

TERADATA PORTFOLIO FOR HADOOP

TERADATA

BRINGING HADOOP TO THE ENTERPRISE

In recent years, Apache™ Hadoop® has emerged from humble beginnings to disrupt the traditional disciplines of information management. As with all technology innovation, the hype is rampant—and data professionals can be overwhelmed by new Hadoop projects, diverse opinions, and confusing messages about how the enterprise architecture can or should change to harness the value of Hadoop.

Yet, regardless of technology advances, the premise behind data management and analytics hasn't changed. The big data movement simply represents the next phase in the evolution of technology and data cultures within organizations that view information and insights as a competitive advantage. Many organizations are looking for cost-effective ways to capture the value of dark and extraneous data (and other new data types) that in the past were often discarded due to cost, scale, or processing limitations.

So what is the best role for Hadoop in your existing data environment? By using Hadoop as a tightly integrated data management area in your architecture, you can seamlessly refine data for feeding downstream analytical systems. Hadoop can extend your data architecture to provide new processing capabilities for new data types and to enable new applications. Additionally, with the right set of skills and technologies, big data exploration and single-subject analytical processing on raw data in Hadoop is easily implemented.

What should your organization do? Every business is unique—each has its own data processing and storage needs, as well as different levels of skill sets. Some already take advantage of Hadoop as cheap storage on commodity servers, while others are adding Hadoop to their enterprise data architectures to complement the existing data warehouse, discovery platform, event processing, business intelligence (BI), and data integration systems. But every company, including yours, shares some common requirements regarding their needs for new and better ways to glean more value from data.

ENTERPRISE-READY PRODUCTS AND SERVICES

The premier offerings of the Teradata Portfolio for Hadoop are the **Teradata® Appliance for Hadoop** and the **Teradata Aster Big Analytics Appliance**. These appliances are designed to make Hadoop ready for enterprise deployments and to bring you an array of built-in benefits, including:

Business

- ~ Enables you to focus on your business, not technology.
- ~ Delivers fastest time to Hadoop production, accelerating time-to-value.
- ~ Minimizes risk with proven technology, services, and support.
- ~ Leverages existing skill sets and investments in BI and SQL tools.
- ~ Drives lowest total cost of ownership.

Technical

- ~ Teradata Open Distribution for Hadoop 2.1, built upon 100 percent open-source Apache Hadoop from Hortonworks
- ~ Enterprise-class data integration to and from Hadoop with Teradata® QueryGrid™, the Teradata Connector for Hadoop (TDCH) and Teradata Aster File Store
- ~ Enterprise-class server nodes and storage configured specifically for Hadoop workloads.
- ~ High throughput BYNET™ V5 on a 40GB/s InfiniBand interconnect.
- ~ Easily scalable to petabytes of customer data.
- ~ Highly available and reliable enterprise-class hardware with simplified recovery features.
- ~ Robust data management and proactive monitoring tools.

TERADATA PORTFOLIO FOR HADOOP

TERADATA PREMIER PLATFORMS FOR HADOOP		TERADATA COMMODITY PLATFORMS FOR HADOOP	
TERADATA APPLIANCE FOR HADOOP	TERADATA ASTER BIG ANALYTICS APPLIANCE	TERADATA COMMODITY OFFERING FOR HADOOP	TERADATA SOFTWARE-ONLY FOR HADOOP (HORTONWORKS DATA PLATFORM SUPPORT)
Enterprise-ready Hadoop in an integrated package with single-vendor support.	Big data analytics and storage in an integrated package with single-vendor support.	Pre-designed, cost-effective commodity Dell® hardware configuration with Teradata software support.	Flexible hardware options with Hortonworks Distribution of Hadoop, backed by Teradata software support.

Figure 1.

GETTING VALUE FROM HADOOP WITH TERADATA

Companies are searching for the best ways to get value from *all* of their data. They are examining whether or not Apache Hadoop fits in their data architecture, and if so, how? Considering a trusted partner like Teradata for big data analytics makes sense. Teradata is the world's largest company fully dedicated to integrated data warehousing, data management, big data analytics, and business applications. For more than 30 years, we've had a laser-like focus on helping customers achieve competitive advantage through analytics. We've always taken a best-of-breed approach with workload-specific platforms, and we are now extending that approach to include Hadoop. The **Teradata Portfolio for Hadoop** is a flexible offering for our customers to integrate Hadoop into their data architecture while taking advantage of the world-class support Teradata provides. We offer a selection of products and services to suit every budget, from fully-integrated appliances to consulting and support services on existing customer hardware. Training for Hadoop is also available from Teradata. We are now your single source for all things Hadoop.

