

PCMark[®]05

PC Performance Analysis

Sami Niemelä
Futuremark Corporation
June 15th 2005

Table of Contents

OVERVIEW3

SYSTEM AND COMPONENT BENCHMARKING3

OUR DEVELOPMENT METHODOLOGY5

PCMARK05 OVERVIEW8

TEST SUITES8

 SYSTEM TEST SUITE..... 8

 CPU TEST SUITE..... 9

 MEMORY TEST SUITE..... 10

 GRAPHICS TEST SUITE..... 10

 HARD DISK DRIVE TEST SUITE..... 11

THE TESTS11

 FILE COMPRESSION / DE-COMPRESSION..... 11

 FILE ENCRYPTION / DECRYPTION..... 12

 TEXT EDIT..... 12

 WEB PAGE RENDERING..... 12

 IMAGE DECOMPRESSION..... 13

 AUDIO COMPRESSION AND DECOMPRESSION..... 13

 VIDEO ENCODING..... 13

 PHYSICS AND 3D 14

 GRAPHICS MEMORY..... 15

 MEMORY TESTS..... 15

 2D - TRANSPARENT WINDOWS..... 15

 VIDEO PLAYBACK..... 15

 3D - FILL RATE..... 15

 3D - POLYGON THROUGHPUT 16

 3D - PIXEL SHADER..... 16

 3D - VERTEX SHADER..... 16

 HDD TESTS 17

SCORE CALCULATION.....18

 PCMARK SCORE 18

 CPU SCORE..... 19

 MEMORY SCORE..... 19

 GRAPHICS SCORE..... 19

 HDD SCORE..... 19

 EXAMPLE SCORES..... 19

ONLINE RESULTBROWSER20

SYSTEM INFORMATION21

LIMITATIONS22

CONCLUSION22

Overview

This paper introduces PCMark[®]05, the latest version in the PCMark benchmark series by Futuremark[®] Corporation. PCMark05 is a state-of-the-art benchmarking tool designed to be easy to use and widely accessible. It produces highly reliable and detailed benchmarking results with a simple, intuitive user interface.

PCMark series has become a popular tool for benchmarking home and office PC performance. Its earlier versions, PCMark2002 and PCMark04 have distributed millions of copies worldwide.

PCMark05 builds on Futuremark's strong benchmark development experience by providing a sophisticated tool for measuring PC performance for home usage. PCMark05 supports the complete benchmark cycle – allowing you to benchmark your PC, view the resulting benchmark details, compare your results to those of others and analyze how to improve your PC performance.

PCs continue to improve at a rapid pace. As each PC component (CPU, system memory, graphics card, hard disk, etc.) increases in speed or capacity, home users are finding it increasingly difficult to evaluate the benefits of these changes. Software applications that may or may not take full advantage of the latest hardware add further to the confusion. The goal of PCMark is to solve this problem by providing an easy-to-use tool that outputs easy-to-understand performance figures.

In this paper we discuss our approach to PC benchmarking and the development of PCMark05. We first explore the various facets of successful system and component benchmarking. We next cover our multi-step development process that is central to developing a dependable benchmark. We describe the individual test suites and technologies used in each in detail and present our score calculation approach and the formulae used to generate the results. Lastly, we illustrate our integrated and highly popular online comparison tool, the Online ResultBrowser.

System and Component Benchmarking

System benchmarking allows users to accurately evaluate the overall performance of their PC hardware for a defined usage type. *Component benchmarking*, by contrast, measures the performance of individual PC components such as CPU, memory, graphics card etc. Benchmarking PCs can be a complex process as hardware comes in a tremendous variety. Each PC component – motherboard, CPU, system memory, graphics card, etc. – has multiple possible manufacturers and variations. Furthermore, the rapid pace with which manufacturers introduce improved versions of these components makes system benchmarking a very challenging task.

PCMark05 supports both system and component benchmarking. System benchmarking produces a measure of the PC's overall performance for a specific type of usage. In our case this is home usage. It is a single number that is easily comparable to that of other PCs. However, advanced users and professional testers may want to further explore the performance of individual components such as the CPU, the memory, the graphics subsystem, the hard disk, etc. Component benchmarking helps users to isolate the performance of individual components, exposing the component's strengths and weaknesses. PCMark05 includes an exciting feature - the ability to create custom benchmarks by putting together tests tailored to specific needs.

Usage Profile: Identifying the typical usage being measured – the usage profile – is central to building a benchmark. Average home PCs may predominantly be used to browse the internet, compose emails and view digital photographs. In contrast, gaming PCs will likely have heavy 3D graphics usage for playing the latest games. The set of tasks for the benchmark must closely emulate the chosen usage profile. PCMark05 targets typical home usage.

Benchmarking Approaches: A well-defined usage profile allows us to design the set of tasks that are run and measured by the benchmark – the workload. The workload must have the characteristics of the chosen usage profile. In PCMark05 the usage profile is typical home PC usage. Thus, PCMark05 stresses the PC in a similar manner as PCs would experience in real life-home usage. There are two common approaches for building the workload: *Application* and *Synthetic* benchmarking.

Application benchmarks include complete applications the user would typically run. Such benchmarks would include the complete word processor, email client, graphics authoring software, etc. The primary benefit of this

* Note that PCMark05 Basic does not include all tests.

approach is that the performance results are likely to correlate closely with the actual performance of using the included applications. Application benchmarks have some drawbacks in their usability and accessibility. They typically have a very large installation footprint and require a large amount of disk space making distribution and usage more challenging. Moreover, they may take a long time to run due to the inherent size and complexity of the benchmark. Also, defining the performance of a single PC component might be more difficult, as the applications tend to use the system more widely.

Synthetic benchmarks tend to have the opposite strengths and weaknesses as application benchmarks. They include code fragments written specifically to mimic the work performed by real-world applications. This allows them to be small in file size and quick to run. However, synthetic benchmarks often face the criticism of not being the “real thing” as actual applications are not being used. Despite this, synthetic benchmarks can be powerful tools for isolating performance of certain parts of the hardware.

PCMark05 draws the best parts of both of these approaches by building the workload using an *Application-based* approach. The workload consists of actual applications with freely available public code as much as possible[†]. However to keep the benchmark small and easy to run, we use only relevant parts of these applications instead of complete applications.

Complete Benchmarking Cycle: A good benchmark provides an accurate performance measurement as well as assists the user through the full benchmarking cycle of analyzing the results, comparing results against other PCs, exploring PC improvements, and leading finally to higher performing PCs. PCMark05 and Futuremark’s online tools were designed to support this process.

PCMark05’s system score is a single, globally recognized number that represents the overall PC performance for home usage. PCMark05 also produces several component scores. The complete benchmark data available to the user consists of results from 85 tests and over 300 pieces of detailed PC configuration information. As benchmark results are useless in isolation, the user can take the next step and submit the results to Futuremark’s Online ResultBrowser (ORB). The ORB allows the user to compare the results to millions of other benchmark results in Futuremark’s benchmark database. With this comparison, the strengths and weakness of the user’s PC are revealed. This provides the user unique insights into how much performance improvement each component upgrade may yield. As a result, the user is more educated and can make a more informed purchasing decision.

