TECHNOLOGY BRIEF

CONTENTS

Pentium II Processor Technology

EXECUTIVE SUMMARY

The Intel Pentium II microprocessor is the newest addition to Intel's line of sixth-generation microprocessors. The Pentium II processor is based on an Intel Pentium Pro processing core, with the addition of faster core frequencies, a new form factor, a new cache structure, and multimedia extension (MMX) instructions. The higher core frequencies of the Pentium II processor bring enhanced performance to Compaq's line of workstations, file/print servers, and other systems used in CPU-intensive applications. However, the Pentium II processor is not targeted for use in all application environments. The Pentium Pro processor remains the primary processor for Compaq servers and workstations used in memory-intensive applications and systems employing more than two microprocessors. This paper highlights the similarities and differences between the Pentium II and Pentium Pro processors and suggests some appropriate application environments for each processor.

COMPAQ

Please direct comments regarding this communication to the ECG Technology Communications Group at this Internet address:
tech_com@bangate.compaq.com (WFootnote)Text

NOTICE

Pentium II Processor Technology

SIMILARITIES BETWEEN PENTIUM II AND PENTIUM PRO PROCESSORS

Like the Pentium Pro processor, the Pentium II processor is considered to be a sixth-generation processor. The Pentium II processor is based largely on the Pentium Pro processing core. Most of the major architectural features of the Pentium Pro have been carried over into the Pentium II processor. Both processors

- Employ the same or similar micro-architecture.

- Share binary compatibility with previous x86 processors.

- Use the Gunning Transceiver Logic Plus (GTL+) system bus.

- Use dynamic execution techniques.

- Achieve increased data throughput from a Dual Independent Bus architecture.

- Support Advanced Symmetric Multi-Processing (SMP) configurations.

- Provide on-chip level one (L1) instruction and data caches.

These similarities are discussed in more detail in the following section.

Micro-Architecture

The Pentium Pro and Pentium II processors have similar micro-architecture. Both processors have a reduced instruction set computer (RISC) core coupled with a complex instruction set computer (CISC) decoder. Both processors are capable of issuing and retiring three instructions per clock cycle, making them 3-way super-scalar.

Binary Compatibility

The Pentium Pro processor and the Pentium II processor are fully binary-compatible. All software compiled for Pentium Pro and other x86 processors will operate on the Pentium II processor.

GTL+ Bus

Both the Pentium Pro processor and the Pentium II processors use the GTL+ bus. The GTL+ bus, or host bus, connects the processors to system memory, the advanced programmable interrupt controller (APIC) bus, and the peripheral component interface (PCI) bus. The GTL+ bus is designed to accommodate multiple processors in an SMP configuration, and to support multiple PCI buses.

Dynamic Execution

The Pentium Pro and Pentium II processors use dynamic execution to optimize the flow of instructions and data in and out of the processor. Dynamic execution employs three different techniques — data flow and dependency analysis, multiple branch prediction, and speculative execution — to achieve this optimization.

The data flow and dependency analysis technique analyzes program code to determine the dependencies between instructions. Using this technique, instructions are scheduled to optimize the processor's workload, regardless of the instructions' order in the original program. Program instructions with no data dependencies, such as instructions that do not require reads from memory or the output of previous instructions, are scheduled during periods when the processor would otherwise "stall" waiting for data from the caches or system memory.

On average, programs jump, or branch, every five to seven instructions. Multiple branch prediction uses a static algorithm and a historic algorithm combined with a large branch target buffer to predict program branch locations. Correctly predicting the branches allows the data flow and dependency analysis to schedule instructions even across program jumps.

The Pentium Pro and Pentium II processors will speculatively execute instructions based on the outcomes of data flow analysis and multiple branch prediction and store the results in temporary storage. However, the instructions are always formally retired in the original program order. All dynamic execution techniques are designed to enable the processor to recover in case of an interrupt, trap, fault, or mis-predicted branch.

Dual Independent Bus Architecture

Both the Pentium II and Pentium Pro processors employ a Dual Independent Bus architecture. This design uses two independent buses: a GTL+ host bus that connects the processor to main memory and a cache bus that connects the processor core to the level two (L2) cache. Relocating L2 traffic off the main memory bus frees up bandwidth on the host bus for memory transactions. This crucial additional bandwidth improves scalability in multiprocessor systems. In addition, the dual bus design allows the L2 bus to function at a much higher frequency than the host bus.

