



WHITE PAPER

The Green Revolution in Data Warehousing:

The Power, Cooling and Footprint Advantages of the Netezza Data Warehouse and Analytics Appliance

Introduction

Skyrocketing power costs and finite power availability are forcing many companies to re-examine their IT strategies and budgets. With their intense workload — dozens or hundreds of processors and large numbers of spinning disk drives straining over massive queries — data warehouse systems based on general-purpose designs have long been one of the major consumers of power. As data volumes continue to grow and complexity of the enterprise analytic environment increases, companies will continue striving for a competitive edge through better-performing data warehousing and analytic infrastructures. The Netezza data warehouse and analytic appliance redefines energy efficiency in data warehousing, while providing cost-effective query performance beyond the reach of other systems.

This paper examines the power crisis in the data center, and the environmental impact of general-purpose processor technology (web servers, email servers, OLTP servers, etc.). Today's tightly packed servers consume more power and give off more heat than their predecessors as the latest generations of general purpose computers struggle to keep up with the demand for more compute power and more data.

The paper then examines the Netezza appliance from an energy efficiency perspective. Netezza's family of data warehouse and analytic appliances uses a fundamentally different processing architecture, based on streaming technology originally developed for applications such as video games, industrial process controls and medical imaging that also happens to be perfect for accelerating query performance of massive databases. The Netezza architecture greatly reduces processor workloads, allowing the Netezza system to outperform our competitors at a fraction of their power consumption, heat and footprint. Using the newest generation of more energy-efficient chips and storage combined with Netezza's expertise in developing balanced systems the Netezza appliance is without equal in energy and space efficiency.

To show how technology differences translate into power ratings, the paper concludes with an analysis comparing the energy efficiency of the Netezza appliance with data warehouse systems of equivalent storage capacity. Values are taken from the American Power Conversion Corp. (APC) on-line UPS selector that provides power and cooling ratings for all major competitive configurations. The results are clear: Netezza's analytic appliance is setting the pace in energy efficiency as well as performance.

Power Crisis in the Enterprise Data Center

According to Moore's Law, the processing power of computer chips doubles about every 24 months. This pattern of innovation has been driving information technology forward for decades. Unfortunately, this progress has a downside, one that was less apparent when energy was cheap and components drew much less power than they do today. As processor performance has increased, so too has power consumption and its side effect, heat. Smaller servers packed with more-powerful components not only consume more electricity in their processing operations, they also generate more heat, requiring aggressive use of internal fans and air-conditioning that add a huge overhead to power consumption. According to Gartner, "During the next three years, most CIOs will experience constraints in data center floor space and power that could limit an IT organization's ability to grow as the business grows."¹

As processing technology becomes smaller, faster and hotter, corporate data centers are consuming more electricity than ever before. Coupled with rising energy costs and a finite power grid, the result is an energy crisis affecting both individual companies as well as the power utilities that supply them. Recent industry studies have examined the surging demand for power from various perspectives, and the figures are startling:

- Space that's leased for \$12 to \$20 per square foot can cost \$60 per square foot to cool.²
- At the Gartner Data Center Conference in December 2008, an attendee poll showed the greatest facility problem with data centers continues to be insufficient power and cooling. Of those polled at the session, 96% said they will expand/upgrade, relocate or renovate their facility to accommodate power and cooling needs over the next two years.³
- Data center professionals have to deal with many challenges, including power and cooling issues, and rapidly evolving technologies, all on a tighter budget than in previous years.⁴
- A soaring amount of data, combined with the necessity for quick access to stored data for legal and business purposes, highlights the need for efficient data storage techniques and technologies.⁴
- The server installed base has been growing at an approximately 11% to 15% compound annual growth rate during the past four years, and this rate will only accelerate.⁵
- U.S. data centers are consuming 1.5% of total electricity consumption, a figure predicted to double in the next five years.⁶
- Survey results suggest that most issues facing European and North American data center facilities are the same, although some aspects of energy efficiency and use may be bigger factors in Europe.⁷

Increasingly, power utilities are informing corporate customers that they can no longer provide the additional electricity needed to run their latest processing-intensive applications, a development that would have been unheard of only a few years ago. Companies then find themselves at a power crossroads: either scale back growth plans, or pay the astronomical cost of building an additional facility or relocating the data center to a region where power is less restricted.

