Intel Unveils "iCOMP" Performance Index

Composite Performance Rating Aimed at Unsophisticated Buyers

By Michael Slater

In an attempt to give unsophisticated computer buyers an easier way to evaluate the relative performance of the ever-increasing range of x86 microprocessors, Intel has developed a new performance index called iCOMP (Intel COmparative Microprocessor Performance). iCOMP is intended to provide a single number that roughly characterizes the performance of a microprocessor, replacing the ambiguous but ubiquitous "MIPS."

Intel was motivated to create the iCOMP rating by research which showed that many computer buyers assumed that the clock speed-the "MHz" rating-was indicative of performance, regardless of the processor type. Many buyers thought that a 33-MHz 386, for example, was faster than a 25-MHz 486. The P5 will exacerbate this problem.

Unfortunately, although Intel's goals in creating iCOMP are understandable, the concept is inherently flawed. Despite what appears to be a good-faith effort to create a useful metric, iCOMP is likely to create more problems than it solves. The most fundamental problem with iCOMP is that it associates a specific performance rating with a microprocessor type, independent of the system implementation. In reality, performance varies considerably in various system designs with a given processor, so the iCOMP rating can be quite misleading. It will only distract attention from application-level benchmarks, which are the only true measures of performance.

Category	Benchmark	Weight
16-bit Integer	ZDbenchCPU	52%
16-bit Floating-Point	16-bit Whetstone	2%
	ZDbenchCPU	1%
16-bit Graphics	ZDbenchCPU*	10%
16-bit Video	ZDbenchCPU*	5%
32-bit Integer	SPECint92	15%
32-bit Floating-Point	SPECfp92	5%
32-bit Graphics	SPECint92*	5%
32-bit Video	SPECint92*	5%

Table 1. Components of the iCOMP rating and their weights. Benchmarks marked with an asterisk (*) will be replaced by specific graphics and video benchmarks in the next release of the definition. The combined weights for this release are 68% ZDbenchCPU, 2% 16-bit Whetstone, 25% SPECint92, and 5% SPECfp92.

iCOMP Derivation

While Intel will provide iCOMP ratings only for its own microprocessors, it is based on standard benchmarks and the formula for calculating iCOMPs has been made public. This will allow other x86 microprocessor makers to provide their own iCOMP ratings, should they choose to do so.

iCOMP is not a new benchmark; rather, it is a performance metric based on a composite of existing benchmarks. Each of the benchmark components is scaled so that a 486SX-25 scores 1.0, and each is then weighted according to Intel's estimate of how widely used that particular function is.

Table 1 shows the eight benchmark categories, the benchmark program used, and the weight. Intel chose a 70%/30% split of 16-bit and 32-bit code based on International Data Corp.'s estimate of the mix in the installed base as of 1995. The split between integer, floating-point, graphics, and video is based on Intel's own estimates.

The largest component is the integer CPU benchmark from Ziff-Davis Labs (ZDbenchCPU), which is derived from the earlier PC Labs benchmarks. Whetstone (as implemened in PowerMeter) is used for 16-bit floating-point, and SPECint92 and SPECfp92 are used for the 32-bit components.

For the initial iCOMP release, no specific benchmarks have been identified for graphics or video, so ZDbenchCPU is used for 16-bit video and graphics, and SPECint92 is used for 32-bit video and graphics. These will be replaced with more applicable benchmarks in the next release of the iCOMP specification.

The future video and graphics suites will be designed to measure just the CPU component of these functions, independent of any particular graphics or video hardware. Intel plans to use a WinMark-like benchmark for graphics, with a null display driver that eliminates the actual writes to the display controller so it will measure only the processor aspect of graphics performance and not the speed of the display subsystem. For video, Intel is considering video decompression using both MPEG and RTV (one of Intel's DVI algorithms). The plan to include these components suggests that Intel intends to add features to future microprocessors that will accelerate such graphics and video functions.

