DEFENSE SOLUTIONS

A Comparison of OpenVPX™ System Bandwidth Between Serial RapidIO® and 10 Gigabit Ethernet

cwcdefense.com

CURTISS WRIGHT Controls
Defense Solutions

CURTISS WRIGHT Controls Defense Solutions

A Comparison of OpenVPX™ System Bandwidth Between Serial RapidIO® and 10 Gigabit Ethernet

Introduction

The VITA 65 (aka OpenVPX) specification has introduced a number of standardized backplane topologies. The major addition brought by VITA 65 is the standardization of systems that employ one or more centralized switch cards. For systems based on Ethernet, a central switch is the norm. For Serial RapidIO (SRIO), VITA 65 adds central switched architectures to the list of possible topologies that includes switch-less mesh implementations.

The cornerstone of the VITA 65 central switch topology is the 16-slot backplane with 14 payload cards and two switch cards. Known as BKP6-CEN16-11.2.2-n, this is the maximum size backplane that accommodates 1" pitch cards (the de facto standard) within a 19" rack enclosure.

This paper compares the bandwidth available to processors resident on the payload boards using the CEN16 switched architecture for Serial RapidIO and 10 Gigabit Ethernet (10GbE). The analysis compares two different "benchmark" data flow models, the classic all-to-all case typical of a corner turn operation, and a pipeline case.

The analysis will show that the SRIO systems have more than three times greater bandwidth than 10GbE, and that half of that improvement is due to the more flexible routing afforded by actual hardware implementations that can be achieved with SRIO silicon.

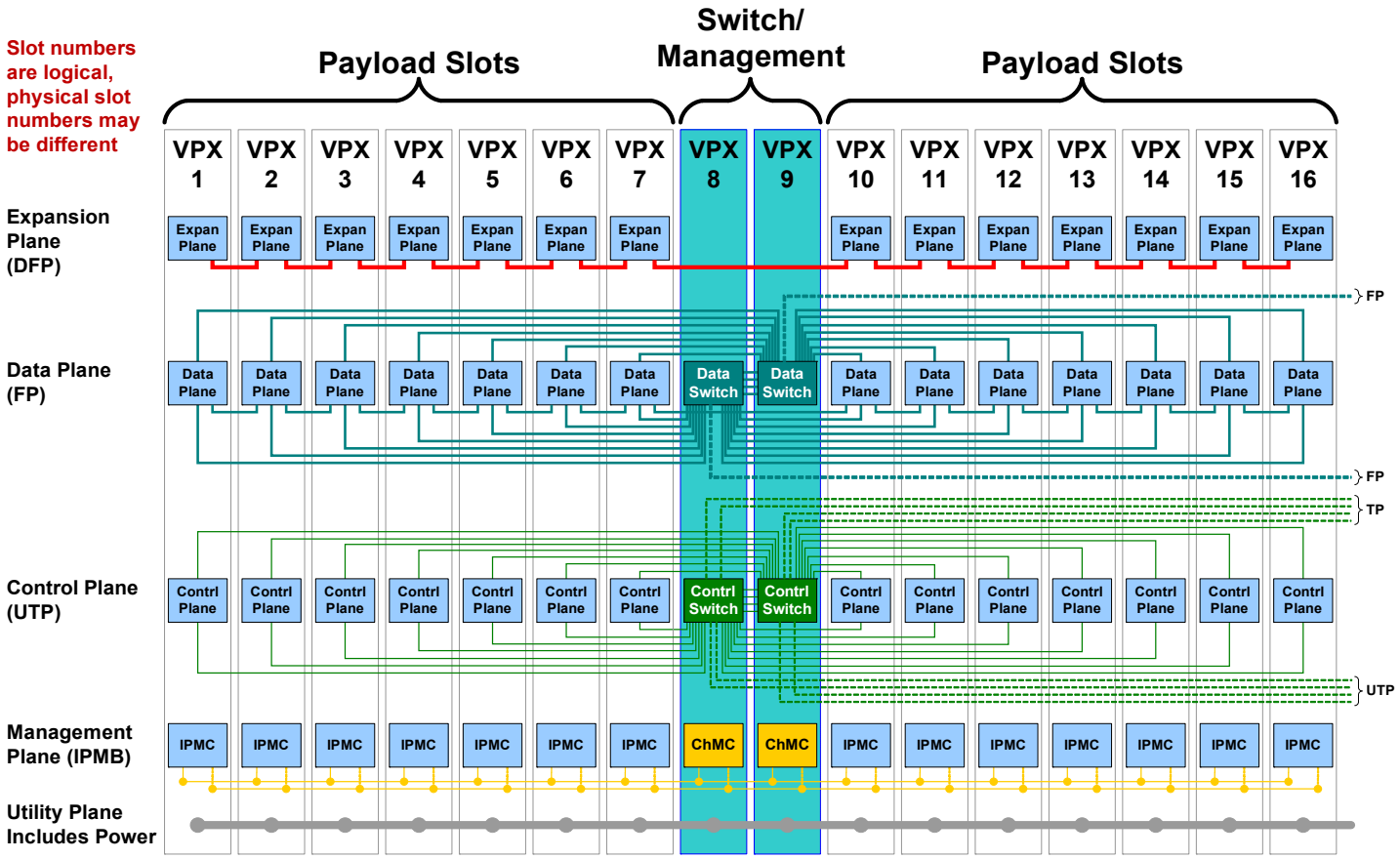
Figure 1: Curtiss-Wright HPEC system based on OpenVPX CEN16 backplane with SRIO-based CPU, FPGA and fabric switch modules.

