SITERA SAMPLES ITS FIRST NPU

IQ2000 Programmable Network Processor Targets Edge Routers

By Tom R. Halfhill {5/29/00-02}

Programmable network processors (NPUs) are the newest rage, and one of the latest examples is from Sitera, a four-year-old startup based in Longmont, Colo. Sitera recently began sampling a multiprocessor chip called the Prism IQ2000 and plans to start production in 4Q00.

As with similar NPUs from C-Port, IBM, and Intel, the IQ2000 is intended to replace some of the dedicated ASICs found in routers, switches, and network-gateway devices. Off-the-shelf programmable NPUs are more flexible than ASICs and can dramatically shorten hardware-development cycles. However, these advantages come at the cost of lower performance and longer software-development cycles.

Acquisitions of network-processor companies are even more popular than network processors, and Sitera is keeping up with the Joneses in this way too. Vitesse Semiconductor recently bid $750 million in common stock for Sitera in a deal that's expected to conclude this quarter. The announcement came shortly after Motorola's successful $430 million bid for C-Port (see *MPR 3/6/00-03*, "Motorola Buys C-Port: Smart Move").

Vitesse, based in Camarillo, Calif., qualifies as a veteran in the communications-chip business, having been founded in 1984. The company makes a wide range of digital and analog devices, including framers and PHYs (physical-layer chips). The merger seems like a good fit, because it will allow Vitesse to offer a much broader product line to router vendors and similar customers. Vitesse also has two fabs—one in Camarillo and another in Colorado Springs, Colo. In contrast, Sitera is a fabless company.

Vitesse won't manufacture the IQ2000 in its own fabs though, because both plants use highly specialized GaAS processes. Instead, Sitera will go ahead with plans to use Taiwan-based UMC as its foundry. Vitesse also relies on independent foundries for CMOS manufacturing (about half of Vitesse's designs are built in CMOS). UMC will fabricate the IQ2000 in a 0.25-micron CMOS process and package it in a 673-pin BGA. Sitera is pricing the 200MHz chip at $250 in 1,000-unit quantities.

Among Sitera's first customers are Quarry Technologies and Nortel, although the announcement concerning the latter doesn't specifically name the IQ2000 as the design win.

Living on the Edge

So-called edge routers are Sitera's primary targets. Unlike core routers, which manage traffic deep inside a network, edge routers handle packets at the boundaries or aggregation points between different parts of a network. One edge might be the connection between a metropolitan-area network and a local loop. Another might be the point where a large enterprise LAN enters the public data network.

Sitera's approach is similar to that of most other companies making NPUs. The IQ2000 doesn't attempt to usurp the control functions of the RISC microprocessors that typically accompany dedicated ASICs on a line card. Instead, Sitera designed the IQ2000 to handle packet-switching data. The IQ2000 operates as a coprocessor that's optimized for the specialized tasks of packet filtering and forwarding. A conventional microprocessor (usually a MIPS or PowerPC chip) executes the control code, while the IQ2000 handles the heavy lifting.

While core routers are generally designed to forward packets toward their destinations as quickly as possible, with little or no regard to the packets' contents, edge routers are increasingly expected to do some intelligent packet filtering in addition to routing. This activity requires edge routers to look deeper into the packets instead of simply glancing at the headers. That means an edge router's ASICs or NPUs must be fast enough to analyze packets at wire speed—which varies according to the edge for which the router is designed. The closer to the core of the network, the higher the wire speed.

Sitera designed the IQ2000 for edge routers whose wire speeds range from a pedestrian 64Kb/s (DS0, the base rate for a single analog voice channel over POTS) to a much more demanding 2.488Gb/s (OC-48, a SONET standard for optical-fiber backbones). To span such a wide range of performance, a line card can link as many as eight IQ2000 chips (each with four cores), along with a control CPU, up to 1G of 800MHz Rambus RDRAM, and additional coprocessors and peripherals. The VxWorks operating system that runs on the IQ2000 currently limits the number of processors to four, however, and Sitera says no customer to date has needed more than that. Figure 1 shows how the IQ2000 works together with those components.

