Big Analytics Insights in a Powerful, Enterprise-Ready Solution
Executive Summary

Technology and organization leaders like you wrestle every day with the challenge of capturing and exploiting huge and growing volumes of data. Perhaps you’ve mastered the collection and use of traditional relational data from transactions and business software. But the volume of new unstructured and semi-structured data – from web interactions, mobile devices, customer call center records, machine-generated data and video images – is exploding.

You’re sure this new data holds business value, if only you could readily analyze it. Now you can.

Introducing the Teradata Aster Big Analytics Appliance. The industry’s first unified big analytics appliance provides a powerful, ready-to-run big analytics and discovery platform that is pre-configured and optimized specifically for big data analysis.

This appliance gives you the power and performance to create new high-value analytics. With the patented SQL-MapReduce® framework, the appliance lets your analysts work quickly and effectively, generating insight using existing SQL tools and knowledge. Integrated hardware and software solutions, including industry-leading Teradata software and hardware management tools, make it easy to manage the appliance and gain business value quickly.

The Teradata Aster Big Analytics Appliance is unlike any other product on the market. It provides the power you need from a 21st-century big data platform, while respecting the budget, staffing, and training needs of your enterprise.

Here’s how:

- By providing an analytics-optimized environment, the appliance helps your users perform rapid, on-the-fly big analytics.
- The appliance includes Apache Hadoop for big data capture, staging, and refining and Aster MapReduce Platform for powerful, iterative analysis. It also offers a comprehensive analytic library with over 50+ pre-built analytical functions for advanced analytics, such as graph, pattern, and path analysis.
- You can process and analyze multi-structured data with this tightly coupled hardware and software solution, which contains the Aster Database and Apache Hadoop.
- It is compatible with your existing business intelligence (BI) and extract, transform, and load (ETL) tools, and ANSI-standard Structured Query Language (SQL) applications for analysis.
- You can manage the appliance with enterprise-grade system management solutions.

With the Teradata Aster Big Analytics Appliance, you get exactly the right combination of intelligence, simplicity, and power you need to gain rich business insights from complex big data sets, using your existing BI tools and resources. It’s the fastest, surest path to quick value, with a low total cost of ownership.

CONTENTS

2 EXECUTIVE SUMMARY
3 TERADATA ASTER: BRINGING BIG DATA INSIGHT TO YOUR ORGANIZATION
   3 Unified Data Architecture for the Enterprise
   3 Teradata Aster Big Analytics Appliance
4 USE CASE: MULTI-CHANNEL CHURN ANALYSIS DRIVES NEW CUSTOMER INSIGHT
7 TERADATA ASTER BIG ANALYTICS APPLIANCE OVERVIEW
   8 New High-Value Analytics, with Superior Analyst Productivity
   9 Use Case: Multi-Touch Attribution Optimizes Marketing ROI
   9 Innovative, Unified Solution to Easily Extract Insights from Big Data
12 Use Case: Product Affinity Analysis Boosts Cross-Sell and Up-Sell
   12 A Complete, Enterprise-Ready Solution
13 TERADATA ASTER: DELIVERING THE VALUE OF HIGH-PERFORMANCE ANALYTICS
14 LEARN MORE
Teradata Aster: Bringing Big Data Insight to Your Organization

For the first time in 2011, data in the digital universe exceeded a zettabyte. According to industry analyst IDC\(^1\), data volumes are expected to more than double every two years between now and 2015. What does this mean for organizations like yours?

To keep pace with your competition, it’s not enough to simply collect operational data from your business systems. You must be able to store, manage, and analyze exceedingly large volumes of traditional and non-traditional data – everything from your customer transactions to their interactions with your organization. With the right tools, you can use this data to develop a 360-degree view of your customers, one that reveals not just their purchases, but also their intent, preferences, behaviors, and relevant life events.

But building that view of your customers and your organization requires huge volumes of data, much of which is unstructured or semi-structured – what many experts call “big data.” Many companies store big data volumes on open source platforms such as Apache™ Hadoop. For business analysts with SQL expertise to gain insight from Hadoop systems, they often need to work through engineering teams and data scientists, who understand how to develop in Hive or Pig for processing and querying data within Hadoop.

This can be an inefficient, time-consuming process. To help your users gain insight from big data volumes, you need a big analytics and discovery platform and powerful technology solutions that provides seamless SQL-based access to data in Hadoop or a database.

