

M I C R O P R O C E S S O R

www.MPRonline.com

THE INSIDER'S GUIDE TO MICROPROCESSOR HARDWARE

CAVIUM STALKS STORAGE

Coming Soon: The First Octeon Storage Processors

By Tom R. Halfhill {7/16/07-01}

Cavium Networks is entering the mainstream storage-processor market with two families of Octeon chips based on the company's successful networking and communications processors. When the new storage processors debut late this year, they will bring the same

high integration and programmability to networked storage systems that Cavium's existing processors have brought to routers, broadband-access devices, and many other networking products.

The new Octeon Storage Services Processors will have two to twelve MIPS-compatible processor cores per chip, as much as 2MB of L2 cache per core, configurable I/O interfaces, and hardware acceleration for critical tasks. Clock frequencies will range from 600MHz to 1.0GHz, with power consumption ranging from 8W to 30W. Sampling is scheduled to begin in 3Q07, with production following one quarter later.

Cavium's entry into this market is an important development. Although the company is only six years old, it has quickly become a force in network/communications processors—a market it entered in earnest less than three years ago. (See [MPR 10/5/04-01](#), "Cavium Branches Out.") Cavium has won more than 100 customers, including all the top-tier manufacturers of networking and communications equipment worldwide. At a stage when many fabless semiconductor startups are still burning through their private funding, Cavium is debt free, has accumulated \$100 million in cash, and has completed a successful public offering.

New Demands for Networked Storage

There are five reasons for Cavium's rapid rise. First, Cavium's Octeon processors are muscular multicore designs with as many as 16 MIPS-compatible cores per chip. Second, these

chips are highly integrated for their target applications, sporting numerous acceleration engines and flexible I/O interfaces that offload tasks from the main processor cores. Third, Cavium typically introduces broad families of chips derived from a common design, so customers have numerous choices at different prices. Fourth, Cavium's processors are realistically programmable in high-level C or C++, thanks to optimized APIs that handle low-level drudge work. And fifth, Cavium is a productive company—fewer than 200 employees churn out a surprising amount of chips, boards, and software.

Now Cavium is bringing those strengths to bear on the fast-growing storage-processor market. Once the exclusive domain of general-purpose processors, networked storage systems are rapidly becoming more specialized and commoditized. There seems to be an insatiable appetite for more storage, more-secure storage, and faster storage.

New government regulations are mandating the archiving of financial records, corporate emails, and even text messages for certain periods of time. Privacy laws and attacks by malicious hackers require the handlers of sensitive information to carefully guard their data. Mass storage is almost useless without fast access, so storage subsystems must be thoroughly indexed and searchable. High reliability and availability are paramount. Offline storage is becoming impractical, so the storage systems must have fast connections to a network. And the storage market is broad, reaching from multinational enterprises all the way down to

