StarLink

Switched Fabric Data Interconnect PMC

Features

- Switched Fabric Interconnect PMC
- 400Mbytes/sec total sustained bandwidth
- Zero protocol, PCI to PCI transfers converted to packets which automatically route through the fabric
- 64-bit, 66MHz PCI Interface
- PCI to StarFabric bridge (2 ports)
- 6 port StarFabric switch
- 4 ports via PMC Pn4 connectors or
  - 2 ports via front panel RJ-45's & 2 ports via Pn4 connectors
- 800 Mbytes switching cross-sectional bandwidth
- Quality of service features for real-time determinism
- Robust LVDS physical layer
- High availability features
- Low power operation
- VxWorks drivers
- Air-cooled and conduction-cooled versions available.

Description

Many Signal Processing problems demand the use of multiple processors to achieve real-time throughput and response times. For applications of this type, it is invariably necessary to share large amounts of data between the processing nodes. The StarLink PMC adaptor provides a modern technology solution to this problem with many specific benefits for the military aerospace systems designer.

The StarLink PMC card provides the user with a flexible, switched fabric board interconnect system that easily scales from a few to many boards. The flexibility of the system is a virtue of the underlying packet switching technology, where data is automatically and transparently routed through a fabric network of switches to it's destination.

StarLink is based on StarFabric technology from Stargen Inc. StarFabric is a high-speed serial, switched-fabric technology. The system is based on two types of device, a PCI to StarFabric bridge, and a StarFabric switch. The elegance of StarFabric is in it's simplicity, as a network of bridges and switches presents itself to the application as a collection of bridged PCI devices. Memory attached to one node, is made visible in PCI address space to the other nodes in the network. This is an existing architecture in many systems (i.e. cPCI), so from a software interface perspective, a group of cards linked through StarFabric network, appears the same as if they were connected to each other through non-transparent PCI to PCI bridges.

Mission-Critical Features

StarFabric was designed to meet the many demanding requirements of high-availability systems used in the telecom industry. Many of these requirements have parallels in military real-time computers, such as fault detection and recovery, redundancy, quality of service and low power consumption. The StarLink interconnect system provides a rich set of these features, combined with a low latency, high throughput data flow capacity.

Features and Performance

The StarLink PMC card is implemented with two StarFabric devices, a PCI to StarFabric bridge and a six port StarFabric switch. The bridge provides two ports which are connected to the switch. The remaining four ports of the switch are accessible externally. Systems are constructed by simply interconnecting between ports on the cards involved. See Figure 1, StarLink Block Diagram.

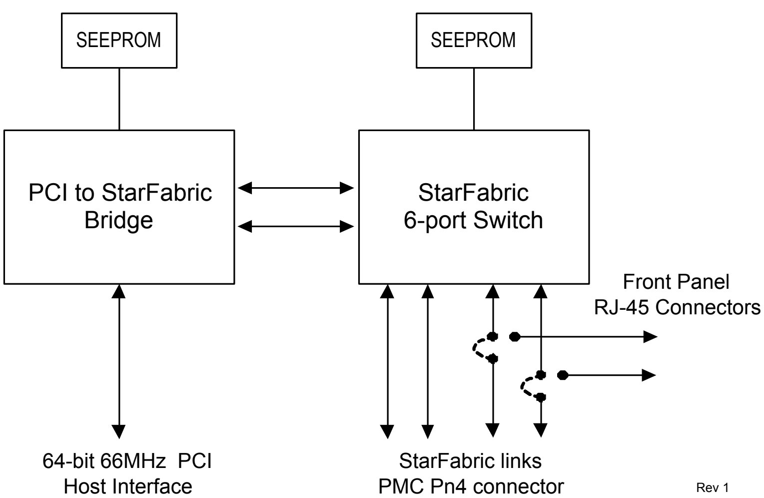

Figure 1: StarLink Block Diagram

StarFabric links are point to point, full duplex, and operate at a link rate of 2.5Gbps. Accounting for the overhead of 8B/10B encoding and packet headers, each link is capable of sustained transfer rates of 200Mbytes/sec, in each direction simultaneously. It is possible to logically bundle two links together to create 400MB/sec connections between nodes. The fabric will automatically recognize the parallel path and permit two links to behave logically as a single, higher bandwidth connection.

