Advantages of the Intel® Extended Memory 64-bit Technology (Intel® EM64T®) in Workstation Applications

The introduction of 64-Bit Extension Technology into personal workstations gives users expanded problem-solving capabilities while maintaining compatibility with existing IA-32 applications.

Executive Summary 2

Introduction 2
- Applicability to Workstation Market Segments 3

Why 64-bits? 4
- Processor Performance and Problem Size 4
- Workstation and Application Trends 5

EM64T Technology 7
- Overview 7
- Operating Modes 7

Advantages of EM64T 8
- Compatibility 8
- Flexibility 8
- Multiprocessing Performance 8
- Application Performance 9

Practicalities 9
- Expected Performance 9
- ISV Acceptance 10
- Operating Environments 10
  - Operating Systems 10
  - Development Environment 10

Conclusions 10

For more information 12

invent

Executive Summary

The ever-expanding data and performance needs of business, academic, engineering, and scientific organizations are pushing the limits of existing workstation technology. Competitive demands require higher quality CAD models, increased realism of digital content and the demands of HDTV, the large-scale analyses of the oil & gas industry, and the exploding amount of data processed by financial traders are just a few examples.

To address these needs, Intel introduced an extension to their existing IA-32 processor technology, called EM64T, or Extended Memory 64-bit Technology. EM64T not only provides increased performance and problem-solving power, it also provides a smooth transition for today's 32-bit UNIX or Windows applications.

Employing EM64T will allow users to generate much larger models and to manage a much greater amount of data in memory, instead of working the data back and forth to disk. Personal workstations will efficiently utilize tens of gigabytes of physical memory instead of the 4 GB limitation that is in place today as a matter of routine.

EM64T will be integral to the next generation of workstations based on Intel processors and chipsets. Customers in all market segments will see increased performance across the board—because of architectural improvements and increased capabilities, as well as the opportunity to employ 64-bit applications and operating systems. And because the systems are completely backward compatible, today's users of 32-bit applications will not lose any performance or functionality by acquiring these platforms.

In the second half of 2004, Microsoft will introduce Windows XP 64-Bit Edition of the Windows operating system. The 64-bit Edition of Windows provides a platform on which both new 64-bit applications and current 32-bit applications may be co-resident. This provides users with tremendous flexibility in running both classes of applications; they can take advantage of today's 32-bit Windows-based applications while preparing their IT environments for next-generation 64-bit applications as they become available. In addition, many ISVs have already begun the conversion of their major applications to leverage 64-bit technologies. As these applications come on-line, users will be able to step up to new levels of problem-solving power. Also, 64-bit Linux is available now for the extended 64-bit platforms.

IT departments that are considering any kind of personal workstation upgrade should consider processors that are EM64T-capable. There are numerous upside benefits, with very little extra investments required.

Introduction

With the introduction of Extended Memory 64-bit Technology (Intel® EM64T®), workstation users have the potential of solving much larger problems than are possible today on 32-bit platforms. This is because the 64-bit capabilities of EM64T allow processors to access larger amounts of physical memory, and applications to address larger amounts of virtual memory. As we will discuss later, workstation users are now routinely defining problems that require more memory than is accessible with the 32-bit systems of today, thus the requirement for an expanded address space.

The key words here are "more memory"—as shown in Figure 1, this means <u>much</u> more memory. Today's processors use a four-byte (32-bit) word for addressing, making them capable of addressing 2^{32} bytes of memory, or 4 gigabytes[1]. By using an eight-byte word (64 bits), processors theoretically can address 2^{64} bytes (18 exabytes!). (An exabyte is one *billion*

[1] Note that this includes everything—operating system, drivers, applications, etc., thus a user is actually limited to much less than 4 GB.

gigabytes. Someone once estimated that five exabytes is roughly equivalent to the number of words ever spoken by all human beings[2]).

Figure 1. Address space as it relates to number of address bits.

