AMD Aims To Take Over x86 Leadership

AMD – Not Just An Imitator

In most industries the market leader would refer to the other companies participating in that market as competitors – but not Intel. It is so arrogant and self-focused that it publicly proclaims the x86 universe consists of itself and several “imitators”. Advanced Micro Devices (AMD) is one competitor that has long ago decided imitation was the road to oblivion and it needed to take its destiny into its own hands. As a result two years ago AMD made two surprising announcements.

The first announcement was that its future K7 processor would use the mechanical form factor of Intel’s slot 1 connector and cartridge. Instead of using the P6 bus (which it agreed never to do as part of a wider legal settlement with Intel) AMD announced it had licensed the bus technology of the next generation Alpha high end RISC processor, the 21264 or EV6. This was a very bold gamble. Though Alpha technology is widely viewed as leading edge, chipsets developed for the EV6 would likely be far too expensive for the PC market. AMD was basically committing to building an entire infrastructure of chipsets and motherboard reference designs for the K7 with or without initial support from third part chipset vendors.

The second surprise was AMD’s 3DNow extensions to the x86 instruction set architecture (ISA). Up to that point the x86 instruction set was basically considered the exclusive property of Intel. Any move to change or add to x86 would be tantamount to proclaiming incompatibility with Intel devices. Fortunately, the 3DNow extension were welcomed by a small but important niche of x86 customers, game developers and game players. More importantly in the long run, Microsoft blessed the 3D now extensions with support for them in the DirectX graphics application programming interface (API). It would be about a year before Intel could answer the 3DNow extensions delivered to customers in AMD’s K6-2 MPU with its own streaming SIMD extensions or SSE, first incorporated in the “Katmai” Pentium III.

To Boldly Go Where No x86 Has Gone Before

A few weeks ago AMD presented a few details about its next generation K8 processor code-named “SledgeHammer”. The most surprising disclosure was that it would extend the x86 ISA to 64 bits (x86-64). Intel long ago decided against extending x86 to 64 bits and instead, in conjunction with Hewlett Packard, developed a completely new, RISC-like 64-bit ISA, called IA-64. Although IA-64 has an x86 compatibility mode it will very likely operate at a much lower performance level than native 64-bit code. Also, with IA-64 it is very difficult and inefficient to mix x86 and IA-64 code within a single application thread. It certainly seems that Intel intends to wean its customers off the x86 platform in favour of IA-64 starting first at the high end (servers and workstations). This contrasts sharply with the AMD’s evolutionary approach which can support 32 and 64 bit x86 code nearly

equally and apparently allow efficient mixing of code at the same level as can be done with 16 and 32 bit x86 code since the Intel 386.

If AMD succeeds in eliciting support for x86-64 from Microsoft then it will be quite obvious that it has taken over leadership of the x86 architecture. This is likely for several reasons. The biggest reason Microsoft would want to strongly support x86-64 is Compaq's recent public announcement that it would not support development of a 64 bit version of Windows as a product offering for Alpha. Microsoft has great ambitions for Windows challenging the traditional proprietary versions of Unix for the role of running the largest and most prestigious computer systems, the high end, multiple processor Enterprise class server. But these systems are almost exclusively based on 64 bit RISC processors.

With support for both IA-64 and Alpha, Microsoft seemed to have the two most powerful 64-bit platforms in its stable. The loss of Alpha makes 64 bit Windows an operating system exclusively targeted for IA-64. This leaves Microsoft's ambitions increasingly vulnerable to Intel, which has plans of its own. Intel is in the driver's seat for IA-64 because it can point customers to Monterey, Solaris, HP-UX, AIX and Linux as alternatives to Microsoft's 64 bit OS offering. As testimony has shown in the DOJ antitrust trial against Microsoft, there is very little love lost between these two ruthless and highly competitive giants.

This is where AMD's x86-64 comes in. I wouldn't be surprised at all if knowledge of AMD's plan to extend the x86 ISA to 64 bits was a major factor in Compaq's decision to pull Alpha out of the Windows arena and focus on TRU64 and Linux. In a two way competition for 64 bit Window platforms between IA-64 and Alpha, Compaq likely thought it could stay a ahead on native performance and match or beat IA-64's x86 compatibility mode with its FX!32 binary recompilation tool, while offering smaller and cheaper processors and chipsets. In a three-way match, Alpha would likely be squeezed out on price and x86 performance by AMD's x86-64 based processors.

So What Does a 64 bit x86 Look Like?

Calling a computer architecture '64 bits' means different things to different people. For a marketing person this is quite a flexible attribute and has in the past meant that a CPU has at least one data path or register in its implementation that is 64-bits wide. For computer designers, a 64 bit architecture is one that supports 64-bit general purpose registers, instructions for arithmetic and logical operations on 64 bit integers, and the ability to perform 64 bit logical addressing for pointer manipulation and array index calculation.

