

MICROPROCESSOR *report*

Insightful Analysis of Processor Technology

ADESTO EXECUTES IN PLACE

New EcoXIP Memory Simplifies IoT Design

By Linley Gwennap (October 10, 2016)

Adesto is a small memory supplier with big plans. After introducing an innovative nonvolatile memory earlier this year, it has taken standard NOR flash and added a new high-speed interface designed specifically for streaming instructions—a technique that designers call execute in place (XIP). The new EcoXIP product is now sampling in 32Mb (4MB) capacity, with production expected in 1Q17. Additional capacity options will follow.

Many small systems employ a microcontroller with embedded flash memory that holds the application code. When these devices add a radio for IoT capability, they require larger storage to hold the wireless protocol and IP stack as well as security software. MCUs normally top out at 1MB of internal flash, but IP-based IoT devices often need more code space, requiring an external flash device.

Commodity flash chips connect through a low-speed SPI, requiring the MCU to copy the code into a large internal SRAM to maintain reasonable performance. Many systems execute directly from the external flash (XIP); they are not only slower but may require a second flash chip to support over-the-air (OTA) code updates. In an XIP design, even writing data (such as log information) to flash can be challenging.

Announced at the recent Linley Processor Conference, EcoXIP solves these problems by enabling simultaneous read and write transactions. It can deliver instructions at a sustained rate of 156MB/s (266MB/s peak), which is fast enough for most MCUs and better than other XIP memories. Because it uses a modified SPI to improve performance, however, the new memory works only with compatible MCUs. At the conference, Adesto CTO Gideon Intrater disclosed that NXP, a leading MCU supplier, will support the EcoXIP interface in future MCU products.

Accelerating the Bus

To more efficiently implement XIP, Adesto redesigned the basic SPI protocol to better handle the typical access patterns. SPI is designed for random accesses; it returns the requested cache line (e.g., 16 bytes), then waits for the next request. This approach works well for data storage, but instruction fetches tend to be sequential as the CPU proceeds through a block of code. Therefore, EcoXIP continues to provide sequential bytes until it receives a new command. It calls this approach “command fusing.”

This approach can double the bus throughput, as Figure 1 shows. Using an octal (8-bit) SPI, a transaction typically requires 1 bus cycle for the command, 2 cycles for the address, about 14 cycles to wait for the response, and 8 cycles to transmit 16 bytes of double-clocked data (DDR). Fetching the next 16 bytes requires another 25 bus cycles, or 50 for the two transactions. Using EcoXIP, the first 16 bytes take the same number of cycles, but data then continues to flow, delivering four cache lines in 49 cycles.

The benefit of Adesto’s approach comes when the CPU executes a sequential set of instructions. For a 32-bit MCU, each line holds four instructions, and a branch occurs about every seven instructions; about half of all branches are taken. The company estimates that the average number

Octal SPI at 133MHz



EcoXIP at 133MHz



Figure 1. EcoXIP bus timing. C/A=command/address. By chaining data responses using “command fusing,” the Adesto design can deliver twice as many cache lines in the same number of bus cycles. (Source: Adesto)

Price and Availability

Adesto is currently sampling a 32Mb EcoXIP product to lead customers; it expects to enter production in 1Q17. The company withheld pricing. To download a free copy of the Adesto presentation from the Linley Processor Conference, access www.linleygroup.com/processor-conference. For more information on EcoXIP, access www.adestotech.com.

of line fetches per instruction-cache miss is 3.84, or nearly 14 instructions (the first cache line may have fewer than four useful instructions if the target is in the middle of the line). Using this average, Adesto calculates the sustained throughput of its 133MHz EcoXIP at 156MB/s and the average latency at just 57ns. When the CPU reaches a taken branch, it sends a new request to the EcoXIP, which then begins transmitting data from the new address.

Most flash chips have a quad SPI to reduce pin count and cost. These chips generate as little as 58MB/s of sustainable throughput at an 80MHz bus speed. More-expensive parts offer an octal SPI and operate at up to 200MHz, but even they fall well behind the 133MHz EcoXIP in throughput and latency for XIP applications, as Figure 2 shows.

Adesto plans to increase the EcoXIP bus speed to 200MHz in order to boost this performance further. In addition to modifying the protocol, the EcoXIP interface has an extra data strobe signal, which simplifies the implementation of designs that operate at speeds above 80MHz. Current high-speed designs require a dynamic delay line to synchronize the DDR transfers, but the strobe allows the MCU to capture data using a simple fixed delay.

