AN ANALYSIS OF THE ATHLON MP PROCESSOR

Submitted by

LAVANYA SITA TEKUMALLA

00320994

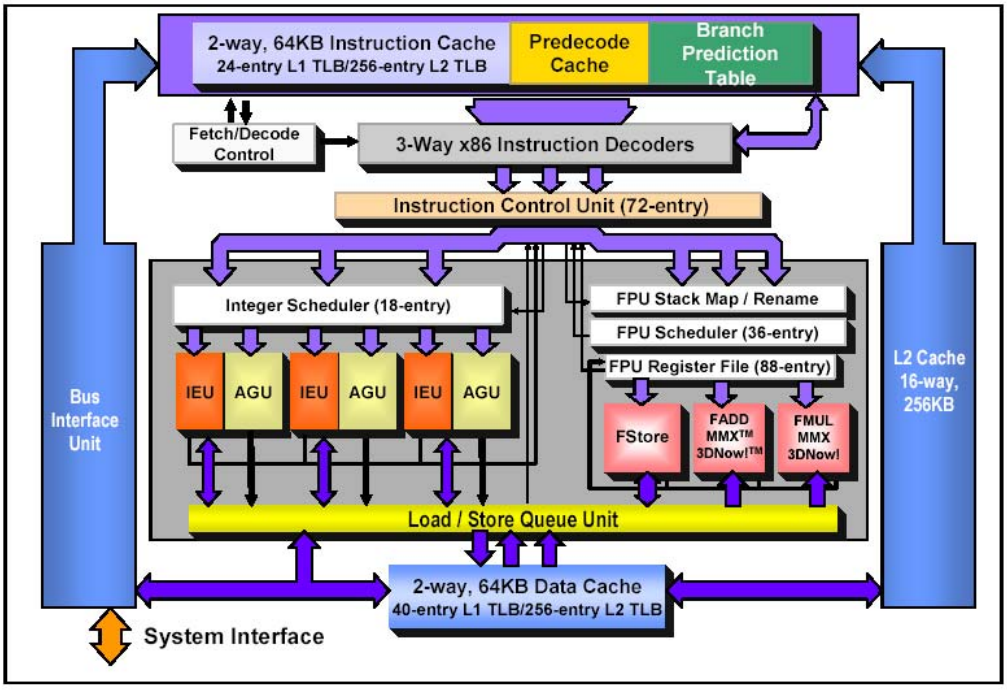

Figure 1: *AMD Athlon™ MP Microarchitecture Block Diagram*

CONTENTS

PART1- FEATURES

PART II- ANALYSIS

THE ATHLON MP PROCESSOR

1. OVERVIEW

The Athlon MP processor is specially designed for multiprocessor environments. It is particularly tailored to optimize the execution of multithreded and mission critical applications and to meet the computation intensive workloads in the server and workstation domain.

The key features of the Athlon MP processor are.

1. Quantispeed architecture
2. Smart Multiprocessing technology
3. Extended Instruction Set for superscalar SIMD and 3D graphics generation.
4. 266 MHz system bus which gives higher system bandwidth
5. 384 K Full speed on-chip Cache

2. INSTRUCTION SET:

The Athlon MP basically uses the X86 architecture. This processor's register file consists of general-purpose, segment, floating-point, MMX/3DNOW, task, debug, memory-management registers. There are 8 32-bit x86 general-purpose registers (EAX, EBX, ECX, EDX, EDI, ESI, ESP, EBP) used to hold the integer data or memory pointers used by the instructions.

The instructions in Instruction Set can be identified as Arithmetic and Logical, Data Transfer, Control, System, Floating-point, MMX instructions, Graphics (pixel operations etc).

However, it introduced some extensions like 3DNow and MMX. These include,

21 original 3DNow!™ instructions enabling superscalar SIMD

19 additional instructions to enable improved integer math calculations for speech or video encoding and improved data movement for Internet plug-ins and other streaming applications

5 DSP instructions to improve soft modem, soft ADSL, Dolby Digital surround sound, and MP3 applications

52 SSE instructions with SIMD integer and floating point additions offer compatibility with Intel's SSE technology

.

3. THE MICRO ARCHITECTURE

The AMD Athlon MP processor micro architecture approach implements the x86 instruction set. Complex x86 instructions are converted into simpler operations and executed. The processor is fully pipelined and there are nine execution units. There are three Integer units, three Floating Point units and three address calculation units.

