



## Product Brief

### IBM's iDataPlex & BLADE's RackSwitch: A Game Changing Combination of Energy Efficiency & Dramatically Lower Cost for Web 2.0 Data Centers

#### Introduction

For almost a decade a computer design chasm has existed between traditional commercial computer manufacturers and evolving “Web 2.0” (information/entertainment providers and social networking) companies:

- Traditional server makers have focused on building enterprise-class systems that offer solid reliability, availability, and security (RAS) — and that have been designed to be deployed in highly-available (HA), no-failure-allowed, the-transaction-must-go-through environments. And all of this RAS and HA stuff came at a significant price premium...
- On the other hand, Web 2.0 companies do not focus on high-RAS and high availability — except when conducting buy-sell transactions or when running their own run-the-business applications. The computing architectures of these companies readily tolerate system/network failures and simply and transparently fail-over Web requests to other active servers. Downed equipment can be fixed whenever technical maintenance crews can get around to it...
- In addition, some financial market and entertainment applications make use of *high performance computing clusters* that follow the same pattern as Web 2.0 architectures: it's more cost-effective to ignore the loss of a few nodes than to build in RAS and HA into a 512 node cluster.

*This major conflict in design has forced Web 2.0 companies (such as Yahoo, Google, Amazon, AOL, and MSN), as well as large financial service companies (Merrill Lynch, Morgan Stanley, et al), and large entertainment providers (such as Sony) to build their own computing data complexes (dataplexes) from scratch.*

This situation must have really annoyed IBM. First, the sales opportunities in the Web 2.0 market are huge (orders for 10,000+ servers plus are common). Second, IBM knows it has the engineering expertise needed to design computing solutions in any computer market (IBM builds systems for a broad range of computing markets including high-performance computing[HPC], for secure transaction processing, and for traditional business applications processing — using technologies that range from miniscule nanoprocessors to comprehensive mainframe architectures). And third, IBM has a major purchasing advantage over these Web 2.0 companies (IBM buys millions of computer components each year at highly-discounted volume rates so IBM knows that it can make decent margins on the sale of commodity designs to this market). Yet IBM could not crack the Web 2.0 market. At least until now...

On April 23, 2008, at the Web 2.0 Expo in San Francisco, IBM introduced a new class of server known as the “iDataPlex” solution. This new design is not focused on RAS and HA — it is focused on the top three elements that drive Web 2.0 IT buying decisions: *density, energy*

## **BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center**

*efficiency, and cost.* In this *Product Brief*, *Clabby Analytics* (that's me) examines IBM's new iDataPlex server — and also the new, energy-efficient, patent pending rack switches that Blade Network Technologies (BLADE) has designed to service iDataPlex. BLADE's "RackSwitches" contribute heavily to driving down iDataPlex networking costs while also greatly reducing energy demands. This combo, *Clabby Analytics* believes, will help IBM finally open previously-closed doors in the Web 2.0 marketplace.

### ***Homework Time...***

Over the past year-and-a-half, IBM spent almost \$2 million on Web 2.0 market analysis — including hundreds of face-to-face interviews with the IT buyers at all of the leading Web 2.0 companies. What this market analysis and these interviews revealed is that Web 2.0 IT buyers have distinctly different RAS/HA requirements from those of their counterparts in the enterprise computing world.

Prospective iDataPlex customers told IBM that their buying criteria centers on:

- Total application cost (capital expense);
- Total cost of ownership (operational expense);
- High energy efficiency with minimal power draw;
- Usable density (meaning that manufacturer specs may indicate a system is capable of holding ten or twelve or more rack servers or blades — but Web 2.0 customers want to know how many can really be deployed without generating too much heat);
- Single point of management and open infrastructure (Web 2.0 buyers require IPMI [intelligent platform management interface] standard adherence);
- Flexible configurability (the ability to mix and match building blocks to meet each Web 2.0 company's specific requirements);
- Rapid scalability; and,
- Global deployment capabilities.

IBM also learned that these companies have built their own type of "RAS/HA" software applications (essentially based on failover algorithms); that each company has differing configuration requirements (requiring great flexibility in design); and that energy consumption characteristics are paramount in Web 2.0 systems designs (and are measured in kilowatts per floor tile).

A closer look at Web 2.0 high-availability and RAS buying criteria showed that Web 2.0 companies have generally written their code to deal with server/component failures. In the Web 2.0 world, system failures are expected across a large pool of servers. To handle these failures, automatic load-balancing kicks in — and if an IP address drops out, Web 2.0 fail-over software simply finds a new IP address in order to carry-on with a given Web-page request or process.

