

Direct RDRAM Sustains 1.5 Gbytes/s

Controller Optimally Schedules Future RDRAMs' Every Move



by Peter Song

With unanimous backing from major DRAM chip makers, Direct RDRAM is poised to become the mainstream memory solution for 1999 and beyond. At the recent Microprocessor Forum, Allen Roberts of Rambus disclosed the first technical details of the revolutionary new interface, which narrows the widening gap between the bandwidth required by high-end systems and the bandwidth existing DRAMs can deliver. The interface offers more bandwidth than competing solutions by using a higher data rate and a conflict-free memory pipeline, allowing the controller to schedule every operation in the memory core.

Using an 18-bit data bus and both edges of a 400-MHz clock, a single Direct RDRAM chip can deliver up to 1.6 Gbytes/s, 2.5× more than existing RDRAMs and 8× more than current ×16 SDRAMs. To sustain as much of the peak bandwidth as possible, the latest Rambus standard uses row- and column-command buses separate from the data bus, 16 banks per RDRAM to reduce bank conflicts, and a late-write scheme that integrates reads and writes into a conflict-free memory pipeline. The new interface requires a simpler controller than previous Rambus interfaces, yet it sustains bandwidth closer to its peak data rate.

With Intel's endorsement (see MPR 4/21/97, p. 12), Direct RDRAM has garnered wider industry support than the second-generation Rambus standard, known as Concurrent RDRAM. The top 13 DRAM vendors, which shared more than 90% of DRAM sales in 1996, have all licensed Direct RDRAM, and more are likely to join by the time the first Direct RDRAMs appear in the second half of 1998. While most DRAM vendors plan to introduce the new Rambus interface in their 64- and 72-Mbit DRAMs, some plan to introduce it first in 32-Mbit DRAMs, which are suitable for cost-sensitive graphics applications.

New Interface Uses Wider, Smaller-Swing Buses

With the addition of Direct RDRAM, Rambus now offers three memory interfaces, as Table 1 shows. Beyond the 16-Mbit generation, its original standard known as Base RDRAM is being replaced by Concurrent RDRAM, which keeps the same physical and electrical interfaces but improves the protocol efficiency. Rambus's latest standard uses entirely new physical, electrical, and logical interfaces and will coexist with Concurrent RDRAM in the 64-Mbit generation.

The current RDRAM interfaces use a single high-speed bus to transmit both control and data packets, limiting the effective memory bandwidth. Direct RDRAM doubles the

data-bus width to 18 bits and moves the control packets onto dedicated control buses; as the 64-Mbit generation of standard DRAMs migrates to a wider bus, keeping the RDRAMs to 32 pins becomes less necessary and provides an obstacle to achieving higher bandwidths.

Direct RDRAM's electrical interface better accommodates 0.25- and 0.18-micron processes, which will be common when Direct RDRAM products ship in volume in 1999. The supply and termination voltages for the channel are lowered to 2.5 V and 1.8 V, respectively. Centered on the reference voltage of 1.4 V, signal swing is reduced from 1 V to 0.8 V, reducing power consumption and signal-switching time. (Although 600 mV is specified in the original Rambus signaling logic, actual implementations use a 1-V signal swing). In addition, the clock loop that provides transmit and receive clocks now uses differential signaling, resulting in sharper clock edges and better timing margins.

Direct RDRAM introduces clock domains to manage propagation delays longer than a clock cycle, making it physically scalable. The data and the clock travel in the same direction; this minimizes the skew between data and the associated clock but does not reduce the propagation delays. The propagation delays cause reads from RDRAMs further away to take longer than reads from those closer to the controller. Concurrent RDRAM, therefore, limits the length of the Rambus Channel to about 4 inches at 600 MHz, enough for 32 RDRAMs or 10 RModules of 32 RDRAMs each.