Advanced SMP Support

Unlike the Pentium-class processors, both the Pentium Pro and the Pentium II processors contain the necessary pins and circuitry to directly tie processors together in a completely symmetric manner. This includes circuitry to ensure cache coherency and interface directly with the APIC interrupt bus. Integrating this circuitry into the processor reduces the need for additional components on the motherboard to manage these functions.

On-Chip L1 Data and Instruction Cache

The Pentium Pro and Pentium II processors include on-chip, non-blocking L1 instruction and data caches. Non-blocking caches allow the processor and cache to process cache hits while concurrently servicing cache misses to the L2 cache and main memory. The L1 caches operate at the same frequency as the processor and provide quick access to frequently used data and instructions.

DIFFERENCES BETWEEN PENTIUM II AND PENTIUM PRO PROCESSORS

The Pentium II processor does have some significant differences from the Pentium Pro processor. The following are the major distinctions:

- Higher processor frequencies

- New packaging

- Different cache architecture

- Limited multiprocessing capabilities

- Multimedia extension instructions

- Lower processor voltage

These new features of the Pentium II are outlined in the following section.

Processor Core Frequency

One of the most exciting differences between the Pentium II processor and the Pentium Pro processor is the increased core frequency of the Pentium II processor. Pentium II processors operate at 233, 266, and 300 MHz. Currently, the highest Pentium Pro core frequency is 200 MHz. Higher core frequencies typically increase the performance of systems in CPU-constrained applications.

Single Edge Contact (SEC) Cartridge

With the Pentium II processor, Intel is introducing a new packaging form factor called the single edge contact (SEC) cartridge. The SEC is illustrated in Figure 1. By contrast, Pentium Pro processors are packaged in pin grid array (PGA) packages.
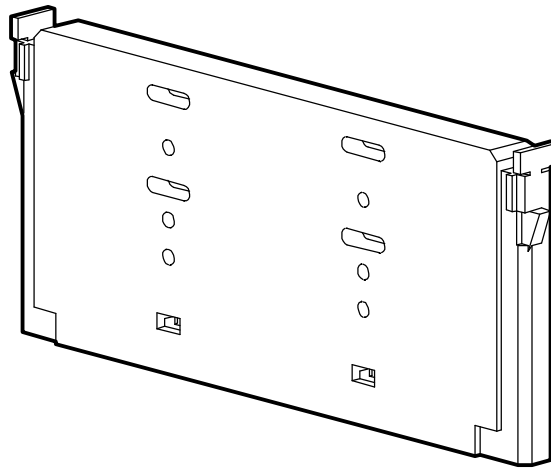
Figure 1. Outline of the Pentium II Single Edge Contact cartridge

The SEC cartridge measures 4.9" x 2.1" x 0.5" and resembles a video game cartridge. The SEC processor cartridge fits into a card-edge connector socket on the system motherboard, designated as Slot 1. Intel has selected the SEC as the form factor for its next-generation processors.

Cache Architecture

The Pentium II processor cache architecture differs significantly from that of the Pentium Pro processor architecture. In the Pentium Pro processor, the L2 cache is located in the processor cavity. With the Pentium II processor, the L2 cache memory moves out of the immediate processor package and onto a processor printed circuit card within the SEC cartridge, as shown in Figure 2. This configuration allows Intel to implement the L2 cache using commodity TagRAM and burst pipelined synchronous static RAM (BSRAM).

Processor Core   L2 Cache
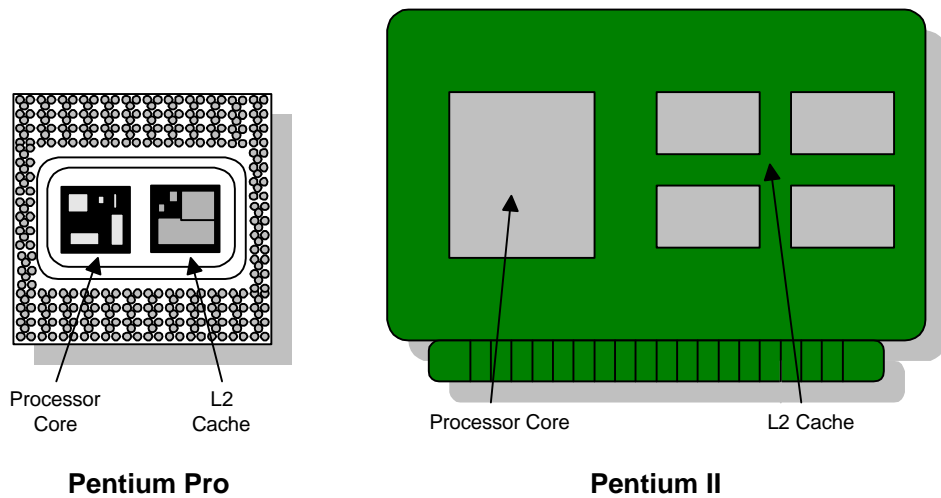
Pentium Pro   **Pentium II**

Figure 2. Representation of L2 cache locations in Pentium Pro and Pentium II processor packages. The Pentium Pro processor contains the L2 cache within the pin grid array package, while the Pentium II relocates the L2 cache further away onto the processor printed circuit board.