Other Benchmarks: There have been several efforts in PC benchmarking. Some prominent programs are listed below:

- SPEC (www.spec.org), a not-for-profit consortium of companies, produces CPU integer-workload and floating point-workload component benchmarks.
- SYSmark[®], MobileMark[®] and Webmark[®] series from the BAPCO[®] Corporation (www.bapco.com) are application benchmarks for internet content creation and office productivity usage. They include a large set of complete applications such as the Microsoft[®] Office suite and Adobe[®] Photoshop[®].
- SiSoftware[™]’s Sandra[™] (<http://www.sisoftware.net/>) consists of a large set of different diagnostic tools. It also includes several components-level benchmarks.

With PCMark05, it has been our goal to create a benchmark that is accessible to all users. We want to help users understand their PC performance. For this reason, we have made the benchmark very intuitive to run. With a single click, the user is able to get performance results in minutes. We also have tried to keep the installation size small to enable quick downloads and faster access to results. We have done this by only including the relevant parts of real applications and not the complete application. Furthermore, we have worked hard to make the benchmark highly accurate. We have worked in cooperation with all key players in the industry to ensure that the workload accurately reflects real-world usage. Lastly, we provide online tools that allow users to put their PC performance results in context and make intelligent purchasing choices.

[†] Note that certain tests in PCMark05 have been written entirely by Futuremark. These tests, e.g. the memory test suite, require us to isolate the performance of a particular subsystem. Such code is not easily extractable from existing applications and requires us to implement the tests ourselves.

Our Development Methodology

Futuremark approaches all of the benchmarks it creates with a standard development methodology. We believe that the process we follow is central to the development of a successful and dependable benchmark.

The key part of the development process is cooperation with the key PC technology developers. This cooperation ensures that our benchmark is designed to measure the right things correctly. It ensures that the benchmark will contain tests that are relevant and that the measurements are meaningful. PCMark05 was developed in cooperation with the following members of Futuremark's Benchmark Development Program (BDP): AMD, Dell, Intel, Microsoft and NVIDIA.

These companies are willing to cooperate with us because they share the vision that strong, objective benchmarks are in everyone's interest. The Benchmark Development Program allows these companies to participate in designing leading benchmarking standards in the PC industry. The cornerstones of our design process are *transparency* and *neutrality*. We make a strong effort to document all processes that make up the benchmark; we continuously strive to make these documents better. Also, we always maintain the highest standards of neutrality, neither favoring nor dis-favoring any party.

The figure below depicts a high-level view of our benchmark development methodology.

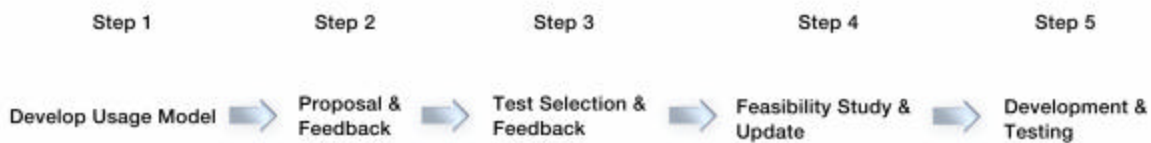


Figure 1: Benchmark Development Methodology

Step One. The benchmark development process begins with identification of target usage profile – the type of PC usage we want to benchmark. For PCMark05, we chose Home PC Usage. Using a variety of resources, we developed a usage model consisting of several categories of tasks. The usage model for home PC is shown in Table 1.

Table 1: Home PC Usage Model

Task Category	Task	Weight		
		Low	Medium	High
Productivity	Text Editing		x	
	Spreadsheet Calculations	x		
	Presentations	x		
	Picture Viewing			x
	High-End Document Editing	x		
Internet	Electronic Mail		x	
	Internet Browsing			x
Entertainment	Music Playing & Recording			x
	Video Playing & Recording			x
	Game Play			x
Other	Desktop Usage			x
	Compression			x
	Encryption			x

Within each category, we list a set of representative tasks. Each task is further assigned a weight reflecting its importance to the chosen usage.

To develop this model, we drew upon our own experience with previous benchmarks and the feedback we have received from the millions of users of these products. Communications with our benchmarking community also provided valuable input. BDP member companies were another source of insights.

Step Two. The document produced in the previous step is a proposal. It is designed to present features and implementation options in a format amenable for getting constructive feedback. This proposal was circulated to our BDP members. The feedback received aided us in choosing the tests and implementation methods.

Step Three. In this step, we incorporated the feedback and modified the task list as was appropriate. At this point some lower priority tasks were dropped. We next selected the application fragments to use for each task in the usage model. The applications chosen depended on variety of factors: licensing, size, popularity, etc.

We then created a written benchmark specification. Each of the workload tests is specified in detail with exact versions of the technologies used. For PCMark05, this is summarized below:

Table 2: PCMark05 Tests Used

Task Category	Task	Test(s) Used	Applications Used
Productivity	Text Editing	Text edit	Microsoft® WordPad
	Picture Viewing	Image processing	JPEG
Graphics memory		DirectX®	
Internet	Internet Browsing	Web page rendering	Microsoft® Internet Explorer 6
Entertainment	Music Playing and Recording	Audio compression and decompression	OGG Vorbis
	Video Playing and Recording	Video encoding and playback	WMV
	Game Play	Physics Calculation	Havok
		3D Graphics	DirectX®
General	Desktop Usage	Transparency, moving windows	Windows® 2D graphics API
	Compression	File compression and decompression	Zlib
	Encryption	File encryption and decryption	AES
	Accessing & saving files	HDD	RankDisk

The specification was circulated to the BDP members. Their feedback was analyzed and incorporated after thorough review by Futuremark, assuring that our cornerstone of neutrality was not in conflict.

Step Four. In the next step, we implemented prototype code to see if the available technology would support our plans. This is the stage where we may discover that certain tests are not possible, or may be surprised to discover that more can be achieved. The results were then incorporated into the specification and again circulated to the BDP members.

Step Five. The fifth step consists of implementing the workloads or tests. Periodic releases were made to the BDP members; after each step their feedback was taken into account.

PCMark05 Overview

PCMark05 uses public domain applications in several tests which enables users to examine the source code of those tests where applicable. In addition to the open source code used, also commercial programs or portions of such are used.

PCMark05 measures home PC usage. The workload is designed to stress the PC in the same manner as typical home usage does. The workload includes a test suite that gives an overall PCMark score as well as several component test suites that give individual scores measuring the CPU, memory, graphics and hard disk drive. Some of the tests run concurrently in separate threads. Multithreading is commonly used in software design to maximize performance and resource utilization. Lastly, PCMark05 includes the ability to define custom test suites that are tailored to your own specific needs. The user can choose various tests and compose them in multithreaded sets. This allows the user to model real-life usage where several applications running in a multithreaded fashion compete for PC resources.

We also allowed vendors to examine any code that was written by us, for example the memory tests. The vendors suggested optimized code paths, which we reviewed for accuracy and implemented at our discretion.