Using up to three clock domains, the Direct Channel can be as long as 16 inches on conventional (FR4) PC boards. Based on their distance from the controller, the slave devices are placed into clock domains, as Figure 1 shows. When RDRAMs in the inner domains are programmed to use additional cycles for read latency, all RDRAMs in the system

	Base	Concurrent	Direct
Memory Density	4M, 16M	16M, 64M	32/64M-1G
Frequency	600 MHz	700 MHz	800 MHz
Peak Bandwidth	600 MB/s	700 MB/s	1.6 GB/s
Data Bus	9 bits	9 bits	18 bits
Dedicated Control	2 bits	2 bits	8 bits
Protocol Efficiency	<60%	80%	95%
Transfer Size	8-256 bytes	8-∞ bytes	16-∞ bytes
RDRAM Pin Count	32 pins	32 pins	58 pins
V _{dd}	5/3.3 V	3.3 V	2.5 V
V _{term}	2.5 V	2.5 V	1.8 V
Signal Swing	1.0 V	1.0 V	0.8 V
Clocks	Single ended	Single ended	Differential

Table 1. The Concurrent RDRAM interface replaces the Base RDRAM interface from the 16-Mbit generation and will coexist with the Direct RDRAM interface in the 64-Mbit generation.

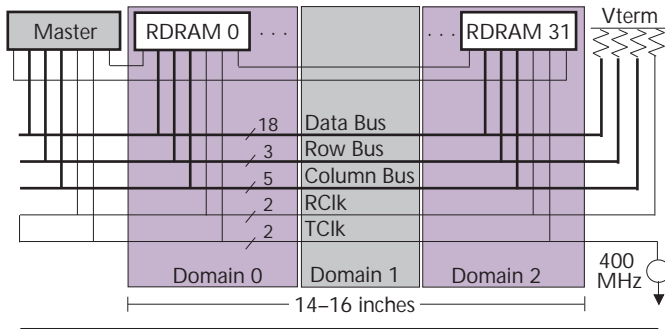


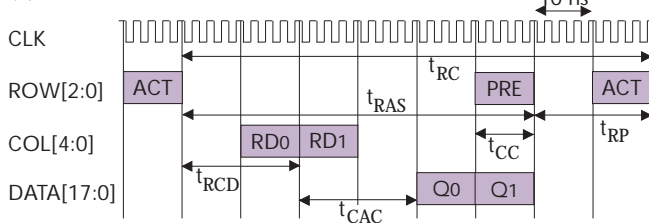
Figure 1. The Direct RDRAM interface uses separate control and data buses, differential transmit and receive clocks, a lower termination voltage, and smaller signal swings than the Concurrent RDRAM interface. The three clock domains support a larger area at higher data rates than previously possible.

respond in the same number of cycles, simplifying controller designs. Whereas Concurrent RDRAM supports only one domain, Direct RDRAM supports one, two, or three domains—requiring zero, one, or two additional cycles of latency, respectively. Because these extra latency cycles do not reduce bandwidth, they have only a minor performance impact. The extra cycles can be omitted in single-domain designs, which can have 10–14 RDRAMs.

Controller Manipulates Memory Core

A key innovation is the Direct RDRAM controller, which remotely directs every aspect of the DRAM. The controller moves bits among the memory core, the sense amps, and the I/O buffers while obeying the DRAM timing requirements. What makes this much control bandwidth economical are the independent row- and column-command buses that operate at eight times the speed of the RDRAM core.

(a) Read Transaction



(b) Write Transaction

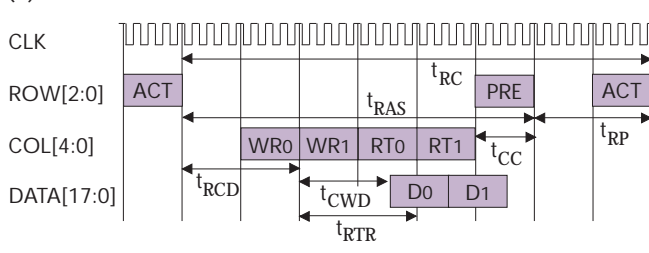


Figure 2. The timing diagram of a read transaction resembles that of page-mode DRAMs. For a write transaction, however, the controller delays sending data packets to avoid a bus conflict with an earlier read.

Since each packet occupies four clocks (eight transfers using both clock edges), the 3-bit row and the 5-bit column packets transfer 24 and 40 bits of control information, respectively, in 10 ns. Since 64-Mbit DRAMs are also expected to have 10 ns of CAS cycle time—the time needed to access the sense amps—these 64 bits of 10-ns control signals are fast and plentiful enough to control every aspect of the DRAM core. In comparison, existing DRAM interfaces, such as fast-page-mode, EDO, SDRAM, and even Concurrent RDRAM, provide fewer than 16 bits of control information in each CAS operation.

