Proton Single Event Effects (SEE) Testing of the Myrinet Crossbar Switch and Network Interface Card

James W. Howard Jr.[1], *Senior Member, IEEE*, Kenneth A. LaBel[2], *Member, IEEE*,
Martin A. Carts[3], *Member, IEEE*, Ronald Stattel[3], Charles E. Rogers[3], and Timothy L. Irwin[4]

1. Jackson and Tull Chartered Engineers, Washington, DC 20018 USA
2. NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA.
3. Raytheon ITSS, Lanham, MD 20706 USA.
4. QSS Group, Inc., Seabrook, MD 20706 USA.

***Abstract* – As part of the Remote Exploration and Experimentation Project (REE), work was performed to do a proton SEE evaluation of the Myricom, Inc. network protocol system (Myrinet). This testing included the evaluation of the Myrinet crossbar switch and the Network Interface Card (NIC). To this end, two crossbar switch devices and five components on the NIC were exposed to the proton beam at the University of California at Davis Crocker Nuclear Laboratory (CNL). Official description of the Myrinet standard appears in its entirety in an ANSI document [1].**

I. Introduction

The Remote Exploration and Experimentation Project (REE) was part of the NASA's High Performance Computing and Communications (HPCC) Program. An effort was made to place a commercial-off-the-shelf (COTS) supercomputer in space. The architecture being investigated was a multi-processor system connected to a prime control processor that was hardened for the space environment. The network system that was being evaluated for this application was Myrinet.

For this system to be useful in the space environment, the network electronics should not be the limiting radiation susceptibility factor in the overall system architecture. To evaluate the Myrinet system for space use, crossbar switches and network interface cards were exposed to protons at the University of California at Davis Crocker Nuclear Laboratory (63 MeV incident energy).

The following sections will describe an overview of Myrinet, the details of the devices that were tested, the hardware and software systems utilized to evaluate these devices, and the results of the testing.

II. Background

The concept of the REE project was to employ a cluster computer architecture to achieve supercomputer computing capabilities with commercial state-of-the-art microprocessors. A cluster architecture is a type of parallel processing system, which consists of a collection of interconnected stand-alone computers, working together as a single integrated computing resource. A typical cluster architecture is shown in Figure 1.

At the heart of this architecture is the high-speed network that ties all the computing resources together. The Myrinet networking system is a high-performance, packet-communication and switching technology that is used to interconnect clusters of workstations, PCs, servers, or single-board computers. Conventional networks, such as Ethernet can be used to build clusters, but do not provide the performance or features required for high-performance or high-availability clustering. Characteristics that distinguish Myrinet from other networks are its high-speed data rate, flow and error control, and switch networks that can scale to tens of thousands of hosts.

The network system for a cluster architecture will always consist of two pieces: the network switching and the network interface hardware. Within the Myrinet nomenclature, these are referred to as the crossbar switch (Xbar16, where the 16 is the number of nodes) and the Network Interface Card (NIC). Figure 2 shows an example of the cluster architecture using Myrinet terminology. This figure shows the REE concept of using a prime control processor that talks to all the individual nodes of the cluster via the Myrinet network. In this framework, the prime control processor would develop computational tasks to be performed, determine which node will perform that task, and then transfer that task to the computational node via the Myrinet network.

There are different Myrinet speed standards. The bandwidth of Myrinet is described as the data rate available in the “forward” direction plus the bandwidth available in the opposite direction. The Xbar and the NIC types tested here are capable of operating at the Myrinet-2000 data rate of 2 gigabits per second (GBPS) in both directions simultaneously (full duplex). Thus, the data rate is expressed as 2000 + 2000.

This work was performed at NASA/GSFC for the Remote Exploration and Experimentation Project. This project is part of NASA's High Performance Computing and Communications Program, and is funded through the NASA Office of Space Sciences. Additional resources were provided by the NASA Electronic Parts and Packaging (NEPP) Program under the auspices of the NASA Office of the Chief Engineer.

Each single-direction 2000 megabits per second (MBPS) link is referred to as a channel. The opposite-direction pair of these channels is referred to as a link.

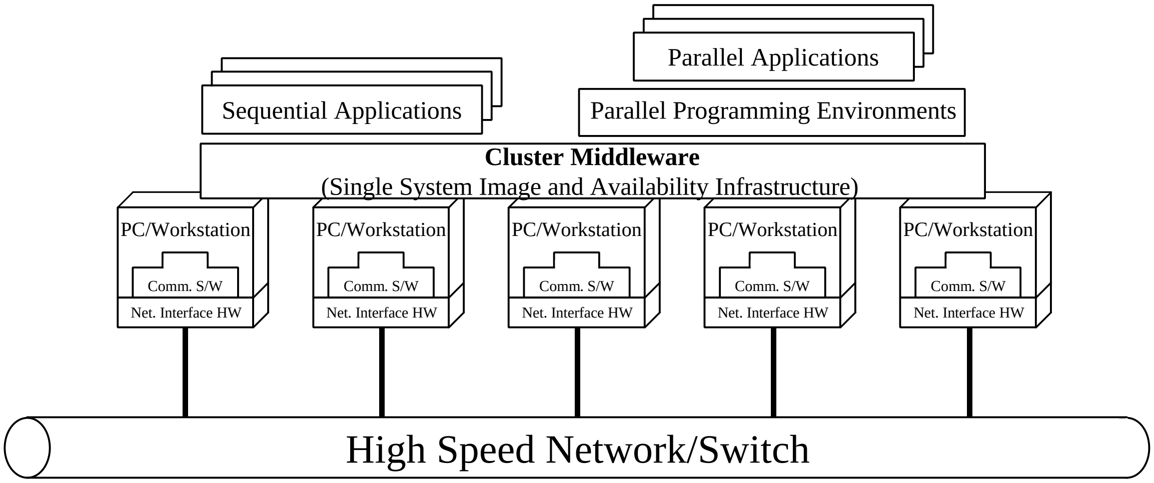

Figure 1. Cluster computer architecture. [2]

III. System Tested

A. Introduction

As shown in Figure 2, there are two components to the Myrinet network system. The crossbar switch (Xbar) device type is the essential component interconnecting devices residing on the network. The 16-port Xbar tested allows 16 devices to connect with any one other device on the network and will be discussed first. The Network Interface Card (NIC) allows each of the nodes access to the network. The NIC in its PCI-64 form-factor will be addressed second. This will then be followed by the network protocol for package construction description and a discussion of the test methodologies employed for this testing.

