

March 2002
16G1-0302A-WWEN

Prepared by:
Workstations Division
Compaq Computer Corporation

Contents

Introduction	3
Differences Between Intel Xeon Processors	4
Intel Xeon .18 Processors	4
Intel Xeon .13 Processors	4
Benchmark Analyses	4
Business Winstone.....	5
SYSmark 2001	7
Cadalyt	8
ProE 2001i2	10
Summary	13

Analysis of Intel Xeon Processor Frequency Grades and Cache Sizes on Performance Benchmarks

***Abstract:** Compaq Evo Workstations W6000 and W8000 refreshes will feature the leading edge Intel Xeon .18 processor with Hyper-Threading Technology. The new processors are fabricated with the latest .13 μ technology, 512KB-L2 cache, the ability to support frequencies ranging from 1.8 GHz to greater than 2.6 GHz and also includes support for multi-threaded execution.*

The purpose of this paper is to study the performance benefits in terms of processor frequency grades and larger cache sizes of the new Intel Xeon .18 processors versus the previous Intel Xeon .18 processors. This will be done using different industry-standard benchmarks.

Notice

The information in this publication is subject to change without notice and is provided “AS IS” WITHOUT WARRANTY OF ANY KIND. THE ENTIRE RISK ARISING OUT OF THE USE OF THIS INFORMATION REMAINS WITH RECIPIENT. IN NO EVENT SHALL COMPAQ BE LIABLE FOR ANY DIRECT, CONSEQUENTIAL, INCIDENTAL, SPECIAL, PUNITIVE, OR OTHER DAMAGES WHATSOEVER (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF BUSINESS PROFITS, BUSINESS INTERRUPTION, OR LOSS OF BUSINESS INFORMATION), EVEN IF COMPAQ HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

The limited warranties for Compaq products are exclusively set forth in the documentation accompanying such products. Nothing herein should be construed as constituting a further or additional warranty.

This publication does not constitute an endorsement of the product or products that were tested. The configuration or configurations tested or described may or may not be the only available solution. This test is not a determination of product quality or correctness, nor does it ensure compliance with any federal, state or local requirements.

Compaq and Evo are trademarks of Compaq Information Technologies Group, L.P. in the U.S. and/or other countries.

Microsoft, Windows, and Windows NT are trademarks of Microsoft Corporation in the U.S. and/or other countries.

Intel, Pentium, and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.

All other product names mentioned herein may be trademarks of their respective companies.

©2002 Compaq Information Technologies Group, L.P.

Analysis of Intel Xeon Processor Frequency Grades and Cache Sizes on Performance Benchmarks
White Paper prepared by Workstations Division

First Edition (March 2002)

Document Number 16G1-0302A-WWEN

Introduction

Computer performance is highly dependent on key system features. Processor speed, memory bandwidth, graphics cards, and disk drives all play important roles in determining system performance. Highly computationally intensive tasks can benefit from higher frequency processors. A larger cache on the processors can reduce the number of memory accesses and increase system performance on applications that have small data sets (those that can reside on the processor cache). Applications that require very large files use many disk accesses and require more optimization in that area. Memory bandwidth is a crucial factor in getting the data quickly to the processor. This paper will focus mainly on system processor performance.

A good measure of performance is the amount of time it takes to execute a given application. Contrary to popular belief, clock frequency (MHz) and the number of instructions executed per clock (IPC) are not fair indexes of performance by themselves. True performance is a combination of both clock frequency (MHz) and IPC.

Performance = Frequency x IPC

The formula: Performance = Frequency x IPC means that performance can be improved by increasing frequency, IPC or both. Frequency is a function of both the manufacturing process and the micro-architecture. At any given clock frequency, the IPC is a function of processor micro-architecture and the specific application being executed. Although it is not always feasible to improve both the frequency and the IPC, increasing one and holding the other close to constant with the prior generation provides a significantly higher level of performance.

In addition to these two methods for increasing performance, it is also possible to increase performance by reducing the number of instructions that it takes to execute a specific task. Single Instruction Multiple Data-Stream (SIMD) is a technique used to accomplish this. This is done using 128-bit SIMD single-precision floating-point Streaming SIMD Extensions (SSE).