The row packets typically specify a device, bank, or row operation, such as entering a low-power mode, precharging a bank, or accessing a row. The column packets typically specify reading or writing a column as well as precharging a bank. Both row- and column-packet formats have many reserved bits to define additional commands and address bits, making Direct RDRAM functionally scalable.

A read transaction takes place in two phases, as Figure 2 shows. The ACT (activate) row command first copies an entire row into a row cache—storage within the sense amps—during the t_{RCD} interval. The RD (read) column command then moves 16 bytes from the row cache to the output buffers during the t_{CAC} interval. A row must remain active for the minimum of t_{RAS} interval, which includes the time to activate the row as well as to restore its contents from the row cache, since accessing a row in any DRAM is destructive. The interval t_{RP} is needed to precharge a bank before another row in the same bank can be accessed.

Unlike Base RDRAM, a single read (or write) transaction no longer transfers variable-length bursts. A read transaction instead transfers 16 bytes, matching the required bandwidth of 1.6 Gbyte/s with the expected CAS cycle time of 10 ns. Similar to page-mode DRAMs, a series of RD commands can be issued to read an arbitrarily long burst of 16-byte blocks. The blocks, however, can be from any active page, not just from one as with page-mode DRAMs.

Late Write Avoids Data-Bus Conflict

Unlike Concurrent RDRAMs and conventional DRAMs, which use different timing for read and write transactions, Direct RDRAMs use nearly identical timing for both reads and writes. The controller sends the ACT and the WR (write) commands in the cycles that correspond to similar commands in read transactions. Instead of sending the data packets as soon as possible, it sends them in the cycles in which data would return from RDRAMs, had this been a read transaction. The controller sends the write data packets after t_{CWD} cycles following the WR command. By doing so, it can interleave read and write transactions without creating bus conflicts.

In most systems that support multiple drivers on a bus, protocols generally require one idle cycle when switching from one driver to another to prevent bus contention. Direct RDRAM controllers must insert an idle cycle before sending

a write transaction when the write follows a read, but not when a read follows a write or when consecutive reads are from different devices. Its current-mode bus and the clock distribution scheme eliminate the need for an idle cycle for the latter cases.

Within Direct RDRAMs, the process of writing 16 bytes of data into the row cache occurs in two steps. The first step is to transfer the write command, along with the column address and the write data, to the write buffer. The second step is to retire the write buffer into the row cache while applying an optional byte mask. In most cases, the write buffer is retired automatically by an RT (retire) or by another column command that is received t_{RTR} cycles after the WR command. The byte mask must be specified in the column command that retires the write buffer.

Pipeline Sustains Peak Bandwidth

Although Rambus touts high data rates, its previous two generations delivered sustainable bandwidths well short of their data rates for moderate-length transfers. Due to its shared bus, the original interface can sustain less than 60% of its peak data rate. By eliminating the ACK/NACK protocols and thereby requiring controllers to keep track of active pages within RDRAMs, Concurrent RDRAM improves the sustainable bandwidth for 32-byte transfers to 80% of its peak rate. Direct RDRAM, in contrast, can sustain 95% of the peak rate of two RDRAMs for random 32-byte reads and writes, as well as for mixtures of both. The efficiency improves further with longer bursts.

Because the controller can precharge and activate arbitrary pages during data-transfer cycles, it can also avoid page misses, transferring blocks from multiple pages anywhere in the system in a single burst. Conventional DRAMs, in contrast, restrict a burst to a single page for two reasons. The primary reason is that their single-bank organization supports only one active row. The other is that their multiplexed address bus cannot receive a new row address during a burst. Direct RDRAMs avoid these problems by having 16 banks in each chip and a separate set of row and column commands. Since each bank shares sense amps with the adjacent bank, no two adjacent banks can be simultaneously active, so the number of independent banks is reduced to eight.