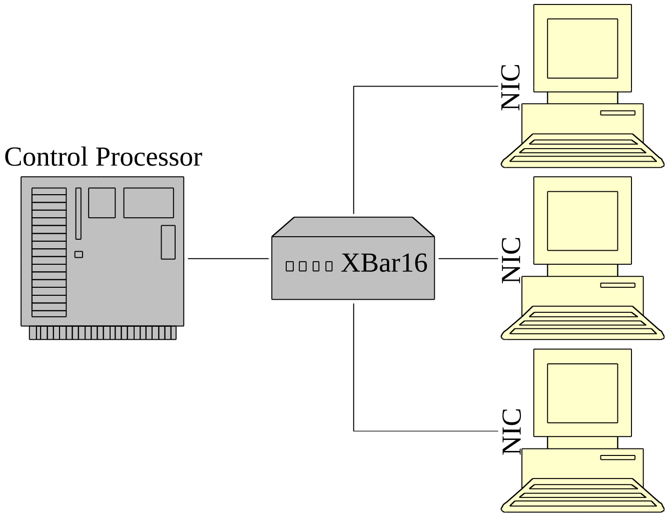

Figure 2. Cluster computer network with Myrinet hardware.

B. 16-port Crossbar

The Xbar Device Under Test (DUT) is a 0.25 µm commercial CMOS Application Specific Integrated Circuit (ASIC) manufactured for Myricom. Each DUT acquired was mounted to a printed circuit board (See Figure 3). No other components on the DUT board were directly irradiated (as with any proton test, secondary neutron scatter is possible).

Each Xbar Integrated Circuit (IC), shown in Figure 4, has 16 System Area Network (SAN) ports. The Xbar IC and SAN specification for Myrinet is described in [3]. Briefly, the SAN is a parallel data and control signals format for short haul (*i.e.*, components no more distant than a board that shares the same backplane). Eight of these SAN links are brought to the front panel through a serializer/deserializer (SerDeSer) for connection to external components. The other eight ports are connected to the backplane connector for SAN connection to other components within the chassis that hold these cards.

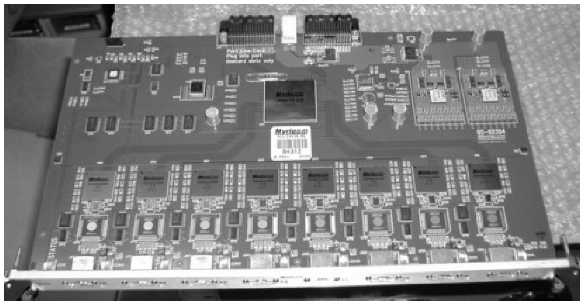

Figure 3. The Myrinet Xbar card showing the backplane System Area Network (SAN) ports at top and serial front panel ports.

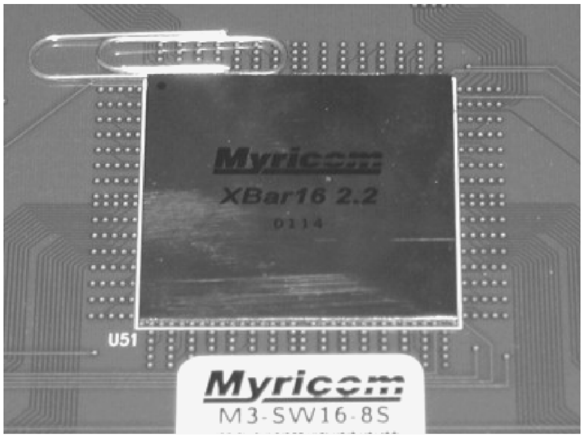

Figure 4. Close-up of the Xbar IC. The total height above the board of the IC is 1.4 mm.

A simplified block diagram is shown in Figure 5. Shown here are the eight frontplane ports (numbered eight through fifteen) linking into the Xbar DUT. The eight backplane ports (numbered zero through seven) with power and control signals are shown at the backplane interface connector (recall that only the frontplane ports have the SerDeSer conversion).

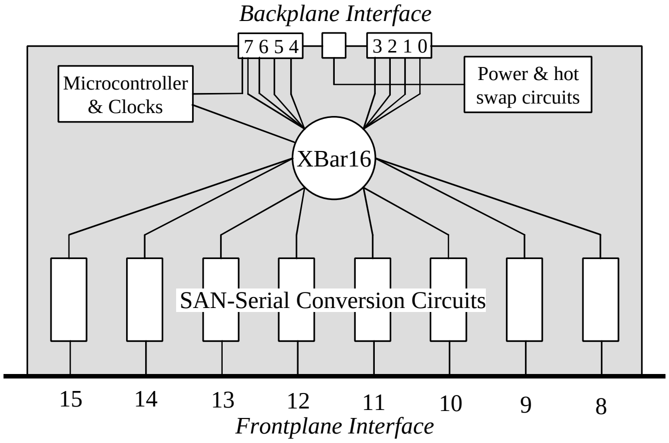

Figure 5. Block diagram of the Myrinet crossbar card.

Providing 16 serial ports requires more than one Xbar card; the other eight backplane ports must be made available at the front panel. This is accomplished with a "spline" card, which does not contain an Xbar but merely converts the eight SAN channels to/from serial format (See Figure 6). An inefficient arrangement of two Xbar cards could be used in place of one Xbar and one spline. A single Xbar card can be used if the eight backplane ports are not required. Some of this testing involved a single Xbar card and the rest involved two Xbar cards because the proton irradiation was sufficiently penetrating to hit both Xbar ICs, when two cards were stacked simultaneously in the system test chassis. This provided data on many more paths (described below in the DUT System description).