PCMark05 is compiled with Microsoft[®] Visual Studio[®] .NET 2003 with the following compiler switches: /Ox /Ob2 /Ot /G7

Note that vendors are not allowed to detect a running PCMark05 instance in any of their drivers or software and use that to reduce the workload or alter the behavior. Our policy for generic optimizations is available on Futuremark website (www.futuremark.com).

Test Suites

The tests in PCMark05 have been divided in different test suites depending on what part of the PC they measure. There is a test suite for the whole system, and separate test suites for CPU, Memory, Graphics and HDD. In this chapter the test **suites** are described. Individual tests are described in their own chapters later in this document.

System Test Suite

The system test suite is a collection of tests that are run to generate an overall PCMark score. Three of the test scenarios are run multithreaded – two of those include two simultaneous tests and one is with four tests running simultaneously. The remaining eight tests are run single threaded. The setup for the default run is shown in Figure 2.

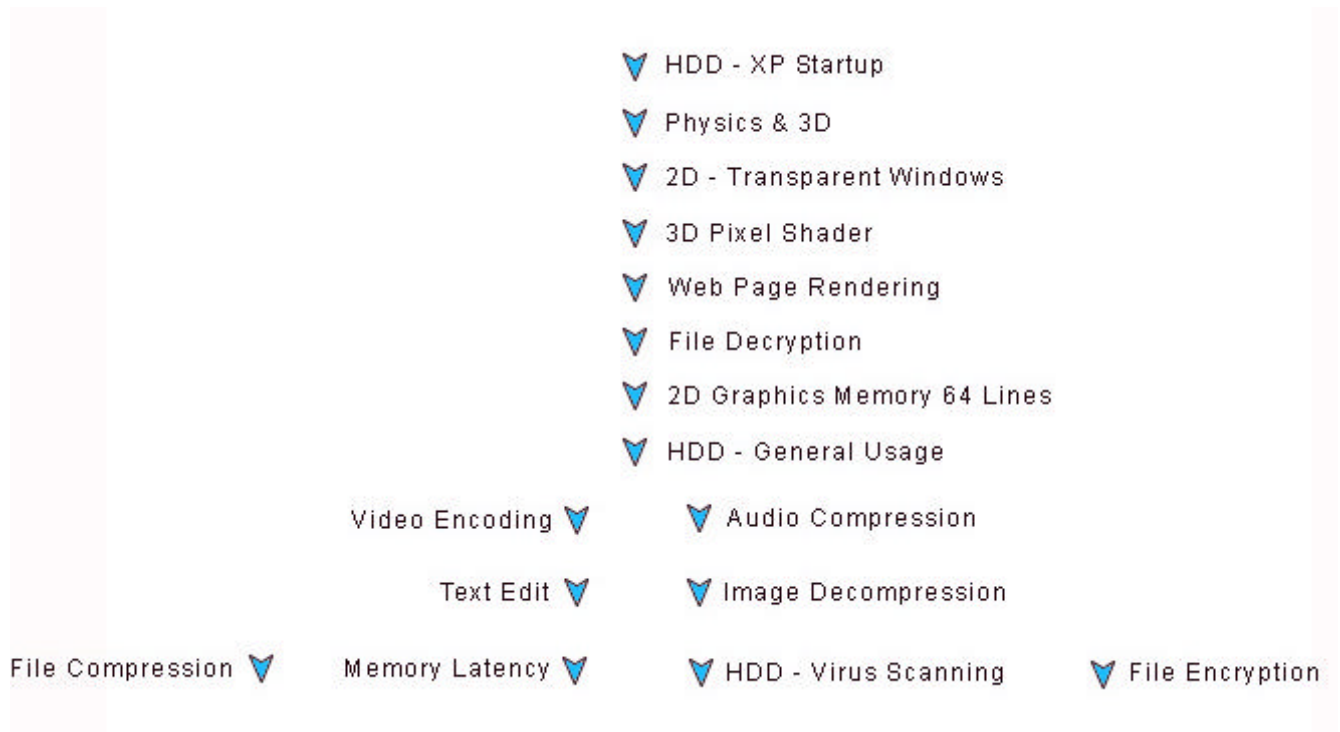


Figure 2: System Test Suite Default Setup

System tests stress different parts of the PC. These tests include some tests which stress mainly the CPU, some tests stressing both the CPU and memory, as well as some tests stressing the HDD.

CPU Test Suite

The CPU test suite is a collection of tests that are run to isolate the performance of the CPU. The CPU Test Suite also includes multithreading: two of the test scenarios are run multithreaded; the other including two simultaneous tests and the other running four tests simultaneously. The remaining six tests are run single threaded. The setup for the default run is shown in Figure 3.



Figure 3: CPU Test Suite Default Setup

Memory Test Suite

The memory test suite is a collection of tests that isolate the performance of the memory subsystem. The memory subsystem consists of various devices on the PC. This includes the main memory, the CPU internal cache (known as the L1 cache) and the external cache (known as the L2 cache). As it is difficult to find applications that only stress the memory, we explicitly developed a set of tests geared for this purpose. The tests are written in C++ and assembly. They include:

- Reading data blocks from memory
- Writing data blocks to memory
- Performing copy operations on data blocks
- Latency

Depending on the size of the data blocks, the tests use the main memory, L2 cache or L1 cache. Before each test, cache cleanup procedures are performed to prevent the cache contents from affecting the results.

The tests in the Memory test suite are described below:

- Memory Read 16 MB block size
- Memory Read 8 MB block size
- Memory Read 192 KB block size (L2 cache)
- Memory Read 4 KB block size (L1 cache)
- Memory Write 16 MB block size
- Memory Write 8 MB block size
- Memory Write 192 KB block size (L2 cache)
- Memory Write 4 KB block size (L1 cache)
- Memory Copy 16 MB block size
- Memory Copy 8 MB block size
- Memory Copy 192 KB block size (L2 cache)
- Memory Copy 4 KB block size (L1 cache)
- Memory Latency 16 MB block size
- Memory Latency 8 MB block size
- Memory Latency 192 KB block size (L2 cache)
- Memory Latency 4 KB block size (L1 cache)

Graphics Test Suite

The graphics test suite is a collection of tests that isolate the performance of the graphics subsystem of the PC. With each revision of Microsoft® Windows® and related productivity applications, the 2D graphics capabilities of the home PC are pushed further. PCs are being used to look at increasingly sophisticated content such as high-resolution images and video. 3D game play also continues to get more complex fueled by improvements in 3D software and hardware platforms. With this in mind, we created a set of test to expose the graphics performance of your PC.

In the graphics test suite, we have both 2D and 3D graphics presented, as well as video playback. The tests in the suite are described below:

- 2D – Transparent Windows

- 2D – Graphics Memory 64 Lines
- 2D – Graphics Memory 128 Lines
- Video Playback
- 3D – Fill Rate
- 3D – Polygon Throughput
- 3D – Pixel Shader
- 3D – Vertex Shader

Hard Disk Drive Test Suite

The hard disk drive test suite is a collection of five tests that isolate the performance of the hard disk. The hard disk is an important part of the PC providing quick access to large amounts of data. Today's computers come with hard disks that can store several billion bytes (gigabytes) of data.