The System Architecture

Figure 2 shows the topology of the OpenVPX 16-slot centralized switch backplane (VITA 65 BPK6-CEN16-11.2.2) There are 14 payload slots and two switch slots. It is the Data Plane connections that are relevant to this analysis. Each card has two connections to the switch cards and a connection to each adjacent slot.

OpenVPX™

Figure 2: OpenVPX Central Switch Backplane Architecture



Board Architectures

The focus of this paper is high-performance, multi-processor systems where bandwidth available to the CPUs can have a large impact on the system computing throughput.

The analysis compares fabric bandwidth available to the CPUs in such a system. To that end, the paper is based on board architectures that represent current (Sept 2011) state of the art 6U OpenVPX multi-processor cards.

Figure 3 illustrates a board architecture based on Intel® processors and SRIO. This is the architecture of the Curtiss-Wright Controls Defense Solutions CHAMP-AV8 multi-processing card.

There are two CPUs, each having a pair of PCIe to SRIO bridges. These operate with 5Gbps signaling. The resulting four SRIO interfaces connect to a local SRIO switch, with four ports connecting to the OpenVPX Data Plane.

Figure 3: Serial RapidIO Board Architecture

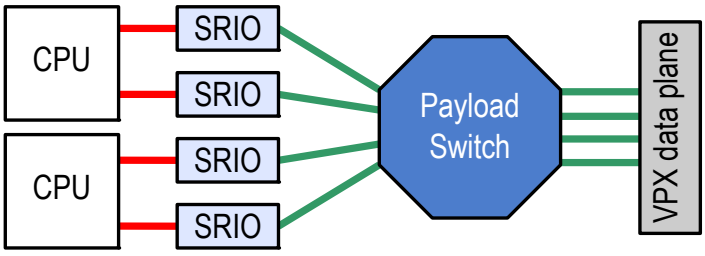
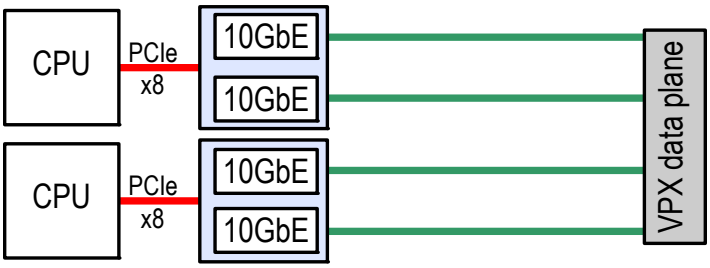


Figure 4 illustrates the architecture of a board based on identical Intel processors and 10GbE. Each CPU has a dual 10GbE NIC. The resulting four 10GbE interfaces connect to the OpenVPX Data Plane. There is no local switch.

Figure 4: 10 Gigabit Ethernet Board Architecture



Assumptions and Data Rate Arithmetic

For the analysis we make the following assumptions and simplifications.