Numerous Interface Options

The key to the IQ2000's scalability is a multitude of I/O interfaces. Each chip has a 16-bit 800MHz RDRAM channel that
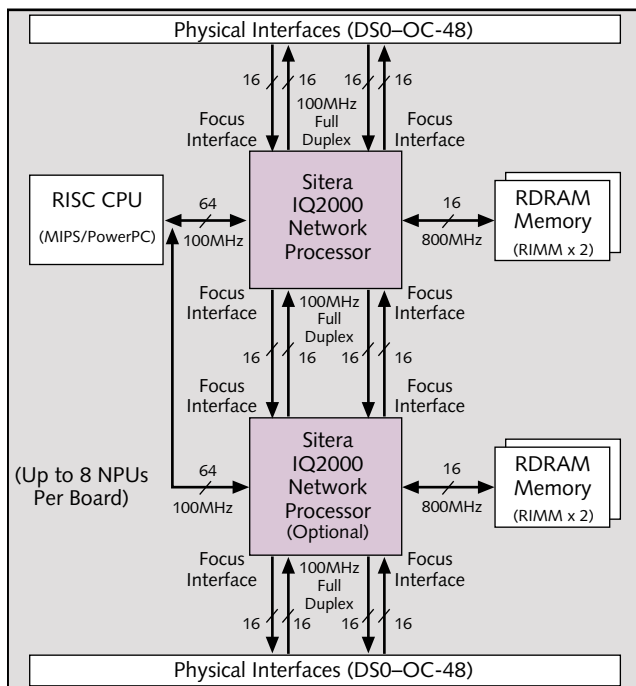
Figure 1. Sitera's IQ2000 offloads some packet-processing chores from a general-purpose CPU, which runs the control code. Multiprocessor boards can link up to eight IQ2000 chips together, although four is the practical limit with existing software. Packet payloads wait in RDRAM while the IQ2000 processes the packet headers.

supports one or two RIMMs of any size; a 64-bit 100MHz CPU bus that gluelessly supports the MIPS SysAD interface; and up to four of Sitera's own Focus interfaces.

Focus is a Utopia-like 16-bit full-duplex interface (16 bits per direction) that runs at half the 200MHz core speed. This provides up to 12.8Gb/s of bandwidth. Although Focus is a proprietary interface, Sitera openly publishes the specifications and doesn't levy any licensing fees or royalties on companies that want to design Focus-compatible coprocessors and peripherals. One partner is Fast-Chip (Los Altos, Calif.), which is designing a Focus interface for its policy-engine coprocessor. Sitera says other customers are building Focus interfaces into their product-specific FPGAs and ASICs.

Ideally, NPUs such as the IQ2000 would use an industry-standard interconnect that makes it easy to link chips from different vendors. Unfortunately, none exists that exactly fits the technical requirements and has broad industry support. RapidIO (see *MPR 5/8/00-01*, "RapidIO Expands Narrow-Bus Options") is a strong candidate, but it wasn't even announced until after Sitera began sampling the IQ2000. If RapidIO or another specification ever becomes a popular standard for connecting components in networking equipment, it would make more sense for Sitera to switch than to continue moving in an entirely different direction with future generations of the processor.

The IQ2000 is actually a family of three chips with configurable Focus interfaces. The standard configuration has four 16-bit Focus ports. Other configurations combine pairs of those channels into one or two 32-bit Focus ports; still other configurations substitute one or two Gigabit Ethernet media-access controllers (MACs) for one or two Focus ports. Table 1 shows the different combinations possible with the members of the IQ2000 family.

By offering three similar chips with several port-configuration options, Sitera can target many different applications with the IQ2000. For instance, the Gigabit Ethernet ports (which require external PHY chips) make the IQ2000 suitable for server load-balancing tasks within LANs. An IQ2000-based router could distribute incoming packets among several servers, according to programmable priorities. A router at an e-commerce business could steer incoming packets from preferred customers to the fastest available Web server while relegating packets from anonymous window-shoppers to a slower server. Or the router could steer packets carrying secure HTTP traffic (representing an online credit-card transaction in progress) to a special server dedicated to that purpose.

Similarly, an IQ2000 chip could be configured with two 16-bit and one 32-bit Focus ports in an edge router or gateway device that aggregates multiple lower-bandwidth channels into a single higher-bandwidth pipe. The Focus ports are also the channels for hooking multiple IQ2000 chips together (see Figure 1), allowing many different port configurations on multiprocessor boards.

Familiar Cores Ease Programming

All members of the IQ2000 family share the same basic core, which actually consists of four R3000-like CPU cores surrounded by special task engines. Sitera patterned its CPU cores after the MIPS-I architecture, because MIPS is well understood; widely used in networking; and amply supported by development tools that need only a few modifications to work with the IQ2000.