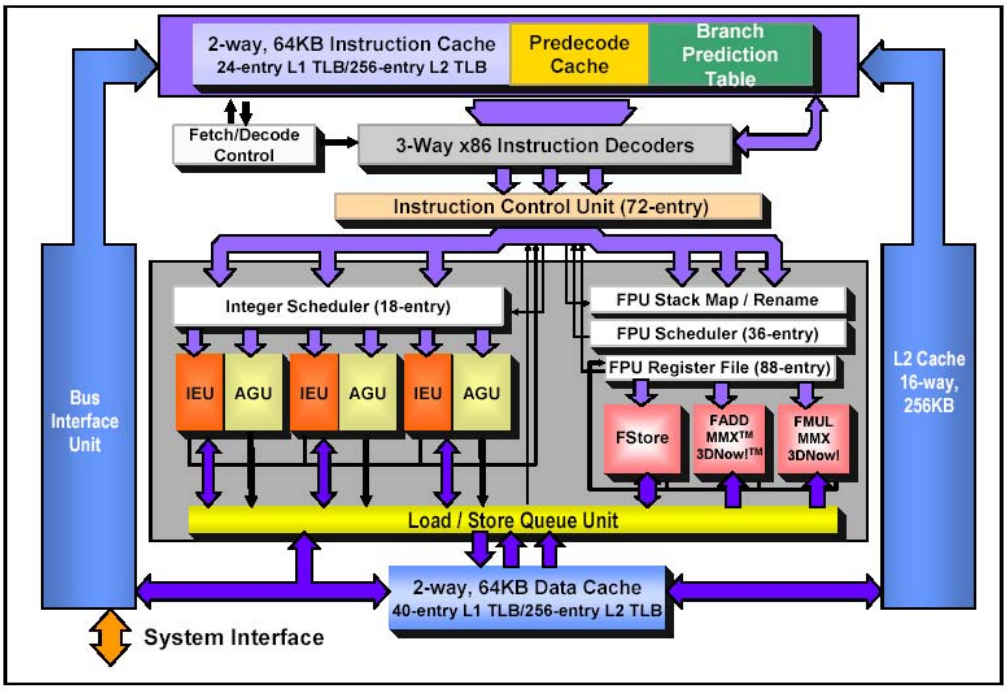

Figure 1: *AMD Athlon™ MP Microarchitecture Block Diagram*

3.1 INSTRUCTION FETCH

The AMD Athlon MP processor instruction fetcher reads 16-byte aligned code from the instruction cache. The instruction bytes are then merged into a 24-byte instruction queue.

3.2 .INSTRUCTION DECODE:

The AMD Athlon MP processor includes three X86 instruction decoders.

Each decoder is capable of decoding there instructions per clock cycle.

All instructions (x86, x87, 3DNow, and MMX instructions) are classified into two types of decodes

Direct Path : DirectPath instructions are common instructions that are decoded directly in hardware. Up to three DirectPath instructions can be selected for decode per cycle.

Vector Path : VectorPath instructions are more complex instructions that require the use of a sequence of multiple operations issued from an on-chip ROM. Only one VectorPath instruction can be selected for decode per cycle. DirectPath instructions and VectorPath instructions cannot be simultaneously decoded.

3.2.1 EARLY DECODING:
The DirectPath and VectorPath decoders perform earlydecoding of instructions into Macroperations in program order. Early decoding produces three MacroOPs per cycle from either path. The outputs of both decoders are multiplexed together and passed to the next stage in the pipeline, the instruction control unit.
When the target 16-byte instruction window is obtained from the instruction cache, the predecode data is examined to determine which type of decode is necessary.

DIRECT PATH DECODER:
A DirectPath instruction is limited to those x86 instructions that can be further decoded into one or two OPs. The length of the x86 instruction does *not* determine DirectPath instructions. A maximum of three DirectPath x86 instructions can occupy a given aligned 8-byte block. 16-bytes are fetched at a time. Therefore, up to six DirectPath x86 instructions can be passed into the DirectPath decode pipeline.

VECTOR PATH DECODER:
Instructions requiring two or more MacroOPs proceed down the VectorPath pipeline. The sequence of MacroOPs is produced by an on-chip ROM known as the MROM. The VectorPath decoder can produce up to three MacroOPs per cycle. Decoding a VectorPath instruction may prevent the simultaneous decode of a DirectPath instruction.

