



# Real World Benchmarks 2005

This month, we introduce a brand-new set of benchmarks, designed to push both single and dual-core processors to the limit

In July 2002 (see issue 93, p56), we introduced the last version of PC Pro's custom-developed benchmarks. Back then, we used a cutting-edge 2GHz Pentium 4 Dell Dimension PC, setting its result as a baseline of 1.00. We've seen scores creep ever higher – topped by Gladiator's overclocked Athlon 64 FX-57 system, the Titan Extreme PC (see issue 131, p52), scoring a colossal 3.95. In other words, it was almost four times as fast as the Dell.

Using industry-standard applications, we run scripts to give each program a set amount of work and measure, to the nearest one-hundredth of a second, how long it takes to complete the task. To show what a system is capable of, we disable any intermittent background tasks and then run through the full set of tests six times, rebooting between each test to prevent any caching.

Not only does this reflect actual performance as meaningfully as possible, but it's done to a level of plus or minus 0.5 per cent accuracy, putting other



We use industry-standard applications to measure real-world performance. Each test is run six times to ensure accuracy.

benchmarks to shame. And unlike the benchmarks you'll see used elsewhere, they're completely independent – the development costs aren't met by any hardware or software manufacturers, allowing us to get to the core of a system's true performance without any external influence or bias.

We've divided the tests into four categories to reflect real-world usage:

Office, 2D Graphics, Encoding and Multitasking. We record the time taken to complete each test and, using our Pentium D test rig (see graphs, opposite) as a baseline, calculate a relative score. So, a machine scoring 1.50 is 50 per cent faster than our reference machine. The overall benchmark score is then simply calculated as an average of each category.

## DUAL CORE

Last month, at a Microsoft research summit, Bill Gates announced that he expected processors consisting of 32 cores, each running at 5GHz, within the next five years. And as unlikely as that may sound, it may just happen – multicore processing is undoubtedly the future.

Nonetheless, single-core processors will be with us for some time, particularly at the budget and mobile ends of the market. And while Windows XP will divide up the work as far as possible, it's going to be several years before the majority of applications really take advantage of more than one processor.

As a result, not every one of our benchmark tests will benefit from extra processing cores – so despite an Athlon 64 X2 4800+ containing effectively two Athlon 64 4000+ cores, it won't score double. The notable exception is 3ds Max, which shows upwards of a 95 per cent speed increase.

## INTRODUCING THE TESTS

**OFFICE** Our Office test actually consists of several Office 2003 applications: Word, Excel, Access and PowerPoint. In real life, you wouldn't be working in Word and Excel simultaneously, so we don't ask the PCs to run our Word and Excel scripts simultaneously either. However, we do keep all the applications running in memory at the same time.

We start by generating a 10,000-record Access database, simulating a massive collection of contacts and including a number of linked tables. We then make an Excel workbook that extracts data from our Access database, then performs a

## Intel vs AMD

The principle choice for many when considering a new PC is still which of the major processor manufacturers' offerings should form the basis of their machine. In the desktop space, it's clear that in the battle between the Pentium 4 and Athlon 64, it's AMD's chip that delivers better performance per clock cycle – an Athlon 64 4000+ running at 2.4GHz scores 0.99, against 1.05 for a 3.8GHz Pentium 670.

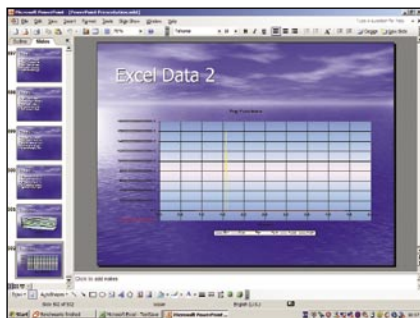
The integrated memory controller of the Athlon 64 provides a more efficient model for any tasks that require large amounts of data to be shuffled between RAM and the CPU – evident in tasks such as photo editing and 3D rendering. The Pentium 4, however, uses the south bridge as a communications hub for all data passing back and forth, potentially inducing a delay as data queues up to pass through.

But Intel does manage to scrape back some of the edge in tasks such as video and audio encoding. This is due to the

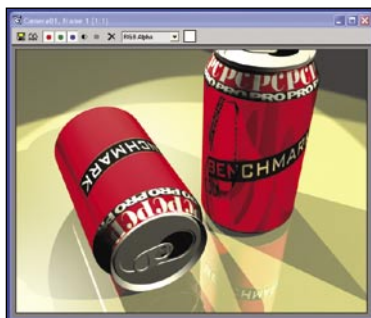
longer pipeline and more complex branch prediction of the Pentium 4. This pre-empts and prioritises calculations, helping it power more quickly through repetitive or predictable tasks.

These principles also hold for Intel and AMD's dual-core processors, the Pentium D and Athlon 64 X2 respectively. As there's no fundamental change to the architecture, just the implementation of two cores rather than one, we see a similar pattern of performance.

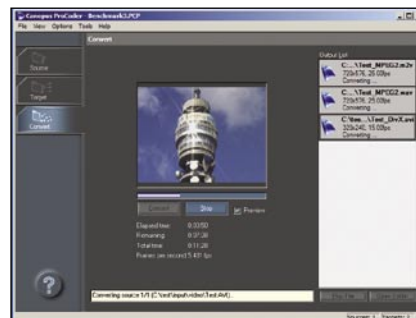
It's a slightly different story for mobile processors. The Pentium M is a remarkably nippy chip given its power consumption and clock speed, but it also has a comparably short processing pipeline, which restricts raw computational performance. AMD's Turion, based on the same architecture as the Athlon 64, shows similar advantages to its desktop cousin, comfortably handling more complex data-manipulation tasks.