StarLink supports multicasting in hardware. Data packets are automatically replicated by the switches (in hardware) as needed and sent to multiple nodes. Applications that need to share common data between many processing nodes can be greatly accelerated. Applying this feature is done by providing for independent routes from the transmitting node to the multiple receiving nodes. Up to 32 different multicast groups can be defined. See Figure 5, Multicast Example.

StarLink supports "quality of service" features which are beneficial to ensuring correct real-time behavior in a system. In most real-time systems, there is a mix of critical "real-time" data, and non-real time control messages that flow in the system. StarLink provides mechanisms to ensure priority of the former over the latter. This means a developer who has achieved a correctly functioning real-time system, has the option to further exploit the remaining unused bandwidth in that system for non-critical data, without fear of disrupting the real-time design of the system. In existing systems it is common for developers to employ alternate data paths such as the VMEbus or Ethernet to act as secondary low performance data channels, to avoid the risk of mixing both types of traffic one a single system. StarLink eliminates this risk, thus allowing the simplified model of using a unified data transfer system.

Flexibility

The StarLink interconnect system can be configured in many different topologies. The system supports simple point to point connections, basic mesh topologies, and more elaborate topologies with redundant connections for high availability consideration. The system supports routing of packets through up to seven switches, so systems can be scaled to extremely large numbers of nodes. A high-availability system can be constructed by providing for redundant routes between end points. A failure in any one connection will be detected and automatically re-routed over the remaining good connection. Failures are reported so that the application software can take appropriate action.

The StarLink PMC adaptor provides both the StarFabric bridge and a switch. A interconnected DSP system is constructed solely of these components, with associated interconnecting wiring or backplane. There are no active backplane overlay modules, active hubs, or special cards required. As a result, the logistics costs of maintenance and sparing are the minimum possible. During a development project, reconfiguring the system requires little more than re-arranging standard category-5 cables and re-initializing the software to perform the network discovery process.

Reliability/Availability

StarLink provides a number of features which permit the construction of high availability systems. The link layer incorporates 8B/10B encoding and CRC checking between nodes, and will re-transmit frames as needed. This checking is done on the fly by the hardware, incurring no performance penalty. Transmission failures are reported to the software layer. It is not necessary to add additional "test channel" software to monitor the integrity of the network.

A StarLink fabric can be arranged in many different topologies, so as to suit the data flow and redundancy requirements of the system. For high availability systems, it is possible configure a network without any single points of failure. This is done by ensuring that redundant paths exist between nodes in question. During initialization, the fabric will pre-determine alternate routes between nodes, and automatically re-route data to accommodate a failure in wiring, or in a switch.

Physical Layer

The physical layer of StarLink is based on Low Voltage Differential Signaling (LVDS) connections, operating at 622Mbits/second. Connections are full-duplex, and consist of four pairs in each direction. (16-pins total). Note the natural fit to PMC and VME standards that provide 64-pins of connectivity. (Pn4 and VME P2). The 8B/10B encoded LVDS link layer is well suited to the electrical environment found in a typical embedded computer system. StarLink can be used with off the shelf category-5 unshielded cables up to 5M in length. (same as used for 100BaseT Ethernet). StarLink can be used between cards in a backplane, and also chassis to chassis. This physical layer has many benefits applicable to the design of rugged deployed systems such as:

- Use of conventional PWB materials and tracking techniques in backplanes
- Use of standard VME and cPCI connectors
- No need for coaxial cabling routed to the backplane connectors
- No need for expensive co-axial 38999 connectors at the chassis
- No extra termination networks
- No TTL signal quality issues, with edge rates affected by temperature
- Use of standard Category 5 cables for development

