JOURNAL

INSIDE THIS ISSUE:

FROM THE PRESIDENT

EEMBC Celebrates Launch of Digital Entertainment Benchmarks and Scores

DENmark™

It's been a long time coming . . . EEMBC has released the official version of its new Digital Entertainment benchmark suite (DENbench™), and we also have four influential members that have certified and published scores already on these new benchmarks. Before digging into the details of these benchmark scores, corny as it may sound, I am very proud to be able to offer this suite to EEMBC's membership. It is, by far, the most comprehensive, sophisticated, and challenging benchmark suite we have offered to date. The suite's more than 60 discrete tests include MPEG, MP3, JPEG, and cryptography. And as an added bonus, our cryptography benchmarks have already secured a place in EDN's top 100 technology products of 2004!

These benchmarks offer an innovative departure from previous EEMBC practice with the use of multiple datasets. For example, the MPEG-4 encode benchmark has five unique video streams that get fed in. Each video stream represents a different bit-rate, screen size, and picture complexity, allowing us to demonstrate how a digital entertainment system handles different workloads. In other words, one of the video streams will be applicable to the small screen on a mobile phone, while another video stream is representative of the quality you'd find on a high-end set-top box.

Another interesting twist that we've applied to these benchmarks (specifically, the MP3 player and video encode/decode benchmarks) is the use of a quality measure. This quality measure is based on a signal to noise ratio (PSNR) to provide "an estimate of the quality of a reconstructed image compared with an original image" (a summary of PSNR is available here). PSNR is a fairly common industry practice that ECL integrated into our Test Harness to provide a uniform approach that all of EEMBC's members could adhere to.

(continued on page 4)

Benchmark Scores
✓ Analysis
✓ Comparison
✓ Performance

for the latest in benchmark scores
www.eembc.org

NEWEST BENCHMARK SCORE REPORTS

TMS320C64x 1 GHz
Production Silicon
Out-of-the-Box
Telecom

C Optimized
Telecom

Assembly Optimized
Telecom

IGNITE 2FX 600 MHz
Simulation
Out-of-the-Box
Consumer

TC1130 150 MHz
Production Silicon
Out-of-the-Box
Automotive/Industrial

Visit the new DENbench web page at www.eembc.hotdesk.com/digital_entertainment.html

Processor Name-Clock	ADSP-BF533 594 MHz	AMD Geode NX1500@6W - 1 GHz	MPC7447A 1.4 GHz	IBM 750GX 1 GHz
MPEG Decodemark™	N/A	785.1	1506.3	1054
MPEG Encodemark™	355.5	587.4	1281.5	967.9
Cryptomark™	N/A	509.3	1263.3	903
Imagemark™	N/A	918.9	1709.4	1090.4
MPEG2 Encode™ (Floating Point)	N/A	30.2	67.2	N/A
DENmark™	N/A	131.7	257.6	173.6

From the lab

The Value of Reference Platforms for Comparison Purposes

Alan R. Weiss, EEMBC Certification Laboratory (ECL, LLC)

Benchmarking is inherently a comparative, and competitive, activity. While absolute scales are fairly common in science (for example, the Kelvin scale for temperature), business and engineering people like to measure directly against others. So when EEMBC develops a new benchmark suite, until there are at least a few benchmark scores available, it is difficult to know *a priori* if a score of, say, 25.5 DENmarks is good — or very low. How does this compare against the fastest processors on the market? How does this compare to other processors in its class (price, power consumption, temperature range)?

Engineers may want to analyze each and every benchmark score in a given benchmark suite, but marketeers seem to want a single figure of merit. EEMBC uses *iterations per second* as a figure of merit for many of its performance benchmarks, but these must be translated into a single number, in a range easily understood. A score of 3,238,560 on one benchmark kernel, followed by a score of 594 on another kernel in the same suite, results in some funny arithmetic when you try to average them together.

Reference platforms can solve this problem by offering a measuring stick against which one can gauge another processor, and in helping to calibrate the single-number consolidated score(s) (such as the Telemark, Consumermark, and OAmark) that EEMBC reports for each of its benchmark suites. But it is no simple matter to decide which

(continued on page 4)

EEMBC Calendar

February 22–24, Nuremberg

EEMBC exhibits at Embedded World 2005 (Nuremberg, Germany) within Green Hills Software stand 10-321. For show information, visit www.embedded-world-2005.de.

February 23, Nuremberg

Markus Levy presents "Analyzing the Tradeoffs Between Performance and Power When Designing Embedded Applications" at 4:30 p.m. as part of the Embedded World 2005 conference in Nuremberg. For conference details, visit www.elektroniknet.de/termine/ew2005/index.htm.

