Subatomic Particle Sensitivity of Gigabyte Networks

Anubhuti Khare, Manish Saxena, Rajesh Kourav

***Abstract*—We have presenting subatomic particle radiation testing of the 57710 network controller. It shows that there is a SEFI mode that could cause the internal network to become unavailable every two to 1136 years with the TCP/IP protocol and every five to 2276 years using the UDP protocol based on location and solar activity. To use intersystem networks, devices will need network controllers and switches. These devices are likely to be affected by single-event effects, which could affect data communication. In this paper, we will present radiation data and performance analysis for using a Broadcom network controller in a neutron environment.**

***Index Terms*— Networking, WLAN, Security**

I. INTRODUCTION

Innovations in computational systems have focused on intra system networking. For high-end, commercially available computing systems, the system consists of several processors, coprocessors, and peripherals connected though an internal network. One example of this type of computing system is the Los Alamos National Laboratory (LANL) Road- runner [1], which is a supercomputing cluster where each node has two processors and four coprocessors that are connected through an internal network.

There are several advantages of using intrasystem networks to transmit data between devices rather than using dedicated traces. First, the network switch provides a flexible framework for data communication between devices, so computation can be routed around unavailable devices. Second, as data com- munication speeds increase, reliably passing data through par- allel buses becomes difficult. Third, as system sizes increase, intrasystem networks are more scalable than dedicated routing.

Finally, using industry-standard devices and protocols allows designers to leverage the commercial networking industry and decades of research in information theory on fast, robust, and reliable communication. Because of these reasons, intrasystem networks are attractive for both on-payload and interpayload. communication for both satellites and airplane applications .

Manuscript received November, 2011

Dr. Anubhuti Khare, Department of Electronics and Communication, University Institute of Technology Rajeev Gandhi Technical University, Bhopal, India, Mobile: +919826526247, (Email:- anubhutikhare@gmail.com).

Manish Saxena, Head of Electronics and Communication Department, , Bansal Institute Of Science and Technology, Bhopal, India Mobile: +919826526247, (, Email:- manish.saxena2008@gmail.com).

Rajesh Kourav, Student, M.Tech (Digital Communication, Bansal Institute Of Science and Technology, Bhopal, India, Mobile:-+919893890199, (Email-rkourav2000@gmail.Com).

Al- ready, many designers are using either gigabit or 10-Gb Ethernet for their payloads. Because of this trend, we are seeing an in- creased use of network devices in radiation environments.

Radiation-induced failures, such as single-event functional interrupts (SEFIs) that cause network availability issues or single-event upsets (SEUs) that cause data loss or corruption, must be mitigated so the performance and reliability of data communication do not suffer. While there are a number of pa- pers and application notes published by Cisco [3]–[5] regarding radiation effects in their products, these papers focus on the processors and memories in their switches. A 2003 paper [6] on the use of commercial network protocols for space application states that was no radiation test data for Ethernet. This paper included radiation tests of a cross-bar switch, which showed that the network switch was susceptible to both SEFIs and bit errors. The paper also stated that the errors increased with the number of connections. These conclusions informed our research of a network controller and that the most common failure modes would be a loss of availability from SEFIs and data reliability problems from SEUs.

In this paper, we will present data on the Broadcom 57710 converged network interface card, which is a 10-Gb Eth- ernet controller chip that has been used in many servers. We tested this controller at the Los Alamos Neutron Science Center (LANSCE) neutron accelerator. We found from this testing that there was a SEFI mode that caused network avail- ability issues and radiation-induced performance degradation.

In this paper, we will present our test setup in Section II, as well as a background information on the network protocols used in the test. During this test, we found a SEFI mode that causes the network to crash, which is discussed in Section III, and performance and data loss issues, which are discussed in Section IV.

II. TEST SETUP AND BACKGROUND

We tested the Broadcom 57710 NetExtreme II 10-Gb dual-port Ethernet controller chip at LANSCE. This particular net- work controller is available in the BCM957710A1022G_AC network interface card (NIC) that is available from Dell. As shown in Fig. 1, the 57710 controller includes on-chip memory, two four-lane XAUI SerDes, an eight-lane PCI SerDes, and native support for three different protocols. In another version of this figure from the BCM57710/BCM57711 Programmer's Guide [7],

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

the

three Protocol engines and the processor are shown as very large instruction word (VLIW) RISC processors. The protocol engine processors allow a microprocessor to off-load the computational work of sending and receiving data to the controller's computational elements.