Just to be clear, it is important to note that Web 2.0 companies do have a need to run enterprise-class systems. Transaction processing must be secure; transaction failures need to be capable of roll-back. In addition, Web 2.0 companies need to run their respective businesses on reliable, available, secure systems. But the lion's share of Web 2.0 processing involves handling Web page content, streaming files, and supporting on-line gaming, social networking,

## BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center

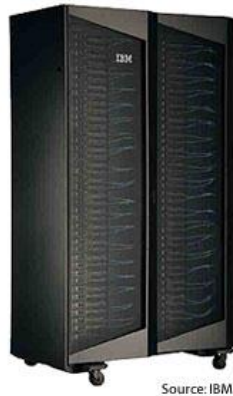
virtual worlds, multi-tenant hosting and the like — activities that are mostly free to the user that don't require guaranteed high-availability, redundant systems components, and other “baggage”.

*The composite picture that can be derived from these requirements is that Web 2.0 buyers want low-cost, extremely energy efficient, dense systems that adhere to formal management standards. These systems need to be easily scalable and rapidly deployable across the globe. These buyers are not willing to pay for redundancy in their systems architectures — in fact, they expect system/component failures and have software that they have written to deal with that eventuality.*

### ***IBM's Answer: The iDataPlex Design***

To address these requirements, IBM designed a completely new rack-based solution that has deemphasized redundant components and RAS/HA in the server — instead concentrating on density, energy efficiency, and low cost across the entire infrastructure. That solution is IBM's iDataPlex (shown in Figure 1).

### **Figure 1 — IBM's iDataPlex**



This new design:

- More than doubles the number of systems that can run in a single IBM rack;
- Uses 40 percent less power while increasing the amount of computing that can be done by 5X;
- Can be outfitted with a liquid cooling door on the back of the rack that enables it to run at “room temperature” -- no air conditioning required,
- Uses all industry standard components as well as open source software such as Linux to help lower costs.

*Clabby Analytics* had the privilege of previewing this server solution before its official announcement — and as I inspected this rack and server solution inside and out, four design features immediately stood-out when compared to traditional rack-servers or blades:

1. First, this system is wide, not deep (as is the case with racks and blades). The iDataPlex design features side by side servers and switches that are 15 inches deep (blades are usually double this deep). Shorter servers and switches dissipate heat more easily, making it possible to pack more (and more powerful) servers into a dense cabinet. iDataPlex enables 84 servers to be densely packed in the same geometry of a rack environment.

## **BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center**

2. Second, the use of redundant power supplies, redundant fans, redundant network interface modules, and other redundant components has been reduced. Example: in iDataPlex, two servers share one very energy efficient power supply (that operates at 92% efficiency). Further, the use of many, small fans has been replaced by the use of fewer, larger fans (that use 66% less power and are quieter). By eliminating redundancy, IBM has been able to reduce system cost while improving energy efficiency.
3. Third, the iDataPlex generates heat — but it does an excellent job of handling cooling as compared with equivalently-configured racks or blades (due to superior airflow design, server/switch depth, and the use of fewer components). Plus, when a water door heat exchanger is installed on the back of an iDataPlex, data center cooling costs can actually drop significantly because water has the capability of drawing 23X more heat from a heat source than air (so chilling and air conditioning costs can thus be lowered — in many cases very significantly).
4. And fourth, Ethernet switches that support the internal servers are housed in the iDataPlex cabinet along with the servers. By handling networking at the cabinet level, Web 2.0 companies can take advantage of the existing air or water cooling structure of the iDataPlex in order to efficiently cool both servers *and* network switches. Further, by using internal rack switches, iDataPlex servers do not need to rely on external, centralized, costly, energy-hogging central Ethernet switches!

The following section explores these two points in greater detail.

### ***iDataPlex Combined with BLADE RackSwitches = Very Low Cost, Extreme Energy Efficiency***

As described previously, the three most dominant buying criteria for Web 2.0 companies are density, energy efficiency, and cost. Blade Network Technologies' RackSwitches help address each of these criteria. Here's how:

- *Density* --- by incorporating the switches into the iDataPlex rack, users do not need to bring one or more cables per server to a central switch location. (Consider a typical case where 84 or 168 networking cables need to be tied to a central switch. The deployment of all of this cabling is an inefficient, expensive and cumbersome way to connect servers).