March 20-22, Austin

At the 2005 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Markus Levy will deliver the keynote address "EEMBC and the Purposes of Embedded Processor Benchmarking." For location details and further information, visit www.ispass.org/ispass2005.

April 18, San Jose

Markus Levy kicks off the plenary session of the Portable Power Developers Developer's Conference at 10 a.m. with "Analyzing Power Efficiency in Portable Applications." Further information is available at www.darnell.com/ppdc.

Visit the EEMBC "In the News" page to read the latest in editorial coverage.

Automotive real-time benchmarks, by Markus Levy. *Embedded Computing Design,* December 2004.

EEMBC Adds a Metric, by Don Tuite. *Electronic Design - Powerful Ideas for Design*. November 29, 2004.

Microprocessor consortium to add energy spec. *EE Times*, Novemnber 9, 2004.

Editorial: Can benchmarking be rational?, by Kevin Krewell. Microprocessor Watch, September 24, 2004.

Networking benchmarks score Freescale and IBM chips, by Graham Prophet. EDN Europe, September 2, 2004.

NEWS BRIEFS

Tokyo-based **IPFlex Inc.** has joined EEMBC as a full member of the Consortium's Board of Directors, with full voting rights on all major issues decided by the consortium, plus full access to all of the EEMBC benchmarks. IPFlex's flagship product, the DAPDNA-2, is a dynamically reconfigurable processor with the ability to switch functions by dynamically and instantly changing the circuit configuration within the chip. The processor is for commercial use, suitable for the areas of telecommunications and security where mass capacity high-speed processing is required, and for industrial and medical image data processing. www.ipflex.com.

EEMBC's Cryptography benchmark, part of the new Digital Entertainment benchmark suite, was among the **Hot 100 Products of 2004** announced in the December 17 issue of *EDN*. "Although how innovative or 'hot' a product is never ensures market success, these products are the 100 we think are the most promising toward that end among the thousands that vendors introduced in 2004," said John Dodge, EDN editor in chief. The award was made in the Embedded Tools category. www.edn.com/article/CA486570.html

EEMBC mourns the loss of **Kaivalya Dixit**, long-time president of Standard Performance Evaluation Corp. (SPEC), who died November 22 at age 62. Dixit had headed SPEC since 1990. A 30-year veteran of the computer industry, Dixit worked most recently at IBM in Austin, Texas. He also held engineering, marketing and management positions with Sun Microsystems and Ford Aerospace. Dixit is survived by his wife Evelyn of Austin, and son Raj of Bakersfield, Calif. Memorial contributions can be made in memory of Kaivalya Dixit to the American Heart Association or the American Diabetes Association. www.spec.org/spec/kaivalya/contributions.html

Members Only

Following a series of briefings with analysts and editors in the U.S., Europe, and Asia, EEMBC announced publication of the first scores based on its **DENbench Digital Entertainment benchmarks** on January 24 (see scores on page 2). "EEMBC's release of these new benchmarks and the publication of these scores from AMD, ADI, Freescale, and IBM is very timely given the rapid growth of the digital entertainment market," said industry analyst **Will Strauss**, founder and president of Forward Concepts. "These benchmarks provide a comprehensive interpretation of a processor's system performance running a wide variety of digital entertainment code, and they have been constructed to allow apples-to-apples comparisons among a wide range of processors."

At its meeting in Menlo Park, Calif. on December 7, the EEMBC Board of Directors voted to create a **Multiprocessing Subcommittee** to address benchmarking of multiprocessing/multithreading (MP/MT) devices. The group, headed by John Goodacre of ARM Ltd. and Mike Uhler of MIPS Technologies, discussed the approach that would be taken in developing the new suite, which will initially focus solely on symmetric multiprocessing (SMP) implementations, where the software assumes homogeneity across processing elements, threading is the technique to express concurrency, and an OS or RTOS is available to abstract the hardware into an agreed-upon software threading model. Consumer, mobile, and PDAs were mentioned among the applications into which the devices proposed for benchmarking are being deployed.

The **next EEMBC Board of Directors meeting** is tentatively scheduled for March 23-24 in Austin, Texas. Location details will be announced in February.

The Value of Reference Platforms for Comparison Purposes

(Continued from Page 2)

processor, on which board, and at what clock speed, should serve as the reference platform for a given benchmark suite. One issue is finding a reference platform that can serve as a typical example of the expected performance range. Another issue is the availability of a volunteer member company to offer its platform as a reference platform.

Fortunately, EEMBC has had some volunteers.