Bits	Notation	Address Space (bytes)	Name	Column1
8	2^8	256	256	bytes
16	2^16	65,536	65	kilobytes (KB)
32	2^32	4,294,967,296	4	gigabytes (GB)
64	2^64	18,446,744,073,709,600,000	18	exabytes (EB)

Servers and some workstations (e.g., the HP's PA-RISC workstations or the Intel® Itanium 2-based HP Workstation zx2000 and HP Workstation zx6000) have permitted 64-bit addressing for several years. Complementing Itanium2-based products, EM64T brings a 64-bit capability to all mainstream workstation users and applications that use processors based on the Intel IA-32 instruction set (e.g., Pentium 4® and Xeon® processors).

To implement a 64-bit address space, applications (and operating systems) must be redeveloped (or at a minimum, recompiled—more on that later). The most important feature of EM64T is that it allows for execution of existing IA-32 32-bit binaries—at native performance levels—without changes. This allows the coexistence of 32- and 64-bit applications on the same system without penalizing the performance of either one. It also facilitates a smooth transition period, where applications may be migrated to the 64-bit environment at a practical pace.

Applicability to Workstation Market Segments

Specific areas where users in the following market segments will benefit from a 64-bit capability include:

- Mechanical Computer-Aided Design (MCAD)—MCAD users in areas such as automotive and aerospace are increasingly facing the need to conceptualize larger and larger designs. Today, the workaround is often to break large models up into subassemblies, perform modeling and simulation operations, and then perform the tedious task of combining the results. By employing 64-bit applications, entire assemblies may be modeled on a single workstation.
- Digital Content Creation (DCC)—Animation and rendering, video editing, and game development are three major areas of the Digital Content Creation (DCC) segment that benefit from a 64-bit architecture. The ability to manipulate fully rendered 3D models, instead of a wireframe representation, or to render multiple video layers simultaneously are examples. The ability to store larger files, such as those generated in HDTV, in memory will minimize the need for render time, even with complex effects and animations
- Financial Trading and Analysis—The explosion in data that is manipulated by financial traders and analysts requires larger amounts of memory for database caches and computationally intensive in-memory modeling. In addition, these users tend to engage many applications simultaneously; keeping these applications in memory (instead of transferring them back and forth to disk) will result in tremendous performance increases.
- Oil and Gas—Both upstream and downstream modeling applications in the Oil and Gas industries are requiring higher degrees of accuracy and realism. For example, reservoir modeling employs huge arrays of 64-bit data; increasing reservoir size and model accuracy quickly causes the application to outgrow the addressability of a 32-bit operating environment.

[2] Roy Williams, "Data Powers of Ten", Caltech. See http://www.sims.berkeley.edu/research/projects/how-much-info-2003/

- Geographic information systems (GIS)—Tasks that benefit from a 64-bit environment in GIS applications include image processing and database management. Both of these applications can employ 64-bit addressing to hold more (and more accurate) data in memory, reducing application run time and response times.
- Electronic Design Automation (EDA)—Many EDA applications require 64-bit addressing and are already available on proprietary UNIX platforms (e.g., HP-UX and Solaris). The introduction of 64-bit capabilities on cost-effective Intel-based platforms brings users of these applications a very attractive alternative; not only do they get the benefits of a 64-bit architecture, but are able to use applications that are now supported on the Windows operating environment.

This paper will describe EM64T, how it is implemented in the current (and upcoming) families of HP workstations, and a practical look at why users would use this technology.

Why 64-bits?

The obvious question becomes, "why do I need to use all of this memory?" There are several ways of answering this (below), but the quick answer is "because we can!"

