SCRAMNet GT

A New Technology for Shared-Memory Communication in High-Throughput Networks

Introduction

A number of today's most advanced networking applications require the fast transport of large amounts of data, and the real-time sharing of that data among the nodes on that network. Such applications include simulation systems (aircraft simulators, missile simulators, power generation plant simulators, mission planning simulators, force-level training simulators, virtual reality simulators and location-based entertainment systems), as well as feedback loop-controlled systems (range and telemetry systems, flight testing, wind tunnels, test stands, robotics, particle accelerators and industrial process control systems). In each case, these networks require the high-speed, real-time data distribution and sharing of a mixture of data–including video–by all the various computers connected to that network.

This paper will describe SCRAMNet GT, a new high-throughput technology for connecting multiple computer platforms to form a single, real-time, distributed processing system in which memory is shared among the computers.

The Challenge of Real-Time Data Sharing

Systems such as those described above, may consist of a number of different computing platforms, each performing computations or analyzing portions of the same real-time problem. These platforms can be running at different speeds and be running different operating systems, yet all must have access to the same data in real time. In other words, all the nodes must have access to all the data, all the time.

To be successful, such a shared-memory, distributed processing system must have a number of high-performance characteristics, including:

<u>High Throughput</u>. Sustained high data transfer rates (hundreds of megabytes per second) must be achievable to handle a mixture of high-speed video data and control data.

<u>Low Latency</u>. Data transfer from platform to platform, or from application to application, must take place very quickly. It is unacceptable for Operating Systems (OS) or network overhead to create time delays while requesting writes to memory or constructing memory packets.

Instead, it must be possible for data to be written directly to a memory location.

<u>Deterministic System</u>. Availability and location of data must be known precisely at every moment. Again, this means that no OS (or File System) intervention is acceptable, as the OS decides when data reads and writes occur. The resulting delays in data access are undesirable in a system that shares data in real-time.

<u>Platform Independence</u>. The various nodes in a real-time distributed processing system are likely to be running different OSs and different hardware. Systems

running Windows®, Linux®, IRIX, VxWorks®, etc., should be accommodated. Hardware may be utilizing PCI or VME bus architectures, supplied by any number of manufacturers.

High Reliability. The system should be able to manage data errors, as well as the failure of a node, fiber, or connector, without adding overhead to the system.

The Development of a Solution

In the 1980s, Curtiss-Wright Controls Electronic Systems developed the first-generation SCRAMNet Network specifically to solve the communication needs of distributed real-time computing applications. SCRAMNet was based on the innovative "replicated shared-memory" concept, which places identical computer-addressable memory into each network computer node (see Figure 1). This additional memory is known as replicated shared memory because it contains all data to be shared among the interconnected computers.

