# Exploiting Virtualization on eServer p5

*by N. Dai Vu*
*Product Marketing Manager, eServer® p5 Virtualization*
*IBM Corporation*

**Challenge of Today's IT Organizations**
Traditionally, most applications have been developed to support a specific business area or function. These applications generally run on their own dedicated physical servers and storage devices. There is very little, if any, sharing of IT resources between different applications. This has led to fragmented infrastructures and often results in a lack of flexibility within an organization's IT environment. Companies struggle with servers that have low utilization rates, as the unused system resources sit idle and wasted. Costs for today's implementations are driven even higher as IT departments overprovision systems to ensure that processing power is available to meet demands for the specific business area being supported. The related operational costs continue to rise as the number of physical resources and complexity continues to grow as shown in Figure 1.

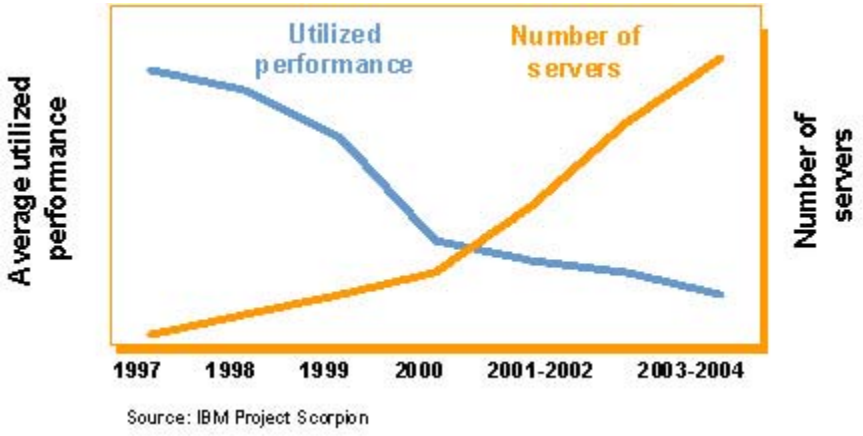


**Figure 1: Estimated server utilization and proliferation history[5]**

**Simplifying the IT Infrastructure with Virtualization Technologies**
Over the past few years many IT organizations have pursued physical/centralized systems consolidation -- taking several instances of servers/storage and consolidating them into larger systems to reduce complexity, improve manageability, and lower the total cost of ownership (TCO).

Virtualization-- the abstraction of physical properties of hardware resources (e.g., microprocessors, memory, I/O devices, or storage) in a way that allows a more flexible usage model -- is the logical next step to systems consolidation. Fine grain virtualization permits near instantaneous matching of workload to resources allocated, avoiding the wasted resources common to the "single application per server" model of computing. Furthermore, virtualization is designed to help clients with cost savings and higher return on investment (ROI) by improving efficiency and resiliency, easing integration (a large component of IT costs), and enabling dynamic provisioning of resources.

**Virtualization on eServer p5 with POWER5 processors**
IBM has been delivering systems (with mainframes and now, eServer zSeries) with virtualization technologies for many years. With the recent introduction of the IBM eServer p5 with POWER5™ processors, mainframe-inspired virtualization technologies have arrived for the UNIX® world. These virtualization technologies are enabled by POWER5 processors and the Advanced POWER Virtualization (APV) feature on eServer p5 systems.[1]

The key POWER5 technologies include:

> ***Simultaneous Multi-threading*** allows two instruction paths to share access to the POWER5 execution units on every clock cycle. Each instruction path is abstracted, by the operating system, so that each POWER5 virtual processor will appear as two logical POWER5 processors, one for each instruction path. With no modifications to the applications, simultaneous multithreading will enable SMP-scalable applications to benefit from significantly increased system level performance.[2] This is demonstrated in Figure 2.