Various aspects affect the HDD workload. For example one task may consist of mostly writing a large file, whereas another task might be mostly reading small files on the disk. The PCMark05 HDD test suite has been created to include different workloads for the disk to widely cover the use. The HDD test suite is described below:

- HDD – XP Startup
- HDD – Application Loading
- HDD – General Usage
- HDD – Virus Scanning
- HDD – File Write

The Tests

File Compression / Decompression

With increasingly powerful PCs, the ability of applications to handle larger, more sophisticated documents, and sharing files over internet being highly popular, file compression has become a very common task. Compression algorithms are also embedded in many applications; very often the home user may use compression and decompression without realizing it.

For file compression and decompression tasks we use the zlib library. Zlib is a free, general-purpose data compression library available on a large number of platforms. It is used in over 500 applications (listed on <http://www.gzip.org/zlib/apps.html>) including several Microsoft® Office products, Microsoft® Internet Explorer, Symantec™ Norton AntiVirus™, Adobe® Photoshop®, and Macromedia® Shockwave™.

Compression and decompression is integer computation and primarily stresses the CPU. Zlib library version 1.2.2 is used. For this test, the input data files are:

- A 3 Megabyte executable
- A 2 Megabyte document
- A 10 Megabyte video file
- A 3.5 Megabyte audio file

The test is run continuously for 20 seconds. The result is expressed in Megabytes processed per second.

File Encryption / Decryption

Encryption is the conversion of data into a form (called a ciphertext) that cannot be easily understood by unauthorized people. Decryption is the process of converting encrypted data back into its original form, so it can be understood. Although the use of encryption/decryption is as old as the art of communication, it is particularly important for current day networks which may be easier to tap into and can carry highly sensitive information. Many companies already enforce policies where all emails must be encrypted.

For the File Encryption and Decryption task we use the Rijndael/AES Algorithm (<http://www.esat.kuleuven.ac.be/~rijmen/rijndael/>). In the year 2000, the US National Institute of Standards and Technology (NIST) announced that Rijndael was selected as Advanced Encryption Standard (AES). AES is unpatented and license-free, and is available for free for all uses.

Encryption and Decryption stress the CPU with integer operations. For both we use the Crypto++ library version 5.0 (www.cryptopp.com). This is a free C++ class library containing implementations of various cryptographic algorithms including Blowfish. The input files are:

- A 3 Megabyte executable
- A 2 Megabyte Microsoft® Word document
- A 1.1 Megabyte JPEG image
- A 2 Megabyte WAV audio file

The test is run continuously for 20 seconds. The result is expressed in Megabytes processed per second.

Text Edit

Text editing is something every PC user needs to do in various occasions. In addition to plain text typing, searching for keywords in the text is a common way to utilize a text editor.

For the text editing test, we use the WordPad application provided with the user's Windows installation and perform search & replace operations on the document.

Text Edit is integer computation and primarily stresses the CPU. The input for this test is an 85 Kilobyte text document. The test is run continuously for 20 seconds. The result is expressed in Kilobytes processed per second.

Web Page Rendering

Web browsing is becoming one of the most popular uses of many home PCs. The Internet has proven a successful way to generate revenue both for internet-only businesses and well as traditional stores. The highest traffic sites have moved cautiously towards making their web pages increasingly complex. Internet connection speeds still vary among users and these sites must balance providing rich content and acceptable response times.

For the Web Page Rendering task, we use Microsoft® Internet Explorer 6. This is clearly the predominant web browser. We require that IE 6 be previously installed on the PC. Our goal for this test is to measure the performance of rendering of web pages after they have already been downloaded. We do not measure the speed of downloading web pages. Internet connection speeds can vary due to several factors – type of connectivity (dialup, DSL, cable, etc.), proximity to provider, proximity to web site, etc. – and may not accurately reflect the performance of the PC.

Web page rendering is integer computation and stresses both the CPU and the memory subsystem. The input data for this test are four html pages and related images. These are:

- A 130 Kilobyte web page representing a typical company main page
- A web page containing 7 images (total size approximately 2.9 Megabyte) representing a page containing photos
- A 310 Kilobyte web page containing text representing rendering of documentation
- A web page representing a discussion board, sized approx. 430 Kilobytes.

The test is run continuously for 20 seconds. The result is expressed in Pages processed per second.

Image Decompression

Image operations for home PC usage primarily include viewing photographs taken with a digital camera or shared on a web site. It may also include viewing larger images embedded in a web page or document. All of these involve the process of decoding a large image. The most common format found for large images is JPEG. There are two reasons for this. First, JPEG is particularly well suited for rich color photographs and artwork as it uses 24 bits per pixel to store color data. Second, JPEG compresses well allowing images files to be quite small and hence more suitable for storing and transmitting over networks.

For the Image Decompression task, we use the Standard JPEG library from the Independent JPEG Group. It is production quality software and is freely available for both commercial and non-commercial use at <http://www.iijg.org>.

Image processing primarily stresses the CPU. The JPEG decoding pipeline uses fixed-point IDCT and RGB-24 output pixel format. The input image is decoded one scan line at the time to a buffer to provide cache coherent memory usage. The input data files are:

- A 130 Kilobyte file with compression ratio of 18
- Two 900 Kilobyte files, compression ratios 4 and 10
- A 1.1 Megabyte file, compression ratio 12

The test is run continuously for 20 seconds. The result is expressed in millions of pixels (MPixels) per second.

Audio Compression and Decompression

The use of audio in PCs is growing beyond simply replaying one's favorite songs. Audio is becoming an important part of games, web sites and speech enabled applications. The process of recording music involves encoding audio into a music file. Similar to images, music files come in various formats. Ogg Vorbis and MP3 are some examples. Both of these are similar in that they are considered "lossy" – they discard data in order to compress better than would be otherwise possible. This makes them small and more amenable to transmission over networks.

For the Audio Compression and Decompression tasks, we use Ogg Vorbis libraries developed by <http://www.xiph.org/>. Ogg Vorbis uses a comparable approach to MP3. However, unlike MP3, it is a non-proprietary format and comes with free, open-source libraries.

Audio Compression and Decompression use floating-point operations and stress mostly the CPU. Ogg Vorbis libraries libogg 1.0 and libvorbis 1.0 are used. The Audio Compression test encodes a 3.5 Megabyte uncompressed WAV audio file into Ogg Vorbis format, and the Audio Decompression test decodes a 2.2 MB Ogg format audio file. The results are reported as Kilobytes processed per second.

Video Encoding

With digital video cameras gaining increasing popularity, video editing is becoming a popular task on the home PC. For this test, we concentrate on the process of compressing movies into formats more amenable to use on PCs. As with images and audio, various file formats exist for storing video. These formats similarly vary in how much they compress.

We have chosen Windows Media Video (WMV) as the target format for the video encoding task. The process for Video Encoding involves encoding video data from one format to a more compact format. Video compression primarily stresses the CPU with floating point operations.

