Reed Electronics Group

MICROPROCESSOR REPORT

www.MPRonline.com

THE INSIDER'S GUIDE TO MICROPROCESSOR HARDWARE

# TENSILICA XTENSA V HITS 350MHZ

## *Customizable CPU Achieves Highest EEMBC ConsumerMark Score*

*By Tom R. Halfhill {9/16/02-01}*

Tensilica is shipping the fifth version of its customizable soft microprocessor core since the debut in 1999, adding new features and enhancing some existing ones. According to Tensilica's simulations, the new Xtensa V core should run at 350MHz (worst case) when fabricated in a 0.13-micron CMOS process. Tensilica says some Xtensa V configurations will run even faster and that improvements to the company's proprietary hardware-design language and C/C++ compiler can boost actual software performance by 50% over the Xtensa IV.

The target clock frequency of a configurable microprocessor can vary widely, for several reasons. The designs are fully synthesized; customizations greatly affect the size of the core; and complex extensions may force the whole processor to run more slowly to avoid clock skew. Tensilica's estimate of 350MHz is for an "average" configuration of Xtensa V with 50,000 to 70,000 gates. The base configuration of 25,000 gates could push 400MHz, and Tensilica has simulated a 48,000-gate configuration optimized for EEMBC's office-automation benchmark suite at 360MHz. At the other end of the spectrum, Tensilica simulated a much larger 263,000-gate configuration, optimized for EEMBC's consumer benchmark suite, at the considerably slower clock rate of 260MHz.

All those estimates are based on Tensilica's post-layout 3D-extraction simulations using 0.13-micron design rules. They assume a high-performance process such as TSMC-LV, which is faster than the lower-cost generic TSMC-G process but slower than the highest-speed TSMC-HS process. Tensilica's worst-case conditions assume 0.9V core power (–10% $V_{dd}$) at 125°C. Although the difference in clock frequencies among different Xtensa V core configurations is significant—260MHz to 360MHz for the three different EEMBC configurations—the extra performance enabled by custom extensions is often well worth the trade-offs in gate count and clock speed, as we'll discuss later in this article.

Tensilica specified a worst-case clock frequency of 200MHz for the older Xtensa IV core in a larger 0.18-micron process. That means Xtensa V is about 75% faster, largely because of the process shrink. According to Tensilica, the smaller process provides a 50% speed boost, with improvements to the automated scripts for Synopsys Physical Compiler accounting for the remaining 25%. Apparently the new scripts are somewhat less effective with 0.18-micron design rules; Tensilica says Xtensa V customers that prefer to target the larger, more mature process (which offers lower mask costs) can expect a clock frequency only 10–20% higher than Xtensa IV in the same process.

Existing Xtensa IV binaries are compatible with Xtensa V. However, to take advantage of some new features that extend the instruction-set architecture (ISA), programmers will have to modify their code.

### Changes Affect Core and Tools

The Xtensa V revisions are not a major overhaul, but they do affect the architecture at every level. There are new instructions and features in the processor core; new functions in the Tensilica Instruction Extensions (TIE) language, a proprietary Verilog derivative; and new optimizations in the Xtensa C/C++ compiler (XCC).

What's unchanged is the basic Xtensa design flow. Developers begin by optimizing the Xtensa core for their

## Wiggle Room In EEMBC's Simulated Benchmarks

Has the temperature in Hades plunged below 32°F? Perhaps, judging by *MPR*'s amazing discovery that a CPU vendor deflated its own EEMBC benchmark scores by 25% to offer a more realistic estimate of actual performance.

