The z10 is Not Your Father's Mainframe

A modern assessment

Terry Keene
CEO/President
Integration Systems LLC

INTEGRATION SYSTEMS

terry.keene@e-isys.com
678-641-1722

9/1/2009

The z10 is not your father's Mainframe

IBM has been building and shipping mainframe computers for over 40 years. Every generation of the mainframe gets better, faster, more efficient, and carries more and more of the data load for business and industry all over the world. In the early 90's there were no mainframes in India or China either for that matter. Now there are a growing number of mainframes in China running the major banks and financial institutions in that emerging country.

Mainframe installations and skills are growing in India, not only running financial markets but also in place to support millions of lines of programming code that are being developed for the rest of the mainframes circling the globe.[1] And yet, year after year since the first "minicomputers" were developed and shipped by Digital Equipment, Data General, and others, we have continuously heard about the mainframe being replaced by smaller, cheaper, easier to deploy technology of every kind.

Admittedly the mainframe has not assumed the role of personal computer, mobile appliance, small departmental server, nor supercomputing grid, but in fact most of the world funds clearing and transfers are still routed through mainframe connections. It has been estimated that the mainframe manages some 70% of the world's critical data. And today, with new departmental workloads in Java, C and C++, the mainframe is assuming many of these roles with Linux on System z.

So why is it then that year after year, someone writes a paper, article or book about the demise of this most capable machine? How can rational technologists believe that there is any technology in the market today that can do what the mainframe has done, and continues to do? What would make someone believe that a hundred Intel processors, ganged together in a loosely connected environment, could replace the capabilities that the mainframe has provided all these decades? The only answer that comes to mind is that those who espouse the demise of the mainframe don't understand either the problem, or the solution.

It's about moving "dirt"

To illustrate this point, let's assume that a group of contractors are bidding on a development job where millions of cubic feet of earth have to be moved to make way for a major construction project. Moving millions of $feet^3$ of dirt takes some very big machines. We call these "earth movers". They weigh tens of tons and carry hundreds of cubic feet in a single load. Even with that it will take thousands of these loads to move that much dirt. Presumably the contractors bidding on this project will have that type of equipment and have the experience to move that much dirt efficiently.

[1] http://www.mainframe-exec.com/articles/?p=107, **Mainframe Executive,** "Mainframe Sales Increasing in Southeast Asia, India and China"; Joe Clabby, March 19, 2009

Big earth movers are specialized, not very common, required special training and cost more than the pick-up trucks that you see roaming the streets in the suburbs. However, it is not inconceivable that contractors who have little experience with major projects like this might think that the task could be done with hundreds of ½ ton pickup trucks at a much lower cost. It will take more trips, require less skilled workers and could be competitive if managed correctly.

I know it sounds a bit far fetched, but that's the same thing that we see in the IT industry. There are many IT professionals that are managing IT organizations today who have never worked on a mainframe, don't really understand the capabilities of a mainframe, and truly believe that hundreds of commodity processors loosely connected can do the job at a lower cost. It will take more trips, require less skilled workers and could be competitive if managed correctly. In the land development business it's about moving the dirt. In the IT business it's about moving the data.

Why did your Father have a mainframe?

The mainframe was born some 50 plus years ago when the government needed a machine to support advanced research. Since that time it has taken its rightful place in the world's financial transactions, transportation management, health care and industry management. Most of that code was written in COBOL and VSAM and ran on systems from CICS and IMS to DB2. Today it still fulfills those venerated roles but it has evolved into the most powerful, flexible, reliable server in the industry, running new workloads from Java to Linux. This is the machine that supports tens of thousands of users working with hundreds of terabytes of data, with response times still in sub-seconds. This is ***not*** your father's mainframe.

Why then do many IT professionals, and as a result industry leaders, believe that migrating these trusted applications and data from a mainframe to rooms of distributed commodity servers can provide the same levels of service, security and reliability that the mainframe still excels at today? The truth is these folks don't really know what a mainframe is, much less what makes it so powerful. Let's take a short tour through this new mainframe, the System z10, and see what separates it from the rest. And the best way to review these differences is to look at one of those "mainframe is dead" white papers, in the form of "The Story about the IBM Mainframe Makeover" that HP has posted on their web site[2].

HP's "Fact 1": IBM Mainframe hardware investment protection falls short of HP's Superdome

What constitutes investment protection? Applications and data! The same applications that ran on MVS circa 1980 still run on the latest IBM z10. New applications written in

[2] http://h71028.www7.hp.com/ERC/cache/582432-0-0-0-121.html?ERL=true&jumpid=reg_R1002_USEN

C, C++, and Java running on Linux run on the same z10 platform. The same data that was used in the 1980's is still available on the z10, if you really want that data.

HP Application and data investment protection:

1980 – HP was running Motorola 68000 processor with HP-UX UNIX and proprietary processors with MPE.
Mid 1980's to 1990 – binary compatibility was broken with move to PA-RISC.
2000 – binary compatibility was broken again with move to Itanium IA-64.
2012 –binary compatibility broken again with the potential failure of Itanium IA-64 and move to Intel x86 (32/64) Nehalem and beyond.

Applications that ran on the Motorola 68000 did not run on the PA-RISC. Applications that run on the PA-RISC Superdome do not run on the Itanium Superdome. Both are still shipping today. The PA-RISC is end of life. The move just from PA-RISC to Itanium is at a minimum a recompile, and potentially a complete rewrite of the code, if the source code is still available. Which part of that investment is protected?

HP's "Fact 2": IBM's Big Green consolidation of 3900 servers to 30 mainframes doesn't add up.

Performance is the key – how could 3900 servers even be considered to consolidate to 30, or 50, or even 100 servers? Even if it were reduced to 100 servers that's a 39 to 1 reduction. Is that possible? It is if you understand how applications are deployed today in our distributed commodity server environments. Distributed commodity computing is a product of low cost personal computers and inexpensive servers that can support a single instance of an operating system, running typically a single instance of an application.