Direct RDRAMs also provide more banks (and active pages) per device than do most DRAMs of the same size. For example, current 64-Mbit SDRAMs have four independent banks, only half as many as in 64-Mbit Direct RDRAMs. One of the few exceptions is Mosys's Multibank DRAMs, which are constructed of 32-Kbyte banks, one-sixteenth the size of

Direct RDRAMs' 512-Kbyte banks (see MPR 12/25/95, p. 17). Although the added number of banks and pages reduces page conflicts and miss rates, it also increases the area overhead and the cost premium over conventional DRAMs.

Due to Rambus's byte-serial memory organization, which assigns each RDRAM to a contiguous address range, the number of independent banks increases proportionally with the number of RDRAMs in a system. In memory systems that arrange conventional DRAMs in parallel, the number of independent banks is reduced by the number of DRAM chips used in parallel. For example, a 64-Mbyte system constructed of $4M \times 16$ SDRAMs—four in parallel—has only 8 independent banks, whereas the same system constructed of 64-Mbit Direct RDRAMs has 64.

Controller Dictates Performance

Requiring the controller to keep track of active pages within RDRAMs and issue row and column commands in a timely manner binds the performance of the memory system to the intelligence of the controller. To deliver the best performance, the controller must monitor the state of 16 banks in each RDRAM, or 128 banks in a 64-Mbyte system. In addition, the controller must be designed for the maximum memory capacity that can be supported in the system.

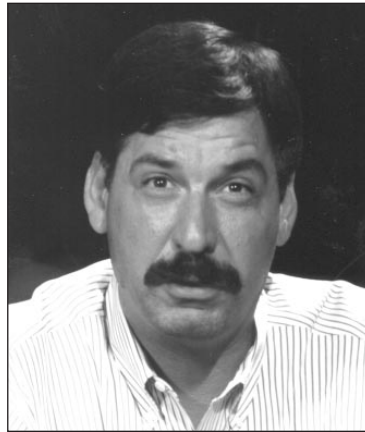
Simple controllers that keep track of a fixed number of active pages, however, can deliver good performance. Because activating a row is faster when the associated bank is already precharged than when another row is active, keeping all banks active may not yield the best performance. For references that lack locality, closing pages—updating the memory array if the row caches contain modified data or discarding the contents otherwise—and keeping the banks precharged can yield

better performance than keeping the banks active during idle cycles. In the banks that are being precharged, there are no active pages to monitor.

According to Rambus, a controller that leaves open only four of the most recently used pages is easy to design and can actually deliver better performance than one that leaves all banks active. Such a controller requires only 10,000 gates and is suitable for the PCs of 1999, which will have only four threads, from PCI, AGP, CPU code, and CPU data. Since the row cycle time—the shortest time to access different pages within a bank—is equivalent to eight packet-transfer cycles (or 32 cycles), a controller that interleaves four requests can deliver the peak bandwidth, provided that each burst is at least 32 bytes long.

Direct RDRAMs to Compete With SDRAMs

Roberts stated that Direct RDRAMs will in 1999 be competitive with 100-MHz SDRAMs in system cost while delivering



MICHAEL MUSTAGHI

Allen Roberts explains how Direct RDRAM sustains 1.5 Gbytes/s of bandwidth.

three times their effective bandwidth. The 16-Mbit RDRAMs use a 10% larger die than EDO DRAMs of the same generation, because they have more banks in the memory core and more logic and faster transistors in the interface. In the 64-Mbit generation, however, Direct RDRAMs are up to 5% larger than $\times 16$ SDRAMs in die size but smaller than $\times 32$ SDRAMs, according to Rambus. Compared with EDO DRAMs, direct RDRAMs have lower area overhead than SDRAMs because SDRAMs also use more banks in the memory core and more transistors in the interface than do EDO DRAMs.

Die size is the most important factor that affects the manufacturing cost of DRAMs, and the RDRAM area overhead adds a modest premium over the cost of conventional DRAMs. The Direct RDRAM's high-speed interface, however, creates additional cost. Testing RDRAMs at 800 MHz requires expensive test equipment—equipment that DRAM vendors are currently not using—that must be depreciated. (HP recently announced a digital IC tester that can test eight RDRAMs in parallel, at a starting price of \$1.2 million.) The high-speed tester doesn't reduce overall test time, because most of the DRAM test time is spent waiting for the DRAM cells to leak most of their charge.

