Product Brief

Overloaded Networks and the Evolution of "Rackonomics"

Executive Summary

The existing model for data center network scaling is broken. New technologies (such as multi-core/multi-threading CPU architectures), new practices (such as server/storage consolidation and virtualization), and new application designs (such as input/output [I/O] intensive Web 2.0 applications and message-heavy Web services applications) have all contributed to a monstrously huge increase in demand for network bandwidth. And to deal with this demand, IT (information technology) managers are being forced to continually buy and deploy super-expensive, power-hungry, high speed core switches. This is no way to scale network bandwidth — it necessitates and perpetuates the continual deployment of costly, distributed core switches. Clearly, there's got to be a better way…

BLADE Network Technologies (BLADE), a market-leading supplier of Gigabit and 10G Ethernet network infrastructure solutions, believes that it has found a better, more economical way to scale network bandwidth for consolidated rack/blade server/storage environments. Rather than tying racks and blades into expensive, external core switches, BLADE has architected two new switch modules that can easily fit into existing blade server/rack environments — enabling networking bandwidth to be scaled cost effectively *within* a blade or rack cabinet, rather than externally. Using this approach, data center architects can standardize on a unified and affordable rack-level network infrastructure to cost-effectively provision and scale out Web 2.0 environments, high-performance computing (HPC) clusters and virtualized data centers. BLADE refers to this rack-based, economically differentiated approach as "rackonomics".

In this *Product Brief*, *Clabby Analytics* (that's me) examines BLADE's "rackonomics" concept (the idea that putting switches into blades and racks is more economical than using external core switches) — and its two new RackSwitch offerings (the G8000 and G8100) designed to implement its rackonomics view. What I find is that BLADE's put-it-in-the-rack strategy is valid and well conceived — and will enable IT managers to save tens-of-thousands of dollars in switch hardware, deployment, and energy costs for every external core switch they are able to avoid deploying. Further, I find BLADE's new RackSwitch designs to be highly innovative (especially when it comes to air flow design and energy efficiency); highly integrated (especially when it comes to the integration of disparate network fabrics); highly-virtualizable (if there is such a word); and extremely cost effective (especially when compared to CISCO core switching options).

A Closer Look at Rackonomics

The fundamental concept of rackonomics is extremely straightforward. Tying scale-out blades, racks, and storage subsystems into expensive external networking switches is costly

and inefficient. By linking servers into localized switches that reside inside blade and/or rack enclosures, IT buyers can save money by reducing acquisition costs, and by increasing utilization by virtualizing network resources. Operational costs can be reduced through automated network bandwidth provisioning (as opposed to manual provisioning), automated firmware update distributions, and decreased energy costs (smaller switches draw less power draw than larger, core switches — and, accordingly, smaller switches generate less heat and therefore require less cooling).

> ***Simply stated, the "rackonomics" concept simply identifies that placing low cost, energy efficient, easily managed and virtualized switches inside rack cabinets is more economical than continuing to buy expensive core switches.***

Are These Rackonomics Claims True?

BLADE argues that using a rackonomics approach is the best and most economical way to scale networks that serve blade and rack servers. And they claim that rack switches have distinct advantages over core switches in the areas of virtualization, cooling, and ease-of-management. Each claim deserves closer scrutiny.

Cost Advantages?

There are many ways to measure cost advantages when comparing core switches to rack switches. Some of these include comparing:

- The hardware prices when purchasing the same number of ports (examined below);
- The energy consumption/heat dissipation characteristics of each switch (examined below);
- Deployment/cabling costs;
- Service/maintenance costs; and,
- So on…

The Cost Claim

In this subsection, I examine raw hardware costs. As a core switch example I chose CISCO's very popular Catalyst 6509. The following (Chart 1) are three examples of pricing for this switch: 1) a 2:1 oversubscribed 10 Gb 56 port system; 2) a 4:1 over-subscribed 10 Gb 112 port switch; and, 3) a non-blocking, 1 Gb, 336 port switch.