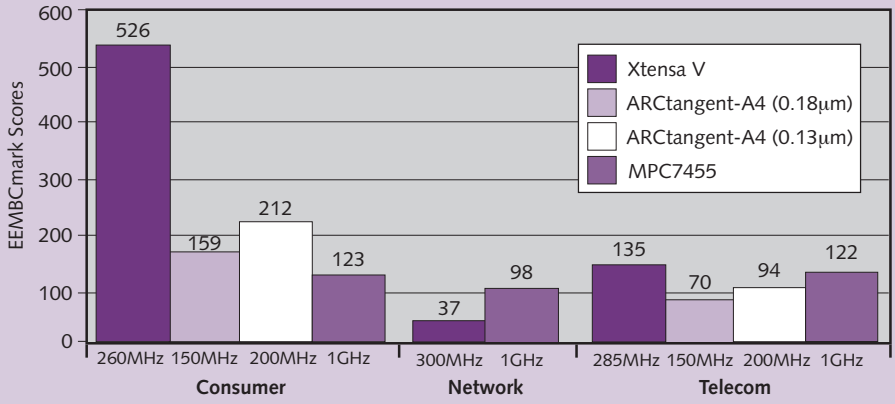
Most EEMBC members benchmark real microprocessors, but EEMBC allows soft-IP vendors to test cycle-accurate simulations of their processor cores. It's too costly and time-consuming to spin custom silicon for the tests. This compromise is acceptable, given the accuracy of modern simulation tools. But what if—for whatever reason—a processor can't reach the expected clock frequency in actual silicon?

The figure below shows some EEMBC scores in the consumer, networking, and telecommunication suites. Tensilica tested three different optimized simulated configurations (at 260MHz, 300MHz, and 285MHz) of the new Xtensa V microprocessor core—a specially optimized configuration for each suite—and racked up impressive scores. ARC International tested two different simulated configurations of the ARCtangent-A4 microprocessor core, optimized for the consumer and telecommunication suites, and scaled the benchmark scores to 150MHz and 200MHz. (ARC has not published results for the network suite.) For comparison, we included a Motorola PowerPC MPC7455 microprocessor—the only CPU on this chart that exists in silicon—because it's the highest-scoring real chip in these suites.

As the figure shows, an optimized Xtensa V simulated at 260MHz posted the highest-ever EEMBC ConsumerMark score, shaming the 1GHz MPC7455 by a factor of 2.5x. At 285MHz, a different Xtensa V configuration bested the same PowerPC chip in the telecommunication suite, even though Motorola optimized the benchmark code with AltiVec extensions. The PowerPC did manage to defeat the Xtensa V in the network suite by a factor of 1.7x, probably because those tests depend more heavily on high data throughput, which allowed the 1GHz PowerPC to shine. However, at 100x the Xtensa's power consumption, it was a Pyrrhic victory.

ARC's scores, though not as high as Tensilica's, are also impressive. The 150MHz and 200MHz ARCtangent-A4 processors beat the 1GHz PowerPC in the consumer suite and finished close behind in the telecommunication suite.

EEMBC Certification Laboratories (ECL) certified ARC's results earlier this year and stated the raw scores in iterations per 1MHz, as it does with all simulated processors. CPU vendors are supposed to scale the 1MHz ECL-certified scores to a clock frequency that's a realistic estimate of the simulated processor's clock speed in silicon. With one exception, every vendor that




applications, using Tensilica's graphical design tools and (optionally) the TIE language. Next, they submit their design over the Internet to Tensilica's Web-based automated processor generator, which converts the processor's HDL model and user-defined TIE extensions into RTL (VHDL or Verilog). The generator also creates synthesis scripts and software-development tools optimized for the custom processor. Developers can simulate the processor's RTL model to verify the design and repeat the first two steps as often as necessary. When they're satisfied, they generate a gate-level netlist with their own synthesis compiler, targeting a fabrication process of their choice. Finally (or in parallel), developers write software by using XCC or a customized GNU compiler that emits native code for the custom processor.