Software Support

StarLink is supported with a VxWorks driver. The applications interaction with the driver is primarily to initialize and configure the various operating modes such as quality of service, shared memory allocation etc. The driver, in conjunction with the hardware performs the entire network discovery process. This includes determining the routing tables which govern how packets flow from source to destination. Once the initialization of the network is complete, processing nodes can transfer data between themselves with conventional memory to memory mechanisms, including, but not limited to DMA transfers. There is no protocol required for data transfers, but one is not restricted from implementing a protocol on top of the shared memory mechanism that StarLink provides. Other operating system support is planned - contact your local representative for further information.

Options

The StarLink PMC is available in two configurations, one with all four ports connecting to the Pn4 connector for backplane I/O usage, and the other with two ports on the front panel and two ports on the Pn4 connector. A rear-panel transition module will be available for VME systems. This module plugs into the P0 and P2 connectors at the rear of the backplane. It provides access to the four StarFabric ports via RJ-45 connectors, and supports one or two StarLink PMC cards installed on the host.

Application Example #1

An example application is presented in the following section, with three diagrams representing the example system from different perspectives. In Figure 2, Application Example: Network Topology, a network is constructed from the two fundamental StarFabric components, edge nodes and switches. The edge nodes are the termination points for packets crossing the fabric. The StarLink PMC uses a StarFabric to PCI bridge that contains two edge nodes. The other component shown is a StarFabric switch. The system is comprised of 10 edge nodes and 5 switches, corresponding to the use of five StarLink PMC modules.

The network is configured in a mesh topology where each switch has a direct connection to each of the other four switches. This a generic topology, which by virtue of it's symmetry is suited to a random data traffic patternwhere each node is sending equally to all the other nodes. For this scenario it can be shown that the fabric has much more capacity than needed to manage the throughput available from the two StarFabric ports on each PMC.

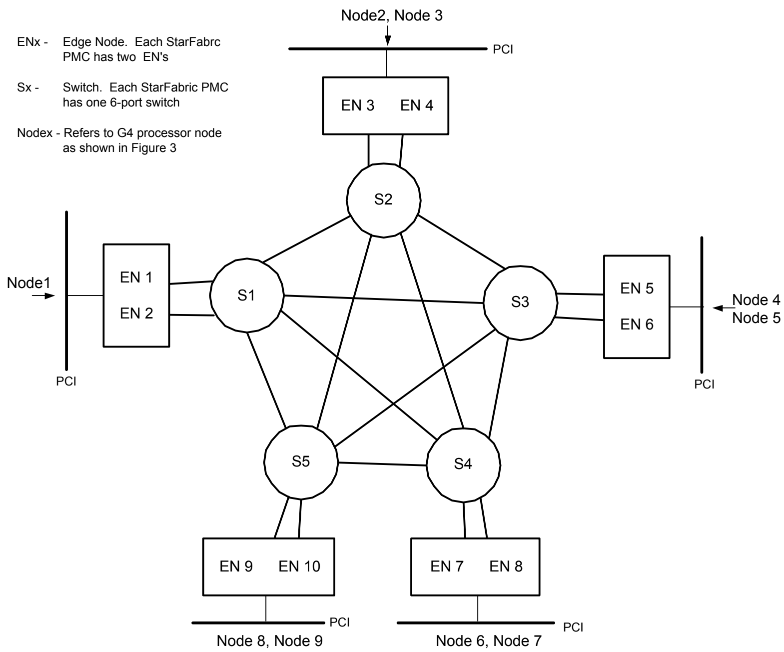

Figure 2: Application Example: Network Topology

Take for example the link between S1 and S2. Each edge node sends 1/9 of it's traffic to each other node. (2.5Gbps/9 = 0.278Gbps). The traffic on the link is then.

