IBM PureData System for Analytics N3001-001: the Mini Appliance

Powered by Netezza technology

Highlights

- Designed to deliver fast performance for complex analytics in a powerful and cost-effective solution
- Easy to deploy as a rack-mountable appliance, which simplifies your data warehouse and analytic infrastructure
- Arrives ready to go with IBM Fluid Query, plus data integration, business intelligence and Hadoop starter kits
- Utilizes self-encrypting drives to ensure protection of data at rest
- Powered by IBM Netezza technology

The IBM® PureData™ System for Analytics N3001-001 brings the power and simplicity of Netezza technology to small and midsize organizations for big outcomes. Available as a rack mountable appliance, the N3001-001 is a high-performance, massively parallel processing system that helps you gain insight into your growing data volumes. It is a patented solution built to simplify operations and increase the performance of business analytics. The system is designed specifically to run complex analytics on terabyte data volumes, orders-of- magnitude faster than traditional custom systems.[1]

The IBM PureData System for Analytics N3001-001, along with the entire IBM PureData System for Analytics N3001 appliance family, promotes high performance based on its unique asymmetric massively parallel processing architecture that combines with IBM Netezza data filtering. This combination delivers fast query performance on analytic workloads in close support of your business intelligence and data warehouse users. This appliance family requires minimal ongoing administration or tuning, while allowing customers to realize a low total cost of ownership (TCO).

Best practice: Routing the query to the data

IBM Fluid Query is the capability that unifies data access across the logical data warehouse. Users and analytic applications need access to data in a variety of data repositories and platforms without concern for the data's location or access method or the need to rewrite a query. IBM Fluid Query is the capability for a data store to route a query (or even part of a query) to the correct data store within the logical data warehouse so that the query can flow to the data, not the data flow to the query.

IBM Fluid Query 1.0, included with IBM PureData System for Analytics, provides access to data in Hadoop from IBM PureData System for Analytics appliances. IBM Fluid Query 1.0 enables the fast movement of data between Hadoop and IBM PureData System for Analytics appliances. Enabling query and data movement, IBM Fluid Query 1.0 connects those appliances to common Hadoop systems: IBM BigInsights for Apache™ Hadoop®, Cloudera, and Hortonworks. IBM Fluid Query 1.0 allows queries against PureData System for Analytics, Hadoop or both by merging results from PureData System for Analytics database tables and Hadoop data sources thus creating powerful analytic combinations.

IBM PureData System for Analytics N3001-001 system will:

- Deliver an agile platform for responding to the challenges of rapid deployment of business intelligence and advanced analytics.
- Provide high performance for complex analytic and business intelligence applications at a lower total cost of ownership.
- Offer protection of all data from unauthorized access with self-encrypting drives.
- Include a flexible analytic environment that supports multiple languages (C/C++, Java™, Python, Perl, Lua, Fortran), frameworks (MapReduce), and tools (Open R, SAS, PMML, IBM SPSS).
- Simplify analytic development with ready-to-use, fully parallelized execution for SAS, SPSS, and Open R.
- Deliver more than 200 prebuilt, scalable, in-database analytic functions.
- Be easy to install and manage, in a rack mountable solution for your data warehouse and analytics infrastructure.
- Leverage most market-leading analytics and business intelligence tools, applications, and infrastructures.
- Support industry-standard interfaces (SQL, ODBC, JDBC, OLE DB).
- Deliver enterprise-class reliability and high availability.
- Require low power and cooling in a compact footprint.
- Use these included Software Entitlements:
  - IBM Cognos Business Intelligence — 5 Analytics User licenses, 1 Analytics Administrator license
  - IBM DataStage (280 PVUs) — 2 concurrent Designer Client licenses and IBM InfoSphere Data Click (with PureData System for Analytics as a source or target).
  - IBM BigInsights for Apache™ Hadoop® software licenses to manage around 100 TB of Hadoop data[2].
  - 2 non-production user licenses for IBM InfoSphere Streams Developer Edition.

Features are delivered with the same simplicity and ease-of-use that have always distinguished the IBM PureData System for Analytics product family. Based on Netezza technology, the IBM PureData System for Analytics is designed to deliver simplicity, ease-of-development and rapid deployment, and delivers high performance automatically with no data modeling, indexing or tuning required.

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

IBM PureData System for Analytics is architected for high availability from the ground up. All components are internally redundant for a robust, production-ready environment from the moment the appliance is installed, rack mountable, into your data center. The N3001-001 is different from the N3001 appliance family, in using software and server processing to emulate Field Programmable Gate Array (FPGA) data filtering functionality.

Software specifications

Database
IBM Netezza Platform Software v7.2.0.1 or greater

Operating system
Red Hat Enterprise Linux Server 6.5

IBM Fluid Query 1.0
Run PureData System for Analytics queries against Hadoop data. Move PureData System for Analytics data quickly to Hadoop file systems, and quickly move Hadoop data to PureData System for Analytics.

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, Fortran, C/C++, Perl, Lua

Netezza Analytics foundation
In-database analytics, R[3], Open Source R Matrix, MapReduce, Geospatial Analytics with embedded Esri technology

High-speed load/unload
Interoperable with ETL and EAI tools at rates of 1 TB/hour

Backup and restore
Interoperable with IBM Tivoli®, EMC Legato and Symantec Netbackup, with parallel, multi-thread exaction.

