Pund-IT

PUND-IT RESEARCH

Weekly Review

In this issue:

- Shades of Green – Next Steps for Eco-Friendly IT
- IBM's Energy Efficiency Smoking Gun
- SanDisk Vaulter Disk and Spansion's Server Flash Portend Long Term DRAM Trouble
- Managing Risk in Evolving Markets – IBM's Global CFO Study 2008

November 28, 2007

Pund-IT, Inc.
2776 Sulphur Drive
Hayward, CA
U.S.A. 94541

Phone: 510-909-0750
Fax: 510-886-4937
charles@pund-it.com
www.pund-it.com

Shades of Green – Next Steps for Eco-Friendly IT

By Charles King, Pund-IT, Inc.

It is likely that 2007 will be remembered as the year of "green" computing, with vendors performing both as solo vocalists and as members of disciplined chorales. Group efforts such as The Green Grid, Climate Savers Computing Initiative, and Blade.org stood out from the crowd, but what has been most remarkable about the evolution of ecologically friendly computing is the growing number and variety of available products. From miniscule embedded processors to sophisticated datacenter cooling products to massively scalable virtualized server and storage consolidations, these efforts reflect the systemic nature of broader ecological issues that do not lend themselves to "silver bullet" solutions.

This may frustrate those intent on trumpeting one organization or another's greener-than-thou credentials, but it is a critical point for businesses hoping to maximize their green IT investments. Along this line, we would like to consider how companies can benefit both from "top down" solutions designed green from inception to deployment and "bottom up" strategies that aim to leverage existing IT assets in the most ecologically intelligent way.

The automotive industry offers an interesting analogy to consider here. Whether you are an ecologically concerned consumer or just looking to improve your mileage, you can consider alternative vehicles, including new hybrid cars like Toyota's Prius, high-mileage conventional fuel options such as Volkswagen's diesel-based Jetta, or emerging alternatives like the swanky all-electric Tesla sports car. Along with providing better fuel efficiency, such vehicles also leverage innovative, ecologically friendly, energy-efficient technologies.

IT solutions can offer similar benefits. New x86 servers from Dell, HP, IBM, and Sun utilize the latest/greatest Intel and AMD processors to good effect, as well as other energy-friendly features and add-ons. The same can be said for enterprise class UNIX systems from IBM, Sun, HP and others, which have all profited from next-generation CPUs and other component upgrades during 2007, as well as IBM's System i and System z platforms. Green IT also extends well beyond servers, as EMC has demonstrated with its well-designed and executed strategy around energy-efficient storage solutions and services.

But do you need to buy new to get green? Not necessarily. So far as automobiles go, you can significantly improve gas mileage with time-honored techniques such as making sure your car is well-tuned and your tires are properly inflated. Another option; just drive at the speed limit? According to the Department of Energy (DOE), slowing down from 65 mph to 55 mph can increase miles per gallon by as much as 15 percent. In addition, the DOE says each 5 mph you drive over 60 mph equates to paying an additional $0.21 per gallon for gas (at $3.00 per gallon – if you can find it). Not green enough? Why not ditch your car entirely, ride a bike, use a skateboard, or just take public transportation whenever possible?

While your mileage may vary, taking a common sense systemic approach to IT can also deliver notable energy benefits. For example, numerous consultants we have spoken with report that they find easily fixable cooling issues in virtually every datacenter they inspect, with problems ranging from mistakenly set thermostats to working systems being inadvertently turned off. Bigger problems obviously require more sophisticated solutions. For example, server and network virtualization solutions from market leaders like VMware holds particular resonance among green IT fans and with good reason since they allows compa-

nies to maximize server utilization while reducing their total number of servers and the amount of energy they consume.