For the Encoding task, we use Windows[®] Media encoder 9 which must be pre-installed on the PC. The test itself consists of three usage scenarios:

- 1) Encoding mpg format to WMV
 - Source video resolution is 1920*1080 and output is 1280*720 with 5Mbps CBR.
- 2) Encoding DV avi format to WMV
 - Source video resolution 720*480, output same resolution with 5Mbps CBR.

3) Encoding WMV HD format to WMV

- Source video resolution 1440*1080, output 640*480 with 2.13 Mbps CBR.

The three tasks are run after each other and the total frames per second processed is reported as the result.

Physics and 3D

With continual improvement in 3D graphics hardware and software, 3D games are becoming very sophisticated. Besides 3D graphics, games makers are employing complementary technologies like real-time physics, artificial intelligence, and 3D sound. We include this test in PCMark05 to represent 3D game usage on home PCs. Note that Futuremark also produces another benchmark, 3DMark series, which focuses solely on 3D graphics performance.

For this test, we use Microsoft® DirectX® 9 and Havok Physics engine 2.3. DirectX® is the dominant 3D API on PCs. Havok is a highly popular real-time physics engine, used e.g. in Valve's Half-Life® 2 and Remedy Entertainment's Max Payne. This test is a real-time rendering of several falling blocks. As they fall they collide with each other and their surroundings. Note that this is not an animation or playback of recorded events. All 3D graphics and physics computations are being performed in real-time – exactly as in 3D games.

For this test, the 3D graphics part stresses the 3D graphics subsystem with integer operations and the physics part stresses the CPU with floating point operations. For the 3D graphics workload we use multi-texturing and high-polygon models. The physics workload consists of rigid body calculations and collisions. There are 800 blocks and 4 light sources in the scene. Frame-based rendering – an approach that forces the same number of frames to be generated on all PCs – is used in this test. This ensures the same workload regardless of the hardware used. The result of this test is expressed in Frames rendered per second.

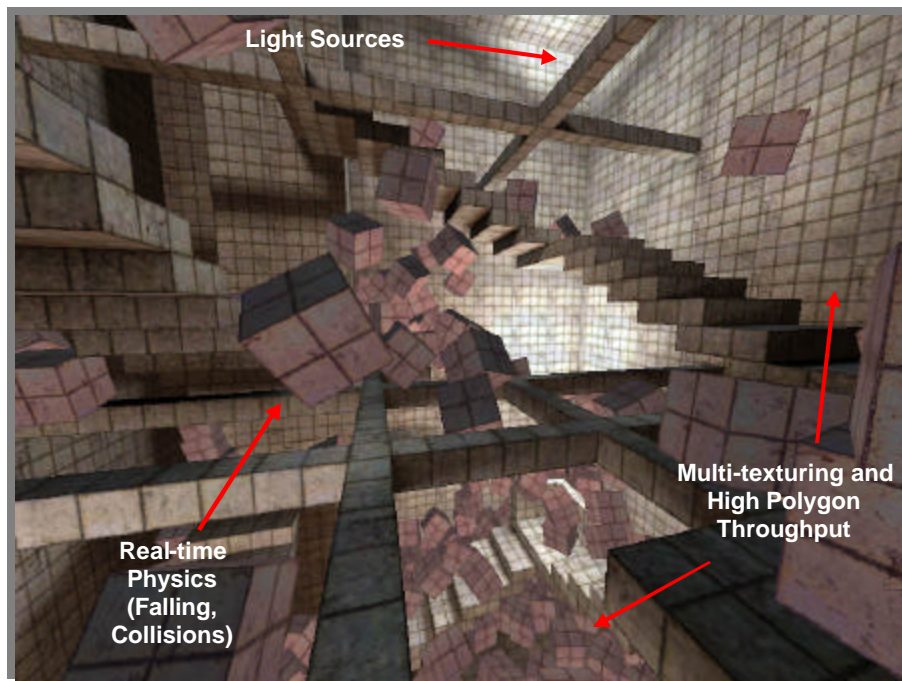


Figure 1: Physics and 3D Test

Graphics Memory

The Graphics Memory tests are designed to stress the video memory subsystem in same manner as typical Windows desktop applications. In home usage, common activity affecting the video memory is e.g. scrolling through documents.

The Graphics Memory tests stress CPU, memory, AGP graphics bus and graphics memory subsystem. These tests are written by Futuremark and do not include any public domain or commercial code. They use the Microsoft® DirectX® APIs. The graphics memory tests are also present in the Graphics Test Suite, although they are not pure graphics tests, but instead the CPU and memory performance have some effect on the result. This is, however very often the situation with real-life 2D operations; very rarely are 2D operations done solely on the graphics hardware.

The test creates a back-buffered primary surface with 1024 x 768 (32 bit) resolution in DirectX-exclusive mode. A second work surface is created off-screen that is twice as high. For each frame, the work surface is updated by transforming data through the AGP bus with a scrolling speed of 4/64/128 scan lines per second (three separate tests). The work surface is then copied every frame to the displayed primary surface to stress the internal memory bandwidth.

The test is run continuously for 20 seconds. The result is expressed in Frames per second.

Memory Tests

The read, write and copy tests in the default Memory Test Suite use data blocks of 16 Megabytes and 8 Megabytes for the main memory, 192 Kilobytes for the L2 cache and 4 Kilobytes for the L1 cache. The tests are run continuously for 10 seconds and Megabytes of data processed per second is given as the result. These memory tests are optimized differently depending on the platforms to report the best result.

The random access test uses a large list of data items. The size of each data item in the list is 64 bytes. The list is unsorted (or sorted using a pseudo-random key). Because of this setup, accessing each consecutive item in the list represents a random access operation. The list is iterated through as many times as possible in 5 seconds and Megabytes of items accessed per second is given as the result.

The latency tests are theoretical tests and are meant to isolate the performance of a particular part of the memory subsystem. They measure the time it takes to access data memory. In latency tests, a linked list is created and its order is then randomized. The test then walks through the randomly linked list, and the time it took is measured. In linear latency tests, the order for the list created is linear (sequential). Unit for the latency tests is MAccesses/s which means million accesses per second in the memory subsystem. The more usual unit ns can be derived from this by dividing one thousand by the value: $1000 / \text{value [MAccesses/s]} = \text{latency [ns]}$. The latency tests in the default Memory Test Suite use data blocks of 16 and 8 Megabytes for main memory, 192 Kilobytes for the L2 cache and 4 Kilobytes for the L1 cache.

2D - Transparent Windows

This test measures the performance of typical windowing operations of your PC. The test uses a workspace of 1024 x 768 with 32-bit color precision. 30 windows are created, each with its own moving and fading speed. Each of the windows is created fully visible and its alpha-blending value is first reduced and then increased to make a fading effect. Fading is a newer Windows® feature, which enhances the visual content of programs. Users can set their own preferences for applications for highlighting certain dialogs. The result for this test is the Average Number of windows drawn per second.

Video Playback

This test measures the performance of playing back a video file. The test uses WMV content.

The WMV file has a resolution of 1920 x 1080 and it is decoded through Windows Media Player 10.