**UNIFIED DATA ARCHITECTURE FOR THE ENTERPRISE**

Teradata offers a Unified Data Architecture for the enterprise. This data architecture integrates relational database systems based on SQL and SQL-MapReduce software framework with Apache Hadoop.

In this architecture, each component plays an essential role in helping you realize new insight from big data volumes:

- **Hadoop** offers a low-cost platform for landing, staging, and refining raw data in batch. With Hadoop, you can capture extensive volumes of historical data and perform massive processing to refine data.
- The **Teradata Aster Big Analytics and Discovery Platform** tightly integrates SQL, MapReduce, and a library of analytic functions, allowing your business analysts to explore and uncover new insights through on-the-fly interactive analytics. Aster Database and SQL-MapReduce provide processing flexibility for big data volumes.
- **Teradata** provides the best-in-class integrated data warehouse, which supports your enterprise across a variety of analytic workloads and stores valuable data and metadata. Once insights are discovered in Aster, they can be operationalized by codifying in the schema and provided throughout the enterprise in Teradata.

By bringing together these components into an integrated solution, the Unified Data Architecture bridges the gap between the business language of SQL and the extreme processing power of Hadoop. Together, the unified data architecture provides a best-of-breed system that allows users to leverage the right tool for the job.

**UNDERSTANDING MULTI-STRUCTURED DATA**

To gain maximum insight, you must be able to analyze multi-structured data – that is, data in a variety of formats and types. This can include structured or relational data that you would find in a typical database, such as ERP data, inventory records, and supply chain records; as well as semi-structured data such as web logs, call detail records, sensor logs, social profiles and Twitter feeds; or unstructured data such as images, videos, web pages, and PDF documents.

**TERADATA ASTER BIG ANALYTICS APPLIANCE**

To help you meet your big data analytics needs, Teradata Aster introduces the Big Analytics Appliance, the industry’s first unified big analytics appliance. This appliance provides a powerful, ready-to-run big analytics and discovery platform that is pre-configured and optimized specifically for big data analysis.

The hybrid architecture includes Aster Database, Aster SQL-MapReduce, and Hortonworks Data Platform, providing a complete big data solution that helps you to store and analyze all types of data (See Figure 1.). The appliance provides a complete solution for big data analytics and discovery through unified interfaces for information discovery, analytics, and system management.

---

\(^1\) Zettabyte = 1 Trillion Gigabites

This product is the first appliance in the industry to combine big analytics technologies, such as Hadoop, MapReduce, graph, pattern, and path analysis, with ecosystem compatibility with BI and ETL tools, business-friendly ANSI-standard SQL, out-of-the-box MapReduce analytics, and mature, enterprise-grade system management solutions.

Offering the industry’s only patented SQL-MapReduce interface, the appliance lets your users perform interactive analysis on the fly. Business analysts can use their existing SQL tools and skill sets while benefiting from the processing power of MapReduce.

With this appliance, analysts can perform high-value analytics on new data types, productively generating new business insight. Unlike other appliances that require specialists and separate interfaces to access and analyze data in Hadoop, the Big Analytics Appliance provides a truly unified approach through Aster SQL-H™ for both business analysts and administrators. This integration is unique. The appliance deploys a new approach using Apache HCatalog and the Aster SQL-H technology. SQL-H is a conduit that allows new analysis to be created and schema changes to be adopted within Hadoop from Aster. The Hadoop environment is completely abstracted from analysts, so they can focus on what they do best, performing analysis and generating insights. This is enabled by an innovation provided by HCatalog in the Hortonworks Data Platform, which enables this metadata exchange.

The appliance offers purpose-built hardware and software, with prepackaged nodes, disks, software, and networking. A fast, dedicated, fault-tolerant 40 Gb/s InfiniBand network connects the Aster and Hadoop nodes; it can also be used to connect multiple Big Analytics Appliances and you can integrate with the Teradata Analytical Ecosystem via 10 GbE. The solution also offers simplicity by offering comprehensive management tools, such as Teradata Vital Infrastructure system and integration with Teradata Viewpoint.

By minimizing the number of moving parts to deploy, it offers easy, integrated management of an enterprise-ready MapReduce solution with the benefits of optimized performance, continuous availability, and linear scalability. As a result, you can bring data science to the organization by simply plugging the Big Analytics Appliance into your infrastructure. Like all Teradata appliances, the Big Analytics Appliance is engineered and fully supported by the most trusted name in enterprise data management and analytics, Teradata.

