A High-Performance Dual PCI Bridge for PowerPC™ with 100 MHz Bus and PC-100 SDRAM

Gerard Boudon, Claude Sitbon, and Herve Mougel

Introduction

PowerPC is a premium choice for embedded applications because, as a RISC processor, it offers a very good ratio of computational power versus price and it offers extremely low power dissipation — even beyond 400 MHz. However, the chipset that bridges the processor to the system memory and to the I/Os is just as important to performance. The PowerPc support chip was designed to provide a 500 MHz PowerPC processor high-speed memory access and to permit attachment of peripherals on the popular PCI bus.

State-of-the-art system architecture is based on three level of memory hierarchy:

- L1 integrated in the processor chip
- L2 cache implemented in fast SRAM
- L3 system memory using low-cost and fast SDRAM

Currently, to increase bandwidth in most systems, the L2 cache is directly attached with a dedicated bus (gunning transceiver logic [GTL], point-to-point transmission) to the processor. With the copper PowerPC 750 running at 400 MHz, a throughput of 1.6 GBps in burst and pipeline mode is achievable at 200 MHz on the L2 bus.

The use of an L2 cache support function is no longer justified in the chipset support, and the saved pins can be used for other functions.

Today, the main memory — fast SDRAM with high throughput, but with several system clock cycles of latency — is accessed by the processor via a processor bus running at 100 MHz. To go beyond 100 MHz bus speed, the choice of the electrical interface has to evolve to a lower voltage swing than the low voltage transistor logic (LVTTL) used for many years.

The PowerPC support chipset was designed for embedded systems, especially those in telecommunication applications, such as network infrastructure equipment. Please note that this chipset was developed to meet the specific needs of an IBM customer and has limited availability at this time.

PowerPC Bridge Features

Typical embedded system can be designed around a PowerPC processor, and the proposed support chipset as shown in Figure 1. Four buses are supported by the bridge circuit:

- PowerPC bus @ 100 MHz o Memory Bus PC100 SDRAM 64 bits + ECC
- PCI 32-bit @ 33 MHz
- PCI 64-bit @ 66 MHz

A conventional host system is built with a single PCI primary bus and a secondary PCI bus, which can be accessed through a PCI-PCI bridge.

In a telecommunication application — such as a router or a switch — a lot of data is exchanged between the I/Os sitting on the PCI buses and the memory. Due to the hierarchical approach in conventional systems, the I/Os on the secondary bus have a very low chance to access the memory, since they first have to gain access to the secondary and then the primary PCI bus.

Figure 1. Typical embedded system with PowerPC chipset architecture with two independent PCI buses.

To overcome this architectural limitation, the PowerPC support chip has two primary PCI buses:

1. 64-bit PCI 66 MHz bus to support high-speed graphic and real-time video computation. At 33 MHz, this bus can also be used as a CompactPCI controller, allowing seven slots in the backplane.
2. PCI 32-bit 33 MHz bus.

Both the PCI32 and the PCI64 bus have full master/slave and PCI bus controller capability.

Why a PCI 64-bit bus?

Improvements to the CPU's processing capability so that it can transmit more data, more voice channels, better quality images, etc., result in requests for more bandwidth and low latency. For example, the Ethernet LAN adapter card for the 10/100 Mbps uses the 32-bit PCI bus for data
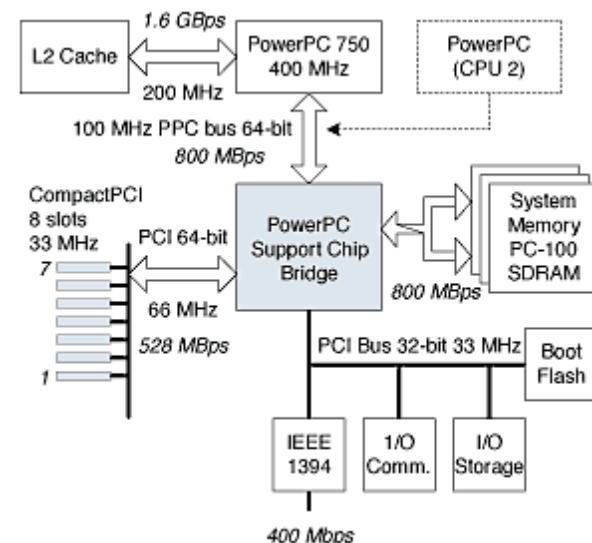
transmission to the host CPU, but gigabit requires higher bandwidth that can be handled only by a 64-bit wide PCI.

- PCI 32-bit 33 MHz   132 MBps   1.0 Gbps Peak
- PCI 64-bit 66 MHz   528 MBps   4.2 Gbps

The 64-bit PCI is fully backward- and forward-compatible with 32-bit devices. At 66 MHz, the load is limited and the number of slots are reduced.