The Teradata Portfolio for Hadoop adds business value by making Hadoop enterprise-ready with integrated hardware, software, services and support, as well as management and connectivity improvements. By partnering with leading Hadoop distribution vendors such as

Hortonworks, Teradata seamlessly integrates Apache Hadoop as a key component within its **Teradata Unified Data Architecture™**, an enterprise-class architecture that creates an environment designed for volumes of data—both the mission-critical data used to run the business as well as big data to discover new insights.

CHOOSE THE PERFECT PLATFORM FOR HADOOP

Whatever your specific requirements, there's a Teradata portfolio product to meet your Hadoop needs. Our premier offerings are the Teradata Appliance for Hadoop and the Teradata Aster Big Analytics Appliance. Both are integrated hardware and software stacks, delivered ready to run, with complete Teradata support and service. Our commodity offerings are a predesigned Dell hardware stack and Hadoop-only software, both with Teradata support and service.

Teradata Appliance for Hadoop

The Teradata Appliance for Hadoop is a tightly integrated hardware and software stack optimized for enterprise-class big data storage and management. Performance, ease-of-use, security, and reliability are greatly improved from the typical do-it-yourself Hadoop. Delivered as a fully-integrated system that's ready to plug in and use, the Teradata Appliance for Hadoop is purpose-built for big data storage on multi-structured

data. The appliance can hold up to 205TB uncompressed customer data space per cabinet, with the entire system scaling up to 13PB. The appliance is networked by Teradata's fabric-based computing on the 40GB/s InfiniBand network for fast data exchange among Teradata, Teradata Aster, and Hadoop appliances. It features Teradata Open Distribution for Hadoop based on 100-percent open-source Apache Hadoop from Hortonworks, on a proven Teradata hardware platform with dual six- and ten-core Intel® processors and enterprise-class storage—all preinstalled into a power-efficient unit. That means you can have the system up and running live in just a few hours for rapid time-to-value.

For the complete Teradata Appliance for Hadoop datasheet, please [click here](#).

Teradata Aster Big Analytics Appliance

The Aster Big Analytics Appliance is a powerful discovery platform optimized specifically for big data analysis and storage. A purpose-built, integrated hardware and software solution for analytics at the big data scale, the appliance runs patented Teradata Aster SQL-MapReduce®, Teradata Aster SQL-GR™, and Teradata QueryGrid™ technology on a time-tested, fully supported Teradata hardware platform. Depending on workload needs, it can be configured exclusively with Teradata Aster nodes, Teradata Open Distribution for Hadoop nodes, or a mixture of the two. Integrated backup nodes are also available for data protection.

For the complete Teradata Aster Big Analytics Appliance datasheet, please [click here](#).

COMPLETE CENTRALIZED MANAGEMENT, SERVICE, AND SUPPORT

Specific enterprise enhancements to the Hadoop experience have been added to both appliances to make administering your Hadoop cluster less time-consuming. **NameNode failover** is an important high-availability feature that works out of the box for faster recovery and higher overall system reliability. Intelligent start and stop is a graph-based approach for automated launching of all Hadoop services, which traditionally had to be manually implemented. Typically cumbersome hardware procedures, such as adding a bare data node to the cluster or replacing single drives, have been simplified.

As with all Teradata platforms, the appliances feature simplified administration, control, and monitoring through the single operational view of Teradata Server Management. All Hadoop nodes are centrally monitored through a single Teradata Viewpoint interface, making it easier for the system administrators by improving scalability, reliability, and manageability. It combines intuitive tools for centralized management with powerful capabilities for facilitating and automating administration, and managing data and applications to minimize administrative work, even as the system scales. Moreover:

- ~ **Integration with Teradata Server Management and Teradata Vital Infrastructure** enables proactive monitoring of hardware and software events, such as disk or node failures. System warnings are sent to Teradata's global customer support team so issues can be addressed quickly, minimizing system downtime.
- ~ **Viewpoint Portlets**, a suite of management and monitoring applications, gives business users access to crucial system information. Hadoop software monitoring is available in Teradata Viewpoint 14.10.

In addition, Teradata offers a full array of support advantages, including:

- ~ Industry-certified regional and global **Customer Care Centers** with experienced service representatives available 24x7.
- ~ **End-to-end support** for the entire solution, from the hardware and OS to the Hadoop software.
- ~ **Secure remote connectivity options** to pinpoint problems and react rapidly.