Wasted Real Estate

There's another, rather ironic, dimension to the power crunch: inefficient use of data center space. As phenomenal processing power in smaller form factors outpaces the ability of server vendors to keep their equipment cool, workarounds are needed that leave extra space for air to circulate around miniature components. Rather than filling racks to capacity, the customer is required to skip rows or slots to allow heat to dissipate, giving up valuable real estate on the data center floor. Figure 1 illustrates the relationship between available floor space and rack density.

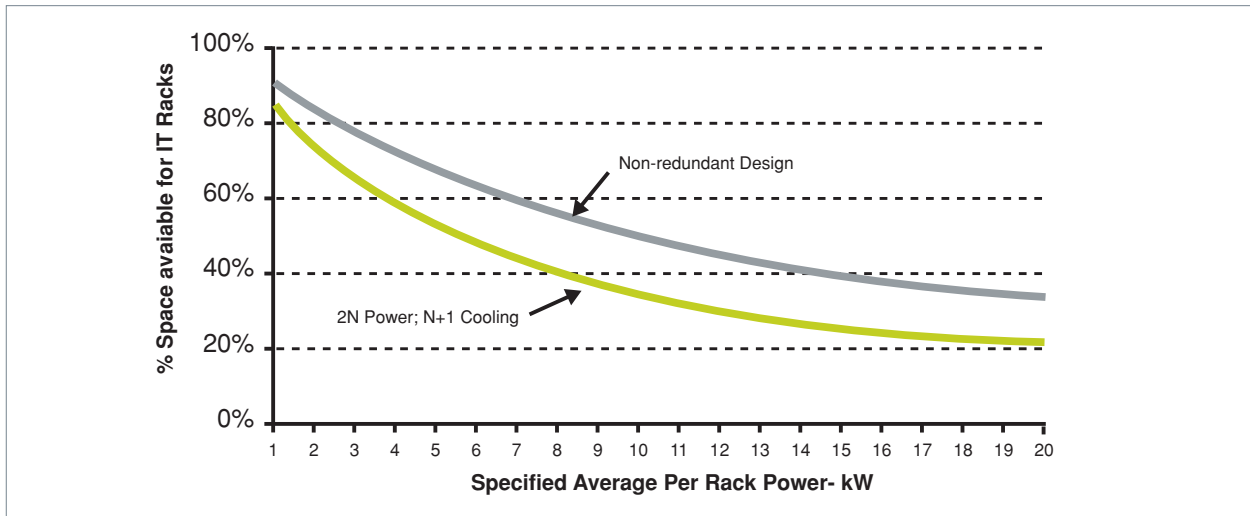


Figure 1. Effect of average rack density specification on the fraction of available space for IT racks

Hardware manufacturers use complex configuration tools that limit the number of blades that can be installed in a rack, and reserve much-needed space for cooling. Most hardware manufacturers also look for new and innovative technologies to cool their over-heated systems, but these new products cost more money, require additional power and, naturally, occupy more space. In a Gartner survey published in February 2007, 35% of respondents expect their data center racks to consume an average of 7 to 10 kWh each in the next two years, and 23% expect even higher.⁹

Green and Greener

“As IT and Facilities budgets receive increased scrutiny, efficiency becomes a more business critical issue for data center managers globally,” said Tom Brey, a director of The Green Grid and senior technical staff member in IBM's systems and technology group at the Green Grid's 2ND Annual Technical Forum.

As power consumption becomes increasingly problematic, the industry is responding from various directions. A quick search of the term returns reams of information about ventilation systems, power management software and consulting services aimed at making the data center more energy-efficient. Chip manufacturers are announcing high-performance, low-power chips that deliver twice the performance but draw 30% less power. But with ever increasing demand for larger systems the net result is that data warehouse systems relying on this technology still tend to be extremely power-hungry with individual processing nodes requiring nearly 100 Watts of power. The high power requirements of these systems are further extended when accounting for their high-performance RAM memories and separate, multi-terabyte disk storage arrays. Over the next few years, the IT industry will be battling to bring power consumption under control. But for data warehousing, much greater energy efficiency, along with system performance, has already arrived. Netezza's balanced architecture makes the difference.

