WHITE PAPER

Dinesh Maheshwari,
CTO, Memory Products Division
Cypress Semiconductor Corp.

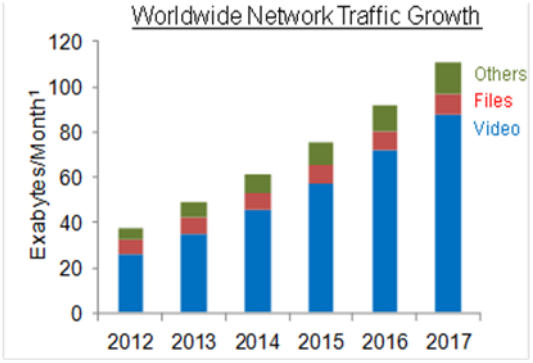
QDR®-IV: The Next Generation of the Highest-Performance Memory Standard for Leading-Edge Networking Systems

Abstract

As networking and computing equipment provider's scale their solutions to meet the demands of the Internet, memory capabilities increasingly govern system performance gains. Routers, switches, and high-performance computing systems all require memories that can transact with processing elements at ever-increasing rates. Random transaction rate (RTR), the number of random read or write accesses that a memory can support, therefore becomes the key memory metric. This white paper discusses how Cypress's QDR®-IV provides the highest RTR and standards-based solution to meet these system challenges.

Demands on Current Networking Systems

It is projected that global networking traffic will triple over the next four years with the proliferation of Internet-connected devices and bandwidth-hungry video services. By 2017, 3.6 billion Internet users will drive 19 billion network connections. In response, network equipment providers are deploying switch and router line cards, which double in performance, measured in line card rate (Gbit/s) or packet rate (MP/s), every generation. Figure 1 and Figure 2 depict these trends. Note that packet rates are directly proportional to line card rates: packet rate = line card rate / (minimum packet size + inter-packet gap). The minimum packet size is 64 bytes, with an interframe gap of 20 bytes. Therefore, the packet rate for 100-Gbit/s line cards is approximately 150 MP/s.



[1] Network traffic is measured in exabytes (10^{18} bytes) per month

Figure 1. Worldwide Network Traffic Growth

Source: Cisco Visual Networking Index (VNI) Forecast (2012–2017)

.

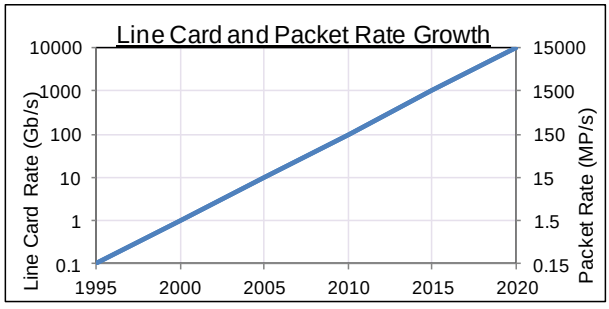


Figure 2. Line Card and Packet Rate Growth

Source: Ethernet Technology Summit

Random Transaction Rate

Random transaction rate (RTR) represents the number of random read or write accesses that a memory can support in a given time. It is measured in multiples of transactions per second (for example, MT/s or GT/s). The RTR required in networking systems is determined by the packet rate and the number of random memory accesses required for each packet, which varies with the line card function that the memory supports.

Line Card Memory Functions

The criticality of RTR in networking systems is best viewed in the context of the six key functions of memory on a line card. As detailed in the following sections, these functions determine the performance of the entire system. The major performance bottleneck is not the switching speed, but the time taken to examine the packet header to determine its type and destination, look up the route in the forwarding table, classify the packet for scheduling, and update various statistics and state.

Packet Buffering

Each line card on the router has an ingress buffer and egress buffer. The ingress buffer holds the packets from incoming physical links (after classification and forwarding decisions are made) as they wait their turn to be scheduled and dispatched to the switch fabric. The egress buffer stores packets coming from the switch fabric to be scheduled and dispatched to outgoing physical links. The higher the packet rate, the greater the need for packet buffer size, so the choice of the packet buffer memory is dictated primarily by density. If a high-density memory does not have the requisite RTR to keep pace with packet rates, system manufacturers often architect a hierarchical structure in which they use a “head-tail cache” to store the heads and tails of the packet buffer. Head-tail caches can require up to four to six transactions per packet.