3.3 ISSUE: This processor can issue 9 instructions at a time.

3.4 THE INSTRUCTION CONTROL UNIT:
The following are the functions of the Instruction control Unit.

1. Control the centralized in-flight reorder buffer,
2. Control the integer scheduler, and the floating-point scheduler.
3. MacroOP dispatch and MacroOP retirement
4. Register and flag dependency resolution and renaming
5. Execution resource management, interrupts, exceptions, and branch mispredictions.

3.5 REORDER BUFFER

The reorder buffer in the Athlon MP is a centralized, fixed-issue one. The ICU takes the three MacroOPs per cycle from the early decoders and places them in the reorder buffer. This buffer is organized into 24 lines of three MacroOPs each. The reorder buffer allows the ICU to track and monitor up to 72 in-flight MacroOPs (whether integer or floating-point) for maximum instruction throughput.

3.6 INTEGER OERATIONS
3.6.1 INTEGER SCHEDULER:
The reservation stations feed three integer execution positions or pipes. They are six entries deep, for a total queuing system of 18 integer MacroOPs. Each reservation station divides the MacroOPs into integer and address generation OPs, as required and sends them to the appropriate execution unit.

3.6.2 INTEGER EXECUTION UNIT: The integer execution pipeline consists of three identical pipes. Each integer pipe consists of an integer execution unit and an address generation unit. Each of the three IEUs are general purpose in that each performs logic functions, arithmetic functions, conditional functions, divide step functions, status flag multiplexing, and branch resolutions. The AGUs calculate the logical addresses for loads, stores, and LEAs.

3.7 THE SUPERSCALAR FULLY PIPELINED FLOATING POINT UNIT:
The floating-point execution unit (FPU) is implemented as a coprocessor that has its own out-of-order control in addition to the data path. The FPU handles all register operations for x87 instructions, all 3DNow! operations, and all MMX operations. The FPU consists of a stack renaming unit, a register renaming unit, a scheduler, a register file, and three pipelined out of order floating point units, each with one cycle throughput. So three floating point instructions can be issued at a time. Using a data format and single-instruction multiple-data (SIMD) operations based on the MMX instruction model, the AMD Athlon MP processor can deliver as many as four 32-bit, single-precision floating point results per clock cycle.

The FPU also encorporates other features such as 36 entry instruction scheduler(Described Next) with an 88 entry register file for independent out of order speculative execution of floating point instructions.

3.7.1 FLOATING POINT SCHEDULER: The floating-point scheduler handles register renaming and has a dedicated 36-entry scheduler buffer organized as 12 lines of three MacroOPs each. It performs OP issue, and out-of order execution.

3.7.2 EXECUTION UNITS
The floating point pipeline of the Athlon MP processor contains the following execution units:

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.

3.8 THE LOAD STORE UNIT
The load-store unit (LSU) manages data load and store accesses to the L1 data cache and, if required, to the L2 cache or system memory. The 44-entry LSU provides a data interface for both the integer scheduler and the floating-point scheduler.

3.9 EXCLUSIVE AND SPECULATIVE TRANSLATION LOOK-ASIDE BUFFERS :

The AMD Athlon MP processor has advanced, two-level Translation Lookaside Buffer (TLB) structures for both instruction and data address translation. The AMD Athlon MP processor's Level 1 (L1) Instruction TLB holds 24 entries, the L1 Data TLB holds 40 entries.

The Level2 Instruction-TLB and Data-TLB each hold 256 entries. 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.

4 MEMORY ARCHITECTURE,

4.1 *L1 CACHE*

The AMD Athlon MP processor's ,on-chip cache architecture includes a dual-ported 128K (two separate 64K) split-L1 cache with separate snoop ports, and an integrated full-speed, 16-way set-associative, 256K L2 cache using a 72-bit (64-bit data + 8-bit ECC) interface. The integrated L1 cache is comprised of two separate 64K, two way set-associative data and instruction caches.

4.1.1 INSTRUCTION CACHE: There is a 64-Kbyte L1 instruction cache. The L1 instruction cache is organized as a 64-Kbyte, two-way, setassociative array. Each line in the instruction array is 64 bytes long. Functions associated with the L1 instruction cache are instruction loads, instruction prefetching, instruction predecoding, and branch prediction. Requests that miss in the L1 instruction cache are fetched from the L2 cache or, subsequently, from the local memory using the bus interface unit (BIU).

