Reed Electronics Group

MICROPROCESSOR REPORT

www.MPRonline.com

THE INSIDER'S GUIDE TO MICROPROCESSOR HARDWARE

TENSILICA PREVIEWS VIDEO ENGINE

Synthesizable Dual-Core Decoder Is Optimized for Digital Video

By Tom R. Halfhill {11/28/05-01}

For seven years, Tensilica has steered close to its original strategy of licensing general-purpose configurable-processor cores that developers can customize for specific designs. Since introducing its first configurable processor in 1998, Tensilica has steadily improved the cores, greatly improved the development tools, and offered a few packages of extensions for popular applications.

Now Tensilica is exploring a slightly different course—without abandoning its original strategy or significantly changing its business model. At the recent **Fall Processor Forum** in San Jose, hardware-design manager Gülbin Ezer previewed a high-performance video-decoder engine based on two of Tensilica's Xtensa LX configurable-processor cores. Tensilica is preconfiguring the cores by customizing them with application-specific extensions, adding local memory and other intellectual property (IP), and licensing the whole synthesizable design as a drop-in module for SoCs needing video acceleration.

In short, the new video-decoder engine does for digital video what Tensilica's preconfigured HiFi-2 package already does for digital audio. (See *MPR 9/23/03-01*, "Tensilica Makes Music.") Tensilica is taking a higher-level approach to IP licensing than its usual strategy of offering general-purpose processor cores and leaving all configuring to the customer.

Target applications for the new video engine will include portable video players, digital camcorders, personal video recorders, standard-definition digital TV (SDTV), set-top boxes for satellite TV, and Internet Protocol TV (IPTV). Those are hot markets, and Tensilica faces boiling competition. In the same session at FPF in which Tensilica previewed its video engine, archrival ARC International introduced new audio/video extensions for the ARC 700 configurable processor, and German startup Videantis introduced two synthesizable video-acceleration engines similar to Tensilica's. (See *MPR 11/21/05-01*, "ARC Shows SIMD Extensions," and *MPR 11/7/05-01*, "Videantis Chases Digital Video.") Other competitors include synthesizable-IP vendors ARM and MIPS Technologies. As *Microprocessor Report* has noted in recent articles, there's a virtual stampede toward digital video.

Nevertheless, Tensilica has found a way to differentiate. Tensilica's video decoder integrates two Xtensa LX processor cores in a heterogeneous multicore design. To boost performance, Tensilica has defined more than 200 new instructions for accelerating popular digital-video codecs. Local memories and a special pixel transposer are tuned for streaming video. A proprietary interconnect ties the processors together and links to a multichannel DMA controller. This highly integrated design appears to be thoughtful, not a quick knock-off.

Tensilica plans to ship the video decoder in 2006. The company is also working on an even more powerful video-*encoding* engine that will integrate four heterogeneous processor cores. That engine will run popular video codecs entirely in software without additional hardware acceleration, and it will be capable of simultaneously encoding, decoding, and transcoding multiple video streams.

Multicore Engine Conserves Power

Gülbin Ezer's presentation at FPF was a technology preview, not a product announcement. Nevertheless, it provided

enough information for an early evaluation. *MPR* will analyze the video decoder in more detail next year, when the final design is ready to ship and more information is available.

Building a video engine around two processor cores instead of one might suggest that power consumption isn't a high priority. But, in fact, low power was an important design goal, because some of the target applications are battery-driven portable products. By combining the throughput of two Xtensa LX processor cores—each configured for a different role—the processors can run at a lower clock frequency than a single core performing both roles. Tensilica's cores are relatively small for 32-bit embedded processors, and even with their new video extensions, they don't require much silicon.

One core in the video engine is configured as a stream (control) processor, and the other is configured as a pixel processor. Ezer said her design team arrived at this partition after profiling the codecs on a cycle-accurate instruction-set simulator while decoding video streams at various bit rates and resolutions. A homogeneous dual-core design would have duplicated some resources in the processors, inflating the gate count. Instead, Ezer's engineers settled on a heterogeneous design that carefully partitions the codec's algorithms between the stream processor and the pixel processor. Figure 1 is a block diagram of the video decoder.

The stream processor is configured for serial-processing tasks, such as motion prediction, entropy decoding (symbol-based data decompression), and running control code. It has 30KB of instruction memory, 32KB of data memory, and a 4KB data cache. The instruction memory is a stable repository for critical code, not a conventional L1 cache, which would be prone to thrashing in a streaming-video application. Likewise, the data memory holds important information (such as data for creating prediction blocks from video reference frames) that a conventional data cache might evict at inconvenient moments.