One way of predicting what AMD's x86-64 will look like is to examine how the x86 was extended from 16 bit to 32 bit starting with the Intel 386 and apply the same basic principles to reach 64 bits. The basic 16 and 32-bit programmer's model of the x86 along with one possible 64 bit implementation are shown in Figure 1. The floating point (FP) component of ISA will be dealt with later.

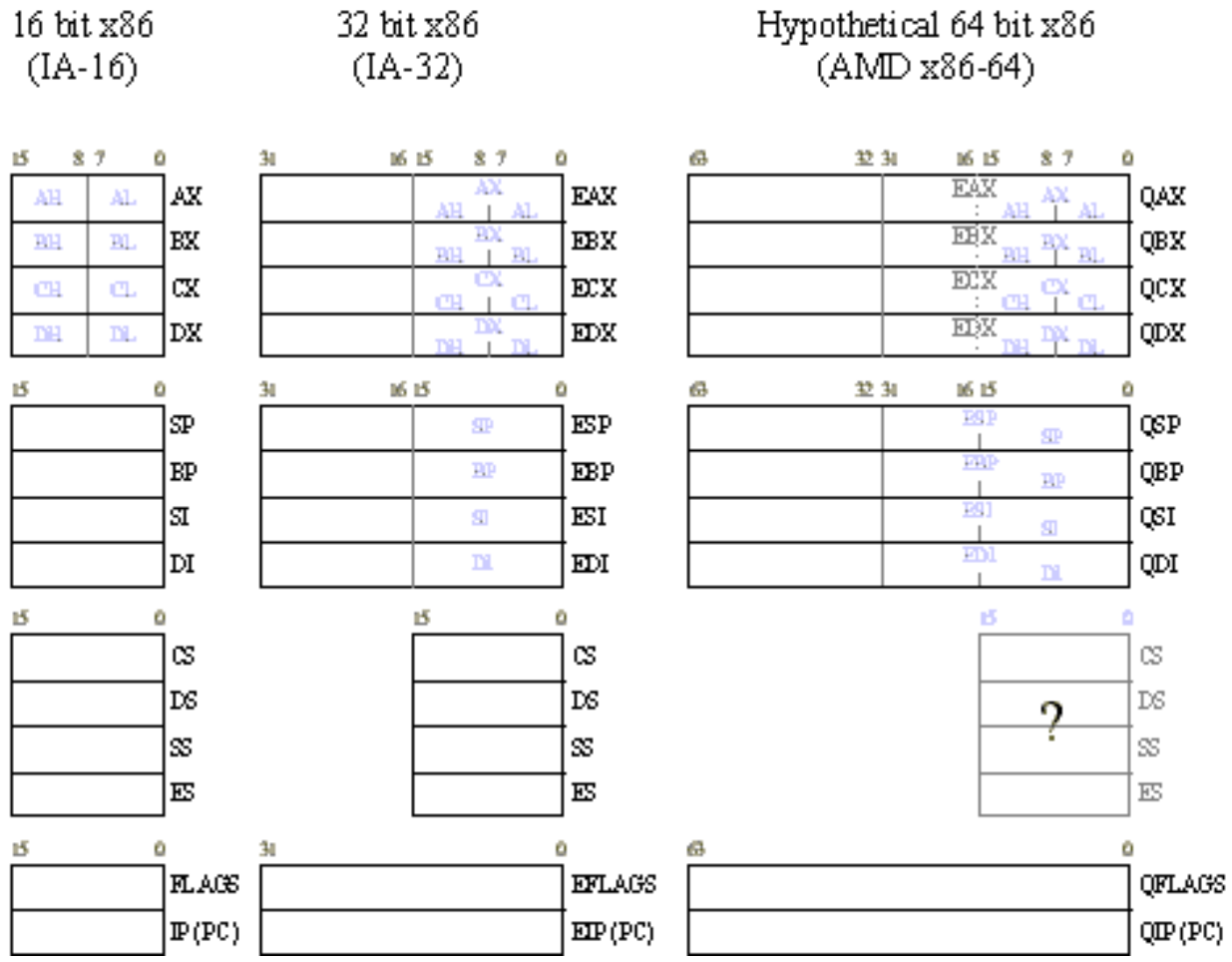

Figure 1

According to AMD, addressing in x86-64 will be 64-bit by default, with the option to override for 32-bit addressing. Also, x86 data manipulation instructions will default to 8 and 32-bit with the option to override for 64 and 16-bit. The x86-64 ISA specifies that the results of 32-bit operation will be sign extended to a full 64 bits before updating the target register. This is similar to what RISC designs such as 64-bit MIPS and Alpha do when executing 32 bit data manipulation instructions. Immediate and displacement fields in x86-64 instructions will be either 8 or 32-bits in length to limit code bloat, however, there will be a new instruction to load a 64-bit immediate data value into a register. AMD has said x86-64 will have “no segment base or limit registers”. This seems to indicate that the segment registers CS, DS, SS, and ES will not be used in 64 bit mode since there are no segment descriptors to select.