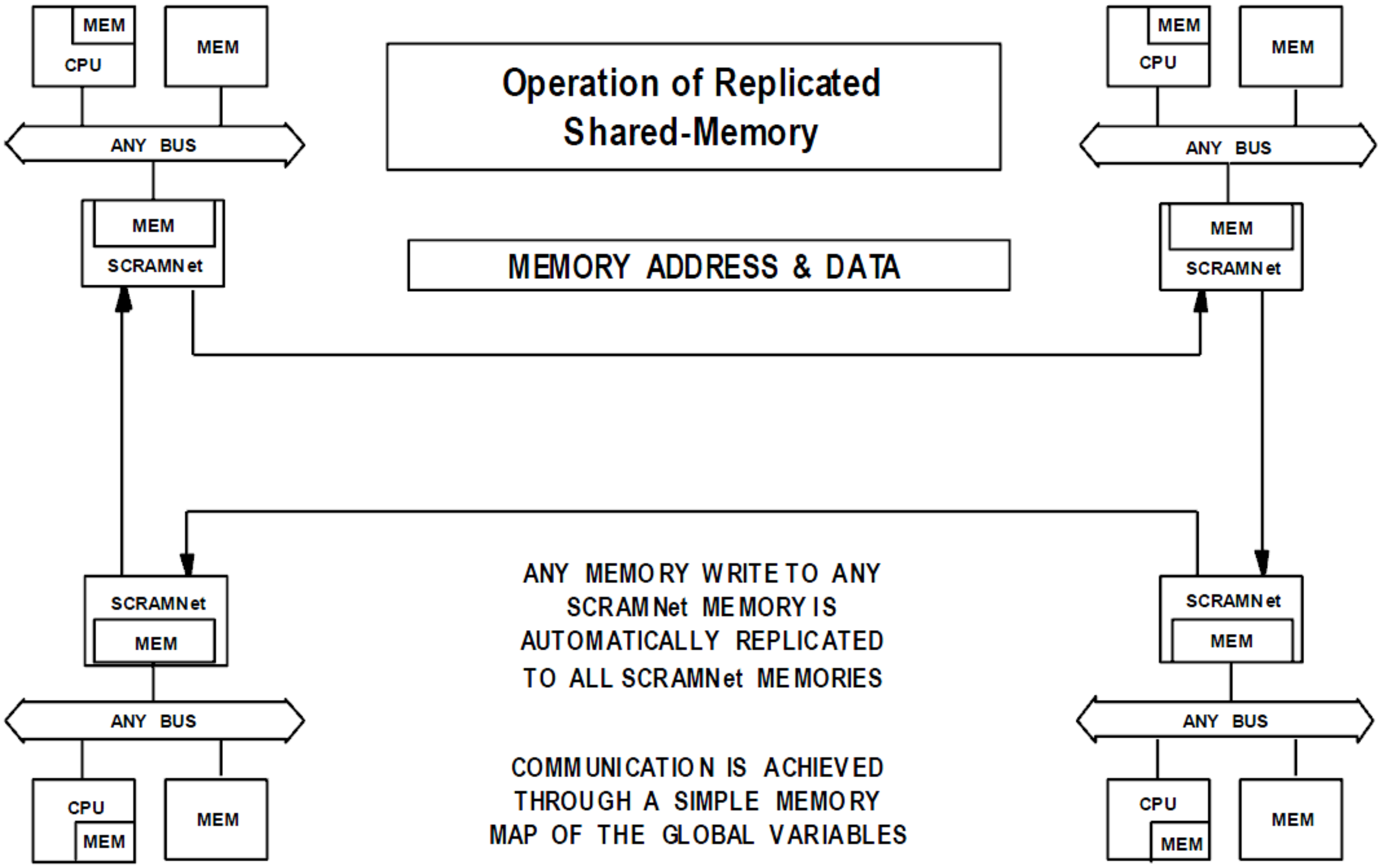
With SCRAMNet, memory writes to the replicated shared memory at one computer are instantly sent to all other replicated shared memories on the network via a high-speed connection. There are no bus contentions, time-consuming protocols or nondeterministic message crashes. SCRAMNet enables the distributed computer system to appear as one large, powerful "virtual" multi-processor to real-time applications.

For nearly two decades SCRAMNet, and its later generations SCRAMNet LX and SCRAMNet+, became the de facto standard real-time network interconnect. Literally thousands of SCRAMNet cards are in operation today, embedded in a variety of real-time applications worldwide.

The Evolution of a Proven Solution

As computers have become more powerful and processing capability has increased, real-time system designers have taken advantage of these advances to place even greater demands on a real-time

Figure 1: Example of Replicated Shared-Memory Configuration

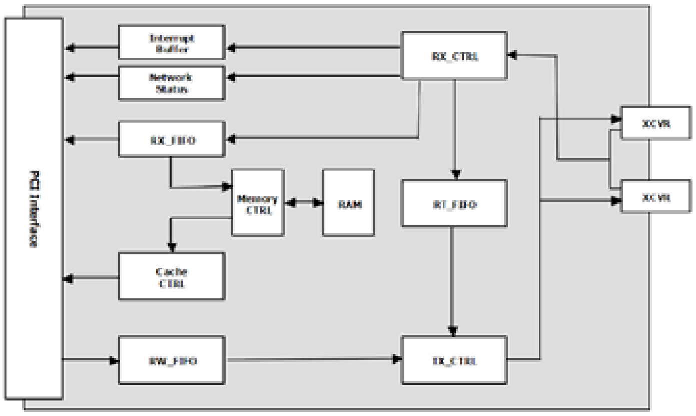


interconnect design. The data requirements of these powerful new real-time applications require more memory and greater network bandwidth than ever before. To specifically address these new requirements Curtiss-Wright Controls Electronic Systems has developed SCRAMNet GT (Greater Throughput), a next generation replicated shared memory network that incorporates the fundamental strengths inherent in prior SCRAMNet products while adding more network bandwidth, more shared memory, and additional network features that enable system designers to simplify overall design while significantly increasing performance.