The Netezza Alternative: Balanced Design, High Performance

As members of the Green Grid, Netezza and our suppliers have worked to create an appliance that brings new levels of power and space efficiency to the computer room. Netezza's patented Asymmetric Massively Parallel Processing™ (AMPP™) architecture takes a different approach to processing massive queries than architectures developed for general-purpose computing. This not only provides proven performance gains, it also dramatically lowers power consumption, heat and space requirements. The Netezza architecture is unique in delivering streaming analytic processing, leveraging high-end embedded technology used primarily in areas such as video games, industrial process controls and medical imaging, and adapted by Netezza for data warehousing and advanced analytics. The approach uses commodity FPGAs (field programmable gate arrays) to do the bulk of the filtering, with low-power, multi-core x86 processors handling the remainder. These components are integrated together to form a powerful and compact processing node called a Snippet Blade (or S-Blade). The innovative balanced architecture is what provides the performance difference, along with substantial energy savings. Unlike traditional systems that rely on general purpose processors and software for processing high speed data streams, the Netezza appliance streams data off the disk and through query logic loaded into the FPGA. By performing initial filtering as data streams off the disk, the FPGA reduces the workload on the CPU cores by 95-98%. Since the S-Blades main processors are not expected to handle gigantic workloads, they can handle multiple streams of data far more efficiently.

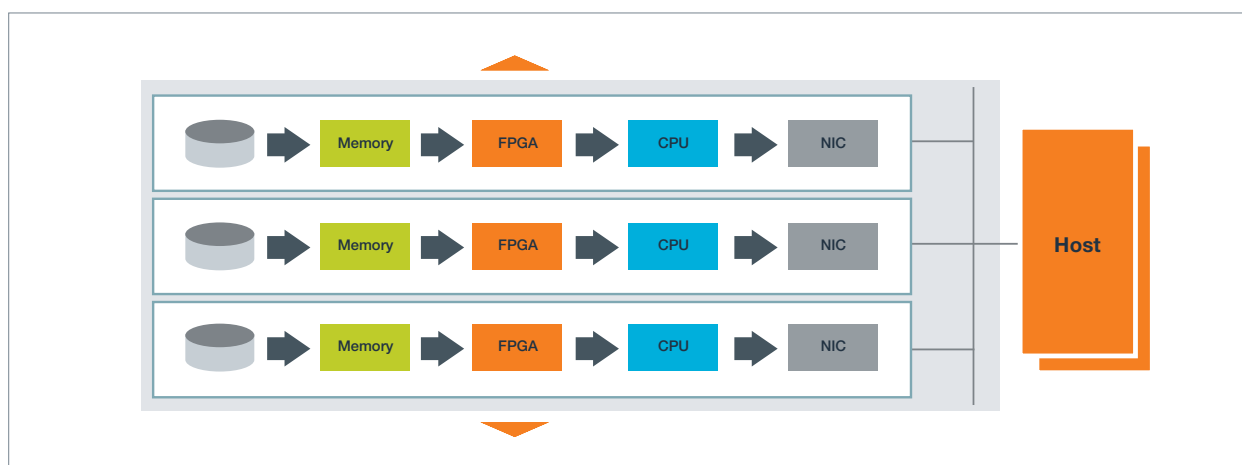


Figure 2 Data Flow in the Netezza Appliance

A Linux host server in the Netezza appliance aggregates S-Blade results and manages the query workload. Peak Power and cooling requirements for the Netezza TwinFin™ (fully populated one-rack - 1.8 tiles) appliance, including the host and 12 S-Blades (with 32 TB of uncompressed user data capacity) are only 7,635 Watts together with only 26,100 BTU/hour of cooling. Compare this to two similarly sized configurations:

- Oracle Exadata requires two racks (3.6 tiles), 11,500 watts of power and 39,300 BTU/hour cooling.
- Teradata 5500H with 2X dual Intel Xeon processors and NCR 6843 storage array requires five racks of equipment, 18,500 Watts to power and produces 63,000 BTU/hour that needs to be ventilated and cooled.