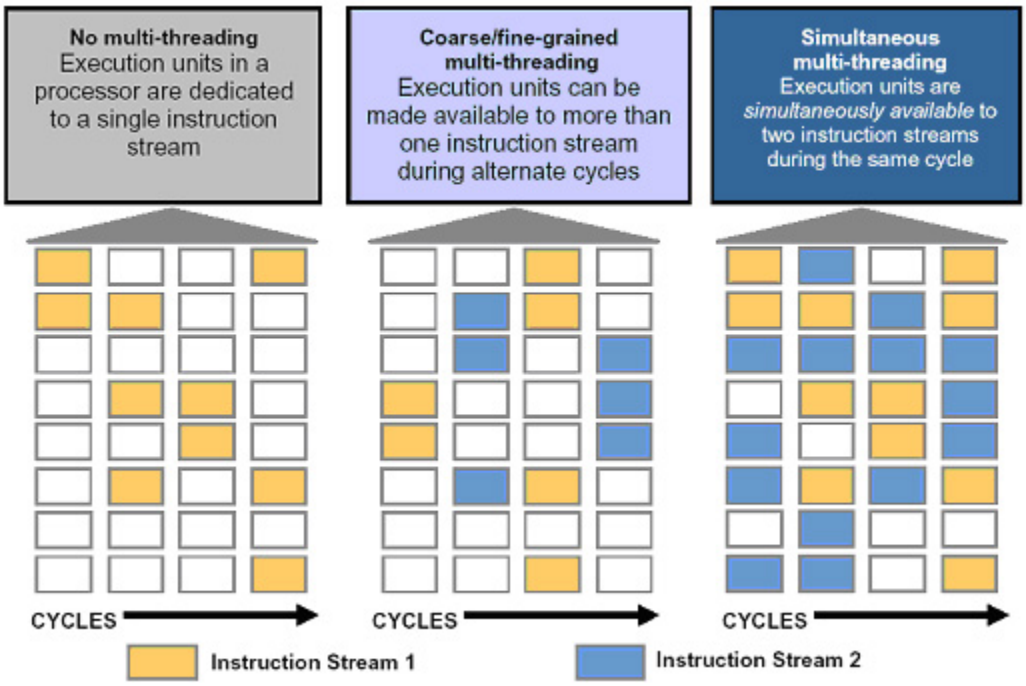


**Figure 2: Improving processor utilization with Simultaneous Multi-Threading on POWER5**

***Design advantages of POWER5:*** Random or unpredictable transaction behavior will have low cache affinity (i.e., data and instructions required for transaction execution not found in high performance cache) and puts greater stress on the memory subsystem to deliver good application performance and transaction response time. POWER5 has key *architectural* advantage:

- POWER5 has a rich cache hierarchy (64/32KB L1; 1.9MB L2; 36MB L3) which provides extremely high probability of cache affinity
- The POWER5 memory controller is resident on the POWER5 chip. A fetch for data or instructions can be executed every second clock cycle of the processor. A memory fetch is issued at the same time the L3 cache is being checked which minimizes the latency of memory access
- Cache and memory performance scale directly with POWER5 processor clock speed
- POWER5 has a very robust data delivery interconnect. IBM has a long history of awareness that the server's effectiveness at getting the right data and instructions to the executions units is the key to overall throughput.

The key APV technologies include:

***Micro-Partitioning™ Technology and Shared Processor LPARs*** – Shared processor logical partitioning (LPAR) allows customers to "slice up" a machine into virtual partitions and provides the flexibility to dynamically change the allocation of system resources for those environments. Micro-Partitioning technology on eServer p5 provides the capability to create multiple virtual partitions *within a processor*-- to a granularity of 1/10$^{th}$ of a CPU. Spare capacity can be re-allocated to virtual partitions in granularity 1/100$^{th}$ of a CPU within 10ms. Any of the virtual servers may run on any of the physical processors, meaning that the processor resources are fully shared, which makes it possible to run the physical server at very high utilization levels.

***Multiple operating system (OS) support***[3] is available on eServer p5 products. Support includes: IBM AIX™ 5L Version 5.2, IBM AIX 5L Version 5.3, SUSE Linux ® Enterprise Server 8, (SLES8), Red Hat Enterprise Linux 3 (RHEL3), and (planned) i5/OS™ operating system.

***Virtual LAN and Shared Ethernet Adapter*** allows clients to create virtual Ethernet connections to provide high-speed inter-partition communication between logical partitions on an eServer p5 server without the need for network I/O adapters and switches. Connectivity outside of the eServer p5 server can be achieved using an I/O server partition that acts as an IP forwarder to the Local Area Network through an Ethernet I/O adapter.