Forwarding Information Base Lookup Table

The forwarding information base (FIB) lookup table stores the exit port address on the switch/router that corresponds to a destination IP address. The lookup is an iterative process, involving multiple accesses to the memory. Lookups are performed at both Layer 2 and Layer 3 in the network stack. Layer 2 lookups are typically implemented via hashing algorithms that look for exact matches to the search index and usually require four random accesses per packet. Algorithmic Layer 3 lookups using the longest prefix matches can require up to eight random accesses per packet.

Packet Classification or Access Control List

In this processing step, the processor examines the characteristics of incoming packets (for example, a 5-tuple composed of source address, destination address, source port, destination port, and protocol) and makes decisions on whether or not to allow the packet through. Once the packet has been classified or policed, it is temporarily stored in the buffer subsystem for scheduling. Algorithmic approaches to access control list (ACL) lookup also require more than 10 random accesses per packet.

Scheduling

Scheduling is the process of deciding when to send a packet onto the switch fabric and is determined based on the destination of the packet and quality of service (QoS) or class of service (CoS) required. Packets are grouped into several classes, each of which relates to a tiered service offering (revenue segments for service providers). The scheduling application requires at least one read plus one write transaction per packet, making RTR critical.

Statistics and States

Routers maintain statistics on a per-packet and per-flow (stream of related packets) basis. This is accomplished by counters, which are used to store information on prefix, flows, and packet classification. Updating a counter requires a read-modify-write operation (two random memory transactions). The RTR requirement scales with both packet rates and the number of counter updates required. The same memory in a line card can be shared for statistics and states.

RTR, the Key Metric

Each of the functions demands a high RTR. Other memory metrics are less important. Bandwidth increases proportionally with RTR for any given data width. Note that bandwidth can also increase due to an increase in data width. Density is a critical attribute for packet buffering, but head-tail caches illustrate the need for high RTR in this function as well.

RTR Requirements by Packet Rate and Memory Function

Table 1 shows the RTR requirements for 100-Gbit/s (150-MP/s) and 200-Gbit/s (300-MP/s) line cards.

Memory Functions	Transactions Per Packet	100 Gbit/s 150 MP/s	200 Gbit/s 300 MP/s
Head-tail cache/scheduler	4	600 MT/s	1200 MT/s
Lookup	4–8	600–1200 MT/s	1200–2400 MT/s
Statistics/states	8–6 (4–8 counters)	1200–2400 MT/s	2400–4800 MT/s
Packet buffering	4–6	600–900 MT/s	1200–1800 MT/s

Table 1. RTR Requirements for Line Cards

RTR: Beyond Networking

The value of RTR performance is not limited to networking systems. RTR is also a key requirement in high-performance computing, general-purpose servers, and image-processing applications. High-performance computing systems—clusters of compute and storage resources aggregated to solve complex problems—require memories with high RTR to process data from several streams. General-purpose servers, which are increasingly built with “multicore” and “many-core” processors with shared last-level cache, require high-RTR memories to support the demands of the CPUs. Additionally, high-performance image-processing systems implement digital signal processing algorithms that demand fast random access to image data, so they also benefit from high-RTR memories.

Selecting Memory Devices With High RTR

Designers must understand the differences in RTR provided by various memory types to optimize system performance. The core memory technology, memory device structure (banks), and memory interface determine the RTR of any memory device.

Memory Technology

The most critical factor in determining the RTR of a memory device is the core memory technology; for example, synchronous DRAM (SDRAM), reduced latency DRAM (RLDRAM), or QDR® Synchronous SRAM (QDR SRAM). In SDRAM and RLDRAM, true random access is largely limited by random cycle time latency (tRC). The maximum RTR is approximately the inverse of tRC (1/tRC). SDRAM tRC has not evolved substantially over the past 10 years (nor is it expected to evolve going forward) and stands at ~48 ns, which correlates to a 21 MT/s RTR. Other DRAM-based memory devices have been designed to improve tRC at the expense of density. For example, RLDRAM 3 has a tRC of 8 ns, which correlates to a 125-MT/s RTR. QDR SRAM is specifically optimized for random access. The QDR-II+ Xtreme family (the predecessor of QDR-IV) delivers a 900-MT/s RTR.

