RapidIO delivers low latency and high bandwidth for multi-computers

By Thomas Roberts

The RapidIO interconnect design accommodates the next-generation performance and transport requirements for chip-to-chip and board-to-board communications within embedded computer systems. Targeted for an initial bandwidth of between 1 and 12 Gbytes/sec per device pair, RapidIO technology is particularly well suited to applications that must meet stringent latency constraints and exceed the performance capabilities of a single processor.

RapidIO offers tremendous potential for increasing the bandwidth and, hence, overall performance of computing systems using multiple processors. This article will examine tests on a multi-computer system to exploit the potential of RapidIO. The test data validates the performance claims made by the RapidIO community and demonstrates the architecture’s maturity for use in real-world applications.

Unlike other next-generation I/O technologies that augment or displace PCI or Ethernet, RapidIO technology fills the full spectrum of interconnect needs within an embedded system, starting with the microprocessor bus. Designers can integrate RapidIO interfaces directly into communications processors, digital signal processors (DSPs), Field Programmable Gate Arrays (FPGAs), and host processors enabling low-latency, tightly coupled multi-computing. The steady and predictable increase in microprocessor clock speeds has driven system bandwidth requirements beyond the capabilities of any bus architecture. RapidIO provides a direct memory-mapped interface, evolved from microprocessor buses, which designers can implement into part of an FPGA or into a small corner of a microprocessor.

The RapidIO Trade Association, a non-profit corporation governed by member companies, administers RapidIO as an open standard and directs its future development. RapidIO provides high data bandwidth as well as message passing services in a shared memory model. It is transparent to application and operating system software, does not rely on device drivers, and is amenable to software-managed programming techniques.

RapidIO specification layers
The RapidIO specification comprises three partition layers: logical, transport, and physical. This multilevel partitioning scales and accommodates future enhancements while maintaining compatibility with legacy RapidIO technology.

Logical layer
The logical layer specification defines the overall protocol and packet formats. This layer is physically independent and can transmit data over anything from serial to parallel interfaces, from copper to fiber media. The logical layer contains the information necessary for end points to initiate and complete a transaction, such as the destination address and the size of the transaction.

Transport layer
The transport layer specification provides the necessary route information for a packet to move from end point to end point.

Physical layer
The physical layer specification describes the specifics of the device level interface, such as packet transport mechanisms, flow control, electrical characteristics, and low-level error management.
Physical interfaces
The RapidIO specification defines two physical interfaces: parallel and serial.

Parallel interface
The specification refers to the parallel physical interface as the 8 or 16-bit, link-protocol end-point specification. The 8/16 LP-LVDS, Low Voltage Differential Signaling (LVDS), typically has 8 or 16 data bits in each direction along with clock and frame signals in each direction. The parallel interface is suitable for chip-to-chip and some board-to-board communication across standard printed circuit board technology at throughputs around 10 Gbits/sec.

Serial interface
For pin-sensitive implementations, RapidIO defines a serial interface that uses differential current steering drivers based on those defined in the 802.3 XAUI specifications. Engineers developed this signaling technology to drive long distances over backplanes.

RapidIO performance in a fully functional system
Engineers can now make true system performance tests of RapidIO. For example, Mercury Computer Systems’ ImpactRT 3100 (Figure 1), a CompactPCI-based multi-computer, implements the RapidIO parallel physical interface over an active backplane. The peak data transfer performance of a parallel RapidIO link as implemented in that system is 622 Mbytes/sec in one direction, or 1.244 Gbytes/sec in two directions. Designers ran a series of tests to determine how actual measured performance compared to this theoretical value.

Designed to address high-end signal and image processing applications, the ImpactRT 3100 system served as a system-level RapidIO test harness. The system consists of 19, 6U RapidIO-enabled slots for quad-processor boards, one RapidIO-enabled slot to serve a RapidIO-PCI bridge board, and one standard CompactPCI slot for the system host board running Windows 2000. Each PowerPC processor in the system has a direct connection to the RapidIO switch fabric, resulting in a theoretical bisection bandwidth of roughly 10 Gbytes/sec. Up to 40 fiber-optic interfaces in the rear of the system support the input data stream for an aggregate peak of up to 10 Gbytes/sec. However, designers did not exercise these fiber-optic I/O modules during the RapidIO performance testing.