***Virtual I/O (Disk Sharing)*** provides virtual storage capability to partitions on an eServer p5 system. It allows a physical adapter with attached disks on a special "hosting partition" to be shared by one or more "client" partitions, enabling

customers to consolidate and potentially minimize the number of physical adapters. To a partition/OS, Virtual I/O is practically indistinguishable from physical I/O.

***Partition Load Manager (PLM)*** provides automated processor and memory distribution between dynamic LPARs and Micro-Partition-capable LPARs running AIX 5L. The PLM application is based on a client/server model for the sharing of system information, such as processor or memory events, across concurrent present LPARs.

***Capacity on Demand (CoD)*** features optionally available on select eServer p5 systems, enable clients to respond quickly to their workload demand by permanently or temporarily activating processors or memory when required. One form of temporary activation, Reserve CuOD,[4] enables users to add reserve processor capacity to the shared processor pool when the base shared pool capacity is exceeded.

Figure 3 below shows how POWER5 and APV technologies come together to create an extremely powerful and flexible resource model for the IT organization.

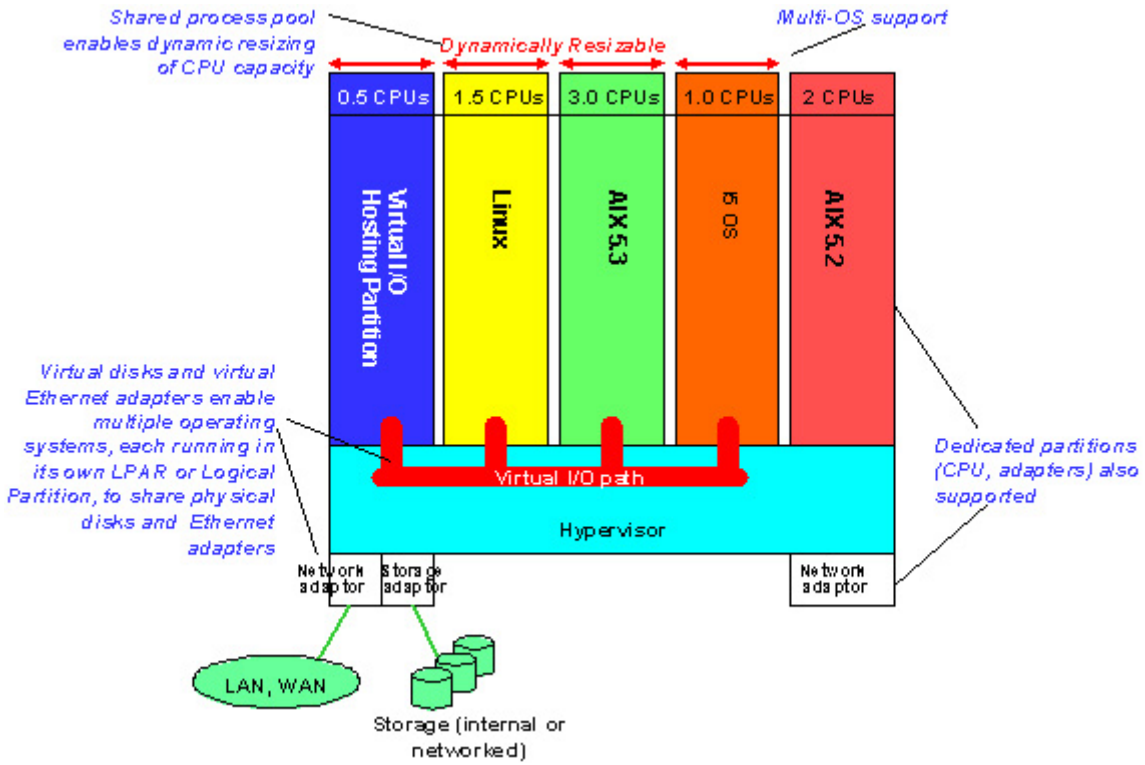


**Figure 3: Virtualization technologies on eServer p5 servers**

**Representative scenarios for customers**

With POWER5 processor-based eServer p5 systems and virtualization, there are many opportunities for consolidation and simplification. Broadly speaking, there are two major value propositions: 1) rapid response to changing workload profiles with shared processor pools and therefore, enabling a more flexible/agile organization; and 2) improved utilization/TCO by consolidating many small underutilized servers using multiple sub-CPU Micro-partition LPARs and shared Ethernet and Fiber Channel Adapters and Virtualized disks. Here are a few representative scenarios:

