The Linley Group

Tech Insights

MICROPROCESSOR report

Insightful Analysis of Processor Technology

CEVA'S PENTAG2 REDUCES 5G POWER

New "Lite" Version Targets Industrial Systems

By Linley Gwennap (March 28, 2022)

For its second-generation 5G baseband, Ceva is going down, not up. The original PentaG design already scales to 10Gbps, the maximum speed we expect 5G to reach. So for PentaG2, the company focused on reducing energy per bit, limiting heat and stretching battery life for its smartphone customers. The new design is also suitable for use beyond phones, in products such as laptop PCs, tablets, cellular hot spots, and automotive telematics. It can scale down to a "lite" configuration that fits into low-data-rate devices such as asset trackers, smart meters, and remote sensors.

Ceva introduced PentaG in 2018; it remains the only licensable 5G baseband design that provides complete Layer 1 hardware and software. The new version extends the original with additional hardware accelerators that can completely offload data-plane processing from the DSP. These accelerators use less power than running the equivalent functions in software. Initial customers are already designing with both PentaG2-Max and PentaG2-Lite; tapeouts are expected late this year. Production RTL should be generally available in the third quarter.

Unisoc and Samsung licensed Ceva's LTE baseband, but no smartphone customers have announced they're using PentaG. Thus, the DSP licensor is expanding its role beyond handsets, where vendors such as AutoTalks and Sequans already integrate its 5G designs.

As Figure 1 shows, the non-handset cellular market consumed more than 500 million units last year—nearly half as many as the handset market—and is projected to approach 800 million in 2026. Most of the current units employ standard cellular modems, but a growing number use low-data-rate LTE protocols such as Category 1, Category M, and NB-IoT. PentaG2-Lite targets 5G reduced-capability (RedCap) modems, which we expect to replace Category 1 in some devices starting in mid-decade.

PentaG2 Automates Data Plane

The original PentaG featured several hardware accelerators, including polar and LDPC codecs, vector multiply-accumulate (MAC) engines, and an accelerator for FFTs and DFTs (see *MPR 3/12/18,* "Ceva PentaG Adds AI to 5G Baseband"). Although these accelerators offloaded many common 5G operations, the DSP software still had to exe-

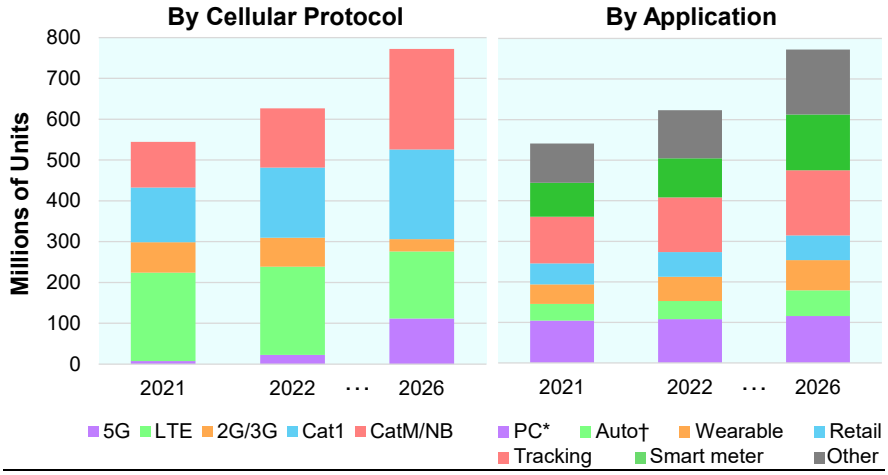

Figure 1. Non-handset cellular-device shipments (worldwide). Growth in asset trackers, vehicle trackers, and smart meters will drive increasing shipments of low-data-rate LTE Category 1, Category M, and NB-IoT devices. *Includes tablets, hot spots, and CPE; †telematics only. (Data source: Techno Systems Research)

Price and Availability

Initial PentaG2-Max and PentaG2-Lite RTL is available to lead customers; Ceva expects general RTL availability in 3Q22. The company withheld license fees. For more information, access *www.ceva-dsp.com/product/ceva-pentag*.

cute the data plane. As a result, the baseband required two high-performance Ceva-XC4500 DSPs to attain 10Gbps throughput.

For PentaG2, Ceva designed a set of hardware units that perform all the basic 5G data-plane functions. This structure creates an autonomous data plane that can process incoming data directly from the radio-frequency (RF) subsystem and convert it to a data stream, or reverse the process when transmitting data. These units pass data from one to the next without software supervision; the transmit and receive units even have DMA engines that move data directly to and from memory.

To start the demodulation (receive) process, the equalizer evens the frequency response for any channel matrix of up to 64x64 elements. The bit demodulator converts the incoming symbols to bits by performing QAM demapping, maximum likelihood decoding (MLD), descrambling, and deinterleaving. The HARQ unit checks for packet errors and indicates when retransmission is necessary. The LDPC and turbo decoders handle forward error correction (FEC). The transmit process reverses this flow, adding FEC codes and modulating the bit stream into symbols for the RF front end.

Customers can use software to configure these units for proprietary techniques; they can change the data flow or even bypass a unit while executing that function in software.

