

Introduction to the wire-speed processor and architecture

H. Franke
J. Xenidis
C. Basso
B. M. Bass
S. S. Woodward
J. D. Brown
C. L. Johnson

In this paper, we introduce the wire-speed processor (WSP) project, an advanced development project led by IBM Research and the IBM Systems and Technology Group. The WSP represents a generic processor architecture in which processing cores, hardware accelerators, and I/O functions are closely coupled in a system on a chip. The first implementation of the WSP architecture targets applications operating at "wire speed" (i.e., speeds in which the data are transmitted and processed at the maximum speed allowed by the hardware). These applications include those that involve routers, firewalls, intrusion-prevention systems, and other network analytics. The WSP combines 16 multithreaded IBM PowerPC® cores with special-purpose dedicated accelerators optimized for packet processing, security, pattern matching, compression, Extensible Markup Language (XML) parsing, and I/O for networking that provides four 10-Gb/s bidirectional network links. In this paper, we describe the various system components, the underlying design philosophy involving close integration of these components, and the special system features that were developed to achieve this close integration.

Introduction

Users and manufacturers of computing systems have come to realize that various limitations exist with respect to traditional serialized computing approaches and the homogeneous performing of tasks. This has led to several chip technology transitions that are currently in development. Due to power use challenges, there are diminished benefits of increasing single-thread performance through frequency scaling and extracting instruction-level parallelism through highly effective out-of-order core technology [1]. Instead, technologists are creating processor designs that focus on throughput rather than single-thread performance. In particular, current chip designs at the 45- and 32-nm scales are now decreasing or maintaining current core frequencies to maintain constant power densities. The increase in density due to decreased sizes in nanometer technology has provided for significant transistor growth on the die. More transistors mean more cores and larger functional units that can be shared between more processor threads. At one conceptual extreme, technologists at IBM [2] and other companies [3] have explored the use of multicore processors for high-end servers,

focusing on high-performance single-threaded computation and increased concurrence. More recently, at the other extreme, large numbers of smaller simpler cores, originally designed for graphics processing, have gained some popularity for scientific and general-purpose computing [4, 5]. Located somewhere in the middle of this range of concepts, we find the heterogeneous multicore Cell Broadband Engine** (Cell/B.E**) processor [6] with its programmable and asynchronously operating synergistic processor elements (SPEs), demonstrating its value for graphics and scientific computing [7]. A second transition is emerging due to the offloading of core cycles to domain-specific hardware accelerators, which can provide significantly more computation performance and chip density at a lower power budget but at the cost of decreased flexibility [8–10].

Historically, system-on-a-chip (SoC) designs were created to combine popular platform components into a single chip, thereby decreasing the platform cost and increasing the value of the final chip. Although most SoCs exploited a simpler integrated design that removed the requirement for external pins to standardized buses to connect the components, software access to the components remained as if the components were off-chip and physically "distant." This

Digital Object Identifier: 10.1147/JRD.2009.2036980

©Copyright 2010 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the Journal reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied by any means or distributed royalty free without further permission by computer-based and other information-service systems. Permission to republish any other portion of this paper must be obtained from the Editor.

0018-8646/10/\$5.00 © 2010 IBM

perceived physical “distance” is apparent in the programming model in which the software running on the CPUs must interact with devices through their device-specific exposed I/O channels and registers or through invoking a device driver, which can decrease the overall performance. On the other hand, devices interact with the software through interrupts that can create significant performance overhead for the application software. Furthermore, memory accessed by devices typically needs to remain resident (i.e., pinned down) to ensure that direct memory accesses (DMAs) by the device operate properly.

In this paper, we present the wire-speed processor (WSP) architecture, which provides a generic heterogeneous architecture by integrating multiple generic cores with domain-specific acceleration and I/O functions in an SoC. The WSP provides the performance advantages of accelerators while reducing the software development costs associated with the heterogeneity of accelerators. It accomplishes this by providing streamlined and uniform interfaces to accelerator functions and a uniform memory addressing scheme across software threads and accelerators. In particular, accelerators use the same addresses that the software uses. We refer to this as the “principle of uniform addressability in a heterogeneous environment.” By exploiting key technologies and expertise in scalable coherency fabrics, custom coprocessing units, high-end networking and IO components, and the ability to customize an existing and mature instruction-set architecture (ISA), the WSP removes many of the constraints discussed above, dramatically decreases the distance between the components, and increases the level of integration such that the software is presented with a closely integrated model.