New and improved features at the CPU level enhance Xtensa V's I/O capabilities. For the first time, the processor can act as a slave while an external device (or an external function block in an SoC) uses DMA for reads and writes through the core's processor interface. This will improve the efficiency of data transfers between other devices (including tightly coupled coprocessors) and local data memory, because the CPU can execute other instructions while handling the incoming read/write request. The Xtensa processor interface is configurable, supporting data read/write bus widths of 32, 64, and 128 bits.

The Xtensa Local Memory Interface (XLMI) now supports devices that require multiple cycles to access or that cannot respond in a real-time deterministic fashion. Previously, any device attached to the XLMI had to respond in a single cycle; variable-latency devices had to be attached to the processor interface. When a device attached to the new XLMI needs more than one cycle, it stalls the CPU for the duration

**EEMBC's Simulated Benchmarks** *continued from Page 2*

has published certified EEMBC benchmark scores in simulation (including ARM, Improv, Intrinsity, and Tensilica) has stipulated an estimated clock frequency in its EEMBC benchmark report. The one exception was ARC.

ARC did scale the 1MHz scores to 200MHz in a March 10 press release announcing the scores. ARC says the estimate was based on expected performance if the tested configuration of the processor were fabricated in a 0.18-micron CMOS process.

But the size of a configurable processor can greatly affect its clock frequency. The optimized configurations of the simulated ARCtangent-A4 processors contained 120,000 gates (telecommunication suite) and 170,000 gates (consumer suite). Both configurations are much larger than a 62,000-gate unoptimized configuration that ARC tested with EEMBC's out-of-the-box benchmarks. A minimal ARCtangent-A4 core is about 8,500 gates.

Since issuing the March 10 press release, ARC has found that the optimized ARCtangent-A4, as configured for the EEMBC tests, probably cannot achieve 200MHz in a typical 0.18-micron process. ARC therefore quietly scaled down the scores to a more realistic clock rate of 150MHz and reported the 25% lower score on the EEMBC page of its Web site.

ARC's action must be unprecedented. We've never heard of a CPU vendor voluntarily deflating its own benchmark scores to such a large degree. Imagine if AMD described a 2GHz Athlon as a PR1500 processor.

However, according to EEMBC president Markus Levy, it's against EEMBC rules to arithmetically scale a certified benchmark score to a different clock frequency, whether the scaling is up or down. The vendor is supposed to apply to ECL for recertification, a process that takes several weeks and costs thousands of dollars. ARC would therefore seem to be in violation of the rules. But it's a gray area, because ARC's benchmark report never stipulated a target clock frequency to begin with, and neither EEMBC nor ECL noticed the omission until *MPR* brought it to their attention. EEMBC recently tightened up its rules—members must prominently specify the clock frequency when reporting and publicizing benchmark results.

ARC says the ARCtangent-A4 should easily reach 200MHz in a smaller 0.13-micron process, which is why we're showing both sets of scores (150MHz and 200MHz) in the figure. *MPR* isn't a member of EEMBC and isn't bound by its reporting rules.

On the one hand, ARC deserves credit for voluntarily restating its performance estimates at a more realistic clock frequency. On the other hand, ARC deserves criticism for overestimating the clock frequency in the first place. But the episode provides a useful lesson for potential customers of any customizable microprocessor core. Depending on the configuration, your mileage may vary, and there are many trade-offs to consider.

Note that the optimized configurations of the ARCtangent-A4 processor still scored 19x higher in the consumer suite and 41x higher in the telecommunication suite when compared with their unoptimized configurations—easily offsetting the 25% drop in clock speed. Likewise, Tensilica's optimized Xtensa V processors showed gains of 23x, 4x, and 15x in the consumer, networking, and telecommunication suites, respectively, when compared with their unoptimized configurations. (All these comparisons are based on raw scores at 1MHz and shouldn't change significantly at any clock frequency.)

Are those huge performance gains worth the extra gates, higher power consumption, and additional design time for a customized core? The answer depends on your application and time-to-market window. What matters is asking the right questions.

of the exchange, but at least it won't drop any data. As Figure 1 shows, chip designers often use the XLMI to attach memory, coprocessors, and application-specific logic to an Xtensa core.

