

# IT TAKES AN ARCHITECTURE TO MAKE A MULTI-CORE PROCESSOR

Part I in a Series of AMD White Papers on 64-bit Computing

On March 6, 2000, AMD became the first supplier to introduce a microprocessor that ran at the then-astounding speed of 1GHz. AMD was proud of that achievement, but even as it basked in the glory of the moment, its architects and engineers were laying the groundwork for the next era of computing performance—an era that would include 64-bit multi-core processors that leave a 1GHz AMD Athlon™ processor in the dust. An era that would focus on true performance, rather than megahertz. Back then, many scoffed at this vision. Today, looking back over the past five years, most have come around to AMD's way of thinking. In April 2003 the company fulfilled the first part of this vision as AMD delivered its first AMD Opteron™ processors; exactly two years later in April 2005, AMD fulfilled the second part as it delivered its first Dual-Core AMD Opteron processors. To carry out its plan, AMD needed to develop an advanced architecture, an ecosystem to support that architecture, and the advanced manufacturing technology to produce the silicon implementations of that architecture. This document lays out the architectural underpinnings of the 64-bit multi-core vision, and explore some of the ramifications of that vision with regard to performance and datacenter operations. Future documents will explore issues regarding the ecosystem built around the AMD64 vision and the process technology used to deliver AMD64 products, so stay tuned for what's to come.



## Why 64-bits Matters

### MORE EFFECTIVE USE OF MEMORY CONFIGURATION

Before AMD began shipping 64-bit AMD Opteron processors in 2003, users in need of 64-bit computing were forced to pay dearly for the privilege. The problems they needed to solve were so big they did not fit inside 32-bit computers based on the 32-bit AMD Athlon MP or 32-bit Intel Xeon processors. These users were forced to resort to one or another of the proprietary 64-bit RISC processors offered by HP, IBM, Silicon Graphics, or Sun, and usually had to pay in excess of \$50,000 for entry-level systems with enough memory to handle their problems. But DRAM prices decrease at highly predictable rates, and over time users can often afford larger memory configurations. Today, users can buy 4GB of DDR DRAM, in the formats needed for most servers, for about \$500, but their 32-bit operating systems and applications rarely have the ability to address memory configurations this large. 64-bit architecture is the key to the effective use of large memory configurations on servers and clients. Applications that can benefit from the use of more memory need to run on 64-bit processors to fully capture those benefits. Fortunately, the AMD Opteron processor arrived in the market just in time to make 64-bit computing affordable.

## 64-BIT ARITHMETIC ENABLES HIGHER PRECISION MATH

Some big problems may not require huge memory structures but still can benefit from a move to 64-bit processing. Users who need absolute precision in their calculations often resort to integer (fixed point) representations of their data, in order to avoid the inevitable rounding errors that occur with floating point numeric representations. 32-bit processors can only represent integers as large as 4,294,967,295, and thus cannot be used to store very large numbers like the U.S. national debt (\$8,210,569,853,980.08 as of 1/06/06). 64-bit processors like the AMD Opteron™ processor have the ability to handle integers as large as 18,446,744,073,709,551,615. This means that even if the national debt continues to mount at the current rate, as long as the U.S. Treasury sticks with AMD Opteron processors, it won't have to worry about running out of capacity to store these figures as 64-bit integers for another 308,152 years.<sup>1</sup>

## CUSTOMERS PREFER AN ORDERLY TRANSITION TO 64-BIT ENVIRONMENTS

AMD's 64-bit vision included the realization that not all programs benefit from a move to 64-bit architecture; some may even find such a move counterproductive. The key decision points are whether programs need to address large data structures in main memory or represent very large numbers in fixed-point formats. If a program's aggregate memory requirements exceed 2GB, it's a good candidate for conversion; if it operates without constraint in less than 2GB, a move to 64-bits won't help its performance very much, if at all. One of the key attributes of the AMD64 architecture is its ability to run a mix of 32- and 64-bit x86 codes without sacrificing performance in either mode of operation. AMD would like to claim credit for inventing the notion of extending a 32-bit architecture to 64-bits in a compatible manner, but others, including HP (with PA-RISC), IBM (with PowerPC), and Sun (with SPARC) all applied the same transformations to their 32-bit architectures long before AMD did. Intel even used this approach when it extended its 16-bit 286 to create the 32-bit x86 architecture used in its 386, 486, Pentium®, Pentium II, Pentium III and (until recently) Pentium 4