**Use Case: Multi-Channel Churn Analysis Drives New Customer Insight**

How can the Teradata Aster Big Analytics Appliance help companies like yours to gain greater insight from big data analytics? Companies in various industries can use the Teradata Aster solution to address critical business problems. Let’s consider some typical examples.

A telecommunications provider notices an increase in early cancellation of service by customers. It is significantly more costly to acquire a new customer than to retain an existing customer. The company wants to learn more about customers who terminate a contract after interacting with one or more of the company’s channels (including online, call center, or a store location), even when additional fees or penalties are levied.

**Other Sample Industries that Benefit from Multi-Channel Churn Analysis**

- **Retail banking** – to understand and prevent customers from closing their account with the bank
- **Media and content providers** – to determine why an increasing number of customers are cancelling their paid monthly access to content.
ENTERPRISE-READY HADOOP

The Teradata Aster Big Analytics Appliance is configurable with nodes that run the Hortonworks Data Platform. The Big Analytics Appliance is an engineered solution that combines the power of Teradata Aster and Hadoop to deliver unprecedented power, performance, and flexibility for the enterprise. It is easy to deploy and support through:

- **Hortonworks Data Platform** - an enterprise-class, 100% open-source distribution of Apache Hadoop that is complete, stable, and reliable. Hadoop MapReduce, Hive, Pig, and other projects are fully enabled and supported.

- **Joint engineering** - a top-to-bottom integration between platforms, from query and analytics to storage, monitoring, and management. This includes key innovations in data access methods and metadata management, such as Apache HCatalog and Aster SQL-H, which simplify analytics and lower administration time.

- **World-class Teradata support** - backed by the unparalleled Teradata support team as well as integration with Teradata Vital Infrastructure, which offers proactive system monitoring and alerting to mitigate the risk of production deployments and increase uptime and reliability.

Teradata Aster eliminates the hard parts of adopting big data analytics and Hadoop, so you can focus on discovering new insights from your data and adding value to the business.

The company wants to identify the top 10 common paths and analyze specific patterns of behavior for these customers. By understanding the paths and patterns customers follow, the company hopes to take the right actions at the right time to prevent an at-risk customer from cancelling service.

To fuel this analysis, the company plans to use multi-structured data from three separate hardware platforms and database technologies:

- **Unstructured call center data stored in interactive voice recognition logs on Hadoop.** Hadoop is useful for landing and storing data, especially data in file formats with no schema. Hadoop can also be used to transform data, for example, in data cleansing and refinements operations.

- **Semi-structured web log data from user activity on online store web sites, which is stored on Aster.**

- **Structured retail store data from the Teradata Warehouse.** Teradata is most often used as a high-performance data warehouse platform for structured data.

Individually, each of these data sources provides some useful insight into customer behavior. By combining this data, the company hopes to better understand behavior in all of its channels, thereby gaining insight into customer attrition behavior.

By using the Teradata Aster Big Analytics Appliance, the company can store and analyze the data from all three sources. The appliance integrates data across the Teradata Warehouse and Aster Database with other technologies, such as Hadoop.

Using Aster SQL-H, a new feature of Aster Database 5.0, the company’s business analysts and data scientists can access data stored in Hadoop and perform iterative analysis in the Aster Discovery Platform. SQL-H empowers business analysts to directly analyze vast amounts of Hadoop data without requiring complex MapReduce programming skills or an understanding of how data is stored within the Hadoop Distributed File System (HDFS). The appliance also provides pre-built analytic functions, such as nPath, that help identify patterns in customer behavior.

To transform the call center data in Hadoop to a more usable format, an analyst working from the Aster Database creates a view into Hadoop using a SQL-MapReduce function load_from_hcatalog. Hive_
callcenter is the metadata listed in the Hcatalog in Hadoop that corresponds to the call center log files.
The analyst uses nPath to perform path and pattern analysis. Iterative analysis enables the analyst to discover interesting patterns and trends. To reduce the noise and find the signal, the analyst analyzes all service cancellations that involve one or more customer complaints.

![Figure 2. Performing Path and Pattern Analysis with nPath](image)

**CREATE VIEW hcat_telco_callcenter AS**
select “customer_id”, “sessionid”, “channel”
from “nc_system”.“load_from_hcatalog”
(on “public”. “mr_driver”
server (‘hadoop.asterdata.com’) port (‘9083’) 
dbname (‘default’) 
tablename (‘hive_callcenter’) 
username (‘hive’));

Next, the analyst creates a three-way view or join of the data as input to nPath, a pre-built analytic function in the Aster SQL-MapReduce Platform. In the SQL code below, store data is from the Teradata Warehouse, online data is from Aster Database, and call center data is stored in Apache Hadoop.

