



Author: Tero Sarkkinen
Executive Vice President of Sales and Marketing for Futuremark Corporation
Date: February 14th, 2003

Futuremark's Response to 3DMark[®]03 Discussion

3DMark[®]03 launch has been a great success and it has been received enthusiastically by the computer industry worldwide. There is active discussion about the product. We have analyzed the concerns that have been raised and found that many of those are without merit. In this document, we review the key facts of 3DMark03.

Synthetic Benchmark Versus Games as Benchmarks

3DMark03 is a synthetic benchmark that is designed for the sole purpose of enabling **objective performance measurement of DirectX[®] 9 compatible hardware today**. There are no other significant DirectX 9 applications published yet, and the most awaited DirectX 9 game is most likely at least six months away. 3DMark03 is a forward-looking tool that provides unique value to consumers in the form of impartial information to support their purchasing decisions today.

We continue to recommend to independent testers to complement their analysis by using published games in measuring performance. However, it has to be noted that those results are valid only for that game, whereas 3DMark03 can provide a forward-looking overall view of the performance and features of the hardware.

Additionally, benchmarks in games have not necessarily been created with the same diligence and attention to detail that Futuremark puts into ensuring the benchmark's independence and reliability. Games may contain specifically created code paths for different vendors' products; either for marketing or application compatibility reasons, invalidating at least the generalization of measured results to other games.

Well-built synthetic benchmarks measure computer hardware performance within a specific usage area in an impartial way. Only good synthetic benchmarks enable true apples to apples comparisons. 3DMark03 fulfills both criteria.

Our recommendations for correct benchmarking are the following:

- Use game benchmarks when you want to find out how fast a certain game runs on your computer;
- Use 3DMark2001 for a comparable overall performance measurement of DirectX 7 or first generation DirectX 8 compatible hardware;
- Use 3DMark03 for a comparable overall performance measurement of DirectX 9 compatible hardware

Driver Optimization

A concern has been expressed, that synthetic benchmarks force hardware manufacturers to optimize drivers for that specific benchmark.

Futuremark's recommendation has always been that default WHQL certified drivers should be used for benchmarking purposes. This is because any specific driver tuning might produce results that are not genuinely comparable. According to 3DMark03 license agreement, any review to be published **has to use** generally available, shipping versions of products and drivers.

Furthermore, 3DMark03 includes advanced analytical tools to enable independent observers to catch any potential questionable driver optimizations. By taking a tough stance against any kind of driver optimization, the media can discourage this practice.



Why Does Game Test 1 Not Use Multitexturing for All Visible Surfaces?

The use of single texturing in game test 1 has been criticized. It is claimed that the test does not measure DirectX 7 generation game performance.

Just like game developers, Futuremark develops its game test content so that the end result looks good and works optimally. Most games use single texturing for a skybox because more texturing does not necessarily make the end result look better. Examples of popular games that use single-textured skybox include Crimson Skies, IL-2 Sturmovik, and Star Trek: Bridge Commander.

As this issue was brought up already during 3DMark03 development, we did a test by adding a second texture layer to the skybox. The performance difference stayed within the error margin (3%), and in our opinion the additional layer did not significantly add to the visual quality of the test. Thus, there were no game development or technical justifications for implementing a multitextured skybox.

We would like to refer to our white paper about this test:

“This test is not meant as a definitive evaluation of DirectX 7. It is not designed to give the average performance of DirectX 7 3D graphics usage. Typical DirectX 7 games use fixed vertex processing, whereas this game test uses 1.1 vertex shaders. We believe this is the future of vertex processing on both graphics cards and CPUs. The overall goal of game test 1 is to complete the collection of the four game tests as a test that can run on DirectX 7 hardware and one that requires a lower fill-rate. To fully evaluate DirectX 7 performance, the previous version of the benchmark, 3DMark2001 SE, is more appropriate.”

It is important to note, that our Beta members suggested and supported the integration of a DirectX 7 compatible test in 3DMark03.

Why Do Game Tests 2 And 3 in 3DMark03 Only Use Pixel Shader 1.4 or 1.1?

According to the DirectX 8 specification, there are 4 different pixel shader models. In order to do a fair benchmark, you want any hardware to do the minimum number of shader passes necessary to render the desired scene. We analyzed all 4 shader models and found that for our tests Pixel Shader 1.2 and Pixel Shader 1.3 did not provide any additional capabilities or performance over Pixel Shader 1.1. Therefore we provided two code paths in order to allow for the broadest compatibility.

