IBM PureData System for Analytics N1001

Powered by Netezza technology

Highlights

- Easy to deploy and manage; dramatically simplifies your data warehouse and analytic infrastructure
- Arrives ready to go with expert integration
- Powerful platform for unifying business intelligence and advanced analytics
- Support for thousands of users and complex analytic workloads
- Simplified analytic development with default parallelized analytics and modules; no need for parallel programming
- Powered by Netezza® technology

To gain an edge over the competition, organizations must rely on deep, sophisticated analytics on large volumes of historical data. Yet many companies are challenged by both time-to-value on new analytical capability as well as maintaining service level agreements on existing analytics.

IBM PureData™ System for Analytics is a high-performance, scalable, massively parallel system that enables clients to perform analytics on enormous data volumes. Big data volumes are made simpler, **faster and more accessible**. This system, powered by Netezza technology, is designed specifically for running complex analytics on very large data volumes, orders of magnitude faster than competing solutions. The PureData System for Analytics delivers the proven performance, scalability, intelligence, and simplicity your business needs. It is a low cost option requiring minimal ongoing administration or tuning for a low total cost of ownership (TCO).

The IBM approach to data analysis is patented and proven. Minimize data movement, while processing it at physics speed. Do this in parallel, on a massive scale, inside an easy-to-use data warehouse appliance—extremely fast and at a low cost. And run business intelligence (BI) and advanced analytics that were previously impossible or impractical.

IBM PureData System for Analytics is a purpose-built, standards-based data warehouse and analytic appliance that architecturally integrates database, server, storage and advanced analytic capabilities into a single, easy-to-manage system. Designed for rapid and deep analysis of data volumes scaling into the petabytes, it can help deliver a low cost of ownership in the marketplace.

The analytics opportunity

Deep, sophisticated analytics on volumes of historical data are integral to enterprises in an intelligent economy, giving them an edge over the competition. However, most organizations are challenged by both the time-to-market on new analytical capability as well as maintaining service level agreements on existing analytics. The time has come to shift the focus from managing complexity to simplicity and agility.

Strategic analytics should not be complex to deliver and manage.

The PureData System for Analytics is a scalable, hardware accelerated, massively parallel system that enables clients to gain insight from enormous data volumes, 10 - 100 times faster than they can with traditional custom systems.[1]

Fast. Scalable. Smart. Simple. Completely integrated

PureData System for Analytics is designed specifically for running complex analytics on very large data volumes, orders of magnitude faster than competing solutions. It delivers the proven performance, scalability, intelligence, and simplicity that organizations need to dive deep into their data.

Fast

The PureData System's orders-of-magnitude performance advantage over other analytic options comes from its unique asymmetric massively parallel processing (AMPP™) architecture that combines open, IBM blade servers and disk storage with IBM's patented data filtering using field programmable gate arrays (FPGAs). This combination delivers blistering fast query performance on analytic workloads supporting tens of thousands of BI and data warehouse users, sophisticated analytics at the speed of thought, and petabyte scalability.

Increased performance and concurrency

With the new Directed Data Processing algorithm, the IBM PureData System for Analytics delivers 20 times greater throughput for tactical queries than previous generation Netezza appliances.[2] The system automatically distributes workload to only those processing nodes containing the relevant data, thus enabling greater efficiency of processing for both mixed and pure analytic workloads while maintaining appliance simplicity.

Scalable

With the PureData System, organizations can deploy the right-sized environments for their data volumes and workloads, and be confident that as data volumes grow, larger IBM PureData Systems can be deployed quickly and easily. The PureData System for Analytics starts with a raw data capacity of 8 terabytes for a quarter-rack system, and can grow to well over 300 TB. With built-in hardware compression, usable capacity exceeds 1.2 petabytes for the largest systems.

The PureData System for Analytics provides near linear performance scalability as the size of the appliance grows, which means that organizations can pick the appropriate sized appliance to meet both their data volume and performance SLA needs, all with predictable, scalable performance and no need to add significant additional resources to manage the appliance as data volumes grow.