Chart 1 — Three Cisco 6509 Switch Configurations

2:1 Oversubscribed 10Gb 56 ports						
Part number	**Type**	**Watts**	**List**	**Units**		**Extended**
WS-C6509-E	Chassis	0	9500	1	$	9,500
WS-SUP720-3B	Supervisor modul	282	28000	2	$	56,000
WS-X6708-10G-3C	8 port 10GbE	444	37500	7	$	262,500
WS-C6509-E-FAN	Fan	5	495	2	$	990
WS-CAC-6000W	Power supply		5000	2	$	10,000
	total watts	3682		Total Price	$	338,990
	watts per port	66		per port	$	6,053

4:1 oversubscribed 10Gb 112 ports						
Part number	**Type**	**Watts**	**List**	**Units**		**Extended**
WS-C6509-E	Chassis	0	9500	1	$	9,500
VS-S720-10G-3C	Supervisor modul	282	28000	2	$	56,000
WS-X6716-10G-3C	16 port 10GbE	457	47500	7	$	332,500
WS-C6509-E-FAN	Fan	5	495	2	$	990
WS-CAC-6000W	Power supply		5000	2	$	10,000
	total watts	3773		Total Price	$	408,990
	watts per port	34		per port	$	3,652

Non-blocking 1G	336 ports					
Part number	**Type**	**Watts**	**List**	**Units**		**Extended**
WS-C6509-E	Chassis	0	9500	1	$	9,500
WS-SUP720-3B	Supervisor modul	282	28000	2	$	56,000
WS-X6748-TX	48 port GbE	407	15000	7	$	105,000
WS-C6509-E-FAN	Fan	5	495	2	$	990
WS-CAC-4000W	Power supply		5000	2	$	10,000
Total	total watts	3423		Total Price	$	181,490
	watts per port	10		per port	$	540

Source: Cisco's Global Price List (generally available on the Internet)

Please note that the above prices are list prices. Resellers such as CDW generally discount CISCO products somewhere in the 30% range for new customers — and more aggressively for established enterprise customers and government buyers.

> ***Now, compare the above prices (and discount them if you desire) to BLADE pricing for its rack switch environments. BLADE's G8000 rack switch lists for $5,495 for 48 ports of 1 Gb Ethernet. THIS REPRESENTS $114/PORT AT LIST vs. CISCO's $540! And BLADE's G8100 costs $11,950 at list — for 24 ports of 10Gb Ethernet. THE BLADE COST IS $498/PORT AT LIST vs. CISCO's $6,503! And bear in mind that the CISCO switch is oversubscribed in this example — probably carrying an average of 5Gb per second per port. The BLADE product is non-blocking with full bandwidth on every port.***

Power Usage Comparison

Chart 1 also contains power consumption data. Note that BLADE's rack switch uses 125 watts for 48 ports, (or only 2.6 watts per port). CISCO's best case is 10 watts per port and its worst case is 66 watts per port. Further, BLADE's 24 port 10gig switch uses 6 watts per port vs. 66 for the 2:1 blocking Cisco solution. Now consider this: in many parts of the world some data centers just plain don't have power available to waste in this manner. And also consider that more power use generates more heat — resulting in the need to use even more power to cool a data center environment. One way to break this viscous cycle is to not waste power in the first place…

> ***The raw hardware cost comparison shows that BLADE's rack switch approach is significantly less expensive than the core switch approach. After considering this pricing data, it appears that BLADE's first claim (significant hardware cost advantage) is wholly justified.***

Virtualization Advantages?
Virtualization refers to the pooling of resources. Servers, storage, networks, applications, databases, I/O and other information systems components can all be virtualized to create large resource pools. Unused, pooled resources can then be shared — enabling IT buyers to get more utilization out of the equipment and software found in their information systems environment.

In the case of network virtualization, unused network bandwidth (network I/O) can be found and pooled — and then be made available to applications that need that bandwidth. By virtualizing network resources, two things happen:

1. Utilization rates increase as unused resources are found and exploited (enabling IT buyers to get better return-on-investment); and,
2. The cost of network switch acquisition drops for a certain amount of time because IT buyers get better utilization out of existing switches (and can thus forestall the purchase of additional switches).

Large core switches (such as CISCO's Nexus 7000 — chosen at random from CISCO's core switch product offerings) can virtualize I/O at the device level using what CISCO calls virtual device contexts. And rack switches such as BLADE's new RackSwitch G8000 and G8100 can also virtualize I/O at the device level. So clearly, device level differentiation does not create a major advantage for BLADE.

So why does BLADE seem to think it has a virtualization advantage over centralized core switches?

First, BLADE believes that by distributing virtualization to the rack/blade level, IT systems administrators can gain better control over power, cooling and space requirements. By so doing, IT systems administrators can build pay-as-you-grow distributed switch environments versus conventional centralized networking infrastructures that are more expensive to provision and require an overhaul and rebuild of the data center network.