C-Port's C-5 processor also contains multiple CPU cores patterned after the MIPS-I architecture. Neither Sitera nor C-Port has a MIPS license. But they don't support MIPS Technologies' patented unaligned load and store instructions or claim to be fully MIPS compatible, so perhaps they won't suffer the same legal difficulties as Lexra, which is battling MIPS in federal court (see *MPR 12/6/99-03*, "MIPS vs. Lexra: Definitely Not Aligned").
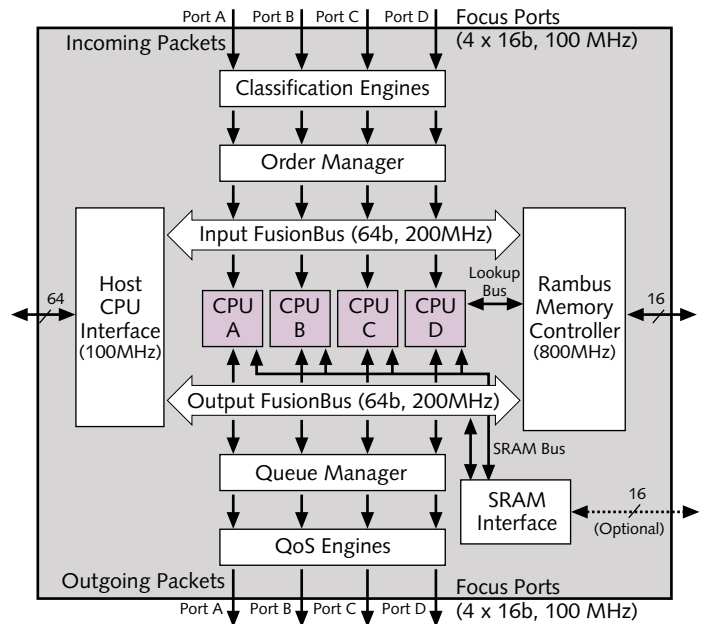
Figure 2 is a flow-oriented block diagram of the IQ2000 that shows the CPU cores and supporting logic. The classification engines are function units that can classify incoming packets according to programmable priorities, but they aren't intended to replace separate chips that do more sophisticated chores, such as encryption and decryption for virtual private networks (VPNs). The order manager tags the packets so the IQ2000 can retire them in their original order. (The IQ2000 can process packets out of order but always retires them in order.)

The Streaming Fusion bus is Sitera's on-chip bus. It's 64 bits wide, runs at the full core frequency, and handles multiple transactions simultaneously. It links the CPU cores to the external CPU interface and the integrated RDRAM controller, which is a glueless interface to external RIMMs. The IQ2000 temporarily stores the packet payloads in RDRAM while processing the headers (typically 64–80 bytes long). Other NPUs park the payloads in SRAM or SDRAM, but Sitera believes RDRAM has a better combination of speed and cost.

On the output side of the core is the queue manager, which restores the packets to their original order, and the quality-of-service (QoS) engines, which have programmable policies for packet forwarding. This is where some packets get preferential treatment. For instance, a router in a corporate WAN might give higher priority to internal email

Figure 2. The IQ2000 has four CPU cores, highlighted in purple in this block diagram. The on-chip Fusion bus connects the CPU cores to the RDRAM controller, host CPU interface, and packet-processing function units. Packets go in and out over the four Focus ports, which appear in this diagram as separate groups of ports but are actually full-duplex interfaces.

than to messages arriving from outside the firewall. The QoS engines also reunite the packet payloads with the headers before moving them off chip over the Focus ports.

Transparent Multithreading

At the center of the action are the four CPU cores. They are 32-bit scalar cores with 64-bit memory interfaces, so each CPU can perform a double-word load or store in a single clock cycle.

Each core has five identical register files (32 registers, 32 bits wide, triple-ported). This arrangement allows each core to run five concurrent threads of execution with fast context switching, because they don't have to save their register states in memory. The CPUs automatically switch threads (transparently to application programmers) while waiting for RDRAM accesses. One thread is set aside for a kernel process, leaving the other four threads for packet processing.

Other resources are duplicated as well. Each CPU core has 4K of triple-ported SRAM that's divided into a 2K header buffer and a 2K data buffer. The header buffer has room for 16 entries of 128-byte headers. The CPUs share their data buffers with each other—each buffer has a 512-byte region for each CPU. It's polite to share, because the IQ2000's CPUs operate as a pool of processors for incoming packets. Any available CPU can handle a packet from any input port.

Sitera says pooling is more efficient than dedicating a CPU to each port, as C-Port's C-5 does. Pooled CPUs stay busy more often. One drawback, however,

IQ2000 Network Processors			Focus Port Configurations			
S21100	S21102	S21132	Focus A	Focus B	Focus C	Focus D
•	•	•	16-bit I/O	16-bit I/O	16-bit I/O	16-bit I/O
•	•	•	16-bit I/O	16-bit I/O	32-bit I/O	
------	•	•	GigaMAC	16-bit I/O	16-bit I/O	16-bit I/O
------	•	•	GigaMAC	16-bit I/O	32-bit I/O	
------	•	•	GigaMAC	GigaMAC	16-bit I/O	16-bit I/O
------	•	•	GigaMAC	GigaMAC	32-bit I/O	
------	------	•	32-bit I/O		16-bit I/O	16-bit I/O
------	------	•	32-bit I/O		32-bit I/O	

Table 1. Sitera makes three different versions of the IQ2000 that support different configurations of the Focus ports. The S21132 part supports all configurations, while the S21100 and S21102 are more limited.

is that pooling creates dependencies between the CPUs, because a stream of related packets can flow through multiple CPUs at the same time. Dedicating a CPU to each port ensures that independent packet streams stay apart. Both approaches have their merits; Intel's IXP1200 (see *MPR 9/13/99-01*, "Intel Network Processor Targets Routers") pools its CPUs in a way that's similar to the IQ2000's.

