IBM Netezza Analytics: The Simple Appliance for Serious Analytics

By making more sophisticated questions possible, IBM Netezza Analytics delivers better-targeted answers

Highlights:

- **Serious Analytics** – Answer questions that were previously too complex, required too much data and took too much time to analyze

- **Data Exploration** – Discover subtle patterns within ever-growing data volumes to answer interrelated and complex business questions

- **Analytics On-Demand** – Quickly respond to dynamic business conditions to choose the best course of action

- **Simple to Use** – Build models and score data using existing analytics technologies and skill sets, including out-of-the-box functions, popular analytic packages and languages

- **High-Performance** – In-database, parallelized algorithms execute quickly against an organization's largest data sets and take full advantage of IBM Netezza's data warehouse Asymmetric Massively Parallel Processing™ (AMPP™) architecture

Today enterprises are confronted with growing data volumes and increasing demands for speedy answers to increasingly complex and interrelated business questions and issues. Advanced analytics helps businesses get the answers they need to make better decisions and stay competitive. IBM® Netezza® Analytics is a purpose-built advanced analytics platform that empowers analytic enterprises to meet and exceed those business demands.

IBM Netezza Analytics' advanced technology fuses data warehousing and in-database analytics into a scalable, high-performance, massively parallel analytic platform that is designed to crunch through petascale data volumes. This lets a large number of users ask questions of the data that could not have been contemplated on other architectures. IBM Netezza Analytics is designed to quickly and effectively provide better-targeted answers to the most sophisticated business questions.

IBM Netezza Analytics is IBM Netezza's most powerful advanced analytics platform that provides the technology infrastructure to support enterprise deployment of in-database analytics. The analytics platform integrates a robust set of built-in analytics with leading analytic tools such as R and Hadoop™, with IBM Netezza's core data warehouse appliances. IBM Netezza pioneered the modern data warehouse appliance and has customers worldwide that have realized the value of combining data warehousing and analytics into a single, high-performance integrated system. IBM Netezza Analytics allows analytic enterprises to realize significant business value from new business models and helps companies realize both top-line revenue growth and bottom-line cost savings.

IBM Netezza offers a distinctive and simple-to-use approach for serious analytics. Traditionally, analytics are built and deployed on separate analytic servers. This elongates the end-to-end time from model inception to deployment. This also requires data movement, from a data

warehouse or other data sources to the analytic server. In addition to this process taking too much time, it is inefficient, limits the data used, constrains the scope of the analytic modeling, and impedes the ability to experiment iteratively.

With IBM® Netezza® Analytics, analytic models can be built and deployed right where the data resides – in the data warehouse. By doing this, the time it takes to build and deploy analytic models throughout an enterprise can be significantly reduced. By shrinking the time from model inception to model deployment, companies can move to enterprise-wide fact-based decisions by infusing more of their decisions with insightful on-demand analytics.

Every IBM Netezza data warehouse is delivered with a library of in-database analytic functions that can be accessed via any SQL compliant interface. Additionally, customers can develop new capabilities using the platform's user-defined extensions. It is the easiest, most extensible platform that supports multiple tools, languages, and frameworks.

Use IBM Netezza Analytics for:

- Data exploration and discovery
- Data transformation
- Model building
- Model diagnostics
- Model scoring

