CLABBY ANALYTICS

Advisory

Workload-Optimized Systems For Superior Business Value

Executive Summary

Computing systems exist to support and respond to changing business demands. You may be considering offering a new service to your clients: a new customer self-service application, a new type of savings account, or a new "do-it-for-them" service when your company is in the do-it-yourself business. Often, information technology is your company's face to your clients, prospects, partners, suppliers and the community at large. It's up to you to present your business innovation and value effectively and efficiently every time, and workload-optimized systems offer the best approach to doing so

> ***What is a workload-optimized system? Simply put, it is one that matches architecture to the needs it is purchased for. Systems performance is no longer defined only by the fastest processor or other leading benchmarks, but rather by the ability to deploy new services faster, with higher quality, and within financial objectives. Commoditization has led some to think that financial objectives will lead to Intel-based systems for all uses. In this Advisory, we show that this is not the case, and by way of example, we demonstrate how Power Systems can be less expensive for ongoing operational costs and initial outlay.***

To maximize your investment in systems platforms, you need to:

1. Increase their utilization of resources (using virtualization);
2. Find new ways to achieve higher qualities of service (often through special-purpose systems components); and,
3. Take advantage of superior economics.

Information technology buyers sometimes approach the solving of these challenges by gravitating toward x86 multi-core architecture without fully evaluating other, superior and less-costly architectures that meet these challenges. We firmly believe in the concept of workload-optimization (deploying solutions on systems architectures best suited to serve those solutions) as a means to meet the goals stated above.

In this *Advisory*, *IT Market Strategies* and *Clabby Analytics* examine why making the right server choices can help solve the above mentioned problems. To do this, we contrast x86 multi-cores with Power Systems and mainframes. All are offered by IBM as platform alternatives, and all have their place. We show how making the right server choices can help an enterprise realize greater ROI while also helping extracting greater business value from the servers that it buys.

Background

Merv Adrian of *IT Market Strategy* and Joe Clabby of *Clabby Analytics* recently concluded an eight city road show in the United States where we discussed maximizing the business value of information systems. Attendees at these seminars consisted of C-level executives, line-of-business managers, and information technology (IT) executives (primarily systems architects, managers, and administrators). The focal point of these seminars was to discuss the role that workload optimization plays in helping enterprises improve operational efficiency while also realizing greater business value from their server investments.

> ***What we learned in our discussions with attendees was that few audience members were familiar with the concept of workload optimization — and also that many were not aware of the additional business value that can be extracted by choosing the right system to execute particular workloads. The remainder of this report focuses on explaining these two concepts.***

"Workload Optimization" and "Business Value" – A Closer Look

We define workload optimization as "*matching applications and workload to the systems architectures that can best serve them.*" Describing the business benefits of workload optimization is a bit tougher: it necessitates familiarity with the concepts of virtualization and provisioning — and it demands an understanding of some of the feature/functional differences between server architectures.

Workload Optimization

The focus in describing workload optimization is on how to choose the right system to perform a specific task. Our key observations include:

- **Workloads use servers and their resources in different ways.** For instance, some applications benefit from having large amounts of available memory in which to execute; others require only a little memory to do their jobs. Some read large amounts of data from storage in repeated passes, but never write to it; others search out small amounts of data and make changes to it. Further, some applications require very strong security (for instance, bank transactions); while others require lower levels of security (for instance, email).
- **Servers offer different processing** characteristics and provide varying levels of quality-of-service (QOS)**.** For instance, some servers process floating point calculations, encryption, or other processing at the hardware level, while others process do so in software. This can have a dramatic impact on utilization characterisitcs, memory requirments, and ultimately performance.
- **Servers also offer different levels of QOS.** Some offer very strong security (for instance, a mainframe is the only server in the world to achieve EAL Level 5 security certification), while other servers offer lesser degrees of security.
- **Servers typically have differing utilization rates.** In the field, x86 servers generally max-out at about 55% utilization, while Power Systems can operate at about 80% utilization. Mainframes, on the other hand, can run at 100% utilization for long periods of time. This is a function of the chip architectures themselves and the software typically deployed to manage them.