EN1 to EN3 , EN3 to EN1 = 0.556Gbps

EN1 to EN4, EN4 to EN1 = 0.556Gbps

EN2 to EN3, EN3 to EN2 = 0.556Gbps

EN2 to EN4, EN4 to EN2 = 0.556Gbps

Total = 2.224Gbps.

The link has a 5 Gbps capacity which is more than twice the data traffic load, demonstrating that for the random data distribution case, the fabric has more than sufficient capacity.

In Figure 3, Application Example, System Block Diagram , an actual hardware implementation of the network is presented. This example is of a small processing system comprised of a Single Board Computer and two quad PowerPC DSP boards. The DSP boards each carry two StarLink PMC's. Doing so, provides each DSP with the highest possible I/O bandwidth. In many systems, one StarLink PMC will be sufficient to manage the data I/O requirements of the application. In Figure 4, three memory maps, (simplified), one for each of three processors are presented to illustrate how the network presents itself to an application program. One processor from each card is selected for the purpose of discussion.

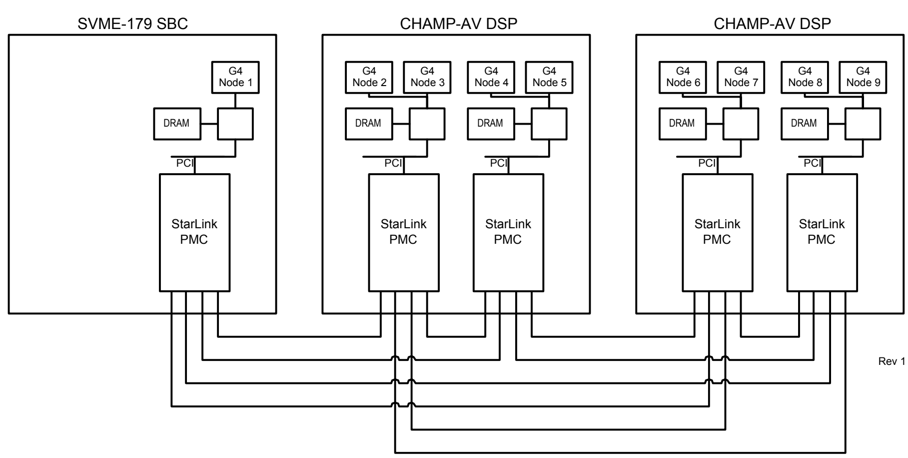

Figure 3: Application Example: System Block Diagram

In the upper half of each memory map, is the PCI space as seen from the local processor. Within that address range, are blocks that map to physical SDRAM on another card somewhere in the fabric. In this example, Node 1 (on the SBC) has regions which are mapped to all of the other nodes in the system. Up to 1024 nodes may be mapped in this manner. When the Node 1 processor reads or writes into this PCI space, the StarFabric network routes the data to the appropriate node across the fabric. The simplicity of this from a software perspective is noteworthy. The processor can read and write single memory locations, or for high performance applications, the data can be moved by DMA controllers in large blocks.