The Cache line replacement is based on a least-recently used (LRU) replacement algorithm. The L1 instruction cache has an associated two-level translation look-aside buffer (TLB) structure. The first-level TLB is fully associative and contains 24 entries (16 that map 4-Kbyte pages and eight that map 2-Mbyte or 4-Mbyte pages). The second-level TLB is four-way set associative and contains 256 entries, which can map 4-Kbyte pages.

The instruction cache contains pre-decode data to assist multiple, high-performance instruction decoders.

4.1.2: DATA CACHE: 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 L1 data cache contains two 64-bit ports. It is a write-allocate and write-back cache that uses an LRU replacement policy. It is divided into 8 banks where each bank is 8 bytes wide. In addition, this cache supports the MOESI (Modified, Owner, Exclusive, Shared, and Invalid) cache coherency protocol and data parity.

The L1 data cache has an associated two-level TLB structure. The first-level TLB is fully associative and contains 32 entries (24 that map 4-Kbyte pages and eight that map 2-Mbyte or 4-Mbyte pages). The second-level TLB is four-way set associative and contains 256 entries, which can map 4-Kbyte pages.

Both instruction and data caches are dual-ported and contain dedicated snoop ports to eliminate all system coherency traffic, common in systems with many devices, from interfering with application performance.

4.2 L2 CACHE

The AMD Athlon MP processor also includes an integrated, full-speed, 16-way set-associative, exclusive 256K L2 cache. This translates into more possible locations in which 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 L2 cache contains only victim or copy-back cache blocks that are to be written back to the memory subsystem as a result of a conflict miss.
This enables 256K of L2 cache and 128K of L1 cache for a total usable storage space of 384K.

VICTIM BUFFER:
processor victim buffer contains data evicted from the L1 cache. It features up to eight 64-byte entries, where each entry corresponds to a 64-byte cache line.

The AMD Athlon MP processor cache architecture also supports error correction code (ECC) protection

5 SYSTEM BUS:

The AMD Athlon system bus consists of a pair of unidirectional 13-bit address and control channels and a bidirectional 64-bit data bus. It sipports low-voltage swing, multiprocessing, clock forwarding, and fast data transfers. The clock forwarding technique is used to deliver data on both edges of the reference clock, therefore doubling the transfer speed. A four-entry 64-byte write buffer is integrated into the BIU. The write buffer improves bus utilization by combining multiple writes into a single large write cycle. By using the AMD Athlon system bus, the

The features of the system bus can be summarized as follows:

1.266MHz3 operation, designed to deliver a peak throughput of up to 2.1GB/s per system bus

2.Split-transaction architecture

3.Cache-coherency protocol

4.ECC capability

5.64-bit data path

6.Packetized request transactions

5.1 DUAL POINT-TO-POINT HIGH-SPEED SYSTEM BUS

Dual point-to-point high-speed system buses allow two processors to run independently without the overhead of sharing a common system bus.

THE SPLIT TRANSACTION BUS

the AMD Athlon system bus is composed of three separate
buses. These operate independently and concurrently at 2ggMHz: The three busses are as follows:

1. Processor-to-system bus
2. System-to-processor bus
3. Data bus

5.1.1 - THE PROCESSOR-TO-SYSTEM BUS is the channel in which the processor issues
requests/commands (memory read, memory write, etc.) to the system. This is a unidirectional bus controlled only by the processor. Only packets containing command, address , and other information (but not data) are transferred.

5.1.2 - Data information is transferred over the DATA BUS. The data bus is a bi-directional bus used to transfer data packets between the
processor and system elements in response to requests from their respective buses. Each data packet contains an ID tag for association with the corresponding processor-to-system or system-to-processor request.

5.1.3 -THE SYSTEM-TO-PROCESSOR BUS is the channel in which the system issues requests/commands to the processor. This is also a uni-directional bus controlled only by

the system controller. Similar to the processor-to-system bus, only packets—request and command information—are transferred on this bus. Data information is transferred over the data bus.

5.2 BUS SNOOPING CAPABILITY:
MP implements a snooping mechanism in which the processors leverage the independent processor-to-system, system-to-processor, and data channels of the AMD Athlon system bus to create a "virtual" snooping channel. A processor can transfer data while simultaneously receiving snoop information, or a processor can broadcast snoop information while simultaneously receiving data.