The iCOMP rating is calculated as a weighted geometric mean of the eight components, as follows:

Processor	Clock Rate	Relative ZDbenchCPU	Relative SPECint92	Relative SPECfp92	Relative 16-bit Whetstone	iCOMP	SYSmark92	Normalized SYSmark92	System tested for Sysmark92
386SX	16	0.24	0.17	0.14	0.27	22	28.9	27	Compaq Deskpro
	20	0.37	0.24	0.19	0.41	32	42.4	40	Compaq Deskpro
	25	0.43	0.32	0.28	0.54	39	51.6	48	NEC Powermate
386SL	25	0.47	0.30	0.27	0.52	41	51.9	49	Compaq LTE
386DX	25	0.52	0.45	0.29	0.59	49	59.5	56	Compaq/M
	33	0.74	0.59	0.45	0.78	68	77.5	73	Compaq/L
486SX	16	0.62	0.65	0.64	0.64	63	68.1	64	Compaq/M
	20	0.80	0.75	0.74	0.80	78	82.2	77	Everex Tempo
	25	1.00	1.00	1.00	1.00	100	106.4	100	Compaq/50M
	33	1.36	1.37	1.34	1.34	136	129.7	122	Compaq/i
486DX	25	1.02	1.00	15.00	16.76	122	107.9	101	Compaq/50M
	33	1.38	1.37	20.06	22.40	166	131.3	123	Compaq/i
	50	2.08	2.02	30.24	33.60	249	179.9	169	Compaq/L
486DX2/ OverDrive	16/32	1.09	1.08	17.26	21.32	132	108.3	102	Compaq/M
	20/40	1.44	1.19	19.57	26.94	166	127.2	120	Everex Tempo
	25/50	1.94	1.81	27.37	33.60	231	168.1	158	Compaq/50M
	33/66	2.51	2.28	36.10	44.95	297	189.7	178	Intel 403

Table 2. Relative performance results for the iCOMP component benchmarks and the composite result. Note that the SPECfp92 ratings for processors without FPUs are not measured numbers but are the ratings for the same processor with an FPU divided by 15. Also shown for comparison are BAPCo SYSmark92 results, and SYSmarks normalized so 486SX-25 = 100. (Data source: Intel.)

$$\text{iCOMP} = 100 \times \left(\frac{\text{BM}_1}{\text{Base\_BM}_1}\right)^{\text{P}_1} \times \left(\frac{\text{BM}_2}{\text{Base\_BM}_2}\right)^{\text{P}_2} \times \ldots \times \left(\frac{\text{BM}_8}{\text{Base\_BM}_8}\right)^{\text{P}_8}$$

BM_n is the performance on benchmark component n, and P_n is the weight given to that benchmark (with P_n = 1.0 corresponding to a weight of 100%; a 5% weight results in a P_n of 0.05). The Base_BMn figure is the performance on benchmark component n for the base processor, which Intel has chosen as the 486SX-25. Using a geometric mean, rather than a simple average, limits the degree to which exceptional performance on one component skews the final result.

For processors without an FPU, executing the SPEC floating-point suite is problematic. According to Intel, floating-point emulation libraries are not available for most 32-bit compilers, and the trap overhead in the UNIX emulation mechanism imposes an unrealistic burden. To get around this problem, Intel decided to simply measure floating-point performance with an FPU and then divide by 15 as an estimate for the same processor without an FPU. The floating-point aspect of the iCOMP formula is thus only a rough approximation, at best, for processors without an FPU.

iCOMP Results

Table 2 shows the iCOMP ratings for Intel's 386 and 486 microprocessors, along with the individual benchmark results (scaled to 486SX-25 = 1.0), BAPCo "SYSmark92" benchmark results, and SYSmark92 ratings normalized so that the 486SX-25 scores 100 for comparison with the iCOMP results. The BAPCo suite is based on real PC applications running scripts, and it includes disk and display performance as well as CPU performance (see sidebar for information on how to get the BAPCo suite or reports).

As the table shows, the iCOMP ratings have a wider range than the SYSmark92 figures. For example, the fastest x86 processor-the 486DX2-66-has an iCOMP rating of nearly three times the base 486SX-25, but its SYSmark92 rating is only 1.78 times higher. Intel claims that the wider range of iCOMP ratings is due to the inclusion of 32-bit programs; BAPCo uses only 16-bit programs. However, an even bigger factor is that the iCOMP rating measures processor and memory performance alone, while the SYSmark92 rating is decreased by the limitations of disk and display systems. For the same reasons, the range of performance from the fastest to the slowest processor is 13.5:1 for iCOMP, but only 6.6:1 for SYSmark92.

Note that the 386 ratings assume that there is no 387 math coprocessor installed. Intel has not provided iCOMP ratings for the 386/387 combination, but these ratings can be computed from the information in Table 2 by multiplying the 386 floating-point values by 15.

iCOMP is intended to be a microprocessor rating, not a system rating. There is, however, no way to measure processor performance without a memory system, so some system assumptions are implicit. Intel selected what it considers to be "best of breed" systems for each processor to measure the iCOMP ratings.