A significant new enhancement to Xtensa V is the option of adding conditional load and store instructions to the processor. Conditional loads and stores can often improve software performance and reduce code size by eliminating some explicit compares and branches. The new load/store instructions are both optional and configurable. Designers create the instructions in TIE language, specifying which of the processor's 16 status flags will determine whether the load/store should execute. Designers can also create their own application-specific status flags.

Tensilica is promoting two other Xtensa V enhancements as multiprocessing features for SoCs with multiple CPU cores, although both enhancements are relatively minor. One is the addition of a write-back replacement policy for the configurable data cache, which previously had only a write-through policy. As Tensilica points out, write-back caching should reduce data traffic on the internal bus of an SoC with multiple processor cores. But write-back caching is equally beneficial for uniprocessor SoCs and is a common feature of data caches, so the connection to multiprocessing is tenuous. There are no cache-snooping or other cache-coherency mechanisms in Xtensa V.

The other new multiprocessing feature is more relevant, though no more earthshaking. The Xtensa V ISA defines a processor ID register, which allows each CPU core in a multiprocessor SoC to have a unique identifier. This is largely a formality. About 59% of Tensilica's customers are already designing multiprocessor SoCs with earlier Xtensa cores, using an average of 5.1 cores per chip. Presumably, they have

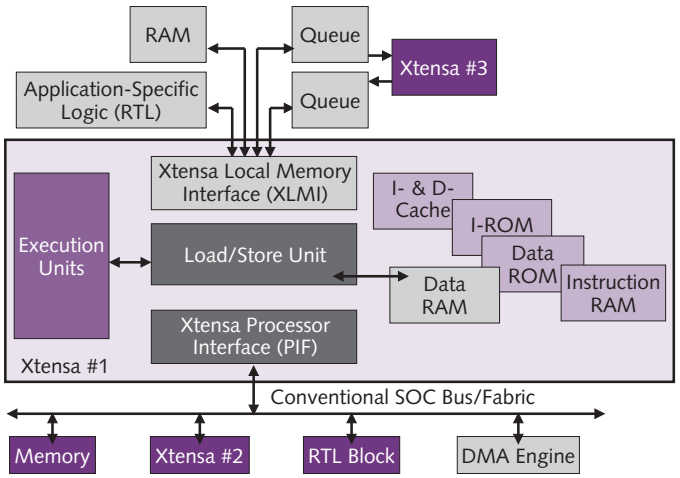



**Figure 1.** Xtensa V now supports external DMA requests through the processor interface (PIF) and variable-latency devices on the local-memory interface (XLMI).

allocated an existing general-purpose register or have created their own extension register for this purpose. The new 16-bit ID register is a convenience and can handle enough identifiers (more than 65,000) to accommodate even the most profligate chip designer.

### Improved TIE and C/C++ Tools

The revised TIE language has new function modules, which are similar to predefined C functions. The modules make it easier for chip designers to create optimized logic that performs n-wide adds, n-way muxes, carry-save adds, and partial-product operations. Designers typically use these functions to extend the Xtensa core with application-specific execution units.

On the software-development side, Tensilica made several improvements to its home-grown XCC compiler, introduced last year to supplement the GNU C/C++ compiler. XCC now does more function inlining (such as cross-file inlining); has improved variable handling; generates better vector/SIMD code for the Vectra DSP engine without manual optimizing; and includes various other code-generation improvements. Tensilica says that, as a result, C programs are 5% to 10% smaller and run up to 50% faster, according to certified EEMBC benchmarks.

To test the improved XCC compiler, Tensilica ran EEMBC's consumer, network, and telecommunication benchmarks on an unoptimized base configuration of Xtensa V, comparing the results to a virtually identical configuration of Xtensa III using the GNU compiler. (Tensilica hasn't published EEMBC results for the Xtensa IV.) Because the unoptimized tests don't use application-specific core extensions, they isolate the effectiveness of the XCC compiler.