6 **CACHE-COHERENCY PROTOCOL**
OPTIMIZED MOESI (Modified, Owner, Exclusive, Shared and Invalidate)

The MOESI protocol is similar to the MESI(Modified, Owner, shared and Invalidate) protocol; But it has an additional state (Owner) is added. The addition of this state allows the processor owning the cache block to directly supply cache data to the requestor. In this respect, the requestor does not have to wait for main memory to be updated to receive the requested data.

When running two processors, both processors must run at the same clock frequencies. Also, the AMD Athlon system bus and DDR memory subsystem are locked in frequency. Whether running one processor or two processors, the AMD Athlon system bus and DDR memory subsystem must operate at the same speed.

7 **: CHARACTERISTICS OF THE CHIP:**
Die Size: Approximately, 37.2 million transistors from 80 mmsquare.

Athlon MP uses the 0.13 micron process technology which helps in gaining higher clock speeds with less thermal power.

8**. SOME CHIPSETS USING THE ATHLON MP PROCESSOR**

8.1 AMD-760MP CHIPSET.

The AMD-760MP chipset uses a Point-to-Point bus protocol to help reduce memory bus traffic as all inner-CPU communication occurs without using the memory bus. Point-to-Point FSB bus is where each CPU gets their own dedicated path to the North Bridge, inner-CPU communication can be done in the North Bridge.

Each Athlon MP gets 2.1GB/s of bandwidth to/from the North Bridge General features of North Bridge AMD-762 are Dual-processor support, SmartMP technology, 266MHz (133

double pumped) FSB for each CPU, up to 4 PC2100 registered ECC DDR-SDRAM DIMM's for 4GB memory, AGP 4X, 32-bit/33MHz & 64-bit/33MHz PCI support.

General features of South Bridge AMD-766 are Dual-channel ATA-100 support, 4 USB 1.1 ports.

8.2 AMD-760MX CHIPSET

The main difference is it features 66MHz 64-bit PCI bus with support for up to two devices off the North Bridge and a new South Bridge with support for legacy 32-bit 33MHz PCI bus.

The new AMD-768 South Bridge connects with the North Bridge via a 32-bit 66MHz PCI bus that gives a total of 266MB/s bandwidth.

PART II

ANALYSIS:

1. *POSITIVE ASPECTS.*

This processor has several features such as parallel and specialized execution units, advanced two-level branch prediction, speculative execution, out-of-order execution, register renaming and data forwarding which enhance its performance

1. Parallel Compute Resources Benefit From Out-of-Order Approach
The extra complexity of creating an out-of-order machine is wasted if there aren't parallel compute resources available for taking advantage of those exposed instructions. Here is where Athlon really shines. The microarchitecture can execute 9 simultaneous RISC instructions.

2. It has a ***high decode bandwidth*** which enables the processor to advantageously utilize the execution bandwidth capabilities of QuantiSpeed architecture, thereby improving IPC.

(Each decoder is capable of decoding three instructions per clock cycle. In comparison, the Xeon processor is designed to decode only one instruction per clock cycle with the resource of only one x86 instruction decoder. Thus, the Xeon processor has only one-third the maximum theoretical decode bandwidth of the AMD Athlon MP processor.)

3. Pre-Decoding Uses Extra Cache Bits
To deal with the complexities of the x86 instruction set, AMD does some early decoding of x86 instructions as they are fetched into the L1 instruction cache. These extra bits help mark the beginning and end of the variable-length instructions, as well as identify branches for the pre-fetcher (and predictor). These extra bits and early (partial) decoding give some of the benefits of a trace cache, though there is still latency for the completion of the decoding.

4. Final Decoding Follows 2 Different Paths
The "DirectPath Decoder" generates MacroOp's that take one or two Ops. The "VectorPath Decoder" fetches longer instructions from ROM. Athlon can supply three MacroOp's/cycle to the instruction decoder , and later they'll enter the instruction scheduler, equating to a maximum of 6 Ops/cycle decode bandwidth. (The actual decode performance of course depends on the type of instructions.)

5. **FPU**: The FPU of the AMD Athlon MP processor incorporates 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 FPU can boost the performance of floating point-intensive

applications varying from commercial applications such as 3D modeling and CAD to consumer applications such as digital video and audio editing for workstations.

(In comparison, the FPU of the Xeon processor only offers two execution units, one for both Fadd and Fmul and one for Fstore. Thus, as an example, the AMD Athlon MP processor can do one floating point addition AND one multiplication per clock cycle, while the Xeon processor can only do one multiplication OR one addition per clock cycle.)