Differences in the system designs used to measure iCOMPs result in some curious ratings. For example, the 50-MHz 486DX has an iCOMP rating that is 2.04 times the 25-MHz 486DX. For a metric that is supposed to be a

pure processor rating, it is odd that performance would increase by a larger factor than the clock rate for a given processor type. Both test systems have a 256-Kbyte write-back cache, so this discrepancy is puzzling. As another example, the 25-MHz 386SX has an iCOMP rating that is 77% higher than the 16-MHz version, even though the clock rate increases by only 56%. In this case, the 25-MHz test system has a cache but the 16-MHz system does not, which explains the disproportionate performance increase.

Is This Useful?

Despite Intel's goal of simplifying the computer selection process, giving the industry a new single-number performance index that ignores system effects only adds to the confusion. There are several ways in which it may mislead buyers as much as it educates them. A buyer looking at iCOMP ratings, for example, might expect that 486DX2-66 system would run applications nearly three times faster than a 486SX-25 system, while according to BAPCo the speedup is only 78%.

Another problem with iCOMP is that it is intended as a processor benchmark, not a system benchmark, so Intel will tell system makers to use language such as "based on a 166-iCOMP 486DX-33," rather than calling the system a 166-iCOMP computer. A computer with a low-cost memory system, however, will not meet this performance rating. Unscrupulous system makers (of which there seems to be no shortage) might promote such systems using the processor's iCOMP rating, even though their design is incapable of achieving that performance level.

System makers might measure iCOMPs on their systems, which would lead to its abuse as a system benchmark. Unless system makers do this, however, iCOMPs will be of no use in selecting a system within a particular category. In addition, display system performance can be more important than processor speed. A 386DX-33 system with local-bus graphics may well perform better on most Windows applications than a 486DX-25 with standard VGA, even though the iCOMP rating of the 486 system is much higher.

From Intel's perspective, it may be useful to have a performance index that can be associated with particular microprocessors, independent of the systems in which they are used. For users, however, this is irrelevant; all that matters is system-level performance. For this purpose, the BAPCo SYSmark92 is a much more useful metric. The iCOMP metric does have the advantage, from Intel's perspective, that it exaggerates the value of a faster processor-surely a good thing when you are trying to move the market from 386 to 486 and then to P5. In the real world, display systems and disk drives keep the full performance potential of faster processors from being realized-a fact that iCOMP obscures.

BAPCo

The Business Applications Performance Corporation (BAPCo) suite costs $795 per copy plus shipping (and sales tax in California). To order, call IDG Books Worldwide at 415/312-0650.

BAPCo results are published in a quarterly report, which costs $250 for a one-year subscription or $100 for a single issue, plus shipping (and sales tax in California). The first issue has just been published, with results for 29 systems plus articles describing the benchmarks. The first batch of results includes systems using most of Intel's 32-bit microprocessors, IBM's 386SLC and 486SLC2, and AMD's 386SX and DX; there are no results for Cyrix or C&T processors. Order from BAPCo at 408/988-7654; fax 408/765-4920.

iCOMP also has the advantage, as compared to MIPS ratings, of including a floating-point component, so it will show a bigger difference between chips such as the 486DX and Cyrix's 486DLC.

From the computer buyer's viewpoint, Intel would be better off spending its resources promoting the use of the BAPCo suite instead of giving the world yet another composite performance rating that makes no real contribution. Benchmarks that don't measure application-level performance have relatively little value, and single-number benchmarks are inherently limited. BAPCo, in addition to providing a rating that more accurately reflects application-level performance, provides a set of six numbers for different application categories. This set of numbers provides much more information than a single, composite number that blends floating-point and integer performance and doesn't consider system-level effects.

One reason why Intel developed iCOMPs is that system makers weren't jumping on the BAPCo bandwagon. System makers aren't eager to publish BAPCo results unless they happen to make the "best of breed" system, so BAPCo isn't getting as much use as Intel hoped. It seems that Intel could have achieved the same objective as iCOMP by simply publishing a list of typical BAPCo results for each processor type. Apparently, this option was not chosen because of fear of alienating system makers; those with slower systems than Intel's "standard" BAPCo ratings would be at a disadvantage.

The iCOMP ratings could create the same problem, and their only real advantage in this respect seems to be the difficulty of measuring them. The bottom line on iCOMPs appears to be that it provides a way for system makers to quote performance numbers without revealing the performance of their particular system implementation.

Albert Einstein once said, "Everything should be made as simple as possible-but no simpler." The iCOMP metric fails to heed the wisdom of this maxim. ♦