In contrast, the pixel processor runs minimal control code; it concentrates on pixel calculations and transformations. It's the processor that sends reconstructed video frames to main memory for display on the output device. The pixel processor has 20KB of instruction memory, 64KB of data memory, and a 4KB data cache.

As with the stream processor, the pixel processor's instruction and data memories are deterministic program-managed stores for critical code and data, not processor-managed caches. Because the pixel processor interacts directly with main memory, it's connected through the Xtensa Processor Interface (PIF) to a five-channel DMA controller with software-controlled arbitration. Programmers can assign different priorities and bandwidths to each DMA channel.

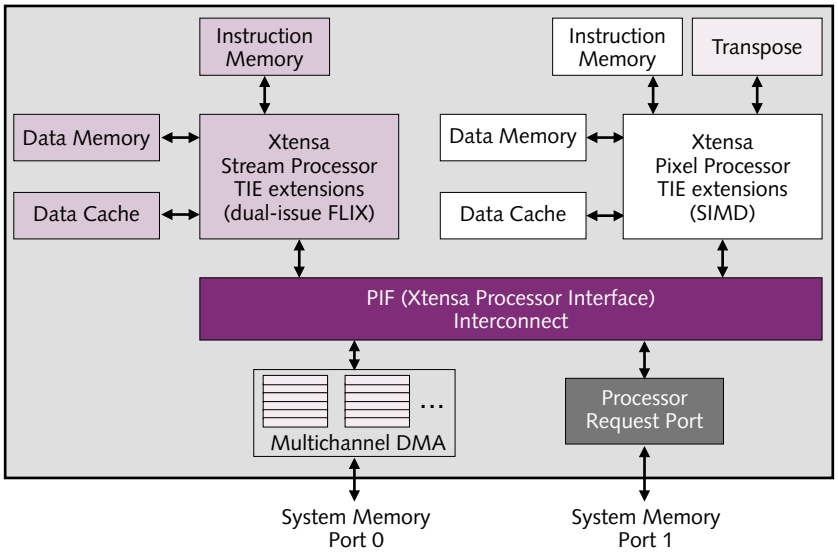

Figure 1. Tensilica's video-decoder engine is a heterogeneous dual-core design with one Xtensa LX processor configured as a stream processor and the other configured as a pixel processor. Each processor has local instruction and data memories in addition to a data cache. The core interconnect is Tensilica's proprietary Xtensa Processor Interface (PIF), which provides an impressive 3.2GB/s of peak bandwidth.

New Extensions Boost Performance

When configuring these processors, Ezer's designers took advantage of new features introduced last year with the Xtensa LX. (See *MPR 5/31/04-01*, "Tensilica Tackles Bottlenecks.") For the stream processor, they used Tensilica Instruction Extension (TIE) language to define new RISC instructions, add a system interface for interrupt and status signals, and create new dual-issue instruction words. The dual-issue instructions, which resemble VLIW, use Tensilica's Flexible-Length Instruction Xtensions (FLIX). (See *MPR 11/25/02-06*, "FLIX: The New Xtensa ISA Mix.")

For the pixel processor, Tensilica used TIE to define new SIMD instructions, build queues for accelerating pixel transpositions, and add system interfaces for interrupt and status signals (like the one in the stream processor). In addition, the pixel processor has a special transposer with a two-dimensional array of dedicated storage elements. The TIE queues interact directly with this array, which can transpose a 16×16 matrix of 8-bit pixels or an 8×8 matrix of 16-bit pixels in 16 clock cycles. Although the transposer is fast, it appears to be less efficient than ARC's new SIMD extensions. ARC's vector-exchange instructions are even faster than Tensilica's transposer, and they operate on pixel values in the SIMD registers instead of using a dedicated storage array. Tensilica's array requires more flip-flops and wiring.

Extensions are indispensable for video processing. Few, if any, general-purpose processor cores can decode digital-video

streams at acceptable frame rates and resolutions without running at impractically high clock speeds or depending on additional hardware accelerators. Many chip designers bolster a general-purpose processor with hard-wired accelerators implemented as register-transfer-level (RTL) blocks. These blocks tend to be specific to a particular design, which hinders reuse, and they aren't as flexible as programmable solutions. Not surprisingly, configurable-processor vendors like Tensilica eschew the hard-wired approach in favor of custom extensions.

In all, Tensilica has defined more than 200 new RISC, SIMD, and FLIX instructions for this video-decoder engine. (The future video-*encoder* engine will add more new instructions; *MPR* hopes to publish a complete instruction table at that time.) Some new instructions control the pixel-transposer and the DMA controller, but most are application-specific instructions for video decoding.