Processor Performance and Problem Size

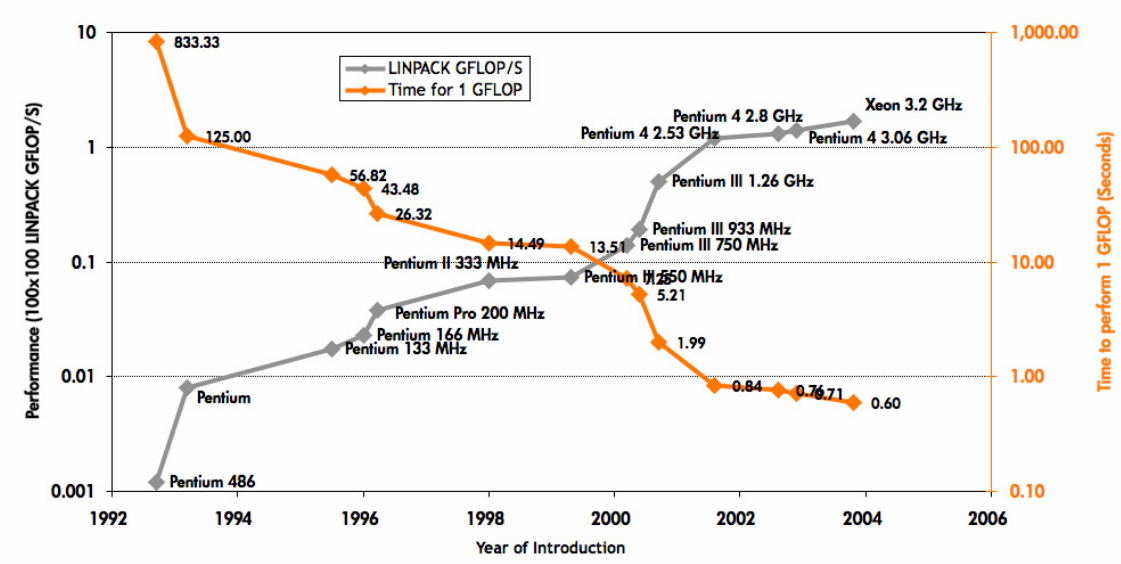
Let's explain this answer in terms of the kinds of problems being solved by personal workstations. Typical problems being solved today by workstations tend to be those that model real-world events (for example, computer-aided design and engineering applications simulate stresses on components for manufacturers, digital content creation simulate images—real or imagined—for film studios).

One dominant characteristic of most simulation applications is that the more processing power available to them, the better the results, or more precisely, the better the results model what happens in the "real world."

And, as shown in Figure 2, these applications are achieving more and more performance. This figure shows performance (primary Y axis) on an industry-standard benchmark, the 100x100 LINPACK[3] on an industry-standard architecture (IA-32). Note especially the logarithmic scale on the Y axis (which shows billions of floating operations per second, or GFLOP/S); performance has increased several thousand times in a decade!

[3] We chose the 100x100 LINPACK because of its long history and applicability to scientific applications.

Figure 2. Processor performance improvements over time as indicated by the 100x100 LINPACK benchmark.



However, our discussion is not to point out the increase in performance, but rather the size of the problems that are now solvable *given this performance*. A quick recalculation of the data in Figure 2 gives us the secondary Y-axis—the amount of time (in seconds) required to perform one billion floating-point operations (one GFLOP). As can be seen, in 1993, it took about 13 minutes, in 2004, it takes a little over 1/2 of a second.

One conclusion that may be drawn is that the amount of data that can be processed (the problem size) in a "reasonable" amount of time is now much larger—in fact, (probably much) larger than a 32-bit address space is capable of addressing. Thus, the need for a larger address space is born out of the tremendous increases in processor performance that have taken place over the last decade.

A perfect example of problem size and processor performance is that found in Digital Content Creation (DCC) applications. Companies that create entertainment-based digital content are facing increasingly larger datasets, the result of the more complex and realistic graphics that audiences expect to see in movies, television programs, and games. For example, applications may be able to render multiple layers simultaneously, instead of a layer-at-a-time approach used to conserve memory. It also allows them to work with fully rendered 3-D models, rather than a wire frame representation of the model. All of this frees up designers and artists to spend more time creating and less time computing.

Workstation and Application Trends

In the 1990s, the engineering workstation market was dominated by proprietary UNIX-based systems, product cycles were slow, and the total cost of ownership for Windows and UNIX workstations was high. Nevertheless, the power and potential represented by these workstations demanded their adoption for engineering.

