The Mobile AMD Athlon™ XP-M Processor with 512KB L2 Cache

Technology and Performance Leadership for x86 Microprocessors

Jack Huynh

AMD
One AMD Place
Sunnyvale, CA 94088

Introduction: Continuing Performance Leadership of x86 Microprocessors

Founded in 1969, AMD has shipped more than 200 million PC processors worldwide. AMD processors are the power behind desktop and notebook PCs, and a new generation of servers and workstations. Since its introduction in 1999, the award-winning AMD Athlon™ processor has been known as an industry leader, enabling one of the highest system performance levels in the PC market. Since its launch in October 2001, the AMD Athlon XP processor and computer systems based on the AMD Athlon XP processor have won more than 30 awards worldwide, including *PC World's* World Class Award for Computer Product of the Year in June 2002. In all, the AMD Athlon processor family and systems based on such processors have won more than 100 awards worldwide, including *PC World*'s World Class Award for overall Product of the Year in 2000 and 2002. The mobile AMD Athlon processor family has provided industry-leading processing power to pave the road for new levels of end-user capability with notebook PC application areas. These areas span from gaming to productivity applications, and include digital photo editing, digital video, image compression, video encoding, audio compression, 3D modeling and animation, and speech recognition.

Engineering and technology leadership are key to performance leadership. AMD's engineering and technology leadership specific to the seventh-generation AMD Athlon processor family includes driving innovations such as instruction set extensions aimed at 3D applications (3DNow!™ Professional technology) at the processor instruction level, DDR memory and HyperTransport™ technology at the platform level, and 0.13-micron process with copper interconnect at the process technology level. With the introduction of the new mobile AMD Athlon XP-M processor with 512KB L2 cache on 0.13-micron process technology, AMD continues its tradition of technology innovation in enabling one of the highest levels of notebook PC performance. The discussion that follows provides an in-depth look at how the mobile AMD Athlon XP-M processor with 512KB L2 cache on 0.13-micron process technology increases the performance scalability of QuantiSpeed™ architecture. The differentiating features as well as the real-world application performance benefits of AMD PowerNow!™ technology and QuantiSpeed architecture will also be discussed.

Manufacturing Technology Leadership with Leading Edge 0.13-Micron Process Technology

The new mobile AMD Athlon XP-M processor, code-named “Barton,” is manufactured on state-of-the-art 130 nanometer (0.13-micron) copper technology and is the newest member to the family of seventh-generation AMD Athlon processors designed to meet the computationally-intensive requirements of software applications running on high-performance mobile systems. The mobile AMD Athlon XP-M processor with 512KB L2 cache based on 0.13-micron process technology increases the performance scalability provided by QuantiSpeed architecture by delivering higher clock speeds. The mobile AMD Athlon XP-M processor with 512KB L2 cache is also available in a low-power, small footprint µPGA package for thin and light notebook designs. The 0.13-micron process technology is designed to provide the thermal headroom necessary to scale frequency within the thermal limits of mobile platforms, thus maximizing overall performance. The mobile AMD Athlon XP-M processor with 512KB L2 cache on 0.13-micron process technology, like all mobile AMD Athlon XP processors, is also pin-compatible with AMD’s established Socket A infrastructure.

With the increased frequency scalability brought about by 0.13-micron technology combined with AMD PowerNow! technology and QuantiSpeed architecture, AMD continues to deliver compelling solutions for high-performance computing, and superb integer, floating point, and 3D multimedia performance for applications running on x86-based platforms.

AMD PowerNow!™ Technology: Optimizing Notebook System Battery Life while Delivering Performance on Demand

The addition of AMD PowerNow! technology within the mobile AMD Athlon XP-M processor builds upon the original AMD PowerNow! technology implemented in the mobile AMD-K6®-2+ and mobile AMD-K6-III+ processors. This technology allows the processor to dynamically modify delivered performance—based on application requirements—to enhance system battery life. To accomplish this, AMD PowerNow! technology-enabled dynamic software controls processor voltage and frequency based on application requirements. AMD is the first to bring a dynamic power-management solution to notebook PCs.

