



Power.org™ White Paper

What's New in the Server Environment of Power ISA v2.06?

Version 1.0 – January 18, 2010

Satish Kumar Sadasivam, Staff Software Engineer, IBM - India

Prathiba Kumar, System Software Engineer, IBM – India

Varun Mallikarjunan, Software Engineer, IBM - India

© Copyright 2010 Power.org. All rights reserved

The Power Architecture and Power.org word marks and the Power and Power.org logos and related marks are trademarks and service marks licensed by Power.org.

Abstract

Additions and Enhancements to the Power ISA v2.06 extend the advantages of the Power Architecture ranging from performance, flexibility and RAS to energy efficiency. Vector-Scalar Extension (VSX) – a major addition in v2.06 extends the edge Power ISA has in HPC and computation-intensive workloads. Power ISA v2.06 continues to provide enhancements to the Server space such as Memory Management, Processor version compatibility features, cache management, etc and also introduces a number of capabilities for the embedded space such as embedded hypervisor, energy management, multi-core and multi-threading.

We will highlight the major changes and their merits in this short paper. The focus here is on the Server ISA.

Contents

Introduction.....	4
Vector-Scalar Extension (VSX).....	4
Processor Compatibility Register.....	6
Extensions to Virtual Page Class Key Protection Mechanism.....	7
Multiple Page-Size Segments.....	7
Cache Management Enhancements.....	8
Strong Access Ordering.....	9
Decimal Floating Architecture Extensions.....	10
IBM POWER7.....	10
Summary.....	10
References.....	11
About the Authors.....	11

Introduction

The Power ISA evolved out of IBM's POWER and later PowerPC architectures. The first version of the Power ISA – v2.03, taken largely from PowerPC v2.02 and PowerPC Book E v1.9, was released in 2006. Equipped with robust RAS, virtualization capabilities, advanced dynamic Memory Management and consistent floating-point architecture, Power ISA stands differentiated from its competitors in the market place. The Single Instruction Multiple Data (SIMD) processing capabilities with the Vector processor and the new Vector-Scalar Extension (VSX) provide extreme flexibility to program for performance. Multiple Page Size Segments (MPSS) provide for efficiency in memory management and enhance performance. These and other enhancements are outlined below.

This white paper presents an overview of the additions and enhancements in the Power Server ISA and highlights the benefits of these changes.

Vector-Scalar Extension (VSX)

Vector Scalar Extension to the ISA has introduced Vector and Scalar Binary Floating Point Operations conforming to the IEEE-754 Standard for Floating Point Arithmetic. The introduction of VSX in the Power Architecture increases the Parallelism by providing Single Instruction Multiple Data (SIMD) execution functionality for floating point double precision to improve the performance of the HPC applications.

Register File:

A 64-entry Unified Register File is shared across Vector Scalar Extension, the Binary floating point unit (BFP), Vector Multimedia Extension (VMX) and the Decimal Floating point unit (DFP). The 32 FPR which are used by the BFP and DFP unit are mapped to the register 0 to 31 of the Vector Scalar Registers, The 32 VR which are used by the VMX are mapped to the register 32-63 of the VSRs.

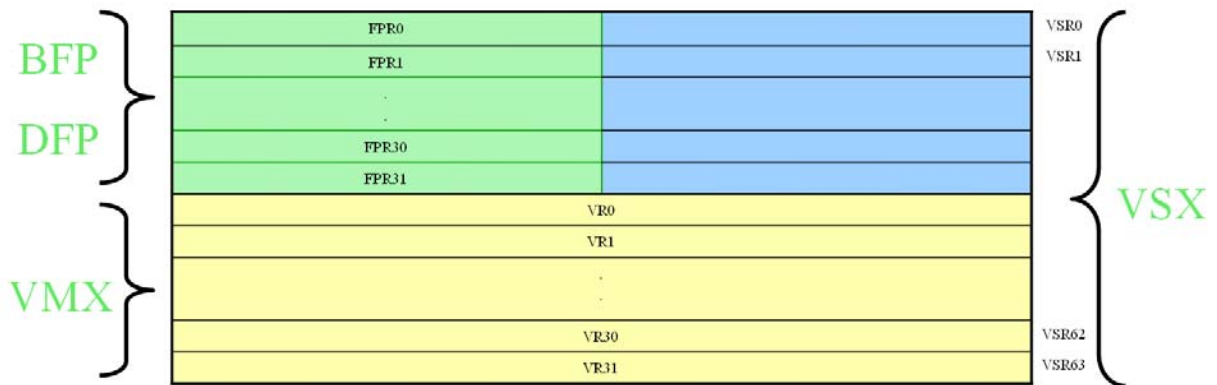


Figure 1. Unified Register File

The Unified Register File facility (128 bit by 64):