lines. AMD realized that customers would appreciate an evolutionary approach like this, and it was the first to apply it to the x86 architecture that powers most of the world's personal computers and servers. And AMD knew customers don't like to be forced into entirely new 64-bit architectures that lack 32-bit antecedents, like the DEC Alpha design of the 1990s that struggled to gain market acceptance until it was end-of-life'd by HP. When AMD first unveiled its concept in 1999, its competitor suggested that x86 customers had no interest in mixing 32-bit and 64-bit applications on the same machine. Eventually that same competitor came around to AMD's way of thinking and added 64-bit features to its own line of 32-bit x86 processors.

## Why Multi-Core Matters

### MULTI-CORE IS ALL ABOUT PROCESSOR PERFORMANCE

From the beginning of the microprocessor age in 1971, advances in processor performance have been closely correlated with advances in the clock frequencies at which those microprocessors operated. Recently that correlation has begun to weaken. Now most processor architects believe there are easier and more cost effective ways to increase system performance. Several factors account for the decreased utility of clock frequency as a source of enhanced performance:

1. The mismatch between processor cycle time and DRAM cycle time (sometimes known as the "memory gap") has grown so large that any benefit from increased clock frequency gets lost when cache misses force reloads from main memory. Pushing on clock frequency causes the processor to spin its wheels, but no forward motion results.
2. Increases in frequency force a chip to use more power, which in turn makes it harder (and more expensive) to cool. Given that this increased frequency results in little if any incremental performance, the futility of the approach becomes increasingly apparent.



<sup>1</sup> Some processors that claim to have 64-bit extensions actually rely on the chip's 32-bit logic to handle 64-bit operations. Their approach delivers the correct answers, but sacrifices performance for 64-bit operations. Authentic AMD64 processors use full 64-bit logic to deliver 64-bit results without impacting performance.

3. The recent shift from 130nm to 90nm process geometries doubled the transistor budgets available to chip designers. This means a dual-core 90nm chip requires about as much silicon as a comparable single-core 130nm design and can be manufactured in a cost-effective manner.
4. Dual-core chips possess the ability to increase performance of many applications by almost 100 percent, while other approaches offer improvements of 10 to 20 percent at best.
5. Many applications, especially those that run on servers, have already been modified to take advantage of dual-processor systems, and can immediately benefit from the presence of dual-core systems as well.

All-in-all, a dual-core approach turns out to be the most effective and most efficient way to enhance processor performance for chips produced on 90nm production lines. The same factors that make dual-core processors attractive at 90nm are likely to make quad-core processors attractive when the industry moves to 65nm process technologies.

### Why Do New Systems Use So Much Power?

Until just a few years ago, the equipment in a typical 19-inch 42U datacenter rack rarely consumed more than a few kilowatts. Today it's possible to fill that same rack with equipment that consumes 20KW and dissipates almost that amount of heat. These new configurations do a lot more, but they've become harder to power and harder to cool. How did the industry get itself into this situation, and how can it get itself out? Two factors account for most of the change. First, AMD has dramatically increased the number of processors (and associated gear) that can be packed into a single rack. A few years ago, users were ecstatic if they could fit 42 IU 2-way servers (with a total of 84 processors) in a single rack. Today, using a bladed server approach, users can cram up to 260 processors in that same space. Second, system power consumption has grown, as CPU frequencies and DRAM content per system

have increased. As Intel's Pat Gelsinger noted at an industry conference last year, "If we let the power density on the chip continue [to increase] unabated, it's a hot plate, it's a rocket nozzle, a nuclear reactor, and eventually the surface of the sun."<sup>2</sup>



### AMD POWERNOW!™ TECHNOLOGY WITH OPTIMIZED POWER MANAGEMENT (OPM) CAN REDUCE CPU POWER REQUIREMENTS BY UP TO 70 PERCENT