This analysis will discuss the performance differences between different speeds of the previous Intel Xeon .18 processor (Foster) and new Intel Xeon .13 Processor (Prestonia). This paper will also analyze how the larger cache on Xeon .13 determines system performance.

Differences Between Intel Xeon Processors

Intel Xeon .18 Processors

The Intel Xeon .18 processor (Foster) builds upon the Intel Netburst micro-architecture, built with the 0.18-micron process and with 256KB L2 cache, which facilitates high-speed critical calculations, memory accesses, and an Execution Trace Cache. The Execution Trace Cache caches decoded x86 instructions (micro-ops), removing the latency associated with the instruction decoder from the main execution loops. In addition, the Execution Trace Cache stores these micro-ops in the path of program execution flow, where the results of branches in the code are integrated into the same cache line. This increases the instruction flow from the cache and makes better use of the overall cache storage space (12K micro-ops), since the cache no longer stores instructions that are branched over and never executed. The result is a means to deliver a high volume of instructions to the processor's execution units and a reduction in the overall time required to recover from branches that have been mispredicted. The trace cache is a micro-architectural design that has a direct impact in the Intel Pentium 4 (P4) core attaining a higher IPC than the Intel Pentium 3 (P3). However, this has a drawback too. When the processor needs to fetch a new instruction, it must rely on relatively much slower instruction decoders—thereby causing the netburst architecture to idle and wait on the slow decoders.

The Level 2 Advanced Transfer Cache is 256KB in size and delivers a high data throughput between the Level 2 cache and the processor core. The Advanced Transfer Cache consists of a 256-bit (32-byte) interface that transfers data on each core clock. As a result, the processor can deliver a data transfer rate of core speed multiplied by 32 bytes, reported in GB/s. This contributes to the processor's ability to keep the high-frequency execution units executing instructions vs. sitting idle.

Intel Xeon .13 Processors

The new Intel Xeon .13 processor now features a 512KB L2 cache instead of the original 256KB cache in the Xeon .18 Processor. The addition of the extra cache reduces the miss rates versus the 256KB cache misses. The size of the execution trace cache has not been changed nor have any of the other units of the P4 core, but the increase in L2 cache will provide some performance increase for most applications, especially newer ones. This will be evaluated in the following sections.

The Xeon .13 processor is built with a 0.13-micron die shrink. The smaller transistors can switch faster and produce less heat than their older counterparts. This can result in higher clock speeds for these processors. All 0.13-micron CPUs use copper interconnects, which also aid in increasing clock speeds.

Benchmark Analyses

These analyses are based on running benchmarks that focus on real-world applications run by typical users running business applications, such as the following:

- Microsoft Word
- Microsoft Outlook
- Users running Internet applications

- Users running 2D and 3D workstation applications on their computers

This analysis compares the same frequency Xeon .18 and Xeon .13 processors and two frequency grades of Xeon .13 (2.0GHz and 2.2GHz) on a *Compaq Evo Workstation W6000* system. Table 1 details the system configuration for the benchmarks.

Table 1. Benchmark Configuration

System	Compaq Evo Workstation W6000
Number of Processors	1 CPU
Memory	512 RDRAM @ 800MHz
HDD	18GB 15K rpm SCSI 3 (U160 SCSI controller on board)
Graphics Card	NVidia Quadro 2 Pro
Operating System	Microsoft Windows 2000 Pro, SP2, 1-4 CPU
Hyper-Threading	Disabled
Graphics Card Driver	12.95
Graphics Setting	Vertical sync = Always OFF

Business Winstone

Business Winstone

Business Winstone is a system-level, application-based benchmark that measures the overall performance of a PC when running today's top-selling Microsoft Windows-based 32-bit applications on Microsoft Windows 98 SE, Windows NT 4.0 (SP6 or later), Windows 2000, Windows Me, or Windows XP. Business Winstone runs real applications through a series of scripted activities and measures the time a PC takes to complete those activities to produce its performance scores. Higher scores mean better performance. When Business Winstone 2001 runs the test, it runs at least a portion of each application through a script that was developed by eTesting Labs. The script automatically executes commands within that application with no input from the user.

