Itanium-Based Solutions in the Digital Content Creation Market

An IDC White Paper

Analysts: Kara M. Yokley, Earl Joseph II, Ph.D., and Christopher G. Willard, Ph.D.

Digital Content Creation Market

According to IDC research, the digital media market has grown 139% in the past three years. Its presence has infiltrated every area of digital output from movies to videogames to corporate presentations. Much of the NT workstation growth in this segment can be attributed to advances in 3D graphics, which have moved at three times the pace of Moore's law. In 1999, branded personal workstation[1] shipments into the digital media space accounted for more than $600 million in revenue, and the market is still growing. The sector is expected to have a revenue compound annual growth rate (CAGR) of 20% from 2000 to 2004. However, animators, game developers, and video producers alike have been pushing the limits of today's platforms to the breaking point. The largest challenges currently facing digital media creators using today's architectures are:

- Insufficient real addressable memory
- Insufficient memory throughput
- Inadequate floating-point performance
- Insufficient overall system performance

Today, digital content creators on the cutting edge are looking for increased headroom, bandwidth, and floating-point performance that can bring their art to a new level.

The digital media market is characterized by five subsegments: 2D/3D animation and rendering, video editing and compositing, streaming media, game development, and multimedia/Web authoring. Of these subsegments, applications for the first four will benefit most from the increased real addressable memory, floating-point performance, and

1 IDC defines branded personal workstations as NT-based systems, using x86 or Alpha processors, which are marketed as workstations, have minimum graphics requirements, and offer a higher level of application certification and customer service than a desktop does.

5 Speen Street • Framingham, MA 01701 • Phone (508) 872-8200 • Fax (508) 935-4015

memory bandwidth. Game developers, although essentially a distinct set of users, employ techniques from both the animation/rendering and video editing disciplines. As such, they face the same challenges as 3D animators and video editors.

A few years ago, the first branded personal workstations emerged on the scene. The price was right, but the technology could not compete with the performance of traditional workstations.[2] For the next three years, the digital media market continued to be dominated by RISC-based machines. Since then, however, usage of NT-based workstations has surpassed that of traditional workstations in the digital media market. Branded personal workstations will continue to drive growth in this space.

Early on, the relatively prohibitive prices of Unix systems kept small animation and game development houses from entering the market. Armed with the relatively less expensive, 32-bit capable, NT-based workstations, content creation houses were able to buy more systems and employ more digital artists.

The landscape of today's digital media market remains dichotomous. The largest and most complex problems, such as film compositing and volumetric rendering, are being tackled by multiprocessor, 64-bit capable, memory-laden, Unix-based workstations and servers. More manageable tasks, such as 3D animation and multimedia authoring, are being completed on NT-based workstations.

Although large, traditional workstations do an excellent job of handling the most challenging of digital media tasks, the attached price tag can be prohibitive. However, while current NT workstations are much less expensive, they are also resource-constrained. This affects their ability to manage large data sets, create models at full resolution, and render effectively in real time.

3D Animation and Rendering

Many animators already feel the restrictions of a 32-bit architecture. Memory availability, floating-point performance, and I/O bandwidth are among their most pressing concerns. However, a 64-bit architecture offers a compelling solution to these problems: increased performance through more floating-point accuracy and registers and more addressable memory because of longer word lengths. Both will benefit the digital media end user.

2 IDC defines traditional workstations as Unix-based systems, utilizing RISC processors, with a minimum level of graphics and a high level of software certification and customer service.

In addition, leading 3D animation packages such as Discreet's 3D Studio MAX, Alias|Wavefront's Maya, Softimage's XSI, and NewTek's LightWave 3D are all being ported to IA-64. Although the 4GB real memory limit currently achievable with a 32-bit architecture is enough for the majority of digital content creation (DCC) applications, those end users tackling the largest 3D animation tasks require more headroom. Independent software vendors (ISVs) employ unwieldy software solutions to handle the lack of addressable memory. The resulting segmentation often lessens overall performance. To compensate, a majority of time is spent working with wireframe renderings, which strain system resources less than rendering at full resolution. However, working with wireframe models hampers animators' peak creativity. Animators require systems that have increased memory addressability, better floating-point performance, and larger bandwidth. The EPIC architecture and 64-bit capability of Itanium represent Intel's solution to these problems, and as a result, many ISVs are porting their software packages to run on it.