Direct RDRAMs use a CSP (chip-scale package) that offers better electrical and thermal properties than the cheaper TSOP. Although the cost of CSP is rapidly decreasing, it is likely to remain more expensive than the more common TSOP. Due to their faster interface and byte-serial memory organization, RDRAMs consume more active power and generate greater thermal fluctuations than do SDRAMs. Memory systems that use RDRAMs, therefore, will need to be designed with better thermal management than those that use SDRAMs.

At the system level, however, RDRAMs consume less power while delivering better performance than SDRAMs, making them a good candidate for SDRAM replacement. Although RDRAMs use faster signals, they also use smaller signal swings and for a shorter period of time. Their larger and better-organized row cache reduces row-cache misses and therefore memory-array accesses. In addition, a row-cache miss causes only one RDRAM chip to access its memory array, but it causes multiple SDRAMs to access their memory arrays. By using nap mode, which takes only 100 ns to exit, power dissipation can be reduced to tens of milliwatts, making RDRAMs well suited for low-power applications.

Competitors May Find Niches

DDR (double data rate) SDRAMs are likely to offer strong competition to Direct RDRAMs in server markets. Because DDR SDRAMs, also known as SDRAM II (see MPR 2/17/97, p. 4), require smaller die, cheaper packages, and lower testing costs, they are expected to have a lower cost than RDRAMs, making them better suited for applications that require large amounts of memory. Because the Direct Channel can support at most 0.25 Gbytes of memory (32 RDRAMs of 64 Mbits each) without

using a repeater chip, it is easier to add more memory when using DDR SDRAMs than RDRAMs. It is also more reliable to store contiguous memory locations in multiple SDRAMs than in a single RDRAM device. Servers that need more bandwidth than DDR SDRAMs can provide, however, are likely to use multiple Direct Channels.

In the 3D graphics market, where small amounts of memory must deliver high bandwidth, Direct RDRAM is likely to face steep competition from embedded DRAM (see MPR 8/4/97, p. 13). By 1999, embedded-DRAM chips are likely to pack 4–8 Mbytes of DRAM and enough transistors to include a powerful graphics controller. Although embedded DRAMs are likely to be more expensive than RDRAMs, they offer more bandwidth and shorter latency. They also offer the added benefits of reducing power consumption, EMI radiation, and board space, making them more attractive than RDRAMs for portable applications.

SLDRAM (see MPR 5/12/97, p. 9) sports several bandwidth-enhancing features that are conceptually similar to those of Direct RDRAM. The major difference is that the 64-Mbit SLDRAMs will support a peak data rate of only 400 MHz, or 800 Mbytes/s of peak bandwidth, because the SLDRAM consortium believes that yields will be high enough at 400 MHz to minimize the cost premium. Without a stronger commitment and better cooperation among the DRAM vendors, however, SLDRAM is unlikely to be a real threat to Direct RDRAM. In spite of the consortium members' devotion, a committee of part-time volunteers offers limited competition to Rambus—a company whose very survival depends on its memory interface.

Direct RDRAM to Dominate PC Main Memory

The cost premium of Direct RDRAMs over SDRAMs is largely unknown—Rambus projects 5%, while some DRAM vendors quote 10–15% to their OEM customers. The agreement between Rambus and Intel caps Rambus's royalty at 2% on RDRAMs, as long as in each quarter Intel ships 20% of its PC controllers using Direct RDRAM and the DRAM vendors together ship more than 25% of their DRAMs with the Direct RDRAM interface. The agreement implies that some, if not many, DRAM vendors will pay more than 2% in the early going, when the volume status is not reached.

As long as the price premium doesn't get out of hand, the exact level is probably irrelevant, at least in the PC market. Once Intel begins deploying chip sets that support only Direct RDRAM (and not SDRAM), the company's dominant position in processors and chip sets should allow it to establish Direct RDRAM as the de facto PC memory standard.

By 1999, PCs will require more memory bandwidth than 100-MHz SDRAMs can provide. DDR SDRAM could be an interim solution, but Direct RDRAM delivers more sustainable bandwidth and better scalability for the future. The transition will take some time, but by the end of 2000, Direct RDRAM could easily become the best-selling memory device for computer systems. 