AMD is doing its part to rein in these server systems' voracious appetite for power. It started down this path several years ago, when AMD added silicon on insulator technology to its manufacturing processes. SOI improves the power efficiency of its processors by almost 20 percent compared to the non-SOI processes others use. Then AMD enhanced its processors with an extended version of AMD PowerNow!™ technology with OPM. This technology adjusts CPU frequency and operating voltage in response to instantaneous changes in the system's workload; under heavy loads it cranks up the frequency to full power (a maximum of 95 watts), but as the load drops off it lowers the frequency and reduces

<sup>2</sup> Pat Gelsinger, IDF Keynote, February, 19, 2004



power consumption by almost 70 percent, to a maximum of 30 watts at a 1GHz “idle” state. Since it’s really hard to keep these fast processors busy for long, your actual power consumption will fall between these two extremes. Granularity is the key to minimizing power consumption; just as jet fighters in formation use terrain-following radar to hug the ground, AMD PowerNow!™ technology with OPM continually adjusts the performance of each processor, where a mere high, medium, and low won’t do.

#### AMD POWERNOW!™ TECHNOLOGY WITH OPM CAN REDUCE OPERATING EXPENSES

AMD PowerNow! technology with OPM can reduce a server customer’s electricity bills. For more information, visit [www.amd.com/power\\_cooling](http://www.amd.com/power_cooling). Processors that consume less power also emit less heat, so AMD PowerNow! technology with OPM helps lower HVAC expenses. All-in-all, AMD PowerNow! technology with OPM is one more example of technology that can be deployed to save power in an economically productive manner.

#### AMD’s Direct Connect Architecture

Let’s face it. Companies that design and manufacture microprocessors love catchy names that capture the essence of the technology they embed in their chips. But what does “NetBurst” really tell you about a Pentium processor? Or “CoolStream” about one from VIA? Some customers even had trouble relating the

name “QuantiSpeed” to the features of AMD’s own AMD Athlon™ XP processors. When it came time to name the architecture that underpins the AMD64 line, AMD vowed that it would be more direct, even if it had to sacrifice marketing glitz in the process. And then it hit them. The essence of AMD64 processors is the manner in which they connect directly with the memory and I/O resources in the system, as well as with each other. Rather than give into the basic marketing urge to invent an ornate title, AMD chose the more plain spoken “Direct Connect Architecture,” a term that tells the AMD64 story in an unambiguous manner that customers can understand.

#### DIRECT CONNECT ARCHITECTURE REDUCES MEMORY LATENCY TO IMPROVE PERFORMANCE

Direct Connect Architecture addresses the two key inhibitors of system performance—memory latency and memory bandwidth. Memory latency measures how long it takes to move data from the system’s DRAM memory to the processor’s Level 2 cache. Cache misses often stall the processor’s pipelines for hundreds of CPU cycles and erode system performance, so chip designers do their utmost to minimize the time it takes to load cache lines from main memory. Part of the delay stems from physics; the basic DRAM cell is inherently slower than CPU logic functions. But much of the delay results from the way the CPU communicates with the DRAM memory devices, and is independent of the DRAM’s internal structure. For many years, memory controller functions resided on a so-called north bridge chip separate from the CPU. Each memory transaction was forced to travel over the processor’s front-side bus to the north bridge and then on to the actual DRAM chips; data then flowed back to the CPU cache via that same north bridge and front-side bus. Needless to say, the front-side bus often bottlenecked, resulting in idle CPU cycles and disappointing performance. Some vendors speed up their front-side bus to minimize these delays. AMD eliminates the bus and bottleneck entirely. Direct Connect Architecture combines the north bridge function with the CPU itself, so AMD64 processors can talk directly to their associated DRAM memories. No middleman. No front-side bus bottleneck.



