

Aster MapReduce Data Warehouse Appliance

The Challenge

The amount of data you need to store, manage, and analyze is growing relentlessly. You are expected to manage this data, adapt to business growth, and provide business users with the analytical tools they need to turn this big data into actionable intelligence—and do all this on a modest budget.

Traditional data warehouse appliances struggle to cost-effectively keep pace with this explosion in data, as well as the analytic depth and performance of data-driven organizations. High-priced proprietary server hardware and interconnects have put the power of MPP out of reach of many organizations and make upgrades and scaling painful for the rest—until now.



The Aster Appliance enables powerful MPP performance on low-cost Dell hardware

The Solution

The Aster MapReduce Data Warehouse Appliance is an easy-to-deploy, massively parallel processing (MPP) database appliance capable of handling the most demanding data warehousing environments. It empowers businesses to make better data-driven decisions on large amounts of data faster and cheaper than any other solution in the market.

With the Aster appliance series you get:

- A fundamentally lower-cost and more flexible architecture
- Lower operating costs
- Analytic power

For businesses with modest data warehousing needs, the Aster MapReduce Data Warehouse Appliance - Express Edition provides the ideal starter solution. At a starting price of \$50,000 it offers unmatched price/performance for up to 1 TB of user data. For businesses with more demanding analytical requirements, the Aster MapReduce Data Warehouse Appliance Enterprise Edition offers a cost-effective, high-performance solution with easy expansion up to 1 PB of data.

Key Benefits

- **Fundamentally lower-cost architecture**
 - ✓ True commodity-grade servers and interconnects
 - ✓ \$50K price for Express Edition: up to 1TB user data
 - ✓ Enterprise Edition: up to 1PB on commodity HW
- **Lower operating costs**
 - ✓ Intuitive admin console: visibility and easy control
 - ✓ Live Administration while system is online
 - ✓ High availability: online fault-tolerance and recovery
- **Analytic power**
 - ✓ Compute-rich hardware with high ratio of CPU and RAM to disk
 - ✓ Analytic expressiveness through both SQL and open-language function support

Fundamentally Lower-Cost Architecture

Aster's novel parallel architecture transforms true commodity-grade servers and interconnects into a high-performance and fault-tolerant MPP data warehouse appliance. In contrast, traditional data warehousing appliances – built on proprietary hardware and interconnects – are expensive to acquire and even costlier to scale. With these appliances, incremental upgrades are rarely possible, resulting in expensive rip-and-replace upgrades and system downtime. With the Aster Appliance you can “grow in place” with your data warehousing needs, resulting in a more economical and flexible long-term solution.

Lower Operating Costs

The Aster Appliance provides a turn-key package from hardware to business intelligence (BI) bundled together with the most open language support in the market that lowers the barrier to adoption by opening up access to a larger set of users with varying skill-sets and allowing for effective reuse of existing IT assets. The Aster Appliance reduces the cost of administration by providing a unifying cluster fabric to ensure high availability, adaptation, and resiliency with minimal effort via pervasive self management.

Analytic Power

The Aster Appliance allows you and your applications to get deep insights from data quickly—answers you can't get quickly with traditional approaches that are overwhelmed by rich queries. Behind the appliance's analytic power is a compute-rich MPP architecture and patent-pending algorithms that efficiently perform rich large-scale queries, reporting, and analysis. The Aster Appliance also comes packaged with a novel In-Database MapReduce framework that combines the power of the MapReduce programming paradigm with the strengths of SQL relational databases. Additionally, Aster's MapReduce framework opens up the data warehouse to a whole new class of developers who can write powerful analytic applications in common development languages such as Java, .NET, Python, Perl, and more, and have them run in a massively parallel fashion – accelerating performance. Fault isolation and resource management features ensure robust execution of analytic applications with firm administrative control over priority policies.