> ***The task-at-hand for IT executives is to figure out what server characteristics their applications and workloads require — and then to enable those applications/workloads to execute on those servers whenever processing capacity is available on those servers.***

Business Value
We encourage our audiences to look beyond systems specifications and characteristics when choosing where to deploy applications and consider *what a particular system is intended to actually do for an enterprise*. For instance, a commodity x86 server might be able to process a business analytics job — but it might not have the scale or memory capacity to handle real-time business analytics on a large scale database in addition to the other workloads it may be pointed at. So while it may be "able to do the job," if it doesn't have enough memory or processing power, using it for business analytics workloads may result in a business not getting the information that it needs to make decisions in a timely fashion. Conversely, there's no reason to over-provision either – a system with a well-understood workload can be acquired with an eye to costs that match needs, and x86 servers can be very cost-effective, and can scale consistent workloads well at modest incremental costs.

> ***The business impact is greatest when the server can execute the work, and deliver the results to the business in a timeframe that has maximum impact, at no greater cost than is necessary to meet business service level requirements.***

Additional business value is derived from other systems characteristics such as reliability, high availability, scalability, efficiency, resiliency, and utilization rate (virtualization/-provisioning capabilities).

An Example: Comparing Power Systems and x86 Architectures
In April 2010, Intel Corporation announced a new generation of its x86-based Xeon server architecture: its Xeon 7500/6500 multi-core processors. These processors manifest the largest jump in performance in Xeon history — and they provide the expandability needed to run advanced scale-up applications as well as to allow (someday) for massive server consolidation.

Still, there are major differences between Xeon and POWER microprocessors and systems designs — and, accordingly, neither server processes all workloads in the most optimal fashion.

IBM's Power System's Website cites the major differentiators for its POWER 7 environment as:

- TurboCore for maximizing per core performance for databases;
- MaxCore for parallelization and high capacity throughput;
- Intelligent threading technology to utilize more threads where needed;
- Intelligent Cache technology to optimize cache utilization;
- Intelligent Energy to maximize performance dynamically when thermal conditions allow; and,
- Active Memory Expansion.

IBM offers a short video at: http://www.youtube.com/ibmpowersystems#p/u/7/CmXfZmplxUA that further describes these features, including the relationship between processor cores, threads, and on-board memory architecture.

Examples of POWER7 Features

POWER7 processor-based systems provide numerous freatures designed to optimize performance:

- ***Memory management***—the way in which memory is managed — and the amount of memory that can be managed — differs greatly between x86 architectures and Power Systems.
    - An IBM Power 795 system can support up to 8 TB of memory. Power Systems offer more memory per core, more memory bandwidth per core, more I/O bandwidth per core; and more cache per core than x86;
    - Power Systems offer active memory sharing— a feature that enables Power Systems to handle more workloads than x86 servers. Essentially, active memory sharing finds available memory and enables it to be shared with other applications that may require additional memory. The system manages this – programmers do not need to;
    - Power Systems offer active memory expansion — a feature that enables Power Systems users to launch more virtual machines or expand memory as needed. SAP, a market leader in business applications, finds this feature to be extremely valuable for handling SAP workloads that can spawn new tasks to handle specific business requirements; and,
    - Finally, Power Systems offer active memory mirroring at the hypervisor level, ensuring greater memory availability by performing advanced error checking functions.

IBM builds its own advanced memory hardware and software — and it is clearly functionally deeper and broader than commercial off-the-shelf memory. The level of sophistication of this memory translates into clear business value as memory is more reliable and better managed on Power Systems. (For instance, larger Power Systems have custom memory DIMMS with special ECC error checking for double-bit and single-bit errors — ensuring high degrees of memory reliability. Equivalent memory features are not found in the x86 world).

- ***Reliability/availability/security (RAS)*** — Power Systems feature advanced RAS (in both systems and software design) as compared with x86 servers.
- ***Advanced power management*** — IBM can monitor and control power use on its Power Systems servers. This issue has become important not just for those shops that want to be "green" – in some cities, more power is simply not available, as a New York bank recently found out when planning a data center expansion.
- ***Advanced service management*** — IBM offers numerous automated service management functions to monitor, control, secure and otherwise manage Power Systems servers.
- ***Thread support*** — the high end Power 795 can run over 1000 simultaneous tasks (4 threads per core, 256 cores.) For organizations suffering from "server sprawl," this opens the door to cost-effective, space-effective, power-saving consolidation.
- ***Advanced virtualization*** — see Figure 1(next page).