Business Winstone 2001 uses the following applications in its tests:

- Norton Antivirus 2000 from Symantec
- WinZip 7.0
- Microsoft FrontPage 2000
- Lotus Notes R5
- Microsoft Access 2000
- Microsoft Excel 2000
- Microsoft PowerPoint 2000
- Microsoft Project 98
- Microsoft Word 2000
- Netscape Communicator 4.73

Content Creation Winstone

Content Creation Winstone is a system-level, application-based benchmark that measures the overall performance of a PC when it is running under a 32-bit operating system, such as Windows 2000 or Windows XP.

Content Creation Winstone 2001 uses the following applications:

- Adobe Photoshop 5.5
- Adobe Premiere 5.1
- Macromedia Director 8.0
- Macromedia Dreamweaver 3.0
- Netscape Navigator 4.73
- Sonic Foundry Sound Forge 4.5

Following the lead of real users, Content Creation Winstone 2001 allows multiple applications to be open concurrently and switches among those applications. Content Creation Winstone 2001 is a single large test that runs the above applications through a series of scripted activities and returns a single score. Those activities focus on "hot spots" or periods of activity when the PC is working but the user is likely to only see an hourglass or a progress bar.

See Figure 1 for the Winstone Benchmark using Business Winstone and Content Creation Winstone.

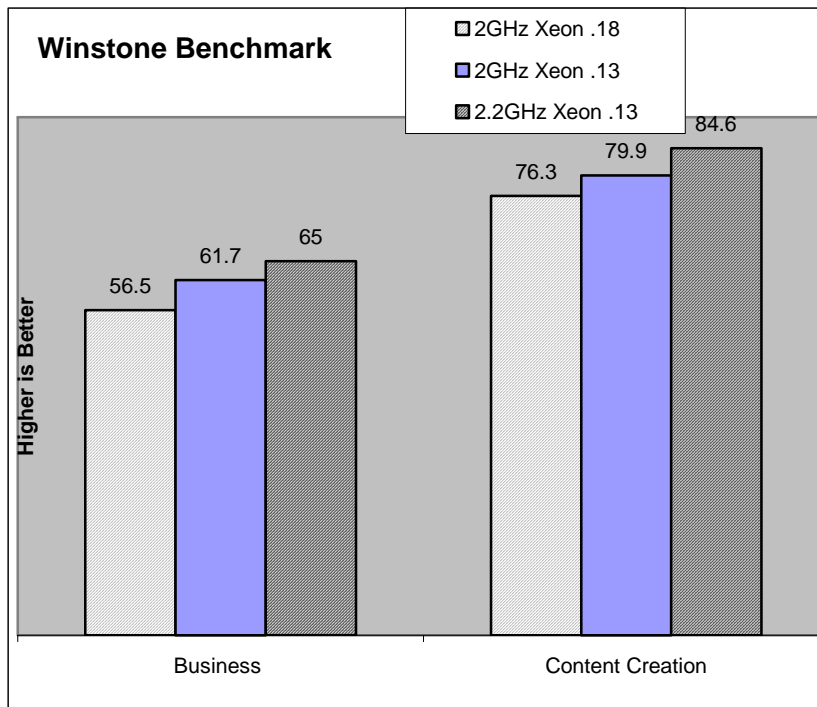


Figure 1

Analysis

Testing shows a 9% performance improvement using higher caches in the Xeon .13s versus the Xeon .18 processor at the same frequency, when running business applications. The business applications take advantage of the larger cache in Xeon .13s due to the reduction of memory cycles, thereby bettering the performance. Only an additional 5% improvement is realized in going to a higher frequency grade Xeon .13s.

In the Content Creation benchmark, a 6% increase is achieved in performance by moving to higher frequency processors. However, only a 4.7% performance boost is realized from the addition of 256KB of L2 cache in the Content Creation test, which could be caused by the disk dependent nature of this test. Winstone Content Creation is not restricted by main memory accesses to the same degree as Business Winstone. It puts a larger strain on the CPU core as is evidenced by the better scaling observed when the core clock is increased. Both of these numbers point to the fact that system benchmarks are not 100% dependent on CPU.

SYSmark 2001

Productivity performance should represent current business usage models of multi-tasking with background computing. The Industry Standard productivity benchmark SYSmark 2001 incorporates mainstream applications for office productivity and Internet Content Creation, as well as the latest business usage models to reflect platform productivity performance.