In the first implementation of the WSP architecture, the WSP is targeted toward the domain of network-facing applications such as intrusion-prevention systems and enterprise service buses. Networks continue to experience significant growth in bandwidth, i.e., toward 40 and even 100 Gb/s. At the same time, more intelligent processing and functionality are being provided by these network applications, significantly increasing the computational demand associated with packet processing. A useful overview of the emerging field of network-optimized computing systems that provides the requirements that resulted in the WSP architecture is described in [11]. In many cases, available processor core cycles are insufficient to achieve data processing at wire speed. Instead, reoccurring functions can and must be offloaded to accelerators, while also considering the tradeoffs between flexibility, performance, and power use [12]. Finally, because the compute-to-network data ratio is still lower for network-facing applications than in traditional server applications, the latency and bandwidth overhead of crossing chip boundaries for I/O and accelerators can become a significant performance bottleneck.

The integrated WSP architecture is designed to reduce this overhead.

WSP overview

The WSP architecture is designed to facilitate concurrence, heterogeneity, and asynchronous processing on-chip, while still retaining a uniform and coherent environment for the software. To meet the demands of existing and emerging network-facing applications, the WSP architecture is an SoC that is composed of four distinct complexes, shown in **Figure 1** (in different colors) and described in more detail later in this paper. The four complexes are described as follows.

- The *interconnect complex* joins the internal processor components (complexes) and other external components such as memory, PCI Express** (PCIe**), and other WSPs.
- The *processor compute complex* is composed of a large set of threads that provide high performance per watt and are optimized for parallel processing.
- The *accelerator complex* is composed of a set of special-purpose accelerators and coprocessors that are frequently used in the targeted application domains. These accelerators are significantly more power efficient than general-purpose processors [10] and will exceed the performance of highly tuned software on general-purpose processors.
- The *network I/O complex* is based on multiport 10-Gb/s Ethernet technology and supports functions such as packet classification, packet scheduling to specific cores and threads, packet ordering, and traffic management, thus easing the burden on the software required to implement these functions.

In contrast to existing systems, in which most accelerators are devices attached to an I/O bus and are programmed through system-level interfaces, the individual complexes in the WSP (compute, accelerator, and packet processing) all operate using the same application address space (i.e., virtual address space). This uniform addressability reduces the amount of specialized code required to allow the units to share storage. In addition, accelerators and network I/O units are all “first-class citizens” in the system, in that they have full access to the caches and participate in the coherency protocol, freeing the application programmer of costly synchronization and system invocations. The intent is to significantly reduce the performance overhead for accessing the accelerator and network functions.

It is not the intent of this paper to describe the specific performance characteristics of the WSP and its individual components. Rather, we focus on the architecture of the WSP and the additions that were made to achieve the first-class citizen status of network and accelerator functions.