The ISVs that are porting their products want to utilize the improved floating-point performance and additional addressable memory afforded by 64-bit architectures. Free of the complexities of memory segmentation solutions, they hope to simplify their software and focus energy on increasing creative functionality. End users need applications to enable them to render at full resolution in real time and special effects software that will enhance existing tools.

Right Hemisphere, a software company that specializes in paint technology, is planning to offer a new version of Deep Paint 3D exclusively for Itanium. The software will have new features for 3D painting and texture mapping, which are designed to take advantage of a higher level of floating-point performance.

Another ISV is taking advantage of these benefits to improve its rendering software. Volume Graphics has ported VGStudio MAX, its volumetric rendering application, and VGL 3.0, its volume rendering toolkit, to run on Itanium. Both VGStudio MAX and VGL 3.0 are built on a scalable, software-based processing architecture meant to capitalize on the parallelism and bandwidth promised by Intel's EPIC architecture. End users will see performance boosts in 3D Voxel graphics and image processing.

mental images, whose rendering software mental ray is utilized by most major digital media ISVs, has already ported and optimized its software for Itanium. Like software from Volume Graphics, mental ray's performance is directly related to floating-point accuracy and the underlying bandwidth of the system.

Faster I/O results in overall system performance gains. Naturally, if a processor can retrieve data from memory more quickly, it is able to perform more calculations, and software works better. With increased I/O and bandwidth, animators will be able to work at full resolution. This will help them to create better, more realistic models.

Larger addressable memory, better floating-point precision, and increased bandwidth will all lead to better overall system performance. However, end users want more than a faster system. They require hardware and software tools that will enhance their productivity. A 64-bit architecture can answer many of the challenges faced by animators using current classes of branded personal workstations.

Video Editing and Compositing

Many features of 64-bit capable machines that are beneficial to animators are also useful for those in the video editing field. Memory addressability, increased front-side bus speed and bandwidth, and improved I/O will all be important. The level of computational complexity runs the gamut in the video editing field. Users in this space range from serious hobbyists and small videographers all the way to large production houses in Hollywood. In each instance, larger addressable memory and increased overall system performance will be welcome.

When filmmakers need to add special effects to a movie, they rely on video compositing software to combine two or more images into a single frame. It has been said that Sony Pictures, in creating an anatomically correct version of Kevin Bacon's body in the movie *Hollow Man*, used video compositing software to create complex shots with multiple layers of texture. A 64-bit capable machine was necessary to handle the large data sets and perform the geometric calculations to make *Hollow Man* realistic.

Most video compositing software packages, which perform computationally intensive operations to render effects in real time, benefit not only from the larger amount of real addressable memory but also from floating-point performance.

Streaming Media

Streaming media is an area within the digital media arena that is still in its infancy but is growing faster than the overall market. Some growth is coming from home enthusiasts, but much of the interest in streaming media creation is from up-and-coming Internet companies that want to lure audiences to their Web pages with slick mini-movies and audio clips. In order to stream media, all data must be encoded and decoded. The encoding process is inherently computationally intensive.

Currently, workstations used for preprocessing and encoding video to create streaming media content are often configured with costly, dedicated graphics acceleration hardware. These digital signal processor (DSP)-based products function as a coprocessor that performs the floating-point arithmetic necessary to convert analog information into a digital format. The very act of digital signal processing requires the use of a data compression/decompression technique. High floating points can accelerate the algorithms responsible for encoding and decoding data.

The next challenge for this market is to remove the need for DSPs entirely and have an on-chip host. Companies are working on media preprocessing automation to further increase the performance of streaming media specifically for Itanium-based platforms.

Itanium's Technical Computing Capabilities

In technical computing, once you get the answer, it is no longer interesting.
— Ancient Cybernetic Proverb

Applications areas such as DCC have historically placed some of the greatest performance demand on computer processors. In technical computing, the solution of one problem inevitably leads to a set of new and more complex problems; that is, the demand for 3D lighting and geometry is followed by the demand for photorealistic graphics. The complexity of the next generation of problems drives requirements for more powerful and complex tools — including computer processors. Thus the terms technical computing and high-performance computing have become virtually synonymous.