Considering the breathtaking growth in most companies' information assets, consolidating existing storage assets via workable Information Lifecycle Management (ILM) strategies and data de-duplication solutions can also help organizations get the biggest bang out of their IT investment bucks. Some add-on technologies are well worth considering if datacenter greening is on your mind. For example, ensuring the efficiency of cooling systems should be high on any business IT users list, particularly those who depend on high-density infrastructures populated with blade servers. High efficiency air conditioning systems such as those offered by APCC and other vendors may be just what Dr. Green ordered, but liquid-cooled heat exchanger solutions (like IBM's Cool Blue) prove that literal bolt-on energy innovation is alive and well.

Finally, a discussion of green technologies would not be complete without mentioning the efforts of interested parties outside the IT industry. Just as many states run programs that aim to get gross polluting vehicles off the road, government agencies and power utilities offer numerous incentives designed to encourage businesses to upgrade out of date equipment and improve datacenter energy efficiency. One example: California's Pacific Gas and Electric (PG&E) has been at the forefront of such efforts in the Golden State, certifying numerous IT solutions for its rebate/incentive program and working with vendors to approve other servers, storage, and networking solutions. Another: The U.S. Department of Energy (DOE) recently announced a collaboration agreement with the Green Grid.

The lesson we see in all this is that there are as many ways for IT customers to become ecologically astute as there are shades of green. Eco-friendly technologies play a role in a growing host of new IT solution, and we applaud vendors for moving so aggressively in these areas. In addition, most are actively pursuing green initiatives in areas including manufacturing, transportation, supply chain processes, and product recycling, and we expect such efforts to continue and expand. We also believe that there will be opportunities for IT vendors to provide green products, services, and guidance well-beyond the desktop and datacenter. We will be interested to see how the IT community responds to Google's announcement this week of its plans to support research exploring renewable energy production. Considering the vast R&D resources that help drive the IT industry's singular innovations, we expect other vendors to launch broader energy research efforts, as well.

At the same time, IT users would do well to understand that newer shades of green are not the only colors available to them. These organizations, like sensible consumers everywhere, should realize that using existing computing assets wisely can deliver benefits that rival many new solutions, and that approaching technology challenges systemically is the surest way to achieve lasting results. In the end, the essential purpose of green IT is not to outfit datacenters in the trendiest available products but to follow mindful, sustainable computing practices. This is a critical and worthy goal to pursue at a time when resources of every sort are increasingly under siege.

About Pund-IT, Inc.
Pund-IT emphasizes understanding technology and product evolution and interpreting the effects these changes will have on business customers and the greater IT marketplace.

IBM's Energy Efficiency Smoking Gun

By Joe Clabby, President, Clabby Analytics

What you're about to read is a review of a recent IBM announcement that I consider one of the most important announcements made in the computer industry in 2007. It is the announcement of a power and cooling management program that IBM calls its "mainframe gas gauge". (Yes, I agree — this could be the MOST BORING program name in history — but stick with me).

IBM now has "the smoking gun" when it comes to proving how scale-up architecture is superior to scale-out architecture when it comes to energy efficiency. What the company announced was that it has begun "gathering and publishing typical energy consumption data for the IBM System z9 mainframe". IBM further states that "the data is derived from actual field measurements of approximately 1,000 customer machines, determining average watts/hour consumed which can be used to calculate watts per unit — similar to automobile miles per gallon estimates and appliance kilowatt per year ratings".

Nap time — right? Time for more caffeine? Ho-hum…

Wrong-O!

With this data in hand, enterprises can now perform comparative analyses that will make it possible to understand the true benefit of moving away from distributed, rack and tower computing designs to more energy efficient scale-up designs. If CEOs, CIOs, CFO, and information technology (IT) executives take the time to understand what this data really means — and if they extrapolate how much energy could be saved in their respective computing environments— they would all be hard-pressed to justify the use of scale-out computing designs to their stockholders, much less to the rest of us who want enterprises to stop wasting greenhouse-gas-producing energy.