During the 1990s, intense competition and commoditization of components and software drove costs down in the 32-bit Windows Workstation space. By 1996, sales of UNIX workstations (both 32 and 64-bit) were outpaced by the 32-bit Microsoft Windows-based challengers[4]. The

[4] IDC and Wall Street Journal, March 6, 1998

Windows workstation became an efficient design platform and was increasingly being used to run more computationally challenging simulation and analysis tasks.

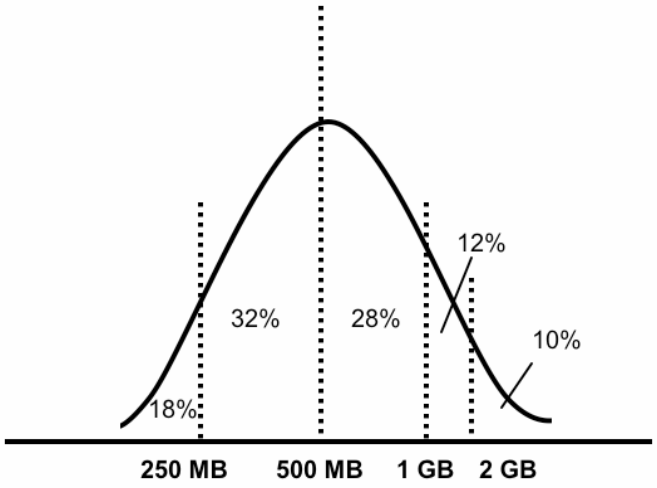
During the same time, falling memory prices, higher density memory modules, and a growth in popularity of 64-bit shared memory architecture computers changed this situation. For example, in 1996 the price of a 32 MB DRAM module was approximately $500, or $15 per Megabyte. A 256 MB DDR-RAM module can now be bought for $80, or $3.20 per Megabyte. Engineers could now afford to load their 32-bit and 64-bit workstations with larger memory configurations more cheaply, enabling them to work on larger and more complex models.

The 32-bit address space constraints are complicated further by the inherent demands of the Operating System. With Microsoft Windows, assuming a fully loaded 4 GB system memory, only between 3 GB and 3.75 GB maximum is available, because of the background processes that require the memory.

The applications have also changed. Application functionality and complexity, as well as the hidden complexity that underlies usability improvements, has also dramatically increased in the last 20 years—although not at the same rate as hardware. Software designers have taken advantage of the frequency of hardware advances, enabling new functionality like analysis and simulation components, improved interoperation with other software and CAM tools, and more realistic display options; all of which require more memory and are computationally expensive.

Many designs are hitting the limits of what is achievable on a 32-bit workstation. Figure 3 illustrates this point with a snapshot of a typical range of saved file sizes for Pro/ENGINEER CAD models in 2003[5]. Note that this file size sample only reflects the state of final files saved to disk, in reality the amount of memory consumed by this data may expand many times depending upon the operations the engineer is performing.

Figure 3. Pro/ENGINEER sample data file sizes based on 2003 sample.



This increase in complexity has another side effect that is influenced by the 32-bit limitations. Currently, many CAD engineers work with parts or subsections of assemblies, as opposed to an entire assembly, because of limitations imposed by their 32-bit workstations. They may have attempted working with larger models, only to find that performance becomes unacceptable because of memory limits that create slowdowns, or vulnerability to memory leaks that cause crashes or data corruption.

[5] The sample size was 600 engineers.

For example, in a 32-bit workstation environment, users often use workarounds like modeling multiple sub-assemblies and stitching individual 2D draft sketches together. This makes it difficult or impossible to complete a full 3D model, or to do easy interference checking. Instead, users spend time reducing model size by removing components and trying to split the problem across multiple workstations.

Obviously, there are multiple advantages to working with the full model; it allows for a holistic approach that can speed up certain tasks, as well as promoting greater confidence in the model for prototyping and analysis.

The convergence of these and other factors sets the stage for a new shift in engineering practices. Large memory configurations are inexpensive, hardware and software allows for more and more complexity in models, and commoditization has driven down overall workstation costs. The creation and mainstreaming of industry-standard 64-bit architectures in these applications is inevitable.