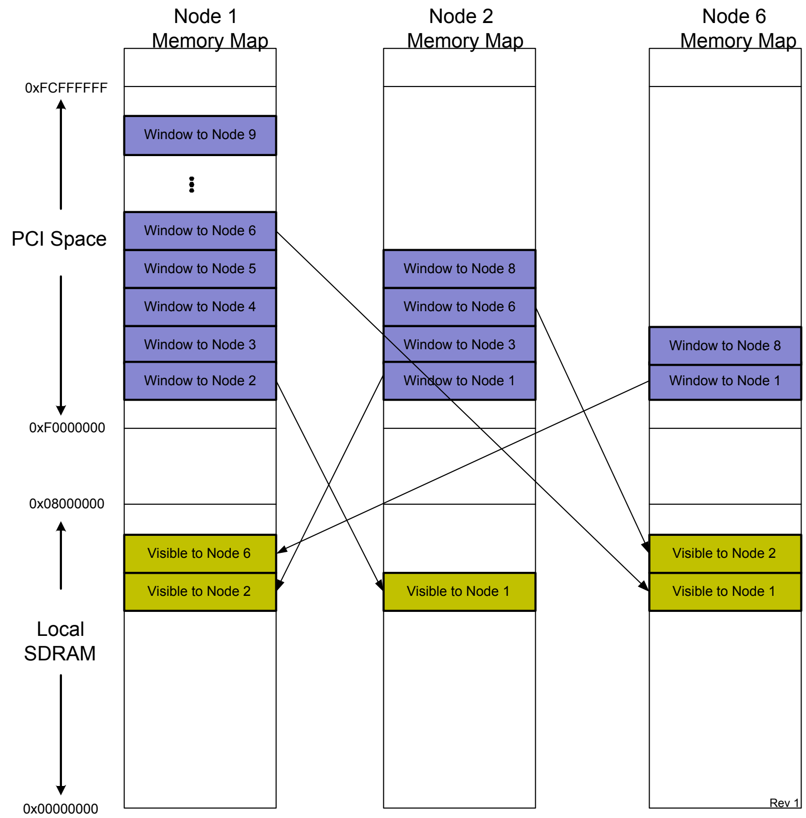

Figure 4: Application Example: Simplified Memory Map

In the lower half of the memory map, is the local SDRAM. In each nodes local SDRAM, are address ranges where another node in the fabric can read and write this memory. If desired, more than one external node can be mapped to the same address.