The switches in both cases are assumed to be non-blocking, meaning that they can handle full bandwidth on all switch ports simultaneously. This is known to be true for Curtiss-Wright Controls Defense Solutions SRIO-based switches and believed to be the case for modern 10GbE silicon.

The analysis is based on raw packet throughput. It ignores packet overheads as the differences between the two types are relatively minor.

Multi-processor cards often have an additional communication path between CPUs. For instance, it is possible to transfer data over PCIe locally on the CHAMP-AV8 card. This analysis excludes consideration of this additional path, based on the assumption that communication middleware is unlikely to support simultaneous use of different hardware transport mechanisms.

The bandwidth properties of the two fabrics are as follows.

Gen2 SRIO operates at 5.0 and 6.25Gbps. The IDT TSI721 PCIe/SRIO bridge device supports 5.0Gbps and so the analysis is based on this signaling rate. Accounting for the 8B/10B encoding, a four-lane interface provides 4lanes x 5Gbps x 0.8encoding **= 16Gbps**.

10 Gigabit Ethernet operates at 3.125Gbps. Accounting for 8B/10B encoding, a four-lane interface provides 4lanes x 3.125Gbps x 0.8encoding **= 10Gbps** (as per the name).

Expanded System Architecture

The following two figures show the system architecture that is achieved with OpenVPX multi-processor cards and fabric switch cards when installed in the CEN16 OpenVPX backplane.

Figure 5 depicts the resulting OpenVPX system with SRIO. There are 28 processors in the system. There are some notable aspects to this architecture with respect to bandwidth. Communications between CPUs on a card will flow through the local switch only. We also make the assumption that communications between adjacent cards takes advantage of the card-to-card connection. This is achievable in an SRIO system as the routing between end-points can be set arbitrarily.

In Figure 6, the equivalent system using 10GbE is depicted. Each processor has a single 10GbE connection to one of the switch cards.