Smart

PureData System for Analytics dramatically simplifies analytics by consolidating all analytic activity in one place, right where the data resides. Moving analytics to the IBM PureData System is straightforward with IBM's embedded analytic platform. With support for PMML 4.0 models, data modelers and quantitative teams can operate on the data directly inside the appliance instead of having to offload it to a separate infrastructure and deal with the associated data preprocessing, transformation and movement. Data scientists can build their models using all the enterprise data, and then iterate through different models much faster to find the best fit. Once the model is developed, it can be seamlessly executed against the relevant data in the appliance. Prediction and scoring can be done right where the data resides, inline with other processing, on an as-needed basis. Users can get their prediction scores in near real-time, helping operationalize advanced analytics and making it available throughout the enterprise.

IBM Netezza Analytics offers a built-in analytical infrastructure and extensive library of statistical and mathematical functions, supporting a breadth of analytic tools and programming languages. It is delivered with a library of more than 200 prebuilt, scalable in-database analytic functions that execute analytics in parallel while abstracting away the complexity of parallel programming from the developers, users, and DBAs.

The Netezza Analytics functionality also includes in-database geospatial analytics that are compatible with the industry-standard Esri GIS formats. This enables easy integration into existing geospatial analytic environments, and analytics across location and corporate data at the same time and speed.

Simple and completely integrated

All of these new features are delivered with the same simplicity and ease-of-use that distinguish all IBM PureSystems™ family of offerings. Based on proven Netezza technology, this simplicity, and ease of development and deployment is what truly sets the IBM PureData System for Analytics apart. The PureData System delivers high performance automatically, with no indexing or tuning required.

As an appliance, all of the integration of hardware, software and storage is done for you, leading to shorter deployment cycles and industry leading time to value for BI and analytic initiatives. The appliance is delivered ready-to-go for immediate data loading and query execution and integrates with leading ETL, BI and analytic applications through standard ODBC, JDBC and OLE DB interfaces.

Included with every system, Netezza Performance Portal provides a web-based GUI that helps administrators to monitor and manage hardware, administer database objects, configure workload management, view active sessions and monitor system resource utilization for capacity planning. The portal provides a consolidated administrative interface supporting one to many PureData Systems for Analytics from one, easy-to-use access point.

The PureData System for Analytics is architected for high availability from the ground up. All components are internally redundant, and the failure of a processing node (S-Blade) causes no significant performance degradation for a robust, production-ready environment from the moment the appliance is plugged into your data center.

Key Capabilities

Database:
IBM Netezza Platform Software 7 or greater

Supported APIs:
SQL, OLE DB, ODBC 3.5, JDBC 3.0 Type 4

SQL standards:
SQL-92 compliant, with SQL-99 extensions

Programming languages:
Java, Python, R^3, Fortran, C/C++, Perl, Lua

Netezza Analytics foundation
In-Database Analytics, R^3, Matrix, MapReduce, Geospatial Analytics with embedded Esri technology

High-speed load/unload:
Interoperable with ETL and EAI tools at rates up to 5 TB/hour

Backup and restore:
Interoperable with IBM® Tivoli®, EMC Legato and Symantec Netbackup, at rates higher than 4 TB/hour

Database portability:
From IBM DB2®, Informix®, Microsoft SQL Server, MySQL, Oracle, Red Brick, Sybase IQ, Teradata

Operating system:
Red Hat Linux Advanced Server 5

Additional tools:
Windows and web-based DB Admin GUI; CLI and high-speed loading/unloading for AIX®, HP-UX, Linux, Solaris and Windows

IBM InfoSphere® BigInsights™ 1.4 – starter edition – 5 users

IBM InfoSphere Streams 2.0 – developer license – 2 users

The IBM PureData System for Analytics is supported by a wide range of market-leading business partners including: complementary technology partners, resellers, systems integrators, and service providers. For a complete list or to find out if a particular company or solution is part of our program, please contact your IBM representative.

IBM eliminates complexity at every step so you can redirect valuable resources to the initiatives that will actually impact the bottom line.

The best value

IBM PureData System for Analytics is a cost effective analytic option. It requires minimal ongoing administration or tuning, minimizing internal resources as well as implementation costs, for an extremely low total cost of ownership.

The performance and scalability of IBM appliances is available out-of-the-box, without requiring tuning, indexing, aggregations, etc.

IBM offers your company fast time-to-value for important BI and analytic initiatives, which will have a positive impact on your bottom line. With IBM on board, your organization is armed with more accurate intelligence to react quickly and accurately to any opportunities or threats the market may present.