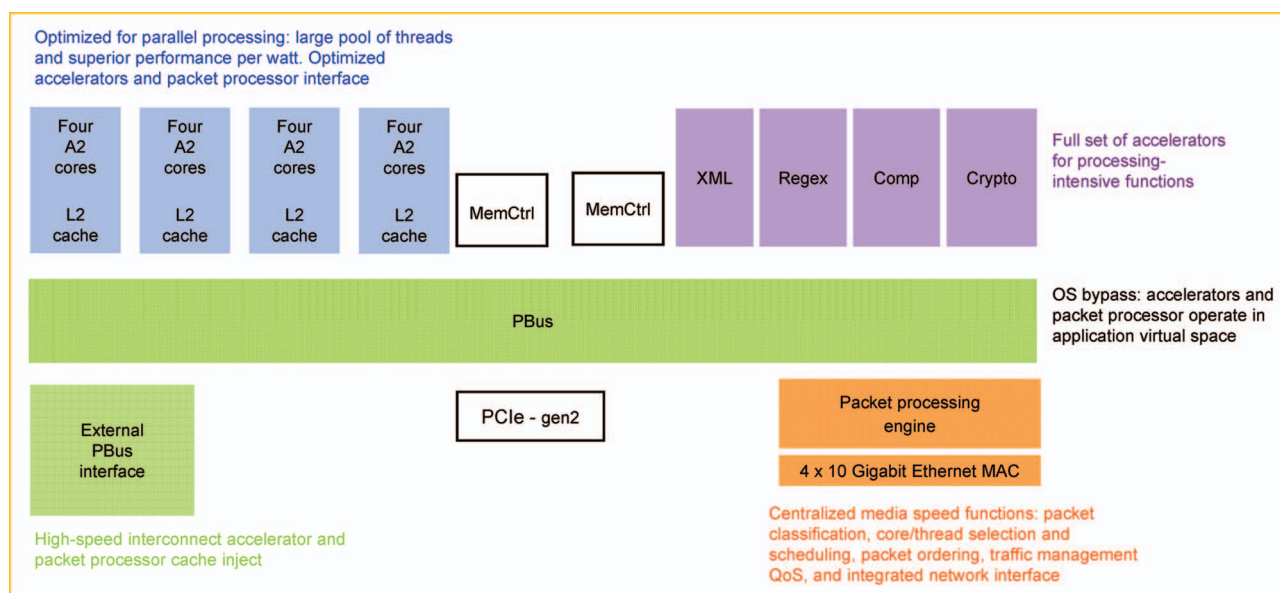


Figure 1

WSP functional diagram and characteristics. The A2 cores are described later in the text. (XML: Extensible Markup Language; Regex: regular expression and pattern-matching accelerator; Comp: compression and decompression accelerator; Crypto: cryptographic accelerator; MAC: media access control, a unique network adapter ID; MemCtrl: memory controller; QoS: quality of service; gen2: generation 2.)

WSP

The WSP is our first chip designed on the WSP architecture, and this section describes the functional details of the complexes introduced in the overview.

Interconnect

Central to every system is the interconnect fabric that joins the complexes. Since the primary goal of the architecture is to provide uniform addressability and coherency throughout the processor, each complex requires a dedicated interface to a coherent fabric. To satisfy these requirements, we developed a coherency fabric called *PBus*, where every component of every WSP complex is represented as a *unit* on *PBus*. Each unit is connected through some form of interface that is optimized for the component. Although the processing complex uses the cache logic to interface with *PBus*, each accelerator and packet processing complex utilizes a generic interface called the *PBus* interface controller (PBIC) described below.

PBus provides a cache coherency and data protocol that can be used by a wide range of processor implementations. The *PBus* architecture is intended to be applicable to system implementations ranging in size from single-processor to heterogeneous multiprocessor configurations. *PBus* provides the basis to support coherent and noncoherent memory accesses, I/O operations, interrupt communication, and system controller communication. The address and data are split into two separate transactions, and the cache coherence is

maintained by utilizing a snooping protocol [13].

Multiple unit-to-unit data paths are supported in order to increase the effective data bandwidth. Each unit has a command bus, a reflected command bus, a partial response bus, a combined response bus, and read and write data buses. The data bus is a partial crossbar switch [13] consisting of four 16-byte vertical buses, two in each direction. *PBus* also extends through three external interfaces for a point-to-point coherent connection with up to three additional WSPs to form a larger coherent system.

A single WSP has two independent double-data-rate (DDR3) memory controllers, each with two independent channels. Up to two registered dual inline memory modules (DIMMs) or unbuffered DIMMs attach to each channel supporting data rates of 800 MHz, 1.066 GHz, 1.33 GHz, and 1.6 GHz.

Processor and compute complex

The processing complex is composed of 16 PowerPC* cores, referred to as A2 cores, operating at 2.3 GHz. Each A2 core has four simultaneous threads of execution [14]. Multiple instructions are issued from different threads in one cycle, and each thread can be viewed as a processor within a four-way multiprocessor. The simultaneous multithreading allows these simple cores to provide high instruction throughput in an area- and power-efficient manner. The threads implement the standard Power* ISA (ISA 2.06 [15]) and are binary compatible with existing software compiled for that ISA.