The goals of the test were to determine if neutron radiation caused the controller to crash or corrupt data. SEFIs

in a network controller are detrimental because the need to reset the device would affect the availability of the network. Understanding how SEUs could cause data corruption or data loss was also necessary because it would determine how much additional software- and protocol-based reliability was necessary for the system. Therefore, the test focused on methods for measuring availability, data loss, and performance loss.

The tests were conduced at LANSCE in July and November 2009. In the July test, we tested the controller for 8 h for a total adjusted fluence of 9.69×10^{10} n/cm In the

November test, we tested the controller for two days for a total adjusted fluence of 3.88×10^{10} n/cm.In both of these tests,the parts were tested at nominal voltages and temperatures. Both the heat sink and the fan were left on the NIC under test as the neutron beam is minimally affected by passing through either materials. In both tests, we used similar hardware and software setups, whose descriptions follow.

The hardware setup involved two machines with BCM957710A1022G_AC NICs connected through a 25-ft Category 6 (CAT-6) cable to create a private network. It should be noted that this particular NIC can only be used with other 10-Gb Ethernet NICs. The NIC that was under test was connected to the computer through a PCIe extender so the July test, the Magma Express Box Pro PCIe extender, which is one lane wide, was used. In the November test, a One Stop System PCIe extender, which is four lanes wide, was used. In the baseline testing, both PCIe extenders provided the same performance. A diagram of the test setup can be seen in Fig. 2. This diagram shows how the NIC under test is separated from its machine, and that only the two cables (one for PCIe and one for Ethernet) go over the concrete partition to the user facility. A picture of the NIC under test and the One Stop PCIe extender at LANSCE is shown in Fig. 3. Please note that the NIC test.

The hardware setup involved two machines with BCM957710A1022G_AC NICs connected through a 25-ft Category 6 (CAT-6) cable to create a private network. It should be noted that this particular NIC can only be used with other 10-Gb Ethernet NICs. The NIC that was under test was connected to the computer through a PCIe extender so the computer could be removed from the beam target area. In the July test, the Magma Express Box Pro PCIe extender, which is one lane wide, was used. In the November test, a One Stop System PCIe extender, which is four lanes wide, was used. In the baseline testing, both PCIe extenders provided the same performance. A diagram of the test setup can be seen in Fig. 2. This diagram shows how the NIC under test is separated from its machine, and that only the two cables (one for PCIe and one for Ethernet) go over the concrete partition to the user facility. A picture of the NIC under test and the One Stop PCIe extender(computers, devices, etc.) before sending data. TCP/IP ensures reliable, ordered, congestion-free data communication, which reduces the need to mitigate data transactions in the applica- tions. There are some detriments to enforcing reliability and ordering at this level, though. Automatic retransmission of lost data packets could cause data packets to become unordered, which have to be reordered in a buffer either on the NIC or in the microprocessor's memory. If radiation effects increase the noise

in the communication channels, it might be necessary to increase the size of receive buffers to ensure that there is enough space for TCP/IP to buffer packets. On top of it, if radiation sufficiently lowers the capability of reliably sending data, the protocol will take down the channel, causing availability issues fixture is positioned behind another test fixture in this setup, which is a standard practice in neutron radiation testing.

To measure data and performance loss during the test, the software test fixture used the Iperf network bandwidth test software [8]. This software supports network performance testing using different protocols, packet sizes, and bandwidths. For both tests, we configured Iperf to send data bidirectionally so that each NIC was both a client and a server. Iperf pro- vides time-stamped logs with information about the bandwidth performance, data loss, and other parameters. It should be noted that there are other network performance measurement programs available, but only Iperf was able to handle the demands of logging performance metrics and information regarding bidirectional data movement at

10-Gb/s speeds. A block diagram showing the setup of the software test fixture is shown in Fig. 4.

