Archives
Columns
Features
Print Archives 1994-1998

Special
BYTE Digest
Michael Abrash's *Graphics Programming Black Book*
101 Perl Articles

About Us
How to Access BYTE.com
Write to BYTE.com
Advertise with BYTE.com

BYTEmark Frequently Asked Questions

BYTEmark Test Suite and Processor Performance

What is the BYTEmark test suite?

The BYTEmark benchmark test suite is used to determine how the processor, its caches and coprocessors influence overall system performance. Its measurements can indicate problems with the processor subsystem and (since the processor is a major influence on overall system performance) give us an idea of how well a given system will perform.

The BYTEmark test suite is especially valuable since it lets us directly compare computers with different processors and operating systems. The code used in BYTEmark tests simulates some of the real-world operations used by popular office and technical applications.

What do the different test scores in the suite mean?

There are 10 tests in the BYTEmark test suite. Each uses a well-known algorithm to analyze the full spectrum of processor performance in the same way real-world applications do. Some tests concentrate on integer performance; others test floating point capabilities. You can read more about the individual tests and the operations they measure in our BYTEmark documentation. It's available on our Website, at http://web.archive.org/web/20080619081129/http://www.byte.com/bmark/bdoc.htm.

Because processors don't always perform the same tasks in the same way, and thus may sometimes give unexpected results on individual tests within a benchmark suite. We've designed the BYTEmark tests to perform statistical checks to confirm the validity of our results, and we calculate separate composite scores for integer and floating point performance using a weighted geometric mean. This guarantees that no single test can overwhelmingly impact the overall index.

Volume 1: Programming Languages

In this issue of *Best of BYTE*, we bring together some of the leading programming language designers and implementors...

What is the advantage of benchmark tests test suite?

PC test centers generally measure performance in two ways: by performing controlled tasks with commercially-available applications, and with custom-built programs designed to test a particular facet of the computer system. Application-based benchmarks perform tasks inside commercially available programs. They're extremely useful for determining relative performance within a given platform, and less useful for ranking performance among machines that use different processor architectures.

The BYTEmark suite forces a system to perform carefully chosen operations natively and in isolation, that is, without assistance from an advanced operating system or other system resources such as a fast hard drive or enhanced video subsystem. Application benchmarks may show that a computer system is fast or slow, but the BYTEmark tests can help determine if the change in performance is due to extra memory or a faster processor, for example. In certain cases they can also be used to demonstrate obscure features (and sometimes flaws) in compilers or in microprocessor architectures. To see one example of this, check out our January, 1998 story on compiler differences, at http://www.byte.com/art/9801/sec12/art7.htm .

Remember, however, that a computer's overall system performance is a composite of many factors**:** the processor, the speed at which it and any L2 cache is clocked, the speed of its physical memory, the speed of its hard drive, the speed of its video subsystem and many other features. Vendors make trade-offs when assembling computer systems in order to achieve desired price points for new machines; these can cause two machines using the same microprocessor to report very different results. For that reason, BYTE Labs also tests computer systems using real application benchmarks as well as the BYTEmark suite. The combination gives us an accurate picture of overall system performance.

When was the BYTEmark test suite developed?

The original code was developed in 1990. Version 2 for PowerPC was released in October, 1995. Version 2 for Intel architecture was released in March, 1995.

Will you be updating the BYTEmark test suite again?

The BYTELabs staff continually examines new technologies and develops meaningful ways to measure them; the BYTEmark suite is only one of our tools. We do plan to expand the suite later this year; as beta versions become available we will announce them on our Website.

Have you "tweaked" results to favor either Pentium or PowerPC processors?

No. In fact, the BYTEmark executables we've used to test both have been constant and unchanged for two years. Despite repeated attempts by manufacturers to "break" them, they've produced remarkably consistent results from processor to processor.

Is the BYTEmark test used by any other organizations?

Yes. The BYTEmark suite's cross-platform compatibility has made it popular among computer manufacturers, developers and product analysts for performance testing. Online, you can find out more about what others think of the BYTEmark suite at these Websites:

- University of Auckland at Tamaki
- Apple Computer
- Linux nbench
- MacSpeedZone
- This standard benchmark proposal includes BYTEmark as a standard element

Where can I obtain a copy of the BYTEmark test suite?

