deliver ultra-fast analysis that scales to terabytes and petabytes of data. It allows analytic applications to be embedded within the analytic platform with data for maximum performance, combines the analytic power of MapReduce with the familiarity of SQL, delivers the first visual development environment for SQL and MapReduce applications, and provides a suite of powerful analytic functions for accelerating development.

- **SQL-MapReduce** – Aster Data’s patented SQL-MapReduce framework makes it easy to leverage the power of MapReduce within the familiarity of SQL. SQL-MapReduce allows developers to write powerful and 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. SQL-MapReduce functions are simple to write and are seamlessly integrated with SQL statements.
- **Embedded Applications** – Aster Data nCluster’s Applications-Within™ approach makes it simple to embed analytic applications within the analytic platform next to data to eliminate the overhead of moving large data sets to analytic applications. Enterprises can easily push their existing applications into Aster Data nCluster to see significant performance benefits.
- **Aster Data Developer Express** – Developer Express makes it easy to develop, validate, and deploy advanced analytic applications with the first visual development environment for SQL-MapReduce applications. It enables visual development through integration with the Eclipse integrated development environment, provides a desktop testing environment for easy testing, and enables one-click push down of applications into nCluster.
- **Aster Data Analytic Foundation** – This suite of powerful, reusable analytic functions accelerates the development of rich analytics with ready-to-use modules optimized to take advantage of the power of Aster Data’s SQL-MapReduce framework for fast analytics.

Efficient Management of Data and Applications

The Aster Data analytic platform delivers rich visibility and control of not only data but also the analytic applications running inside the system. It provides intuitive tools for centralized management combined with powerful capabilities for simplifying and automating administration, streamlining and automating management of data and applications to minimize administrative work even as the system scales to 10s and 100s of servers.

- **Powerful Console for Monitoring and Managing Data and Applications** – The Aster Data Management Console (AMC) 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, and easy drill-down. It also makes administration easy with single-click scaling and point-and-click access to workload management policies.
- **“Always-On” Online Maintenance** – Aster Data nCluster enables simultaneous load and export during queries, online backup and recovery, online restoration, and online scaling to avoid scheduled downtime.
- **Extensibility Framework** – The Aster Data Management Console allows administrators to extend the capabilities accessible through the AMC with custom scripts and programs for common tasks.

About Aster Data

Teradata’s Aster Data analytic platform is the market-leading big data analytics solution. The Aster Data analytic platform embeds MapReduce analytic processing for deeper insights on new data sources and multi-structured data types to deliver analytic capabilities with breakthrough performance and scalability. Aster Data’s solution utilizes Aster Data’s patented SQL-MapReduce to parallelize the processing of data and applications and deliver rich analytic insights at scale. For more information visit www.asterdata.com.

Technical Specifications

- **Server Hardware** – certified x86-based commodity server hardware
- **Operating Systems** – certified Linux-based operating systems
- **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 business intelligence and data integration tools
- **Hadoop Connector** for rapid and parallel data transfer between nCluster and Hadoop
- **Security** – compatible with Quest Authentication Services

“We’re very excited to take advantage of the new Aster Data Management Console which provides deep visibility into the performance of the system and provides a deeper level of situational awareness and historical perspective for our entire operations and engineering staff.”

*Don Watters, Chief Data Architect
MySpace*