Have I got your attention? Good! Let's keep going…

Background

A few years ago I started a research report that sought to compare the cost of powering scale-up versus scale-out computing environments. My instincts were telling me that scale-out computing (more specifically: distributed tower servers) would be prove to be significantly more costly to operate because these solutions were:

- Heavily underutilized (according to industry statistics, most operate at a 20% or lower utilization rate); and,
- Heavily under-virtualized (at the time of the report, the virtualization trend had not yet taken-off).

Further, I noted that scale-out computers each had their own power supplies (which are notoriously energy inefficient), and that the distributed nature of scale-out architecture forced a lot of NICs (network interface cards) and external hubs/switches/routers to be used — adding even more energy inefficiency to scale-out designs. I also remember musing about how much additional cooling would be needed to cool tens, hundreds, or thousands of distributed servers that run at only 20% capacity…

I ran into several challenges with that report — and eventually (and reluctantly) canned it. The biggest problem I had was that there was very little empirical data available at the time to compare energy usage in scale-up versus scale-out environments. But finally, this situation has changed. We now have metrics with which we can measure scale-up versus scale-out energy consumption and this data shows how truly wasteful scale-out computing is!

What Does IBM's Mainframe Gas Gauge Do?

IBM's mainframe gas gauge "monitors a mainframe's actual energy and cooling statistics (collected by internal sensors) and presents them in real time [on a system activity display]. With this system, a user can now correlate the energy consumed with work actually performed. Statistics can be observed real time or also summarized weekly when the machine reports its maintenance health. These statistics can also be summarized for project or trend analysis. Energy consumption statistics are used for demonstrating cost savings for electric rebates and programs to reduce Data Center energy consumption".

Further, IBM says its mainframe gas gauge "calculates how changes in system configurations and workloads can affect the entire energy "envelope" – including the power needed to both run and cool the machines. For example, a customer adding a single mainframe processor for Linux applications could project the exact amount of additional energy required before and after the feature is turned on. Normally less than approximately 20 watts are added when an Integrated Facility for Linux (IFL) feature is turned on. A single mainframe processor with zVM virtualization can perform the work of tens to hundreds of x86 processors because of the mainframe's capacity for running multiple, mixed workloads at high utilization rates. This efficiency is the key to the mainframe consuming much less energy than many x86 servers which have many more power consuming components This translates into a simplified infrastructure and cost savings".

What IBM's "Gas Gauge" Shows So Far

A single mainframe running Linux can perform the same amount of work as approximately 250 x86 processors WHILE USING AS LITTLE AS TWO TO TEN PERCENT OF THE AMOUNT OF ENERGY THEY REQUIRE! Further, IBM has been able to show that the data it collected in August and September shows that typical mainframe energy use is normally 60% of the a model's "label" or maximum rating. In other words, mainframes are performing 40% better than they were previously rated.

In an age when energy prices are going through the roof, and where there are real concerns about environmental damage caused by the fossil fuel many states use to produce electrical power, this kind of message should resonate throughout the datacenter…

Summary Observations

I find it curious how little public attention was paid to IBM's announcement, but I expect that most industry observers — as well as many prospective listeners — stopped reading the announcement when they saw "Mainframe Gas Gauge". Despite being the most energy efficient commercial computing system in the marketplace while running at a 90%+ utilization rate, IT executives still cringe when they see "mainframe" in any title (because most IT executives think mainframes are boring, old architecture).

But look beyond the “mainframe” to what is really happening. What IBM has implemented is a monitoring/measurement system that clearly shows how much energy can be saved by using *scale-up systems designs*. IBM should take this program a step further — and compare/contrast its four distinct server lines across various workloads. Standards should also be developed using this technique to compare the servers of Sun, Dell, and Hewlett-Packard with those of IBM.

Let’s also take a minute to talk about blades. As you may have noticed I’m a big proponent of scale-up architecture when it makes sense. If you must scale-out, and if you can centralize your computing infrastructure, blade architectures make a heck of a lot of sense. They share some mainframe design points such as minimizing power supplies used to service multiple servers; an internal networking backplane to service inter-server communications (as opposed to tens/dozens/hundreds of NICS and external hubs/switches/routers).