## DIRECT CONNECT ARCHITECTURE ADDS MEMORY BANDWIDTH TO IMPROVE SCALABILITY

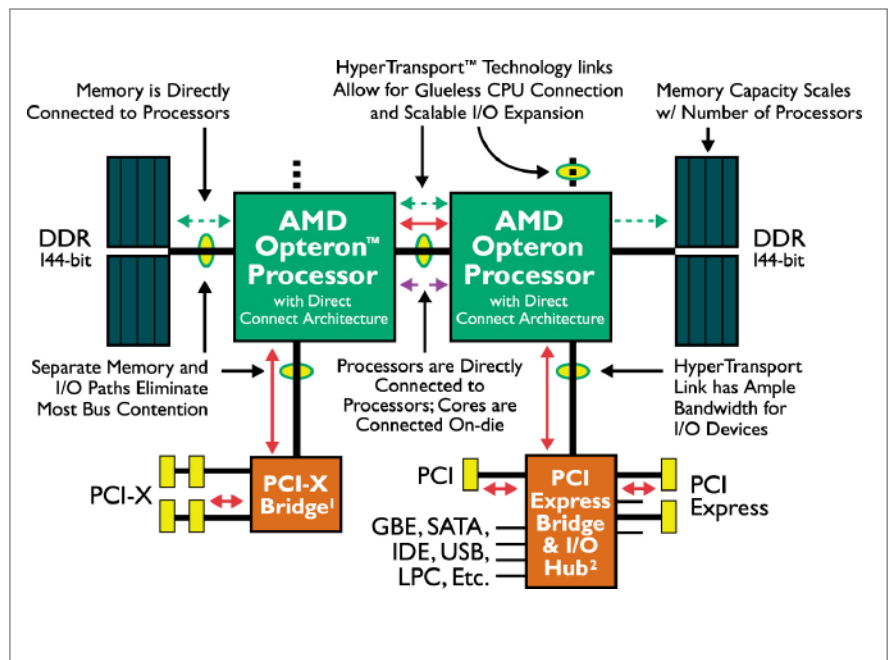
Direct Connect Architecture also gives the AMD Opteron™ processor family its scalable memory bandwidth in multiprocessor systems. Fast processors have a voracious appetite for data that resides in system memory; the more processors in the system, the greater the demand on memory bandwidth. The old-fashioned north bridge-based memory controllers used in some x86-based servers supply a fixed amount of memory bandwidth, regardless of the number of processors in the system. With Direct Connect Architecture, each processor includes its own memory controller; dual-processor systems have twice the memory bandwidth of a single-processor system, and four-way systems have four times the memory bandwidth. This means that memory bandwidth is not a significant constraint on the performance of AMD Opteron processor-based systems.

Although Direct Connect Architecture helps eliminate the classic front-side bus bottleneck, AMD64-based processors still need a way to exchange data with one another, and with external I/O subsystems. To address these requirements, AMD engineers invented HyperTransport™ technology, a high-bandwidth (10 GB/second) low-latency (under 10 ns) chip-to-chip interconnect mechanism. After inventing this technology, AMD worked to share it with the industry. AMD formed the HyperTransport Technology Consortium, a group that owns the specification for the version of HyperTransport technology used in I/O subsystems and licenses it freely to all comers<sup>3</sup>. A few licensees have developed I/O devices that interface directly with HyperTransport technology, but most have created bridges that tie HyperTransport technology links to conventional, lower performance buses like PCI, PCI-X and PCI-Express.

AMD's engineers also created an enhanced version of HyperTransport technology known as "Coherent HyperTransport technology" that it uses to coordinate the contents of on-chip caches in multiprocessor AMD64 configurations. AMD Opteron 200 and 800 Series processors include three HyperTransport technology links that can be deployed in I/O or coherent modes of

operation. The coherent links directly connect the memory controllers on each of the processors to one another, and add only 20ns to the time needed to access DRAM data from memory attached to a system at the other end of the link. Two-way configurations based on the AMD Opteron 200 Series maintain one coherent link, and can use the remaining four HyperTransport technology links for I/O, providing up to 40GB/second of I/O, more than enough for many IT infrastructure applications. Four- and eight-way configurations typically use all three links on AMD Opteron 800 Series processors for coherent traffic, but reserve a few links for I/O operations.

These larger configurations are best suited for executing computationally intensive applications and database management. These engineers optimized the AMD Opteron I00 Series for use in low-cost applications that need only one HyperTransport technology link, so they left out the other two to make these processors more affordable.



<sup>3</sup> The Consortium consists of over 40 industry-leading member companies, including founding member AMD.