SYSmark 2001 is a suite of application software and associated benchmark workloads developed by the Business Applications Performance Corporation (BAPCO), a non-profit consortium of leading computer industry publications, independent testing labs, PC hardware manufacturers, semiconductor manufacturers, and software publishers. SYSmark 2001 is a tool that measures system performance on popular business-oriented applications in using the Windows operating environment.

SYSmark 2001 contains fourteen application workloads that are divided into two categories:

- **Office Productivity** suite runs applications, including Dragon Naturally Speaking Preferred Version 5, McAfee Virus Scan 5.13, Microsoft Access 2000, Microsoft Excel 2000, Microsoft Outlook 2000, Microsoft PowerPoint 2000, Microsoft Word 2000, Netscape Communicator 6.0 and WinZip 8.0.
- **Internet Content Creation** suite runs through Adobe Photoshop 6.0, Adobe Premiere 6.0, Macromedia Dreamweaver 4, Macromedia Flash 5, and Microsoft Windows Media Encoder 7.

See Figure 2 for the SYSmark performance benchmark.

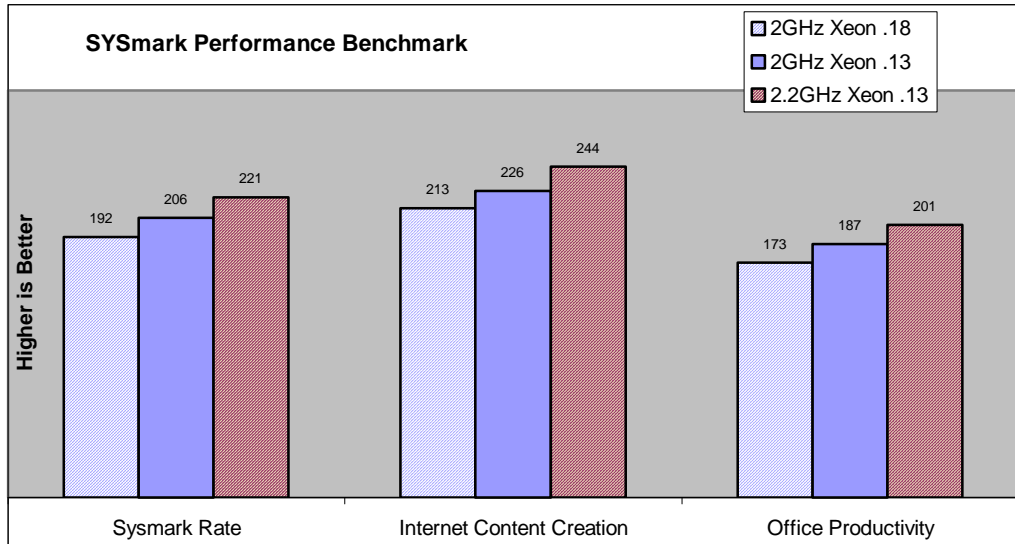


Figure 2

Analysis

SYSmark 2001 is a horizontal benchmark that is a good measure of performance using a multitasked load. This type of application has many branches in the code. BAPCO constructed SYSmark 2001 to run in real time, meaning that the benchmark actually pauses for user input and runs just as quickly as the tasks would if a normal user had been sitting at the keyboard performing all of them.

The Office Productivity test is much more intense than Business Winstone and combines conventional office tasks with virus scans and archive compression tasks, using WinZip.

The added cache results in an 8% boost for the Xeon .13 processor. An additional 7.5% performance improvement is achieved using higher frequency Xeon .13.

The Internet Content Creation test is very memory bandwidth intensive since a large part of the test is composed of encoding a video using Windows Media Encoder. A 6% performance boost is achieved with the larger cache Xeon .13 processor and about an 8% boost in system performance when going to a higher frequency Xeon .13 processor.

Overall, the additional L2 cache of Xeon .13 results in a 6% improvement in performance for the Xeon .13 processor versus the Xeon .18. Additionally, about 6-8% performance boost is noted with increased clock speeds in equivalent processors.

Cadalyst

For a more professional tool, the Cadalyst Labs C2001 benchmark for AutoCAD 2002 was used. This test is comprised of two parts: 2D and 3D.