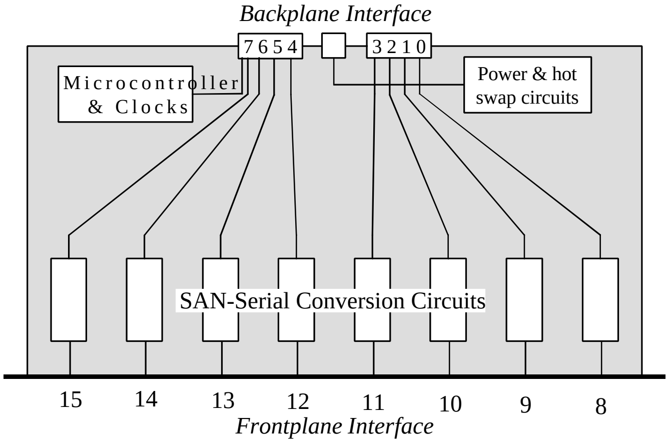

Figure 6. Block diagram of the Myrinet spline card.

C. Network Interface Card (NIC)

The NIC provides functionality for a device (PCI-bus computer) to communicate via the Myrinet-2000 standard [4]. It is a PCI-64 form-factor card that can operate from 32-bit PCI bus as well. It operates at either 5 Volts for PCI-32 operation or 3.3 Volts for either PCI-32 or PCI-64 operation (PCI-64 is a 64-bit bus that can also accommodate 32-bit bus hardware).

The card operates at either 33 MHz or 66 MHz PCI bus speed with ICs that provide bus interface (including PCI Direct Memory Access (PCIDMA)), protocol processing, and serialization and deserialization (SerDeSer) functions. A block diagram is shown in Figure 7.

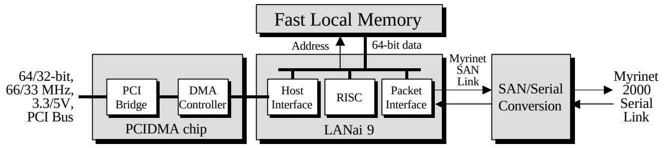

Figure 7. Block diagram of the Myrinet Network Interface Card.

This block diagram shows the main blocks chosen to be exposed to the proton beam. The primary focus was on the protocol processing IC, which is called the "LANai 9 processor". The other two Myricom ASICs exposed are the PCIDMA and the SerDeSer (SAN/Serial Conversion) blocks, shown to either side of the LANai 9. The Fast Local Memory is an SRAM and a serial Transceiver is what drives the Serial Link, shown at the right of Figure 7.

Operating in typical PC situations (33 MHz PCI-32 bus), a data rate of 132 MB/s is typically seen. Operation at maximum data rate (66 MHz PCI-64 bus) is still limited by either the host computer's capabilities or the PCI bus limits. 523 MBPS is the theoretical limit; the actual rate during testing was near that limit. To achieve full 2 GBPS rate, 64 bit 200 MHz bus and high–speed processors are required. These high-speed processors and bus speeds were not readily available and not used in this testing.

Figure 8 shows the NIC. The additional signals for PCI-64 operation can be seen hanging to the right of the white PCI-32 socket. The black Myrinet cable can be seen at far left.

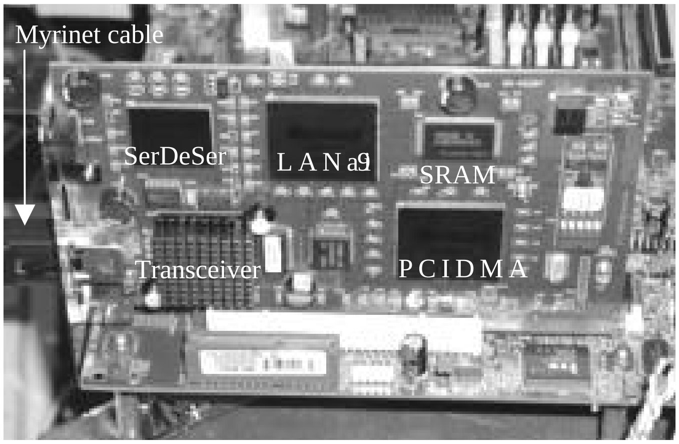

Figure 8. Myrinet NIC installed in PCI-32 motherboard without extender.

D. Network Routing and Packet Format

To accomplish the testing, information must be generated in the test computer and sent to the crossbar via the NIC. This is accomplished via the Myrinet cable that can be seen in Figure 8. The other end of this cable plugs into the front panel of the crossbar card into one of the frontplane interface ports (Shown in Figure 9 coming into port #8).

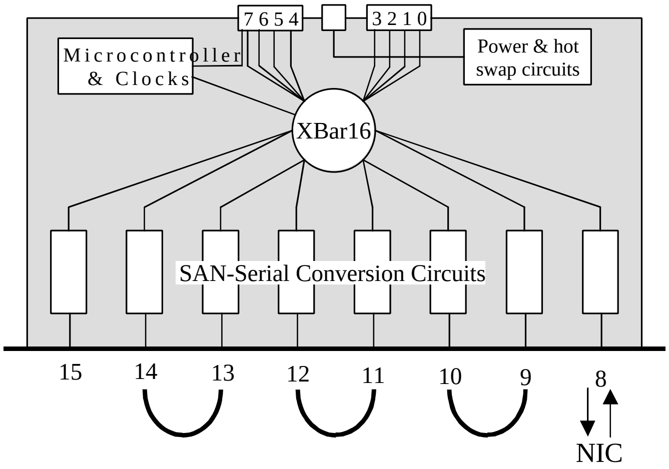

Figure 9. Crossbar block diagram detailing the routing for a seven switch case.

Messages are transported across a Myrinet as one or more packets. The packet format for Myrinet is shown in Figure 10. A packet consists of three components: a header, the payload and the trailer. The header consists of a four-byte packet type (Myrinet is capable of generating many data protocols and the packet type carries information on which protocol is used) and as many bytes as are required to appropriately route the packet. The second portion of the packet is the payload, which is simply the data to be communicated. Myrinet is capable of handling payload portions from zero bytes to at least 4 Mbytes. The final piece of the packet is the trailer, which contains one byte of Cyclic Redundancy Check (CRC) code.