Second, BLADE believes that it has distinct competitive edges in virtualization management. To this end, BLADE offers two virtualization feature sets in their switch-resident software packages: SmartConnect and VMReady — both of which simplify the management of virtualized network and server resources. SmartConnect enables I/O to be shared transparently across a blade chassis or rack (core switches don't operate this way). Additionally, SmartConnect can be used to improve fault tolerance by ensuring that network connections are always available. And finally, SmartConnect has been network optimized to simplify network management for systems administrators. Meanwhile, VMReady is a unique software offering that enables network characteristics to accompany virtual servers if they are moved from one location to another (this process is known as virtual mobility).

I found no directly equivalent SmartConnect or VMReady products on CISCO's web site. In these respects, BLADE, with its rack switches, does have an edge over CISCO's core switch environment.

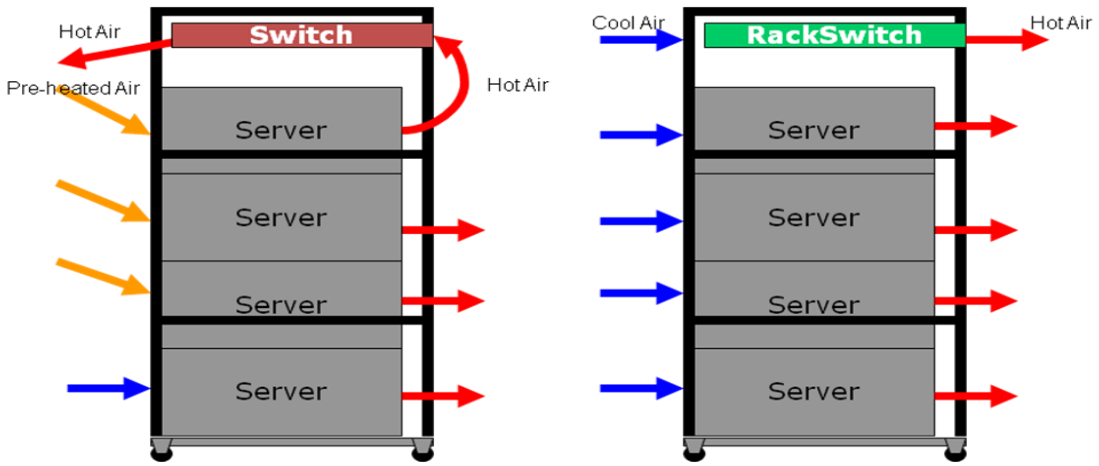
Cooling Advantages?

Core switches are very dense — and with power supplies that deliver up to 6000 Watts of power to core switch components, core switches can become very hot. This same argument, however, can also be applied to rack switches when considering that rack switches can sit inside very dense rack or blade server configurations that also generate a substantial amount of heat. So where are BLADE's advantages when it comes to switch cooling/heat dissipation?

BLADE rack switches have two distinct advantages in airflow/heat dissipation compared to both traditional rack switch designs as well as to core switches. These advantages are manifest in a unique, patent-pending airflow design; and in advantages that can be gained by exploiting water cooling at the rack or blade level.

With respect to airflow/heat dissipation, BLADE has designed an innovative new approach to circulating air over its rack switches. Most rack switches have been designed with airflow that flows in the opposite direction of server airflow (left side of Figure 1 below) — generating a condition known as "hot loops". These hot loops actually serve to increase cooling requirements by drawing warm air over already hot blade or rack servers (see Figure 1). BLADE's RackSwitch products are designed to draw cool air from the front of the rack or blade cabinet (right side of Figure 1 below) — enabling front-to-back airflow that allows for cool aisles and hot aisles (a more energy efficient approach to heat dissipation).

Figure 1 — Hot Looping vs. Front-to-back Heat Dissipation



Hot Looping **Front-to-Back**

Source: Blade Network Technologies — April, 2008

With respect to airflow/heat dissipation advantages over core switches, core switches are not water-cooled. But many blade and rack cabinets are now being manufactured to support water cooling. Water has a thermal conductivity of 0.6062; air is only 0.262 (making water almost three times more efficient than air for removing heat).

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.

BLADE adds one final point to consider when comparing RackSwitch power consumption vs. core switch power consumption. According to BLADE, RackSwitch products consume just six watts per 10 GE port compared with 66 watts per 10 GE port for conventional chassis-based switches.

If this is true, core switches are using more than three time the power — and therefore require three times the cooling — as compared to rack switches!

Ease of Management Advantages?
BLADE's claim that its rack switches are easier to manage is based on its BladeHarmony Manager blade management software. BladeHarmony Manager provides a data center view of where I/O resources are, and allows IT managers and administrators to provision that I/O). Further, IT managers and administrators can use Blade Harmony manager to update the software/firmware of thousands of switches with a single mouse click. And finally, BladeHarmony Manager can be integrated easily under Hewlett-Packard (HP) OpenView and IBM Tivoli management environments.