The 2D test opens drawings, runs array calculations, performs pan/zoom/erase /save operations, and operations with orthogonal lines and cubical splines. The 3D test suite runs tests, which include standard rotation/rendering, OpenGL rotation, DXFouts, in addition to some of the 2D tests. Quadro2 Pro and Wildcat 5110 graphics cards were used to run this benchmark.

See Figure 3 for the Cadalyst performance benchmark.

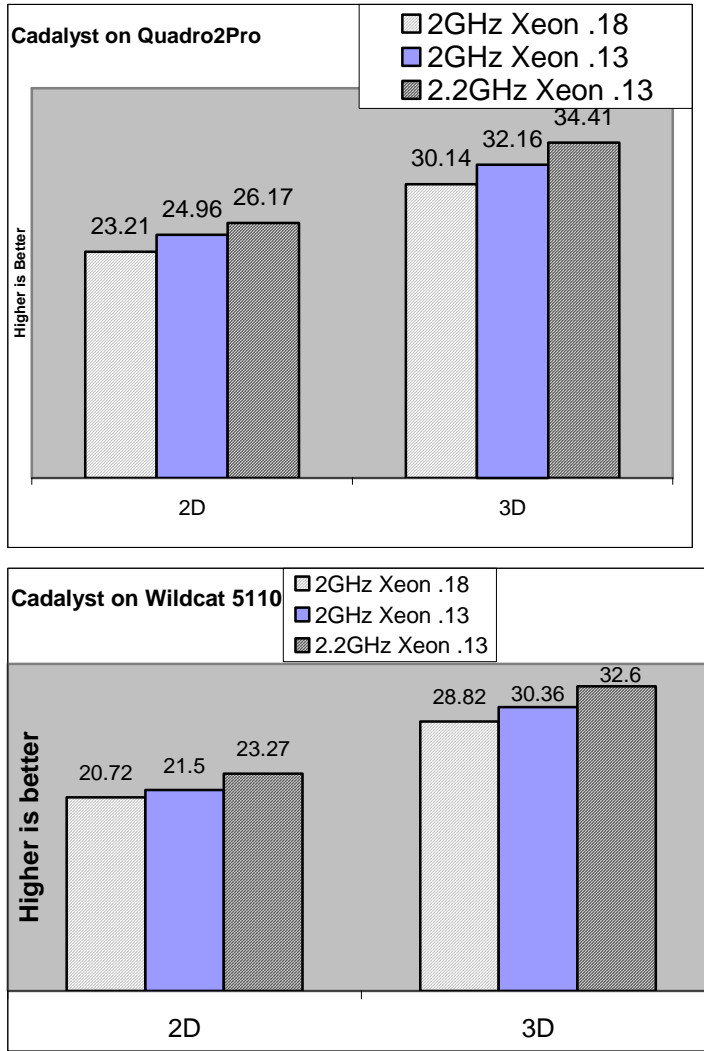


Figure 3

Analysis

In both 2D and 3D tests, the additional cache in Xeon .13 provides a 7% increase in performance using the Quadro2 Pro graphics card and about 3-5% increase using the Wildcat 5110 graphics card.

The 2D tests are less graphics card intensive and more CPU intensive, thus the CPU performance greatly influences AutoCAD 2002 performance. From these scores alone, the 2.2GHz Xeon .13 shows about a 4% increase in performance from their 2GHz counterparts using the Quadro2 Pro card and about an 8% increase using the Wildcat 5110 card.

In 3D tests, there is about a 7 % increase in performance going to higher speed Xeon .13s in both cards. Quadro2 Pro performs almost 13-16% better than Wildcat 5110 in 2D tests and about 4-5% better in 3D tests. This is because Quadro2 Pro drivers are more DirectX optimized. This points to the importance of balancing system components versus focusing on raw CPU speed only.

In the next section, the performance of the system using more graphics-intensive tests using OpenGL tests like ProE is examined by comparing the performance of the same processors with two different graphics cards.