EM64T Technology

Overview

The Intel 64-bit extension technology is an enhancement to the existing Intel IA-32 architecture. An IA-32 processor equipped with 64-bit extension technology is compatible with existing IA-32 software, enables future software to access larger memory address space, and allows co-existence of software written for 32-bit linear address space with software capable of accessing 64-bit linear address space.

A processor with EM64T fully supports existing IA-32 features. In addition, it introduces a new operating mode, referred to as IA-32e mode. The IA-32e mode includes two sub-modes. Compatibility Mode enables a 64-bit operating system to run most existing legacy 32-bit software unmodified. The second sub-mode, 64-bit Mode, enables a 64-bit operating system to run applications written specifically to access a 64-bit address space.

Operating Modes

The result is three operating modes:

Legacy Mode—support for 32-bit O/S and 32-bit applications. No application software changes are required, however the user gets no benefit from the 64-bit extension technology.

Compatibility Mode—support for 64-bit O/S and 32-bit applications. In this mode, the O/S will see the 64-bit extensions, but the 32-bit application will not. This usage also requires all 64-bit device drivers.

Existing 32-bit applications do not need to be recompiled, and may or may not benefit from the 64-bit extensions. Applications will likely need to be recertified by the vendor to run on the new 64-bit extended O/S.

Like 64-bit mode, compatibility mode is enabled by the operating system on an individual code segment basis. This means that 64-bit applications may be running on the processor (in 64-bit mode) at the same time as legacy 32-bit applications (not recompiled for 64-bits) are running in compatibility mode. Compatibility mode also permits legacy 16-bit and 32-bit applications to run, without recompilation, under a 64-bit operating system (with some qualifications).

64-bit Mode—support for 64-bit O/S and 64-bit applications. It also requires applications to be modified for 64-bit operation and then recompiled and validated. This usage also requires 64-bit device drivers.

Advantages of EM64T

As we have seen, a 64-bit processor has multiple advantages, especially in problem size, scalability and performance. Other specific advantages of the Intel EM64T include:

Compatibility

Clearly, the major advantage of EM64T is that it runs existing IA-32 applications untouched, and at native performance. This means that systems employing processors that are EM64T-enabled may be seamlessly introduced into an IT environment, and as 64-bit applications (and operating environments) come online, users may migrate to them on their existing systems.

Processors enabled with EM64T are also compatible with software written for AMD Opteron™ and Athlon™ processors. Even though the hardware microarchitecture for each company's processor is different, the operating system and software ported to one processor will likely run on the other processor due to the nearly identical instruction set architectures. Note, however, that there are additional features supported on the Intel architecture, like Hyper-Threading and the latest streaming SIMD extensions (SSE3) instruction set.

Flexibility

The compatibility with IA-32 and ability to run familiar operating systems (like Microsoft Windows) adds flexibility and choice to a workstation IT environment. An IT department may choose to deploy EM64T workstations directly into 32-bit environments, simply taking advantage of new levels of price/performance and new performance technologies. Alternatively, existing 32-bit applications may be mix and matched with new 64-bit applications using a 64-bit operating system (like Linux or 64-bit Windows). This allows users to leverage their familiarity with existing operating environments while taking advantage of increased scalability and performance of the 64-bit technology.

An additional level of flexibility comes from complementing the existing 64-bit Itanium2 processors and products. Source-level development work already done to redeploy applications on Itanium can be leveraged in moving these applications to EM64T-enabled workstations.

This flexibility is of particular interest to UNIX workstation customers who are considering a migration to a Microsoft Windows platform, but require 64-bit computing. Today UNIX customers need to use Microsoft productivity applications via two methods:

- On a UNIX workstation via Citrix Windows application server
- Two workstations on the desktop – a UNIX Workstation for 64-bit work and a Windows system for productivity

EM64T eliminates these necessities simply and easy. One workstation can do it all.