To accommodate the greater distance between the processor and the L2 memory components, and the higher frequencies of the Pentium II processor, Intel reduces the bus speed between the processor core and the cache. In the Pentium Pro processor, the L2 cache connects to the processor via a 64-bit data bus operating at the processor core frequency. However, the Pentium II processor connects the L2 cache to the processor core at one-half the processor core frequency. This means a 266 MHz Pentium II processor will have a 133 MHz L2 frequency. By comparison, the 166 MHz Pentium Pro processor has a 166 MHz bus connection to the processor. Cache bus speed is an important factor in cache performance. Cache performance can greatly impact overall system performance, especially in multiprocessor configurations and in systems used in very memory-intensive applications such as large database applications.

To compensate for the slower L2 cache, Intel increases the size of the L1 cache in the Pentium II processor to 32 KB, compared with a 16-KB L1 cache in the Pentium Pro processor. The Pentium II L1 cache is divided into a 16-KB instruction and a 16-KB data cache. The larger L1 cache potentially reduces the number of L2 accesses, improving overall memory performance. The current Pentium II processor has a cacheability limit of 512 MB of main memory. The cacheability limit is a result of the TagRAM, which provides only enough address lines to locate data within a 512-MB address space. Thus, while the processor can address up to 4 GB of physical memory, only 512 MB can load into the L2 cache. Because of this limitation, customers with servers and workstations deployed with more than 512 MB of system memory could

experience substantially lower performance from Pentium II processors than from Pentium Pro processors.

For more information on the performance of processor cache subsystems, please refer to the Compaq technology brief *Performance of Pentium Pro and Pentium II Processor/Cache Combinations*, Document 436A/0597.

SMP Support

Advanced SMP support is provided through the use of arbitration pins on the processors. The Pentium Pro processor provides four arbitration pins and can directly support up to four processors in an SMP configuration with no additional circuitry. In contrast, the Pentium II processor provides two arbitration pins and can therefore support only two processors in an SMP configuration. In addition, the internal termination scheme of the Pentium II dictates that no more than two Pentium II processors can connect to the GTL+ bus.

MMX

The Pentium II processor offers MMX instructions, a set of 57 new instructions. The MMX instructions use the single instruction multiple data (SIMD) technique. The instructions are optimized for audio, video, graphics, and other applications having the following characteristics:

- Small integer data types

- Small, highly repetitive loops

- Frequent multiplies and accumulates

- Compute-intensive algorithms

- Highly parallel operations

The SIMD instructions use eight 64-bit MMX registers overlaid on existing floating-point registers. MMX instructions can perform parallel operations on all data in an MMX register. Figure 3 illustrates the potential efficiency of using MMX instructions instead of standard instructions. In this scenario, eight bytes of data from array A are multiplied by eight bytes of data from array B, and the result is stored in a result array. With non-MMX instructions, one byte of data from array A is multiplied by one byte of data from array B per program loop. Processing all eight multiplies requires executing the program loop eight times. With MMX, eight bytes from arrays A and B are packed into two MMX registers. The result is stored in another MMX register. Only one MMX instruction is executed to multiply the packed values together and store the results in the third array. This parallelism can greatly speed up some graphics, video, and audio applications, as well as some file compression/decompression algorithms. However, programs must be specifically written to take advantage of the MMX instructions.
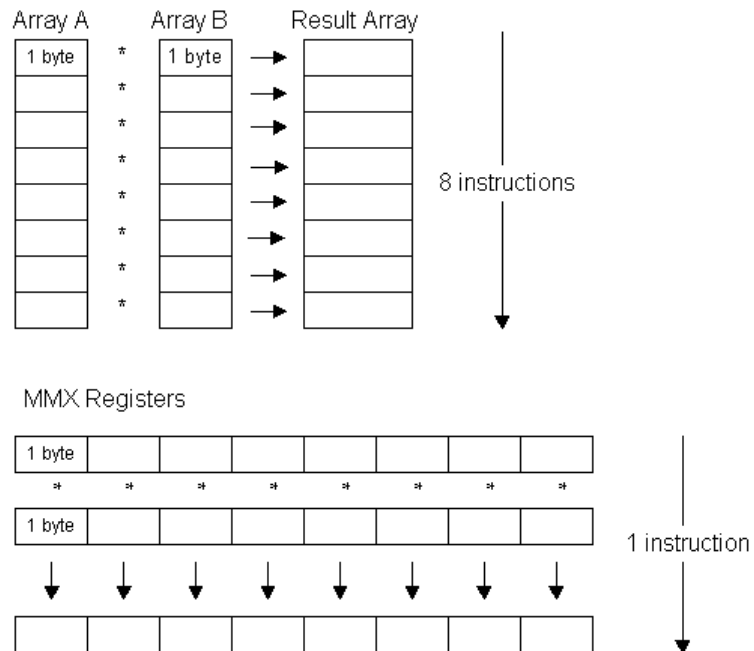
Figure 3. Example of MMX instructions. Using non-MMX instructions, a program loop executes multiple times to process the data. With MMX, the data is processed with one instruction call.