- Saves space and power on the chip.
- Eliminates need for interchange of data between scalar and vector registers.

The Vector Scalar Extension has multiple advantages:

- Increased performance for HPC Application doubling the Max FLOPS rate compared to V2.05 specifications by providing IEEE Standard Vector Double Precision Floating Point operations.
- Increase in number of registers for the Vector Floating Point operations can support aggressive compiler optimization for both Scalar and Vector operations. This helps increase in the performance with reduced power consumption because of sharing of register file across various execution units while improving the performance/watt ratio of the system.
- Providing Non-IEEE Mode support which is permitted to produce results not in strict compliance with IEEE standard largely reduces the latency of the instruction especially for divide and sqrt type of operations.

Vector Scalar Extension supports Double Precision Scalar and Vector Operations and Single Precision Vector Operations. VSX instructions numbering 142 are broadly divided into two categories which can operate on 64 vector scalar registers.

- **Computational instructions:** addition, subtraction, multiplication, division, extracting the square root, rounding, conversion, comparison and combinations of these operations.
- **Non-computational instructions:** loads/stores, moves, select values, etc.

The detailed description of the instructions can be found in the Book I of the Power Server ISA V2.06.

Processor Compatibility Register Support

Processor Compatibility Register (PCR) has been modified to allow turning on/off all the features added to v2.06 (as well as to selectively control the availability of specific facilities like VMX and VSX.) This allows the processor to run in a v2.05 compatible mode and enables hosting v2.05 and v2.06 partitions simultaneously on v2.06 platform. This facility benefits enterprise systems and servers by opening up new options for increased flexibility and RAS:

- Facilitates Heterogeneous Partition Mobility i.e., from v2.05 server to a v2.06 server and vice versa.
- Provides for load balancing across v2.05 and v2.06 servers.
- Utilization of v2.05 servers for v2.06 server maintenance windows.