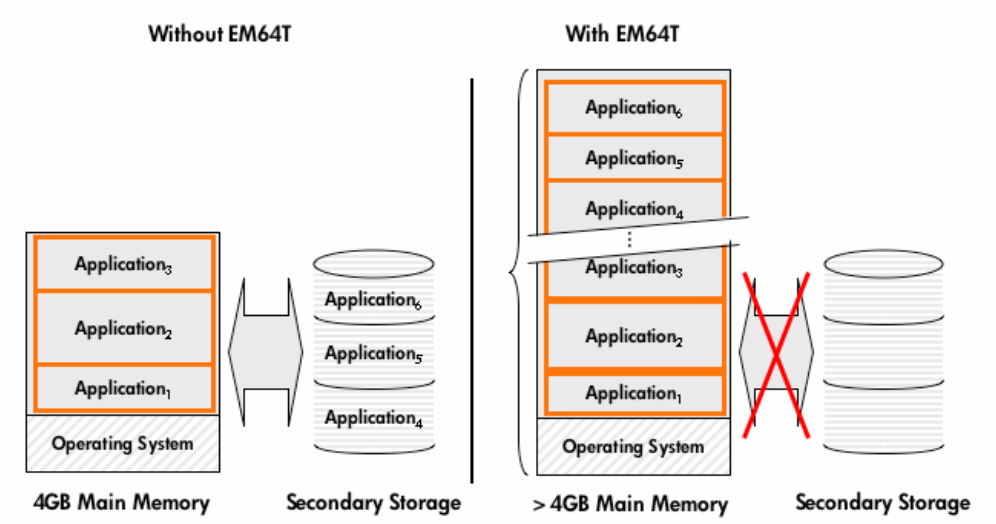
Multiprocessing Performance

Another very important aspect of being able to address more than 4 GB of memory is that of multiprocessing. Many workstations are employed in environments where a relatively large number (many tens) of non-OS applications are executing simultaneously (e.g., on financial trading floors and in electronic design automation).

With a 32-bit (4 GB) addressability limit in a workstation, once this limit is reached, the operating system must page (or swap) chunks of applications to disk and back (Figure 4). This is especially severe when employing multiprocessor workstations, since multiple processors are manipulating multiple processes in memory at the same time.

The additional physical memory and address space of a 64-bit processor enables multi-processing with sufficient memory per process to maximize the benefits of multiprocessing.

Figure 4. Extended addressing reduces secondary storage accesses when multiprocessing.



With the 64-bit extensions (and running a 64-bit enabled operating system), the processor is able to address more than 4 GB of memory, and more processes will fit into physical memory. This mitigates time-consuming paging, with possibly dramatic increases in overall performance for the end user. Importantly, these applications may quite likely be unmodified IA-32 applications straight off of a 32-bit workstation.

Application Performance

Processors enabled with EM64T provide additional performance extensions beyond 64-bit addressability. Applications that are recompiled with EM64T-aware compilers (see below) will take advantage of:

- 8 new general-purpose registers (GPRs)
- 8 new registers for streaming SIMD extensions
- 64-bit-wide general-purpose registers and instruction pointers

32-bit applications that are simply recompiled to take advantage of these new registers and extensions will see performance increases—even without any source-level conversion to 64-bit.

Practicalities

Now that we've reviewed all of the positive effects of migrating to a 64-bit environment, and the numerous architectural extensions to EM64T-enabled processors, we must look at the practicalities involved in implementing this new technology.

Expected Performance

It is important to note that simply redeploying an application with 64-bit technology does not in itself result in increased performance. The extended address space provides more of the application to be resident in memory, and the new architecture extensions will provide opportunities for increased performance, but 64-bit addressability does not provide any inherent performance improvement.

Having stated all of this, EM64T is supported on a heavily redesigned processor core, a Xeon version of Intel's Prescott technology. Major changes in the processor itself include larger on-chip caches, improvements in FSB (Front-Side Bus) frequency and memory performance, and a more

efficient branch prediction mechanism. Many applications could see performance increases—even without recompiling, and at the same clock speed—because of these improvements in the implementation of the IA-32 architecture.