The IBM Netezza data warehouse appliance – a powerful parallel computing platform – is fully exploited by IBM Netezza Analytics to deliver high-speed, scalable analytics processing. The appliance uses the high-speed throughput of the Asymmetric Massively Parallel Processing™ (AMPP™) architecture to maximize speed and efficiency for the in-database analytics processing. The AMPP architecture is a blade-based streaming architecture that uses commodity blades and storage, combined with IBM Netezza's patented data filtering using Field Programmable Gate Arrays (FPGAs), to deliver large data, high speed analytics. IBM Netezza has consolidated all analytics activity in a powerful and simple appliance.
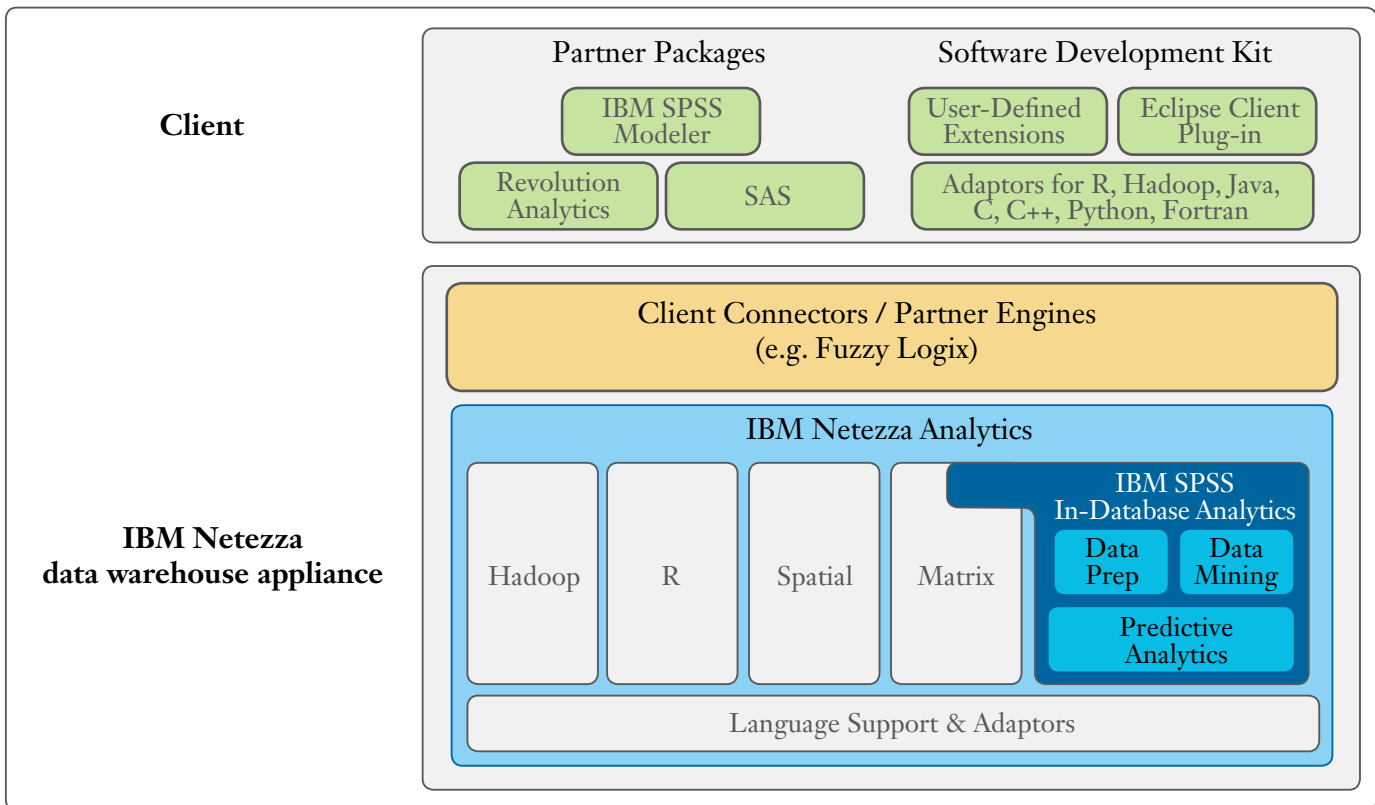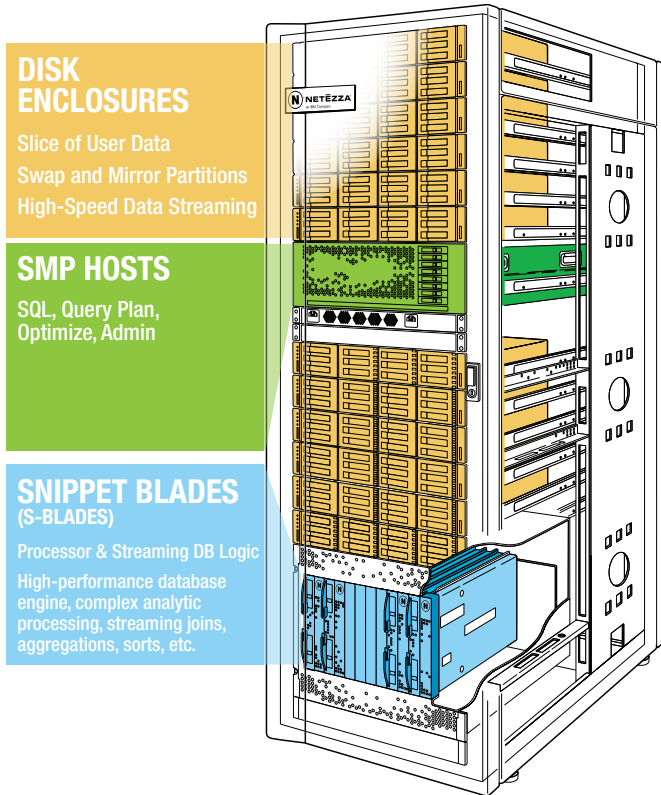