Figure 5: System Architecture with OpenVPX and SRIO

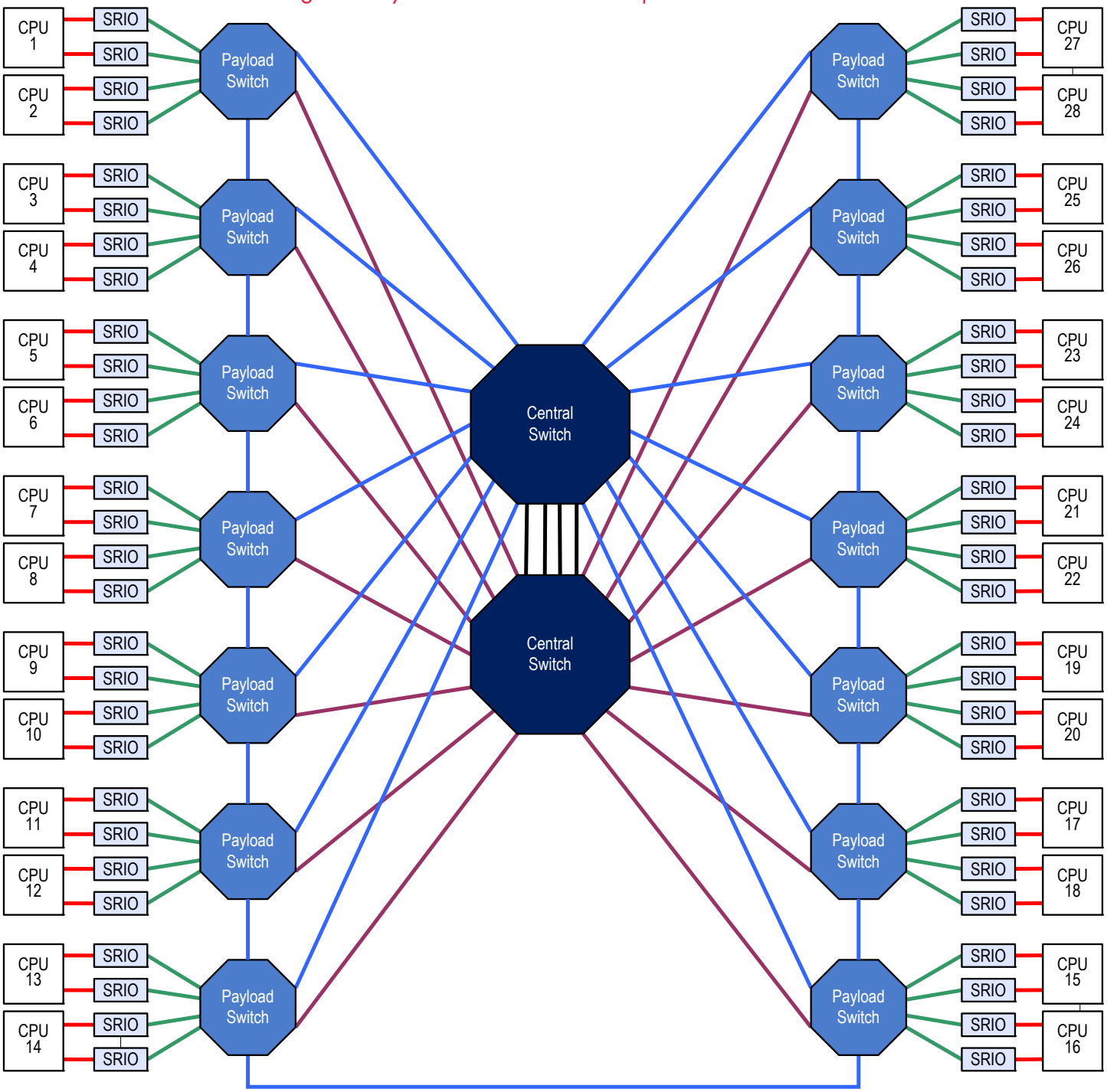
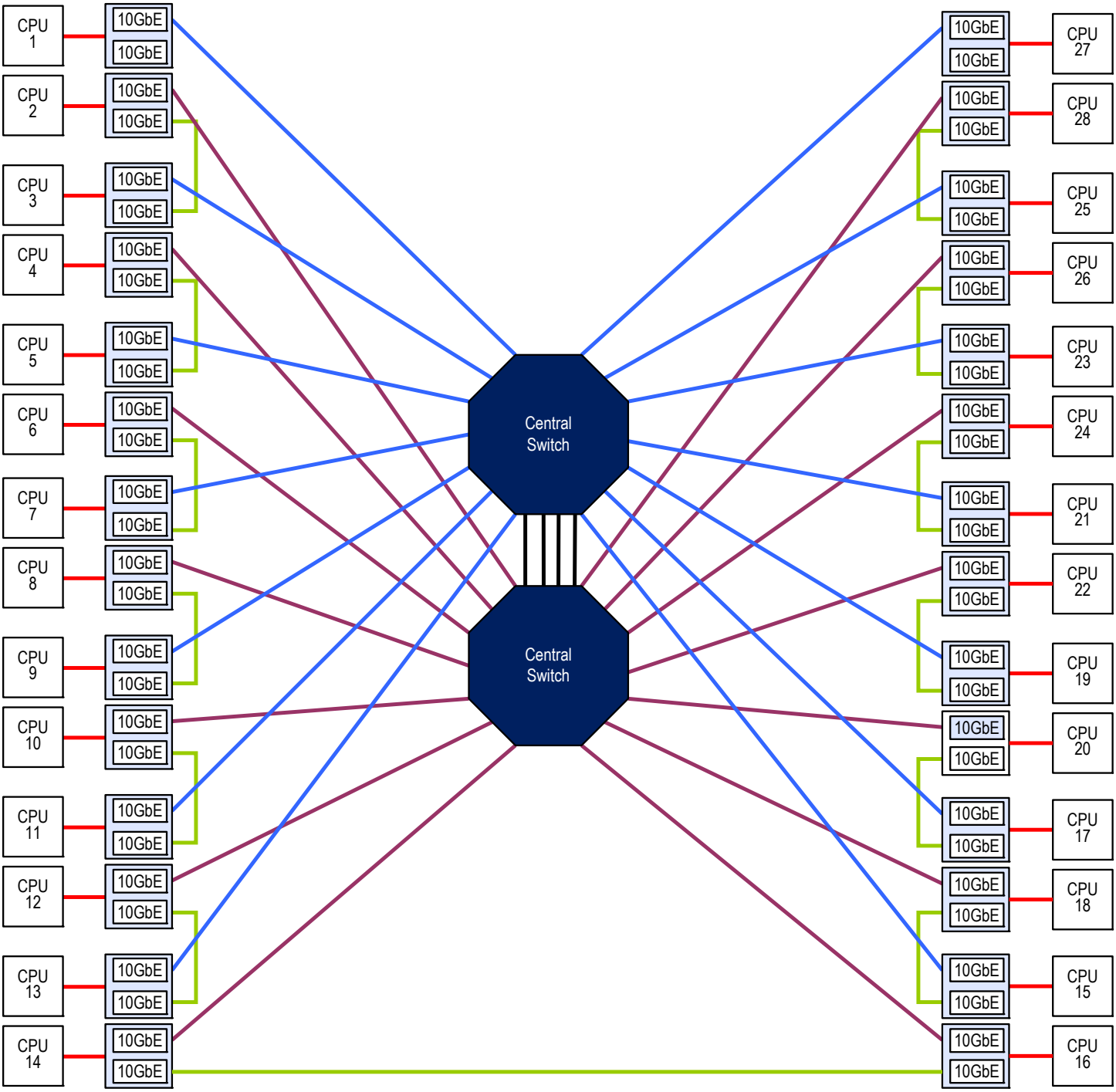