IBM PureData System for Analytics Specifications	**Single rack systems**			**Multiple rack systems**	
IBM PureData System for Analytics N1001	IBM PureData System for Analytics N1001-002	IBM PureData System for Analytics N1001-005	IBM PureData System for Analytics N1001-010	**2 Racks**	**3+ Racks**
Racks	1	1	1	2	3-10
Active S-Blades	4	7	14	28	# Racks x 14
CPU cores	32	56	112	224	# Racks x 112
FPGA cores	32	56	112	224	# Racks x 112
User data in TB (assumes 4X compression)	32	64	128	256	# Racks x 128
Power/rack (Watts maximum/rack)	2,820	3,960	7,635	7,400	7,000
Cooling/rack - BTU/hour	9,600	13,500	26,100	25,500	24,000
Weight/rack kg	453.6	589.7	907.2	907.2	907.2
Height/rack cm	202	202	202	202	202
Depth/rack cm	101.6	101.6	101.6	101.6	101.6
Width/rack cm	64.8	64.8	64.8	64.8	64.8
Power	200-240VAC 50/60Hz, Single Phase 16A	200-240VAC 50/60Hz, Single Phase 24A	200-240VAC 50/60Hz, Single Phase 48A 200-240VAC 50Hz, 3-Phase WYE 24A 200-240VAC 60Hz, 3-Phase Delta 32A	200-240VAC 50/60Hz, Single Phase 48A 200-240VAC 50Hz, 3-Phase WYE 24A 200-240VAC 60Hz, 3-Phase Delta 32A	200-240VAC 50/60Hz, Single Phase 48A 200-240VAC 50Hz, 3-Phase WYE 24A 200-240VAC 60Hz, 3-Phase Delta 32A
Drops/rack	2	2	2	2	2
Safety	US/CSA/EN60950				
Emissions	FCC Part 15, ICES-003, AUS/NZ C-Tick, VCCI and EN55022 Class A; European Immunity: EN55024				

At a time when companies need to be as agile as possible to react to changing market conditions and growing analytic demands, an uncomplicated, easy-to-install system that runs blistering fast and analyzes peta-scale data makes a lot of sense.

How is this possible?
Patented hardware acceleration

IBM PureData System for Analytics adheres to IBM's basic principle of moving processing to the data, not moving the data to the processing. Each PureData System contains multiple Snippet Blades or S-Blades, where SQL query code segments (or "snippets") and complex analytic processes are executed. The S-Blades are intelligent processing nodes that make up the massively parallel processing engine of the appliance. Each S-Blade is an independent server that contains powerful multi-core Intel CPUs, IBM's unique multi-engine FPGAs and gigabytes of RAM as well as dedicated storage devices—all balanced and working concurrently to deliver peak performance.