Target codecs are MPEG-2, MPEG-4, H.264, and VC-1 at D1 resolutions and frame rates: 720 × 480 pixels at 30 frames per second (NTSC), and 720 × 576 pixels at 25 frames per second (PAL). The video engine is fast enough to decode MPEG-2 main profile at main level (MP@ML); MPEG-4 advanced simple profile at level 5 (ASP@L5); H.264 main profile at level 3 (MP@L3); and VC-1 main profile. Target algorithms include motion compensation, pixel deblocking, various transforms, and two methods of digital-video compression: context-adaptive binary arithmetic coding (CABAC) and context-adaptive variable-length coding (CAVLC).

Optimizing CABAC is a notable achievement. CABAC is a high-ratio compression scheme for H.264, but it's too demanding for some video-decoder engines, which must outsource the task to a hard-wired accelerator or another processor. Decoding a CABAC stream at D1 resolution can force a general-purpose RISC processor to run at 700–800MHz, even if it's doing nothing else. Tensilica's solution: 64-bit-wide dual-issue FLIX extensions. They slash the CABAC overhead to 13MHz and require only 27,000 gates in the preliminary design, which Tensilica says will shrink to about 20,000 gates in the final optimized design. Table 1 shows the results of Tensilica's internal benchmarking with a video stream compressed using CABAC and CAVLC.

(For more detailed information about CABAC, see "Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard," published in *IEEE Transactions on Circuits and Systems for Video Technology*, July 2003.)

Dual-Core Design Stays Small

Because Tensilica intends to license its video-decoder engine for use in portable consumer electronics, power efficiency is paramount. However, Tensilica hasn't publicly estimated the engine's typical power consumption. A rough guess, based on the estimated gate count in a 0.13-micron process, is 50mW at 200MHz. Peak power could surge much higher when both processors are busy decoding a highly compressed H.264 video stream.

Tensilica says the dual processor cores by themselves require a total of 287,000 gates. Adding the DMA controller, PIF, and pixel transposer expands the design to nearly 400,000 gates, excluding memories. With memories, the whole video engine in Figure 1 occupies less than 10.5mm^2 when fabricated in a generic 0.13-micron CMOS process. That's admirably small for a dual-core video engine with DMA and optimized extensions. Tensilica's cycle-accurate simulations indicate the processors won't need to run faster than 200MHz. That allows the system to use 16-bit DDR-DRAM at 200MHz, further reducing system cost and power.

Based on preliminary specifications, it appears that Tensilica's video decoder will be competitive with the video extensions and engines that ARC and Videantis presented in the same session at FPF. An ARC 750D processor core with new SIMD extensions requires about 268,000 gates, excluding memories. Our previously referenced article about the ARC SIMD extensions shows a floor plan of the enhanced ARC 750D with memories, which occupies 9mm^2 when fabricated in TSMC's 0.13-micron low-voltage (LVLKOD) process and 6.93mm^2 in TSMC's generic (G) process. That's about 86% as large as Tensilica's 10.5mm^2 design when fabricated in a similar process. (Note that ARC's design has a total of 106KB of memory, whereas Tensilica's has 154KB.)

ARC's single-core design targets the same video codecs as Tensilica's dual-core engine, also at resolutions up to D1, and it can handle CABAC decoding, too. ARC is estimating power consumption at 80mW at the 750D's maximum worst-case clock frequency of 533MHz, but that doesn't mean the processor must run that fast to match the performance of Tensilica's video engine at 200MHz. ARC says it will disclose more power and performance estimates when the complete hardware/software package ships later

	Dual-Core Xtensa LX (No Extensions)	Dual-Core Xtensa LX (With Extensions)	Acceleration
CABAC Test Video (Rugby Football) @ 5.4Mb/s			
Stream Processor	1.3GHz	196MHz	6.6x
Pixel Processor	1.2GHz	187MHz	6.5x
CAVLC Test Video (Rugby Football) @ 5.4Mb/s			
Stream Processor	701MHz	188MHz	3.7x
Pixel Processor	1.2GHz	185MHz	6.6x

Table 1. Tensilica measured video performance on two simulated video-decoder engines: one using a pair of base-configuration Xtensa LX cores for the stream processor and pixel processor and another using the same processors optimized with the new video extensions. These results show the required clock speeds for each processor when decoding an industry-standard video stream of rugby football at D1 resolution. The extended processors can decode the video while running at less than 200MHz, a dramatic reduction compared with the performance of unextended processors.