Figure 6:System Architecture with OpenVPX and 10 Gigabit Ethernet



CURTISS WRIGHT Controls Defense Solutions

System Bandwidth Calculations - All-to-all Case

The 10GbE System

The 10GbE system has 28 processors. (14 payload slots, 2 CPUs/payload). In the all-to-all case each CPU has bidirectional transfers of equal throughput with the other 27.

Each CPU has one direct connection to another CPU via the adjacent slot connections. (e.g. CPU2-CPU3). Each CPU has 13 connections through a single switch (e.g. CPU1-CPU3) and 13 connections that traverse both switches (e.g. CPU1-CPU4).

In total, there are 756 connections. (CPU1-CPU2 and CPU2-CPU1 count as 2). There are 368 connections that require two "switch-hops" where the data must traverse the switch-to-switch connections. The total inter-switch bandwidth is 80Gbps (four connections, bi-directional).

The bandwidth available for each of the 368 two-switch-hop connections is therefore 80Gbps/368 = 0.217Gbps. To restate, in this system any of the CPUs will be able to sustain simultaneous 0.217Gbps transfers between itself and the 13 of 27 CPUs that are two switch hops away.

The remaining 368 connections traverse just one switch and can use the un-used bandwidth of their NICs. These connections therefore each have available (10Gbps-13*0.217Gbps)/13 = 0.552Gbps

In the all-to-all case, it is the lower bandwidth connections that will dictate the system performance and therefore the conclusion is that **each CPU has 0.217Gbps available**.

The SRIO System

This system has the same number of CPUs (28). However, the topology differs significantly from the 10GbE system, having more connection paths available to the CPUs.

As before, each CPU has bi-directional dataflow with the other 27. One connection is on-card via the local SRIO switch. Four connections will make use of the adjacent card links (e.g. CPU3-CPU1/2/5/6). This leaves 22 connections that will be made through the central switch. The board has two CPUs sharing a pair of 16Gbps connections to the switch. Therefore each CPU has 16Gbps available to connect with the other 22. These connections though the switch will be the lowest bandwidth and therefore will determine the performance in this example.

Each CPU has 16Gbps/22CPUs = 0.73Gbps bandwidth to each of the other CPUs.

System Bandwidth Calculations - Pipeline Case

We use a pipeline example as an alternate benchmark of the system bandwidth. In a pipeline, each CPU is transmitting data to another CPU in the system, which in turn is transmitting to another, and so on. All 28 CPUs are transmitting and receiving equally. We permit the order of CPUs to be optimized for each system independently since that is what a system designer would do in reality.