A good 3D benchmark must display the exact same output on each piece of hardware with the most efficient methods supported. If a given hardware supports pixel shader 1.4, like all DirectX 9 level hardware does, then that hardware will perform better in these tests, since it needs less rendering passes. Additionally, 1.4 shaders allow each texture to be read twice (total 4 texture lookups in 1.1, but 12 (=6*2) in 1.4 shaders). This is why, not only Futuremark, but also game developers can only implement single pass per light rendering using a 1.4 pixel shader, and not using a 1.3 or lower pixel shader. A 2.0 pixel shader would not have brought any advantages to these tests either. Note that the DirectX design requires that each new shader model is a superset of the prior shader models. Therefore all DirectX 9 hardware not only supports pixel shader 2.0, but also Pixel Shader 1.4, 1.3, 1.2, and 1.1.

Game Tests 2 And 3 Use Same Rendering Technique

The validity of the results is being questioned due to the fact that game tests 2 and 3 use the same rendering technique.

3DMark03 incorporates as many as three rendering techniques, whereas games use only one rendering technique. Our studies show that using three rendering techniques is more than adequate for producing valid results.



Why Is the Vertex Shader Skinning Used Instead of a CPU Skinning?

It has been alleged, that there is a design flaw in the code related to skinning the characters multiple times with the hardware vertex shader. It is an erroneous allegation and an invalid argument based on the following facts.

3DMark03 is designed to measure DirectX 9 compatible hardware. This level of hardware does skinning several times faster than the CPU. This can be confirmed using the vertex shader test of 3DMark03, which is designed to measure above all skinning speed. CPU vs. vertex shader skinning can easily be compared by running this test with and without software forced vertex shaders. For example, a DirectX 9 graphics card and high-end CPU (ATI Radeon 9700 Pro and Intel Pentium4 3 GHz) gets five times lower frame rates with CPU skinning than with hardware accelerated vertex skinning. An older CPU (Intel PentiumIII 800 MHz) skins more than 20 times slower on the CPU than with the hardware acceleration. Only first generation DirectX 8 hardware is likely to benefit from CPU skinning, since the vertex shader performance advantage of those cards, as compared to the CPU skinning speed, is smaller. Then again, first generation DirectX 8 hardware should rather be benchmarked using 3DMark2001 SE.

Since each light is performance-wise expensive, game developers have level designs optimized so that as few lights as possible are used concurrently on one character. Following this practice, 3DMark03 sometimes uses as many as two lights that reach a character concurrently, not five as mentioned in some instances.

Thus, vertex shader skinning will be more efficient than CPU skinning on well-programmed DirectX 9 games on DirectX 9 level of hardware. We believe that rational game developers will opt for vertex shader skinning in forthcoming DirectX 9 games.

Does this mean that 3DMark03 is now completely bottlenecked by the vertex shader performance? No it does not. Try running 3DMark03 in different resolutions. If the benchmark was vertex shader limited, you would get the same score on all runs, since the amount of vertex shader work remains the same despite the resolution change. Game tests 2 and 3 scale very well with the resolution, and are thereby mostly pixel shader limited. Changing the skinning to the CPU would reduce the vertex shader workload, making 3DMark even more bottlenecked by pixel shader performance. We did some simple experiments with this rendering technique and CPU vs. vertex shader skinning. The test results did not change much at all, the overall performance only dropped somewhat using CPU skinning. This was to be expected, looking at the difference in skinning speed between the CPU and the hardware vertex shader.

Why the Game Test 4 Does Not Use All Existing DirectX 9 Features?

The argument here is that game test 4 is not "DirectX 9 enough". Once again, a good application should draw a scene as efficiently as possible. In the case of game test 4 this means that some objects use Pixel Shaders 2.0, and some use 1.4 or 1.1 if a more complex shader is not required. Because each shader model is a superset of the prior shader models, this will be very efficient on all DirectX 9 hardware. In addition, the entire benchmark has been developed to be a full DirectX 9 benchmark:

- The whole test, as is the whole benchmark, is built using the DirectX 9 API. Most importantly, the benchmark is written directly onto DirectX 9, using only a light DirectX wrapper engine;
- The leaves, the sky and the water all use 2.0 vertex and pixel shaders, that are new main features of DirectX 9;
- The workload of the test is clearly designed for the performance that is currently only available in DirectX 9 generation hardware. On average 780,000 polygons are rendered per frame, and well over 100MB of graphics content is used per frame!

Once 3D games start using vertex and pixel shaders in the same magnitude that 3DMark03 does, there will be a clear correlation between 3DMark03 and game benchmark results. It is safe to say that it will



take quite some time until there is a game on the market that is “more DirectX 9” than game test 4 of 3DMark03. And by that time, we will already be developing a new 3DMark.

Conclusion

Consumers and the computer industry at large have a need to be able to make apples-to-apples comparisons to help them determine the benefits of new technology. 3DMark03 meets this need by being a sound and impartial benchmark for measuring Microsoft DirectX 9.0 compliant hardware.

3DMark03 is called “the Gamers’ Benchmark”, because it empowers gamers and consumers to objectively assess modern hardware performance.

Over 1.5 million copies of 3DMark03 have been downloaded within the first 72 hours of its release.