Figure 1: IBM Netezza Analytics Architecture

> *"Top-performing organizations use analytics five times more than lower performers."*

DISK ENCLOSURES

Slice of User Data
Swap and Mirror Partitions
High-Speed Data Streaming

SMP HOSTS

SQL, Query Plan,
Optimize, Admin

SNIPPET BLADES
(S-BLADES)

Processor & Streaming DB Logic

High-performance database engine, complex analytic processing, streaming joins, aggregations, sorts, etc.

IBM® Netezza® Analytics is purpose-built to simplify the building and deploying of models for analytic enterprises that demand the highest performance on large, complex volumes of data.

Serious Analytics

Businesses are collecting and tracking information more than ever before, and are under increasing pressure to operate more efficiently and effectively. The ability to analyze data, foresee outcomes and find ways to improve business is driving companies to fully exploit advanced analytics. Making sense of massive data volumes and turning it into meaningful results can be daunting or even technically unfeasible in companies with traditional database technology. These systems are easily overextended just keeping up with the growth in user and data volumes.

Analytics that once seemed impossible or impractical to run are now possible with IBM Netezza Analytics. With IBM Netezza's simple appliance approach, all of an organization's data can be used to generate a finer set of results, helping to drive new revenue opportunities and gaining a competitive advantage. By using advanced analytics on IBM Netezza data warehouse appliances, the entire organization can realize value – from financial teams, to lines-of-business, to sales, to IT, to the executive office. This offers greater clarity for the entire business, and ensures everyone is leveraging the same data, using all available data.

By using IBM Netezza Analytics, organizations no longer have to make a choice between large data volumes and serious analytics.

Data Exploration

With IBM Netezza Analytics, it is possible to move beyond traditional business intelligence (BI) and ad hoc reporting that is based on historical data, and discover new ways to create value from data. Organizations now have the ability to build and deploy sophisticated analytic models that mirror real-world complexities more easily and effectively. They can continuously experiment, improve and tune analytic models to discover trends and find ways to lower business risk, reduce cost, increase revenues, and make fact-based decisions.

Investigate new ways to create value from your data while doing concurrent, parallel model experimentation with IBM Netezza Analytics. Now, all data can be used instead of just samples or aggregates, thereby improving accuracy and enabling more targeted decisions.

Telecommunications Case Study

In the highly competitive world of telecommunications, the players need accurate and up-to-date information to manage their businesses, exploit new opportunities and keep customer churn to a minimum. The revenue assurance department uses analytics to identify revenue leakage and plug any gaps in the revenue chain.

Product marketing and pricing teams can proactively plan and forecast the effect of telephony tariff changes, and plan how to best react to competitors' offerings. The credit services department uses analytics to spot high telephony usage and proactively manage any potential credit problems before they arise. IBM Netezza data warehouse appliances crunch through large data volumes and process in-database analytics to help telecommunications companies be competitive and streamline processes.

Health Care Case Study

A health care provider was interested in predicting who was at risk for diabetes. By taking a look beyond typical health parameters such as weight and family history, and adding more attributes to their models such as financial background, the provider uncovered that a person's financial situation does indeed impact their diabetes risk. By refining their analytic models, this health care provider was able to determine not only if a person may develop diabetes at some point in their lifetime, but also predict when the onset of the disease would take place (ex., in one year, three years, etc.).

By identifying these trends, outreach and preventive care could be provided to these patients at risk. This health care provider continues to refine and create models to uncover additional trends and improve patient care leveraging IBM Netezza Analytics.

Analytics On-demand

Organizations can get ahead of the competition by using IBM® Netezza® to predict, forecast and optimize business elements. Historically, the analytic process can be expensive and time-consuming – it generally takes weeks to develop a predictive model from the data in a data warehouse. Once the model is developed, it still takes hours – or even days in some cases – to execute on all the data, despite adding expensive hardware to the problem. The issue is further exacerbated with a growth in data volumes.