Teradata Commodity Offering for Hadoop

Teradata, Hortonworks, and Dell have partnered to create a low-cost, standard hardware configuration that is optimized for the Hortonworks Data Platform. You can contract directly with Dell to purchase, install, and support the hardware, and Teradata will support the Hortonworks software.

Teradata Software-Only for Hadoop

For customers who wish to source their own hardware, Teradata offers support of Hortonworks Data Platform. You can easily choose to start your Hadoop journey here, running Hadoop on servers you already have. As you develop your data processing capabilities and applications built on Hadoop, you can go into production

with the existing configuration or move to one of the other higher-performance offerings within the portfolio. Either way, Teradata is there to support you during your Hadoop journey.

IMPLEMENTATION AND CONSULTING SERVICES OFFERINGS

Professional Services

Teradata Consulting Services—the most experienced source for data warehousing, analytics, and BI support—offers a number of new and expanded services to assist you with your Hadoop environment. These services help you realize the benefits of the Teradata Unified Data Architecture™ following a process refined through years of data and analytics expertise. Teradata Consulting Services focuses on the people, process, and technology required to realize your targeted solution, ranging from adding Hadoop capabilities to realizing a broader, seamless unified data architecture.

These services begin with a wide array of Business Improvement Opportunities (BIOs), expanded to include Hadoop. Optionally, an analytic roadmap can be developed to lay out a broader program designed to lead to the realization of multiple BIOs. Focused workshops provide a deeper dive into the nuances and options of various topics and technologies leading to a solution concept refinement that provides the necessary level of detail for the required objective.

Teradata's Analytic Architecture services have been enhanced to support the various design patterns and implementation alternatives enabled through Hadoop capabilities.

Additional Hadoop-focused services can support the Teradata customer in standing up a Hadoop environment, providing a working solution combined with guidance and advice on utilizing and managing the solution.

Once in place, the Teradata customer can leverage Teradata's Performance and Capacity services to optimize and forecast growth of the Teradata Unified Data Architecture™.

To round out the portfolio of services, the Teradata Managed Services offering has been expanded to include Hadoop and other targeted open-source components for administration and monitoring support of the environment.

Customer Support Services

Teradata Customer Support Services has a number of new and expanded services to assist customers with their Hadoop environments. These services help customers realize the many benefits of the Teradata Unified Data Architecture™. You can take advantage of Teradata's world-class support services for the Teradata Aster Big Analytics Appliance, the Teradata Appliance for Hadoop, Hadoop commodity offering on Dell, and software-only implementations of Hortonworks Data Platform. Teradata Customer Services Premier Support and Critical System Management will provide a full continuum of maintenance, support, and operational services on these environments.

Training

Teradata offers two training classes to enhance your knowledge of Hadoop. The first is mainly for administrators who wish to learn to deploy and manage a cluster. The other is for Java programmers who want to better understand how to create Apache Hadoop solutions.

MAKE THE MOST OF YOUR RESOURCES

The Teradata Portfolio for Hadoop makes it easy for you and your business to jumpstart your big data initiatives and projects, while leveraging your current resources and tools. This means your data scientists, systems administrators, DBAs, and application developers can support multiple Teradata and Hadoop systems without the need for additional training.

FOR MORE INFORMATION

To find out more about how the Teradata Portfolio for Hadoop can help you succeed with Hadoop, please contact your local Teradata representative, or visit Teradata.com.

	TERADATA APPLIANCE FOR HADOOP	TERADATA ASTER BIG ANALYTICS APPLIANCE	TERADATA COMMODITY OFFERING FOR HADOOP	TERADATA SOFTWARE-ONLY FOR HADOOP (HORTONWORKS DATA PLATFORM SUPPORT)
Hardware Features	High performance memory/CPU InfiniBand network Virtualized Management Server	High performance memory/CPU InfiniBand network Virtualized Management Server	Standard Dell commodity servers	Customer choice
Software Features	Teradata Open Distribution for Hadoop 2.1, built upon 100 percent open-source Apache Hadoop from Hortonworks NameNode failover Auto installation Hadoop tools Enterprise connectors Teradata Viewpoint Teradata QueryGrid™ Teradata Studio™ with smart loader	Teradata Aster Discovery Platform 6 100+ pre-built SQL-MapReduce® and SQL-GR analytic functions Teradata Open Distribution for Hadoop 2.1, built upon 100 percent open-source Apache Hadoop from Hortonworks NameNode failover Auto installation Hadoop tools Enterprise connectors Teradata Viewpoint Teradata QueryGrid™ Teradata Studio™ with smart loader	Hortonworks HDP Enterprise connectors Teradata QueryGrid™	Hortonworks HDP Enterprise connectors Teradata QueryGrid™
Support Features	Teradata Vital Infrastructure Single-vendor hardware/software support	Teradata Vital Infrastructure Single-vendor hardware/software support	Dell hardware support Teradata software support	Teradata software support
Services	Managed Services for Hadoop Big Analytics Data Staging for Hadoop	Teradata Aster Managed Services Managed Services for Hadoop Big Analytics Data Staging for Hadoop	Managed Services for Hadoop Big Analytics Data Staging for Hadoop	Managed Services for Hadoop Big Analytics Data Staging for Hadoop