Feature	Cavium Octeon SSP CN5520	Cavium Octeon SSP CN5530	Cavium Octeon SSP CN5534	Cavium Octeon SSP CN5734	Cavium Octeon SSP CN5740	Cavium Octeon SSP CN5745	Cavium Octeon SSP CN5750
CPU Architecture	cnMIPS64r2	cnMIPS64r2	cnMIPS64r2	cnMIPS64r2	cnMIPS64r2	cnMIPS64r2	cnMIPS64r2
CPU Cores	2	4	6	6	8	10	12
Core Frequency	600-, 800-, or 900MHz	600-, 800-, or 900MHz	600-, 800-, or 900MHz	600-, 800-, or 900MHz	600-, 800-, or 900MHz	600-, 800-, or 900MHz	600-, 800-, or 900MHz, 1.0GHz
Max. Instr./Sec	2.4–3.6 billion	4.8–7.2 billion	7.2–10.8 billion	7.2–10.8 billion	9.6–14.4 billion	12–18 billion	14.4–24 billion
L1 Cache (I/D)	32K / 16K	32K / 16K	32K / 16K	32K / 16K	32K / 16K	32K / 16K	32K / 16K
Write Buffer	2K	2K	2K	2K	2K	2K	2K
L2 Cache	512K + ECC	512K or 1MB + ECC	1MB + ECC	1MB + ECC	2MB + ECC	2MB + ECC	2MB + ECC
DRAM Interfaces (DDR2-800)	1x64 bits + ECC	1x64 bits + ECC	1x64 bits + ECC	1x64 bits 2x64 bits 1x128 bits + ECC	1x64 bits 2x64 bits 1x128 bits + ECC	1x64 bits 2x64 bits 1x128 bits + ECC	1x64 bits 2x64 bits 1x128 bits + ECC
SERDES	12	12	12	16	16	16	16
PCI Express	2 controllers Max 8 lanes	2 controllers Max 8 lanes	2 controllers Max 8 lanes	2 controllers Max 16 lanes	2 controllers Max 16 lanes	2 controllers Max 16 lanes	2 controllers Max 16 lanes
Packet Interfaces	8x SGMII or 1x XAUI + 4x SGMII	8x SGMII or 1x XAUI + 4x SGMII	8x SGMII or 1x XAUI + 4x SGMII	8x SGMII or 2x XAUI + 4x SGMII	8x SGMII or 2x XAUI + 4x SGMII	8x SGMII or 2x XAUI + 4x SGMII	8x SGMII or 2x XAUI + 4x SGMII
Crypto Engine	Optional	Optional	Optional	Optional	Optional	Optional	Optional
RAID Engine	RAID5/6	RAID5/6	RAID5/6	RAID5/6	RAID5/6	RAID5/6	RAID5/6
TCP Engine	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DMA Engines	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Compress/Decompress Engine	Yes (Huffman, hashing)	Yes (Huffman, hashing)	Yes (Huffman, hashing)	Yes (Huffman, hashing)	Yes (Huffman, hashing)	Yes (Huffman, hashing)	Yes (Huffman, hashing)
Power	8W (600MHz)	n/a	n/a	n/a	n/a	n/a	30W (1.0GHz)
Package	BGA-800	BGA-800	BGA-800	BGA-1217	BGA-1217	BGA-1217	BGA-1217
Price (10k Units)	\$59	n/a	n/a	n/a	n/a	n/a	\$575
Availability	Sample 4Q07 Prod. 1Q08	Sample 4Q07 Prod. 1Q08	Sample 4Q07 Prod. 1Q08	Sample 3Q07 Prod. 4Q07	Sample 3Q07 Prod. 4Q07	Sample 3Q07 Prod. 4Q07	Sample 3Q07 Prod. 4Q07

Table 1. Feature comparison of Cavium's Octeon CN55xx and CN57xx Storage Services Processors. The key differences among these chips are the number of processor cores, their clock speeds, main-memory interfaces, and packet interfaces. The integrated cryptography engine is optional, so there are 14 distinct parts, not counting the different speed grades and cache options. Note that theoretical maximum performance (in the number of instructions executed per second) varies by a factor of ten from the low end to the high end, commensurate with pricing, but power consumption varies by a factor of less than four. (n/a: data not available)

individual users who are accumulating terabytes of digital photos, videos, and music.

Together, these requirements are pushing general-purpose processors aside. But programmability and flexibility are still important, making fixed-function ASICs unattractive. The increasingly popular solution is a new breed of ASSPs optimized for networked storage applications.

Seven Derivatives, One Basic Design

Cavium refers to its new Octeon storage chips as Storage Services Processors (SSP). It's a logical name extension from Cavium's Octeon Network Services Processors (NSP). (For an explanation of Cavium's naming conventions, see [MPR 2/6/06-01](#), "Cavium Expands Octeon Family.") The implication is that these chips provide specialized services for their target markets, using chip-level integration.

Typical services for an SSP include compressing data, encrypting data, and scanning data for telltale signs of malware. Disk virtualization makes multiple drives appear as a single large drive. Another relatively new service is de-duplication, which optimizes storage by eliminating redundant data. For

instance, an email archive might keep only one copy of a message that went to multiple recipients, with pointers replacing message bodies in the duplicate messages. Cavium has designed its Octeon SSPs to accelerate all these services.

Cavium is announcing seven Octeon SSPs in two families, the CN55xx and CN57xx series. Actually, there are dozens of variations, counting the different speed grades, cache sizes, and other options. For instance, the integrated cryptography engine is optional, because some customers don't need encryption, and U.S. export regulations restrict the sale of chips with crypto engines to some foreign countries. However, all devices in the CN55xx and CN57xx families are based on one fundamental design. All use Cavium's cnMIPS64r2 processor core, a custom two-way superscalar core adhering to the MIPS Technologies MIPS64 Release 2 instruction-set architecture. The low-end Octeon CN5520 SSP has two cores, while the high-end Octeon CN5750 SSP has twelve cores. Other members of the Octeon SSP families have four, six, eight, or ten cores. They will be available in speed grades of 600MHz, 800MHz, 900MHz, and 1.0GHz.