We've made source code and binary files for popular platforms available on our Website, at http://web.archive.org/web/20080619081129/http://www.byte.com/bmark/download.htm. Anyone is free to download and run our BYTEmark tests, or to compile the source code themselves.

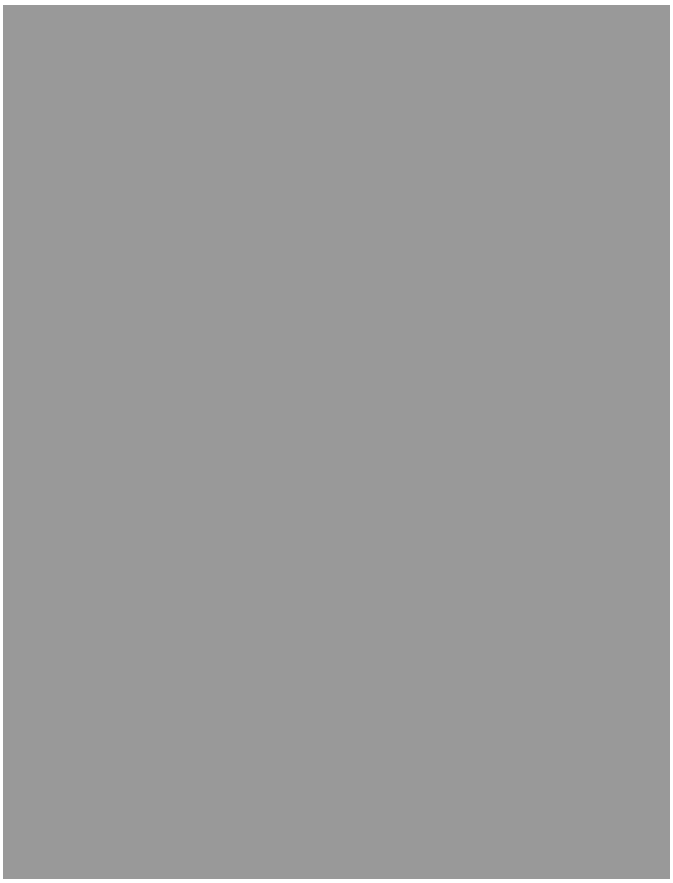
Freely distributing benchmark code allows us to remain completely open about the objectives and capabilities of the BYTEmark suite. We invite comments and suggestions for future upgrades.

So, can your tests show that a Power Mac computer is faster than a Pentium II computer?

Recent Apple Computer ads claim that the processor in a 266MHz PowerPC 750 (known as a G3) machine performs up to twice as fast as a 300MHz Pentium II. Apple based its claims on BYTEmark test results, specifically on the integer tests, which return extremely high performance scores.

On certain tests, the G3 *processor subsystem* returns performance scores that are about twice as fast a Pentium II *processor subsystem*. On other tests, including some of our application benchmarks, the Pentium II edges out the Power Mac. The BYTEmark integer index shows that the G3 is faster than the Pentium II. Even though the Pentium II operates at a faster clock rate, 300 MHz, compared to the G3 at 250 MHz, the PowerPC remains about 60 percent faster on these particular tests. For more details on comparisons of Pentium II to a 266MHz G3, see our January, 1998 issue http://www.byte.com/art/9801/sec3/art8.htm)

Here's a sample of a 300MHZ PowerMac computer tested with the BYTEmarks, versus a 300MHz Pentium II. We invite you to run these tests on your own machines for verification.



Why do the Bitfield test results vary so much between the two processors?

If you'll examine the chart carefully you'll notice that one test, Bitfield, returns very high scores for the G3 when compared to the Pentium II. The reason? The PowerPC compilers used to build these applications can generate code that's significantly different from that of x86 (i.e., Pentium) compilers.

What's the difference between these compilers?

A lot of it has to do with how the compilers generate code for this test. Compilers generate code based on their knowledge of the target processor's architecture. They use this information to produce machine code that implements the algorithm most efficiently.

Do compiler optimizations impact results?