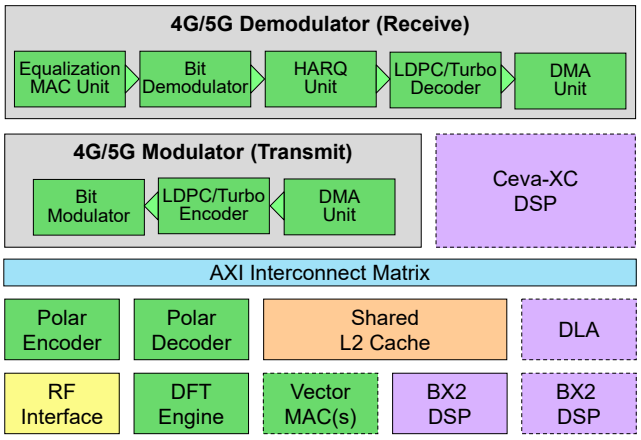

Figure 2. Ceva's PentaG2-Max. The new 5G baseband includes a complete modulator and demodulator in hardware, offloading the data plane from software. The AXI matrix connects all units. PentaG2-Lite omits the dotted blocks and scales down the others.

This approach minimizes power when employing the hardware units, while enabling customer differentiation.

Using AI for 5G

As Figure 2 shows, PentaG2-Max includes a single XC DSP that handles channel estimation in addition to higher functions such as channel tracking and beam management. It carries forward the AI capability of the original PentaG, running a small neural network that monitors channel-state information (CSI) and recommends changes to the channel (such as switching to a different antenna) as needed. To minimize power for this neural network, the design includes a small deep-learning accelerator (DLA) that can run a few billion operations per second. On the transmit side, the XC DSP performs sequence generation before handing off to the hardware data plane.

Two small Ceva-BX2 DSPs act as PHY controllers. Each of the three DSPs has its own instruction and data memories, and they share a larger L2 cache. Customers can configure the size of these memories and even add or remove DSP cores to meet performance and feature requirements. As in PentaG, the DSPs can access additional offload engines: polar encoder/decoders (for FEC on the 5G control channel), a vector coprocessor with 64 MAC units, and a multiradix transform engine that handles FFTs and DFTs of various sizes. PentaG uses the vector and transform engines heavily, but the new design shifts most of their workload to the hardware data plane.

Lite Version Burns Fewer Calories

Whereas the original 5G specification targets high data rates, the 3GPP is now developing specifications for low-cost 5G modems. Specifically, the upcoming Release 17 defines RedCap devices that limit data rates to reduce cost (see *MPR 2/14/22*, "5G Release 17 Simplifies IoT Modems"). For example, whereas most 5G modems must handle 100MHz channels, RedCap modems are limited to 20MHz—the same width as in LTE. Restricting the supported modulations to QAM-64 further simplifies the RedCap design while holding the maximum data rate to 85Mbps, about 1% of the peak 5G data rate.

RedCap offers an upgrade to LTE Category 1 modems, which support up to 10Mbps. LTE Cat-1 deployment is rising, displacing older 2G and 3G devices with faster but inexpensive connections. RedCap can handle HD video and other data-intensive applications, enabling video kiosks, for example.

PentaG2-Lite addresses 5G RedCap designs with a configuration that reduces the die area of the data-plane accelerators to match the lower data rate. It additionally shrinks the shared cache, which provides the HARQ buffer, and it eliminates two of the three DSPs, the DLA, and the vector MAC unit, as Figure 2 shows. The remaining BX2 DSP can not only supervise the data plane but also run the entire 5G protocol stack; to achieve its maximum data

rate, PentaG-Max requires additional CPUs (for example, quad Cortex-A55s) to handle the Layer 2 protocol.

For even lower cost and power, LTE defines Category M and NB-IoT (see *MPR 8/31/15,* "Low-Rate LTE Delivers IoT WANs"), which allow channels as small as 1.4MHz and data rates down to 60kbps. The initial RedCap modems can't compete in this segment, although the 3GPP may address it in Release 18. In the meantime, Ceva offers its Dragonfly baseband design to support these tiny devices (see *MPR 7/9/18,* "Ceva Upgrades NB-IoT Platform").

5G at Lower Power

To support its baseband design, Ceva provides complete reference code for the main data and control channels along with some additional functions, such as channel estimation. This code handles LTE and 5G Release 16, including uplink and downlink plus the main control channel. Licensees must complete the Layer 1 software; they can also modify the reference code to add differentiated capabilities, taking advantage of Ceva's extensive DSP code library and development tools. The vendor provides a baseband simulator model and test code as well.

PentaG and the new PentaG2 are the only licensable 5G basebands on the market. Ceva's main competition is in-house designs; companies such as Qualcomm and MediaTek have enough volume to create their own high-speed basebands, combining off-the-shelf CPUs and DSPs with custom hardware and software. These large suppliers have already made their 5G choices, so that competition is over. Qualcomm is also the leading vendor of high-speed cellular modems in PCs, cars, and similar applications.

PentaG2-Max brings major improvements to Ceva's smartphone customers, adding Release 16 features while greatly reducing power consumption. It represents a major redesign from the original PentaG, moving many data-plane features from software to hardware to save power. Yet it still retains flexibility for customer differentiation and future Release 17 extensions. As chip vendors improve their in-house basebands, Ceva must keep pace to maintain its licensees.

PentaG2-Lite brings an opportunity to broaden Ceva's IoT customer base among both existing LTE licensees and new companies. Because of lower shipments and prices in the IoT market, these vendors have much less revenue than the major smartphone players, so they can't afford to design their own hardware or even a full software stack. As a complete but customizable design, PentaG2-Lite is well suited to these customers. Although we don't expect Release 17 RedCap modems to take off until after 2025, the new baseband design lays the groundwork for that shift. ♦