In addition, automated management software on blades is generally easily to deploy and operate than across distributed tower environments. But scale-up doesn’t make sense in every enterprise. Some enterprises (such as distributed banking and retailing) are better served by distributed towers. And mainframes and various other scale-up architectures don’t run Microsoft Windows. So scale-out still has its place.

In the end, IBM’s little heralded, barely audible announcement could and should have major ramifications on the computer industry provided IT executives open their minds and are willing to receive it. IBM’s mainframe gas gauge is not the main point — the main point is that now we have empirical data that shows how efficient scale-up designs can be (or conversely, how stupendously wasteful scale-out tower and rack designs really are). This is an extremely important announcement — one I hope IBM and the marketplace doesn’t underplay!

About Clabby Analytics

Clabby Analytics (www.clabbyanalytics.com) specializes in teaching our readers how to build information systems that can ultimately support the transparent flow of business processes.

SanDisk Vaulter Disk and Spansion's Server Flash Portend Long Term DRAM Trouble

By Jim Handy, Objective Analysis

On November 13 flash card maker SanDisk announced the Vaulter Disk, a NAND flash based PCI Express module which the company says "tag-teams with the hard drive to provide enhanced performance." Similar to Intel's proprietary Turbo Memory modules (formerly known as Robson) the 8-16GB Vaulter device harnesses the speed of embedded flash memory for frequently accessed files while continuing to use a standard hard drive for bulk storage of less frequently accessed data. Vaulter Disk will debut at CES (Jan. 7-10) in Las Vegas.

The next day Spansion, a NOR flash manufacturer, announced an agreement with power-zealot Virident to develop and market memory solutions designed to dramatically reduce power consumption and improve performance in Internet data centers. Citing that Spansion's NOR flash consumes as little as one-tenth the power per GB of conventional DRAM the companies plan to co-develop a "breakthrough" memory product that enables a significant expansion of main memory in data center servers, with fast read performance and high capacity but that consumes much less energy than DRAM.

Our Objective Analysis

Ten years from now we will look at PC and server memory architectures from 1981 to 2007 and say: "Did we really do that?" It will seem odd to us that the entire main memory of systems was made of DRAM. This is not the first such change to main memories – roll the clock back to the 1970s and early 1980s and you see computers using SRAM as main memory. That seems unbelievable today, with DRAM dominating main memory and SRAM relegated to "small" 2MB on-processor caches that, incidentally, are 4,000 times larger than some of those SRAM main memories of yesteryear.

Today, NAND flash is poised to roar into the PC market, and it's not coming in the form of SSDs. In fact, SanDisk's Vaulter Disk and Intel's Turbo Memory even challenge the need for hybrid HDDs. Compared to HDDs, NAND flash is an extraordinarily expensive form of storage which should maintain its current 20:1 cost/GB ratio with HDD for the foreseeable future. But although it's costly, NAND is also fast so it fits well within the memory hierarchy of computing platforms.

Why now? What has changed over the last year or two to make NAND more attractive for PCs than it has been in the past? Well, NAND crossed over DRAM's price per gigabyte in 2005, and is cascading down the cost curve at a rate that is causing the DRAM/NAND price ratio to double from today's 3:1 to 6:1 by 2013.

Now NAND may not be as fast as DRAM, but it's significantly cheaper. On the other hand NAND is more costly than HDD, but it is also faster. This is exactly the argument that makes any memory fit into the memory hierarchy: Faster than the downstream technology but cheaper than the upstream. By 2013 we expect for users or OEMs to look at their memory budget and observe: "For $100 I could either buy 75GB of DRAM or 450GB of NAND." With this in mind, they are likely to refer to some benchmark to determine which should get them a bigger performance boost, and go with that solution. In many or even most cases, that solution is quite likely to be NAND.