We tested two protocols that are supported by Iperf software. The Transmission Control Protocol/Internet Protocol (TCP/IP) was tested in July, and the User Datagram Protocol (UDP) was tested in November. We focused on testing these two protocols because they provide different data sets and insights into the de- vice. In particular, the controller provides on-chip support for TCP/IP, as shown in Fig. 1, and not for UDP. By testing both protocols, we can see if there are changes to the single-event ef- fects when not using the on-chip protocol off-load engine. UDP is also a lighter-weight protocol than TCP/IP, which could have some advantages in payload architectures. In the remaining por- tion of this section, we will provide an overview of the TCP/IP and UDP protocols, including advantages and disadvantages of both protocols. More information on network protocols and net- work concepts can be found in [9].

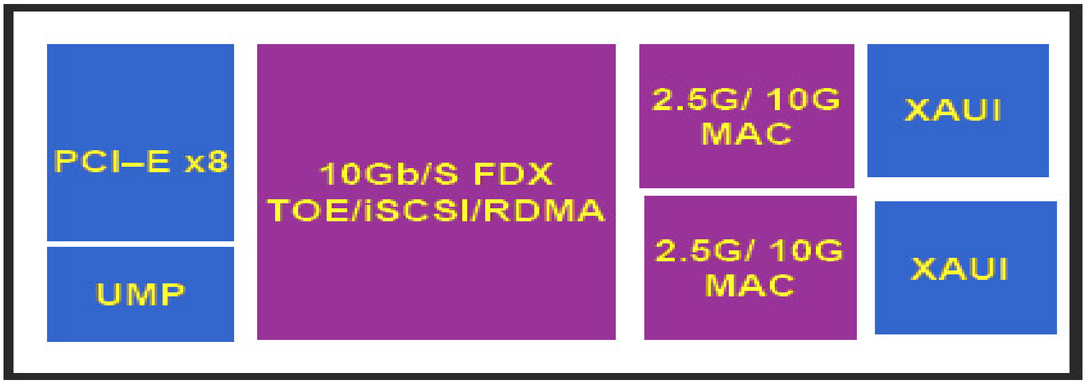

Fig. 1. Block diagram of the Broadcom 57710 controller 2

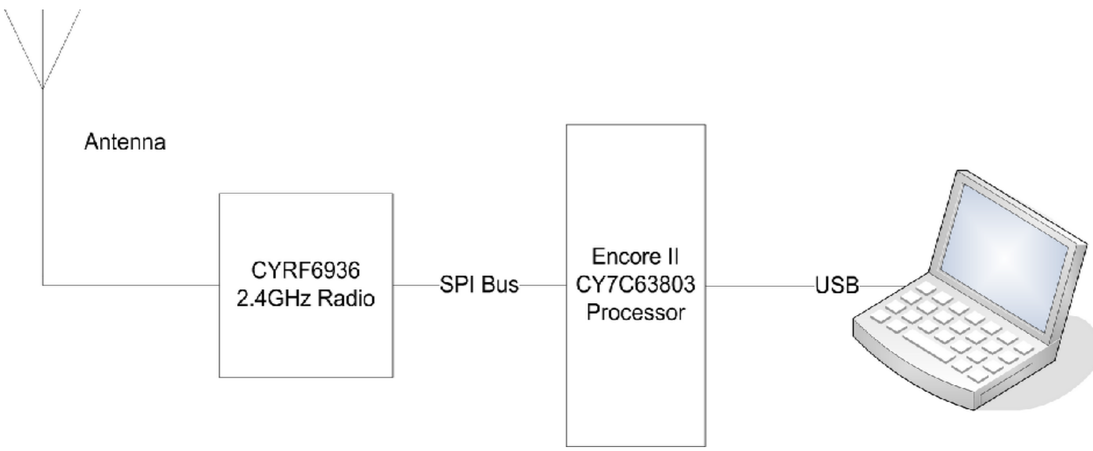

Fig. 2. Block diagram of LANSCE setup.