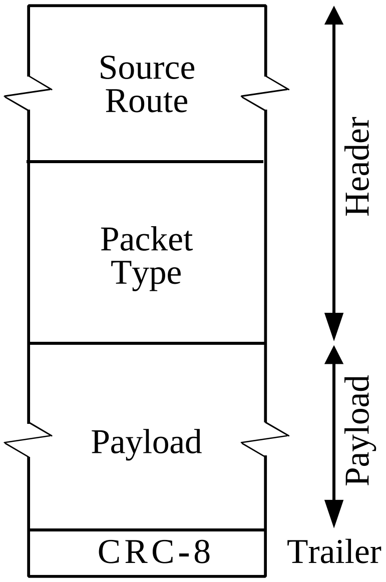

Figure 10. Myrinet packet structure.

Packets are encoded with routing information that allows it to reach the desired destination. Each pass through an Xbar (in a large network many Xbar transits may be required to reach the destination) involves one byte of routing information, which gives a relative (to the incoming port) output port. These routing bytes are removed as they are used, and the CRC is recalculated and appended so that the new packet (1 byte shorter) is correctly formatted. Packets that have inconsistent CRCs are simply dropped. This behavior is hard-wired within each component (within the LANai9 processor on the NIC, and within the Xbar IC). That is unfortunate for SEE testing—events are detectable only by their failure to arrive. No examination of erroneous data is possible. Thus the ability to determine the types of errors is lost (*e.g.,* single bit, multiple bit, burst, etc.)

Figure 9 also shows the routing scheme that was used for one case (the case using seven switches). In this figure, the NIC is shown entering and exiting through port #8 and three Myrinet cables loop the data from ports 9 to 10, 11 to 12, and 13 to 14. The complete loop for this test case would then be: NIC to in #8, out #9, in #10, out #11, in #12, out #13, in #14, out #14, in #13, out #12, in #11, out #10, in #9 out #8 to NIC (Recall that the switches allow for full speed full duplex routing).

E. Test Methodology

In this test, the main objective was to observe what effects would be induced by proton irradiation, with specific concern to latchup sensitivity of any parts. Therefore, to achieve this goal, the main devices of both the crossbar switch and the network interface card (NIC) were placed in the proton beam. All tests were performed at room temperature.

During their exposure, the DUT computer was running software that was generating a fixed data pattern to be passed along the network and watching for the arrival of these packets of information. No direct evidence of upsets was possible. As explained previously, if data within the packet was corrupted, the Myrinet hardware would drop the packet. The missing packet would then be noticed by the DUT software and recorded. This is the main type of error observed. During exposure of the NIC it was also possible to induce errors in the data stream once the NIC accepted a packet as valid. The methodology and software were also in place to observe these types of errors.

The methodology flow was to place the device to be exposed in the beam, start the DUT and test controller systems, turn the proton beam on, and, finally, observe the effects. The proton beam remained on until either a preset amount of fluence was achieved or a functional interrupt or latchup was observed. Initially, the preset fluence was set to a smaller amount due to the uncertainty in the total dose response of any of the devices. As the testing proceeded and the devices appeared to withstand the dose sufficiently well, preset fluences were set to levels such that there was typically a functional interrupt prior to the preset fluence level being reached.

For the errors that were observed, the test software recorded all pertinent information about the errors, including the manner in which they were received (*e.g.*, did a single packet get dropped or were a sequence of packets lost in a very short time span). For functional interrupts, as much information, that could be gleaned from the test system, was recorded. In some instances it was simply that the DUT computer rebooted while in others it was detailed information about which switch in the crossbar devices induced the interrupt. If any latchup currents were observed, the device, the peak current seen at the device, and the functionality after the latch would have been recorded.

F. *Devices Tested*

Two components were tested for this work. The first were crossbar switches manufactured by Myricom, Inc, which provide the interconnectivity in the Myrinet model (as a hub would in star-configuration network model). Secondly, a Peripheral Component Interconnect (PCI)-bus network interface card (NIC) manufactured by Myricom was tested. On the NIC, five devices were chosen (based on being the primary functional blocks for network operation) for exposure and the system evaluated for its response. The listing of all devices used in this testing is given in Table I below.

TABLE I. DEVICE UNDER TEST (DUT) TABLE

Device	Vendor	Location	Model Number
Crossbar Switch 1	Myricom	Switch Board 1	M3-SW16-8S
Crossbar Switch 2	Myricom	Switch Board 2	M3-SW16-8S
NIC	Myricom	NIC	M3S-PCI64B-2
LANai9	Myricom	NIC	9.1
SerDeSer	Myricom	NIC	1.1
PCIDMA	Myricom	NIC	1.3
Transceiver	Vitesse	NIC	VCS7146RH
SRAM	Samsung	NIC	K7N803601M

IV. TEST SYSTEM

A. *Test Hardware*

The test system consists of two subsystems: the Test Controller and the DUT. These subsystems are described below with Figure 11 illustrating the overall test configuration.

1) *Test Controller Subsystem*

The Test Controller hardware is based on the PCI Extensions for Instrumentation (PXI) specification. The PXI subsystem, shown in Figure 12, resides outside the direct irradiation area and is connected to the DUT at the irradiation point by cabling, approximately 40 feet long. It consists of the PXI components, the PXI Computer/DUT System cabling and the user interface.

The PXI components include a PXI chassis, containing an embedded controller (running Win98, Labview™ (LV) environment and a custom LV application), a signal switch matrix, and two digital multimeters (DMMs) in the voltage measurement mode. The switch matrix provides two functions – the multiplexing of analog signals to one of the DMMs, and contact closures (pulling signal levels to ground). One DMM measures all analog values except the value read most frequently or most important minimize be delays by switch settling time. The other DMM is dedicated to monitoring this value.