**CREATE VIEW td_telco_multi AS**
select “customer_id”, “sessionid”, “channel”
from ( ( select * from “td_telco_store”) union all 
( select * from “hcat_telco_callcenter”) union all 
( select * from “telco_online” ) )

The analyst uses nPath to perform path and pattern analysis. Iterative analysis enables the analyst to discover interesting patterns and trends. To reduce the noise and find the signal, the analyst analyzes all service cancellations that involve one or more customer complaints.
This analysis – which would have required hundreds of lines of code, hard-to-find skill sets, and weeks of work – was completed in just three quick steps.

![Figure 3. Common Paths to Service Termination When Customers File a Complaint](image)

Teradata Aster Big Analytics Appliance Overview

The Teradata Aster Big Analytics Appliance is a powerful, ready-to-run platform that supports rapid, on-the-fly data exploration. It is the industry’s first appliance to combine big analytics technologies – such as Hadoop, MapReduce, graph, pattern, and path analysis – with business-friendly ANSI-standard SQL tools for BI and ETL. With this appliance, new data-based insight is available to all of your business users, not just your power analysts or data scientists.

The appliance offers a hybrid architecture that includes Aster Database, Aster SQL-MapReduce and Apache Hadoop. You get a complete big analytics solution that helps you process structured, unstructured, and semi-structured data, all within a single, tightly coupled hardware and software offering.

The following sections describe the appliance features that can help you realize new and enhanced business insight.

**TOP-10 KEY FEATURES**

- MapReduce integrated with and encapsulated in SQL
- Out-of-the-box analytic library, with over 50 functions, including graphing, path, pattern
- Massively parallel processing database for speed-of-thought analytics
- Enterprise-ready Hadoop solution
- SQL-based access to Hadoop
- Integrated, single vendor hardware and software solution is easy to deploy, manage, and troubleshoot
- Tight integration with Teradata Analytical Ecosystem
- Lower skill ramp-up requirements
NEW HIGH-VALUE ANALYTICS, WITH SUPERIOR ANALYST PRODUCTIVITY

The Teradata Aster Big Analytics Appliance gives you the best-in-class analytics technologies you need to explore and gain insight from a wide variety of new data sources. Through the patented SQL-MapReduce framework and the Aster MapReduce Analytics Portfolio of pre-built functions, the appliance lets business users perform interactive exploration and analysis through successive, quick iterations. It’s the fastest way to extract unprecedented business insight that can directly impact your bottom line.

SQL-MapReduce

SQL-MapReduce is Teradata Aster’s patented parallel programming framework which combines the analytic power of MapReduce with the familiarity of SQL. This framework lets business analysts leverage the power of MapReduce from any SQL statement without needing to learn MapReduce programming or parallel programming concepts. (See Figure 4.)

SELECT * FROM nPath (
    ON td_telco_multi
    PARTITION BY customer_id
    ORDER BY datetimestamp
    MODE (NONOVERLAPPING)
    PATTERN
        ('REVIEW_CONTRACT_ONLINE.(CALLCENTER_COMPLAINT)[2].
         STORE_COMPLAINT.CANCEL_SERVICE')
    SYMBOLS (channel = 'ONLINE' AND action = 'REVIEW_CONTRACT' AS
        REVIEW_CONTRACT_ONLINE,
        ....
    )

Figure 4. Analysis Using the Power of MapReduce with the Familiarity of SQL

Huge data volumes are no problem for this powerful appliance. The SQL-MapReduce framework embeds and automatically parallelizes analytics applications, so users can perform ultra-fast, deep analytics on massive data volumes. Aster Database automatically parallelizes the application processing using MapReduce so that any in-database application runs in a massively parallel processing environment. Using SQL-MapReduce, developers can write powerful and highly expressive functions in a variety of languages, including Java, C, C++, C#, Python, and R, and push them into the database for advanced in-database analytics.

Using SQL-MapReduce, any analytics code running in database or any MapReduce function can be incorporated into an analytic application through a SQL statement. Business analysts can analyze all types of data from a single interface, leveraging existing investment in SQL-based BI and ETL tools and resources. Pre-packaged analytic toolsets for business intelligence or data mining that use standard SQL can natively access a MapReduce-enabled analytic application without any code changes, making the power of MapReduce easily and transparently accessible to business analysts.