In first-generation mobile systems, the system processor runs at a fixed voltage and frequency. ACPI 2.0 (Advanced Configuration and Power Interface) processor power management is based on idle detection. When the operating system detects that the system is idle, it places the processor in a "Halt" or "Stop Grant" state. In the Halt or Stop Grant state, the processor is not executing any operating system or application code and its power level drops to a significantly lower level. ACPI 2.0 processor power management is effective for traditional applications such as Microsoft® Word, Outlook, Excel, PowerPoint, and other applications that are normally idle and waiting for user input. However, ACPI 2.0 processor power management is not generally effective with newer "real-time" or "time-dependent" software applications such as DVD players, soft modems, voice recognition, Internet telephone applications, video conferencing, Internet conferencing, audio recording, CD writing, MPEG-4 encoding, video editing, the playback of streaming audio or video such as MP3 and Windows® media, and games. For acceptable performance, these real-time or time-dependent applications require a steady stream of execution by the processor. At the same time, these types of applications often do not require the full performance of the processor. But because these applications are not characterized by periods of inactivity, ACPI 2.0 processor power-management techniques are not effective. For example, DVD playback might require only 50% of the processor's bandwidth, but ACPI 2.0 power management forces the processor to operate at 100% performance, thus wasting a significant amount of energy and, therefore, battery life.

Using dual speeds is another common technique for extending battery life. In dual-speed systems, the processor operates at full speed on AC power and operates in a low-performance mode with a static, lower frequency when the system uses battery power. While somewhat effective at extending battery life, this power-management technique can have an impact on performance and has major drawbacks. The main issue with dual-speed processor power management is that when a system uses battery power, a compromise must be made between performance and power consumption. The performance level when the system is battery-powered can be set high enough to provide an acceptable user experience for most applications. This performance level sets the processor performance higher than necessary for some applications, increasing power and reducing battery-powered run time. The performance level when the system is battery-powered can be set low to optimize battery life, but this setting produces a poor experience for the user with some applications. Dual-speed processor power management is optimal for neither performance nor extending battery life.

AMD has augmented the capabilities of ACPI 2.0 processor power management and has solved the problems associated with dual-speed processor power management by developing AMD PowerNow! technology. AMD PowerNow! technology allows a processor to dynamically switch between different processor performance states. A performance state is a combination of both processor core frequency and voltage. Software controls the transitions between the different performance states. This innovative approach to power management dynamically adjusts the performance level of the processor to match the level of performance needed by any application.

AMD PowerNow! technology has the following key benefits:

1) Extends system battery life through dynamic management
2) Provides performance on demand when required by the application
3) Operates automatically in the background

The processor executes instructions at a frequency dictated by the processor's utilization of the applications in operation, so real-time applications are not starved for processor cycles and the user's experience is optimal. AMD PowerNow! technology is designed to provide performance on demand and to run the processor in its most efficient performance state for the current applications. This power-management technique provides the right performance level while enhancing system battery life. Therefore, AMD PowerNow! technology is able to optimize battery powered run-time based on the processor performance required by the applications in operation.

These characteristics of AMD PowerNow! technology are designed to be particularly effective at prolonging system battery life because the processor's performance level should be increased above its most efficient power level for very short periods of time, only when necessary. Unlike the ACPI 2.0 or dual-speed power management, AMD PowerNow! technology is designed to provide performance on demand and power savings that scale with the application's demands.

An important aspect of AMD PowerNow! technology is that like ACPI 2.0, it is transparent to the user. There are no complicated settings, confusing application preferences, or artificial system modes. AMD PowerNow! technology helps to optimize processor performance *and* system battery life, simultaneously and transparently to the system user.

Another important feature of AMD PowerNow! technology is that it is designed to respond rapidly to the demands placed on the processor by various applications. AMD PowerNow! technology can update the processor's performance level as often as 30 times per second, thus quickly responding to the needs of the applications and the user with very few demands to the system's resources. AMD PowerNow! technology software is designed to offer versatility because it is not confined to work in conjunction with specific applications. Applications do not have to be modified to work with AMD PowerNow! technology. AMD PowerNow! technology measures the performance needs of the applications by directly monitoring the behavior of the core operating system.