Memory Controller
A 100 MHz memory bus permits the attachment of PC100 SDRAMs, widely used since 1998 in PCs and normalized by Intel [1]. By using an interleave scheme, up to eight slots of dual in-line memory modules (DIMMs) can be connected, which results in a 2 GB configuration with the new IBM 256 MB (32Mx72) DIMMs. The SDRAMs are ECC protected with correction of all single-bit errors and detection of double-bit errors.

The next version of the chipset will have to cope with the PC133 specification before supporting the double-data-rate (DDR) SDRAM, and later the SLDRAM open standard or Rambus™ DRAM (RDRAM).

PC100 SDRAM

Setup time	2 ns
Hold time	1 ns after clock
Clock-to-clock delay	6 ns

Figure 2. PC-100 timing constraints.

The PC100 timing constraints (Figure 2) are very important, and the design of the SDRAM controller was not easy. The most difficult was to verify the set-up and hold times — 2 ns and 1 ns respectively, corresponding to timing established for 125 MHz SDRAM modules. As for the 60X bus, the 3.3 V LVTTL with a large signal swing (0.8 to 2.4 V) is used here at its limit. The next generation of SDRAM with an interface such as the series stub terminated logic (SSTL-3) with a reduced voltage swing of ± 0.4 V around the 1.5 V reference and terminated lines should permit an increase of bandwidth to the memory.

Interleaving
The first advantage of the interleave scheme is the capability to double the memory size to 8 DIMMs. The second advantage is that only the clock CLKE and the chip select are 100 MHz-critical signals; the rest are on a 50 MHz period, relaxing some of the AC timing constraints.

High Speed Data Transfer Through the Bridge
The memory and the PowerPC host bus allow 800 MBps of data transfer. The maximum theoretical bandwidth of 32-bit PCI at 33 MHz is 132 MB, and the 64-bit PCI can reach 528 MBps if it is clocked at 66 MHz. To optimize the bandwidth utilization of these buses, special care has been taken on the internal data paths of data. Transfers from one bus to the other have been optimized by the use of wide buffers: 8 deep x 64-bit wide, as shown on Figure 3. These buffers are fed from the input pin at the bus rate, and data are sent to all four interface circuits in the chip that can receive data. In each receiving circuit, a multiplexer is controlled to permit gating of the data.
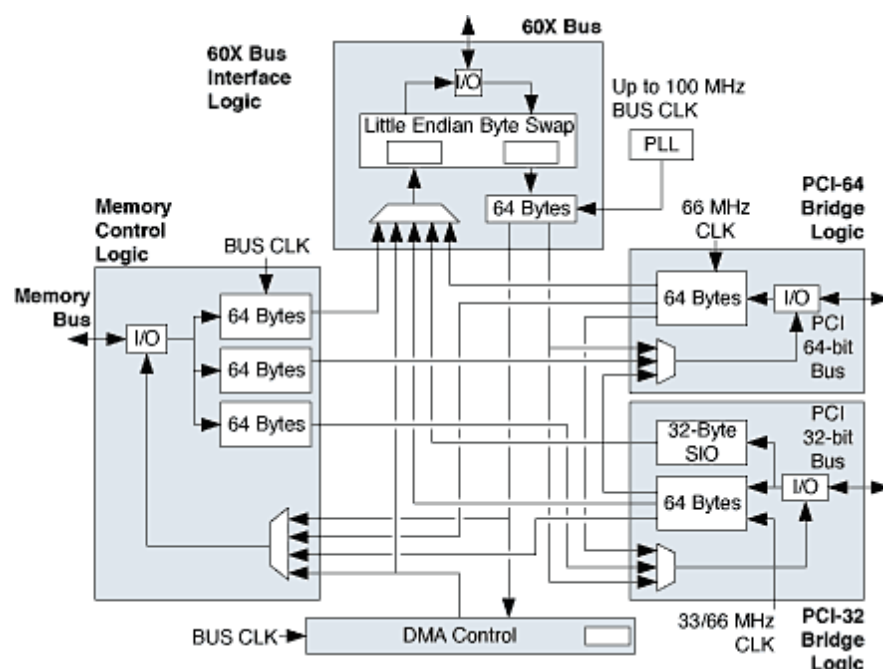

Figure 3. Bridge fast internal data paths.

The internal buses are also 64-bits wide, and there is always a path from one unit to the other. A round robin arbitration is performed at the input of each block. This arbitration controls the N-way

multiplexer such that only one data path is open to provide the data on the output pin of the block.

Data Gathering: The 60X logic circuit provides a data gathering capability for CPU stores, and it transfers data to the PCI bus bridges. It is a mechanism whereby multiple single beat stores from the CPU are gathered — up to as many as 32 bytes — before sending to the PCI bus bridge unit. The resulting benefit is that it operates more efficiently by bursting data on the PCI bus.

Concurrent Operations: The PowerPC chipset has been designed for performance, and its large number of I/Os provides concurrent operations on separate buses:

1. PCI64-to-memory read, while the CPU is snooping on its 60X bus
2. PCI32-to-memory write, while the CPU is writing to PCI64
3. CPU-to-memory read, while the PCI64 is writing to PCI32

Signal Noise Induced

The operations that increase overall system bus bandwidth, and thus performance, are also sources of induced noise in power supply wires. Up to 147 drivers can switch addresses or data to the SDRAM simultaneously, which creates a high demand for current — 14.7 A/ns.