Virtualization Delivers Functional and Business Value

One of the easiest ways to demonstrate the functional differences between x86 servers and Power Systems is to consider how virtualization works in each environment. Figure 1 illustrates this by showing how PowerVM virtualization features deliver greater business value for scalability and performance, flexible operations, and infrastructure economics compared with VMware (the leading virtualization software on x86-based servers).

Figure 1 — Power VM Compared with VMware: Delivering Greater Business Value

Attribute	VMware ESX 4.0	PowerVM
Scalability and Performance		
Real CPU sharing	Up to 20 VMs per CPU (workload dependent)	Micro-partitiong allows dynamic adjustments of 1/100th of a CPU between running VMs
Architected maximum number of VMs	320 per copy of VMware	1000 per physical server using PowerVM
Practical maximum number of VMs	Tens per copy of VMware	Hundreds per server using PowerVM
Real CPU and memory capacity on demand	No	Yes, non-disruptively
In-memory support	Shared virtual memory pages (detected via background operation)	Active memory sharing dynamically flows memory between running VMs
Virtual Machine (VM) scalability	Up to 8 CPUs, 255 GB of memory, modest I/O bandwidth	Up to 256 CPUs, 8 TB of memory, extensive I/O bandwidth
Flexible Operations		
Command and control, monitoring, automation infrastructure	Modest, yet easy to use	Extensive, robust
Resource over-commitment support (memory, CPU, network, I/O)	Modest	Extensive
Virtual Machine mobility support	Yes; essential for workload mgmt across multiple copies of VMware	Yes, live partition migration supported across (and between) POWER6 and POWER7 servers and blades
Infrastructure Economics		
Cost-efficient disaster recovery	No; typically requires a duplication of hardware and software license fees	Yes, including PowerHA and VMControl system pools
Cost-efficient technology refresh	No; typically requires re-purchasing new hardware and application verification	Yes, including live migration of VMs from POWER6 to POWER7 servers

Source: IBM Corporation, October, 2010

Anecdotally, Power Systems customers relate that they can achieve 80% utilization using Power Systems virtualization features as compared with about 50% utilization on x86 servers. A customer who buys a Power System gets an additional 1/3 of a system in additional processing power, or a customer who buys three Power Systems essentially gets the fourth server free. Those who argue that the cost of acquisition is less on x86 architecture should pay particular attention to this little tidbit of information…

Why do Power Systems perform better for virtualization? Some of the major differences in systems designs can be found in:

- ***Balanced design*** — Power Systems processors, I/O subsystems, and memory bandwidth are architected for balanced performance — and have been designed to work together to deliver the highest performance possible on various workloads.
- ***PowerVM granularity*** — PowerVM manages virtual machines. It can find unused processing cycles down to 1/100 of a processor — and make those cycles available for other applications and workloads to exploit. Further, a Power System does this dynamically (without operator involvement) — helping customers get increased utilization automatically.
- ***Priority Schemes*** — Power Systems can be made aware of workload priorities, ensuring that high priority applications are processed more quickly than lower priority applications.

> ***All of the features and functions described in this section illustrate how virtualization drives higher utilization rates and deliver greater business value to Power Systems users. Clearly, x86 architecture can be used to process many different types of applications — but customers who require greater performance, RAS, and utilization may wish to consider Power Systems when running workloads that need more "oomph" than delivered on x86 architecture.***

Customer Scenarios

Energen, an independent oil and gas exploration and production company, provides a great example of the kind of return-on-investment that can be recouped by taking the extra step of tuning and managing applications and databases — as well as by paying close attention to software license management.

Many enterprises use multiple systems architectures to process varied workloads — from infrastructure on x86 Windows to database and applications on Unix — and Energen is no exception. Energen recently underwent a major server consolidation and virtualization transformation, moving from 20 physical Sun Unix servers to two large IBM Power Systems (formerly System p servers). Further, Energen consolidated and virtualized its x86 environment, moving from 60 physical servers to just 4 servers.