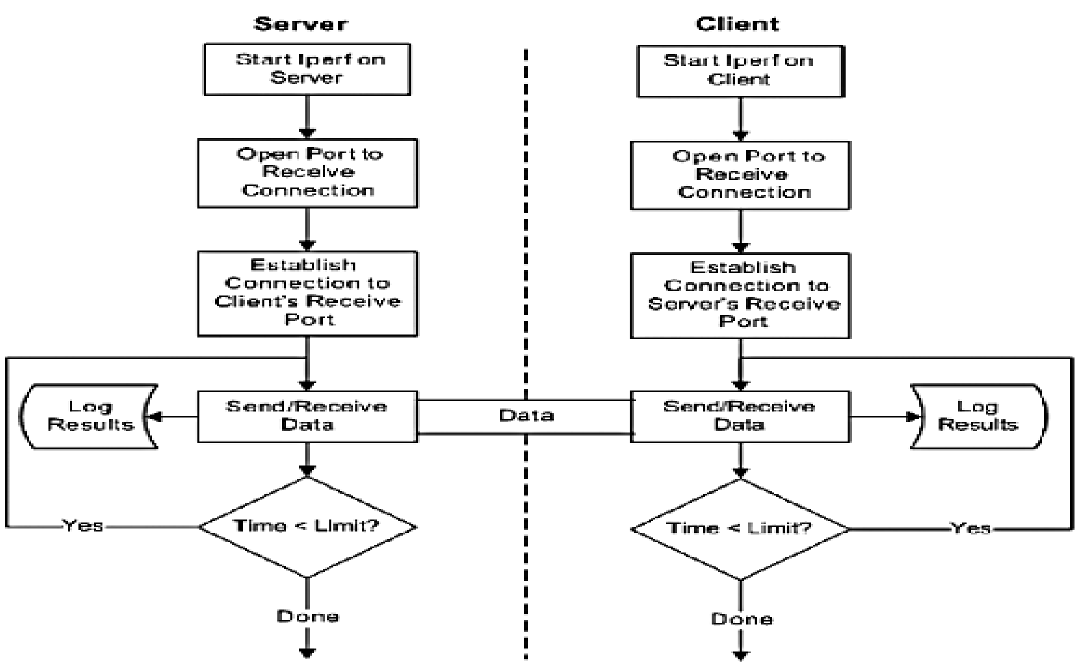

Fig. 4. Procedure for the Iperf software test fixture

Finally, the slow-start algorithm, which mitigates congestive failure by negotiating data communication speeds when the connection is first established, is a disadvantage for communication transmissions that are short, fast, and/or infrequent.For aerospace applications, UDP might be more useful than TCP/IP. In particular, if the communication channels are going to be dedicated and reliable, then the data communication does not need the added overhead of TCP/IP to ensure the communication channel or the reliability. On top of it, if data is sent infrequently or only small data transmissions are made, UDP will be faster than TCP/IP. Finally, the application or architecture design can incorporate methods for reordering packets or requesting packets to be retransmitted. In this way, the payload designers can determine the amount of reliability they want to add to their systeminstead of inheriting the protocol's definitions of reliability. Because of all of these reasons, it is very likely that a number of payloads will adopt the UDP protocol.

III. SINGLE-EVENT FUNCTIONAL INTERRUPTS

While testing the TCP/IP protocol, the controller experienced

18 SEFIs. The effect of the SEFI mode in TCP/IP was that the network controller would crash. In TCP/IP, this meant that both communication channels would come down and could not be restarted. While testing the UDP protocol, the controller experienced a similar SEFI 36 times. In comparison to TCP/IP, UDP had a more varied respond to the SEFI mode. During the UDP tests, we found that the server-to-client communication path failed 19% of the time, the client-to-server communication path failed 33% of the time, and both the paths failed simultaneously

47% of the time. We found in the last case that both communication paths were more likely to completely fail when the client side of the test was being irradiated. For both proto-cols, the controller did not indicate a loss of link, despite being unable to restart the network.

During the tests, we were only able to reset the communication paths and the network for both protocols through rebooting the machine hosting the NIC under test. This behavior might be a side-effect of PCIe. A more graceful method of recovering from this SEFI mode, such as a soft reset of just the controller, is likely possible on a custom board. The application will also need to be either restarted completely or be able to restart the data transmission.

Given the differences in fluence, this means that the con- troller is two times more likely to fail using the TCP/IP protocol than the UDP protocol. For the TCP/IP protocol, the SEFI cross section is 1.86×10^{-9} cm^2/device with a 95% confidence interval of (1.10×10^{-9} cm^2/device, 2.93×10^{-9} cm device). For a UDP protocol, the SEFI cross section is 9.28×10^{-10} cm device with a 95% confidence interval of (6.47×10^{-10} cm device, 1.28×10^{-9} cm^2/device).