6. ***Enhanced Instruction Set***: 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. With the advent of 3DNow! Professional technology, the AMD Athlon MP processor can allow SIMD-optimized software 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 software applications that use industry standard feature flags, provided in the CPUID instruction to automatically recognize SSE support and run the optimized code path.

7. It is compatible with all types of high speed memories including PC100 and PC133.

8. It has MMX media enhancement technology and 3DNOW! Instruction set. It has Accelerated Graphic Port (AGP) support.

9. ***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. This it actually reduces the time it takes to feed the processor critical data, increasing work throughput. As a result, application performance is automatically enhanced

Benefits of the AMD Athlon MP processor's hardware data prefetching are observed more in high-end, data-intensive server applications that access larger arrays of data. Performance also benefits by not occupying processor instruction execution bandwidth required by software prefetching instructions.

10. ***Smart MP Technology***

a. *Dual point-to-point high-speed system buses* - Allows two processors to run independently without the overhead of sharing a common system bus

b. *Innovative bus-snooping capability* - Offers high speed communication between processors in a multiprocessing system

c. *Optimized MOESI cache-coherency protocol* - Reduces memory traffic and allows faster access to cached data

11. The multiprocessor architecture of the Athlon MP including its point-to-point bus protocol also lends itself quite well to a high performance database or web serving environment.

12. The Athlon's excellent performance in Floating point heavy scientific applications makes AMD more suitable than even Intel in a number of niche distributed computing cluter environments. (It is believed that the amount of performance you can get out of a 1U 2P Athlon MP server easily eclipsed anything in the price range)

13. Cache Architecture Emphasizes Size to Achieve High Hit Rate

AMD has chosen to implement large L1 caches. The L1 instruction and data caches are each 2-way, 64KB caches. The L1 instruction cache has a line-size of 64 bytes with a 64-byte sequential pre-fetch. The L1 data cache provides a second data port to avoid structural hazards caused by the superscalar design. The L2 cache is a 16-way, 256KB unified cache, backed up by the fast bus.

14. Integrated L2 cache makes the processor faster.

15. ***L1 and L2 caches are exclusive***. So effectively, essentially 384 KB of onboard, full speed cache is available. This is again a great step forward in processor efficiency, as data is not duplicated between the two caches. This gives more effective cache space, which leads to more data being stored directly in the processor, instead of outside in the main memory.
    The Athlon's powerful execution units, very large L1 cache and short pipeline make it an excellent CPU for servers

16. ***An exclusive and speculative 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.

17. ***Branch Prediction Logic is a Combination of the Latest Methods***

There is a 2048-entry Branch Target Buffer that caches the predicted target address. This works in concert with a Global History Table that uses a "bimodal counter" to

predict whether branches are taken. If the prediction is correct, then there is a single-cycle delay to change the instruction fetcher to the new address

18. AMD works closely with leading software and O/S vendors to provide good performance of AMD processors on the applications you use most. AMD Athlon processors are designed for compatibility with more than 60,000 applications and are compatible with 32-bit x86 operating systems including Microsoft® Windows®, and multiple Linux and Unix operating systems.

2. *NEGATIVE ASPECTS*

Integrated Instruction and Data cache: Sharing code and data in the same 64-byte cache line , may cause the L1 caches to thrash (unnecessary castout of code/data) in order to maintain coherency between the separate instruction and data caches.

The processors in Athlon family are without any built in thermal protection features. If a working Athlon loses contact with its heat sink for more then a few seconds catastrophic failure could result, leaving the CPU irreparably damaged

Athlon are cooled by direct processor die contact with the heat sink, as opposed to using an integrated heat spreader like the latest P3 and P4 designs. This design choice results in a more fragile package and creates a situation where the interface between the CPU and the heat sink is more critical.

The Xeon MP processors have a 256KB on-die L2 cache and an on-die L3 cache. The on-die L3 cache is available in 512KB and 1MB sizes. The athlon MP processors do not have an L3 cache.

3. *CONCLUSION:*

The AMD Athlon MP provides a good solution to serve the needs of multiprocessing servers and workstations. It has various features that outweigh its few inadequacies and give it a performance edge.

REFERENCES:

Information has been collected from the following websites

http://www.amd.com

http://www.anandtech.com

http://www.siliconacoustics.com/athlon.html

http://www.extremetech.com