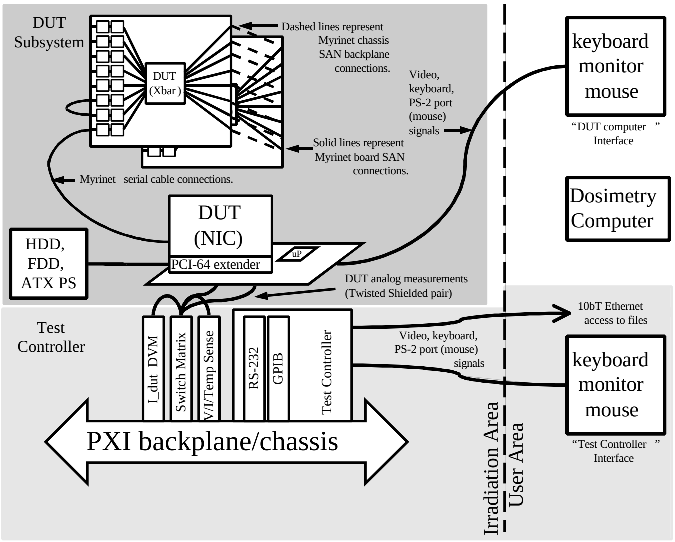
DUT Subsystem
DUT (Xbar)
Dashed lines represent Myrinet chassis SAN backplane connections.
Video, keyboard, PS-2 port (mouse) signals
Solid lines represent Myrinet board SAN connections.
Myrinet serial cable connections.
keyboard monitor mouse
"DUT computer " Interface
Dosimetry Computer
DUT (NIC)
PCI-64 extender
uP
HDD, FDD, ATX PS
DUT analog measurements (Twisted Shielded pair)
Test Controller
I_dut DVM
Switch Matrix
V/I/Temp Sense
RS-232
GPIB
Test Controller
10bT Ethernet access to files
Video, keyboard, PS-2 port (mouse) signals
keyboard monitor mouse
"Test Controller " Interface
PXI backplane/chassis
Irradiation Area
User Area

Figure 11. Block diagram of the test system.

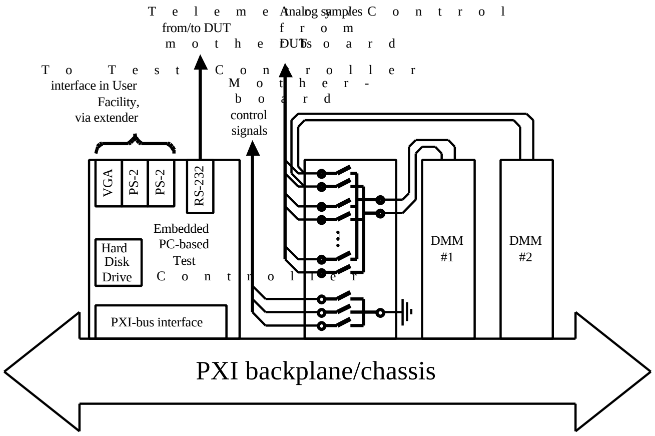

Figure 12. Block diagram of the PXI subsystem.

The PXI Computer's user interface, network connectivity (for data file access) and AC power feed are also components of the PXI Computer. An extended keyboard, monitor, and mouse user interface (via a CAT5 cable extender from Cybex, Inc.) provides user control of the PXI computer from the user facility (which in this instance is located in the hallway outside the regular, restricted access, user area).

Most of the Controller-to-DUT subsystem cabling leaves the PXI subsystem from the switch matrix. Exceptions are the AC power cable to power the DUT System and a serial (RS-232) cable for telemetry/command of the DUT System computer (telemetry originates within the DUT System computer, commands originate within the PXI Controller).

2) *DUT Subsystem*

The DUT Subsystem consists of the computer housing the NIC and the Xbar it is connected to. It includes components mounted directly to the motherboard, components located nearby (*e.g.*, disk drives) and connected via cables, and a Cybex extended keyboard, monitor, and mouse user interface.

The DUT system computer motherboard resides in the test chamber, positioned just below the particle beam when the NIC is exposed. The NIC plugs into an extension socket that raises it up by approximately two inches. The dual 1 GHz Pentium III processors on the motherboard are Flip-Chip Pin Grid Array (FC-PGA) form-factor so they lay very low and well out of the particle beam, as do the low profile (< 1 inch) RAM modules. This does not completely eliminate these devices from interfering with testing, as there is a stray neutron environment in the irradiation area.

Located nearby (approximately 6 feet) are a modified standard PC ATX power supply (PS), a floppy and/or hard disk drive, and a Cybex user interface extension identical to the one used to extend the PXI computer. The motherboard is modified to allow power cycling and reset via the PXI switch matrix. The ATX power supply is modified to allow manually perform power shutoffs. The PCI-64 extension board, which the NIC plugs into, is modified to sample DUT current via the PXI switch matrix and DMMs.

The NIC is connected via a Myrinet cable to the rest of the DUT system, the Xbar switches. These are housed in a chassis containing its own AC power supply. The motherboard is modified to allow connection to two controlling signals, both momentary contact closures. The motherboard front panel power on/off (MotherPonoff) input signal is controlled by the PXI switch matrix, as is the motherboard front panel soft reset (MotherSR) input signal. ATX PS on/off state is normally controlled by a constant signal from the motherboard (The ATX SP supplies a standby +5 Volts, to power such motherboard functions). This signal (PS_ON#) is, approximately, a latched toggle of the front panel signal, MotherPonoff. This motherboard PS_ON# signal is disconnected from the ATX power supply's PS_ON# input so that it can be controlled directly from the PXI. This additional control is necessary because the computer can hang to the extent of not responding to the normal on/off commands. The ATX PS AC power is extended back to the user facility.

The DUT Computer runs the Windows-NT™ operating system and a software application that access Myrinet NIC drivers. Commands from the PXI computer are received via an ethernet cable and responses are transmitted back via the same link.

Currents and voltages, from as many as three devices (one NIC and two Xbars), were monitored. System cabling was designed to allow four current/voltage samples in one subD 15-pin connector cable. A cable assembly was added to trifurcate three signals to separate locations.