To accelerate the tasks of separating headers from payloads and analyzing the headers, Sitera added some byte-manipulation and bit-extraction instructions that aren't found in the MIPS-I instruction set. Another twist on MIPS is that branch instructions can be adjacent—if a program takes the first branch, the CPU nullifies the second branch instruction in the delay slot. A static branch predictor assumes branches won't be taken, because branches in packet-processing code usually test for errors or special conditions that rarely occur.

Comparing NPUs, CPUs, and ASICs

Performance comparisons among NPUs, CPUs, and fixed-function ASICs are almost impossible, due to the lack of applicable benchmarks and the difficulty of defining the problems. Although the EDN Embedded Microprocessor Benchmark Consortium (EEMBC) has both a networking suite and a telecommunications suite, no vendor has yet published test results for an NPU, and the existing tests are limited in scope (see *MPR 5/1/00-02*, "EEMBC Releases First Benchmarks").

Some vendors quote the maximum number of packets per second their NPUs can process. But that number depends on how deeply the processor is looking into the packets and what kind of filtering it's doing. The number is higher for Layer 2 and Layer 3 routing, but those layers are more applicable to backbone routing than to the edge-routing tasks for which Sitera designed the IQ2000. The industry needs to define benchmarks for Layer 4+ routing and for common packet-filtering tasks, such as server load balancing and quality-of-service routing.

One objective comparison is the I/O bandwidth available on an NPU. The IQ2000 has significantly more bandwidth than Intel's IXP1200 or C-Port's C-5. The IQ2000's four full-duplex Focus ports provide 1.6GB/s of packet bandwidth, which is 3x more than the IXP1200 and 2.5x more than the C-5. The IQ2000's Rambus interface provides 1.6GB/s of memory bandwidth, which is 2.4x more than the IXP1200 and 8x more than the C-5. (The Intel and C-Port chips use SDRAM instead of RDRAM.)

Although the IQ2000 has four CPU cores, their 200MHz scalar pipelines make them only about one-fourth as fast as the standalone superscalar microprocessors that will be available to router vendors at about the same time. For example, QED's new RM7000A runs at 400MHz; has a dual-issue superscalar core; is compatible with the MIPS software widely used by router vendors; is scheduled to ship a few months before the IQ2000; and costs 10% less (see *MPR 4/3/00-05*, "QED's RM7000A Gets Faster, Cooler"). But a general-purpose microprocessor like the RM7000A lacks the specialized packet-processing engines and other enhancements found in the IQ2000.

Sitera and QED obviously don't perceive each other as direct competitors. At the recent Networld+Interop trade show in Las Vegas, the two companies demonstrated a reference board that combines an IQ2000, a 450MHz RM7000A, and a packet-classification coprocessor from a third partner (Fast-Chip). The reference board also uses some of Sitera's peripheral chips, such as the OctalMAC (which provides eight 10/100Mb/s Ethernet ports) and a Gigabit Ethernet controller. The RM7000A runs the control code, while the IQ2000 runs the packet-processing code.

Router vendors that are using general-purpose CPUs for packet processing must balance the potential advantages of NPUs against the costs of designing new boards around unfamiliar chips and rewriting at least part of their software. Sitera says a typical edge router based on the IQ2000 would require about 4,000 lines of assembly-language code. That assumes the IQ2000 is handling the packet processing while a general-purpose CPU handles the control functions. Sitera provides reference code that customers can modify and extend for their own applications.

More factors come into play when one compares NPUs against fixed-function ASICs. In routers that don't need the highest possible performance, programmable NPUs such as the IQ2000 reduce hardware-development costs and provide more flexibility. Programmers can change an NPU-based router's priorities at any time—to reflect either changing business strategies or evolving network requirements. A fixed-function ASIC that carries out the same tasks in hard-wired logic would almost certainly be faster but is less versatile. And, of course, ASICs take a year or more to develop, at no small expense, whereas programmable NPUs are available off the shelf from a growing number of vendors.

Although NPUs have established a beachhead, it's not yet clear how much of the market they'll take away from CPUs and ASICs. Sitera's IQ2000 is a good example of the breed and should be able to hold its own against the competition. ◇