Figure 2. Differentiation of features and services within the Teradata Portfolio for Hadoop.

TERADATA APPLIANCE FOR HADOOP DESCRIPTION

The Appliance for Hadoop features:

- Teradata Open Distribution for Hadoop (TDH) 2.1, Teradata certified version, built upon 100 percent open-source Apache Hadoop from Hortonworks
- SUSE Linux 11 64-bit Operating System
- Integrated Cabinet
 - Dual Intel Ten-Core Xeon® processors @ 2.5GHz per node (Hadoop Master Nodes)
 - Dual Intel Six-Core Xeon® processors @ 2.1GHz per node (Hadoop Data Nodes)
 - 40GB/s InfiniBand Node Interconnect

- Disk Capacities—uncompressed (3x compression assumed typical)
 - Data nodes for Hadoop with 4TB drives: 13TB user space
 - Full Hadoop cabinet (2 Masters/16 Data/1 optional Edge Node) - 205TB user space
- Scalability to petabytes with expansion cabinets and larger network switch
- Teradata software developed and/or optimized for Hadoop
 - System Management Infrastructure, Teradata Administration, and Teradata Viewpoint
 - High-speed data integration technologies between Hadoop,

the Teradata data warehouse and Teradata Aster includes Teradata QueryGrid, Teradata Connector for Hadoop, Teradata Studio™ with smart loader, and Teradata Aster File Store

- Third-party software
 - Informatica PowerCenter Big Data Edition provides data integration optimization and refining for Hadoop
 - Protegrity Big Data Protector provides comprehensive Hadoop data protection from the file to the application
 - Revelytix Loom provides dynamic data set management, metadata, and data lineage for Hadoop

SPECIFICATIONS

Cabinet

- Height: 80.5 in. (204.5 cm)
- Width: 24 in. (61 cm)
- Depth: 49 in. (124.4 cm)
 - 47 in. (119.4 cm) w/o the front and rear doors
- Weight: 2,228 lbs. (1011 kg) fully loaded with crate
- Installed Weight: 1,828 lbs. (830 kg) fully loaded w/o crate

Operating

- Operating Temperature: Allowable: 59°F to 90°F (15°C to 32°C); Recommended: 64.8°F to 80.6°F (18°C to 27°C)
- Relative Humidity: Allowable: 20% to 80% (non-condensing); Recommended: Low-end moisture: 5.5C DP (41.9F), High-end moisture: 60% RH and 15C DP (59F DP)

Electrical

- North America Voltage Range: 200 - 240V, 3-+PE
- Current:
 - 30A, 3-p, 4-wire, 4-cord
 - 60A, 3-p, 4-wire, 2-cord

International Voltage Range: 220 - 240/381 - 415, 3- +N +PE with three phase

- Current:
 - 32A/30A, 3-p, 5-wire, 2-cord, All continents (including North America with EU-style power)
- Frequency: 50Hz - 60Hz
- Maximum Power: 12.25kW
- Dual AC: Standard
- Compliant with U.S. and International Safety and Emissions Standards



10000 Innovation Drive Dayton, OH 45342 teradata.com



Unified Data Architecture, Query-Grid, and Teradata Studio are trademarks, and Teradata, the Teradata logo, BYNET, and SQL-MapReduce are registered trademarks of Teradata Corporation and/or its affiliates in the U.S. and worldwide. Apache is a trademark and Hadoop is a registered trademark of the Apache Software Foundation. SUSE is registered trademarks of Novell Inc., in the United States and other countries. Intel and Xeon are registered trademarks of Intel Corporation. Protegrity is a trademark of Protegrity USA. 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. Consult your Teradata representative or Teradata.com for more information.