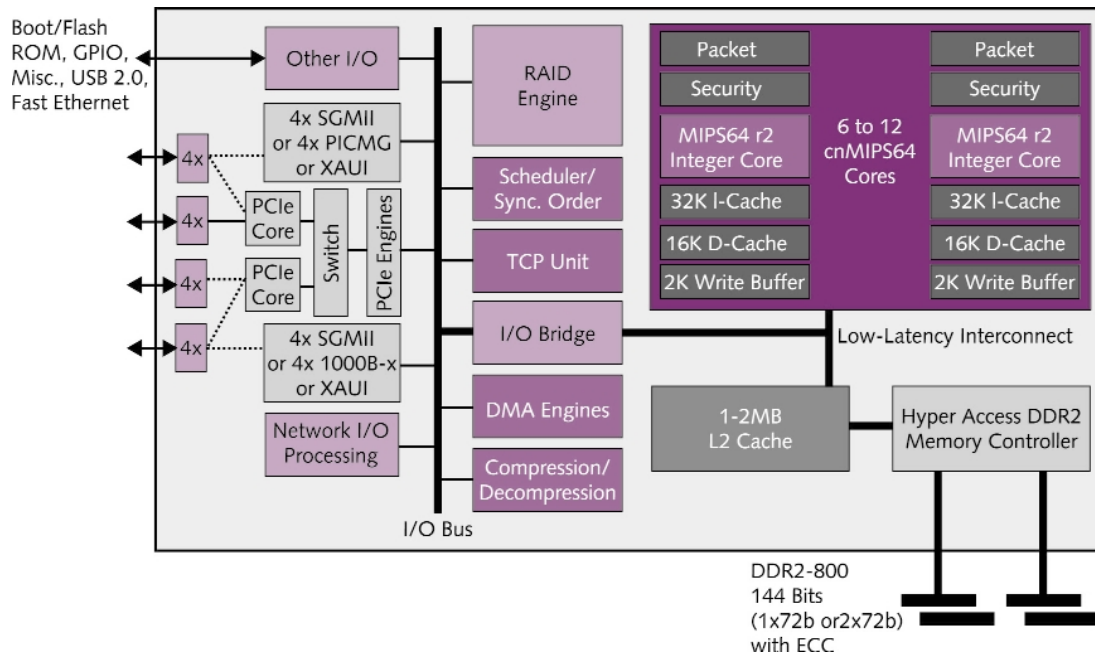


Figure 1. Oocteon SSP CN57xx block diagram. Note the exceptional integration of these devices—several special-purpose engines supplement the MIPS CPUs. The PCI Express, serdes, SGMII, and XAUI interfaces are highly configurable. They can support eight PCI Express lanes in each direction, or eight lanes in one direction, four lanes in another, plus four Gigabit Ethernet ports or one 10-gigabit Ethernet port. Another possible configuration allows four PCI Express lanes in each direction, plus eight Gigabit Ethernet ports or two 10-gigabit Ethernet ports. The CN55xx family is similar to the CN57xx but has fewer MIPS cores, smaller L2 caches, a single 64-bit DDR2 memory controller, fewer PCI Express lanes, and less I/O flexibility.

The number of processor cores and their clock speeds are the most important features distinguishing these chips. At the low end, the 600MHz dual-core CN5520 can execute a theoretical maximum of 2.4 billion instructions per second (2 cores \times 2 pipelines/core \times 600MHz). At the high end, the 1.0GHz 12-core CN5750 can execute a theoretical maximum of 24 billion instructions per second (12 cores \times 2 pipelines/core \times 1.0GHz). That's a performance spread of 10x. Although the processors will rarely achieve these peak throughputs in real-world systems, the general case is that more cores and higher clock rates should improve performance in a roughly linear fashion.

Other features distinguishing Oocteon SSP chips from each other are their L2 caches (512KB to 2MB), main-memory interfaces (one or two DDR2-DRAM controllers at 800MHz), serializers/deserializers (12 or 16 serdes), PCI Express interfaces (maximum 8 or 16 lanes), and network-packet interfaces. The PCI Express and packet interfaces are particularly flexible—programmers can configure them in several ways. All the chips have two integrated PCI Express controllers, each supporting up to four or eight lanes. All the chips have eight Serial Gigabit Media-Independent Interfaces (SGMII), which can provide up to eight Gigabit Ethernet ports. CN55xx chips have one X-Attachment User Interface (XAUI), which supports a 10-gigabit Ethernet port. CN57xx chips have two XAUIs. Table 1 summarizes the features of the Oocteon SSP CN55xx and CN57xx families.