Next, we translated the SEFI cross sections into potential error rates. The neutron environment for airplanes is

dependent on many factors, including the location of the airplane and the solar activity. As pointed out in [10], the neutron flux depends on latitude, longitude, and altitude. For the purpose of illustrating SEFI rates for these devices, we have picked a few ranges of neutron flux based on three altitudes, three latitudes, and two solar activities. These values are shown in Fig. 5. This figure shows that there is an exponential increase in neutron flux with altitude. The figure also shows that the higher altitudes are more affected by solar activity and latitude. For example, the flux at 10 000 ft varies by as much 3.3 times, whereas the flux at 60 000 ft varies by 14.0 times.

The SEFI rates for TCP/IP are shown in Fig. 6. The SEFI rates for UDP are shown in Fig. 7. These SEFI rates indicate that, for many subpolar flights, an airplane could be flying this controller continuously for decades without a SEFI and for a few years for polar flights. Therefore, while this SEFI mode could disrupt intrasystem networks, it is not likely to happen often for individual airplanes. Finally, the SEFI error rates for both protocols are reasonable and should not be used to determine which protocol to use.

IV. PERFORMANCE AND DATA LOSS

During both of the tests, the NIC experienced a wide variety of performance-related issues. These issues ranged from intermittent communication problems with the UDP protocol to performance degradation in both protocols. As UDP ensures less reliability than TCP/IP, it was possible to see how often packets did not arrive at all or arrived out of order. We will discuss these results.

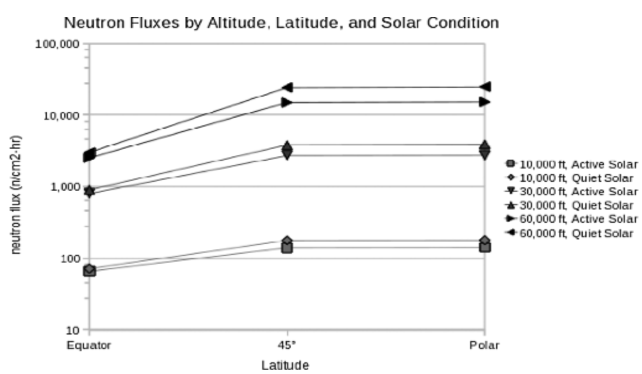

Fig. 5. Ranges of neutron fluxes for different locations.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

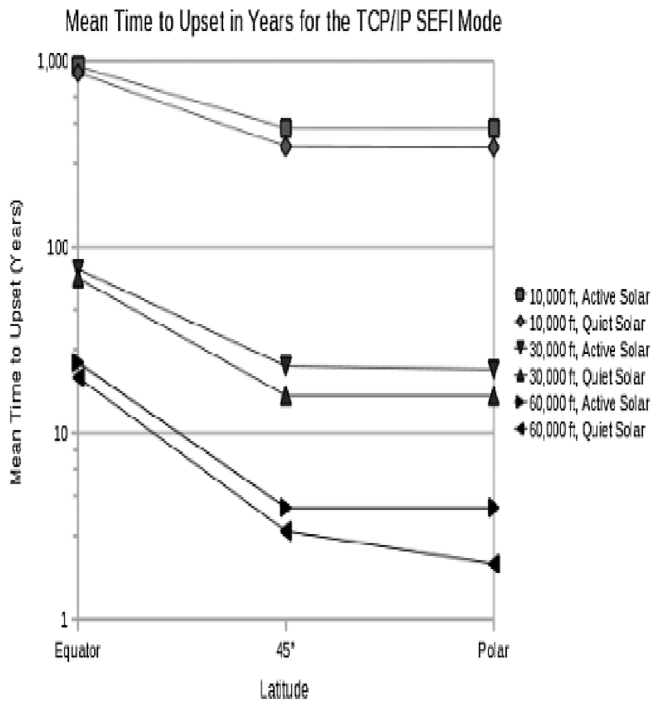

Fig. 6. Mean time to upset in years for the TCP/IP SEFI mode.