AMD PowerNow! technology is truly designed to enhance the battery life of notebook PCs while delivering performance on demand.

QuantiSpeed™ Architecture: An Optimally Balanced x86 Microarchitecture for Real-World Application Performance

The microprocessor is a key component in determining the effectiveness of a computer system to execute specific tasks in the shortest amount of time. The amount of time required to complete specific software tasks is referred to as real-world application performance. Application performance is the function of two elements:

1) Clock speed of the processor, measured in Megahertz, or Gigahertz
2) The amount of work the processor can accomplish in a given clock cycle, measured in instructions per clock cycle (IPC)

Application Performance =
[work completed per clock cycle] x [clock speed] = IPC x Frequency

Different approaches can be taken to optimize the processor for application performance. AMD has worked to maintain a more balanced microarchitecture with a shorter pipeline designed for much higher IPC than competitive PC processors available in the market. Although other competitive processors enable deeper pipelines with fewer gates per clock to drive frequency improvements, deeper pipelines alone translate into less work per clock cycle. This reduced work per clock cycle or reduced IPC can only be offset by improvements in other areas, such as branch prediction and cache hit rates. Taken to the extreme, processor performance

can actually be reduced by forcing frequency improvements at the expense of IPC improvements.

This key point can be illustrated in PC office applications which tend to be branch-code intensive resulting in lower performance for deeper pipelines that must be flushed with a much greater performance penalty. As reaffirmed in the *Desktop Performance and Optimization for Intel Pentium® 4 Processor* paper (http://developer.intel.com/design/pentium4/papers/249438.htm), "Integer and basic office productivity applications, such as Word and spreadsheet processing, tend to have many branches in the code, thus reducing overall IPC capabilities. As a result, the associated branch penalties and performance on these applications does not generally scale as well with frequency and are more resistant to improvements in micro-architectural means, such as deeper pipelines."

The mobile AMD Athlon XP-M processor with QuantiSpeed architecture implemented on 0.13-micron process technology continues to exhibit the AMD Athlon processor family's balanced combination of improving clock frequency without compromising the amount of work done per clock cycle and therefore the IPC. The end result is a processor design that produces a high IPC as well as high operating frequencies, the optimum combination to deliver one of the highest levels of PC performance in real-world application environments.

QuantiSpeed architecture consists of four key differentiating features that enhance the application performance of the mobile AMD Athlon XP-M processor:

1) Nine-issue, superscalar, fully pipelined microarchitecture
2) Superscalar, fully pipelined floating point unit (FPU)
3) Hardware data prefetch
4) Enhanced Translation Look-aside Buffers (TLBs)