DCC applications, much like traditional technical applications, begin with mathematical models of physical phenomena, products, artistic images, and so on, and then they simulate the effects of different physical forces, perspectives, and so forth on the model in an effort to determine how the modeled object will behave in the real world.

This modeling and simulation process leads to a number of requirements for computer processor architectures in such areas as floating-point performance, memory performance, and support for large data sets. This section briefly reviews these requirements and provides an overview of Intel's strategy for meeting these requirements with its Itanium processor.

Floating-Point Performance

Floating-point performance equates to raw speed for precise calculations — how many adds, multiplies, divides, and so forth the system can perform in a given amount of time. Floating-point performance is a function of the ability to perform floating-point operations in parallel and the cycle time for the processor to execute these operations. On the Itanium this number is increased by functional units that can perform two operations in a single pass. The Itanium is configured with two floating-point functional units known as floating-point multiply add calculations (FMACs). These units can multiply two values and add that result to a third value. (Multiply/add operations are at the heart of many technical calculations such as matrix multiplies.) Thus, an Itanium running at 800MHz can produce four floating-point results a cycle for a peak 64-bit performance rating of 3.2 billion floating-point operations per second (GFLOPS).

The Itanium architecture also includes two single-precision (32-bit) FMACs that are tuned for 3D graphics performance and can each perform an additional four floating-point operations per cycle for a 6.4GFLOPS single-precision rating on an 800MHz processor.[3]

It is important to note that these performance numbers will automatically increase with each step-up of clock rates in the Itanium processor family. In addition, the Itanium architecture is designed to allow future versions of the processor to be configured with additional FMACs.

The above analysis presents a best-case scenario in which the functional units are always busy. Although computer processors can maintain peak performance for only brief periods, Intel has incorporated a number of features in the Itanium architecture that help to maximize sustained performance. These include:

- **Pipelined functional units.** Arithmetic operations generally require more than one machine cycle to complete. A pipelining scheme is used to allow the FMACs to produce results each cycle. The arithmetic operations are broken into a set of independent steps, each requiring one machine cycle to complete. The FMACs perform arithmetic operations in an assembly-line fashion, with each step accepting data from the previous step and sending results to the next step. Thus, after the pipeline is full, a result is produced each cycle.
- **Dual-function arithmetic units.** A secondary benefit of the dual-function FMAC strategy is that the processor is able to use both functional units even when the distribution of adds and multiplies is biased toward one operation. For example, if a section of code performs only additions, both FMACs can be employed on the task. In contrast, a system with separate addition and multiplication functional units would use the adder but would have to leave the multiply unit idle.
- **Large register sets.** The Itanium is configured with 128 floating-point registers. The more data that is directly available to the FMACs, the less likely a functional unit will stall due to lack of data. In addition, the large register sets provide a buffer for the memory system to move data in and out of memory.
- **Internal parallelism.** The Itanium can issue up to six instructions per cycles in a fixed set of combinations of four integer arithmetic/logical operations, two load/store operations, two floating-point operations, and three branch operations. The ability to execute

3 The advantage of double-precision or 64-bit operations over single-precision or 32-bit operations is that the former allow larger sets of calculations to be performed before accumulated round-off errors begin to affect the accuracy of final results. 64-bit systems produce full 64-bit results each cycle; a 32-bit system running at the same clock speed generally requires two cycles to produce the same 64-bit results.

multiple operations not only keeps as much of the processor working as possible but also allows for the pre-fetching of data from memory into registers and cache, thus minimizing processor stalls due to data unavailability. The processor also enables a load-double pair instruction to feed the processor with a balance of a memory operation per floating-point operation.

- **Compiler support for parallelism.** The Itanium was designed to allow for closer coordination between the processor and the compilers that generate the machine instructions for the processor. Three instructions are bundled along with a template field where the compiler can provide "hints" to the hardware on the interactions between the instructions. These hints are used by the processor to schedule instructions in real time and for pre-fetching of data for future operations.