All four high-speed buses can switch simultaneously. The results shown in Figure 4 were obtained with SIMUN, an IBM simulation software tool. These results show that in spite of severe conditions, no induced noise exceeded the 640 mV limits of a receiver circuit that could be tied to the noisy driver.
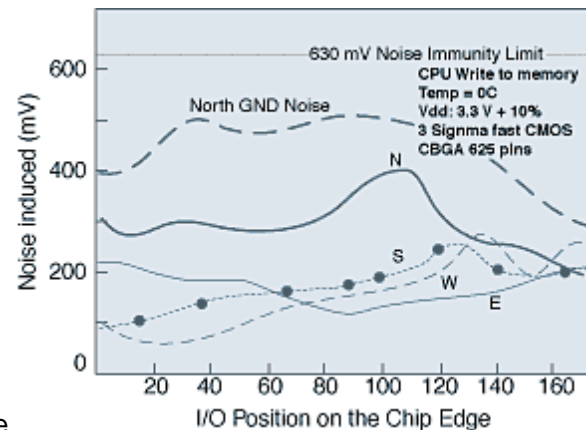

Figure 4. Induced noise in quiet drivers versus its position on the chip edge. The results show that in spite of severe conditions, no induced noise exceeded the 640 mV limits of a receiver circuit that could be tied to the noisy driver.

Design Challenges

The PowerPC bridge (Figure 5) has been implemented on a large 81 mm2 ASIC in 0.35 μm technology based on standard cells. Due to the large number of I/Os placed on the peripheral of the die, only a low percentage of the total available cells in the ASIC are used. With the next 0.2 μm ASIC generation, the silicon area can be greatly reduced, because of availability of I/Os anywhere on the die.

Design Optimization
The 10 ns cycle, which is latch-to-latch delay, was achieved on the design after fine analysis of critical paths and numerous synthesis runs with the IBM BooleDozer™ tool.

Clock Distribution
At 100 MHz, a phase lock loop (PLL) clock distribution circuit is necessary to minimize jitter between the internal clock and I/O clocks [2]. To reduce its skew, the master clock is distributed with one signal phase only along the total area of the chip and then for each cluster of registers, for example 32 latches, the two phases are generated. These two phases drive in sequence the master and the slave latches defined in the Level-Sensitive Scan Design (LSSD) methodology used in all IBM ASICs.

Architecture	- PowerPC 100 MHz Bus - PC-100 SDRAM 64 bit + ECC - PCI 64-bit 66 MHz - PCI 32-bit 33 MHz - DMA function 1 channel - JTAG
Technology	- 0.35 μm CMOS - Four levels of metal - 300K cells
Temperature	0° to 105° C junction
Signal I/Os	480
Power Supply	3.3V ± 5% (support for 5V I/Os)
Power Dissipation	2.1 W @ 100 MHz (typical)
Packaging	625-contact, 32.5 mm, ceramic BGA

Figure 5. PowerPC 100 MHz bus single chip bridge.

Five stages of buffers were necessary to drive the 522 clock splitters of the 60X and the memory controller circuit common for several latches. The number of stages and their fan out to build the clock tree is computed such that it takes into account die size and number of latches.

Asynchronous Clocking
The main CPU clock running at 100 MHz is used for the 60X bus and for the SDRAM memory controller. The 32-bit PCI and the 64-bit PCI each have their own clock, providing a flexible choice of system and PCI clocks. The boundary between PCI and 60X where re-synchronization must occur has been designed with metastability free registers.

Packaging
The 32.5 mm ceramic ball grid array (CBGA) package was chosen because it allows the high pin count necessary to meet the large simultaneous switching requirement with minimum inductance.

Summary

A high-performance PowerPC support chipset has been designed for 100 MHz bus operation with dual PCI buses and a PC-100 SDRAM port. The use of 0.35 μm CMOS technology and of complete static full CMOS circuits has permitted the achievement of a very low power dissipation necessary for embedded applications.

The chip has been developed in Verilog™ language and has been implemented on an ASIC. This approach provides the flexibility to very quickly remap the logic in a more advanced CMOS technology to reach 133 MHz, and even allows it to be integrated with other functions, such as an interrupt controller, IEEE 1394 bus or PowerPC processor core.

References

1. *PC-100 Specification,* November 1997.
2. M. Horowitz, "High-Speed Electrical Signaling; Overview and Limitations," *IEEE Micro,* February 1998.

Gerard Boudon is a program manager in Embedded Systems, IBM France, Corbeil-Essonnes.

Claude Sitbon is project leader VLSI chips for Embedded Systems, IBM France, Corbeil-Essonnes.

Herve Mougel is project leader VLSI chips for Embedded Systems, CETIA, Toulon France.