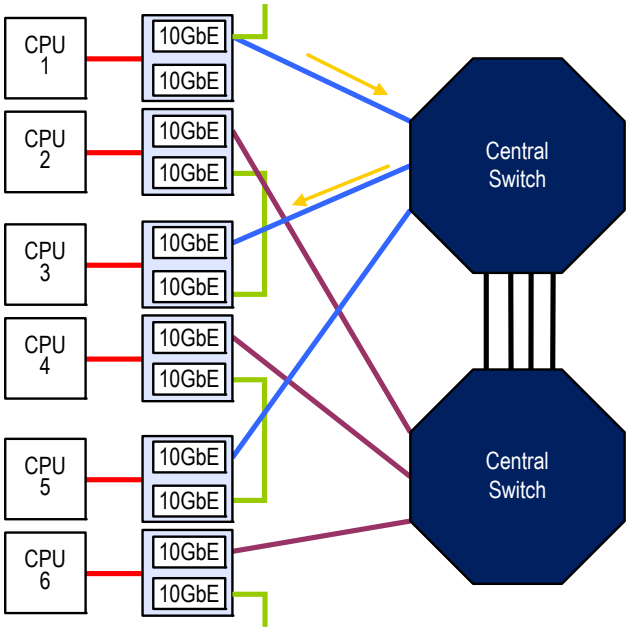
The 10GbE system

The dataflow paths for the 10GbE system are shown in Figure 7. The optimal ordering of CPUs is CPU1-CPU3-CPU5. Trying to route from CPU1-CPU2-CPU3 results in all the traffic crossing the switch-to- switch links, which would quickly saturate. At CPU27 the data will cross between the switches to get to CPU2 (or any even numbered CPU) and the sequence continues through the rest of the system.

So each CPU is fully utilizing its single link to the central switch and **therefore the bandwidth available to each CPU is 10Gbps**.

Figure 7: 10GbE Pipeline Dataflow 1-3, 3-5

10GbE Pipeline CPU1 to CPU3



10GbE Pipeline CPU3 to CPU5

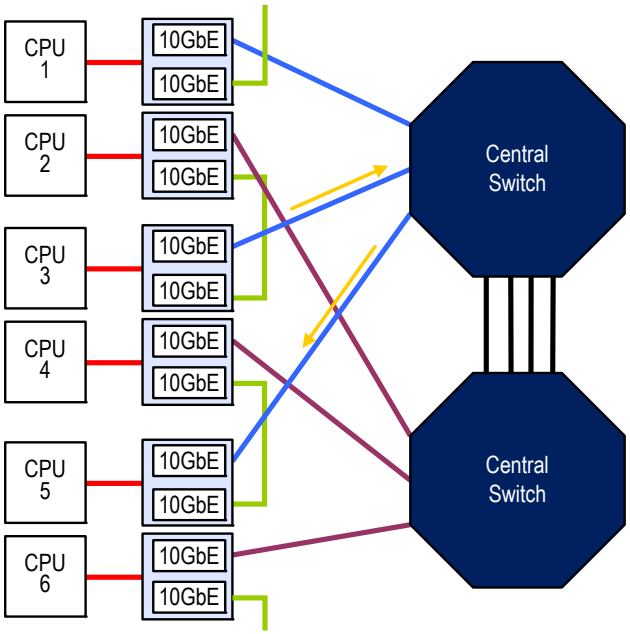
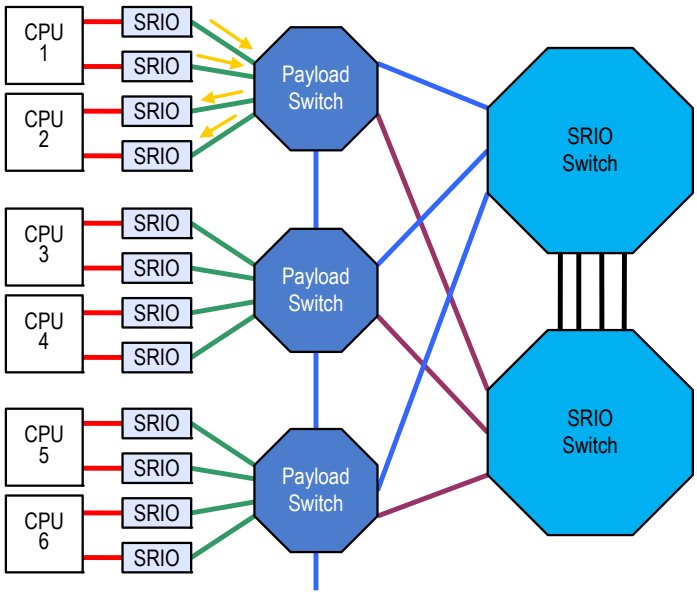
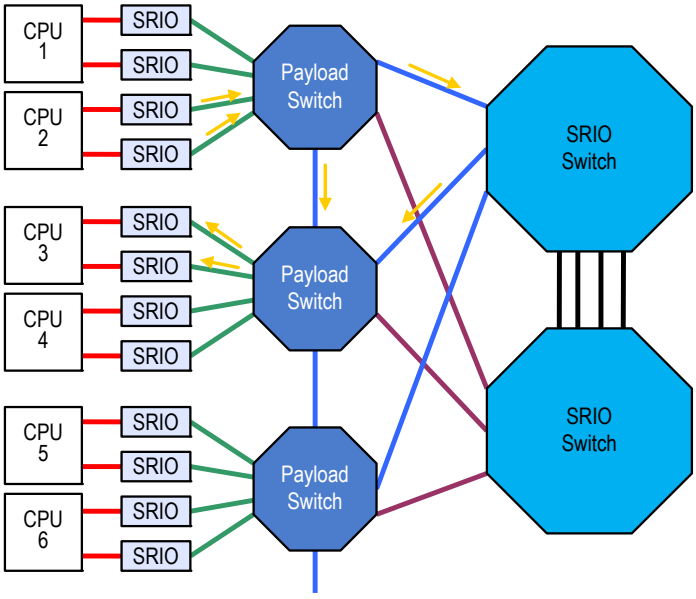