SCRAMNet GT's high throughput allows data such as video streams, data from high-speed radar and/or sonar sensors, lower-speed control signals – or a mixture of these data types – all to be carried over the same network, and processed concurrently by hosts (nodes) distributed throughout the network. Larger amounts of shared memory and network bandwidth, coupled with the efficiency of the replicated shared memory architecture, ensure that all computers in the system are operating simultaneously on the same set of data.

Overall performance capabilities of SCRAMNet GT include a throughput greater than 200 MBytes/sec, with a latency of less than 1 μsec per node. Built-in broadcast capability, (one-to-many and many-to-many), ensures that all nodes have updated information without host or user intervention.

Figure 2: SCRAMNet GT Block Diagram



The heart of SCRAMNet GT is the network card (see Figure 2) installed into each host computer. Available in a variety of form factors, SCRAMNet GT network cards function like intelligent memory/network interface devices. They incorporate an on-board processor, memory, and software that makes the functioning of the board transparent to the user. Initial offerings of SCRAMNet GT cards will include 128 MB of 125 MHz, 32-bit DDR SDRAM, with higher-capacity versions to follow.

Data is written directly to a memory location, without OS or File System overhead. There is no need to packetize the information prior to sending it out onto the network for distribution. The local processor on each SCRAMNet board uses an efficient Register Insertion protocol to put data onto the network. All networking protocol activities are handled without intervention by any host, and are completely transparent to all users.

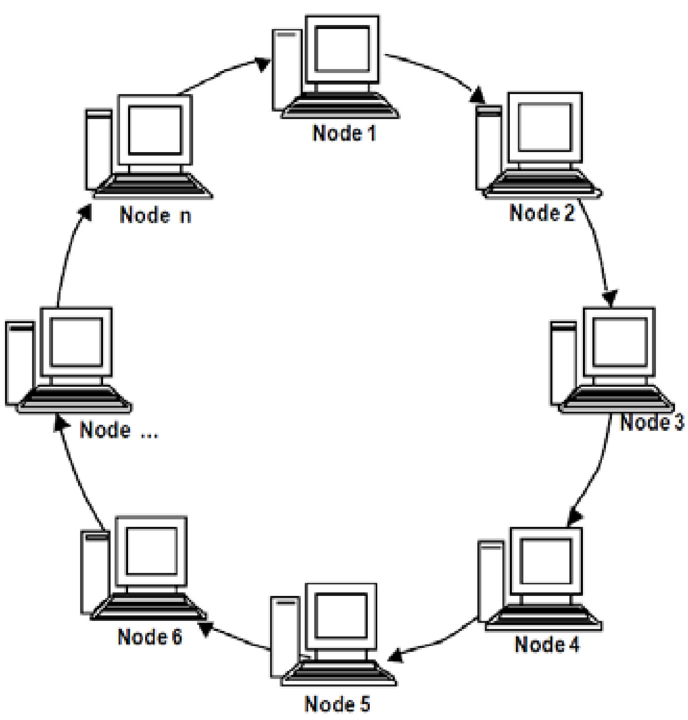
SCRAMNet GT, operating at 2.5 Gb/s, is designed to accept the user's choice of industry-standard Small Form-Factor Pluggable (SFP) transceivers, for physical interconnection to the network. For use at distances up to 300 meters, 850 nm multimode optical SFP transceivers can be used. For longer distances, 1300 nm singlemode optical SFP transceivers can be installed, for operation as far as 10 km. Other SFP transceivers can also be used. The use of COTS transceivers, fiber optic cable, and connectors, contributes to the cost-effectiveness of this system, while still providing state-of-the-art communication speeds.

SCRAMNet GT uses a ring network topology to connect the nodes (see Figure 3). The use of passive switches and cabling options allows other topologies to be created.

Since the memory on the network cards is separate from the host's main memory, each host continues to make its own decisions about how and when to use the shared data, each host continues to run at its own speed–independent of the speed of the network or of the other hosts–and each host continues to manage and use its own main memory independently of the shared network memory.