For the GrinderBench Java benchmark suite, a Sharp Zaurus running Embedix 1.0 on top of an Intel XScale 206-MHz processor was used to help calibrate the GrinderBench score. This was appropriate, because PocketPCs, PalmOS devices, and modern mobile phones were expected to be within a range 10x slower to 10x faster than this platform.

For the second generation of Consumer benchmarks (called Digital Entertainment, or DENbench), Advanced Micro Devices allowed the EEMBC Certification Laboratory (ECL) to use its GeodeNX1500@6W (1 GHz) as a baseline platform and to share these scores with EEMBC members. The Geode is a midrange and highly efficienct processor targeted for set-top boxes, Internet appliances, and thin clients, and it is commonly available in the market. By not using a very high-end PC processor such as AMD Athlon-64 or an Intel Pentium 4, a message was sent that the purpose of this exercise was in calibrating the benchmarks for a specific purpose — not in blowing away the field.

ECL had little trouble producing the scores: when it develops the benchmark code for EEMBC, it regulary uses x86 PC's as a test bed to develop and test the code (using various C compilers), along with a half-dozen other platforms. Building the code was thus a matter of simply doing what every EEMBC member can do: downloading the code from the EEMBC website, unzipping it, making sure a C compiler was installed, and typing `make`. The new makefile system developed by ECL for EEMBC did the rest, including emitting comma-separated files easily imported into Open Office or Microsoft Excel workbooks.

Then came the hard work. Consumer Subcommittee members, led by Sergei Larin of Freescale Semiconductor, worked to develop the arithmetic necessary to calibrate the DENmark scores. Many experiments were tried to ensure that DENbench would be as applicable for lower-end 16-bit fixed-point processors as it would be for high-performance 32-bit and 64-bit processors.

Networking Version 2 (IPmark and TCPmark), as well as the soon-to-be-released Office Automation Version 2 (featuring Ghostscript), will use a similar technique (although reference platforms haven't been selected yet for those benchmark suites). For 2005, EEMBC will continue to use the concept of reference platforms as ECL develops Automotive/Industrial Real Time Version 2 and other benchmark suites.

And what did AMD get for its troubles? A free certification and publication of scores, and the knowledge that they have contributed to the development of an extremely important new benchmark suite.

FROM THE PRESIDENT

(continued from page 1)

Okay, now who are those bold and progressive members who have published the first set of scores? First is AMD with its 1-GHz AMD Geode NX1500. Next is Analog Devices with its 600-MHz ADSP-BF533, based on the Blackfin architecture. There's also the PowerPC-based 1.4-GHz MPC7447A from Freescale. And last but not least is IBM with its 1-GHz IBM 750GX PowerPC machine. These processors represent a wide range of performance, power, and price characteristics, and therefore help to validate the effectiveness of DENbench.

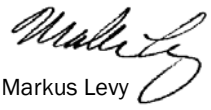
For each processor measured with DENbench, EEMBC can report individual results for up to 69 benchmark algorithms and associated datasets, as well as a series of consolidated scores that provide a snapshot of performance in specific test groups. An overall DENmark™ score provides a single-number performance rating for the entire DENbench suite.

As we've seen with EEMBC's other benchmarks, system designers will interpret the DENbench scores depending on how they prioritize design considerations such as speed, efficiency, power consumption, die size, and price. The various consolidated scores that the benchmarks produce become most useful, in fact, when used to calculate how much a processor's performance 'costs' in terms of several different metrics.

Let's look at a few examples of how the different metrics can apply. Referring to the benchmark scores from the processors listed above (available at www.eembc.org), the MPC7447A is the raw performance leader. Running at 1.4 GHz, this processor cranks through video frames faster than a sous chef chops vegetables. And this doesn't account for the additional benefits that can be provided through the processor's AltiVec engine. This level of performance makes the MPC7447A perfect for high-end encoder applications (i.e. this is the machine that would encode the video streams before they are blasted out over the Internet).

Now if we break it down with an execution efficiency rating (i.e. performance/MHz), the Freescale device still does quite well (due to its beefy microarchitecture), although the IBM device surpasses it on MPEG encode algorithms. On the other hand, look at it from the point of view of performance per Watt, and the Analog Devices and AMD processors really start to shine. And I don't mean 'shine' from being glowing hot. These devices, especially the BF533, are targeted at low-power embedded applications with passive cooling.

It's also possible to make another comparison by applying a performance per dollar metric. I leave it to you to do that math and draw your own conclusions. For EEMBC, though, the bottom line is that there are really four winning processors here. Each has specific features that target different markets, and together they are an excellent showcase for the ability of DENbench to compare processors in a wide range of applications, from MP3 players to mobile phones to set-top boxes.



Markus Levy