What’s the Floating Point?

Probably the worst feature in the x86 architecture is the floating point architecture with its eight 80-bit registers organized as a stack. Code sequences compiled to evaluate an arithmetic expression often require extra instructions to get at operands because of the stack-based computational model. The

x86 stack model also needs extra instructions to preserve intermediate values when data flow analysis indicated these values can be reused later in the program.
AMD has decided to take this problem head on and change the programming model to a flat floating point register file for x86-64. Since this requires the addition of new instructions to use this new flat register file they also decided to utilize RISC-style three address FP instructions. That is, a FP instruction that takes its operands from two registers and writes back the result to a third register. The main advantage of three address FP instructions is that it saves extra instructions that would otherwise be needed to shuffle values around the register file or make copies when the compiler wants to reuse a value or place a final result back into the register reserved for argument passing or a function return value. AMD has not stated how many FP registers x86-64 would have but it seems likely to be 16 or 32 rather than 8. Having a large number of FP registers allows a clever compiler to unroll and software pipeline the computationally intensive loops typically found in scientific and engineering type applications.
The new instructions and flat FP register file seem to only support double precision (64-bit IEEE) operations. At first this may seem surprising but technical code generally eschews single precision format data to avoid numeric problems from loss of accuracy. For 3D graphics and other FP applications that only need 32-bit single precision data the 3Dnow extensions will apparently suffice.

Who is AMD Trying to Sledgehammer?

In his presentation at Microprocessor Forum, Fred Weber (VP of engineering, in AMD's Computational Products Group) made it quite clear that AMD intends to take on existing high end RISC processors and Intel's new IA-64 product line with x86-64. They have addressed the two biggest disadvantages the current x86 ISA has – lack of 64-bit addressing, and disadvantageous floating point performance architecture. Through conveniently selective comparisons he tried to show that AMD's K7 family had caught up to the fastest RISCs in integer performance. Mr. Weber claimed that the new architecture will allow future AMD processors to close the gap in floating performance.
Apparently the extension to 64 bits will come without too much pain either. AMD claims that stretching 32 bit x86 to x86-64 only adds 5% to processor die size, and since the same processor core executes 32 and 64-bit code the performance will be equivalent. This might be a bit of a stretch since research conducted at DEC indicates that in practice compiling an application with 64-bit addressing instead of 32-bit addressing causes a performance loss of about a 4 – 5% on average. This results from address data taking up more room in the data cache and main memory causing the cache and TLB miss rates to increase.

Conclusion: You Can Teach an Old Dog New Tricks

There is no doubt about it, AMD is taking on a risky and ambitious course of action extending the ancient x86 architecture once more, this time to 64 bits. Even if they can deliver the new x86-64 instruction set in future products without impact to schedule or performance their destiny is in someone else's hands, namely Bill Gates. If Microsoft refuses to support x86-64 with tools, applications, and operating systems then x86-64 will simply fail.

In today's computing world, however, a 64 bit version of x86 might just be the magic bullet that Microsoft needs to offset Intel's growing influence and rebalance the power relationship in the Wintel oligarchy. It potentially provides another platform to support a 64-bit version of Windows. It also represents an insurance policy in case Intel has bet on the wrong technological horse in the way it designed IA-64.

From AMD's point of view they had no choice but to develop x86-64 if they wanted to play in the high-end server and workstation marketplace. Any attempt to clone IA-64 would land AMD in the middle of a legal minefield stretching as far as the eye can see in all directions. The only other alternative would be to develop an analog to the IA-64. That is, a split personality MPU that supports both the 32-bit x86 ISA and a 64-bit RISC ISA. It is likely that Compaq would have licensed the Alpha ISA to AMD for use in a hypothetical bilingual processor if asked. But this approach is both more complicated than simply building a 64 bit x86 extension and would likely run afoul of Intel patents related to dual instruction set capability in IA-64.

Based upon the successful development and roll out of the K7 Athlon processor, it is obvious that AMD is in the top tier of processor designers (although the damage from the recent departures of key personnel could spell trouble down the road). Depending on the details, the new ISA should allow AMD to get closer to the FP performance levels of competing IA-64 and RISC processors. They are unlikely to close the gap completely for either integer or FP performance as they claim – the x86 baggage is just too heavy. However, as history has shown us, x86 doesn't have to be as fast as RISC processors in order to retain the PC desktop market. It just has to stay within a factor of two or so and the software legacy does the rest.

With its x86-64 extension solving the 64 bit addressing issue, AMD should be able to ride the x86 market for at least another 4 or 5 years. Since Intel is unlikely to respond with a 64-bit x86 design of its own (and risk sabotaging market acceptance of IA-64) it seems that AMD is destined to take over the x86 leadership role.