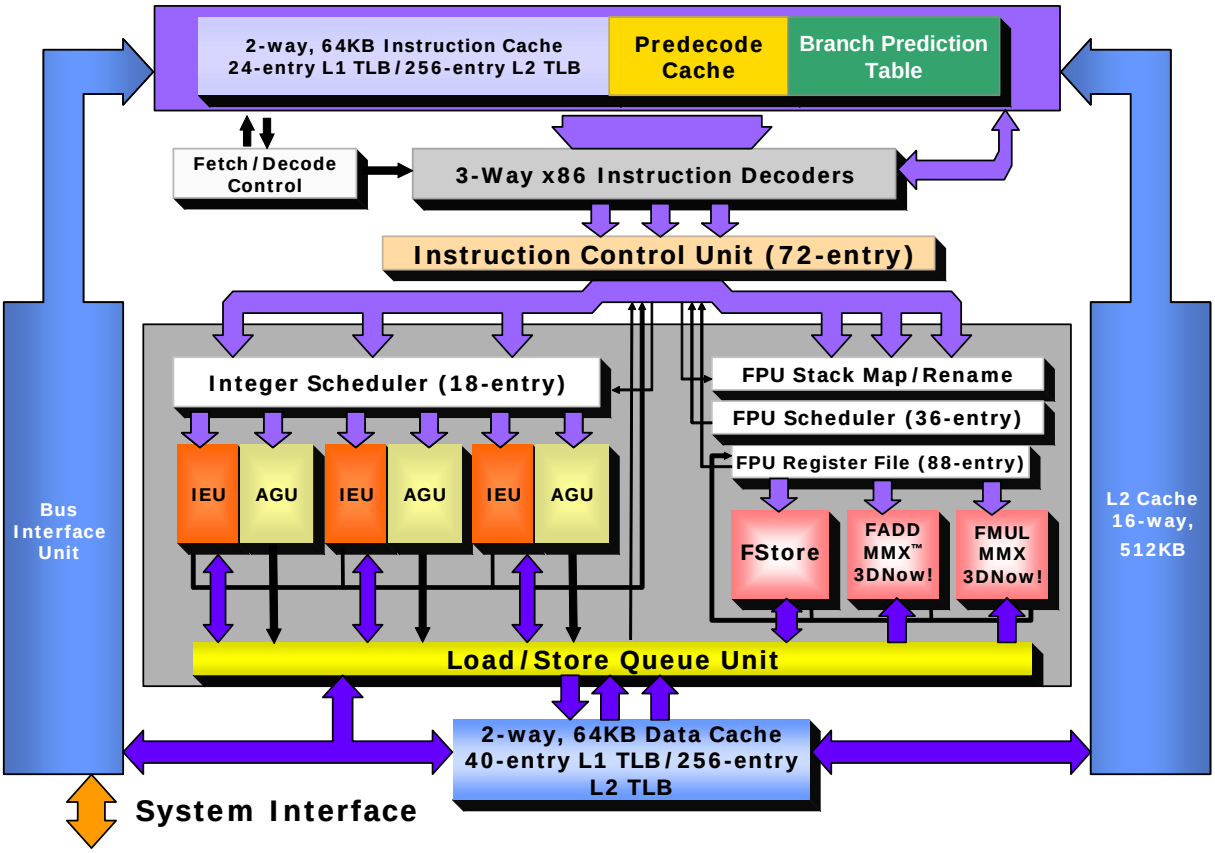

Figure 1: *Mobile AMD Athlon™ XP-M Microarchitecture Block Diagram*

QuantiSpeed™ Architecture: Nine-issue, Superscalar, Fully Pipelined Microarchitecture with High-Performance Cache Memory Architecture, and Three Full x86 Instruction Decoders

At the heart of QuantiSpeed architecture is a fully pipelined, nine-issue, superscalar processor core. The mobile AMD Athlon XP-M processor provides a wider execution bandwidth of nine execution pipes when compared with competitive x86 processors with up to six execution pipes. The nine execution engines are comprised of three address calculation units, three integer units, and three floating-point units.

In order to supply such a highly superscalar microarchitecture, the mobile AMD Athlon XP-M processor implements a very large on-chip cache architecture particularly in the L1 caches closest to the core. The mobile AMD Athlon XP-M processor's high-performance, on-chip cache architecture includes a dual-ported 128KB split-L1 cache with separate snoop ports, and an integrated full-speed,

16-way set-associative 512KB L2 cache using a 72-bit (64-bit data + 8-bit ECC) interface. The mobile AMD Athlon XP-M processor's large integrated full-speed L1 cache is comprised of two separate 64KB, two-way set-associative data and instruction caches which are much larger than the Pentium 4 processor's L1 cache (128KB vs. 8KB + 12KB μop). By featuring a larger L1 cache, applications running on the mobile AMD Athlon XP-M processor perform exceptionally fast because more instruction and data information is local to the processor. Applications exploit the larger caches by benefiting from the increased support of instruction and data set locality. The data cache also has eight banks to provide maximum parallelism for running multiple applications. It supports concurrent accesses by two 64-bit loads or stores. The instruction cache contains predecode data to assist multiple, high-performance instruction decoders. Both instruction and data caches are dual-ported and contain dedicated snoop ports designed to eliminate all system coherency traffic, common in systems with many devices, from interfering with application performance.