Meanwhile, NOR flash is on a steeper price curve than DRAM, and has the added advantage of consuming significantly less power. While NOR is still significantly more expensive than NAND (around twice the price) it is still cheaper than DRAM and the NOR flash interface is far easier to incorporate in existing main memory buses than NAND.

Where does this leave DRAM? While we expect to see the average size of DRAM in PCs and servers continue to grow, this growth will slow eventually. Historically this growth rate has averaged 45% per annum. In the future the growth rate could slow to something like 10%. As a result, DRAM will follow the path of SRAM, becoming the smaller faster expensive bit of memory that augments the speed of the larger NAND or NOR memory. Although today only three companies (Spansion, Intel and SanDisk) offer such solutions, we expect flash-based memory topologies to become widespread over time.

About Objective Analysis

Objective Analysis (www.Objective-Analysis.com) offers third-party independent market research and data for the semiconductor industry and investors in the semiconductor industry.

Managing Risk in Evolving Markets – IBM's Global CFO Study 2008

By Charles King, Pund-IT, Inc.

Turns out the trend toward globalization makes it riskier than ever to be a chief financial officer (CFO). According to IBM's new Global CFO Study 2008, changes in the global marketplace can seriously undermine CFOs' ability to make sound business decisions.

Traditionally, CFOs are responsible for their companies' financial planning and record keeping processes, and have the financial authority to make appropriations and authorize expenditures. Not surprisingly, since CFOs work hard to ensure their organizations perform well and outperform competitors, they are typically closely attuned to and responsible for business-related risks.

In the new Study, IBM, working in cooperation with the Wharton School and the Economist Intelligence Unit, closely examined how current or ongoing events could shift CFOs' relationship to and ability to deal with business risk. During the process of interviewing more than 1,200 CFOs and senior finance professionals, the study found cause for alarm.

For example, most CFOs prefer to let business units and geographies conduct finance activities in accordance with local preferences and specialized standards. But the benefits of such an approach are compromised by expanding risks associated with increasingly dispersed and diverse global business efforts.

Such risks are nothing new. For example, consider how the Taiwan earthquake in September 1999, which disrupted processor, memory, and PC manufacturers for weeks, affected U.S. business partners. Vendors, including HP and Dell, were affected by shortages and higher prices of key components in the run up to the critical holiday shopping season, a situation that eventually affected many vendors' bottom lines.

More recent events, such as the use of lead-based paint and other inappropriate materials in the production of children's' toys in China, offer stark examples of how the behavior of or misunderstandings by distant business partners can damage brands and reputations that companies have spent significant human and financial capital to build.

The IBM Study's participants seemed well aware of the dangers of expanding risk. Nearly two thirds (62%) of companies with annual revenues over $5 billion reported encountering material risk events in the past three years. Of those, 42% reported not being well-prepared, a number that was not particularly surprising given that only 52% of all participants said their companies have instituted formal risk management procedures.

Oddly enough, this anemic approach to risk management also has a significant material impact on a company's overall financial performance. The two of the agenda items that ranked lowest in importance to CFOs (supporting/managing/mitigating enterprise risk, and driving information integration across the enterprise) turned out to be differentiators for companies that outperformed their peers in revenue and stock price growth.

Is there a solution to the expansion of risk? IBM's study believes so, suggesting that CFOs and their companies would do well to institute enterprise-wide information integration and

good governance via common data definitions, a standard Chart of accounts, common standard processes and globally managed standards.

By following these guidelines, businesses would be well on their way to becoming what IBM calls Integrated Finance Organizations (IFOs). This is no new thing. In fact, IBM noted that organizations that have already embraced IFO principles enjoy nearly double the revenue growth rate of their peers.

Due to their central role in recognizing and recognizing the implications of risk, CFOs are in the position of becoming holders and providers of the “truth” regarding an organization’s real performance. Changing times may require changing roles, and by using the IFO blueprint suggested by IBM’s Global CFO Study 2008, Chief Financial Officers should be able to help their organizations better avoid and survive expanding risks in every corner of the globe.