Memory Performance — Keeping the Processor Fed

A large fraction of technical applications is memory-bound rather than compute-bound — that is, the speed of the memory system ultimately determines the speed of the application. Ideally, the memory system will move data in and out of the processors fast enough to keep the floating-point functional units from stalling for lack of data.

Memory performance is measured in terms of both latency (i.e., how many cycles it takes to get data from memory to the processor and, in so doing, fill cache lines for subsequent data requests) and bandwidth (i.e., how many bytes of data can be moved in a cycle). Current systems architectures use memory hierarchies to address both latency and bandwidth issues. Hierarchies consist of a main memory and several layers of caches, and they trade off memory speed for size and cost. A small, fast cache is located "close" to the registers and functional units and can supply data at roughly the rate the processor calls for it. Data is staged through successive levels of cache, with each level holding more data and running somewhat slower until main memory is reached.

At the base of the Itanium hierarchy is main memory, which can vary in size and speed depending on individual system configurations and the system bus (or chipset) that connects the Itanium processors to memory and I/O subsystems. The processor can read or write bytes of data to/from memory every bus cycle; thus, for a 133MHz bus, the memory bandwidth is 2.1GBps. The 460GX chipset, which supports the Itanium processor, also has the ability to write an additional 2.1GBps from I/O to memory, for a total of 4.2GBps memory bandwidth. The Itanium processor uses a 4MB L3 (level 3) cache for quick access to large data structures such as texture maps for digital content applications. The L3 cache communicates with the 96KB L2 cache and the register file, moving data at 12.6GBps (16 bytes per 800MHz system clock) and with a 24 cycle latency for floating-point numbers. The L2 cache feeds data directly into the floating-point registers at a rate of 32 bytes of data per clock tick and with a 9 clock latency.

Although the L1 cache is by-passed by floating-point data, it is worth noting that it is divided into a 16KB instruction cache — L1I — and a 16KB integer data cache — L1D. Both caches operate on a 2 clock latency and provide fast, localized access for integer instructions and data.

Support for Large Data Sets

The Itanium processor is Intel's first 64-bit architecture. As such, it opens up new opportunities for Intel-based systems in technical markets. Scientific and engineering problems grow in two directions. First, more advanced analysis of current problems requires more detailed models, thus requiring larger, more detailed data sets. Second, next-generation problems tend to involve more complex product designs or attempt to model more complex phenomena, and thus require large data sets to describe the problems (e.g., to replace a wind tunnel requires the integration of computational fluid dynamics [CFD] airflow, thermal analysis, physics, and chemistry into a single analysis that is 1,000 times more complex than today's computers are capable of processing).

The requirements to operate on larger data sets generate in turn requirements for computer systems to provide larger real and virtual memories. A computer system's addressable memory is usually determined by the size of its integer or address registers. 32-bit architectures can directly address 4GB of either real or virtual memory. Beyond this limit, some form of memory segmentation must be employed.

64-bit architectures can in theory address about 1019 (10^{19}) bytes of data, a number that is so large that it defeats the authors' ability to describe it in meaningful terms. The important point is that 64-bit systems allow computer systems to expand memory virtually indefinitely without having to resort to some form of segmentation. This large memory space has two major advantages for technical users:

1. **Increased applications performance.** A major bottleneck for many technical applications is the time spent in swapping data between disk and memory — access to data on disk is roughly an order of magnitude slower than memory. The large memories provided by 64-bit systems — up to 1.8TB on the Itanium (the 460GX enables 64GB of physical memory; other original equipment manufacturer [OEM] systems can enable larger memory) — allow applications to keep more of the problem set in memory, thus reducing the amount of time spent reading and writing disk files. In the best case, the entire problem can be moved into memory transforming an "out of core" problem into an "in core" problem.

2. **Simplified programming model.** The larger real and virtual memories afforded by 64-bit architectures enable applications developers to design programs without having to divide the problem into smaller, memory-sized segments and then develop a code to manage those segments.

IDC Analysis

Itanium is just the first iteration of Intel's EPIC architecture. Although it is not the only 64-bit architecture in the workstation market, it does provide some interesting solutions for the digital media market: explicit parallelism, speculation, and floating-point arithmetic. An additional benefit is that it maintains backwards compatibility with existing IA-32 processors.