Aster MapReduce DW Appliance Series - Overview

Specifications	Express Edition	Enterprise Edition						
		MRDW36	MRDW312	MRDW325	MRDW350	MRDW375	MRDW3100	MRDW31000
Usable space	Limited to 1TB	6.25TB	12.5TB	25TB	50TB	75TB	100TB	1PB*
# Worker nodes	3	8	16	32	63	94	125	330
# Loader nodes	0	0	1	1	2	2	2	2
# Queen nodes	1	2	2	2	2	2	2	2
MicroStrategy node (optional)	1	1	1	1	1	1	1	1
Aggregate Worker CPU cores	24	64	128	256	504	752	1000	2640
Aggregate Worker RAM	72GB	192GB	384GB	768GB	1.5TB	2.2TB	2.9TB	7.9TB
Compression support	No	Optional	Optional	Optional	Optional	Optional	Optional	Optional

* User data; assuming 3.75x compression

Technical Specifications

Software	
Operating system	Linux-based
Supported APIs	SQL, JDBC, ODBC
SQL support	ANSI SQL-92 with SQL-99 extensions
BI interoperability	MicroStrategy, Business Objects, Cognos, Pentaho, Jaspersoft and other leading BI platforms
Database portability	Netezza, Teradata, Oracle, Microsoft SQL Server, MySQL, Sybase, IBM DB2, and other relational databases
MicroStrategy bundle (optional)	BI Intelligence Server, MicroStrategy Web, MicroStrategy Desktop
AquaFold (optional)	Aqua Data Studio

Node Hardware	
Server	Dell PowerEdge R710
Processor	Dual quad-core 2.0GHz Xeon
Memory	24GB
Raw storage capacity	2.4 TB (8x300GB SAS drives)
Network	Gigabit Ethernet

About Aster Data Systems

Aster Data Systems is a proven leader in high-performance analytic database systems for data warehousing—the first DBMS to tightly integrate SQL with MapReduce—providing deep insights on data analyzed on clusters of low-cost commodity hardware. The Aster nCluster database cost-effectively powers frontline analytic applications for companies such as Coremetrics, MySpace, aCerno (an Akamai company), and ShareThis. Running on low-cost off-the-shelf hardware, and providing ‘hands-free’ administration, Aster enables enterprises to meet their data warehousing needs within their budget. Aster is headquartered in San Carlos, California and is backed by Sequoia Capital, JAFCO Ventures, IVP, Cambrian Ventures, and First Round Capital, as well as industry visionaries including David Cheriton and Ron Conway. For more information please visit <http://www.asterdata.com>, or call 650-232-4400.

Key Features

- **Raw compute power** – highest ratio of CPU and RAM per terabyte of data compared to other appliance offerings in the market
- **MPP architecture** – a network-optimized massively parallel processing architecture (MPP) enables efficient use of commodity hardware resources and provides maximum system performance
- **Ease of deployment** – plug-and-play appliance with a fully integrated stack from hardware through BI software provides seamless integration with existing IT systems
- **Ease of development** – support for standard development interfaces (SQL, JDBC/ODBC connectivity) along with most open language support in the market (Java, C#, C++, Python, Perl, etc.)
- **In-Database MapReduce** – powerful computing framework that reduces the time and effort required to implement analytic and transformation tasks operating on large structured datasets
- **Single-system manageability** – unifying fabric provides single-system interface to administrators. Automation of routine appliance administration while preserving visibility and control
- **Online fault tolerance and recovery** – highly resilient system continues query processing, even in the event of failures. Predictive heuristics and healing algorithms provide fast, automated recovery from both hardware and transient software failures
- **Fast queries** – patent-pending algorithms that maximize parallel processing performance and network efficiency to deliver enhanced query performance
- **Efficient parallel loading** – parallel loading scales linearly with additional nodes to meet any required SLA. Loading is concurrent with query processing – queries can continue uninterrupted during loading due to workload isolation
- **Online replica restoration** – in the event of node failure, system recovery initiates replica restoration in the background with no system down time. Change-tracking restoration algorithms offer ultra-fast recovery times