

Faster Results at Less Cost The Data Warehouse Appliance Powered by Sun and Greenplum



Business intelligence is key to creating competitive advantage. Unfortunately, it often comes at a steep price from vendors such as Teradata. The Data Warehouse Appliance Powered by Sun and Greenplum provides a fresh approach that can deliver 10 to 100 times the performance of traditional data warehouse solutions at a fraction of the cost. Uniting open source software with best-in-class compute and storage resources, the appliance offers scalable data capacity and query throughput. Available for as low as \$17,500 per terabyte, this innovative appliance delivers ground-breaking price/performance for strategic decision support applications.

Greenplum Database

Central to the Data Warehouse Appliance is the Greenplum Database, the industry's first open source powered database server that scales to meet the demands of multi-terabyte data warehouse environments. Unlike closed and proprietary solutions from other vendors, such as Teradata, the Greenplum Database stores and processes terabytes of data using clusters of low-cost commodity servers.

A faster way to process data

Parallel processing is a necessity when it comes to supporting large-scale data warehousing demands. By using a shared-nothing architecture, Greenplum Database is able to move computations as close to the data as possible, and distribute data across compute nodes. Database operations, from data loading and query processing to backup and administrative tasks, are executed in parallel to help provide superior performance, scalability, and reliability.

Breakthrough performance of the Greenplum database is achieved by dividing storage into small segments on individual low-cost servers with multiple dedicated, independent, high-speed channel connections to local disks. Data segments are distributed across disks and I/O channels on servers which act as data execution hosts for database queries.

A database optimizer directs hosts to perform work in parallel using simultaneous disk connections. Single SQL statements are broken down into smaller components that are worked on by all nodes in the system, concurrently, to deliver a single result. The high-performance, software-based intercon-

nect in the Greenplum Database allows the system to move data among nodes during query execution time — a capability that is fundamental to the solution's throughput, flexibility, and scalability. Consequently, the database system acts as a set of self-contained parallel processing units, helping enable the solution to perform and scale seamlessly for even the most complex business intelligence workloads.

Sun Fire™ X4500 server

The underlying architecture of the Sun Fire X4500 server complements the parallelism in the Greenplum Database, creating a cost-effective, large capacity solution that features fast load speeds and data scans. Each Sun Fire X4500 server includes two dual-core, Next-Generation AMD Opteron™ processors and 24 TB of storage, providing enterprise-class compute power and high density storage in a compact, seven-inch, four-rack unit (4 RU) form factor. High bandwidth networking and multiple I/O channels in each server help speed data access in the appliance.

Sun Fire X4500 servers feature Integrated Lights-Out Management, providing full local or remote access for setup, maintenance, and ongoing monitoring. These tools work in concert with Greenplum administrative tools that supply real-time performance monitoring and streamlined management functions.

Sun Fire™ X4100 server

Each Data Warehouse Appliance includes a low-cost Sun Fire X4100 server, which acts as a dispatch server in the appliance to optimize query distribution. Greenplum software running on this system authenticates users and



Highlights

- Leave expensive Teradata systems behind with a high-performance, low-cost solution from Sun and Greenplum
- Select from 20 TB, 40 TB, and 100 TB configurations that can easily grow as needs dictate
- Create a better environment with an appliance based on Sun Fire™ X4500 servers and the Greenplum Database
- Scale deployments by adding nodes to appliances, or adding appliances to the environment
- Take advantage of the innovation found in industry standard hardware and open source software
- Experience 10 to 100 times the performance of traditional data warehouse solutions
- Spend a fraction of the price, with solutions available as low as \$17,500 per terabyte

connects to remote segment instances. Once SQL requests are parsed, the software forms an optimal parallel query plan, distributes it to all segment instances, coordinates execution, and returns the result to the user.

Solaris™ Operating System

All servers in the Data Warehouse Appliance run the Solaris™ 10 Operating System. Traditionally known for being open and scalable, the state-of-the-art Solaris 10 OS also boasts indisputable performance advantages for database services, with world-record benchmark results.

Massive data capacity and scalability

Solaris 10 offers new capabilities that help enhance the performance, scalability, reliability, and management of database solutions like the Data Warehouse Appliance. For example, Solaris ZFS is a revolutionary new file system that provides virtually unlimited file system scalability and increased data integrity to large-scale data warehousing applications. Solaris ZFS helps protect data with 64-bit checksums aimed at error detection and correction. Because data is mirrored, corrupted data can be automatically repaired, helping protect against costly and time-consuming data loss due to hardware or software failures.

One of the first commercially available 128-bit file systems, Solaris ZFS simplifies data management by eliminating the need for a volume manager, and executes system administration tasks within a matter of seconds with no file system downtime. Complicated storage management tasks are automated and consolidated, helping to reduce administrative overhead.

Virtual storage pools make it easy to expand or contract file systems simply by adding more disk drives, streamlining storage administration and enabling easy reallocation of resources.

Sun and Greenplum

Greenplum provides the high-performance software that helps manage massive amounts of data and turn it into useful information. Sun provides the products and technologies that make storing and processing that information possible. By bringing these innovative solutions together in the Data Warehouse Appliance, Sun and Greenplum give companies a visionary solution that transforms data warehousing environments and helps improve data analysis without breaking IT budgets.

Learn More

For more information, visit: sun.com/bidw, or contact your local sales representative.

Feature	DWAPP-DW20	DWAPP-DW40	DWAPP-DW100
Usable capacity			
• RAID-Z capacity	20 TB	40 TB	100 TB
• RAID-10 capacity	10.5 TB	21 TB	52.5 TB
Number of racks (42 RU per rack)	1	1	2
Rack units used	12	20	44
48-port 10/100/1000 switches	2	2	2
Sun Fire X4100 Server (Dispatch server)	1	1	1
Sun Fire X4500 Servers (Execution servers)			
• Data servers	2	4	10
• Total CPUs	4	8	20
• Total CPU cores	8	16	40
• Total RAM	32 GB	64 GB	160 GB
Greenplum Database licenses	4 CPU license	8 CPU license	20 CPU license