“It is transparent to application and operating system software, does not rely on device drivers, and is amenable to software-managed programming ...”
Processor board
The quad-processor board itself represents a complex packaging challenge. Each 6U processing board contains four PowerPC G4 Compute Nodes (CNs) and a RapidIO switch. Each G4 CN consists of a 1-GHz PowerPC 7445 with 256 or 512 Mbytes of SDRAM and a CN ASIC, which provides the interface to the RapidIO fabric for each CN. The board also includes three, off-board RapidIO links from the RapidIO switch. Two of these links use the J4 pins of the board’s CompactPCI connector, where they can join with a RapidIO interlink module. The third RapidIO link uses the J5 pins, where it can connect with a rear transition module for very high-bandwidth fiber-optic I/O. (See Figures 2 and 3.)
Designers built the active backplane around an 8-port RapidIO switch. These switches reside both on the physical backplane and on each board in the system. Figure 4 illustrates the RapidIO switch fabric topology on the backplane in the test harness system. Passing data between processors on different boards involves routing it through the switch on the board to a switch onto the active backplane, and then either to a switch on another board or to another backplane switch, which in turn routes the data to a board. Thus, the data packet can parse through up to four switches.

![Figure 4](image-url)

**Measured RapidIO system performance**

Mercury conducted unidirectional and bidirectional tests exchanging 4000 chained 1-Mbyte data transfers between PowerPC end points. Two conditions were tested: the first was a packetized memory write that did not require a response from the receiving node (NWRITE), while the second condition did require a response (NWRITE_R). Engineers conducted these tests in such a way as to minimize the impact of other activities that can take place within fabric controllers. Thus, the experiment assumed that each command packet fetch consumed approximately 250 nsecs, no snoop activity affected the DCACHE, and the impact of DDR refresh was minimal.

The unidirectional NWRITE test achieved a measured performance of 585.4 Mbytes/sec, or 94 percent of the theoretical raw-data capacity of the RapidIO link. Bearing in mind that the potential overhead consumes 3 percent of the link’s bandwidth, this measured performance of 94 percent raw capacity is even more impressive. The bidirectional test achieved 73 percent of theoretical performance. The decreased data rate is due largely to the additional overhead of bidirectional traffic and falls well within the expected range of performance.

The second test employed a memory-write function that waits for a response from the receiving node. This state represents real-world conditions most closely, since applications typically must verify the receipt of data transferred between end points. The RapidIO test harness provided the option of setting the number of outstanding responses allowed before shutting down the memory-writes. This is a coarse method of flow control that permits up to 16 writes to go unacknowledged.
Table 1 lists the performance obtained under each of four outstanding-response settings. Predictably, the one-outstanding-response case provides the lowest performance, since the sending node must wait for confirmation following each transfer that it sends. Conversely, performance varies little between the 4, 8, and 16-response case, in both the unidirectional and bidirectional tests.

<table>
<thead>
<tr>
<th>Outstanding Responses</th>
<th>Unidirectional Data Rate</th>
<th>Bidirectional Data Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>146.39 Mbytes/sec (23%)</td>
<td>283.32 Mbytes/sec (23%)</td>
</tr>
<tr>
<td>4</td>
<td>558.63 Mbytes/sec (90%)</td>
<td>732.10 Mbytes/sec (59%)</td>
</tr>
<tr>
<td>8</td>
<td>576.41 Mbytes/sec (93%)</td>
<td>744.70 Mbytes/sec (60%)</td>
</tr>
<tr>
<td>16</td>
<td>576.66 Mbytes/sec (93%)</td>
<td>745.90 Mbytes/sec (60%)</td>
</tr>
</tbody>
</table>

The most notable finding from these tests is that unidirectional performance in the 8 and 16-response conditions comes very close to the theoretical maximum performance of the RapidIO link, especially when overhead is considered. Thus, tests have proven that the RapidIO link can deliver the capacity anticipated by the specification.

**Unique capabilities with RapidIO**

For designers with high-end bandwidth and processing challenges, a RapidIO-based system offers unique capability solutions. The RapidIO interconnect was introduced with the promise of ample data transfer capacity for the most demanding embedded-real-time multiprocessing applications. RapidIO systems can deliver data transfer rates very close to the architecture’s theoretical maximum performance, in real-world conditions.

**Thomas Roberts** joined Mercury Computer Systems in 1999 and has more than 20 years of experience in systems engineering and technical marketing with IBM, Nixdorf, Data General, Digital Equipment, and Compaq. Tom holds an MBA from the University of Kansas and a B.S. in engineering from Cornell University.

**Mercury Computer Systems** provides high-performance embedded, real-time digital signal and image processing computer systems. Its products transform sensor data to information for analysis and interpretation for defense electronics, medical diagnostic imaging, and semiconductor imaging applications. For more information, contact Mercury Computer Systems directly.

---

Reprinted from *Embedded Computing Design* / Winter 2003