Aster MapReduce Analytics Portfolio

The appliance includes Aster MapReduce Analytics Portfolio, the industry’s largest library of pre-built MapReduce functions. It includes more than 50 out-of-the-box MapReduce functions and integrated development tools, which help users get started quickly with big data analytics to uncover new insights from multi-structured data. Pattern, path, market basket, graph, text, and advanced statistical analysis are as simple as writing a single SQL statement to call the appropriate pre-packaged module embedded within the Aster MapReduce Platform.

For example, nPath is a sequential and trending pattern analysis function built on SQL-MapReduce. It discovers relationships between rows of data that usually cannot be expressed through SQL. Applications include customer shopping sequences, telephone calling patterns, stock market trading sequences, and more.
Some examples of analysis enabled by the pre-built analytic functions are:

- **Path and Pattern Analysis** — Discover patterns in rows of sequential data
- **Marketing Analytics** - Analyze customer interactions to optimize marketing decisions
- **Segmentation/Clustering Analysis** — Discover natural groupings of data points
- **Graph and Relational Analysis** — Discover important relationships between data points
- **Sentiment and Text Analysis** — Derive patterns and extract features from lengthy descriptive fields and text data
- **Statistical Analysis** — Process common statistical calculations with exceptionally high performance
- **Data Transformation** — transform raw data for more advanced insights
- **Data Parsers** — Parse raw multi-structured data sources like Apache web logs and heavily nested XML feeds

The appliance offers the industry’s highest performance and scalability. Using an optimized hardware configuration that connects the Aster and Hadoop nodes and a 40 Gb/s hyper-scalable InfiniBand back end, the product provides rapid analytics performance. Compared with commodity Apache Hadoop solutions, the appliance has been proven to deliver up to 35x faster speeds for interactive SQL and MapReduce analytics.

**INNOVATIVE, UNIFIED SOLUTION TO EASILY EXTRACT INSIGHTS FROM BIG DATA**

The Teradata Aster Big Analytics Appliance offers a single unified platform that lets you store and analyze multi-structured data from a wide variety of sources. You can use data collected from transactional and operational systems, as well as data collected from non-traditional sources such as call center records, web logs, social media, videos and images, and warranty information. The appliance is built on a hybrid architecture that includes Aster Database, including Aster’s patented SQL-MapReduce® and SQL-H™ technology, and Apache Hadoop. As a result, business analysts can easily access, join and analyze multi-structured data from a variety of sources from a single SQL interface.

**Other Sample Industries that Benefit from Multi-Touch Attribution**

- **Consumer products** – to understand how customers are engaging with their brands
- **High tech** – to gain insight into who influences buying decisions for products
- **Public sector** – to improve the effectiveness of public education and government service campaigns

**USE CASE: MULTI-TOUCH ATTRIBUTION OPTIMIZES MARKETING ROI**

To better understand the effectiveness of their campaigns, many marketing organizations want to move away from single-touch attribution methods such as last-click and first-click.

For example, an electronics retailer wants to understand the customer’s complete buying journey with more clarity and more detail. What information do consumers consider when they are researching a purchase? Where do they go to get information? Who do they talk to and who influences their decision?

Teradata Aster eliminates the hard parts of adopting big data analytics and Hadoop, so you can focus on discovering new insights from your data and adding value to the business.

With the Teradata Aster Big Analytics Appliance, the retailer gains a richer understanding of the many touch points – such as social, mobile, search, web sites, and offline channels – used by customers as they engage with a company or a product.

Most attribution solutions look at multiple touch-points within a single channel, such as an ad network or web visitors. The Teradata Aster Big Analytics appliance helps companies blend more channels into the mix to identify customer connections.

In addition, the appliance allows analysts to quantify channel effectiveness in driving revenue and pinpoint the best-performing channels. Analysts can use the solution to support time-sensitive campaigns and promotions with increased understanding of which channels are most likely to drive immediate or short-term revenue. This analysis is derived using the pre-built Attribution function in Aster. A marketing department or a chief marketing officer could look at such a report on a weekly or monthly basis to track which channels are more effective in...
Easy Access to Hadoop Data for Business Analysts Using Existing Tools and Skill Sets

Aster SQL-H™, a new feature of the Aster Database, is a major innovation on Aster’s integrated query capabilities. SQL-H empowers business analysts to directly access vast amounts of data in Hadoop for advanced analytics, a feature previously restricted to highly trained data scientists.