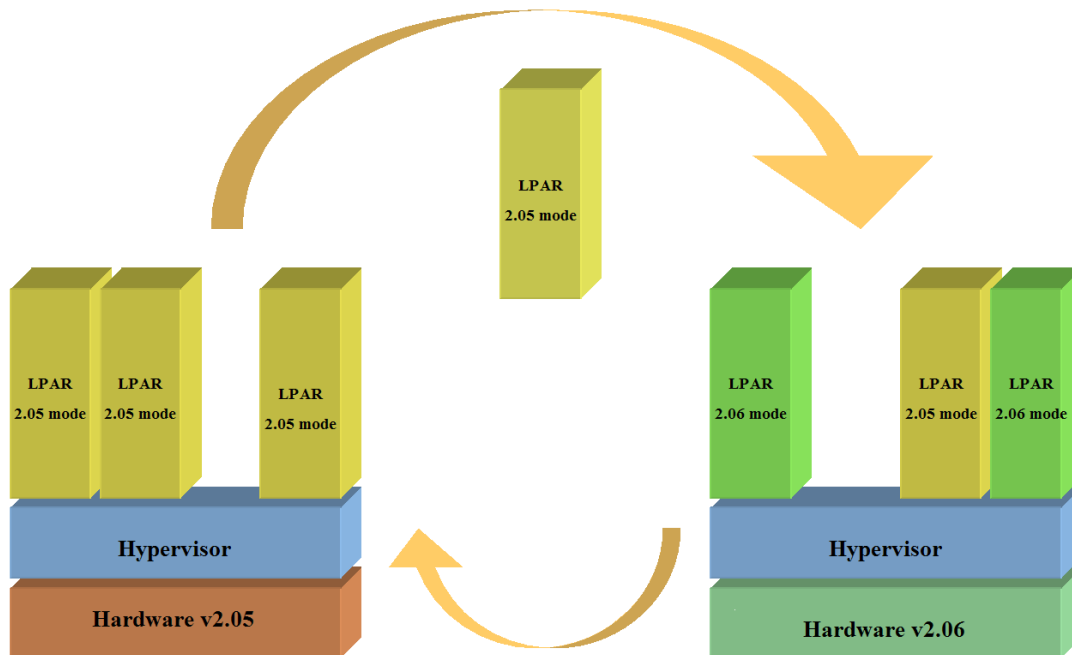


Figure 2. Heterogeneous Partition Mobility

Extensions to Virtual Page Class Key Protection Mechanism

Virtual Page Class Key mechanism has been extended to allow operating systems and application programs to appropriately use the Authority Mask Register (AMR) as opposed to only the Hypervisor servicing the accesses. Two new registers have been added that behave as masks indicating the permissions to modify AMR allowed to OS (privileged programs) and user applications.

- **Authority Mask Override Register (AMOR):** specifies AMR bits allowed to be modified by the OS.
- **User Authority Mask Override Register (UAMOR):** specifies AMR bits allowed to be modified by the user applications.

AMR has also been assigned a SPR number accessible by user applications to allow leveraging of this facility.

AMOR facilitates sharing of privileges between a client partition and a lightweight partition (adjunct partition) that services the client partitions by hosting device drivers, etc. A virtual class used by the adjunct partition can be masked from being modified by the client partition using the AMOR and vice versa.

UAMOR permits the user applications to change protections to the virtual class it is allowed to make changes to without issuing system calls or hcalls.

Multiple Page Size Segment Extension

Multiple Page Size Segment provides the flexibility of mixing pages of various sizes in one effective segment. MPSS introduced in v2.04 and extended to include more combination of page sizes in v2.06 is a feature with far-reaching benefits in database applications and applications with huge datasets.

Large pages provide multiple technical advantages:

- **Reduced Page Faults and TLB Misses:** A single large page being constantly referenced remains in memory. Eliminates the possibility of several small pages often being swapped out.
- **Unhindered Data Prefetching:** A large page enables unhindered data prefetch (which is constrained by page boundaries).
- **Increased TLB Reach:** Saves space on the TLB by holding one translation entry instead of n entries.

MPSS allows data/instructions that are densely referenced in a segment to be referenced as a single large page and sparsely referenced to be held in small-size pages. Memory Management can be programmed to provide extreme flexibility by dynamic reconfiguration of page sizes in a segment as the application behavior changes.

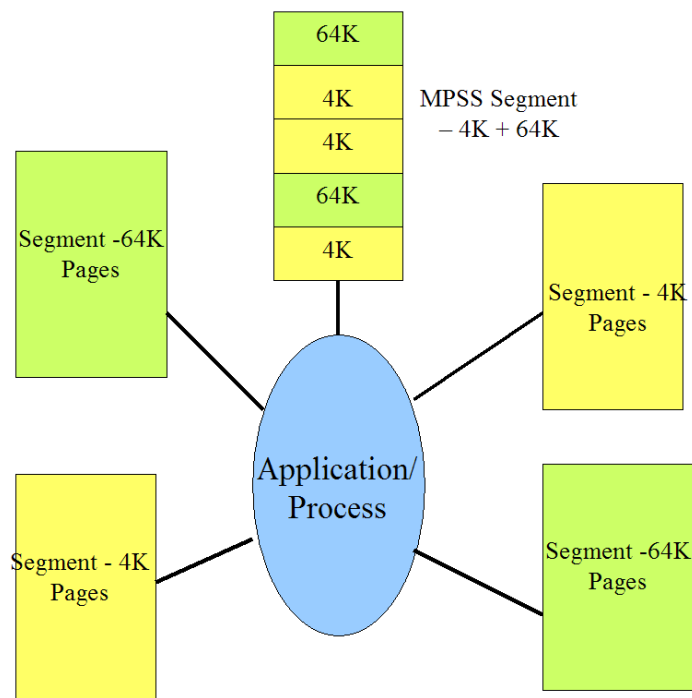


Figure 3. MPSS Illustration

Cache management enhancements

Data Cache Block Touch (dcbt) and **Data Cache Block Touch for Store (dcbtst)** are instructions that can be used by the program to provide hints to the processor that a specified address/stream of data will soon be required. This allows prefetching of data to the cache reducing/eliminating stalls in the processor waiting for the data.