## HYPERTRANSPORT TECHNOLOGY PLAYS A KEY ROLE IN DELIVERING DIRECT CONNECT ARCHITECTURE'S BENEFITS

HyperTransport™ technology gives AMD's OEMs unprecedented flexibility when it comes to implementing innovative servers. HP's BL45p four-way AMD Opteron™ processor-based blade server consists of two two-way blades connected by HyperTransport technology links. Appro's XtremeBlade system allows users to reconfigure two two-way blades into a four-way blade by simply connecting the HyperTransport technology channels on one blade to the HyperTransport technology links on the other. Cray, one of the more venerable names in the history of supercomputing, uses HyperTransport



technology inside its XDI Supercomputer. Its RapidArray Interconnect System ties the HyperTransport technology links coming from hundreds of AMD Opteron processors into a high-speed non-blocking 96GB/second switch fabric, giving its SDI supercomputer startling price/performance characteristics. Cray only markets its RapidArray technology as part of its SDI supercomputer, but PathScale developed a similar technology that works with many standard AMD Opteron processor-based server platforms. PathScale's InfiniPath HTX adapter

plugs into HyperTransport technology-based HTX slots on AMD Opteron processor-based servers, and uses a low-latency InfiniBand switch fabric to connect nodes in a high-performance cluster configuration. PathScale even sells the chips that power InfiniPath HTX to others who want to include this feature in their own custom-designed systems.

## Dual-Core Designed In, Not Added-On

Back when AMD's processor designers were thinking about the best way to add 64-bit capabilities to the x86 architecture, and add memory controllers to processors, they also examined the ways semiconductor process technology would impact future implementations. It's no secret that each process technology generation gives designers twice as many transistors to work with as the one before. These designers realized that if they could fit a Single-Core AMD Opteron processor on a reasonably sized piece of silicon manufactured on AMD's 130nm lines in Dresden, then they could put two cores on a similarly sized chip when Dresden moved to 90nm technology, and four cores when AMD moved to 65nm technology a few years later. They also realized that to achieve optimum performance, they needed to share selected elements of the single-core design between the cores of a multi-core chip, rather than mindlessly stamping two copies of that single-core design on a single piece of silicon. A lesser team might have deferred these dual-core design issues for another day, but AMD's team tackled them from the outset.

The secret sauce that gives the Dual-Core AMD Opteron processor its performance edge lies in a feature called the System Request Queue (SRQ). AMD's designers included it in their original single-core design in order to separate those parts of the processor that would be replicated from those that would be shared in a dual-core version. Some companies, even if they were clever enough to invent something like this, would have labeled the feature "top secret" and deleted it from all but NDA presentations. AMD shared its multiple-core SRQ with the rest of the world in a very public



presentation given in 2001,<sup>4</sup> at a time when other companies still thought 10GHz single-core processors were the technology of the future.

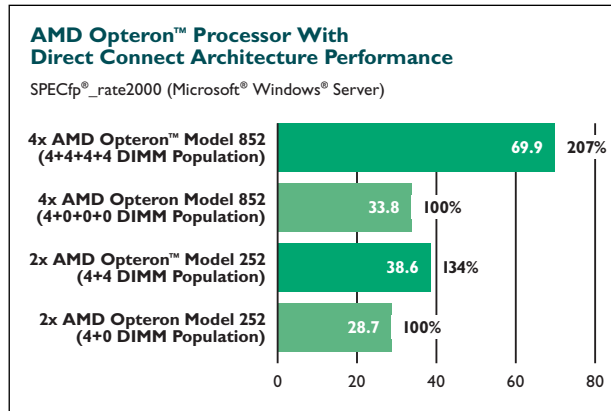
The SRQ passes requests for memory or I/O operations from either core to a high-bandwidth multi-port crossbar switch, a key system element that links the SRQ to the on-chip memory controller and the three HyperTransport™ technology links. This switch shuttles data from one HyperTransport link to another, or from a HyperTransport link to the memory controller without any intervention from either CPU core. When data intended for one of the on-chip cores arrives from the memory controller or a HyperTransport link, the switch passes it to the SRQ, which in turn hands it to the appropriate core.

The combination of the SRQ and crossbar switch provides the integration the two cores need for optimum performance and in some cases allows a dual-core 1-way system to outperform a 2-way single-core system. In either of these arrangements, one of the processor cores often needs to access data residing in the cache attached to the other core. Moving this data from one cache to the other in dual-processor configurations entails a trip over a 10GB/second HyperTransport technology link. In uniprocessor dual-core configurations, the data never leaves the chip; it just goes in one port of the SRQ and out the other. Although HyperTransport technology links are faster than the old front-side bus arrangement, the system operates even faster when it can avoid using them entirely for operations like this.