Yes, because they take advantage of the processor architecture. This is exactly what a software developer wants; it helps speed the performance of the application he's creating.

What is the purpose of the Bitfield Test?

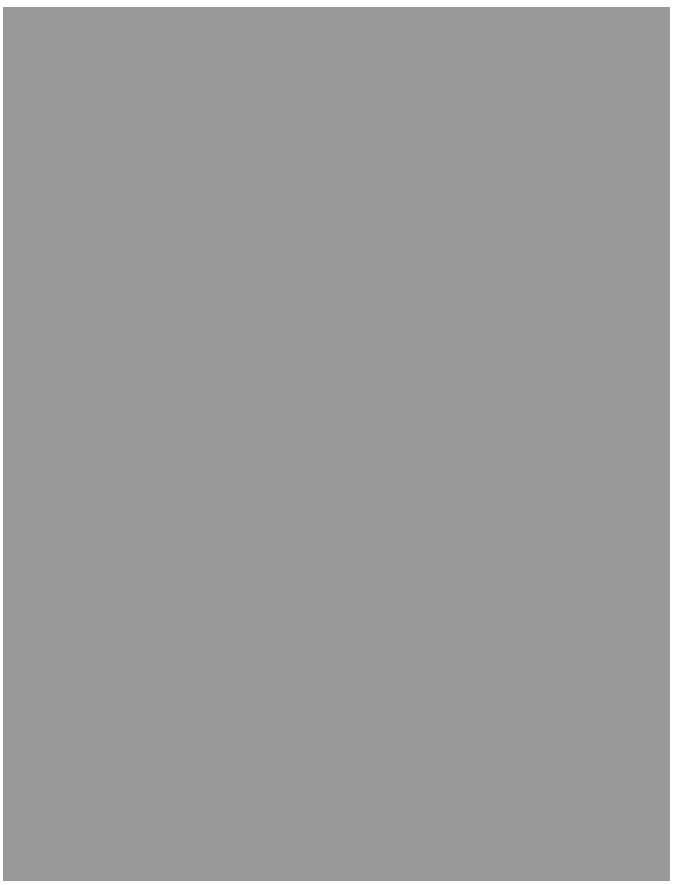
You'll find more complete information on the Bitfield Test on our Website, at http://web.archive.org/web/20080619081129/http://www.byte.com/bmark/bdoc.htm. But briefly, the Bitfield test simulates a technique, disk usage bitmapping, which computer systems often use to speed data storage and retrieval.

An operating system that uses such a bitmap will set a portion of its memory aside (called a "buffer") to act as a kind of inventory of available disk space. Each bit in the buffer represents a block of disk storage space. As information is written to the disk, the operating system sets its corresponding bitmap value to 1, to signal that block is full. Conversely, erased data causes the corresponding bit to be reset to 0. When a file requests storage space, the system searches through the bitmap for empty (0) blocks and returns the number of that block to the application.

A computer's performance in manipulating bitmaps helps predict how quickly it can update stored information. That's an important part of processor performance.

Is it possible, then, to exclude the Bitfield test and look only at other results?

Yes. The BYTEmark was designed so that any one test or group of tests can be run. But even if we exclude the Bitfield test entirely, our results still show the 300MHz PowerPC processor to be about twice as fast in overall integer performance as the 300MHz Pentium II, as seen on the chart below.



Note, however, that the Pentium II's overall score also falls when the Bitfield test results are excluded.

So, are today's G3 Power Macs faster than Pentium II PCs?

In general, yes. Exactly how much faster depends on the applications you're running and how you compare processors. Keep in mind that the BYTEmark suite measures processor subsystem performance only. The overall performance of an entire computer system depends on factors ranging from the speed of the hard drives and memory to the skill of developers in exploiting a particular operating system and processor combination.

Our Adobe Photoshop benchmark is a perfect example of this. Much of Photoshop consists of integer operations and the PowerMac does return better scores in some cases. The Pentium II's MMX capabilities, however, make it the top performer in Photoshop operations that take advantage of special Pentium MMX instruction sets, such as Unsharp Mask and Gaussian Blur, that the PowerPC processor doesn't have. (see chart above).

Questions? Comments? Drop us a line: editors@byte.com .