Processor Voltage

The maximum voltage requirement for the Pentium II processor (2.9 volts) is lower than that for the Pentium Pro processor (3.3 volts). Also, with the Pentium II processor, the APIC bus and clock voltage drops to 2.5 volts. Special voltage regulator modules (VRMs) on the processor board identify the exact power requirement of each processor and provide the necessary voltage directly to the processor. Pentium Pro and Pentium II VRMs are not interchangeable.

APPLICATIONS FOR PENTIUM PRO AND PENTIUM II PROCESSORS

The differences in the Pentium Pro and Pentium II processors dictate the most suitable application environments for each processor. The Pentium Pro processor provides outstanding scalability and performance for memory intensive applications demanding more than 512 MB of system memory and more than two processors. For some very compute-intensive applications, SMP configurations with four Pentium Pro processors may yield the best overall CPU performance. Examples are large database and application servers, and workstations used with Computer Aided Engineering (CAE) or Electronic Design Automation (EDA). The Pentium II processor's higher core processor frequency and MMX capabilities make it an excellent choice for CPU-intensive environments with one or two processors such as file/print servers and most workstation applications. Figure 4 summarizes the recommended application environments for Pentium Pro and Pentium II processor-based systems.
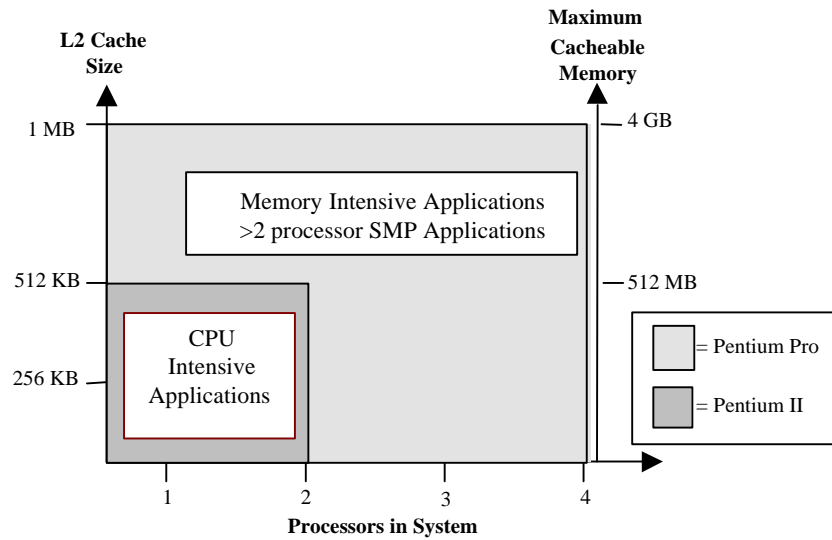
Figure 4. Recommended application environments for Pentium Pro and Pentium II processors.

CONCLUSION

The Pentium II processor has many features that make it an outstanding processor for systems used in CPU-intensive applications using one or two processors. In particular, the higher core frequencies can mean significant performance increases in CPU-bound environments. Pentium Pro-based servers and workstations continue to provide the best system performance for memory-intensive applications and applications needing more than two processors. Users should evaluate their requirements to determine which processor best suits their needs. For more information on Pentium Pro and Pentium II performance and positioning, readers are encouraged to read the Compaq white paper Positioning Pentium II and Pentium Pro in Server Environments, document number 235A/0797.