The video is played at the highest rate possible, so the default frame rate of the video is ignored. The result is expressed in frames per second.

3D - Fill Rate

The fill rate is the speed at which your graphics hardware is capable of drawing textures onto 3D objects. These tests report results in million *texels* drawn per second (MTexels/s). Texels or texture elements are the pixels in

the source texture. We draw a number of large surfaces covering the entire screen and apply textures to these surfaces using fixed function pixel processing.

In the Fill Rate test, 8 screen size quads are mapped with 8 additively blended textures. The maximum amount of fixed function texture layers rendered in a single path is 8; otherwise, there could have been more textures per object. The texture size is 2x2.

3D - Polygon Throughput

Polygon throughput is a measure of the ability of your graphic hardware to process 3D graphics primitive objects. The tests report results in millions of triangles drawn per second (MTriangles/s). The scene is a real-time rendering containing a large number of polygons and no texturing. Eight directional light sources are used in this test.

3D - Pixel Shader

The Pixel Shader test is developed from the one in 3DMark05, showing the rock face shader. In the test, there are no real time shadows but only vertex lighting.

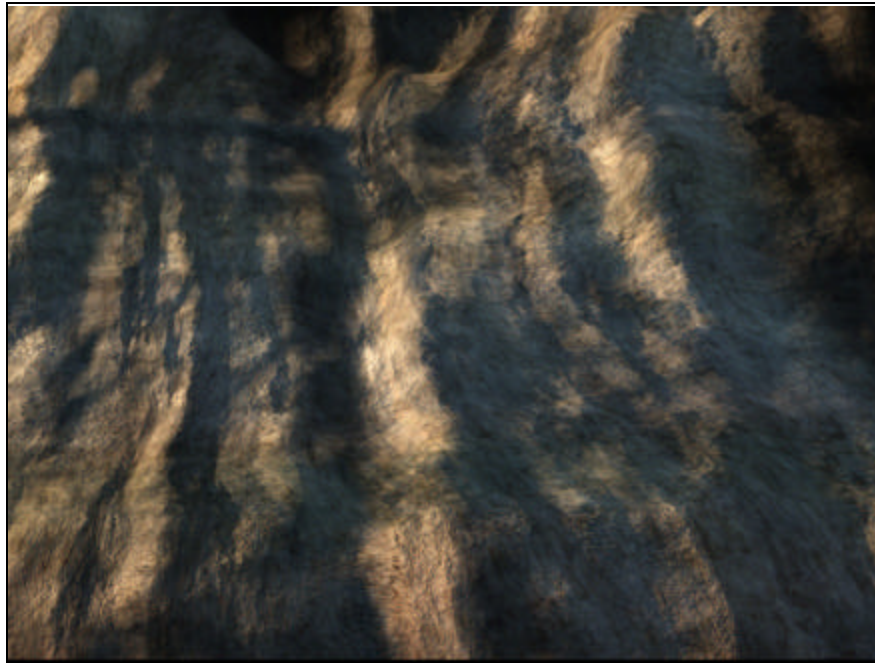


Figure 5: Pixel Shader Test

Filling the screen with a rock face is naturally fairly fast. This test is somewhat bandwidth dependent since any game-like material with a complex shader will also have a number of lookups to large textures. The result is reported as frames per second.

3D - Vertex Shader

The Vertex Shader test is developed from the one in 3DMark05. The test illuminates, and more importantly, transforms a large number of grass straws. Each straw is skinned and bent separately, more towards the tip of the straw, like real grass straws waving in the wind. The straws are waved according to a fractal noise calculated on the CPU, but it is highly optimized to decrease the influence of the CPU performance on the measurement. The grass is kept at a distance from the camera offering a less interesting visual effect. This is necessary to decrease the influence of fill rate to the measurement.

The result is reported as million vertices per second

HDD Tests

For these tests, we use RankDisk, an application developed and copyrighted by Intel®. In our testing, we found RankDisk to be suitable for a neutral benchmark. RankDisk is used to record a trace of disk activity during usage of typical applications. These traces can then be replayed to measure the performance of disk operations for that usage.

RankDisk records disk access events using the device drivers and bypasses the file system and the operating system's cache. This makes the measurement independent of the file system overhead or the current state of the operating system. In replaying traces, RankDisk always creates and operates on a new "dummy" file. This file is created in the same (or closest possible) physical location of the target hard disk. This allows the replaying of traces to be safe (does not destroy any existing files) and comparable across different systems. Due to the natural fragmentation of hard disks over time, they should be defragmented before running these tests.

The traces used for each test were created from real usage. The traces contain different amount of writing and reading on the disk; total ratio in the HDD test suite disk operations is 53% reads and 47% of writes.

The following input traces are used:

Windows XP Startup: This is the Windows® XP start trace, which contains disk activities occurring at operating system start-up. The test is 90% reading and 10% writes. This trace contains no user activity.

Application Loading: This is a trace containing disk activities from loading various applications. It includes opening and closing of the following applications:

- Microsoft® Word
- Adobe® Acrobat® Reader 5
- Windows® Media Player
- 3DMark® 2001SE
- Leadtek® Winfast® DVD
- Mozilla Internet Browser

The application loading trace is 83% reads and 17% writes.

General Hard Disk Drive Usage: This trace contains disk activities from using several common applications. These are:

- Opening a Microsoft® Word document, performing grammar check, saving and closing
- Compression and decompression using Winzip
- Encrypting and decrypting a file using PowerCrypt
- Scanning files for viruses using F-Secure® Antivirus™.
- Playing an MP3 file with Winamp
- Playing a WAV file with Winamp
- Playing a DivX video using DivX codec and Windows® Media Player
- Playing a WMV video file using Windows® Media Player
- Viewing pictures using Windows® Picture Viewer
- Browsing the internet using Microsoft® Internet Explorer
- Loading, playing and exiting a game using Ubisoft™ Tom Clancy's Ghost Recon

The General Usage trace is 60% reads and 40% writes.

Virus Scanning: Virus scanning is a critical task in today's PC usage. As the major bottleneck of scanning viruses is in hard disk activity, it is reasonable to include virus scanning as a HDD test. The test consists of HDD activity of scanning 600MB of files for viruses. The Virus Scanning test is mostly disk reading (99.5%).

File Write: This trace contains disk activities from writing 680MB files on the hard disk and no read operations are involved in this test.

Disk idle times have been compressed to 50 milliseconds to speed up the playback time. Our studies showed that 50 milliseconds was the smallest idle time interval that didn't affect the test results. The results of the HDD tests are reported in Megabytes processed per second.

Score Calculation

PCMark05 provides an overall PCMark score and four component scores. The PCMark score is a measure of the overall performance of the PC and is obtained by running the system test suite. Components scores are obtained by running the corresponding component test suites: CPU, Memory, Graphics, and Hard Disk Drive test suites. Users can also create custom test suites to meet their needs by selecting individually which tests to run. Note that no overall score will be presented for custom test suites. However, using the "Export to Microsoft[®] Excel" feature the user can implement this functionality as an Excel formula[‡].