What has caused this evolution to single application per server, from the world of multi-user, multi-tasking servers that grew up over the 1960's 70's and 80s? Generally the operating systems, notably DOS and Windows, and in some cases early versions of UNIX variants like SunOS, HP-UX, and Xenix, were not sophisticated enough to support the reliability, throughput, security and overall performance that we came to expect from sophisticated "proprietary" OS's like MVS, VMS, GCOS, and MPE. They were indeed much less expensive, but in order to protect our data and our businesses, we limited these lesser servers to single applications, and deployed multiple instances to provide some degree of continuity if a single server failed.

The performance of the Intel, SPARC and PA-RISC based servers increased, the capacity improved from 16-bit to 32-bit to now 64-bit, but the programming model remained pretty much intact. We still today deploy a single application on a single server in the distributed commodity world.

With bigger and bigger manifestations of these servers, it became cost-effective to partition them into smaller, physical partitions to improve utilization. These were still single-application servers, physically running inside of a single enclosure, but using shared components like power, cooling and backplane interconnect. A great example of this type of partitioning is HP nPars on the Superdome. A 64-socket Superdome can be partitioned into multiple servers, each with a minimum of single "cell board" (a four socket PA-RISC or Itanium building block), or with multiple cell boards interconnected electronically.

Even in this partitioned case, a single version of the operating system and a single application are typically deployed on each physical partition. It is not the most efficient use of the technology since the granularity is by single cell boards, This means that if an application needs one or two more CPUs or a bit more memory, it is necessary to add the capacity of 8 cores and16GB of memory of a full cell board at a time, wasting the additional resources that are not required.

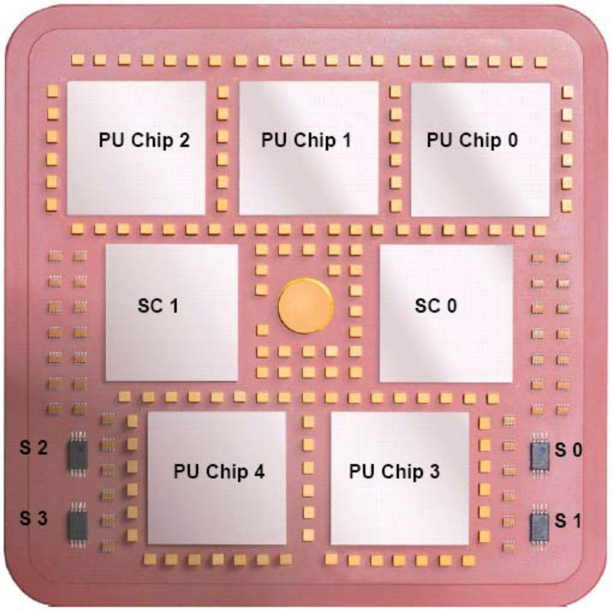
IBM faced this dilemma in the early 60's because the mainframe was very expensive and only a few corporations or institutions could even afford that kind of processing. Time shared systems were developed to allow multiple "users" or programs to share the expensive equipment, partitioning the resources up into time slices for each "user". As the mainframe technology became more sophisticated and the application complexity grew, multi-processing systems became more sophisticated, and the first version of the IBM Virtual Machine (VM) environment was developed and deployed in the early 1970s. Even early releases of IBM VM/370 allowed sharing of physical resources among applications and users.

Versatility and utilization is the key**.** The IBM z10 automatically allocates resources to the applications that need those resources. VMware does similar operations but takes up to 25% or more of the system resources to manage the hosted partitions. That's 25% of the resources that could be used to do useful work in the z10. The IBM z10 runs at or above 80% average utilization of system resources without jeopardizing performance, throughput, security or reliability. Partitioned systems like the Integrity Superdome typically have an average utilization rate at 25% or below, wasting expensive resources.

HP's "Fact 3": The IBM Mainframe continues to be much more expensive than open standard-based systems.

Arguing that one mainframe is more expensive than one Itanium server is a non-starter. One of those earth movers we talked about at the beginning of this discussion is more expensive than a ½ ton pickup. The discussion is not the cost of acquisition. If you need a two core Itanium server, don't buy a mainframe. If you need hundreds, or thousands Itanium servers, you might want to do the math.

In order to understand the difference in value (total cost of ownership) versus price (total cost of acquisition) it is important to understand some of the fundamental differences between the z10 architecture and interconnect and the Itanium Superdome architecture and interconnects. Latency and throughput, and thus the duty cycle or amount of work that can be accomplished, are a function of how efficient and well coupled the critical components are in an SMP server architecture. Let's start with the processor.



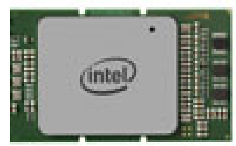


Figure 1: IBM MCM and Intel Itanium dual-core processor

In the photographs of Figure 1 above you see two different processor types. On the left is the IBM Multi-chip Module (MCM) that contains 5 quad-core processors and two cache controllers. On the right is the Intel Itanium Montecito dual-core processor. The Itanium processor includes two cores, each with 1.2MB L2 cache and 12 MB of L3 cache. There is a single interface from the two cache hierarchies through a synchronizer to the memory and I/O interface.

The IBM z10 processor is mounted on an MCM along with four other processor chips and two integrated L2 cache and cache controller chips. Each core on each quad-core processor has 3MB of dedicated L1.5 (similar to the Intel L2 cache) and a direct connection to each of the SC cache/cache controller chips on the MCM. The SC chips each have 24MB of L2 (similar to Intel L3) cache that are accessed through a cross-point interconnect from L1.5 cache from each of the cores on the MCM, plus a cross-point interconnect into three additional processor books that can scale the z10 to 80 cores. The Itanium processor plugs into a mother board that provides an interconnect between the chips for memory and I/O support, not to mention power and cooling.