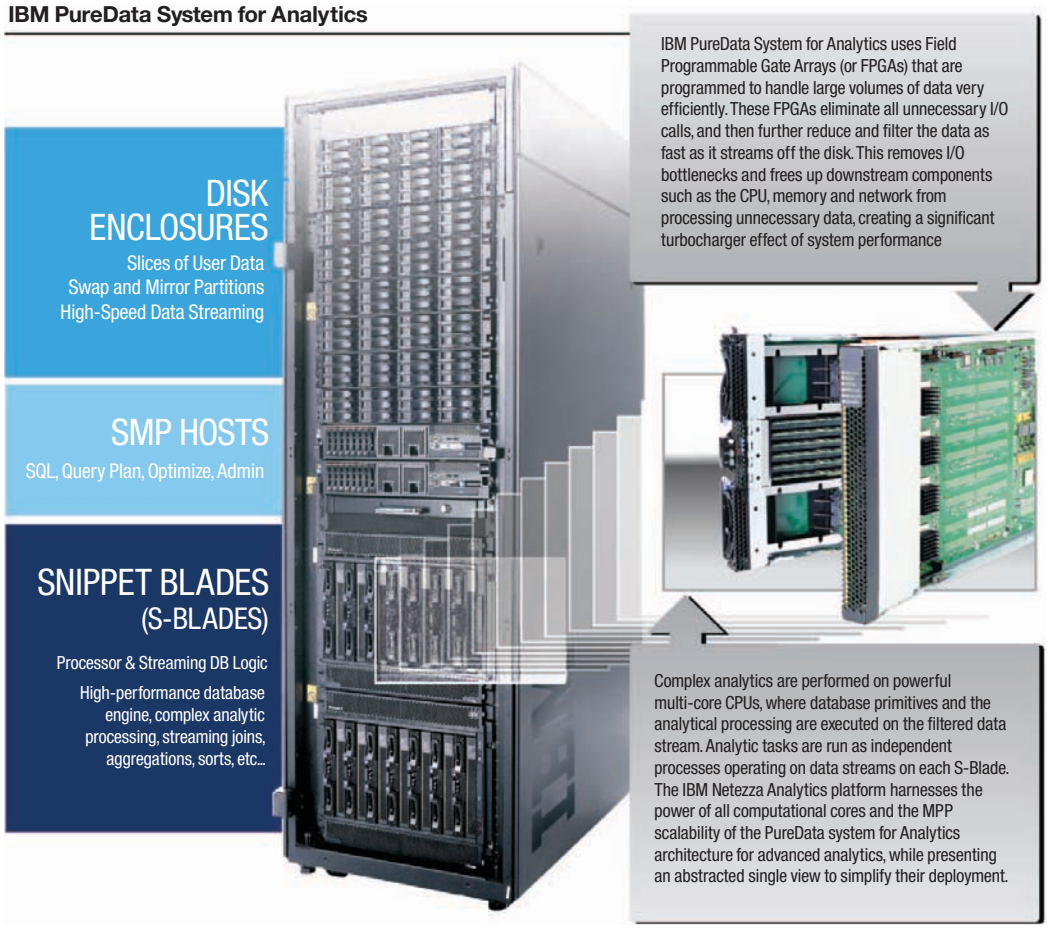

Figure 1. IBM PureData System for Analytics N1001

See for yourself—take a test drive at no charge

Organizations can try out the PureData System for Analytics with the IBM PureExperience Program. This program is available at no charge to allow you to test drive the system with your own data. The program offers onsite installation and demonstration of business value, education and data migration services, use of the system for a specified period and a single line of support. For details on this program and to see what is available in your area, visit: **ibm.com**/PureExperience or contact your IBM representative

About IBM PureData System for Analytics

The IBM PureData System for Analytics, powered by Netezza technology, integrates database, server and storage into a single, easy-to-manage appliance that requires minimal setup and ongoing administration while producing faster and more consistent analytic performance. The IBM PureData System for Analytics simplifies business analytics dramatically by consolidating all analytic activity in the appliance, right where the data resides, for industry-leading performance. Visit: **ibm.com**/PureData to see how our family of expert integrated systems eliminates complexity at every step and helps you drive true business value for your organization

About IBM Data Warehousing and Analytics Solutions

IBM provides the broadest and most comprehensive portfolio of data warehousing, information management and business analytic software, hardware and solutions to help clients maximize the value of their information assets and discover new insights to make better and faster decisions and optimize their business outcomes.

For more information

Help IT make the shift to the strategic center of your business, and leverage proven expertise to take the lead. To learn more about the PureSystems family and the PureData System for Analytics, contact your IBM representative or IBM Business Partner, or visit the following website: **ibm.com**/PureSystems/PureData/

Additionally, IBM Global Financing can help you acquire the software capabilities that your business needs in the most cost-effective and strategic way possible. We'll partner with credit-qualified clients to customize a financing solution to suit your business and development goals, enable effective cash management, and improve your total cost of ownership. Fund your critical IT investment and propel your business forward with IBM Global Financing. For more information, visit: **ibm.com**/financing

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
September 2012

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

It is the user's responsibility to evaluate and verify the operation of any other products or programs with IBM products and programs.

Actual available storage capacity may be reported for both uncompressed and compressed data and will vary and may be less than stated.

Please Recycle

[1] Based on IBM customers' reported results. "Traditional custom systems" refers to systems that are not professionally pre-built, pre-tested and optimized. Individual results may vary.

[2] Based on IBM internal performance benchmarking of previous version to current version.

[3] With Revolution Analytics R Enterprise for IBM PureData System for Analytics

IMD14400-USEN-01