With SQL-H, analysts can use common BI and reporting tools, leveraging their existing business knowledge and SQL skills. They can access data in Hadoop directly, easily join it with data as needed in Aster Database, and use the analytical power of SQL-MapReduce and business-ready analytic functions and applications, such as click-stream analysis, marketing attribution, and graph analysis. (See Figure 5.)

Aster SQL-H and HCatalog Workflow

- Step 1: Copy data to the HDFS
- Step 2: Create metadata table mapping within HCatalog with direct access to HDFS
- Step 3: Run user queries using SQL-H against the HCatalog metadata

Figure 5. Analytics on Hadoop Data with Aster SQL-H

Aster SQL-H interfaces with Apache HCatalog, a table and schema management layer for Hadoop. HCatalog is a new open source metadata and storage management service in the Hadoop environment. It provides a shared schema and data type mechanism for data across the Hadoop Data Platform as well as a table abstraction function.

The HCatalog table abstraction presents users with a relational view of data in the HDFS. Users can work with Hadoop data without needing to understand where or in which format data is stored. HCatalog supports RCFile, CSV text, JSON text, and SequenceFile formats. (See Figure 6.)

Within Hadoop, users can employ different data processing tools - Pig, MapReduce, Hive, Streaming or a database like Aster - to more easily read and write data on the grid. HCatalog uses Hive-like DDL commands, and it supports tables, views, and partitions.

With this access, users gain unprecedented visibility and flexibility into the available Hadoop data. They can iteratively include more data and new data in their analyses, without any manual tasks or help from data scientists.
The appliance also provides the industry’s deepest stack integration across the Teradata Aster and Apache Hadoop platforms. InfiniBand connections between the Aster and Hadoop nodes support high performance. Using the SQL-H integration, organizations can retain archival data on less costly Hadoop storage resources while enabling analytics and discovery capabilities provided by the Aster nodes.

**Unified Hardware Architecture**

The heart of our industry-leading functionality and innovation is the hardware that supports the Teradata Aster Big Analytics Appliance. Within this single unified solution, you can choose the nodes that make sense for your usage.

The solution provides up to 18 MPP nodes per cabinet, with Intel Sandy Bridge dual-8 core 2.6GHz processors. (See Table 1.) Each Aster data node can store up to 5.5TB of data, with 9.5TB of data on each Hadoop node. Depending on workload needs, the appliance can be configured with Aster nodes exclusively, only Hadoop nodes, or a mixture of Aster and Hadoop nodes. This flexibility lets you create precisely the appliance configuration that will most effectively meet your business needs.

### Model: Aster Big Analytics Appliance

<table>
<thead>
<tr>
<th>Description</th>
<th>Fully-integrated cabinet design including:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nodes</strong></td>
<td>Up to (18) MPP Nodes Per Cabinet</td>
</tr>
<tr>
<td></td>
<td>- Intel Sandy Bridge Dual Eight Core 2.6GHz Processors (Aster Queen/Worker/Hadoop Master)</td>
</tr>
<tr>
<td></td>
<td>- Intel Sandy Bridge Dual Six Core 2.0GHz Processors (Aster Loader/Backup/Hadoop Data)</td>
</tr>
<tr>
<td><strong>Cabinet Configurations</strong></td>
<td>All Aster, All Hadoop, or Mixed. Optional Aster Backup nodes, optional Aster Loader nodes. Hadoop nodes ordered in 9 node increments</td>
</tr>
<tr>
<td><strong>OS</strong></td>
<td>SuSE Linux 11 64-bit operating System</td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>(432) 900GB Internal Drives per Cabinet (Aster Queen/Worker)</td>
</tr>
<tr>
<td></td>
<td>(216) 3TB Internal Drives per Cabinet (Aster Backup/Hadoop Data)</td>
</tr>
<tr>
<td><strong>Compression</strong></td>
<td>Up to 3X software compression available</td>
</tr>
<tr>
<td><strong>User Data Capacity (uncompressed)</strong></td>
<td>Aster Worker: RAID 5 – 5.5 TB / node</td>
</tr>
<tr>
<td></td>
<td>Hadoop Data: RAID 0 – 9.5 TB / node</td>
</tr>
<tr>
<td></td>
<td>Aster Backup: RAID 6 – 25.1 TB / node</td>
</tr>
<tr>
<td><strong>Scalability</strong></td>
<td>Scales up to 1000 cabinets; Scales up to 5PB for Aster only, with larger network switches; Scales up to 10PB for Hadoop only, with larger network switches;</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>Aster Queen / Worker / Hadoop Master: 256 GB / node</td>
</tr>
<tr>
<td></td>
<td>Hadoop Data: 128 GB / node</td>
</tr>
<tr>
<td><strong>Interconnect</strong></td>
<td>40 Gb/s InfiniBand node to node</td>
</tr>
<tr>
<td></td>
<td>10 GbE to Teradata EcoSystem</td>
</tr>
<tr>
<td><strong>System Management</strong></td>
<td>Teradata Server Management for hardware &amp; software alerts</td>
</tr>
<tr>
<td></td>
<td>Viewpoint 14.01 for monitoring and management</td>
</tr>
<tr>
<td></td>
<td>Aster Management Console for configuration</td>
</tr>
<tr>
<td><strong>Enterprise Integration</strong></td>
<td>High speed Aster-Teradata Connector</td>
</tr>
</tbody>
</table>