IBM Netezza offers companies fast time-to-value for important predictive initiatives, resulting in a positive impact on the bottom line and in top-line growth. With IBM Netezza on board, organizations are armed with the most accurate intelligence to react more quickly and accurately to any opportunities or threats the market may present. Models can be quickly deployed, tweaked as needed, and dispensed concurrently in multiples, while taking advantage of IBM Netezza's parallel in-database technology. This enables the largest data volumes to be handled quickly and efficiently.

At a time when companies need to be as agile as possible to react to changing market conditions and demands, an easy-to-use system that runs blisteringly fast and analyzes petascale data makes a lot of sense.

Simple to Use

IBM Netezza's data warehouse appliance is easy-to-use and dramatically accelerates the entire analytic process. The programming interfaces and parallelization options make it straightforward to move a majority of analytics inside the appliance, regardless of whether they are being performed using tools such as SAS and R, or written in languages such as Java, Python or Fortran. Additionally, IBM Netezza data warehouse appliances are delivered with a built-in library of parallelized analytic functions, purpose built for large data volumes, to kick-start and accelerate any analytic application development and deployment.

The simplicity and ease of development is what truly sets IBM Netezza apart. It is the first appliance of its kind – packing the power and scalability of hundreds of processing cores in an architecture ideally suited for parallel analytics. Instead of a fragmented analytics infrastructure with multiple systems where data is replicated, IBM Netezza Analytics consolidates all analytics activity in a powerful appliance. It is easy to deploy and requires minimal ongoing administration, for an overall low total cost of ownership.

Simplifying the process of exploring, calculating, modeling and scoring data are key drivers for successful adoption of analytics companywide. With IBM Netezza, any business user can run their own analytics in near real-time, which helps analytics-backed, data-driven decisions to become pervasive throughout an enterprise.

Focus on the business, not the process. Let your IBM Netezza data warehouse appliance do the heavy-lifting for you.

Retail Case Study

A global leader in behavior-based marketing solutions, providing manufacturers and retailers with the ability to execute targeted marketing programs on the fly based on real-time market basket analysis and historical purchasing patterns, has seen great results from introducing IBM Netezza into its business. For example, in one campaign relying on POS information to issue individualized coupons to customers, coupon redemption increased by 30 percent using IBM Netezza.

By leveraging the simplicity of IBM Netezza, this company has reduced the number of DBAs required to maintain its DW and has realized a significant increase in productivity and its IT is moving 5-10x faster than ever before. In addition, the company's primary database storage space has decreased by almost 80TB since migrating to IBM Netezza, due to the elimination of all aggregate tables and indices. This reduction in storage also decreased the corresponding data center footprint.

Digital Media Case Study

A major digital media company can provide right-time data, drive insightful decision-making, correctly measure marketplace dynamics and effectively bridge the gap between retailers and manufacturers by leveraging IBM Netezza Analytics and IBM Netezza data warehouse appliances. The company can accommodate more customers now than ever before, and these customers can run custom-defined, ad hoc market analyses, whereas before they were limited to static views of the market. Also, more data history can be retrieved and analyzed than ever before and new data can be made available in near real-time. Analyses are unconstrained and have greater functionality and flexibility.

IBM Netezza has helped reduce IT costs, thereby creating a more effective business model and with its advanced analytic capabilities, the company has a competitive advantage over its rivals.

High-performance

IBM® Netezza® has created an extremely flexible analytic platform that offers high performance at petascale. By bringing analytics to the data, modellers and quant teams can operate on the data directly inside the appliance, instead of having to move it to a different location and deal with the associated data pre-processing and transformation.

Analysts and modelers can take full advantage of the AMPP™ architecture of IBM Netezza Analytics to ask the most complex questions on all the enterprise data, without the infrastructure getting in the way. Practitioners can iterate through different models more quickly to experiment and find the best fit.

Once the model is developed, it can be seamlessly executed against all the relevant data in the enterprise. The prediction and scoring can be done right where the data resides. Users can get the results of prediction scores in near real-time, helping operationalize advanced analytics and making it available throughout the enterprise.

IBM Netezza data warehouse appliances use Field Programmable Gate Arrays™ (FPGAs) which have been programmed by IBM Netezza to handle large volumes of data very efficiently. These FPGAs filter out extraneous data as fast as it streams off the disk. This removes I/O bottlenecks and frees-up downstream components such as the CPU, memory and network from processing unnecessary data, creating a significant turbocharger effect on system performance.