BLADE also argues that the management of networking environments will become even easier in the future as network fabrics converge. BLADE is working toward creating lossless Ethernet environment that will support Converged Enhanced Ethernet (CEE). CEE is an enhanced version of Ethernet for data centers that adds flow control and congestion notification across multiple lanes of high-performance data and storage traffic on a single, unified Ethernet fabric. It will someday enable several different network fabrics to be managed in a consistent fashion — including IP SAN (NAS, iSCSI) and Fibre Channel over Ethernet (FCoE).

In addition to its CEE efforts, BLADE continues to work on reducing latency (in order to improve performance in high-performance computing (HPC) environments); on reducing power requirements by building products with fewer, more efficient components; and on lowering overall cost using single attach (NIC/Adapter) for servers and storage.

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). These switches and their characteristics are illustrated in Figure 2 (next page).

Along with the cost, virtualization, and cooling advantages described earlier in this report, BLADE's RackSwitch products also feature:

- "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.

- Each RackSwitch has a non-blocking internal switching fabric and a complete suite of switching features, proven in the networks of half the Fortune 500 companies and matured over 5+ years of network deployments.
- 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.
- BLADE's RackSwitch G8100, a 1U top-of-rack switch equipped with 24 lossless, low-latency 10 Gigabit Ethernet (10 GE) ports, is designed for emerging high-volume 10 GE application environments, high-performance clusters that require latency of 300 nanoseconds or less and/or as a 10 GE aggregation switch. BLADE's RackSwitch G8000, a 1U top-of-rack switch equipped with 48 Gigabit Ethernet ports and four 10 Gigabit Ethernet ports for uplinks and/or stacking, is designed for rack-level server connectivity, Web 2.0 cloud clusters and/or as a Gigabit aggregation switch.

Figure 2 — Blade Network Technologies' RackSwitches G8000 and G8100

RackSwitch G8000
1/10G Ethernet Switch

- 48 ports, Gigabit Ethernet with 10G Uplinks
  - Top of Rack, 1U
  - 44 ports 1G, RJ-45 copper
  - 4 ports 1G, SFP fiber
  - 4 ports 10G, CX4 or SFP+ for uplink and/or stacking
- Air flow optimized for hot-aisle / cold aisle server / storage racks
- Applications
  - 1G server/storage connectivity
  - Clusters with 1G endpoints
  - Web 2.0 scale-out
  - Aggregation for 1G Blade switches

RackSwitch G8100
10G Ethernet Switch

- 24 ports, 10 Gigabit Ethernet
  - Top of Rack, 1U
  - 20 ports CX4 – the proven lowest latency
  - 4 ports SFP+ for fiber uplinks
- Patented CX4 cable retainer solves cable issues
- Air flow optimized for hot-aisle / cold aisle server / storage racks
- Applications
  - 10G server/storage connectivity
  - HPC – low latency (300ns) fabric
  - Aggregation switches

Source: Blade Network Technologies — April, 2008

Summary Observations

For midsized and large enterprises, the days of deploying expensive, energy-hungry, heat-generating centralized core switches are coming to an end. Energy and cost-efficient rack-based switches with device level virtualization capabilities are starting to make the scene. These new rack-based switches provide a means to more granularly scale a virtualized I/O environment, substantially reduce energy consumption/cooling costs, while also reducing management costs as compared to centralized core switches. And accordingly, these switches are ushering in an era of "rackonomics".

BLADE is leading this rackonomics charge with its new G8000 and G8100 rack switches. Features and functions of these products that are especially worthy of note include:

- These switches provide low cost, granular scalability;
- They have distinct airflow design advantages when compared to other switches;
- They can capitalize on water cooling at the blade/rack enclosure level; and,
- Advanced virtualization/virtualization management features (device level virtual-ization combined with BLADE's BladeHarmony, VMReady, and SmartConnect software) provide IT managers and administrators with simplified management.

Clabby Analytics expects these products to be hugely successful for BLADE. Blade servers represent the fastest growing systems environment in the IT industry — and BLADE is already one of the market's leading suppliers of Gigabit and 10G Ethernet networking infrastructure solutions that reside in blade servers, scale-out server racks, and storage racks. These rack switches position BLADE to compete even more effectively in this market segment.

What will be most interesting to watch will be whether BLADE's RackSwitches can unseat hugely expensive CISCO centrally controlled core switches. My bet is that BLADE is onto something — and that the enterprise marketplace will find BLADE's rackonomics argument to be highly compelling.

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