Power Math: You Make the Call

So when it comes to power savings, what's the bottom line? The following table compares power and cooling of a Netezza TwinFin 12 appliance with well-known systems of equivalent processing power and uncompressed storage capacity. Figures are based on APC's online calculator that provides power and cooling ratings for major blade and rack server configurations (www.apc.com/tools/ups_selector/index.cfm). Netezza encourages companies to use the APC calculator to make their own independent assessment of power requirements to validate all vendor claims.

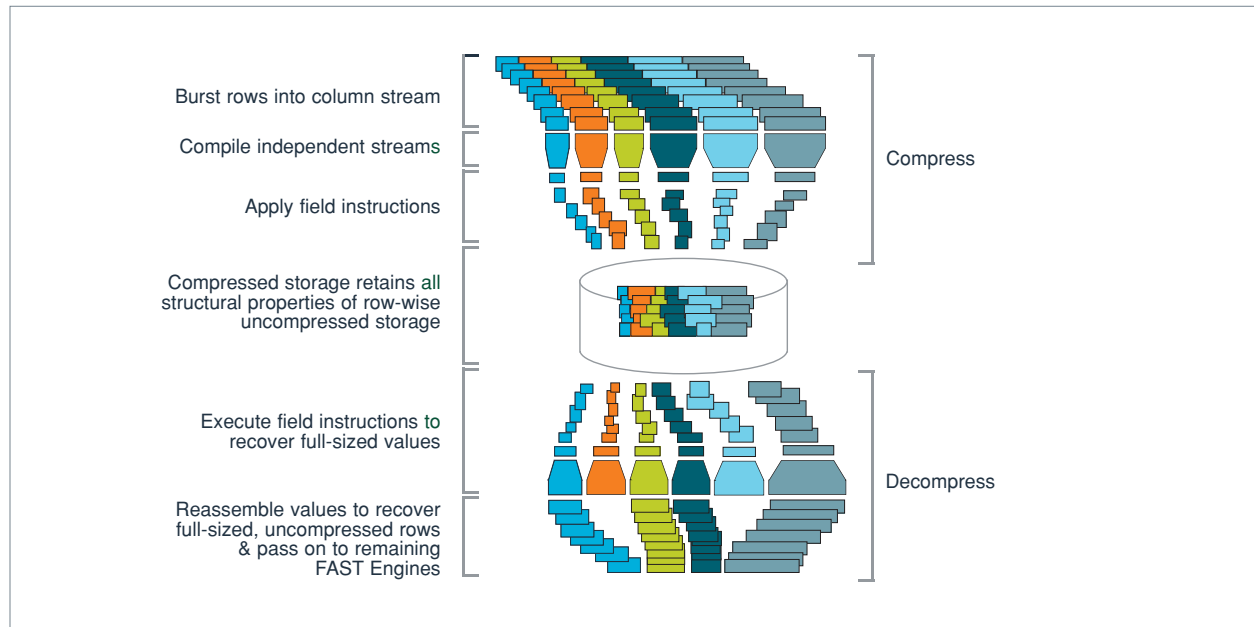
Data Size= 25 TB	Netezza	MPP Solution	SMP Solution	Packaged Appliance Solution	Cluster Solution
# of Data Center Tiles	3 Tiles	12 Tiles	11.5 Tiles	20 Tiles	14.5 Tiles
Power – Watts	7,500	38,000	28,600	32,000	30,600
Cooling – BTU/hr	25,600	129,900	97,600	109,200	104,500
Cooling Watts ²	7,700	39,000	29,300	32,800	31,300
Total Watts	15,200	77,000	57,900	64,700	61,900
Netezza Power Efficiency (%)		80%	74%	77%	75%

Performance in a Small Footprint

Orange UK, the UK's most popular mobile phone service, chose the Netezza system to analyze billions of Call Detail Records (CDRs) in a fraction of the time required by other systems. The need for in-depth analysis about the success of targeted products and services is increasingly important within the telecom industry, and companies like Orange can leverage the Netezza system to ensure that CDR data can be turned into actionable information. Even as the amount of data stored in Orange's data center continues to grow, it has managed to reduce its equipment footprint. The new infrastructure in the data center has seen the number of cabinets spaces drop from 26 to 9. With complexity and floor space now seen as two of the biggest costs facing IT departments, Netezza has helped Orange to prepare itself for future growth. As a Senior Data Center Engineer at Orange explains, "Space and power are always at a premium within our data centers. The Netezza solution saved large amounts of both while providing increased performance. The speed and ease of deployment was an added bonus."