ISV Acceptance

Like most major advancements in processor technology, the changes take some time to implement, especially with complex applications that require extensive testing and qualification. While the impacts of migration are greatly reduced by the extensive compatibility with existing 32-bit technologies, ISVs will still need to redeploy their applications to take advantage of the new architectural features.

Compatibility with AMD's AMD64 technology greatly helps with application migration. Many ISV applications, especially those where 64-bit addressability is a critical requirement, have already been ported to run under Linux on the 64-bit AMD Opteron™ processor; these applications will also run on EM64T-enabled workstations running Linux.

In addition, many vendors are now porting their applications to Microsoft's Windows® XP 64-Bit Edition, which will be released in the second half of 2004 (see "Operating Environments" below).

Operating Environments

Operating Systems

All major workstation O/S vendors have announced product support for the 64-bit extension technology. These include:

Microsoft—Microsoft has announced their planned availability of 64-bit extended operating systems and compilers available in the second half of 2004. These include Windows XP 64-Bit Edition for 64-Bit Extended Systems and Windows Server 2003 for 64-Bit Extended Systems. Please see http://www.microsoft.com/presspass/press/2004/feb04/02-17ExtendedTechnologyPR.asp for additional information.

Red Hat Linux—Future versions of Red Hat Enterprise Linux will support 64-bit extension technology, including Red Hat Enterprise Linux 3 update 2, which is expected in the second quarter of 2004. Please see http://redhat.com for additional information.

SuSE Linux—SuSE has announced production support for EM64T on SUSE LINUX 9.1 Professional. Please see http://www.suse.com/us/private/products/suse_linux/preview/ for additional information.

Development Environment

Because of the compatibility between the two architectures, vendors that are providing products for AMD64™ will likely similarly support the Intel EM64T. Intel has publicly announced compilers and software optimization tools for EM64T. This includes Rev 8.1 Fortran and C++ compiler and the Version 7.2 Intel® Vtune™ Analyzer, with production availability "targeted for Q3, 2004".

Other compiler and development environment vendors, such as The Portland Group, the GNU Project (The Free Software Foundation), Etnus and Absoft are working on tools that will be released on or near the production availability of the EM64T-enabled processors.

Conclusions

The Extended Memory 64-bit Technology offers users the potential to increase return on IT investment by consolidating applications, increasing efficiency and increasing the size of problems that can be solved.

EM64T enables larger, more accurate, more detailed CAD, financial modeling, and DCC. It will improve workflow processes by reducing the need to break large designs into smaller

subassemblies or to process a layer at a time in video editing in order to fit into available memory.

It allows a smooth and easy transition from today's 32-bit environment to a 64-bit one.

Further, additional features that are incorporated into processors with EM64T provide the next level of performance for personal workstation users.

Migration to this technology is eased by compatibility features including binary compatibility with existing IA-32 and 64-bit AMD64™ applications (when running a 64-bit operating system). Further, it permits a high degree of flexibility in the deployment of 32- and 64-bit applications.

Many ISVs have already begun the conversion of their major applications to leverage 64-bit technologies. As these applications come on-line, users will be able to break the 4 GB memory addressing limitation, and step up to new levels of problem-solving power.

The benefits of x86-64 technology will reduce the cost of 64-bit technology, making it the required feature in workstation class computing.

For more information

http://www.hp.com/workstations/index.html
HP Partitioning Continuum, HP, 2002

http://www.sims.berkeley.edu/research/projects/how-much-info-2003/
"How Much Information? 2003" – University of California at Berkeley. And Roy Williams' "Data Powers of Ten" web page at Caltech

http://www.hp.com/workstations/response.html
HP Workstations response to the Intel® 64-bit extension announcement.

http://www.intel.com/technology/64bitextensions/faq.htm
Intel frequently asked questions about E64MT.

http://www.intel.com/software/products
Intel® Software Development Products main web site.

© 2003 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Itanium is a trademark or registered trademark of Intel Corporation in the U.S. and other countries and is used under license.

5982-6724EN, 06/2004

invent