The network boards take care of all communications activities, a host sending updated data does not need to know which node will utilize the data. Each network board receives all transmitted data, and a decision whether or not to actually use the data is made locally at each node. As a result of this fundamental design philosophy, the network functions in a truly modular fashion. Nodes can be added to, or removed from, the network, with no re-programming required.

Figure 3: SCRAMNet GT Ring Topology



SCRAMNet GT has been designed for a high degree of fault tolerance, with built-in CRC error detection, and the option for redundant communication over a second transceiver. In the event of a network problem, data can be routed around linked fibers or nodes. This capability is controlled by the application.

SCRAMNet GT represents a significant increase in throughput over Curtiss-Wright Controls Electronic Systems' SCRAMNet + system, yet it maintains the same shared-memory approach. This facilitates quick upgrades from SCRAMNet + to SCRAMNet GT.

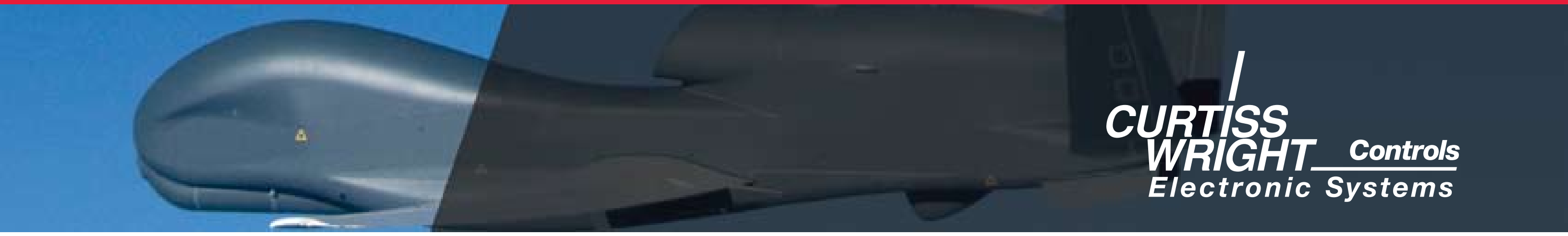
Considering Alternative Technologies

A number of high-speed networking technologies have been introduced in recent years, and at first glance, these might appear to be good candidates for implementing a high-throughput, shared-memory system. These technologies include:

Fibre Channel (FC). FC was designed primarily for computer-to-storage transfers in industrial and commercial applications. Speeds of 1.0625 Gb/sec and 2.125 Gb/sec are supported over fiber, at distances of up to 10 km.

Gigabit Ethernet (GbE). GbE is a the latest generation of the ubiquitous Ethernet networking protocol, very common in all types of industrial and commercial networks. Nominal data rates from 10 Mb/s to 10 Gb/s are now available. GbE provides excellent peer-to-peer communications with TCP/IP being the fundamental protocol used on the Internet. And, GbE can transmit serial data over distances up to 10 km.

These technologies provide high bandwidth communications over considerable distances using industry standard hardware. However, they also depend on the OS to determine the 'right time' to send and receive data. They also suffer from time delays due to network overhead from packetizing the data before it is transported. As a result, neither of these alternatives are deterministic, and the situation gets worse as the traffic on the network increases. Packet collisions on a busy network make it more difficult to predict when information will arrive at its intended destination.



These standard network communication technologies also tend to lack the low latency required of real-time distributed-processing systems. And none of them directly address the need for shared memory.

One other potential solution is the development of a custom design to handle both the memory management and data transport issues. While custom designs can be effective for solving the original problem for which they were created, they often are not transferable to other applications. These custom designs also require a significant investment in time and money. When amortized over a limited number of systems, this non-recurring engineering investment can be difficult to justify.

Conclusion

To successfully handle today's high throughput, distributed-processing applications, a network must be deterministic, have low latency, offer a high degree of platform independence and deliver high reliability. Information must be updated and distributed quickly, so that all data is available to all nodes at all times. With a throughput of over 200 MB/s, SCRAMNet GT meets these requirements, and it allows real-time applications to view the distributed computer system as one large, powerful, "virtual" multi-processor. SCRAMNet GT makes possible the communication of data and control information at the speed of memory itself, eliminating the time delays and overhead of both the Operating Systems and network protocols alike.