Memory Device Structure (Banks)

To overcome DRAM’s fundamental tRC limitations on bandwidth, DRAM device providers introduced the notion of “banking,” in which the device is segmented into banks, each of which can be addressed independently. The RTR within any given bank is still limited by memory technology (as explained in the previous section); however, since each bank can be accessed independently, the

total transaction rate of the device (excluding the effect of memory interface limitations) is the number of banks that can be active at a given time multiplied by the RTR within a bank.

DRAM-based devices impose large restrictions on bank access, and the segmentation of the memory into banks also means that the total transaction rate across the device is not purely random (access is random only within each bank). Enabling purely random read access requires replicating the same data in multiple banks, thereby increasing the read RTR by the replication factor (the number of times the data has been replicated).

However, replication also results in a significant trade-off. It reduces the effective density of the device by the replication factor. Also, replication does not work for applications that require random read and write accesses. In these cases, any given write access has to be repeated for each of the data copies, which requires multiple transactions and thereby defeats the purpose of replication. A "ping-pong" approach can be implemented to enable 2x RTR for read and write, in which write accesses are directed alternatively to each of the two banks, and a map is maintained to designate which of the two banks contains the latest data. Note that this method doesn't result in true replication, since the data sets in each bank aren't exact copies. Also, this approach does not scale to 4x RTR (using four banks), which requires updating at least two banks after each write (again implying multiple transactions and defeating the purpose of replication).

Memory Interface

The final determinant of a memory device's RTR is the capability of the interface, which may not necessarily expose the full capabilities of the memory. For example, consider the RLDRAM 3. The device is configured in 16 banks, enabling a theoretical total transaction rate of 2000 MT/s (16 banks * 125 MT/s). However, the interface uses a single port, which supports a maximum 1066 million random read or write commands (1066 MT/s). So, the RTR for the device is limited by the interface.

This paper uses the following terms to describe RTR: "RTR per bank" for RTR of the memory technology, "interface RTR" for the transaction rate allowed by the interface, and "memory device RTR" for the total random transaction rate. The memory device RTR is the lesser of the interface RTR or the RTR per bank multiplied by the number of active banks.

Introducing QDR-IV: The Highest Performing Networking Memory

QDR-IV, defined in the QDR consortium, is the first memory to be explicitly optimized for RTR, targeted at the most demanding line card functions: FIB lookup table, statistics, state, scheduling, classification, and head-tail cache.

QDR-IV RTR

The QDR-IV family offers two primary device options: QDR-IV HP (high performance), and QDR-IV XP ("Xtreme" performance). QDR-IV HP operates at a maximum frequency of 667 MHz, enabling a maximum RTR of 1334 MT/s (each of the two ports supports an address rate of 667 MT/s). QDR-IV XP, which incorporates banking, operates at a maximum frequency of 1066 MHz, enabling a maximum RTR of 2132 MT/s. This represents a 2.4x improvement over the QDR-II+ Xtreme, which is currently the highest performance, standards-based memory solution.

QDR-IV also contains other features that optimize and simplify system design, as described in the following sections.

Bidirectional Ports

Traditionally, QDR memories have sported two unidirectional ports: one read port and one write port. This served the networking functions that had balanced read-write such as statistics/state, scheduling, and packet buffering. QDR-IV incorporates two bidirectional ports such that both ports can be simultaneously used for read operations. This doubles the RTR available for read-dominated functions like lookup.

ECC
QDR-IV is the first QDR SRAM to incorporate ECC, which enables a reduction in soft-error rate to less than 0.01 failures in time (FIT) per Mbit.

Signal Integrity Features
QDR-IV supports address and data bus inversion, which reduces power and improves signal integrity by reducing simultaneous switching output noise. QDR-IV also incorporates on-die termination and address parity detection, both of which enhance signal integrity.