HP's claim is that the Itanium processor costs $31,500 and the z10 processor costs $1,388,889. That is probably accurate for Itanium. z10 general purpose (GP) processors are not sold on a per processor cost, they are priced based on the work load to be executed in IBM MIPS (millions of instructions per second). Special purpose processors, which are GP with different firmware instantiated at system load, are priced at $125,000 per processor load. One IFL Linux processor on the z10 can execute 100's or more

Linux servers on that single core. If HP runs a single Linux server on a single Itanium core, the value difference between the two is significant. This inequity is one of many apples-to-medicine ball comparisons in this article, but let's not digresses.

The Intel Itanium processor and the mother board are connected in the HP Integrity Superdome into an enclosure called a Cell Board.

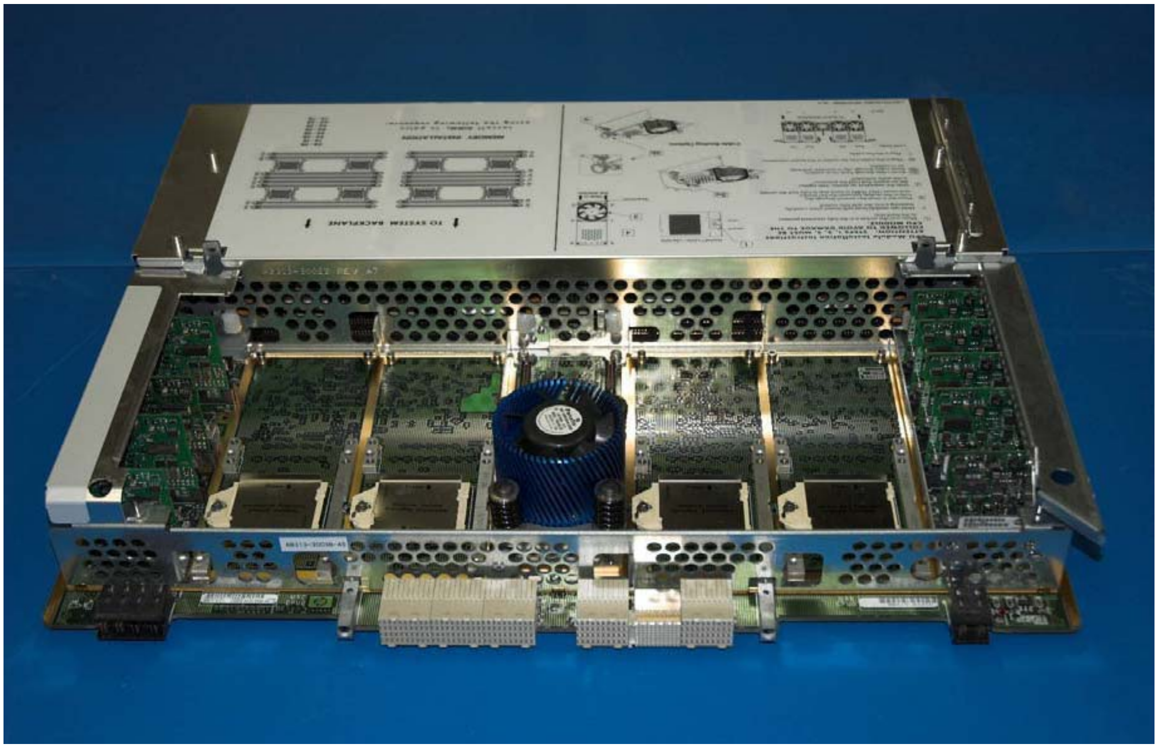


Figure 2: HP Itanium Cell Board – four socket Itanium building block

The Cell Board design shown in Figure 2 looks remarkably like the Intel 8500 North Bridge chipset design as you can see by comparing Figure 3 and Figure 4.