The mobile AMD Athlon XP-M processor also includes an integrated, full-speed, 16-way set-associative, exclusive 512KB L2 cache. When the processor requests data, it first searches the data in its L1 cache. If the processor finds the data in its L1 cache, the result is what is known as a cache hit and the processor retrieves the data from the low latency L1 cache. If the processor can not retrieve the data from its L1 cache, the processor attempts to retrieve the data in its L2 cache and once again attempts to obtain a cache hit. In the event of a cache miss, the processor must then request this data from the slower system memory. With an additional 256KB L2 cache over previous mobile AMD Athlon XP processors, the mobile AMD Athlon XP-M processor with 512KB L2 cache increases the performance of applications such as high-end gaming and digital media by keeping more frequently accessed instructions and data close to the CPU. Higher set-associativity increases the hit rate by reducing data conflicts. This translates into more possible locations where important data can reside in the L2 cache memory instead of system memory. With an exclusive cache architecture, the contents of the L1 caches are not duplicated in the L2 cache. This enables 512KB of L2 cache and 128KB of L1 cache for a total usable storage space of 640KB.

The mobile AMD Athlon XP-M processor cache architecture also supports error correction code (ECC) protection. With these cache architecture features, the mobile AMD Athlon XP-M processor provides reliable, high-performance computing.

When executing software, a processor begins by decoding the program's instructions and translating them into operations (or Ops) that the microprocessor

can execute. In order to continually feed the execution engine with data, the mobile AMD Athlon XP-M processor includes three x86 instruction decoders. Each decoder is capable of decoding three instructions per clock cycle. In comparison, the Pentium 4 is designed to decode only one instruction per clock cycle with the resource of only one x86 instruction decoder. Thus, the Pentium 4 has only one-third the maximum theoretical decode bandwidth of the mobile AMD Athlon XP-M processor. The decode bandwidth of the mobile AMD Athlon XP-M processor enables the processor to advantageously utilize the execution bandwidth capabilities of QuantiSpeed architecture, thereby improving IPC.

QuantiSpeed™ Architecture: Superscalar Fully Pipelined x86 Floating Point Unit (FPU)

The mobile AMD Athlon XP-M processor offers one of the most powerful, architecturally advanced floating-point units (FPU) delivered in an x86 microprocessor. The mobile AMD Athlon XP-M processor's three-issue, superscalar floating-point capability is based on three pipelined, out-of-order floating-point execution units, each with a one-cycle throughput. Using a data format and single-instruction multiple-data (SIMD) operations based on the MMX™ instruction model, the mobile AMD Athlon XP-M processor can deliver as many as four 32-bit, single-precision floating-point results per clock cycle.

FPU Microarchitecture

Three separate execution units in the mobile AMD Athlon XP-M processor's floating-point pipeline support x87 floating-point instructions, MMX instructions, and 3DNow! Professional technology instructions. The three execution units are:

1) **Fstore**—This is the floating point load/store pipeline that handles FP loads, stores, and miscellaneous operations.

2) **Fadd**—This is the adder pipeline that contains 3DNow! Professional technology, add, MMX ALU/shifter, and FP add execution units.

3) **Fmul**—This is the multiplier pipeline that contains an MMX ALU, MMX multiplier, reciprocal unit, FP, 3DNow! Professional technology instruction multiplier, and support for FDIV instructions.

In addition to its superscalar design, the mobile AMD Athlon XP-M processor's FPU is super pipelined. This technique supports higher clock frequencies and enables the FPU to process complex floating-point instructions more quickly and deliver high overall floating-point instruction throughput. In comparison, the FPU of the Pentium 4 processor only offers two execution units, one for Fadd and Fmul and one for Fstore. Thus, for example, the mobile AMD Athlon XP-M processor can do one floating point addition AND one multiplication per clock cycle, while the Pentium 4 processor can only do one multiplication OR one addition per clock cycle. The seventh-generation FPU of the mobile AMD Athlon XP-M processor incorporates additional features such as a 36-entry instruction scheduler and an 88-entry register file for independent, superscalar, out-of-order, speculative execution of floating-point instructions. With three separate execution units, the mobile AMD Athlon XP-M processor's superscalar FPU can boost the performance of floating point-intensive applications varying from commercial applications such as 3D modeling to consumer applications such as high-end gaming for mobile systems.