PCMark Score

PCMark05 scores will initially range between 1,200 and 5,500 PCMarks with generally available PC hardware. They are scaled such that an entry-level system will score approximately 1,200 PCMarks and a high-end system, at the time of product release, will score approximately 5,500 PCMarks. Of course, we expect that over time there will be PCs that will score higher and higher.

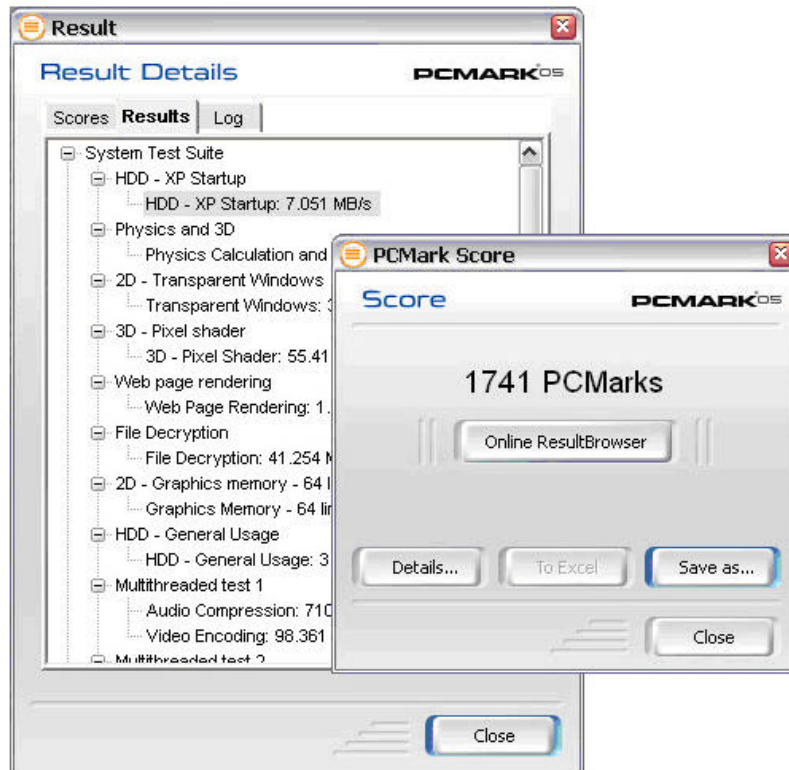


Figure 6: Example PCMark05 Scores

The PCMark score is calculated by using the results of the various tests in the systems test suite. The individual test scores are combined using a geometric mean. The geometric mean provides a fair mechanism to combine a large number of test results as compared to assigning arbitrary weights to individual scores. Geometric means are used in other successful benchmarks such as those from SPEC. The geometric mean is scaled using results

[‡] Note that such custom scores are not comparable to official PCMark scores.

from reference systems to produce the appropriate range of scores. The general formula for the geometric mean is: $\text{geometric mean} = (\text{item1} * \text{item2} * \text{item3} * \text{item4})^{1/\text{number of items}}$.

The formula for the overall PCMark score is:

$$\text{PCMark Score} = 87 \times [\text{geomean of the System test suite test results}]$$

CPU Score

The CPU score is also the geometric mean calculated from the CPU suite test results. The scaling factor is created in a same manner as for the PCMark score. The CPU score formula is:

$$\text{CPU Score} = 82 \times [\text{geomean of the CPU test suite test results}]$$

Memory Score

The memory score is also calculated from the memory suite test results using a geometric mean. The scaling factor is created in a same way as for the PCMark score. The Memory score formula is:

$$\text{Memory Score} = 1.7 \times [\text{geomean of the Memory test suite test results}]$$

Graphics Score

The PCMark05 Graphics score is calculated also using the geometric mean. The formula used is:

$$\text{Graphics score} = 17 \times [\text{geomean of the Graphics test suite test results}]$$

HDD Score

As in the other test suites, also the HDD Score is calculated using the geometric mean:

$$\text{HDD Score} = 300 * [\text{geomean of the HDD test suite test results}]$$

Example Scores

Table 3 below shows sample scores on three PC configurations. Two systems have the same graphics card, and two have the same CPU, other components are the same. It can be seen from the results that the CPU score scales quite well with the CPU speed, whereas PCMark score is less affected by a single component change. Then again changing the graphics adapter to a more powerful one shows improvement in PCMark score, but the CPU score naturally is not affected.

Table 3: Example of PCMark and CPU Scores Scaling

System	CPU	Internal Clock	Graphics Card
	Intel® Pentium® 4	3192 MHz	ATI® Radeon™ X800 XT
	Intel® Pentium® 4	3733 MHz	ATI® Radeon™ X800
	Intel® Pentium® 4	3733 MHz	ATI® Radeon™ X850
PCMark Score	4030	4490	4650
CPU Score	4090	4740	4740

Table 4 shows scaling of the PCMark score, Memory score and the HDD score using again three PC configurations. All the systems have the same graphics card and the CPU, two have the same HDD, and another two have the same memory speed. Note that the overall PCMark score increases both with the HDD performance as well as faster memory type. The memory score rises with faster memory and the HDD score rises with HDD performance.

Table 4: Example of PCMark and HDD Scores Scaling

System	CPU	AMD® Athlon™ 64 FX-53	AMD® Athlon™ FX-53	AMD® Athlon™ FX-53
Memory		1024 MB DDR 400MHz	1024 MB DDR 200MHz	1024 MB DDR 400MHz
HDD		PATA	S-ATA	S-ATA
PCMark Score		2380	3120	3210
Memory Score		4260	3150	4260
HDD Score		960	4870	4890

Online ResultBrowser

The utility of a benchmark result can be rather limited in isolation. Without a mechanism to compare your system to others, it is difficult to say if your system is a high-end state-of-art PC, a mid-end system, or even a low-end beige box. To allow benchmark users to come together to compare and analyze results, we provide a web service called the Online ResultBrowser or ORB.

The ORB has become Futuremark's most popular online service. It provides the users a web application to manage and compare their benchmark results. The ORB database contains over 12 million results. After running the benchmark, the user can choose to *upload* the results and system information to the ORB. We enforce data privacy, so no one except the user will be able to see the individual results. Futuremark also verifies the uploaded results for accuracy. The ORB helps the user by giving context to the benchmark data; it allows the user to compare the results with those from other PCs. The user is able evaluate the PC's relative performance - determine the PC's weaknesses and its strengths.

Users can decide to share their results by explicitly *publishing* them. This allows them to show their results to the rest of the world. For many users, their performance results are a point of pride. For some, their position in our top performing PC rankings is the source of a bit of fame. Many proud PC owners include ORB URL links to their published benchmark result in their email signatures.

The ORB is the user's virtual benchmark laboratory. Users can experiment with different system configurations such as increasing the RAM or upgrading the CPU. They can get the new PC performance by searching through results published by others. Users can also maintain their own performance track record by submitting multiple projects to the ORB over time. This is especially useful for keeping track of what effect new driver versions have on performance. They can assemble custom multi-compare sets to compare their PC to multiple other configurations.