Some digital media users will move from 32-bit branded workstations to ones using Itanium. However, transitions to new architectures are never easy and rarely happen overnight. One very important aspect that will help Intel in this transitional period is that it has forged many partnerships with operating system providers, ISVs, and graphics card vendors. With so many third-party vendors already showing support for Itanium, end users can be assured that if they decide to use IA-64-based workstations, they will have plenty of options from which to choose.

NORTH AMERICA

Corporate Headquarters
5 Speen Street
Framingham, MA 01701
508-872-8200

IDC Canada
36 Toronto Street, Suite 950
Toronto, Ontario, Canada M5C2C5
416-369-0033

IDC Irvine
18831 Von Karmen Ave, Ste 200
Irvine, CA 92612
949-250-1960

IDC Mtn. View
2131 Landings Drive
Mountain View, CA 94043
650-691-0500

IDC New Jersey
120 Wood Ave South, Suite 509
Iselin, NJ 08830
732-632-9222

IDC New York
2 Park Avenue
Suite 1505
New York, NY 10016
212-726-0900

IDC Texas
100 Congress Ave, Suite 2000
Austin, TX 78701
512-469-6333

IDC Washington
8304 Professional Hill Drive
Fairfax, VA 22031
703-280-5161

ASIA/PACIFIC

IDC Asia/Pacific (Hong Kong)
12/Floor, St. John's Building
33 Garden Road
Central, Hong Kong
852-2530-3831

IDC Asia/Pacific (Singapore)
71 Bencoolen Street, #02-01
Singapore 189643
65-226-0330

IDC Australia
Level 4, 76 Berry Street
North Sydney, NSW 2060, Australia
61-2-9922-5300

IDC China
Room 611, Beijing Times Square,
88 West Chang'an Avenue, Beijing,
P.R. China, 100031
86-10-8391-3456

IDC (India) Limited
Cyber House
B-35, Sector 32 - Institutional
Gurgaon - 122002, Haryana, India
91-124-6381673 to 80

IDC Japan
10F The Itoyama Tower
3-7-18, Mita Minato-ku
Tokyo 108-0073, Japan
81-3-5440-3400

IDC Korea Ltd
Suite 704, Korea Trade Center
159-1, Samsung-Dong, Kangnam-Ku
Seoul, Korea 135-729
82-2-55-14380

IDC Malaysia
Suite 13-03, Level 13, Wisma KiaPeng
No. 3, Jalan Kia Peng
50450 Kuala Lumpur, Malaysia
6-03-2163 3715

IDC New Zealand
Level 7, 246 Queen Street
Auckland, New Zealand
64-9-309-8252

IDC Philippines
7F, SEDCCO 1Bldg
Rada Street Corner
Legaspi Street
Legaspi Village
Makati City, Philippines
632-894-4808

IDC Taiwan Ltd.
10F, 31
Jen-Ai Rd, Sec 4,
Taipei 106, Taiwan, R.O.C.
886-2-2731-7288

IDC Thailand
27 Soi Charoen Nakorn 14
Charoen Nakorn Road, Klongtonsai
Klongsan Bangkok 10600, Thailand
66-2-439-4591-2

IDC Vietnam
37 Ton Duc Thang Street
Unit 1606
District-1 Hochiminh City Vietnam
84-8-910-1235

EUROPE, MIDDLE EAST, AND AFRICA

IDC Austria
c/o Loisel, Spiel, Zach Consulting
Mayerhofgasse 6
A-1040 Vienna, Austria
43-1-50-50-900

IDC Benelux (Belgium)
29 Avenue Louis Gribaumont
B-1150 Brussels, Belgium
32-2-779-46-04

IDC Benelux (The Netherlands)
A. Fokkerweg 1
1059 CM Amsterdam
The Netherlands
31-20-669-2721

IDC Central Europe (ECE)
Male Namesti 13
Praha 1 110 00, Czech Republic
420-2-2142-3140

IDC Central Europe (Germany)
Nibelungenplatz 3, 11th Floor
60318 Frankfurt, Germany
49-69-90502-0