3DNow!™ Professional Technology: FPU Innovation of the Mobile AMD Athlon™ XP-M Processor Core

The mobile AMD Athlon XP-M processor with 3DNow! Professional technology adds 51 new instructions to the enhanced 3DNow! technology supported by the original AMD Athlon processor family. These 51 new instructions, along with the SIMD integer additions already included in enhanced 3DNow! technology, are compatible with Intel's SSE technology. Table 1 provides a breakout of the 3DNow! technology instruction set evolution.

Table 1: *AMD Processor Support of SIMD Instruction Set Extensions to the x86 Instruction Set Architecture*

AMD Processor:	AMD-K6®-2 Processor	AMD Athlon™ Processor	AMD Athlon XP-M Processor
3DNow!™ technology version supported:	3DNow! technology	Enhanced 3DNow! technology	3DNow! Professional technology
Description of instructions supported:	Original 3DNow! technology extensions	3DNow! technology plus 19 MMX extensions (part of SSE) plus five DSP/communications extensions	Enhanced 3DNow! technology plus 51 SSE extensions (completing SSE support)

3DNow! technology and SSE are largely complementary architectural enhancements. By implementing them in a variety of ways, software developers are able to determine how they can utilize the advanced architectural capabilities enabled by SIMD instruction set extensions. Examples of applications most able to benefit from the use of these instruction set extensions include speed recognition, video encoding/decoding, and 3D graphics generations.

Many current software applications that are SIMD-optimized use different code paths to benefit from 3DNow! technology or SSE, depending on the processor architecture on which these applications are executed. AMD processor architectures preceding the mobile AMD Athlon XP-M processor only supported 3DNow! or enhanced 3DNow! technology, which yielded the following three code path scenarios for developers:

1) Software optimized exclusively for AMD processor architectures with 3DNow! technology use their 3DNow! technology-optimized code path on AMD processors supporting 3DNow! technology.

2) Software optimized for both AMD processor architectures with 3DNow! technology, and other x86 industry processor architectures supporting SSE, use their 3DNow! technology-optimized code path on AMD processors supporting 3DNow! technology.

3) Software optimized exclusively for other x86 industry processor architectures supporting SSE use the non-optimized code path on AMD processor architectures.

With the advent of 3DNow! Professional technology, the mobile AMD Athlon XP-M processor can seamlessly allow SIMD-optimized software in the third scenario above to recognize SSE support and run the optimized code path for increased performance. The recognition of SSE support in 3DNow! Professional technology is performed automatically by PC software applications that use industry standard feature flags, provided in the CPUID instruction to automatically recognize SSE support and run the optimized code path. This means that with 3DNow! Professional technology's support for both 3DNow! and SSE technologies, the mobile AMD Athlon XP-M processor is able to take advantage of the performance gains offered by SIMD-optimized software applications.

Not only is the mobile AMD Athlon XP-M processor designed to benefit from existing software applications supporting 3DNow! and SSE technologies, but in the future, software developers are planned to have the ability to utilize the strength of both 3DNow! and SSE technology when optimizing code paths for AMD processor architectures that support 3DNow! Professional technology. The mobile AMD Athlon XP-M processor enables this advanced level of SIMD optimization by allowing 3DNow! and SSE instructions to be executed in the same code path.

QuantiSpeed™ Architecture: Hardware Data Prefetch

To further enhance processor IPC and, therefore, processor performance, the mobile AMD Athlon XP-M processor also uses hardware data prefetch technology. This hardware data prefetch technology observes memory accesses, looks for regular access patterns, and speculatively fetches the cache line with the data into the processor's L2 data cache in advance of the actual data access, therefore reducing the average latency seen by the processor in accessing memory. In the past, data prefetch was supported through the instructions introduced in 3DNow! and SSE technologies. However, for the processor to take advantage of this capability, software applications had to be specifically optimized with the 3DNow! and SSE instructions. The mobile AMD Athlon XP-M processor is designed to automatically optimize performance on existing software that had not previously been optimized using the hardware data prefetch instructions supported by 3DNow! Professional technology.