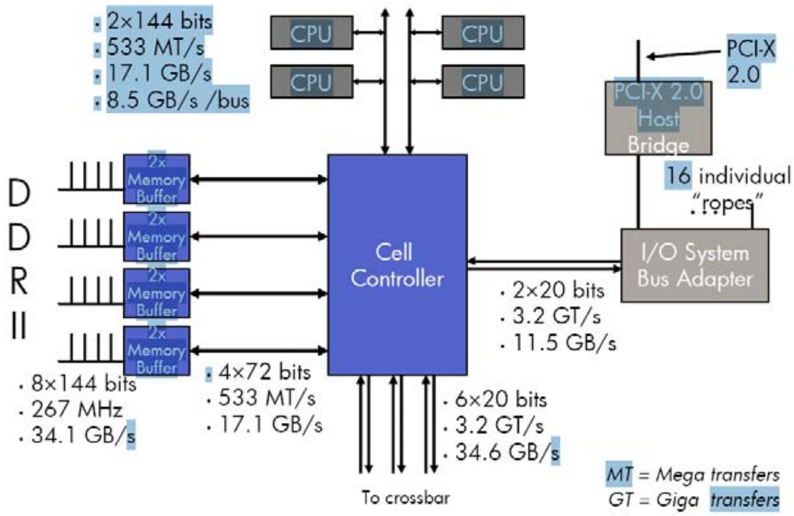


Figure 3: HP Cell Board block diagram

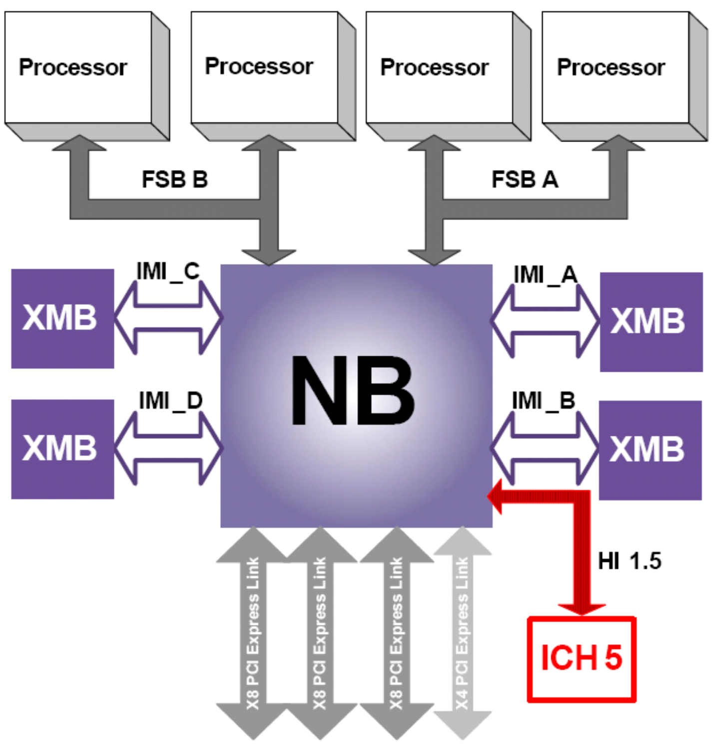


Figure 4: Intel 8500 North Bridge chipset block diagram

The 8500 North Bridge architecture is the "front-side" bus design that Intel has been using and evolving for over 10 years. The latest 7300 chipset front-side bus design from Intel is actually a generation ahead of the HP Cell Board design. Intel is abandoning this design for a new architecture for the Nehalem 5500 processor that is very similar to the architecture AMD introduced with their Opteron in 2004. The Intel front-side bus was too inefficient to compete in the microprocessor market. HP will adopt the new Intel design in the Itanium Superdome if and when Intel can stabilize the next generation of Itanium, the Tukwila, which has been delayed once again[3] until sometime in 2010. That will require a rework of the HP interconnect architecture within the Superdome.

Each four socket cell board holds up to 4 dual-core processors, and plugs into a backplane that interconnects this cell board to others through a hierarchy of crossbar interconnects shown in Figure 5.

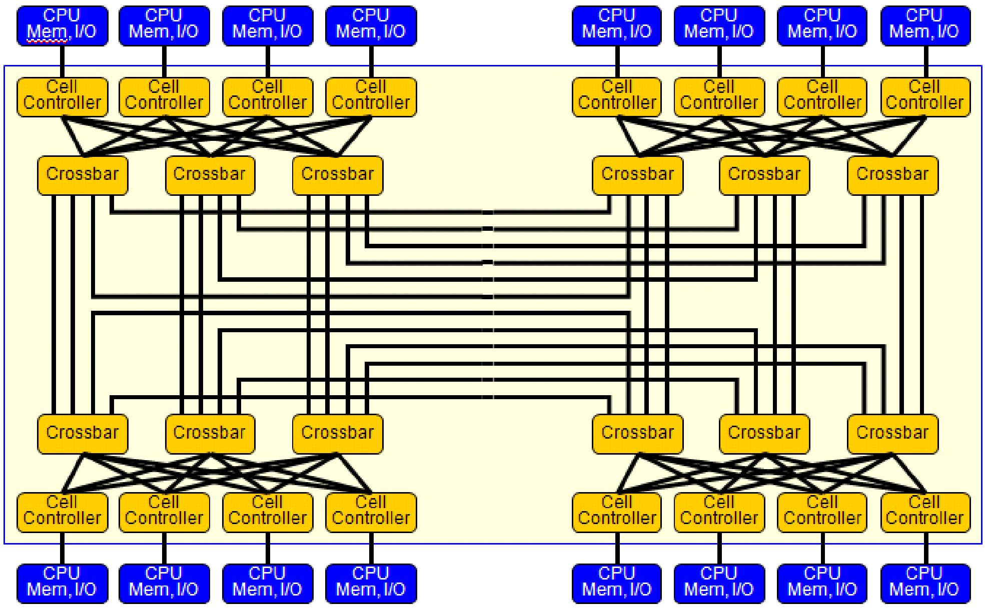


Figure 5: HP Superdome crossbar switch architecture

Each Cell Board is connected to three crossbar switches to reduce latency between the four local cell controllers. This hierarchy creates four different levels of latency depending on locality; cell board local, local switch, remote switch on a single backplane, and remote switch across backplanes. Each of these crossbar backplanes reside in externally separate cabinets. The latency difference between cell local references and remote switch across backplane references is more than 2X. The result is reduced performance per processor as the system scales from one cell board to 16 cell boards, or from 2 processors to 64.

Reliability and availability: Your 3900 Itanium servers are going to be down a lot. According to the survey conducted by the Standish Group sited in "The Real Story about the Mainframe Makeover", the Integrity Superdome is down over 11 hours per year. Maybe you should buy 4900 Itanium servers to make sure you have sufficient spares. Intel reliability might reach 3 or 4 nines (99.99%) if you have enough time, money, facilities and people to make it happen. The z10 will just almost never go down, or maybe minutes a year on average. What is the real cost of downtime? It is not the cost to

[3] http://news.cnet.com/8301-13556_3-10246293-61.html; **cnet news,** "Intel's Tukwila Slips Once Again", Gordon Haff ; May 21, 2009

repair the computer but rather the cost to repair lost customer satisfaction when your customers can't get service.

There are many reasons why z10 reliability is so good. An indication can be seen if we describe the same design and interconnect strategy for the z10 that we just covered with the HP Integrity Superdome.

IBM simplified the z10 interconnect design and integrated it into the processors and Multi-Chip Module (MCM). The same processor-to-processor interconnections that were described for the Superdome with the crossbar switch backplanes in the two enclosures has been integrated into the z10 processor and MCM implementation. That means fewer components, fewer interconnects, fewer failures. Each MCM is housed in a steel cabinet called a book, and is directly connected to memory and I/O inside the steel encased central electronic complex (CEC) without the use of a backplane. There is a significant difference in this architecture as well.