Tensilica Introduces Xtensa 6 Processor Core

Tensilica has introduced a new configurable-processor core that's a hybrid of the Xtensa V core released in 2002 and the Xtensa LX released in 2004. The new Xtensa 6 is available for licensing now. Xtensa LX remains available as the luxury model of Tensilica's product line.

The most important difference between the Xtensa 6 and Xtensa V is compatibility with Tensilica's lastest V6 suite of hardware- and software-development tools, including the Xtensa PRocessor Synthesis tool (XPRES). XPRES is a breakthrough tool that can automatically generate optimized extensions for the configurable processor by analyzing application software written in ordinary C and C++. Until now, XPRES was compatible only with Xtensa LX. (See *MPR 7/12/04-01*, "Tensilica's Automaton Arrives.")

Feature	Tensilica Xtensa V	Tensilica Xtensa 6	Tensilica Xtensa LX
Architecture Width	32 bits	32 bits	32 bits
Uniscalar Pipeline	5 stages	5 stages	5 or 7 stages
FPU	Optional	Optional	Optional
MAC (16-Bit)	Optional	Optional	Optional
Multiplier (16-Bit)	Optional	Optional	Optional
Multiplier (32-Bit)	Optional	Optional	Optional
Xtensa PIF	Optional	Optional	Optional
XLMI	Optional	Optional	Optional
MMU & TLB	Optional	Optional	—
Vectra LX DSP Engine	—	—	Optional
FLIX	—	—	Optional
Load/Store Units	1	1	1 or 2
Custom Ports/Queues	—	—	Optional
Xtensa V6 Dev Tools	—	Yes	Yes
XPRES Tool	—	Yes	Yes
Improved Clock Gating	—	Yes	Yes
Gates (Base Config)	18,000	20,000	20,000
Power (Base Config)*	0.065mW / MHz	0.04mW / MHz	0.04mW / MHz
Max Clock Speed*	350–400MHz	350–400MHz	350–400MHz
Introduction	2002	2005	2004

All of Tensilica's configurable processors are based on the same proprietary 32-bit RISC architecture and are upwardly software compatible. The ability to use Tensilica's powerful XPRES tool is the most welcome improvement of the Xtensa 6 over Xtensa V. (*Tensilica's estimates, assuming fabrication in TSMC's 0.13-micron low-voltage CMOS process.)

Another improvement is lower power consumption. The Xtensa 6 has better clock gating than the Xtensa V and matches the clock gating in the Xtensa LX. (Indeed, the base configurations of the Xtensa 6 and Xtensa LX are virtually identical.) As a result, the Xtensa 6 consumes 25–35% less power per megahertz than the Xtensa V. Specifically, the minimum base configuration of the Xtensa 6 consumes about 0.04mW per megahertz when fabricated in a low-voltage 0.13-micron CMOS process. Of course, the actual power consumption of a user-configurable processor depends greatly on how extensively it's customized—few real-world designs use the stripped-down base configuration.

In other respects, the Xtensa 6 closely resembles Xtensa V. (See *MPR 9/16/02-01*, "Tensilica Xtensa V Hits 350MHz.") The Xtensa 6 lacks several optional features available for Xtensa LX, such as a deeper pipeline, the Vectra LX DSP/SIMD engine, Flexible-Length Instruction Xtensions (FLIX), a second load/store unit, and user-defined ports and queues. (See *MPR 5/31/04-01*, "Tensilica Tackles Bottlenecks.") However, the Xtensa 6 does have an optional memory-management unit (MMU) with a translation lookaside buffer (TLB), a feature offered for the Xtensa V but not for Xtensa LX. The MMU allows the Xtensa 6 to run sophisticated operating systems such as Linux.

The accompanying table summarizes the differences among the Xtensa V, Xtensa 6, and Xtensa LX. Note that Tensilica is moving away from Roman numerals, perhaps to avoid confusion over inconsistent product names. (The "LX" in Xtensa LX is pronounced "el eks" and doesn't represent the number 60.)

this year. Until then, ARC estimates that a 750D processor with video extensions will consume about 30mW at 200MHz. That's about 20mW less than our estimate for Tensilica's video engine, but there's not enough data to conclude which will be more power efficient under real-world conditions. EEMBC recently finishing working on new power-consumption benchmarks, so perhaps these comparisons will require less guesswork in the future.

At FPF, Videantis introduced two synthesizable video engines based on its proprietary v-MP2 processor core. Unlike the ARC and Tensilica cores, which are general-purpose RISC processors, the v-MP2 is designed solely for video processing. For that reason, one might expect it to handily beat the ARC and Tensilica cores. It does appear to have an advantage, but the performance improvements possible with configurable processors definitely narrow the gap.