The requirements for data prefetch can be varied. *dcbt* and *dcbtst* now allow specifying various new parameters for pre-fetching. Two new enhancements introduced in Power ISA v2.06 are highly beneficial to HPC applications that do not allow for even minimal performance penalty and are especially useful in cache-constrained environments.

Data Cache Block Touch – Transient: Data which might be accessed soon but only for a short period of time can be indicated by a special encoding in the *dcbt* instruction. This

helps the processor avoid replacing data that is non-transient i.e., often referenced, from being replaced in anticipation of the data that is expected to be used only for a short interval of time.

Software Initiated Stride-N Prefetching: Data elements to be referenced might be placed at regular displacements called **stride**. Data stream will be required to be prefetched accordingly. The *dcbt/dcbtst* instructions now allow specifying the *stride* and the location of the first element precisely.

An example of the usage of the Stride-N Prefetching: Let us consider an array of data structures and the *i* th element in each structure is to be accessed. Now by can specify the location of the first *i* th element in the block and the size of the structure as the stride precise prefetching can be enabled.

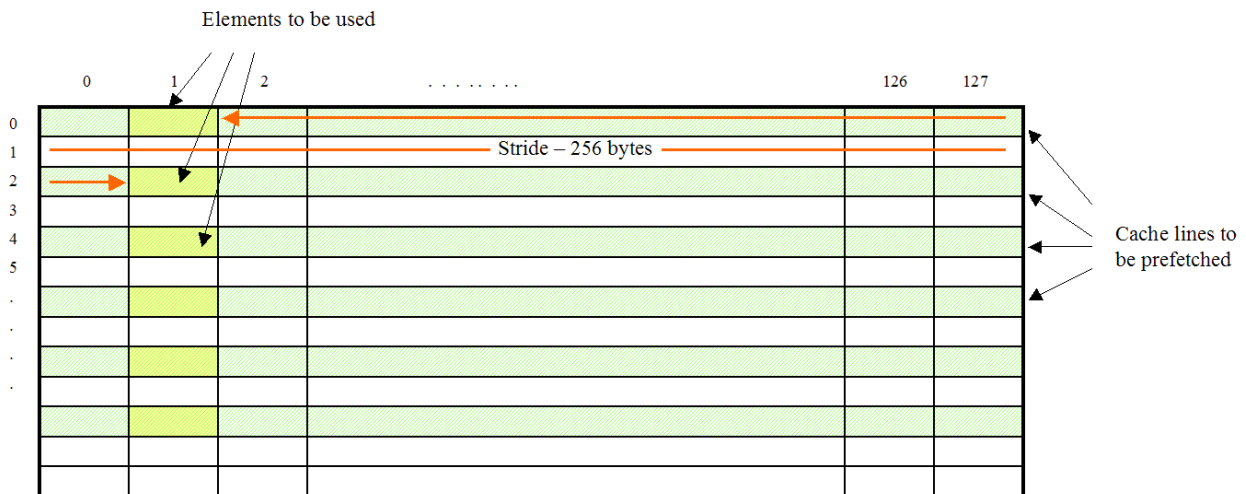


Figure 4. Stride-N Prefetching Illustration

Strong Access Ordering

- Power ISA has introduced a new model for storage accesses – Strong Access Ordering in addition to the Weakly Consistent Model, which primarily allows seamless portability of code to the Power Platform and additionally, supports emulation of SPARC and x86 applications.
- In Weakly Consistent Model, a few mechanisms of combining/reordering accesses maybe performed that can lead to increased performance:
 - Combining loads (or stores) to adjacent memory locations into a single access.