Rather than interfacing processors, and thus memory, I/O and cache, through cell controllers to crossbar switches as is the case with the Superdome, the z10 connects all five processors on an MCM directly to the cache controllers through shared L2 cache. And further connects each cache controller to six other cache controllers on three additional processor complex books. As a result, each core has direct access to every other cores L2 cache which dramatically reduces latency, increases throughput, and adds to the increased availability and resilience, even as the system is scaled to 20 quad-core processors.

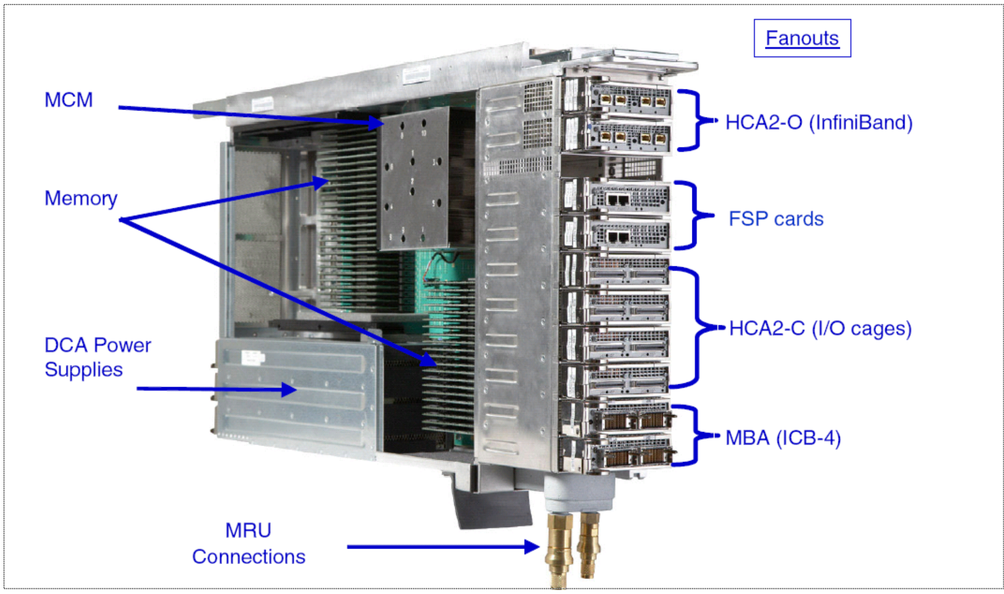


Figure 6: z10 CEC Book

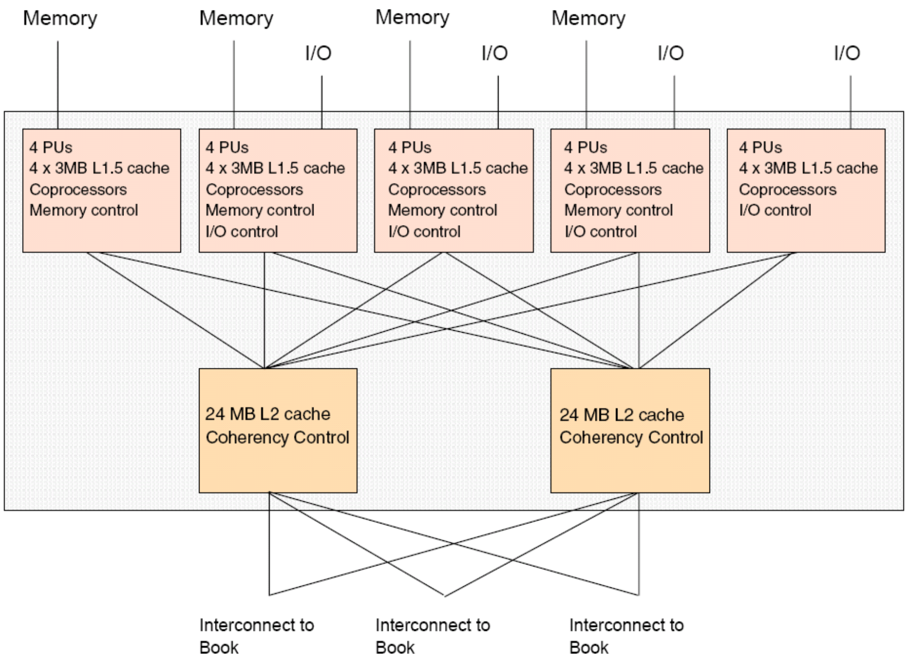


Figure 7: z10 Logical Book Structure

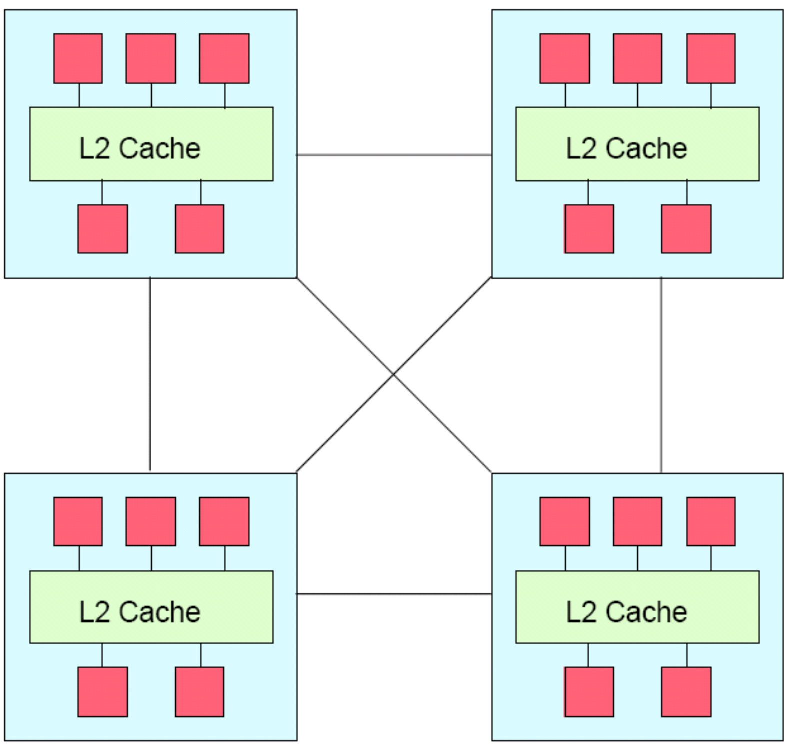


Figure 8: Point-to-point topology for book-to-book interconnect

The result is a system that scales from one to 20 processors with exceptional performance gain per processor and has fewer components, reduced complexity, lower coherence and contention, and thus lower latency and higher throughput. To measure the true value of the two systems, the comparison has to go well beyond the price of the processor.

Provisioning and scalability: What does it take to put a new Itanium server, like a Superdome, into service? Most probably it will take weeks or months? What does it take to provision a new server partition on a z10? Most probably it will take minutes. How big can a superdome get? 128 cores, with less and less performance per core as you scale. The IBM z10 scales to 64 cores, each capable of 10s of times the performance of an Itanium core. And the z10 scales with the workload so you get what you pay for. See the section on core scaling below for evidence of this phenomenon.

And one more thing, there are always two spare cores in each z10 processor book to ensure almost instant recovery from any processor failure. And in addition, each processor can be used for special purpose applications like Java, database and Linux applications with dramatically reduced costs for hardware and software. How can HP measure the real cost of the Itanium processor versus the z10 processor when a single z10 core (four per processor) can be partitioned to run 100s of Linux servers/applications for the $125,000 price that even HP quotes in their article? How many Linux servers will a single Itanium core support? ONE.

HP's "Fact 4": Expanding mainframe staff through zNextGen is falling short.

zNextGen education is much more than just learning to manage MVS (z/OS) and z10 systems. It also includes data center management, reliability criteria, system management, outage management, preventive failure strategies, data security, and data center audit ability. The universities that are participating in this program have a history of producing some of the top IT talent in the industry. Microsoft and Cisco certifications are more predominantly awarded by technical institutes and independent training organizations authorized to award them. That is not to say that either is better or worse, but only to say that education and training to run a large data center is more rigorous than learning to patch Windows and configure a network.

As the mainframe costs drop and interest in the value that the mainframe brings to the industry increases, more and more students are opting to take the mainframe path for their future career progression.

Critical skills covered in zNextGen include the requirements for data security (can you spell SOX?), disaster recovery, data integrity, data center management, SOA deployment, and software as a service.