Clocking
QDR-IV contains one pair of differential command and address clocks (CK and CK#), along with free-running differential output clocks (QK and QK#) for read data (one pair each for every 9 or 18 I/Os, varying by device). QDR-IV adds differential input clocks for write data (DK and DK#, one pair each for every 9 or 18 I/Os). The addition of the input clocks widens the data valid window for read operations.

Banking (QDR-IV XP only)
Traditionally, QDR memories have not supported any banking. QDR-IV breaks away from this trend by supporting eight banks. QDR-IV banking is far less restrictive than banking in DRAM-based solutions (DDR3/4 SDRAM or RLDRAM 3). Each of the eight banks of QDR-IV may be accessed once per clock cycle. Port A may access any bank each clock cycle, while port B may access any bank *other* than the one port A accessed during the same cycle. So, QDR-IV XP lends itself well to algorithms that can split the search or counter database into two distinct groups and provides up to 2132 MT/s for such algorithms without missing an access.

Flexible I/O Signaling
QDR-IV supports HSTL/SSTL (1.2 V/1.25 V) and Pseudo Open Drain (POD) (1.1 V/1.2 V) signaling. Other memory devices also support these signaling standards (DDR4 SDRAM uses POD, DDR3 SDRAM uses SSTL, and RLDRAM 3 uses HSTL).

QDR-IV Performance Comparison with Other Memories
Figure 3 quantifies the QDR-IV advantage in RTR performance versus other memories. QDR-IV delivers a minimum of 2x the RTR performance of any other memory. The chart also illustrates that QDR-IV uniquely meets the minimum RTR required at both the 150-MP/s (100-Gbit/s) and 300-MP/s (200-Gbit/s) packet rates.

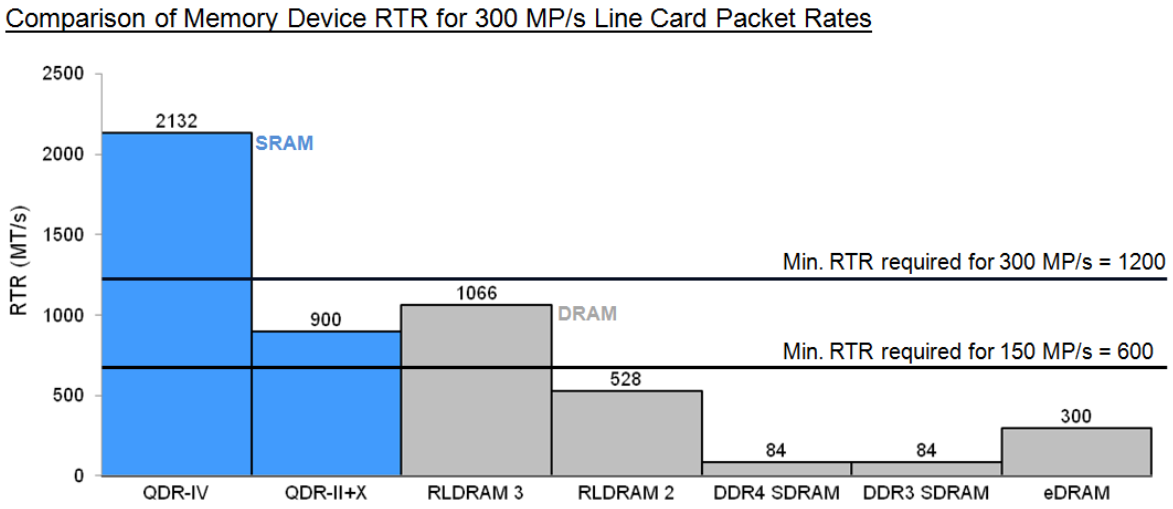


Figure 3. QDR-IV Performance Comparison

Conclusion

QDR-IV, with a maximum RTR of 2132 MT/s, is the highest performance, standards-based memory solution available on the market. Its high RTR, coupled with differentiated features such as dual bidirectional ports, ECC, bus inversion, ODT, and address parity, make it the optimal solution for networking systems. The advantages of QDR-IV also apply to other systems that require high-RTR performance and signal integrity, such as high-performance computing and image processing.

Cypress Semiconductor
198 Champion Court
San Jose, CA 95134-1709
Phone: 408-943-2600
Fax: 408-943-4730
http://www.cypress.com