By consolidating and virtualizing these environments, Energen has been able to reduce operating costs (Energen has been able to save $500,000 per annum on its Unix systems and $600,000 per annum on its Windows [x86] environment). A closer look at these cost savings finds that Energen focused strongly on license management and application tuning to exploit underlying virtualized resources.

A big piece of these cost savings came from reductions in Oracle licensing fees (the company eliminated redundant license costs, and then tightened and tuned its database environment). But it is also noteworthy that Energen took the time to tune its applications to exploit the micro-partitioning feature of PowerVM (micropartitioning allows as little as one-tenth of a single processor to be exploited — keeping all available cycles busy). Further, Energen also takes advantage of the Virtual I/O Server feature of PowerVM. *As a result of these practices, Energen has been able to boost its systems performance by 92 percent, and cut its batch runs from 24 hours to 2 hours.*

> ***Energen's results illustrate how taking extra steps to tune applications — as well as monitoring application/database license usage — can provide enormous paybacks when consolidating and virtualizing systems environments. Failure to tune application environments can leave easy-to-recoup money on the table — something few IT departments can really afford to do.***

Summary Observations

This *Advisory* started with a discussion about how computers should be used to support and respond to changing business demands — and then discussed how computers should be used to provide services for users. We emphasized that computer performance should no longer defined by the fastest processor or the leading benchmarks, but rather by the ability to deploy new services faster, with higher quality, and within financial objectives.

We then expressed our concern that while many IT buyers are gravitating toward Intel x86 multi-cores, Power Systems can be less expensive for ongoing operational costs and in initial outlay, depending on the workloads involved. Our latest research shows that

Power Systems have distinct advantages over x86 servers when it comes to running compute-intensive workloads — and in virtualization (as shown in Figure 1). We expect to publish a follow-on report on Power Systems virtualization and headroom characteristics that describes these Power Systems advantages in more detail next month.

As for suggestions and advice, we suggest that to solve your capacity and utilization problems — and to meet your service level requirements — your enterprise needs to analyze its application workloads and then determine which servers can best support the service level requirements of those workloads. And you need to be able to maximize the business value that can be derived from each type of server your organization uses. (For instance if strong security is a requirement for a workload, then choose the server that meets this business requirement — rather than just the server that has the "lowest cost".

Bottom line: workload optimization can help IT buyers achieve greater business value. We hope that this report showed you that workload optimized servers can help solve numerous operational efficiency problems can be found in datacenters today; and that by optimizing workloads on Power Systems, IT buyers can achieve greater utilization, overcoming simplistic acquisition cost arguments that mask the real costs of poorly chosen platforms.

Additional Reading:

The following reports on Power Systems and Workload Optimization are available (for free) at www.ClabbyAnalytics.com:

On Power Systems:
http://www.clabbyanalytics.com/uploads/POWERXeonFinal.pdf
http://www.clabbyanalytics.com/uploads/POWER7PundITJCfinal.pdf
http://www.clabbyanalytics.com/uploads/ServerMarketViewMarch2010UPDATE.pdf

On VMControl and Systems Director:
http://www.clabbyanalytics.com/uploads/VMControlReportFinalFinal.pdf
http://www.clabbyanalytics.com/uploads/IBM_Systems_Director_Article_Final.pdf

On Business Resiliency:
http://www.clabbyanalytics.com/uploads/BusinessResiliencyFinal__2_.pdf

On Virtual I/O:
http://www.clabbyanalytics.com/uploads/PayattentiontovirtualIOfinalrevisionupdatefinal.pdf

On Applications performance management:
http://www.clabbyanalytics.com/uploads/VirtualizationNOTenoughFINAL.pdf

On Integrated Service Management;
http://www.clabbyanalytics.com/uploads/SMCzFinalfinalfinal.pdf

Clabby Analytics
http://www.clabbyanalytics.com
Telephone: 001 (207) 846-6662

Clabby Analytics is an independent technology research and analysis organization. Unlike many other research firms, we advocate certain positions — and encourage our readers to find counter opinions — then balance both points-of-view in order to decide on a course of action. Other research and analysis conducted by Clabby Analytics can be found at: www.ClabbyAnalytics.com.

POL03079-USEN-00