*Table 1. Teradata Aster Big Analytics Appliance Feature Descriptions*
The appliance also benefits from the ongoing innovation developed by the technical expertise of Teradata Aster. For example, the appliance includes features such as:

- Upgraded hardware and connectivity, including a dual-8 core CPU, which supports increased memory and high-performance networking
- Increased storage capacity
- Enhanced cabinet configuration flexibility
- A 42U rack with multiple system support, so you can share racks with other systems or departments

**USE CASE: PRODUCT AFFINITY ANALYSIS BOOSTS CROSS-SELL AND UP-SELL**

Cross-selling and up-selling products is a surefire way for companies to boost revenues and profitability.

For example, a retail merchandiser wants to know which products customers typically buy in the same shopping experience. Whether the customer makes the purchase in the store or online, understanding product affinities can help the merchandiser gain insight into how to market and position items on the shelf or on the web site.

By using the Teradata Aster Big Analytics Appliance to analyze customer shopping baskets, interaction data, and transaction data, the merchandiser can identify products that tend to be viewed or purchased together.

The appliance helps the merchandiser generate sets or baskets of items that occur together in the transaction records or web page logs. With an understanding of how strongly two or more items are related in customer purchasing behavior, retailers can determine which items to offer customers who are buying or have purchased related items.

Other Sample Industries that Benefit from Product Affinity Analysis

- **Media and content providers** – to up-sell premium content to existing customers
- **Telecommunications providers** – to cross-sell new services
- **Manufacturing** – to better understand relationships between parts failures, improving warranty performance

**A COMPLETE, ENTERPRISE-READY SOLUTION**

To stay ahead of the competition, your organization needs to gain new insight quickly and cost-effectively. The Teradata Aster Big Analytics Appliance provides a single, enterprise-ready solution that helps you collect, store, and analyze new types of data.

This integrated hardware and software solution is engineered and fully supported by Teradata. Because it is integrated with the Teradata Analytical Ecosystem, the appliance provides you with an expanded range of powerful analytical capabilities.

A suite of adapters supports out-of-the-box enterprise integration, for a complete ecosystem of enterprise data management systems. Included are certified ODBC and JDBC support for major business intelligence, visualization, and ETL tools, the Teradata integrated data warehouse high-speed data transfer infrastructure, and native Hadoop connectivity.

Yet the appliance doesn’t skimp on the data-crunching power you need to find that specific insight you need to drive your organization. Using Hortonworks Data Platform 1.1, the appliance delivers the performance of 100% open source Apache Hadoop.

Existing Teradata customers will find the management consoles and capabilities familiar. The appliance delivers comprehensive enterprise-class software and hardware management by providing support for Teradata Server Management, Teradata Vital Infrastructure (TVI) and Teradata Viewpoint.

**Industry-Leading Software and Hardware Management**

When considering a big analytics solution for your organization, you must look not only for performance and affordability, but also for manageability, too. The Teradata Aster Big Data Appliance offers a combination of powerful software and hardware management offerings that ensure your solution delivers maximum value.