HP makes the point that only one of these Universities is listed in the top 500 in the world. Embry-Riddle Aeronautical University in South Florida is the world's leader in aviation and aerospace education but is not among those top 500 universities. The list is built by the Center for World-Class Universities, Graduate School of Education of Shanghai Jiao Tong University and is based on Nobel Prize winners, publications in Nature and Science, and publication index for science and social science, as well as per capita academic performance. No wonder Embry-Riddle didn't make the list.

HP's "Fact 5": Mainframe's TCO comparisons vs. underutilized distributed servers

Most data centers today utilizing clusters of 100s to 1000s of distributed servers are not using VMware. Many who are using VMware are not using it for mission critical applications and data bases. As the uptake of VMware increases, the real value of z/VM will become clear. Most IT professionals, who have grown up in the Windows/UNIX/Linux distributed computing world, have not been introduced to virtualization. Most think that hardware partitioning like blades, Sun Domains or HP nPars constitute virtualization. As they learn about real virtualization like VMware, they will begin to appreciate the value that virtualization brings. When they are introduce to the real value of z/VM, running Linux, they will understand how a single z10 can take the place of 100s of Linux servers in the same data center, all in a single rack.

Even comparing HP nPars or vPars with z10 virtualization is much like comparing z10 to racks of Intel servers. In the HP case, the utilization per server running Linux or UNIX averages 25% - 50% versus 70% plus that the z10 typically provides. z/VM is true shared dynamic resource virtualization for CPUs, memory and I/O.

HP's "Fact 6": When it comes to availability HP Integrity NonStop beats IBM's Mainframe.

HP Integrity NonStop, originally a MIPS microprocessor platform from Tandem Computer, has always been a significant platform for fault-tolerant computing. It was one of a class of other very expensive fault-tolerant systems like Gould, Convex and Stratus. Where are they now? And what happened to them? The IBM mainframe gained more and more reliability at a much lower price and has taken market share from these specialized architectures. As the IBM mainframe became more reliable, ISVs that required extreme reliability stopped developing for Tandem and began moving to other platforms, many to MVS and now z/OS on the mainframe. The result is a very limited number of applications running native on Integrity NonStop.

NonStop has its own proprietary database and transaction processing middleware. The connection to NonStop is primarily through Tuxedo transaction processing or ODBC database messaging. It is still a fault-tolerant system, but the cost and lack of applications make it less than attractive for most extreme reliability ISV platforms. Integrity NonStop systems typically run a single time-critical application in an isolated environment designed specifically for that application. Again this is not good use of an expensive resource. And notice that there was no cost of acquisition comparisons from HP for the Integrity NonStop.

And isn't it interesting that up through Fact 5 the HP comparisons were done with the Integrity Superdome. Once reliability and availability became the issue HP switches to the Integrity Non-Stop. The case study that HP sites in this article also makes it clear that the z10 is more reliable than the Integrity Superdome so HP had to change horses.