- ***Server consolidation:*** Large number of smaller, under-utilized (e.g., 10%), existing server systems are consolidated on a single LPAR-capable system. The Micro-Partitioning feature is particularly attractive in this environment when a fraction of a single POWER5 processor provides equivalent processing capacity as the existing server. Micro-Partitioning technology enables hundreds of these smaller existing servers (e.g., file, print, web) to be replaced by a single POWER5 server.
- ***Virtual blade servers:*** Micro-Partitioning technology enables hundreds of "low cost" individual virtual blade partitions to be defined that mimic a blade server environment. In a blade environment, individual blades must have enough capacity to handle bursts of activity (Web hits), but most blades, in general, are grossly underutilized. Micro-Partitioning features and VLAN are natural fits in this environment, because the idle time of a virtual blade can be utilized by another virtual blade partition and VLANs provide low-cost high-speed communication vehicles for virtual blade servers on the same POWER5 server.

- ***Production and batch/test systems:*** Micro-Partitioning technology provides an optimal environment for the co-existence of production and batch/test systems. Production partitions can be defined with fixed performance requirements such that they receive the processor capacity they require on demand. Batch/test partitions can be defined with minimal resource commitment but with the ability to soak up spare cycles as shown in Figure 4.

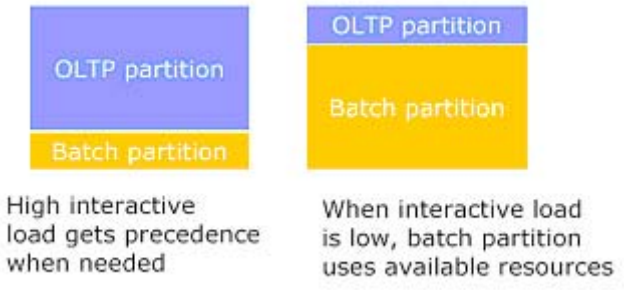


**Figure 4: Batch partition ceded capacity to OLTP when required**

- ***Overlapping production systems:*** This is an environment where system performance is critical, but the workloads of different servers are such that the peaks in demand from one server overlap the valley of demands from another. To some degree this environment can be serviced with dynamic LPAR, however, Micro-Partitioning technology provides a finer grain capability which is much more instantaneous.

**Summary**

IBM's POWER5 processor-based server offerings in the IBM eServer p5 UNIX servers offer many technologies to enable simplification of IT infrastructure. The flexibility of running multiple operating systems on a single POWER5 processor-based server, coupled with the high performance scale-up capabilities of the POWER5 processor, further enhanced with simultaneous multi-threading and the fine grained granularity of Micro-Partitioning features allow clients to consolidate as many disparate applications as possible onto each POWER5 processor-based server. As more applications get consolidated, the POWER5 Virtual LAN and Virtual I/O capabilities can be employed to integrate the surrounding storage and network resources, for each application, into the POWER5 server.

**Notes**:
1. Advanced POWER Virtualization (APV) is an optionally priced feature on eServer p5.
2. Performance improvement from SMT can vary; SMT can be disabled.
3. eServer p5 supports AIX 5L Version 5.2; however to exploit SMT and many APV features, AIX 5L Version 5.3 is required; IBM is planning to make i5/OS support available for select eServer p5 models for a limited number of CPUs on those systems. All statements regarding IBM future directions and intent are subject to change or withdrawal without notice and represent goals and objectives only.
4. Reserve CuOD requires AIX 5L Version5.3 and the APV feature.
5. More information on IBM Project Scorpion can be found at http://www.ibm.com/servers/library/pdf/scorpion.pdf

---