DUT system signals that are controlled by the PXI subsystem, as described above, or by the user from the user facility are shown in Table II. DUT computer signals that are monitored by the PXI or directly by the users in the user facility are shown in Table III.

B. *Test Software*

The DUT software was written in Microsoft C++ Professional version 6.0. It was designed to run in Windows 2000 Professional service pack 2. The driver for the Myrinet network adapter was GM 1.1. This driver was downloaded from the Myricom website (http://www.myri.com/).

The Network Interface Card (NIC) takes data packets from the driver and sends/receives the packets through the cables and network switches. The receive function of this card rejects data packets when errors are detected. The method used for detecting errors is a CRC check byte at the end of each packet.

The DUT Software sends packets with an incrementing packet # and data which is a function of the packet #. If the packet number/16 is odd, then the data is a stream of bytes with the value hex 55; otherwise, it is a stream of bytes with the value hex AA. After each packet is sent, the program waits until either a packet is received or approximately 10 microseconds, whichever comes first. There are two physical setups supported. The first setup uses one NIC for both sending and receiving. The second setup uses two NICs, one for sending and one for receiving.

The DUT software utilizes two methods of data transfer. The first method of data transfer that would be used in normal operations is gm_send_with_callback(). The second is the undocumented gm_raw_send_with_callback().

TABLE II. DUT SIGNALS CONTROLLED BY THE PXI SYSTEM

Name	Destination	Description
PS_ON#	ATX Power supply	Hold low (0 V) for PS on; Open = High = Off
MotherPonoff	Motherboard power switch connector	Pulse low (0 V) to toggle power on and off
MotherSR	Motherboard reset switch connector	Pulse low (0 V) to initiate reset
Command	DUT system computer	CAT-5 cable, ethernet, 10/100 mbps rate. Same cable that carries Telemetry data.
Keyboard/ mouse	DUT system computer	PS-2 keyboard ports

TABLE III. DUT SIGNALS MONITORED BY THE PXI SYSTEM

Name	Source	Description
V_NIC, I_NIC	NIC extender card	Voltage and current samples of the NIC primary supply. Twisted shielded pair (TSP).
V_Xbar, I_Xbar1	First Xbar card	Voltage and current samples of the first/only Xbar switch card supply. TSP.
V_Xbar, I_Xbar2	Second Xbar card	Voltage and current samples of the second Xbar switch card supply, if installed. TSP.
Telemetry	DUT system computer	CAT-5 cable, ethernet, 10/100 mbps rate. Same cable that carries Command data.
GUI output	DUT system computer VGA card.	Video carrying output to the user facility.

The normal method of transfer uses handshaking that ensures that the data is received without any detected errors before the send is completed. If any errors are detected the data is resent until the data is received without detected errors or a timeout of about a minute is reached. Before running the DUT Software in this mode, the GM utility program gm_mapper_service must be executed. This cannot be executed while the DUT software is running. In this mode the speed of data transfer can be set using gm_set_speed. This method was not used in this testing.

The undocumented method of transfer (Raw Mode) uses no handshaking. If the data is received with detected errors it is rejected by the NIC and is never seen by the user software. The user software detects when a packet is skipped or any errors that are not detected by the NIC are received. When packets are skipped, the packet number of the packet received after the skip and the number of skipped packets are recorded. When errors are found within a packet, the packet number, the locations within the packet and the actual values of the bytes in error are recorded.

The DUT communications are controlled through buttons and checkboxes on the DUT console of the custom DUT software. All of these can be manipulated through the keyboard and mouse of the DUT computer. Some of these can be controlled through the TCP/IP connection by the test controller system. These can be controlled from the test controller by sending a one-byte command to the DUT.

The DUT subsystem is connected through a Transmission Control Protocol/Internet Protocol (TCP/IP) socket to the test controller system where the test controller system acts as the host and the DUT acts as a client. The IP address and port used for the test controller connection are hard-coded. When not connected, the DUT tries once every 3 seconds to make a connection. The DUT sends telemetry information to the test controller system and records the same telemetry to a file on the DUT hard disk drive. The telemetry consists of a stream of 4-byte long integers sent LSB first with the following format:

```
// The last byte of 4 is a data code. The
// table below shows the definitions for
// each code:
// FF timestamp and beam info
//    xx xx yy FF
//    xx xx relative timestamp
//    yy 01 for beam on, 00 for beam off
// FE Error in data packet
//    xx xx yy FE
//    xx xx location within packet
//    yy data read
// FD Skipped Packet(s)
//    xx xx xx FD
//    xx xx xx Number of skipped packets
// FC Skipped Packet(s) (Large/-)
//    00 00 00 FC
//    xx xx xx xx
//    xx xx xx xx Number of skipped packets
// FB Buffer overflow
//    00 00 00 FB
// FA Header
//    AA AA AA FA
//    aa aa aa aa
//    tt tt tt tt
//    rr rr rr rr
//    rr rr rr rr
//    rr rr rr rr
//    rr rr rr rr
//    ff ff ff ff
//    ff ff ff ff
//    ff ff ff ff
//    ff ff ff ff
//    ff ff ff ff
//    ff ff ff ff
//    ff ff ff ff
//    ff ff ff ff
//    pp pp pp pp
//    aa aa aa aa Ascii Version
//    tt tt tt tt Time Stamp
//    rr rr rr rr... Route info
//    ff ff ff ff... Filename
//    pp pp pp pp Packet Size
// F9 Reconnect
//    00 00 00 F9
// F8 Packet Number
//    xx xx xx xx
//    xx xx xx xx Packet Number
```

These telemetry streams were stored by the Test Controller for post processing of quantities and types of errors observed.

V. RESULTS

A. Network Interface Card

1) Single Event Latchup

For the test being performed on this system, the NIC current was monitored for the entire board. Therefore, determination of a latchup event in an individual component would have to generate sufficient current to be observable above nominal NIC current. For all five components exposed to the proton beam on the NIC, no high NIC currents or destructive events were observed. There were events on all five devices that led to functional interrupts (to be discussed next). These events could possibly be produced via a high current condition in the respective part, as a power cycle of the DUT computer was required to reset after each interrupt. However, since no events were seen to be destructive, it is impossible to say whether latchup did or did not play any role in these events.