Of the two video engines that Videantis presented at FPF, the one most comparable to ARC's and Tensilica's is the v-MP2000M, a single-core engine designed for mobile video applications. (The other engine is the triple-core v-MP2000HD, which is intended for more-demanding video applications, such as high-definition TV.) According

to Videantis, the v-MP2000M requires a mere 120,000 gates and 2.61mm^2 of silicon in a 0.13-micron CMOS process. Those figures include instruction and data memories configured for their minimum sizes (16KB for instructions, 4KB for data). The v-MP2000M's maximum worst-case clock frequency is 300MHz, but it needn't run faster than 150MHz while decoding H.264 (baseline profile) at VGA resolution. Typical power consumption for that task is 90mW.

At first glance, the v-MP2000M appears to deliver performance similar to that of the video engines from ARC and Tensilica while requiring less than one-third the silicon. However, the v-MP2000M requires a separate stream processor, particularly for CABAC decoding. It cannot handle both stream processing and pixel processing by itself. Indeed, Videantis expects chip designers to integrate the v-MP2000M with a stream processor that's a general-purpose RISC core, such as those licensed by ARC, ARM, MIPS, or Tensilica. Nevertheless, considering the small size of the v-MP2000M, adding a RISC core for stream processing should yield a fully integrated video engine that's smaller than the engines from ARC and Tensilica.

Moreover, the v-MP2000M is capable of video encoding as well as decoding. (Most video codecs aren't symmetrical; they demand more processing power for encoding a video stream than they do for decoding a stream.) Videantis says the v-MP2000M can encode H.264 video streams (baseline profile) at D1 resolution, 30 frames per second, while running at 250MHz—comfortably below its maximum worst-case clock speed of 300MHz. It can encode MPEG-4 (simple profile) at the same resolution and frame rate while running at only 130MHz.

In contrast, Tensilica says its future video-encoding engine will require additional extensions and two more processor cores, for a total of four cores. ARC says its extended ARC 750D processor has enough clock-frequency headroom to handle encoding or decoding. Of course, increasing the clock speed will use more power.

Numerous Alternatives for Video Processing

As *MPR* has noted in recent articles about ARC and Videantis, the options for video processing are growing rapidly. ARM has introduced its Neon Advanced SIMD Extensions for the ARMv7 architecture, as well as the Cortex-A8 superscalar processor core and OptimoDE coprocessor. (See ***MPR 10/25/05-02***, "Cortex-A8: High-Speed, Low Power," which includes our analysis of Neon, and ***MPR 6/7/04-01***, "ARM's Configurable OptimoDE.") MIPS recently announced the 24KE family of 32-bit processor cores with DSP extensions (see ***MPR 5/31/05-01***, "The MIPS 24KE Family") and a deal that brings Sarnoff Corp.'s synthesizable accelerators to MIPS processors.

Beyond synthesizable IP, many more options abound. Product developers can buy video engines as off-the-shelf ASSPs, as Apple did when designing the new video iPod. (Apple chose Broadcom's VideoCore II BCM2722.) IBM Microelectronics claims the PowerPC-based Cell processor, designed with Sony and Toshiba, is adaptable to many applications, including portable products. (See ***MPR 2/14/05-01***, "Cell Moves Into the Limelight," and ***MPR 2/28/05-01***, "Editorial: Cutting Through Cell's Hype.") On November 8, Philips Semiconductors announced the TM3270, the first low-power TriMedia core—which *MPR* will cover soon. And, of course, all the major consumer-electronics companies use proprietary technology to design their own ASICs and SoCs for internal consumption.

Our summary of the daunting competition shouldn't imply that Tensilica faces a hopeless task finding customers for its new video-decoder engine. Instead, it demonstrates why Tensilica, and every other processor vendor, needs a horse in this race. At a time when video is extending its reach into cellphones, pockets, and vehicles, new opportunities are multiplying, and every processor vendor hoping to stay in business must strive to be among the eventual winners. Tensilica's video engine is a worthy entry.

Price & Availability

Tensilica plans to license its dual-core video-decoder engine in 2006. A quad-core video-encoder engine is also scheduled for release next year. The new Xtensa 6 configurable-processor core is available for licensing now. Tensilica delivers all its processor IP as synthesizable Verilog models with software-development tools optimized for the user's custom configuration. At this time, Tensilica has not publicly announced licensing fees for the video decoder or Xtensa 6 processor. For more information about Tensilica, visit *www.tensilica.com*.