As the ORB is an online service, Futuremark continues to improve it by adding new functionality. The benchmark data collected by Futuremark is used for generating statistics and recommendations for the user community. This means that every result submitted helps all users to select reliable hardware upgrades. The ORB aids in making buying decisions; before spending money, the user can validate expectations of different hardware options.

System Information

Associated with each benchmark result is the complete profile of the target PC; we call this the *System Information*. The system information not only provides detailed configuration information (CPU speed, RAM, graphics chipset, etc.) to go along with the performance data, but it also gives the system state (open applications, AGP mode, free system memory, etc.).

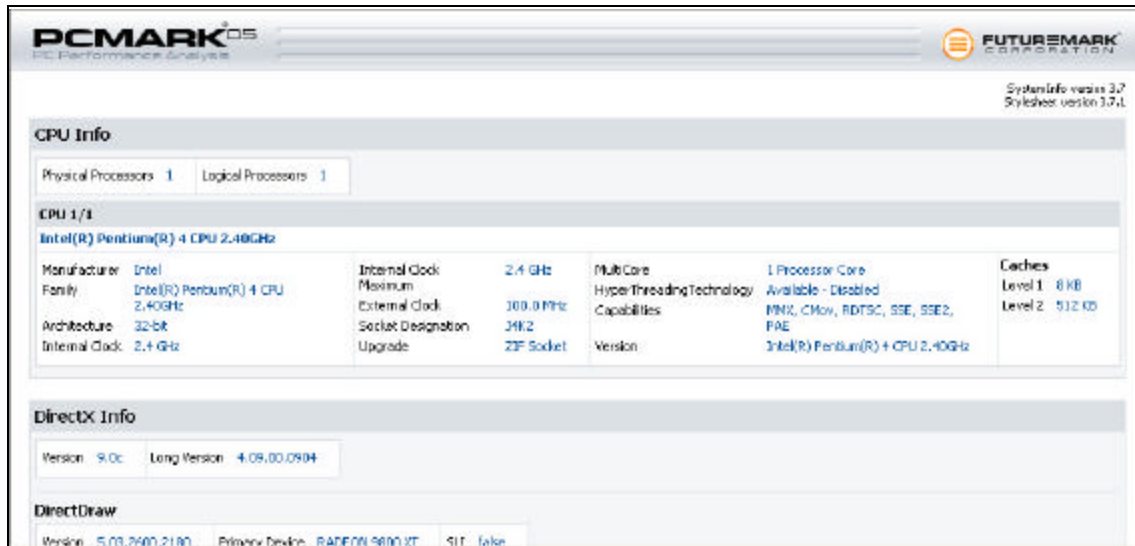


Figure 7: Example PCMark05 System Information

The complete system information consists of over 300 fields[§]. These include:

- **CPU** Information – clock speeds, internal and external caches, amount of physical and logical CPUs
- **DirectX[®]** Information – display drivers, direct draw/show attributes, texture formats, 3D capabilities
- **Memory** Information – memory arrays and modules
- **Motherboard** Information – bios, card slots, system devices
- **Monitor** Information – monitor attributes
- **Power Supply** Information – batteries
- **Operating System** Information – version, installed service packs
- **Open Processes** Information – applications, processes
- **Logical Drives** Information – local and network logical drives
- **Hard Disk** Information – disk drive attributes

The ORB uses the system information to enable the search and compare functionality. System information also allows Futuremark to verify the accuracy of published benchmark results. These increasing numbers of system information records are used to provide information back to the users in the form of lists of most popular and powerful hardware components shown on our web site. We also use system information data to build other tools such as the Futuremark Performance Analyzer – an online tool used for product comparisons.

[§] Note that no private information is ever collected.

Limitations

We have presented the features and benefits of PCMark05. We have detailed where PCMark05 is particularly useful. It is equally important to consider what PCMark05 is not. Below we list what we consider inappropriate uses and limitations of this benchmark tool.

- PCMark05 is not a tool to evaluate the performance of your PC in one particular application. For example, it is not appropriate for comparing the performance of only Microsoft® Word on two PCs. It measures the overall performance of your PC for home usage and performance of individual PC components.
- The system test suite represents average home usage. It may not correlate perfectly with other types of usage. For this, we provide the ability to run custom test suites and to export results to Excel to compose custom scores.
- As some test suites include multithreaded tests, scores will naturally be higher on multiprocessor/multicore and hyperthreading systems. So the scores may not be a measure of pure “number-crunching” performance of the PC. For this, one should use only the single threaded tests.
- Although component test suites isolate the performance of a particular component such as the CPU or memory, they may be nominally affected by the performance other subsystems.
- The graphics tests only measure the performance of outputting graphics and not the quality of images rendered. Futuremark’s graphics oriented benchmarks, like 3DMark05, include image quality tools that can help with this.

Conclusion

PCMark05 is an application based benchmark with both system and component level tests that has been designed to be easy to install and run. We have taken the application-based benchmark approach by including only the relevant parts of real applications. This has allowed PCMark05 to represent real home usage while keeping it small and quick to run. PCMark05 needs neither gigabytes of disk space nor days to run.

Futuremark cooperates with all major PC industry players, following a neutral and transparent development process, to ensure that the benchmark results are highly accurate.

PCMark05 is unique in enabling the complete cycle for benchmarking PCs – running the benchmark, analyzing the results, comparing results against other PCs, exploring PC improvements, leading finally to higher benchmark results and a more powerful PC.

As PCs and PC components continue to improve at a rapid pace, we believe PCMark05 will serve as a highly dependable tool for the benchmarking professional.

3DMark®, PCMark® and SPMark™ trademarks and logos, Futuremark® character names and distinctive likenesses, are the exclusive property of Futuremark Corporation. Microsoft, Windows, DirectX, Visual Studio are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries. Adobe, Photoshop, and Acrobat are either registered trademarks or trademarks of Adobe Systems Incorporated. BAPCO, SYSmark, and MobileMark are either registered trademarks or trademarks of Business Applications Performance Corporation. SiSoftware and Sandra are either registered trademarks or trademarks of SiSoftware Limited. F-Secure and Antivirus are either registered trademarks or trademarks of F-Secure Corporation. Symantec and Norton Antivirus are either registered trademarks or trademarks of Symantec Corporation. McAfee and Antivirus are either registered trademarks or trademarks of Network Associates, Inc. Macromedia and Shockwave are either registered trademarks or trademarks of Macromedia, Inc. Intel and VTune are either registered trademarks or trademarks of Intel Corporation. Electronic Arts and Madden NFL are either registered trademarks or trademarks of Electronic Arts, Inc. Harry Potter, characters and names are either registered trademarks or trademarks of Warner Brothers. Leadtek and WinFast are either registered trademarks or trademarks of Leadtek Research, Inc. AMD and Athlon are registered trademarks or trademarks of Advanced Micro Devices, Inc. in the United States and/or other countries. ATI and Radeon are registered trademarks or trademarks of ATI Technologies, Inc. in the United States and/or other countries. All other names mentioned are trademarks, or registered trademarks of their respective companies.

© All rights reserved Futuremark® Corporation, 2005.