Two Banks, No Waiting

Flash memory retains data even when a device is powered down, but this feat requires a complex and time-consuming write operation. For NOR flash, this operation involves

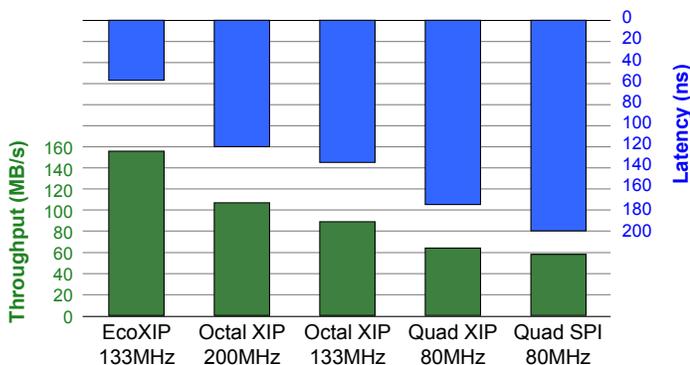


Figure 2. Adesto EcoXIP performance. All numbers are for XIP operation and assume 16-byte instruction-cache lines and an average of 3.84 line fetches per instruction-cache miss. (Source: Adesto)

applying a high voltage (above 5V) to the cell for a period of roughly 1ms. During this period, the MCU cannot fetch instructions from the flash chip. If the flash is solely for code execution (XIP), this situation won't arise, as no writing is necessary. But many systems use flash to store data, such as configuration parameters and event logs. OTA code updates also require writing data to the flash.

Of course, the system can simply stall for 1ms each time it writes to flash, but that delay hampers performance. Another option is to load a small amount of code into the MCU's internal memory before starting an OTA update, but if anything unusual occurs (such as an interrupt), the rest of the code will be unavailable. Thus, many designs include two flash chips, so one can be read while the other is written, but this approach adds cost.

EcoXIP separates its internal flash memory into two banks. Doing so allows the MCU to read from one bank while writing to the other. Designers can adjust the boundary between the banks to split the memory 50/50 or put as little as one-eighth in one bank. The former approach enables OTA updates to store a complete set of code without overwriting the original code; the latter is good for systems that just need a small amount of data memory.

As a further enhancement, EcoXIP implements an automatic power down after a write. Most other flash chips require the MCU to stay awake during the 1ms write so it can power down the flash once the write completes. With EcoXIP, the MCU can "fire and forget," starting the write and immediately going into a low-power mode while the flash chip finishes the write and then puts itself to sleep. The Adesto chip provides a variety of sleep modes that trade off power savings against wake-up time.

A Zippier XIP

Adesto offers two product lines. One is a unique nonvolatile memory, called conductive-bridge RAM, that is CMOS compatible (see [MPR 2/22/16](http://mpr.2/22/16), "Adesto Targets IoT Using CBRAM"). It also acquired a family of standard NOR-flash chips from Atmel in 2012. CBRAM is a lower-power alternative, but NOR flash remains less expensive for storing large amounts of boot code. EcoXIP builds on these standard products, adding a custom interface that improves performance.

Other vendors also offer fast flash memories using custom interfaces. For example, Cypress (formerly Spansion) offers the proprietary HyperBus interface, which can deliver 333MB/s using a 166MHz DDR octal interface that supports arbitrarily long bursts. The Macronix OctaFlash and Micron XTRMFlash have similar capabilities at speeds of up to 200MHz using modified SPI protocols. But all of these parts are designed for fast boot in systems that copy the code into RAM for execution, so they are available only in sizes of 128Mb (16MB) and larger. These systems employ higher-performance processors instead of microcontrollers and often run complex operating systems.

Adesto began with a similar concept but optimized it for XIP applications. Products such as XTRMFlash are designed for long predetermined bursts, whereas EcoXIP allows the CPU to inject a new target address into the burst at any time. Furthermore, Adesto targets applications with more than 1MB of code but less than 16MB, a range that encompasses many MCU-based IoT clients that use a basic real-time OS or no OS at all. The company's dual-bank design is a unique capability that can reduce cost in systems that would otherwise require two separate flash chips.

For MCU-based systems, EcoXIP is less expensive than a large on-die flash memory, since an embedded-flash process adds cost compared with a flash-optimized process. Using XIP eliminates the need for a large and costly on-die SRAM; in fact, EcoXIP can couple with an inexpensive MCU that has minimal on-die memory. The dual-bank chip is also less expensive than two separate flash chips of half the capacity, in part because of package-cost savings. EcoXIP's unique capabilities should help Adesto gain a foothold in the IoT market. ♦

To subscribe to *Microprocessor Report*, access www.linleygroup.com/mpr or phone us at 408-270-3772.