Nuts and Bolts about performance:

Let's talk about performance. In "Fact 5": HP mentions that the IBM claim of thousands of Linux servers on the mainframe cannot be substantiated because IBM does not publish industry standard benchmarks on the mainframe. As anyone in the benchmarking business will tell you, the main difference between two benchmarks is the skill and experience of the person doing the benchmark. If two benchmarks are within 10% of each other then one of the techs probably had more coffee that morning. On the other hand, if they are 100% apart, there is a technological advantage.

A brief discussion about benchmarks. Each benchmark is designed to exercise some part of the computer architecture. Very few benchmarks, aside from running customer code with customer data, actually represent the true application profile that a customer will deploy. Unfortunately there is seldom a footnote on a benchmark report describing what parts of the architecture are actually being tested for each benchmark.

As an example, SPEC CINT2006, the integer component of SPEC CPU2006, is a series of nine C language applications and three C++ applications that are all CPU intensive. CINT2006 claims to exercise the CPU, cache and memory, but in truth the cache is so large in most processors today that the much of the benchmark can be loaded in cache and run from there. The benchmark is run as fast as possible for a period of time and the result are the number of executions. SPEC CINT2006 does not exercise I/O, processor interconnect for SMP applications nor non-uniform memory access (NUMA) across a multi-processing environment so prevalent today. However, if you have a CPU/cache intensive application the CINT2006 might be a good indicator of performance.

Each computer architecture is designed to support a specific application profile that benefits from the system design. In high performance computing (HPC), for example, single-instruction multiple-data (SIMD) applications benefit from very fast CPU-to-cache and memory-to-cache access, with no interaction between CPUs or alternate cache. High throughput transaction processing applications exercise SMP capabilities of computer architectures so cache coherency, cache-to-cache access, memory-to-cache and interconnect contention are critical to optimizing throughput and low latency between transactions.

The current Intel architectures are better fit to HPC, while the mainframe is ideally suited to high throughput transaction processing. Comparing these two architectures with something like a SPECjbb2005 server side Java benchmark designed to exercise the JVM, JIT compiler, garbage collection and thread efficiency would not be fair to either server. SPECjbb2005 is a CPU/cache intensive benchmark that exercises no I/O. Few Java applications are designed to execute HPC SIMD profiles. Again it is hardly a fair or accurate indication of server performance for a customer application.

IBM does not run industry standard benchmarks because they are not representative of the applications that are typically deployed on the z10. It would be fun for IBM to run the SPEC CINT2006 benchmark on the z10 with its 4.4GHz CPUs and 3MB/48MB cache hierarchy but it would be counterproductive and not useful for anything that customers want to do on the z10.

The danger of industry standard benchmarks: Let's look at an industry benchmark that appears to be well suited for the HP Linux/UNIX Superdome and see how it fares. The one benchmark that is an almost certainty for every vendor to run is SAP SD 2-tier. This is an SAP standard benchmark, not industry standard, and SAP uses this benchmark to size hardware requirements for implementation of SAP applications. If a vendor wants to propose their hardware for an SAP implementation it has to have a benchmark point on this test. As a result, most vendors run and publish this benchmark for their hardware.

There are some caveats. Each is run with a specific database, and comparing an Oracle or DB2 implementation to a Microsoft SQL Server database is hardly a fair comparison. Each is run with a specific OS and again comparing Linux/UNIX with Windows is not a fair comparison. And last but not least each benchmark is run with a specific version of the SAP software, from 4.6 to ECC 6.0. These are dramatically different implementations of SAP and should never be compared. It is never easy to line up multiple published results of even this SAP standard benchmark unless every aspect is aligned and relatively equivalent. Let's look at the HP SAP SD 2-tier published results as an example.

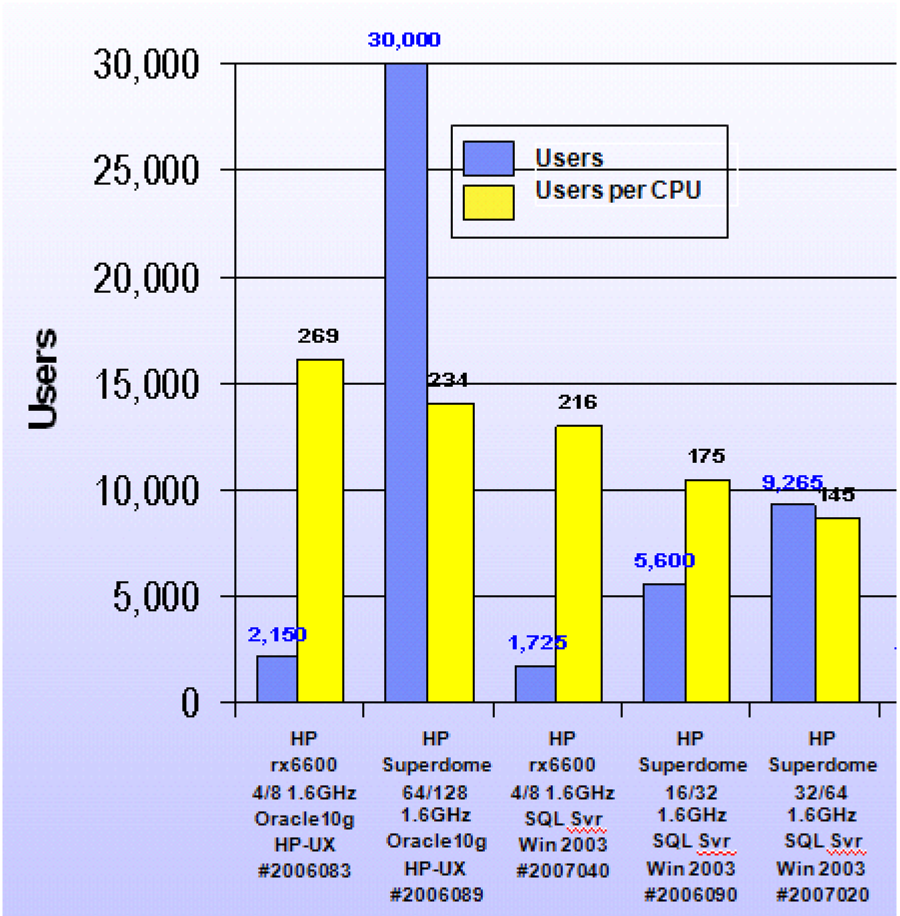


FIGURE 9: SAP SD 2-tier results published by HP[4]