***To further illustrate this point, take an example of 10 iDataPlex racks with two network connections per server (for bandwidth). This would mean that 1,680 Ethernet cables would need to be connected to a central switch, where 7 switches with 240 1G ports each would connect them. Two more switches would connect the first 7 to each other (assuming 8 x 10G uplinks, so 3 to 1 oversubscribed). As a result, these 10 racks would require 9 switches mounted in 5 dedicated racks – at a staggering 15 rack units, 190 pounds, and 6KW per switch! Five networking racks would be needed to support 10 compute racks!***

***A better approach is to put the first level of switching into the iDataPlex racks and use a much smaller core switch. Using the above example, that's 4 switches per rack using 1U each, and 3 core switches. Since the space for the 1U switches is already provided in the iDataPlex racks, only a single switching rack would be needed (instead of five)!***

**BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination  
of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center**

- *Density, part 2* (cooling implications) – Dense servers generate a lot of heat. So cooling components within these servers is important — and air flow design is of paramount importance. Many networking vendors offer switches that have been designed for wiring closets — not for system racks where air needs to flow from front to back. When these switches are deployed in racks, a precious vertical slot is used to house an expensive baffle that redirects airflow in the proper direction. IBM's partner, Blade Network Technologies, uses a patent-pending front-to-rear air flow design that does not require additional baffling, and accordingly does not need to occupy an additional vertical slot.

***BLADE RackSwitch 8000 and 8100 are designed for front to rear cooling in iDataPlex — and take up only 1 Vertical slot. On the other hand other switch vendors have airflow designed for the wiring closet, not the server rack, requiring an adjacent vertical slot for an expensive baffle just to reroute the airflow. BLADE's RackSwitch can cool itself more efficiently while taking exactly ½ the space in the rack. This space is often needed for cabling slack and power distribution units. Further, if four Gigabit Ethernet ports per server are required, then 8 switches would be required — and this is not even possible with competing switches.***

Chart 1 (below) provides a comparison of the centralized, wiring closet approach to networking racks versus the BLADE RackSwitch distributed approach.

**Chart 1 — Centralized Switching vs. the RackSwitch Distributed Approach**

	Central approach	Distributed approach
iDataPlex Racks	10	10
Servers	840	840
Uplinks per server	2 x 1 Gig E	2 x 1 Gig E
Rack-based switches	- none -	40 RackSwitch G8000
Core switches	9 Cisco 6509E	3 Cisco 6509E
Oversubscription	3 to 1	2.1 to 1
Rack space	9 x 15RU = 135 RU	40 x 1 + 3 x 15 = 85 RU
Networking Racks	5	1
Power	9 x 6KW = 54KW	40 x 120W + 3 x 6KW = 22.8KW
Cooling	Conventional air cooling	Rackswitches share iDataPlex cooling

Source: Blade Network Technologies — June, 2008

- *Energy Efficiency* — When judging energy efficiency, Web 2.0 IT buyers need to consider two elements: 1) power consumption characteristics at the device level; and, 2) device-related heat dissipation costs.

From a power consumption perspective, internal, scale-out BLADE RackSwitches need to be compared to a centralized, scale-up core Ethernet switches. In a previous research note (“*Overloaded Networks and the Evolution of “Rackonom-ics”*” located at <http://www.clabbyanalytics.com/uploads/BNTFinal.pdf>) Clabby Analytics found that BLADE RackSwitch products consume just six watts per 10 GE port compared with 20

## **BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center**

watts per 10 GE port for conventional chassis-based switches! Central core switches burn three times as much energy providing the same service!

*When considering this point, also remember that heat dissipation is also an important consideration when judging energy efficiency. Web 2.0 buyers should also factor into their energy efficiency equations the notion that using half as much power means that RackSwitch require half as much cooling! And also factor into the equation that using water cooling is 23X more efficient than using the air cooling that core switches use. (The thermal conductivity of air is 0.0262 W/ (m\*K), as compared with 0.6062 W/ (m\*K).*

*Energy efficiency advantages at the RackSwitch level, combined with energy efficiencies resulting from not having to use as much cooling give internal RackSwitches huge energy efficiency advantages over centralized core switches.*

*By taking advantage of water cooling in the blade or rack systems in which they are installed, rack-based switches have a huge cooling/heat dissipation advantage over air cooled, centralized core switches.*

- *Cost* — In the previously mentioned report, *Clabby Analytics* also discussed the price differences between core switches and rack switches. Succinctly, this report compared BLADE's RackSwitch products for IBM iDataPlex to CISCO's very popular Catalyst 6509. This report showed that BLADE's G8000 rack switch lists for \$6,437 for 48 ports of 1 Gb Ethernet. THIS REPRESENTS \$134/1 Gb PORT AT LIST vs. CISCO's \$540! And BLADE's G8100 costs \$19,437 at list — for 24 ports of 10Gb Ethernet. THE BLADE COST IS \$810/10 Gb PORT AT LIST vs. CISCO's \$6,503! And bear in mind that the CISCO switch, which was oversubscribed in the example used — is probably carrying an average of 5Gb per second per port. The BLADE product described was non-blocking with full bandwidth on every port.

