



## INTEGRATED DATA WAREHOUSE FOR TERASCALE DATA MINING APPLICATIONS

/// FACT SHEET



Netezza provides highly specialised, purpose-built data warehouse machines that are capable of analysing vast amounts of data in near real-time. This greatly enhances overall capability as well as the response of data-intensive applications in logistics, command and control and intelligence.

### FEATURES

- Terascale data management appliance (DBMS software, system hardware, high-speed storage)
- Industry-standard interfaces (SQL, ODBC, JDBC)
- Full compatibility with market-leading analytic tools, applications and infrastructure
- Red Hat Linux Advanced Server operating system
- Support for Latin-9, Unicode (with UTF-8 encoding), with extended support for SJIS, EUC-JP and others

### Overview

The Netezza Performance Server® (NPS®) family of systems provides the basis for an authoritative data source for defence agencies to conduct research and analysis. The NPS system is a purpose-built data warehouse designed to meet the needs of terascale data management and analytics, where billions of rows are the norm, rather than the exception. Architecturally integrating database, server and storage into a single appliance, the NPS system delivers 10 to 100 times the performance of traditional data management systems - in a smaller footprint and at a greatly reduced cost. This level of performance allows complex queries to be executed that simply are not possible with traditional data warehouse systems - and the system's rapid data loading capability ensures that data is always up to date. The NPS system scales up through 100 terabytes of data, so it can easily meet the ever-growing demand for complex analytics. This large-scale capacity also gives agencies a comprehensive and historical view of all of their data - with users able to query against case-level detail rather than misleading summaries or aggregates, in real-time.

The NPS system provides defence agencies with the power to quickly and thoroughly mine volumes of data for tasks such as:

- Logistics e.g. aircraft maintenance, fleet management, defence programming
- Command and Control (a platform for C4I) e.g. mobility command, exploitation analysis
- Intelligence e.g. signal direction, call detail record analysis, pattern identification

### Hardware Components

- **Host** - The host is a high-performance Linux SMP server that maintains database catalogs, generates optimised query plans, coordinates execution on the snippet processing units (SPUs) and aggregates the query results generated by the SPUs. Applications and end users access the data stored on SPU disks indirectly, by communicating with the NPS host. A second host is available for fully redundant, dual-host configurations.

## BENEFITS

- Linear scalability - from a few hundred GB to hundreds of TB of user data
- Flexibility/agility - easy addition of modular processing units for scalability and reconfiguration
- Structural reliability - full mirroring and ACID transaction management
- Streaming data - joins and aggregations enable maximum query speed - 120 TB/hour
- Fast loading/unloading - up to 500 GB/hour
- Enhanced summarisation/aggregation performance - insert, update, delete up to 100x faster
- Low power requirements (< 30 watts per intelligent storage node) deliver more reliable high performance in an efficient footprint
- Integrated package and compact footprint - e.g. 100 TB of user data in 48 square feet (8 racks)
- Low total cost of ownership - fewer support personnel, fewer software licenses, less hardware
- Reduced system complexity - one appliance versus RDBMS, server, storage area network (SAN), fewer software components, etc.

- **Snippet Processing Unit (SPU)** - The SPU is a query processing node that consists of a disk drive and a special purpose computer for accelerating record management and analysis. Information stored in an NPS system is distributed across many SPUs, so that queries can be processed faster in an MPP environment. The SPU is the smallest replaceable unit of storage and computation.
- **Communication Network** - Host-to-SPU and SPU-to-SPU communication is provided by an internal high-bandwidth network of industry-standard gigabit Ethernet switches.

## Software Functionality

- **High-performance SQL Engine** - optimises SQL on the NPS system's massively parallel streaming architecture
- **Database Operating System** - executes highly efficient SQL through intelligent distribution across host and storage-side processing
- **Fast Loading/Unloading** - allows new data to be added rapidly (up to 500 GB/hour) so that ongoing processing can work with real-time data. Integrates externally with leading ETL and EAI tools. Supports "trickle feed" real-time data loading concurrent with query processing
- **Fast Backup and Restore** - provides high-speed backup and restore at data rates as high as 2 TB/hour (with compression). Shortens backup windows and enables rapid recovery, even for large data warehouses. Interoperable with leading backup products
- **Administration Tools** - includes DBA and Systems Management tools that provide easy-to-use GUIs & CLIs for permissions, monitoring, diagnostics, trouble-shooting and other administrative functions

## Industry Compatibility

- Supports full ANSI transactions
- ANSI SQL-92 compatibility and SQL-99 analytics extensions
- ODBC and JDBC Type 4 API compliant
- AIX, HP-UX, Linux, Solaris and Windows operating systems
- Integrates with leading BI and ETL platforms, e.g. Business Objects, Cognos, SAS

**NETEZZA**  
The Power to Question Everything™

### Contact

EDS Defence, 1-3 Bartley Wood Business Park  
Bartley Way, Hook, Hampshire RG27 9XA  
phone: +44 (0)1256 742000  
fax: +44 (0)1256 742612  
visit: [www.edsdefence.com](http://www.edsdefence.com)  
visit: [www.eds.com](http://www.eds.com)



EXPERTISE. ANSWERS. RESULTS.