### Does Architecture Really Matter?

We could wax rhapsodic for many more pages about the elegance of Direct Connect Architecture. Doing so might earn us extra points were we to submit this paper as part of the coursework for a computer science degree, but unless these architectural features lead to superior performance or more robust and reliable operation, most IT users' eyes would glaze over. The real test of any architecture is how it performs. AMD posts the results of many industry-standard benchmarks on its Web site at

[www.amd.com/opteronperformance](http://www.amd.com/opteronperformance), but none of those tests focus directly on the impact of the architectural features discussed. To fill this void, results of some special tests are provided that demonstrate how these features contribute to AMD's overall performance.

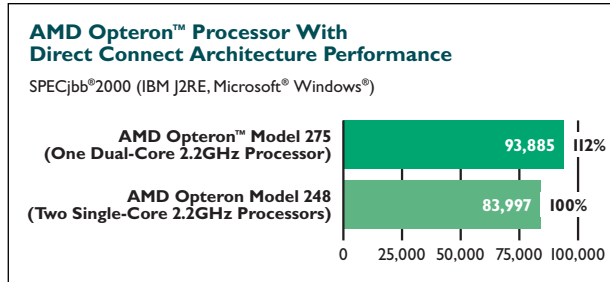


This first test illustrates the performance benefit gained from the on-board memory controller in dual- and quad-processor systems. For maximum performance, it's usually best to distribute physical memory across several processors in order to increase system memory bandwidth. For this test, AMD asked its benchmark team to attach all DRAM modules to a single processor, so that these dual- and quad-CPU configurations would use only one of their memory controllers and have "only" 6.4GB/second of memory bandwidth, regardless of the number of CPUs in the system. When the results of optimally configured systems (with 12.8 and 25.6GB/second of DRAM bandwidth) are compared to these artificially constrained arrangements, the second memory controller boosts overall SPECfp\_rate2000 results by 34 percent in the dual-CPU system, and the additional three memory controllers more than double the SPECfp\_rate2000 performance (an increase of 107 percent).<sup>5</sup> Of course, not all applications benefit equally from increased memory bandwidth, but the odds are good that this one architectural feature can improve your system's performance more than all the radical increases in CPU core frequency or front-side bus bandwidth, as applied to old-fashioned CPU architectures, ever will.

<sup>4</sup> AMD Reveals the "Hammer," Microprocessor Forum, October 2001, Slide 24

<sup>5</sup> In case you are wondering whether the improved performance shown in these tests results from increased memory capacity or increased memory bandwidth, you can rest assured that bandwidth is responsible for virtually all the gains. The SPECfp\_rate tests use a little more than 1GB of system memory; the rest remains idle throughout the test's execution.

And unlike those higher frequencies and faster buses, this approach adds performance with no increases in system power consumption.



The second test shows the impact the System Request Queue and Crossbar switch have on system performance when running SPECjbb®2000 (a Java Business Benchmark) on a single Dual-Core AMD Opteron™ processor and on two single-core processors operating at the same frequency. The dual-core system's ability to handle cache snooping on-chip gives it a 12 percent performance advantage over the dual-processor arrangement. The only other variable in this comparison—memory bandwidth—should have favored the dual-processor approach, since it had twice the available memory bandwidth. Nevertheless, the dual-core wins because of its streamlined cache snooping ability.<sup>6</sup>

<sup>6</sup> This test used processors running at 2.2GHz, but the relative performance advantage of dual-core over dual-processor systems is independent of CPU frequency.

## Summary

We hope this document has convinced you that multi-core technology is not a tide that lifts all processors equally. Architecture matters. AMD is confident its Dual-Core AMD Opteron processors include the most sophisticated architecture of any multi-core x86 processor you can buy today. They started with an outstanding single-core x86 design, and carefully decided which parts of that design to replicate in their dual-core products. AMD endowed those cores with scalable memory and I/O bandwidth to feed their ravenous appetite for data. As a result, Dual-Core AMD Opteron processors can deliver better overall performance than any x86 alternative, as we have demonstrated in benchmark after benchmark.

If you have found this document to be informative, we hope you will eagerly await the next one in our series.