- Load maybe performed before a preceding store (provided the addresses do NOT match).
- In a uni-processor environment it creates no issues, whereas in a multi-processor environment, when a processor loads a datum it previously stored as described above, the load may be performed by the processor before the preceding store has been performed with respect to the other processors. This causes programs running on different processors to see a different interleaving of individual sequences and hence may require memory barrier instructions. Strong Access Ordering provides a mechanism to establish the consistency without using the memory barrier instructions.

Decimal Floating Architecture (DFP) Extensions

Power Architecture provides Decimal Floating Point Architecture support in compliance with IEEE P754 Standard. The introduction of DFP provides extensive support for a wide range of applications in the Banking and Financial sector that greatly improves the performance of the server and helps generate more accurate results. Introduction of DFP in the Power Architecture acts as one of the major differentiators in comparison with other industry standard architectures.

Power Server ISA V2.06 provides some extensions to the DFP instructions to facilitate/improve the functionality of the DFP architecture.

IBM POWER7

- IBM recently announced its next-generation processor POWER7 that implements the Power ISA v2.06. POWER7 comes with the above mentioned features of the v2.06 Power ISA. It has implemented four merged Vector and Scalar Floating Point Pipeline that supports the complete VSX Instruction set and can potentially execute four double precision multiply-add operations delivering up to 8 flops per cycle. Overall POWER7 will deliver scorching performance for the HPC application space with VSX and new cache management features with low power consumption. And will further enrich the Enterprise server systems with advanced memory management and provide for flexible Systems Management through Processor Compatibility Register.

Summary

The Power ISA as an integrated architecture continues to provide for varied advantages including performance, energy efficiency, virtualization and RAS. It allows HPC and compute-intensive server workloads enhance throughput by leveraging VSX. Virtualization and workload mobility for maintenance, consolidation and energy management in enterprise systems and data center environments is facilitated through the PCR. Also, emulation

support for hybrid systems is provided for next-generation systems. Power ISA v2.06 addresses a number of requirements to build future generation systems that couple high-performance with efficient energy management and easier systems management.

References

- Power Instruction Set Architecture Version 2.06:
http://www.power.org/resources/downloads/PowerISA_V2.06_PUBLIC.pdf
- Guide to Multiple Page Size Support on AIX 5L Version 5.3:
http://www-03.ibm.com/systems/resources/systems_p_os_aix_whitepapers_multiple_page.pdf
- Storage Keys for memory protection in AIX6 and POWER6:
<http://www.ibm.com/developerworks/wikis/display/WikiPtype/StorageKeys>
- Power6 and Decimal Floating Point:
http://www.power.org/devcon/07/Session_Downloads/PADC07_Johnson_DFP_final.pdf
- Power6 DFP:
<http://www.ibm.com/developerworks/wikis/display/WikiPtype/Decimal+Floating+Point>

About the Authors



Satish Kumar Sadasivam is a Staff Software Engineer with IBM associated with the Performance Analysis & Compiler Optimization team in IBM Systems and Technology Group. He has wide experience in the area of Post-Silicon Validation and Performance Analysis. He has worked on IBM's Power 5+, Power 6 and Power 7 Processors developing Validation Tools with expertise in Floating Point architectures. He is currently working on Performance Analysis and Optimization of Power 7 and next generation processors. He has various invention disclosures and publications in the area of Post-Silicon Validation and Performance Coverage. Email: satsadas@in.ibm.com



Prathiba Kumar is a System Software Engineer at IBM working for Performance Analysis and Benchmarking for Power 7 Systems in IBM Systems and Technology Group. She is currently concentrating on Power 7 Performance Benchmarking for SPEC CPU. Her work also involves analyzing potential benchmark applications and suggesting compiler

improvements. She has worked on Power Management techniques in the latest Power processors. She has various invention disclosures and publications in the area of Processor Validation and coverage measurement. Email: prathkum@in.ibm.com



Varun Mallikarjunan is a Software Engineer at IBM working for Post-Silicon Validation team at IBM Systems and Technology Group. He has worked for IBM's Power 7 Post-Silicon Validation developing Validation Tools. He is currently working on Next-generation Network Processor Validation tools development. He is proficient in Power Architecture with expertise in Power Architecture Memory Management. Email: varunmallik@in.ibm.com