There are some interesting things to notice on Figure 9. In each case you will note that the more cores in the system, the fewer SAP users per core. That is a result of the interconnect architecture in the HP Superdome that we pointed out in the previous discussion of the hierarchical crossbar switch. As more processors connect to more switches across more interconnects the latency increases by more than a factor of two.

The expert benchmark folks at HP have done a wonderful job of localizing SAP modules on this benchmark to reduce the impact of that increased latency and to reduce the penalty of SMP scaling. The odds of a customer implementation of SAP achieving this level of localization are limited. But even with the best that HP can present they lose 13% over the 8 core rx6600 HP-UX implementation scaling to 128 cores. HP is not as fortunate with the Windows 2003 implementation. Scaling from 8 cores to just 64 cores is a loss of 33%, again with the best benchmark skills available to them.

This brings up another interesting point. At the HP Technology Forum in Las Vegas recently is was reported that HP was handing out a "wheel-like device that allows one to

[4] http://www.sap.com/solutions/benchmark/sd.epx

see alleged savings in various categories for migrating from the mainframe to HP Integrity servers."[5] It was reported that this comparison pitted an IBM z9 EC 2094 Model 720 with two HP Integrity rx8640 and two rx7640s running Microsoft Windows Server 2003. It is nteresting that the z9 comparison is one generation behind the current z10, which is reportedly 2x the performance at the same price points as the z9, e.g. 4.4GHz quad-core processors in the z10 versus 1.7GHz dual-core processors in the z9. And also it is interesting to note that HP would pit their finest Itanium against the mainframe with Itanium sporting Windows 2003 and SQL Server.

As you can see from the chart in Figure 9, comparing the rx6600 running HP-UX and Oracle 10g versus the same server running Windows 2003 and SQL Server, the Windows implementation gives up a 20% performance disadvantage in the Intel sweet spot of four sockets. The rx6600 is a four socket single cell board design. As the Windows servers are scaled to the Integrity Superdome the penalty of Windows on the crossbar interconnect is significantly greater than that of HP-UX. 64 cores of Itanium Superdome with Windows 2003 support 9265 SAP Users, versus 128 cores on HP-UX at 30,000. There might be some marketing trickery going on here to reduce the mainframe advantage and increase the cost difference. That's the advantage of Windows.

Another interesting point about this comparison is that the HP servers are assumed to be running only a single instance of SAP, perfectly configured to fit the multiple Integrity servers. In an actual implementation of SAP, varied workloads will exist in a production environment. The utilization of resources in the HP servers will suffer when the workloads vary. This is typical in comparing benchmark profiles to actual production workloads. The z10 was designed specifically to accommodate variable workloads, utilizing shared resources across production workloads as they change dynamically. That capability will most likely never be designed into the HP Intel based server offerings. It certainly is not there today.

A few closing words about reliability, availability and security:

This paper has focused on the advantages of the IBM z10 as a transaction processing server for heavy transaction workloads with exceptional scaling. Reliability and availability have been touched on but only lightly. Security will not even be broached. Reliability is a topic with many facets. The biggest difference between reliability availability and serviceability (RAS) in the z10 mainframe versus what midrange and departmental server vendors call RAS is really the philosophy of RAS, and thus design points.

The goal of midrange and departmental servers, UNIX, Linux and Windows primarily, is rapid recovery from failure, hot plug replacements, and reduced downtime. The mainframe philosophy is to keep the application running at all cost: even if hardware failure or software failure occurs, maintain sufficient redundancy to ensure that resources

[5] http://itknowledgeexchange.techtarget.com/mainframe-blog/hp-pushes-mainframe-migration-hard-at-tech-forum/

can be launched or repurposed to keep the most mission critical applications running. Every subsystem within the z10 is designed for continuous operations, and for continuous service. The z10 can be segmented and portions taken off line for periodic maintenance and recovery without affecting any operating application.

In fairness, the one UNIX/Linux system that is designed with similar criteria is the IBM Power AIX servers. The reason that they are designed similar to the mainframe is that the original design of Power and AIX included mainframe engineers, developers and systems expertise. IBM continues to add "mainframe technology" to Power development.

The IBM z10 is truly not your father's mainframe. And it is hardly fair to expect the HP Integrity Superdome or any other commodity based system architecture to compare, no matter how much spin is put on the discussion. This is another good lesson: we should only believe part of what we read, some of what we see, and only half of what we think is true. Check it out! Your father knows about mainframes; your son will probably never hear of the Itanium.

1

References:

i. IBM z10 Enterprise Class Technical Guide Redbook, http://www.redbooks.ibm.com/abstracts/sg247516.html
ii. Meet the HP Integrity Superdome, http://h21007.www2.hp.com/portal/download/files/unprot/superdomejan05.pdf
iii. Intel 8500 Chipset Northbridge, http://h21007.www2.hp.com/portal/download/files/unprot/superdomejan05.pdf
iv. SPECInt benchmark information can be found on the SPEC website, http://www.spec.org/cpu2006/CINT2006/
v. Joseph Temple and Roger Rogers , Relative Capacity and Fit for Purpose Platform Selection, Journal of Computer Resource Management #123 March of 2009, pp 20-41