

## THE EDITOR'S VIEW

## Benchmark Abuse

## How to Tell the Truth and Lie at the Same Time

By Michael Slater

Benchmarks have always been a controversial subject. The marketing value of benchmarks creates a strong motivation for "aggressive" marketers to find ways to warp the results to put their product in the best possible light. To steal a line from John Mashey, benchmarks don't lie, but liars benchmark.

Many companies—especially those with the best performing products—make an effort to report meaningful benchmark results. There have always been those who push the limits of honesty, however, and the emergence of non-Intel x86 microprocessors has given them new opportunities to mislead customers.

When Intel was the only vendor of x86 microprocessors (except for direct alternate sources), benchmarks were less important than they are today. Each of the Intel processors occupied a clearly different price/performance point, making it easier to compare systems. Users could generally decide on a class of machine—such as a 33-MHz 486—and then just look for the system that offered the best price and features within that class. The limited needs of this relatively straightforward marketplace, combined with the desire of advertisers to make everything as simple as possible, may explain why simplistic commercial benchmarks such as Norton SI and Landmark "MHz" became the standards, despite efforts by many magazines to establish realistic benchmark suites.

The emergence of new vendors of 386- and 486-compatible processors has dramatically increased the need for good benchmarks to evaluate system performance, since there is now much more overlap among processor types and relative performance levels are not so clear. With 386 processors pushed up to 40 MHz, for example, there is overlap in performance between high-end 386 systems and low-end 486 systems.

Processors such as Cyrix's 486SLC make the situation even more confusing, since this processor has a different CPU microarchitecture and cache structure than any of Intel's designs, and its cache is much smaller than that in Intel's 486. Evaluating the performance of a microprocessor with a tiny on-chip cache by using a benchmark that fits in that cache, as most common PC benchmarks do, produces very misleading results.

One of the most egregious cases of benchmark abuse is now showing up in advertisements for systems using Cyrix's 486SLC. PC Brand, for example, in a full-

page ad headlined "Our 486 Desktops Will Blow You Away," shows Landmark results as the only benchmark. The numbers are impressive: 2.4 times the performance of a 386SX at the same clock speed. This speed is obtained on the Landmark program because it is small enough to mostly fit in the 486SLC's 1-Kbyte on-chip cache. On real programs, Cyrix claims only 1.4 to 1.6 the performance of a 386SX—but you won't find this fact anywhere in the PC Brand ads.

Out of curiosity, I called PC Brand's 800 number and asked if the "2.4 times a 386SX" performance rating was representative of the system's performance on real programs. The salesperson assured me that it was, which reminded me of an old joke: What is the difference between a computer salesman and a used car salesman? The car salesman *knows* when he is lying.

Other than the system name (486/SLC-25), there is no hint in the ad that the microprocessor is not made by Intel, or that there is any difference between a Cyrix 486SLC and an Intel 486. Indeed, the same ad lists 486/33 and 486DX2-50 systems, which are based on Intel processors.

Ads like this that gloss over the difference between an Intel 486 and a Cyrix 486SLC and don't give realistic benchmarks could create a lot of disappointed system buyers and result in a backlash against non-Intel chips. (Cynics might claim that there will be no disappointment because most users won't ever notice that they're not getting the performance the ad claimed, but some users upgrading from 386SX systems are surely aware enough to tell the difference between a 1.4× performance boost and a 2.4× performance boost over their old system.) Users are confused enough without being intentionally misled.

PC Brand isn't alone, either; a press release from Wyse announcing its 486SLC-based notebook computer also uses Landmark as the only benchmark, giving the misleading impression that the system is more than 2.5 times as fast as one based on a 386SX. All the companies involved would no doubt defend their use of Landmark as simply responding to what customers ask for, but this is no excuse. When the numbers customers ask for have become misleading, it is the responsibility of honest vendors to take the lead in switching to more meaningful benchmarks, such as the BAPCo and SPEC suites. By quoting only Landmark figures for the Cyrix processor, vendors are telling a literal truth while lying about the system's real performance. ♦