The serious big math is performed in powerful multi-core CPUs, where database primitives and complex analytics are executed on the filtered data stream. Analytic tasks are run as independent processes operating on data streams across each S-Blade™. IBM Netezza's parallel analytic engine harnesses the power of all the computational cores in the appliance to offer significant performance and scalability for serious analytics.

Eliminate technology hurdles by leveraging IBM Netezza's appliance to make your life simpler.

With IBM Netezza Analytics, you will have an appliance that can manage all of your analytic queries on massive data volumes, all while taking advantage of IBM Netezza's data warehouse appliance parallel processing platform for better performance. IBM Netezza provides you with the simple appliance for serious analytics.

Financial Services Case Study

A financial institution needed to calculate value-at-risk for an equity options desk. The IBM Netezza platform was able to run a Monte Carlo simulation on 200,000 positions with 1,000 underlying stocks (2.5 billion simulations) in less than three minutes. Leveraging an in-database analytics approach allowed the financial institution to use the data where it resided as opposed to building a parallel data-processing platform solely for performing the Monte Carlo simulation.

The combination of execution time and the elimination of latency required to move data between two platforms allowed the financial institution to include more variables in assessing the risk of investment strategies and to perform this assessment with greater frequency.

IBM® Netezza® Analytics Platform

Core Modules:

IBM® SPSS® In-Database Analytics	Execute data preparation, data mining, predictive analytics
R	Run parallel, in-database model building and scoring
Matrix	Leverage parallelized, linear algebra package
Hadoop™	Execute MapReduce™ functions in-database
Spatial	Use Geospatial data types and functions

Software Development Kit (SDK) includes:

Language Adaptors	Develop with R, Hadoop, Java™, C, C++, Python™, Fortran
Plug-in for Eclipse	Build in-database analytics with easy-to-use, standard integrated development environment
User-Defined Extensions	Create custom user-defined functions (UDF), user-defined aggregates (UDA), user-defined table functions (UDTF), user-defined analytic processes (UDAP)

Third-Party Applications

Data Integration	Ab Initio®, BusinessObjects™/SAP®, Composite® Software, DataFlux™ – a SAS® Company, Expressor® Software, GoldenGate® Software, Informatica®, IBM® Information Server, Oracle® Sunopsis, WisdomForce™
Data Analysis	IBM SPSS, Revolution® Analytics, BusinessObjects/SAP, Kalido®, KXEN®, Quest® Software, SAS
BI/Reporting	IBM® Cognos®, IBM® Unica®, Acutate™, BusinessObjects/SAP, Information Builders, MicroStrategy®, Oracle, QlikTech®
Data Visualization	TIBCO® Spotfire®, BIS(2)™
In-Database Analytics	SAS, Fuzzy Logix™ / DB Lytix™
Business Continuity/Compliance	IBM DataMirror, IBM® Tivoli® Storage Manager®, EMC®, Symantec™ Veritas™

Analytic Functions
Built-in analytics library delivered with every IBM Netezza data warehouse appliance (sample)

DATA PREPARATION, VALIDATION AND MANIPULATION

Stored procedures

Tables and views

DATA MINING

Data assessment

Distributions

Information content

Data transformation

Sampling, training/testing sets

Discretization

Principal components

Model building, interpretation, and application

Itemsets (basis for Associations)

Clustering

Classification

Regression

Deployment (PMML export)

STATISTICS

ANOVA

Correlation, covariance, dispersion

Distributions, density, probability

Missing values

Moments, central tendency

Normalization, standardization

Quantiles

Significance

MATRIX FUNCTIONS

General

Matrix algebra, operations, and description

Matrix-table operations

Matrix creation and transformations

Matrix statistics

SPATIAL FUNCTIONS

General

Geometries, single

Geometries, multiple

About IBM® Netezza®

IBM Netezza pioneered the data warehouse appliance space by integrating database, server and storage into a single, easy to manage appliance that requires minimal set-up and ongoing administration while producing faster and more consistent analytic performance. The IBM Netezza family of data warehouse appliances simplifies business analytics dramatically by consolidating all analytic activity in the appliance, right where the data resides, for blisteringly fast performance. Visit netezza.com to see how our family of data warehouse appliances eliminate complexity at every step and lets you drive true business value for your organization. For the latest data warehouse and advanced analytics blogs, videos and more, please visit: thinking.netezza.com.

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 customers maximize the value of their information assets and discover new insights to make better and faster decisions and optimize their business outcomes.

IBM

NETEZZA
an IBM® Company