Other variables are chip packages, power consumption, and prices. The CN55x family will be available in 800-pin BGAs. The CN57xx family requires a 1,217-pin BGA to accommodate the wider DRAM interface, additional serdes, and extra XAUI. Power consumption and price differences are great, as one might expect, given the disparity in processor cores and clock rates. The dual-core 600MHz CN5520 will dissipate 8W (Cavium's estimate) and cost \$59, whereas the 12-core 900MHz CN5750 will dissipate about 27W (MPR's estimate) and cost nearly ten times as much at \$575. Cavium hasn't disclosed power estimates and prices between those extremes, but the distribution should be fairly linear. Samples of the higher-end CN57xx family are scheduled to be available first, in 3Q07, with general availability in 4Q07. Samples of the lower-end CN55x family are scheduled for 4Q07, with general availability in 1Q08.

Integrated Engines Offload the CPUs

Cavium is famous for its high degree of chip-level integration, and the new Oocteon SSPs uphold that tradition. Several hardware-acceleration engines offload routine work from the MIPS processor cores, freeing them for other tasks. In effect, these chips are asymmetric multiprocessors, in addition to being symmetric multicore designs. To ease programmability, Cavium supplies API libraries for the integrated engines. Programmers call high-level functions from C or C++ instead of tediously programming the engines in a proprietary machine language.

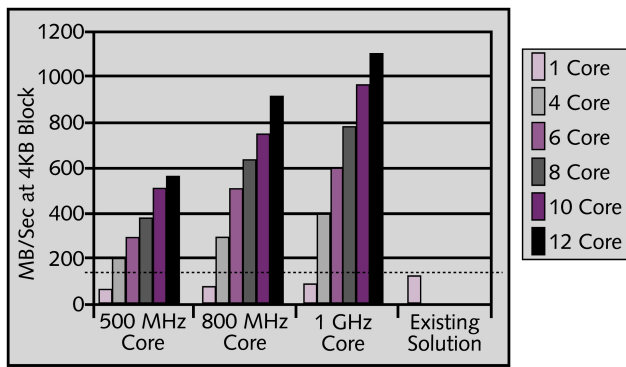


Figure 2. Oocteon SSP sequential-read performance with iSCSI. In this test, throughput is remarkably linear with the number of MIPS processor cores on chip. It's natural for Cavium to choose a favorable benchmark test, but storage processing does offer rich opportunities for parallelism, so these results should be indicative of actual performance on some tasks. (Source: Cavium)

Moving data is the basic job of a storage processor, so multiple DMA engines handle the chores of shunting data from the various network interfaces to the MIPS cores and main memory, then back to the network again. Multiword direct-memory access (MWDMA) allows the DMA engines to rapidly transfer critical data to external nonvolatile memory in the event of a power failure. The DRAM controllers and L2 caches support ECC and other data-protection features not found in Cavium's network processors—dropped packets are replaceable, but lost data is not. Even the chip's internal buses have ECC and parity protection. The L1 caches automatically correct single-bit errors.

The Oocteon SSP RAID engine provides hardware acceleration for RAID5, RAID6, and data de-duplication. (RAID6 allows two simultaneous disk failures without losing data.) Accelerators in the RAID engine speed up exclusive-OR (XOR) operations and Galois-field calculations. Cavium claims its RAID engine is the best available, supporting RAID processing at the full line rate of the network (500MB/s to 10GB/s, depending on the particular Oocteon SSP).

Data compression is a common method of conserving storage capacity and network bandwidth, but it's a compute-intensive job. Cavium's compression/decompression engine offloads most of that work from the CPUs and is programmable for different compression standards and modes. It can optimize for high compression rates (up to 1GB/s) or high compression ratios, depending on the application. Hardware accelerators speed up Huffman algorithms and hashing functions, and special buffers store the token dictionaries created by Huffman-compression programs like WinZip. Cavium says using its compression engine is 20 to 25 times faster than executing the same functions on the MIPS cores or other general-purpose processors.

Data security is a growing concern, so the integrated crypto engine offloads the compute-intensive tasks of encryption and decryption. Oocteon SSPs have essentially the

same crypto engine as Oocteon NSPs—an engine derived from Cavium's standalone Nitrox security chips. (See [MPR 9/6/05-01](#), "Cavium: Security Optional," and [MPR 2/7/05-01](#), "Cavium Expands Security.") One important difference is that the Oocteon SSP crypto engines accelerate the new Advanced Encryption Standard (AES) XTS mode (IEEE P1619), which can encrypt a data file without expanding it. Additional hardware accelerators in Oocteon SSPs offload Transmission Control Protocol (TCP) termination (up to 20Gb/s), checksum calculations, 32-bit cyclic redundancy checks (CRC), and I/O scheduling and synchronization. Figure 1 is a block diagram of the Oocteon SSP CN57xx.