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

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

Included software entitlements

Business intelligence
IBM Cognos® Business Intelligence[4], 5 analytics user licenses, one analytics administrator license where the IBM PureData System for Analytics N3001 must be the data source for Cognos

Data integration
IBM DataStage[4] 280 PVUs—2 concurrent Designer Client licenses and InfoSphere Data Click (with IBM PureData as a source or target).

Hadoop data services
IBM® BigInsights for Apache™ Hadoop® [4] (five virtual servers to manage ~100 TB of Hadoop data; IBM PureData System for Analytics appliances must be a source or target 2)

Real-time analytics
IBM InfoSphere Streams Developer Edition[4]—two users, non-production licenses

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.

Please see the IBM PureData System for Analytics N3001 data sheet to learn more about our other sizes available in the N3001 family.

Specifications	Rack mountable system
IBM PureData System for Analytic	IBM PureData System for Analytics N3001-001
Racks	N/A Rack Mountable 2, 2U slots
Active S-Blades	N/A
CPU cores	40
FPGA cores	N/A
User data in TB (assumes 4X compression)	16
Power (Watts maximum)	1,800
Cooling - BTU/hour	6,140
Weight kg	60 (30 kg/server) Minimum confirmation: 25 kg (55 lb)/server Maximum configuration: 30 kg (65 lb)/server
Height mm	173 (86.5 mm/server) (6.8 in)
Depth mm	746 (29.4 in)
Width mm	445 (17.5 in)
Power	100-127 VAC/220-240 VAC, 50Hz/60 Hz (Single Phase), 10A/5A per drop
Drops	4
Safety	US/CSA/EN60950-1
Emissions	FCC Part 15 Class A, ICES-003, AUS/NZ CISPR 22 Class A and EN55022 Class A; European Immunity: EN55024

About IBM PureData System for Analytics

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

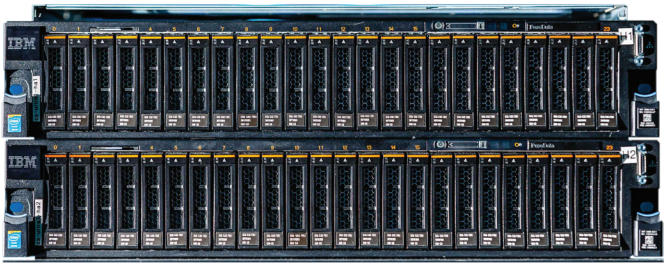

Figure 1: PureData System for Analytics N3001-001, the mini appliance [5]

About IBM Data Warehousing and Analytics Solutions

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

Why IBM

Logical Data Warehouse. IBM is a strategic advisor that can help clients:

- Innovate by effectively addressing a broad and ever-evolving set of data management needs. Through our strategic acquisition strategy and organic development born of IBM Research, we have amassed true best-in-class data warehousing solutions that address virtually every information need.
- Boost the success of their data warehouse initiatives. IBM combines proven innovations, a sharp focus on integration and world-renowned industry experts to help ensure the success of data warehouse initiatives. Plus, IBM is an industry leader in providing a comprehensive solution portfolio that targets key aspects of data warehousing—including data integration, governance and security.
- Modernize their data warehouse to fuel real-time business. IBM's commitment to innovation focuses on designing the right mix of data platforms and integration capabilities for our clients' changing business requirements.

For more information

Help IT make the shift to the strategic center of your business. Leverage proven expertise to take the lead. To learn more about IBM PureSystems 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
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
April 2015

IBM Business Partners set their own prices. Business Partner prices may vary. All client examples cited or described are presented as illustrations of the manner in which some clients have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

1 Based on reported results from IBM customers, "traditional custom systems" refers to systems that are not professionally prebuilt, pretested, and optimized. Individual results may vary.

2 Based on 4 data nodes + 1 master node. 12 TB uncompressed per data node with 4 TB drives. 12 TB x 4 nodes = 48 TB uncompressed. Using 2-2.5x compression yields 96-120 TB compressed data. Capacity will depend on hardware configuration selected.

3 IBM PureData System for Analytics N3001-001 maintains continued compatibility with Revolution R Enterprise, delivering 64-bit processing with high performance, enterprise readiness, and support for the R programming language. Revolution R Enterprise for IBM PureData System for Analytics is available for additional purchase from Revolution Analytics.

IBM PureData System for Analytics N3001-001 supports both open source R and Revolution R Enterprise. Open source R is available from IBM developerworks: **ibm.com**/developerworks.

Revolution R Enterprise for IBM PureData System for Analytics is available for additional purchase from Revolution Analytics.

4 Please refer to IBM for the specific software version of the entitled products included. ibm.biz/N3001_license.

5 The Model N3001-001 requires the customer to supply the following data center equipment to support the appliance:

- Rack with power distribution units (PDUs)
- Keyboard, video, and mouse console (KVM) for user access to host console
- 1 Gbps switch for management (two ports required)

Please Recycle

WAD12370-USEN-02