*CISCOs core switches appear to be using more than three times the power — and therefore require three times the cooling — as compared to the G8x00 rack switches!*

### ***Additional Features: Unique Virtualization Capabilities***

BLADE's RackSwitches are unique in the networking virtualization features they support. Using software resident on each switch, the RackSwitch can reduce the complexity of the network by providing virtual NICs to groups of servers, even for a whole rack. To the network designer, a rack full of servers can be connected as though it's a single, large server, which greatly simplifies the network design. VLANs, ACLs, and other network settings are applied to a whole rack rather than each individual server. Attach points can also be designated per application, so that different network settings can be applied to each different application. For the server administrator, each server can be drag-and-dropped into an application group, and that's it — *there is* no need to consult a network architect for every server add or change.

### ***A Closer Look at BLADE's New RackSwitch Product Offerings***

BLADE's new rack switch product designs include the RackSwitch G8000 (a 1/10G Ethernet switch) and the RackSwitch G8100 (a 10G Ethernet Switch). BLADE's RackSwitch G8100 is a 1U switch equipped with 24 lossless, low-latency 10 Gigabit Ethernet (10 GE) ports,

## **BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center**

designed for emerging high-volume 10 GE application environments, high-performance clusters that require latency of 500 nanoseconds or less and/or as a 10 GE aggregation switch. BLADE's RackSwitch G8000 is a 1U switch equipped with 48 Gigabit Ethernet ports and four 10 Gigabit Ethernet ports for uplinks and/or stacking to create a single large virtual rack switch, is designed for rack-level server connectivity, Web 2.0 cloud clusters and/or as a Gigabit aggregation switch.

These switches also feature:

- Front-to-back cooling (a major airflow design advantage when installed in iDataPlex);
- "Lossless" I/O which is required to carry Fibre Channel over Ethernet (FCoE) storage traffic across Ethernet networks based on the emerging standards for Converged Enhanced Ethernet. Lossless I/O is featured with BLADE's RackSwitch G8100 product offering.
- Redundant power supplies and variable speed fans – because even though Web 2.0 installations are designed to tolerate losing a server or two, you don't want to risk losing connectivity to as many as 48 servers connected to a single switch.
- Non-blocking internal switching fabric and a complete suite of switching features;
- BLADE's RackSwitch G8100 uses deeply recessed CX-4 connectors, removing concerns about insufficiently supported cables, and also allowing rack doors to be shut more easily.

### ***Summary Observations***

Energy efficiency is of paramount importance to Web 2.0 companies. The business plans for most Web 2.0 companies heavily involve the use of computing resources to serve existing and new service products to ever-growing populations of users. Growth is heavily dependent on scaling computer services. And scaling requires energy. Energy is becoming increasingly constrained and expensive in many geographies through-out the world — and in some regions enterprises are being faced with energy caps. To sustain growth, Web 2.0 companies know that they must run highly-energy-efficient datacenters and systems/storage/network infrastructures.

To operate an energy efficient data center, Web 2.0 companies need to pay close attention to costs related to the facility's operation of chillers, air conditioners, power distribution units, uninterruptable power supplies, lighting, and the like. To contain information systems costs, Web 2.0 companies need to pay close attention to energy usage (power consumption) and associated cooling (air conditioning) costs.

By eliminating an overemphasis on RAS and HA, by removing redundant components, and by using energy efficient power supplies and fans, IBM has been able to architect a new class of server: a massively scalable Web 2.0 server solution dubbed iDataPlex. In addition, by partnering with Blade Network technologies, IBM has been able to very significantly reduce networking costs (in one case by a factor of 1:13) — while also significantly reducing energy consumption and cooling costs.

**BLADE's New RackSwitch & IBM's iDataPlex: A Game-Changing Combination  
of Energy Efficiency & Dramatically Lower for the Web 2.0 Data Center**

For Web 2.0 IT buyers, IBM's iDataPlex represents a new architectural approach to delivering industry standard computing platforms in a footprint and energy envelope that makes sense in very dense, massively scaled Web 2.0 data centers. And, through its partnership with Blade Network Technologies, IBM has been able to sweeten-the-pot by not only reducing systems acquisition costs, but also greatly reducing network operating costs for fast growing, cost conscious Web 2.0 companies.

---

**Clabby Analytics**  
**<http://www.clabbyanalytics.com>**  
**Telephone: 001 (207) 846-0498**

© 2008 Clabby Analytics  
All rights reserved  
June, 2008

*Clabby Analytics is an independent technology research and analysis organization that specializes in information infrastructure and business process integration/management. Other research and analysis conducted by Clabby Analytics can be found at: [www.valleyviewventures.com](http://www.valleyviewventures.com).*