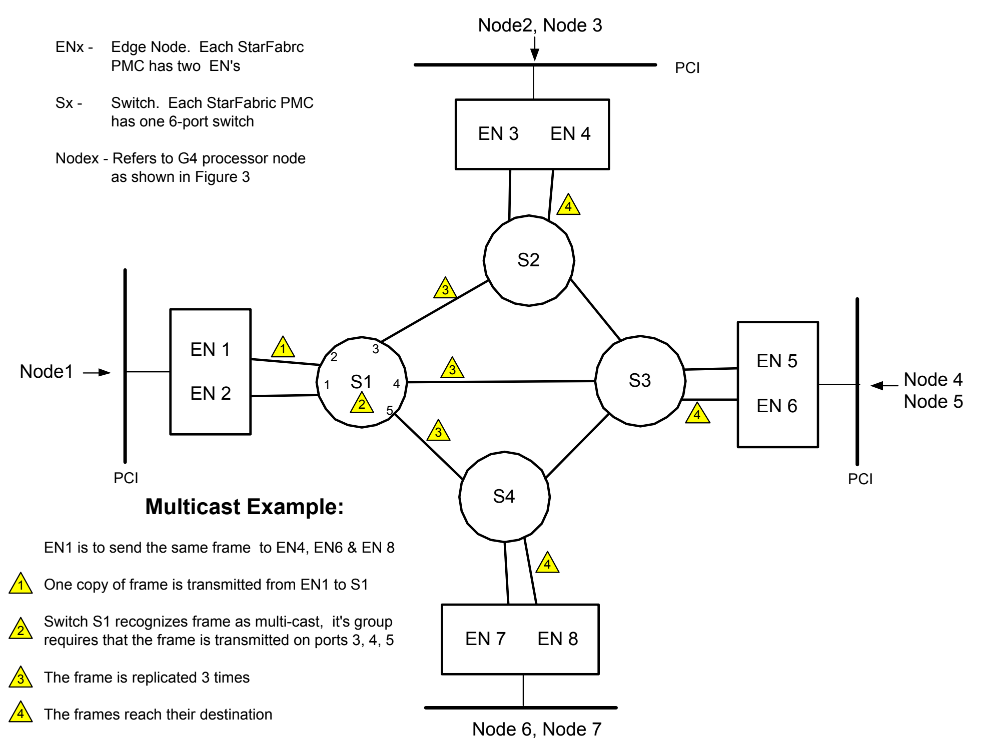

Figure 5: Multicast Example

Application Example #2

The second application example illustrates how a system can be scaled to larger configurations using the StarLink PMC. In Figure 6, a 16 Board Example, a network is constructed using four clusters, of four boards each. Within each cluster, is a high bandwidth fabric. Each cluster is then also to make four connections to other clusters, which in this case are arranged with one connection to the other three, and a two links to one of the three, which depending on the application represent a higher bandwidth requirement, or also a redundant connection existing at a chassis level. As before, any node on the fabric can communicate with any other node. There are also redundant paths for any connection in the fabric, so that the failure of any board, or any external link will not bring down the operation of the rest of the system.

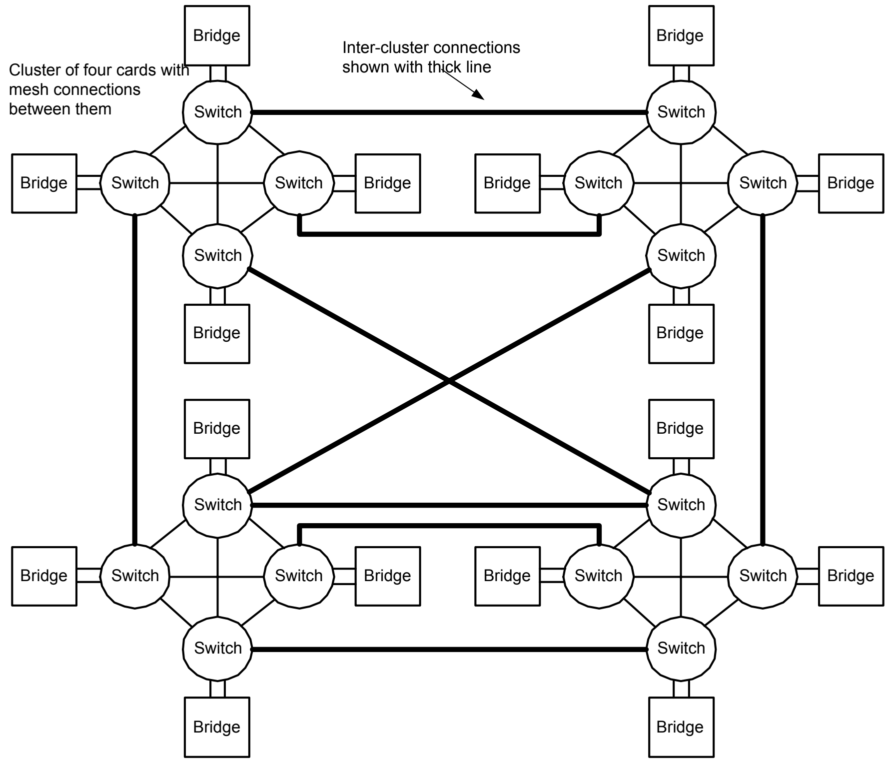

Figure 6: 16 Board Example

Specifications

Table 1:

RUGGEDIZATION LEVELS*		
SPMC card	Available in level 0 and 100	
DPMC card	Available in level 100 and 200	
POWER REQUIREMENTS		
+3.3V	1.3A Typical	
DIMENSIONS		
	Size	**Weight**
SPMC card	per IEEE 1386.1	<150 g (<0.33 lb.)
DPMC card	per IEEE 1386.1 (VITA 20-199x)	<160 g (<0.35 lb.)

* See Ruggedization Guidelines

Contact Information

To find your appropriate sales representative, please visit:

Website: www.cwcembedded.com/sales or

Email: sales@cwcembedded.com

For technical support, please visit:

Website: www.cwcembedded.com/support1

Email: support1@cwcembedded.com