2) Single Event Functional Interrupts (SEFI)

When any of the five devices were exposed to the proton beam, the DUT computer system would experience a SEFI event at some point. This could be seen as the DUT computer either freezing or initiating a self-reboot. In all instances observed for all five devices, a power cycle of the DUT computer that housed the NIC was required to regain functionality. The SEFI cross-sections measured for the five devices are shown in the last column of Table IV.

TABLE IV. NIC RESULTS SUMMARY TABLE

Part	Accumulated Dose (krad(Si))	Upset Cross Section (cm^2)	SEFI Cross Section (cm^2)
Lanai9	59.2	6.81×10^{-12}	1.14×10^{-11}
SerDeSer	53.1	1.52×10^{-11}	5.07×10^{-12}
PCIDMA	45.7	2.94×10^{-12}	1.18×10^{-11}
Vitesse	50.7	4.03×10^{-10}	7.95×10^{-12}
Samsung	9.1	8.84×10^{-11}	7.37×10^{-11}

3) Single Event Upsets

As discussed previously, missed packets are the normal mechanism for errors to display themselves for this test setup. For all but the Samsung device, this is the upset mechanism that was observed. For the Samsung SRAM part, the second upset possibility arose. These are errors that are received that are not detected by the NIC. In other words, data was correctly received by the NIC, processed by the LANai 9 processor and stored into the Samsung SRAM. While in this stored location, it is altered and that difference is detected. The upset cross sections measured for the five devices are shown in the third column of Table IV.

All of the above cross sections are per-device except for the Samsung SRAM. Two of these SRAM parts were exposed during the testing (one on each side of the board). It is not clear from the Myricom documentation how much of both of these parts are used and if their usage is equal. Therefore, those cross sections are left as total cross sections for both parts (not per bit or per device).

4) Total Dose

While total dose testing was not explicitly included in this testing, proton dose is accumulated over the course of the test. No parametric measurements of the devices are feasible with this test setup. However, no functional loss or functional performance degradation was observed throughout this entire test. Therefore, it can be stated that the devices are total dose functionally survivable to at least the maximum proton dose during this test. These dose levels for the five devices on the NIC are shown in the second column of Table IV.

B. Crossbar switch Device

1) Single Event Latchup

For the test being performed on this system, the crossbar switch current was monitored. For both crossbar switch devices exposed to the proton beam, no high currents or destructive events were observed. There were events that led to functional interrupts. These events did require a power cycle of the crossbar power supply to reset after the interrupt. However, since no events were destructive and no high currents were observed for the switches, it is possible to say that latchup did not likely play a role in these events.

2) Single Event Functional Interrupts (SEFI)

Figure 13 and Figure 14 show the per-switch SEFI cross-section for the crossbar switch devices tested. In Figure 13, it is assumed that all switches have the same sensitivity whether they are on the in Xbar frontplane (FP) or backplane (BP), or on either Xbar #1 or #2. This cross section is plotted as a function of the number of switches active during that test (there are different percentages of the switch locations for these four cases). The squares and error bars are the average overall cross section, assuming all switches are the same, and the one-sigma standard deviation. The triangles are the cross sections within each of the four cases. The four data points almost lie within one sigma.

Figure 14 looks at the same data set but with the thought that the SEFI rate could be different between frontplane and backplane switches. The four cases shown here are the two frontplanes of the two Xbars, the backplane switches, independent of which Xbar houses them, and the overall cross section (the same as the squares of Figure 13). While all of the data points lie within the one-sigma error bars of the overall cross section, there does appear to be a difference between the frontplane and backplane switches.

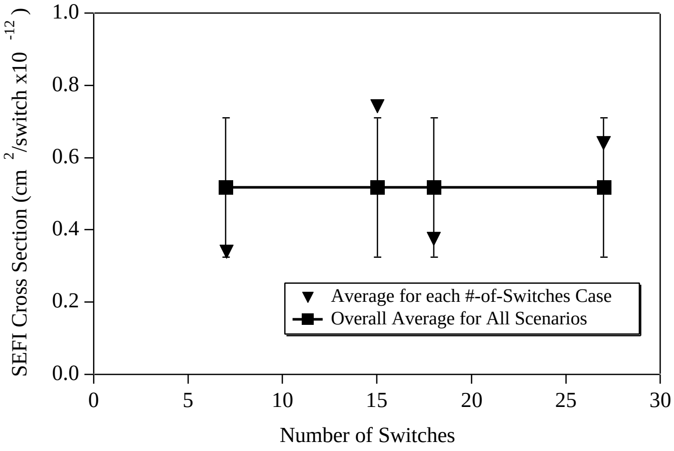

Figure 13. SEFI Cross section as a function of the number of switches.

3) Single Event Upsets – Non-SEFI

Single Event Upsets (SEU), for the crossbar switch devices, are only evident as dropped packets. Data was collected to include the number of dropped packets, whether they arrived as a single dropped packet or in a rapid sequence of dropped packets. This data was collected for four different switch quantities, that also had varying quantities in the frontplane and backplane.

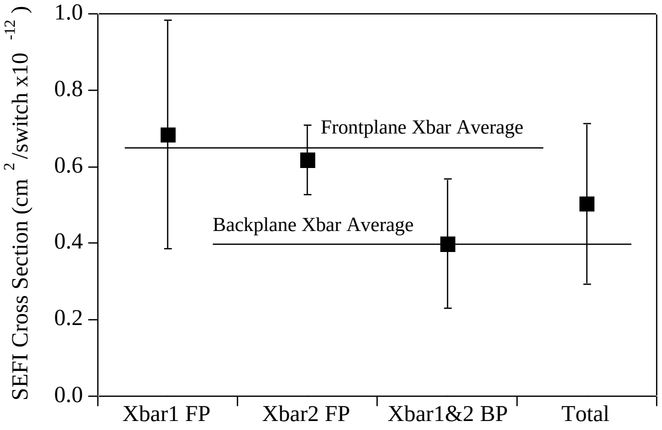

Figure 14. SEFI cross-section as a function of the location of the affected switch.