ProE 2001i2

The ProE 2001i2 benchmark is comprised of 17 tests. The model used in the benchmark is a realistic rendering of a complete photocopy machine consisting of approximately 370,000 triangles. The ProE 16 graphics tests, each of which measures a different rendering mode or feature, include the following, plus a time test:

- The first three graphics tests measure wire-frame performance using the entire model.
- The next four tests measure different aspects of shaded performance, using the same model. Each of these tests executes exactly the same sequence of 3D transformations to provide a direct comparison of different rendering modes.
- The next four tests use a subassembly, and compare the two FASTHLR modes, the default shading mode, and shading with edges. These tests also execute a common sequence of 3D transformations.
- The last five graphics tests use two different instances of the model—the first three without its outer skins (to illustrate the effect of FASTHLR and level-of-detail operations), and the last two to illustrate complex lighting modes and surface curvature display.
- The last test is an aggregate of all time, not accounted for by the previous 16 tests, and is a mix of CPU and graphics operations.

Scores are generated for all 17 tests. Composite numbers are provided for each set of graphics tests (shaded, sub-assembly, wire-frame, and others) and there is an overall composite score for graphics and CPU operations.

See Figure 4 for the ProE performance benchmark.

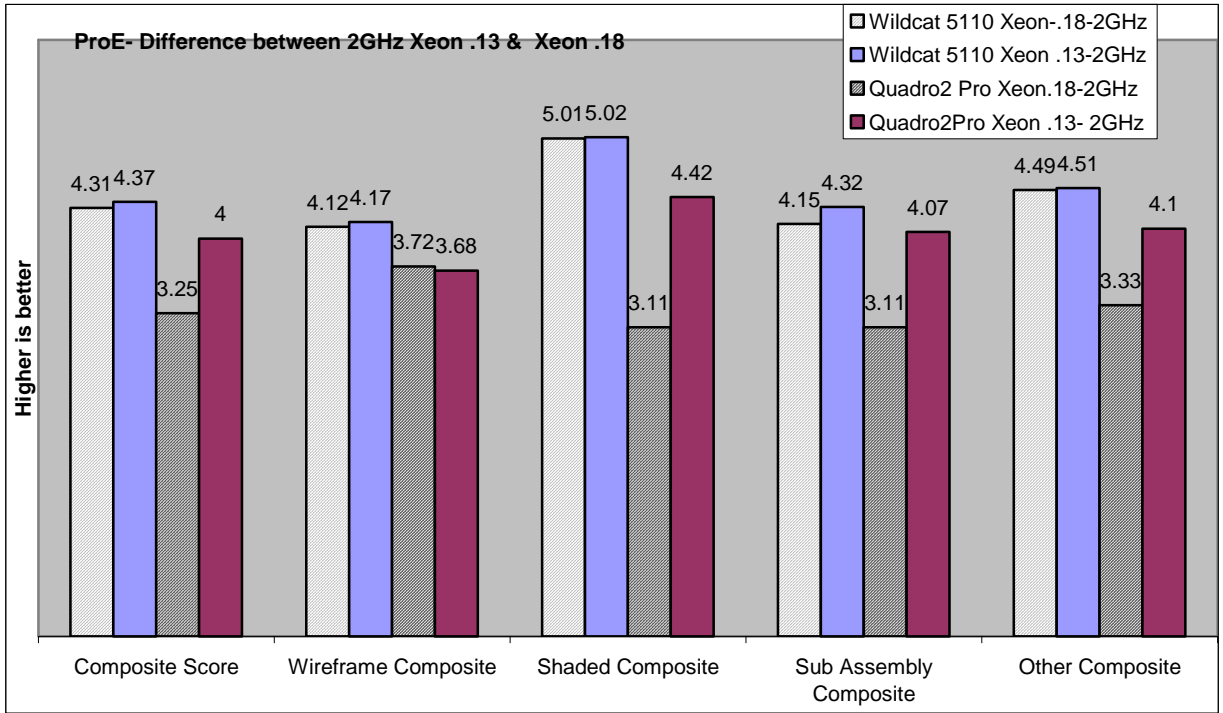


Figure 4

Analysis

Figure 5 shows the results for the Intel 2GHz Xeon .18 and Xeon .13 using two different graphics cards: WildCat 5110 and Quadro2 Pro. The extra cache in the Xeon .13 processor does not seem to make any performance difference in using the Wildcat 5110 card. The Quadro2 Pro card shows a lot more performance improvement using larger cache Xeon .13s. The Wildcat 5110 card performs significantly better than the Quadro2 Pro card in almost all of the ProE benchmarks, which leads to the conclusion that the graphics card performance is a lot more important in determining the system performance in 3D applications, compared to processor performance.