Benefits of the mobile AMD Athlon XP-M processor's hardware data prefetching are observed more in high-end, data-intensive applications that access larger arrays of data. Performance also benefits by not occupying processor instruction execution bandwidth required by software prefetching instructions. The optimization is most effective when coupled with high-bandwidth system memory transfer capability, now available to the processor by platforms such as those optimized to support DDR memory.

QuantiSpeed™ Architecture: Exclusive and Speculative Translation Look-aside Buffers (TLBs)

The mobile AMD Athlon XP-M processor features advanced, two-level Translation Look-aside Buffer (TLB) structures for both instruction and data address translation. The mobile AMD Athlon XP-M processor's Level 1 (L1) Instruction TLB (I-TLB) holds 24 entries, the L1 Data TLB (D-TLB) holds 40 entries, and the L2 I-TLB and D-TLB each hold 256 entries.

To reduce the incidence of TLB entry conflicts, the L1 and L2 TLB structures adopt an exclusive architecture design. With an exclusive TLB architecture, the L1 TLBs can contain entries that are not duplicated in the L2 TLBs, enabling the combination of L1 TLB and L2 TLB sizes for a larger total available entry space on both the instruction and data TLBs. By reducing the number of conflicts caused by holding more TLB entries within the processor, performance increases on high-end, data-intensive applications that encounter instruction sequences that no longer have to wait for TLB entries to be reloaded during execution.

The TLB structures of the mobile AMD Athlon XP-M processor also have the ability to enter data TLB misses in the TLBs speculatively. The mobile AMD Athlon XP-M processor allows TLB entries to be written speculatively before the first instruction is completed, while preserving proper instruction execution ordering which removes the serialization effect and results in improved system performance.

Conclusion: Technology and Performance Leadership of x86 Microprocessors

With these key differentiating features of the mobile AMD Athlon XP-M processor with QuantiSpeed architecture…

- **0.13-micron process technology**—Provides further thermal headroom necessary to scale frequency within the thermal limits of mobile platforms for AMD processors, thus maximizing overall processor performance
- **AMD PowerNow!™ technology**—Industry's first dynamic power-management solution to optimize notebook system battery life while delivering performance on demand
- **512KB L2 cache**—Increases the performance of applications such as high-end gaming and digital media by keeping more frequently accessed instructions and data close to the CPU.

QuantiSpeed™ architecture:

- **Nine-issue, superscalar, fully pipelined microarchitecture**—Provides a wide executing bandwidth to improve overall productivity
- **Superscalar, fully pipelined FPU**—Increasing performance of floating point-intensive applications while offering 3DNow! Professional technology support
- **Hardware data prefetch**—Increasing performance on high-end software applications using high-bandwidth system capability, especially with DDR memory
- **TLB enhancements**—Increasing performance of high-end, data-intensive applications

…AMD continues to accelerate technology innovations while meeting the computationally intensive requirements of software applications including:

- **3D applications**—3D modeling, high-end gaming, etc.
- **Multimedia/digital content creation applications**—Photo and video editing, video encoding and decoding, image compression, soft DVD, MP3 encoding and decoding, etc.
- **High-end applications**—Digital publishing, speech recognition, CAM, digital prototyping, etc.

With industry-leading performance across these and a number of other applications, the mobile AMD Athlon XP-M processor with 512KB L2 cache implemented on 0.13-micron process technology, continues to increase the performance scalability provided by QuantiSpeed architecture by delivering higher clock speeds and industry-leading processor performance. The mobile AMD Athlon XP-M processor with 512KB L2 cache and AMD PowerNow! technology continues in the tradition of the AMD Athlon processor family in providing compelling levels of delivered system performance for today's and tomorrow's applications as well as helping to extend system battery life for users on the go.

AMD Overview

AMD is a global supplier of integrated circuits for the personal and networked computer and communications markets with manufacturing facilities in the United States, Europe, Japan, and Asia. AMD, a Fortune 500 and Standard & Poor's 500 company, produces microprocessors, Flash memory devices, and support circuitry for communications and networking applications. Founded in 1969 and based in Sunnyvale, California, AMD had revenues of approximately $2.7 billion in 2002. (NYSE: AMD).