Figure 15 shows the per-switch cross section as a function of the number of switches in the test configuration. It shows data for both single packet loss and for multiple packet loss. It is evident that the multiple packet loss appears to be within approximately one sigma of an average value for the multiple events. The same cannot be said for single packet loss, as the two higher switch count cases (those using backplane switches) have substantially higher cross sections.

The same data as shown in Figure 15, can be viewed in another way by looking at the total cross section (both single and multiple packet losses and not per-switch). This data is shown in Figure 16.

The two cases, with the lowest number of total switches (the cases with only frontplane switches), have a cross section that is nearly an order of magnitude lower than the two cases with higher number of switches. Both of these higher switch count cases have all sixteen backplane switches incorporated in the path for the data packets. The highest switch count case does have a slightly higher cross section than the next lower case as it does contain nine additional frontplane switches (again, these cross sections are not per-switch).

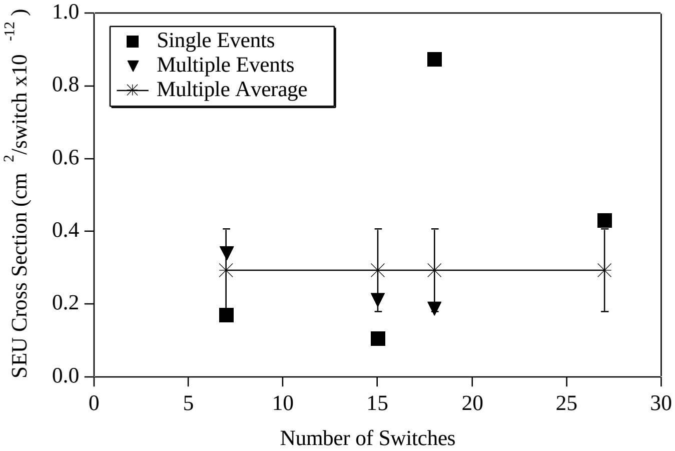

Figure 15. SEU cross section as a function of the number of switches for both single dropped packets as well as multiple dropped packets.

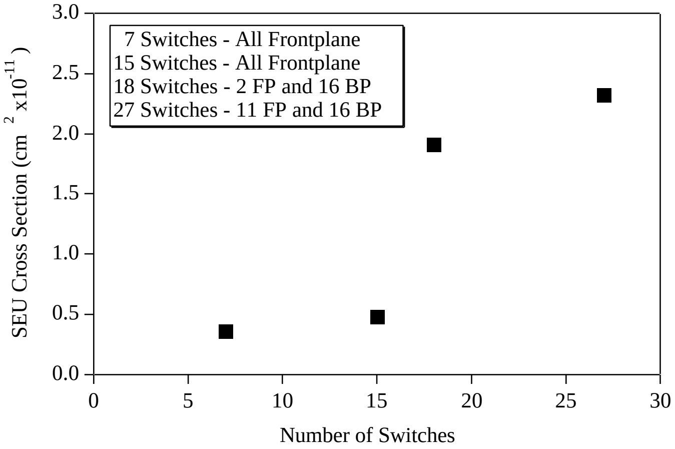

Figure 16. Total SEU cross section as a function of the number of switches with details of switch locations.

This SEU data appears to imply that having the backplane switches in the data path will substantially increase the data packet loss as compared to running without backplane switches. It is possible that there are physical differences between backplane and frontplane switches in dealing with packets that is not immediately evident from the Myricom documentation.

4) Total Dose

As with the NIC, total dose testing was not explicitly included in the testing. However, proton dose is again accumulated over the course of the test on the crossbar switch devices (Xbar). No parametric measurements of the devices are feasible with this test setup. However, no functional loss or functional performance degradation was observed throughout this entire test. Therefore, it can be stated that the devices are total dose functionally survivable to at least the maximum proton dose during this test.

For this test setup, however, some amount of uncertainty exists for the dose levels of Xbar #2. This is because the proton beam passes through Xbar #1 and then the board for Xbar #1 before impinging on Xbar #2. While there is an unknown amount of material between the two switches, it does not appear to be substantial and it is assumed that the incremental doses on Xbar #1 are the same foe Xbar #2 when it is in place (Xbar #2 is only used when more than seven switches are used in the routing). These dose levels for the two crossbar switch devices tested are 400 krads(Si) and 285 krads(Si), for Xbar #1 and #2 respectively.

VI. Summary

The Myricom Myrinet network system was evaluated for proton single event effects response. No indication of latchup was observed for the cross bar switches, and most likely for the five devices tested on the NIC. Functional interrupts and data loss upsets were observed for both crossbar switches and the NIC devices. These cross sections were determined.

Total dose numbers seen during the testing indicate a reasonable tolerance to total dose. Single event upset and functional interrupt rates, however, were substantial and all interrupts required a power cycle to regain functionality. Further testing of this technology is needed before insertion into a space mission would be possible. For future testing, it would be desirable to have access to the bit level information to more accurately assess the single event upset sensitivity and the possibilities of any mitigation techniques.

Acknowledgment

The authors would like to take this opportunity to thank the NASA Remote Exploration and Experimentation Project and the NASA Electronic Parts and Packaging Program for their financial support of this work. The authors would like to personally thank Raphael Some for his technical discussions and insight in the performance of this testing.

References

[1] “American National Standard for Myrinet-on-VME Protocol Specification”, VMEbus International Trade Association, ANSI/VITA 26-1998, November 2, 1998.

[2] “High Performance Cluster Computing: Architectures and Systems,” Rajkumar Buyya (editor), Prentice Hall PTR, NJ, USA, 1999.

[3] “Myrinet M3-SW16-8S Specifications”, Myricom Web site: <http://www.myri.com/myrinet/m3switch/m3-sw16-8s/index.html>

[4] “Myrinet NIC Specifications”, Myricom Web site: <http://www.myri.com/myrinet/PCI64/m3s-pci64b.html>