The office test uses Microsoft Office 2003, with a number of tasks in Access, Word and Excel.



A complex scene in 3ds Max 7 measures a machine's computational performance.



Encoding two streams of video simultaneously makes use of Canopus ProCoder's multithreading.

number of intensive financial and data calculations. We create a number of graphs, apply automatic formatting and place objects into the charts.

Next, we open a large **Word** document – around 180 pages – and apply formatting with styles. We copy some of the graphs from Excel and paste them into the Word document, repaginate it all, run a spellcheck and then print it to file.

Lastly, we create a **PowerPoint** presentation using some of the data from the Access database to populate the slides, change the slide masters and add graphs from Excel. All the documents are saved and the applications closed. It adds up to an excellent guide to a system's overall performance, testing all areas from the speed of accessing data on the hard disk to intensive number crunching.

**2D GRAPHICS** We start with a complex drawing comprising over 4,000 vector objects in **CorelDRAW 12**. Over this, we lay a mass of text lines, each set to have graduated transparency. The fonts and text are repeatedly changed, followed by a forced redraw of the whole document. Finally, we import a large bitmap, again set to partial transparency, and redraw the whole document. This test is so difficult it scares the engineers at Corel: even on fast machines, you see the individual letters being redrawn one by one as they clip and colour the complex objects beneath them.

Using the CS 2 version of **Adobe Photoshop**, we open 30 digital photographs simultaneously, each with a resolution of 1,704 x 2,272 (just over 3.5 megapixels). We then apply auto-levelling, sharpening and colour correction to each in turn, before saving each file out in JPEG, TIFF, GIF and PNG formats. Next, using six high-resolution digital photographs we manually create a large panoramic image – a highly intensive computational task.

Then, using the Picture Package command, we create a typical full-page multilayered poster. Various elements are edited and text layers added. These are then linked and merged before applying several different blend modes in turn,

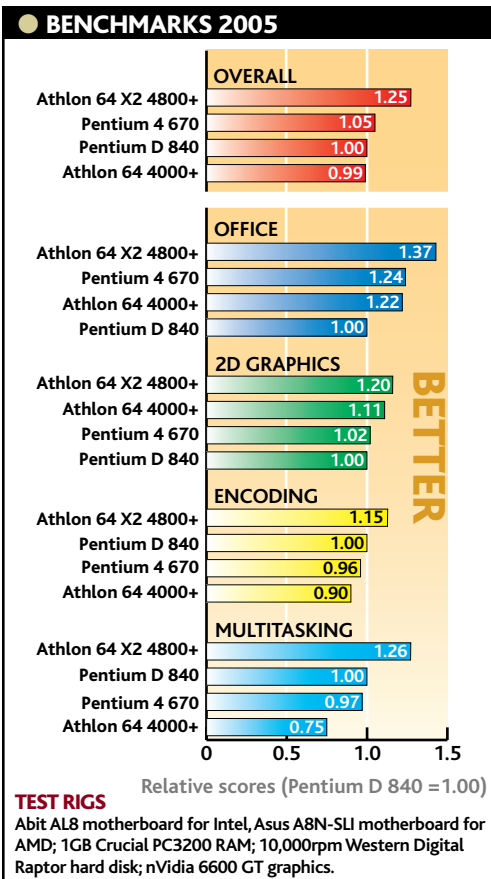
with each result being saved out as a CS Layer composition. The finished version is saved as a PSD format file. We then use the Web Photo Gallery command to create a thumbnail HTML page with 50 high-resolution images. These tasks place particular stress on the hard disk and memory-bandwidth performance.

In **Discreet 3ds Max 7**, we render three frames of a complex animation sequence. Each frame contains several hundred thousand polygons, rendered using photorealistic raytracing and radiosity lighting. Due to the complex maths involved, this is almost exclusively a raw CPU horsepower test and makes use of all available cores, and Hyper-Threading if available.

**ENCODING** Illustrate's **dbPowerAmp** is a popular shareware tool for format conversion, and we use it to encode ten WAV files to MP3, and then to WMA and OGG formats. Media encoding is one of the most computationally intensive tasks that most people will perform on their PCs, and one of the only areas where the Pentium 4 consistently outperforms Athlon 64 processors (see *Intel vs AMD, opposite*).

**Canopus ProCoder** has set the standard for pro-grade video encoding, and is now widely used by video professionals at every level. With support for multithreading, it can take advantage of the latest dual-core processors. Our encoding test loads a 541MB DV AVI file (which is two minutes, 28 seconds long), then compresses it simultaneously to DVD-quality MPEG2 and DivX 5.1.1 at 500Kb/sec with a 352 x 288 resolution. This simulates two real-world video-compression scenarios, which are often performed at once to save time.

**MULTITASKING** In the real world, it's rare that people have just one program running. The typical office Desktop will likely have various Office programs open, as well as several web pages and a virus scanner or firewall running in the background. To really test the mettle of a system, the final test runs



three of our punishing tests at once: the Office, dbPowerAmp and Photoshop tests. As this eeks out every last drop of performance from a machine, it's useful for highlighting any weak links in the component chain, as well as testing stability – overheating can be a real problem here. If a PC can handle the pressure of this test, it can handle almost anything.

## Can I have a copy?

Apologies, but no. Benchmarking a system to any level of accuracy means you have to disable all its security features, and we don't wish to encourage anyone to do this and then suffer a major data loss as a result. Also, you'd need to own precisely the right versions of all the software we use to gain meaningful results.