Next, the test compared the performance between a 2GHz Xeon .13 and a 2.2GHz Xeon .13 using the same two graphics cards.

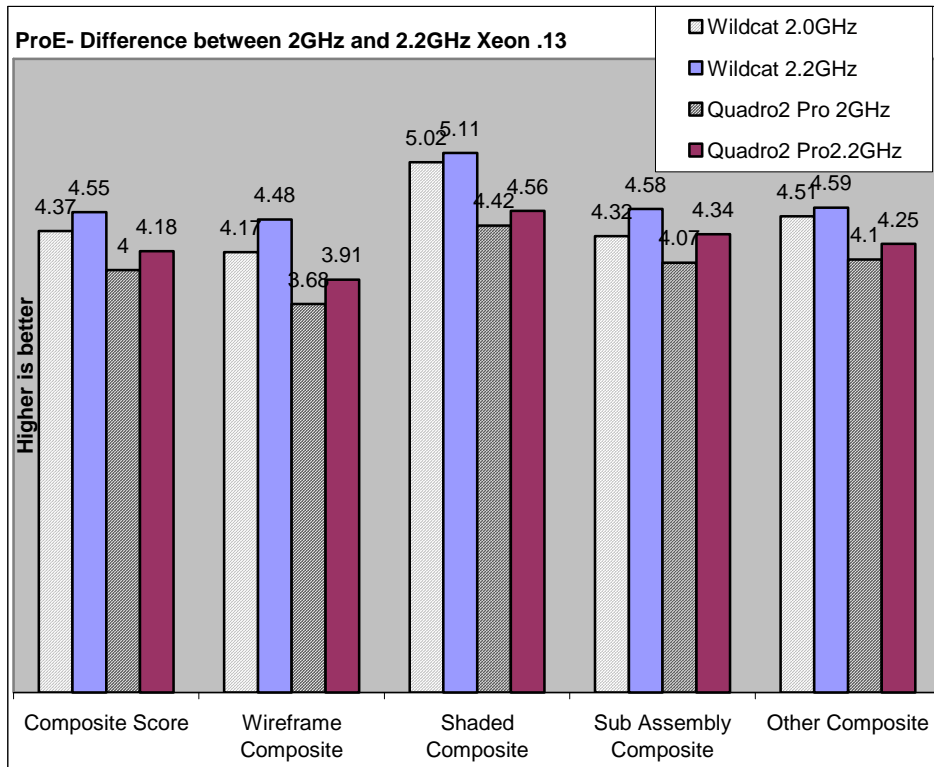


Figure 5

Analysis

Some performance improvement is achieved in moving to a higher frequency processor, but not as significant as the performance difference between the two graphics controllers themselves. The Wildcat 5110 card shows superior performance when compared to the Quadro2 Pro card, using the same processor on the same system. Thus, the graphics controller plays a very important part in determining system 3D performance.

Summary

Applications can be broadly divided into two categories:

- Integer/basic office productivity applications
- Floating-point/multimedia applications

The IPC cycle achieved by each of these different application categories varies significantly. This is dependent on the number of branches that the application code typically takes, and the predictability of these branches. The more difficult the branches are to predict, the higher the possibility of mispredicting and performing nonproductive work, resulting in lower performance.

Integer and basic productivity applications, such as Microsoft Word and Microsoft Excel tend to require several branches within the code that are more difficult to predict and, therefore, reduce the overall IPC. As a result, performance increases on these applications take less advantage of improvements in processor micro-architectures, such as deeper pipelines. This is seen by the 6-8% performance boost going to a higher frequency in SYSmark and Winstone benchmark results. Additionally, smaller datasets can benefit more from the larger cache on the processor.

Floating-point and multimedia applications tend to have branches that are more predictable and, therefore, contain a higher average IPC potential. As a result, these types of applications generally scale very well with frequency and are inclined to benefit greatly from deeper pipelines. Therefore, some performance improvement is seen in going to higher frequency processors using Cadalyst and ProE benchmarks. However, there is not much performance improvement using higher caches in these benchmarks. The graphics controller plays a more important role in determining the performance of the system in running 3D benchmarks. Bigger cache and higher frequencies offer some, but not a significant performance gain in 3D applications.