Integration with Teradata Server Management enables proactive monitoring of hardware and software events, such as disk or node failures. With the Teradata proactive support software available on each Teradata Aster Big Analytics Appliance, support information gathered is automatically routed to Teradata Customer Services’ back-end support infrastructure, so issues can be addressed quickly thereby minimizing system downtime.
Teradata Server Management is a set of components that detect and report hardware and software exceptions (faults), hardware and software asset data, and supporting diagnostic data. It equips support representatives with detailed configuration information from your system and provides electronic notification of faults. Teradata Server Management runs in virtual machines on a series of Virtual Management Server (VMS) nodes/servers. Each VMS manages one or more hardware cabinets to provide fault and asset management for the system. The VMS supports Teradata hardware, Aster Database, and Hortonworks Data Platform. It also provides Teradata Vital Infrastructure support for Teradata Aster and Hadoop.

Teradata Vital Infrastructure is the end-to-end solution for delivering Server Management data to Teradata Customer Service’s back-end support infrastructure. The built-in TVI support software continually collects, retains, and analyzes information about your system. When a fault event is detected, TVI creates automatic incident reports, and sends alert notifications to Teradata support staff.

The VMS consolidates several software servers onto a single physical chassis. Virtualization technology allows multiple operating environments to run on a single VMS. These operating environments are called virtual machines (VMs). Each system contains a single System VMS which can host the following VMs:

- **The CMIC (Cabinet Management Interface Controller)** - VM provides server management services for the system in which it is located.
- **The SWS (Service Workstation)** - VM is the service entry for the system.
- **The Viewpoint** - VM manages a Teradata system.

Each CMIC VM constantly monitors hardware and software systems for exceptions and asset or configuration changes. It applies rules to any exceptions and generates Alerts.

When TVI is enabled, and when warranted, TVI escalates (forwards) the alerts and other types of data to the Teradata Customer Services back-end infrastructure via connectivity software running on the SWS VM. Asset, configuration, diagnostic data and alerts are recorded in a Teradata support repository, and when the Customer Service Contract allows, Incidents are automatically created based on the escalated data. All information is securely maintained at Teradata Service Center.

The diagnostic information collected by TVI assists support personnel in identifying and quickly resolving problems; conversely, omission of this proactive software may prolong the time it takes to resolve incidents. In fact, internal Teradata studies indicate that it uncovers 62% to 70% of all system incidents.

### Integrated Monitoring and Management through Familiar Tools

Integration with Teradata Viewpoint offers a common management console for the Aster Database, Teradata integrated data warehouse, and Hadoop with the Hortonworks Data Platform. Teradata Viewpoint offers simpler, faster, and more comprehensive system management by providing a browser-based portal that delivers management intelligence to DBAs and users alike. Viewpoint 14.01 enables Aster portlets such as the Aster Node Monitor for Aster system monitoring. Apache Hadoop software monitoring will be available in Viewpoint 14.10.

Viewpoint Portlets, a suite of management and monitoring applications, gives business users access to vital system information. The portlets help administrators monitor queries, system health, and Aster related alerts. (See Figure 9.) They enable visualization through features such as capacity heat maps and metrics graphs. Users can check browser-based Viewpoint screens to check the status of their queries or the health of the system.

![Figure 7: System Health Portlet Displays a Summary of States with Drill-Down Ability](image-url)
Teradata Aster: Delivering the Value of High-Performance Analytics

The Teradata Aster Big Analytics Appliance can make your entry into big data enterprise analytics fast, efficient, and cost-effective. The appliance is a powerful, ready-to-run platform delivered by Teradata, the world’s largest company focused on analytic data solutions through integrated data warehousing, big data analytics, and business applications.

Pre-configured and optimized specifically for big data storage and analytics, the appliance runs the Teradata Aster patented SQL-MapReduce and SQL-H technology on a time-tested, fully supported Teradata hardware platform.

Using this appliance, you can expect to realize the following business benefits:

- **New high value analytics for analysts and business users**, who can use traditional SQL-based tools and skill sets along with an analytic library with more than 50 pre-built MapReduce functions for analysis such as path, customer behavior, marketing, graph and text analytics

- **Greater insight across the enterprise**, thanks to big data storage capabilities and integration with the Teradata Analytical Ecosystem, allowing business users to explore a wide variety of traditional and new data sources

- **Faster analytics results**, with an analytics-optimized environment for rapid, on-the-fly data exploration and powerful processing capabilities that help you make the best decisions possible

- **Rapid time to value and low TCO**, through easy installation, configuration flexibility, and manageability

- **High performance and availability**, with the proactive monitoring features of Teradata Server Management, Teradata Viewpoint, and Teradata Vital Infrastructure services

Learn More

For more information about how the Teradata Aster Big Analytics Appliance can bring value to your organization, contact your Teradata or Teradata Aster representative or visit us on the web at