Figure 8: SRIO Pipeline Dataflow 1-2, 2-3

SRIO Pipeline CPU1 to CPU2



SRIO Pipeline CPU2 to CPU3



CURTISS WRIGHT Controls Defense Solutions

The SRIO System

The dataflow paths for the SRIO system are shown in Figure 8. The optimal ordering of CPUs is CPU1-CPU2-CPU3. This routing pattern will allow use of the local card switch, the direct slot-to-slot connections and the central switches.

The CPU1-CPU2 traffic takes place on card. There is a 32Gbps bandwidth between the two.

To understand the CPU2-3 path, it is useful to observe that there are two SRIO endpoints at each end, which would support up to four separately routable paths through the fabric. The logical choice in this case is to route half the traffic through the direct connection and the other half though the central switch. Again this results in a 32Gbps connection between CPU2 and CPU3.

Therefore the SRIO pipeline system provides each CPU with 32Gbps.

Summary

Table 1 summarizes the system bandwidth comparisons. In both of the benchmark dataflow patterns, the SRIO system has more than three times the bandwidth available to the CPUs than the 10GbE system.

The performance advantage of SRIO is attributable to two key effects. The first, obvious benefit stems from the 1.6x faster signaling rate of Gen2 SRIO (16Gbps Vs 10Gbps). However, even if both fabric technologies operated at the same signaling rate, the SRIO systems would still have an approximate 2x performance advantage. This is the due to the presence of the additional switching resources in the SRIO systems, where it is practical to locate an SRIO switch on each payload card. In the all-to-all case, the local switch is the key element that eliminates the switch-to- switch link bottleneck. In the pipeline case, the switch permits the simultaneous use of two connections between every pair of CPUs rather than just one connection.

Table 1: Summary of System Bandwidth Calculations 10Gbe and SRIO

Backplane	Use Case	10 GbE	SRIO	SRIO Advantage
CEN16 14 payload 2 switch	All-to-all	0.217Gbps	0.73Gbps	**3.36x**
CEN16 14 payload 2 switch	Pipeline	10Gbps	32Gbps	**3.2x**

Contact Information

To find your appropriate sales representative:
Website: www.cwcdefense.com/sales
Email: defensesales@curtisswright.com

Technical Support

For technical support:
Website: www.cwcdefense.com/support
Email: support1@cwcembedded.com

CURTISS WRIGHT Controls
Defense Solutions

defensesales@curtisswright.com
cwcdefense.com