Meeting the Multicore Challenge

Multicore processors are ubiquitous these days, but the real challenge isn't slapping down multiple cores on a chip. The industry is wringing its hands over the problem of fully exploiting all the cores. Some applications simply don't lend themselves to parallel processing—or at least, they make parallelism very difficult. Other applications are embarrassingly parallel. Storage processing leans toward the latter category. Many functions, such as compression, can execute in parallel on multiple cores that operate on different parts of a datastream.

One test for efficient parallelism is to compare the throughput of chips that have different numbers of processor cores performing the same task. In Figure 2, Cavium has provided the results of one such test: sequential reads from an iSCSI disk. As might be expected, this particular test shows Oocteon SSPs in a favorable light. Nevertheless, it's good to see throughput scaling in an almost linear progression with the number of processor cores. Although throughput does taper off as the multicore integration climbs from two to twelve, the loss isn't severe, and the test shows that a 12-core Oocteon SSP isn't a fruitless exercise in overengineering. (Oocteon NSPs have as many as 16 cores, but Cavium says the slightly greater performance of a 16-core Oocteon SSP wouldn't be worth the additional power consumption and cost. This statement suggests that packet processing is more scalable than storage processing, at least with this multicore design.)

Another aspect of Cavium's highly integrated designs is that they reduce the system's total chip count. With their numerous I/O interfaces and integrated acceleration engines, Oocteon SSPs will eliminate the need for some auxiliary chips while shrinking the motherboard, saving power, and cutting costs. The only notable omissions in Oocteon SSPs are physical-layer (PHY) interfaces for their I/O controllers. Most other networking and communications processors omit PHYs, too, because PHY chips are small and cheap. But some processors are beginning to absorb the PHYs, further reducing the chip count and permitting smaller boards.

Storage Processors Evolving Fast

Storage processors are relatively new; MPR has covered only a few examples in recent years. (See [MPR 11/27/06-02](#), "Virtually Smart Storage Processing," and [MPR 5/2/05-02](#),

“Storage Processor Leverages LEON.”) As usual when a new application emerges, general-purpose embedded processors met the initial need, because they are widely available, affordable, and programmable. But although they may have multiple superscalar processor cores surpassing 1.0GHz, they are not high-performance solutions in the long term. They lack the hardware acceleration and I/O interfaces that a specialized application demands. (See *MPR 6/11/07-01*, “Trends in General-Purpose Processors.”)

The next step has been to press other types of network/communications processors into service as storage processors. AMCC, Broadcom, Freescale Semiconductor, Intel, Marvell, PMC-Sierra, Raza Microelectronics, and others are doing this to some degree. It’s a step in the right direction, because at least these types of processors have some of the network interfaces and hardware accelerators required. But ultimately, as the market grows, designers will take the final step of creating entirely new processors optimized for storage systems.

Cavium’s new Octeon SSPs are somewhere between steps two and three. They aren’t entirely new designs, but they are based on Cavium’s NSPs, which already have most features a storage processor requires. By adding features for greater data security and reliability, Cavium has created a new product line that can legitimately claim to be specialized for the target application.

Just in time, because other companies are working in similar directions. At the recent Microprocessor Forum in

Price & Availability

Cavium Networks plans to deliver its new Octeon Storage Services Processors late this year and early next year. Samples of the higher-end CN57xx family will be available first, in 3Q07, with production scheduled for 4Q07. Samples of the lower-end CN55x family will be available in 4Q07, with production scheduled for 1Q08. Prices in 10,000-unit quantities will range from \$59 at the low end (dual-core CN5520, 600MHz) to \$575 near the high end (12-core CN5750, 900MHz).

For more information visit:

- www.caviumnetworks.com/newsevents_Cavium-networks_OCTEON-SSP.html

San Jose, AMCC unveiled its new Titan core—a Power Architecture core that will eventually supplant the IBM Power 405 and Power 440 cores that AMCC uses in storage processors today. And Freescale recently announced a new multicore strategy for its future network/communications processors—a strategy that encompasses storage processors. Until those plans come to fruition, Cavium’s new storage processors will be strong competitors that set a new standard for versatility and integration. ♦

To subscribe to Microprocessor Report, phone 480.483.4441 or visit www.MPRonline.com