Results: XCC was the clear winner. The scores improved by 1.51x (ConsumerMark), 1.58x (NetMark), and 2.64x (TeleMark). Tensilica says the TeleMark score showed the greatest improvement because the XCC compiler can automatically vectorize some four-way SIMD operations for Xtensa's Vectra DSP engine.

EEMBC has been a godsend for customizable-processor vendors like Tensilica and ARC International because it validates their long-running propaganda. Despite the handicaps of fully synthesized designs and clock frequencies that often fall off precipitously with the addition of major extensions, the Xtensa and ARCtangent-A4 processor cores have posted some of the highest scores ever seen in EEMBC tests—outpacing microprocessors that dwarf them in clock frequency, die size, cost, wattage, and market share. In particular, the differences between out-of-the-box and optimized scores (often an order of magnitude or more) demonstrate that a custom-tailored processor can dramatically boost application performance (see *MPR 4/9/01-01*, "Stretching Silicon to the Max").

However, one drawback of Tensilica's and ARC's EEMBC benchmarks is that they are based on simulations, not actual silicon. It's too costly for processor-IP vendors to spin a custom chip for the sole purpose of running EEMBC benchmarks. To ensure the accuracy of the benchmark results, the EEMBC certification process for simulated processors is rigorous. It's based on cycle-accurate, bus-functional models that account for such factors as memory latencies at the target clock frequencies. Even so, it's possible for the simulations to stray from reality, as we discovered when preparing this report. (See sidebar, "Wiggle Room In EEMBC's Simulated Benchmarks.")

### Esprit de Core

Xtensa V feels more like a dot-release revision of Xtensa IV than a major Roman-numeral upgrade. Except for a few optional instructions and registers, the microprocessor core and ISA are largely unchanged. However, Xtensa V is really a platform that includes the TIE language for adding custom logic, the automated services of Tensilica's back-end processor generator, and a suite of integrated software-development tools. When considered as a whole, Xtensa V introduces worthwhile features and improvements.

It also demonstrates that Tensilica is eager to please the growing number of chip designers who integrate multiple CPU cores on SoCs. With royalties based on the number of processors manufactured, a chip that has several cores represents a more lucrative revenue stream. Of course, Tensilica was not the first company to recognize the opportunity. Archrival ARC was the first customizable-processor vendor to pursue this market. The ARCtangent-A4 processor has for years supported master/slave configurations, a flexible local-memory interface, DMA, write-back caching, conditional instructions, a multiprocessing RTOS, multiprocessor prototyping/debugging tools, and an optimized C/C++ development environment.

Neither company offers the extensive cache-coherency mechanisms and other features required for server-style

symmetric multiprocessing—that's up to the chip designer to implement. In general, though, designers aren't using the Xtensa and ARCtangent cores in this manner. They tend to create "multiprocessor" designs instead of "multiprocessing" designs, bolting together multiple cores in various configurations to handle different tasks. One core might be the general-purpose CPU for the RTOS and user interface, while another handles signal-processing functions, and a third is configured with application-specific extensions. Some designers are using the cores for parallel-processing applications that break down a complex task in ways that don't require cache coherency.

Both Tensilica and ARC can boast of design wins that cram dozens of cores on a single chip. It's one of the only technology categories in which embedded processors are miles ahead of their more glamorous desktop and server brethren.

### Price & Availability

Xtensa V is available for licensing now. Licensing fees for a single-processor configuration with GNU software-development tools start at $350,000, plus royalties based on the volume of processors manufactured. The XCC compiler, Xtensa instruction-set simulator, and TIE compiler are priced separately.

*(Editor's note: Halfhill worked at ARC International as a technical writer and analyst from 2000 to 2002 before rejoining In-Stat/MDR in August.)*