IDC Central Europe (Switzerland)
Niederlassung Züerich
WTC, Leutschenbachstrasse 95
CH - 8050 Züerich
Switzerland
41-1-307-1000

IDC Egypt
39 Iraq Street
Mohandesseen, Cairo, Egypt
20-2-336-7355

IDC France
Immeuble La Fayette
2, Place des Vosges, Cedex 65
92051 Paris la Defense 5, France
33-14-904-8000

IDC Hungary
Bajcsy-Zsilinszky út. 57
Building 3, Rooms 103-104
H-1065 Budapest, Hungary
36-1-153-0555/ext. 165, 166

IDC Israel
4 Gershon Street
Tel Aviv 67017, Israel

IDC Italy
Viale Monza, 14
20127 Milano, Italy
390-2-284-571

IDC Nigeria
House 2, 'C' Close
403 Road, 4th Avenue
New Extension, Festac Town
Lagos, Nigeria
234-1-883585

IDC Nordic (Denmark)
Jagtvej 169B
DK-2100 Copenhagen, Denmark
45-39-162222

IDC Nordic (Finland)
Jarrumiehenkatu 2
FIN-00520
Helsinki, Finland
358-9-8770-466

IDC Nordic (Sweden)
Box 1096 Kistagången 21
S-164 25 Kista, Sweden
46-8-751-0415

IDC Poland/ProMarket
Wrobla 43
02-736 Warsaw, Poland
48-22-644-4105

IDC Portugal
Av. Antonio Serpa, 36 Piso 9
1050-027 Lisbon
Portugal
351-21-796-5487

IDC Russia
c/o PX Post, RDS 186
Ulitsa Zorge 10
Moscow 125525
Russian Federation
7-501-929-9959

IDC South Africa
c/o BMI-TechKnowledge
3rd Floor, 356 Rivonia Blvd.
PO Box 4603, Rivonia, 2128
South Africa
27-11-803-6412

IDC Spain
Ochandiano, 6
Centro Empresarial El Plantio
28023 Madrid
34-91-7080007

IDC Turkey
Tevfik Erdonmez Sok. 2/1 Gul Apt.
Kat 9D; 46 Esentepe
Istanbul, Turkey
90-212-275-0995

IDC U.K.
British Standards House
389 Chiswick High Road
London W4 4AE
United Kingdom
44-20-8987-7100

LATIN AMERICA

IDC Miami
Latin America Headquarters
5301 Blue Lagoon Drive
Suite 490
Miami, FL 33126
305-267-2616

IDC Argentina
Trends Consulting
Rivadavia 413, 4th Floor, Suite 6
C1002AAC, Buenos Aires, Argentina
54-11-4343-8899

IDC Brasil
Alameda Ribeirão Preto, 130 cj 41
01331-000 São Paulo
SP Brazil
55-11-253-7869

International Data Corp. Chile
Luis Thayer Ojeda 166 Piso 12
Providencia, Santiago 9, Chile
56-2-231-0111

IDC Colombia
Carrera 40 # 103-78
Bogota, Colombia
571-533-2326

IDC Mexico
Select - IDC
Av. Nuevo Leon No. 54 Desp. 501
Col. Hipodromo, Condesa
C.P. 06100 Mexico, D.F.
52-5-256-1426

IDC Venezuela
Calle Guaicapuro
Edif. Torre Seguros Alianza
Piso 6, Ofc. 6-D, El Rosal
Caracas 1060, Venezuela
58-2-951-3270

IDC delivers accurate, relevant, and high-impact data and insight on information technology to help organizations make sound business and technology decisions. IDC forecasts worldwide IT markets and adoption and technology trends, and analyzes IT products and vendors, using a combination of rigorous primary research and in-depth competitive analysis. IDC is committed to providing global research with local content through more than 500 analysts in more than 40 countries worldwide. IDC's customers comprise the world's leading IT suppliers, IT organizations, and the financial community. Additional information on IDC can be found on its Web site at http://www.idc.com.

IDC is a division of IDG, the world's leading IT media, research, and exposition company.

Analyze the Future

IDC
5 Speen Street • Framingham, MA 01701
(508) 872-8200 • Fax (508) 935-4015 • www.idc.com

Sponsored by Intel
00-347SYSTEM2761