Compression, at speed

We have chosen to compare ourselves to our competitors using uncompressed storage numbers. Compression comes standard on our platform, and it provides our customers with a rather unique opportunity to store up to 8 times the data while providing a performance boost as well. We use our FPGA to uncompress data as we read it from the disk, thus decreasing the load on the processors. Employing a patent-pending method for compiling columnar data in all tables of the database, the Netezza Appliance will be able to compress the data to use disk much more effectively while greatly increasing query performance as data streams from the disk. Thus, when scanning data that compressed at 2X the Compress Engine will have the net effect of boosting streaming query performance by 2-3x (100-200%). Is this an unfair advantage? Yes, for our customers.



Conclusion: The Energy-Efficient Data Warehouse

Netezza has earned worldwide recognition for its breakthrough architecture that processes hundreds of terabytes or even petabytes of data in record time. Along with unmatched performance, the Netezza appliance provides an extremely efficient processing approach that requires only a fraction of the power used by traditional architectures. Rather than struggle with growing power and cooling constraints, more and more organizations are discovering the impact that the high-speed, low-power Netezza appliance can have on their businesses. Netezza customers reap the benefits: huge power savings, minimal space requirements and lower cost of ownership—along with performance that provides a competitive edge in the marketplace. ®



About The Green Grid

The Green Grid is a global consortium of companies dedicated to advancing energy efficiency in data centers and computing ecosystems. The Green Grid does not endorse any vendor-specific products or solutions, and will seek to provide industry-wide recommendations on best practices, metrics and technologies that will improve overall data center energy efficiencies. Membership is open to companies interested in data center operational efficiency at the Contributing or General Member level. General members attend and participate in general meetings of The Green Grid, review proposals for specifications and have access to specifications for test suites and design guidelines and IP licensing. Additional benefits for contributor members include participation and voting rights in committees and working groups. Additional information is available at www.thegreengrid.org.

Sources

1. A Message from Data Center Managers to CIOs: Floor Space, Power and Cooling Will Limit Our Growth – Gartner, August, 2006
2. Power Surge – Darrell Dunn, Information Week, February 27, 2006
3. Power and Cooling Remain the Top Data Center Infrastructure Issue – John R. Phelps, Gartner, February 2009
4. U.S. Data Center Conference Focuses on How to Do More With Less, Mike Chuba, Carolyn DiCenzo, Donna Scott, Gartner, February 2009
5. Data Center Power and Cooling Scenario Through 2015 – Rakesh Kumar, Gartner, March 14, 2007
6. From the EPA's report to Congress on August 2, 2007 – Presented by Andrew Fanara
7. Survey Results on European Data Center Expansion Plans – Mike Chuba, Gartner, March, 2009
8. Guidelines for Specification of Data Center Power Density – Neil Rasmussen, White Paper #120, APC
9. 2006 Data Center Polling Results: Power and Cooling – Michael A. Bell, Gartner, February, 2007

SAFE HARBOR

"Safe Harbor" Statement under the U.S. Private Securities Litigation Reform Act of 1995: Certain information contained in this document is forward-looking in nature. Any expectations based on these forward-looking statements are subject to risks and uncertainties and other important factors. These and many other factors could cause delivery of products, features or enhancements to differ materially from expectations based on these forward-looking statements. Netezza does not undertake an obligation to update its forward-looking statements to reflect future events or circumstances.



Netezza Corporation

26 Forest Street, Marlborough, MA 01752

+1 508 382 8200 TEL

+1 508 382 8300 FAX

www.netezza.com

About Netezza

Netezza (NYSE: NZ) is the global leader in data warehouse and analytic appliances that dramatically simplify high-performance analytics across an extended enterprise. Netezza's technology enables organizations to process enormous amounts of captured data at exceptional speed, providing a significant competitive and operational advantage in today's data-intensive industries including digital media, energy, financial services, government, health and life sciences, retail and telecommunications. Netezza is headquartered in Marlborough, Massachusetts and has offices in Northern Virginia, Canada, the United Kingdom, Germany, France, Japan, Korea, Australia and Singapore.

For more information about Netezza, please visit www.netezza.com.