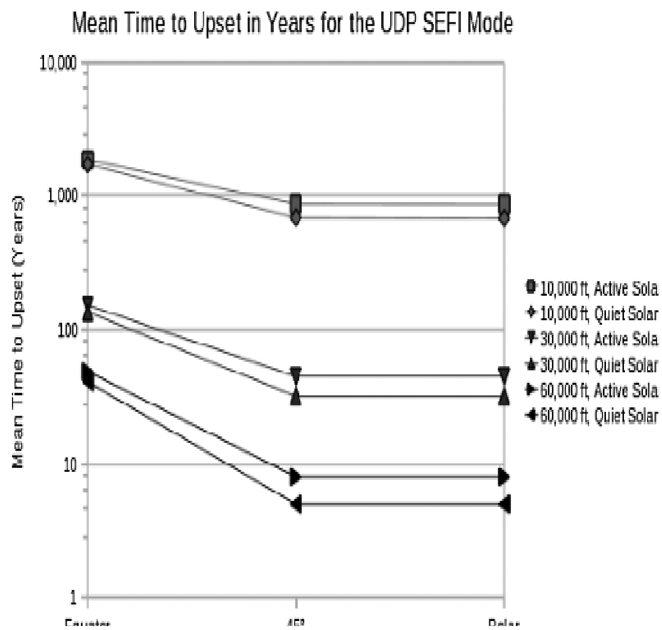

Fig. 7. Mean time to upset in years for the UDP SEFI mode.

Table-I: AVERAGE SERVER-TO-CLIENT UDP STATISTICS

Portion in the beam	Time to Failure (sec)	Total Data Transferred (MB)	Throughput (MB/s)
Server	2,609	40,499	128
Client	2,269	68,327	264
Baseline for 3,600 Sec Tests	N/A	543,744	1,434

Table-II: AVERAGE CLIENT-TO-SERVE R UDP STATISTICS

Portion in the beam	Time to Failure (sec)	Total Data Transferred (MB)	Throughput (MB/s)	Jitter	Percentage Datagrams Lost	Out of Order Arrivals
Client	2,410	8,037	33	4.3	79%	178,695[1]
Server	2,693	11,597	41	3.26	79%	0
Baseline for 3,600 Sec tests	N/A	595,968	1,556	1.4	18%	3

1 While this value is the average statistic for several tests, the data is skewed by one test where, in one 10-s time period, more than 178 000 datagram's were received out of order. All of the other tests did not have any out of order datagram's.

A. INTERMITTENT COMMUNICATION INTERRUPTIONS

In 42% of the UDP tests, either the client-to-server or server-to-client communication channels would not transmit data for up to a minute at a time. From the logs, we were able to determine how many seconds were lost during the test. The percentage of lost communication time was determined by dividing the amount of time the communication was interrupted by the total seconds for the tests. For all tests, the amount of communication time lost from these events is 2.7%. For the

42% of tests that exhibited this behavior, the communication time lost was between 0.87%–32.65%, with an average of

9.07% of lost communication time. In most of these cases, the

temporary communication failures were on the client-to-server transmission side when the client software was in the beam, which indicates possible problems in the controller sending data in a radiation environment. Very rarely did these types of communication problems lead to a SEFI. Therefore, the source of the problem is likely independent of the SEFI problem.

During the communication interruptions, the UDP statistics did not log increases in lost or out-of-order datagrams. There- fore, there is a possibility that datagrams are not lost but being buffered. Further testing with a realistic payload and an application will be necessary to determine whether data is missing or delayed while using UDP.

The TCP/IP protocol did not experience the temporary loss in data transmission that UDP did. This difference is related to the difference in how TCP/IP and UDP send data. TCP/IP es- tablishes a dedicated channel to send packets through and must track the packets to ensure reliable and ordered data communi- cation. Once TCP/IP stopped sending data, the channel never recovered.

B. Performance Analysis

We were also able to determine how the performance was af- fected by radiation. In this section, we will focus mostly on the UDP protocol tests. Because the UDP protocol does not manage the fault tolerance of the data transmission, there is more infor- mation about the reliability that is not hidden by the protocol. As TCP/IP handles retransmission and ordering, this type of in- formation is not available from Iperf. Because TCP/IP ensures that packets arrive in order reliably, there is no concept of either out-of-order or missing packets. Instead the protocol slows down data transmission to ensure that packet transmission is as reliable as possible. During our tests, we set data at 10-Gb/s speeds, but the TCP/IP channel negotiated a channel that was on average 77.75 MB/s for bidi- rectional bandwidth, which is 6.07% of the 10-Gb/s speed. This

77.75-MB/s channel was predominantly dedicated to outgoing data communication, as the incoming data communication was 0.16 MB/s. The average peak communication speed, as measured by IPerf, was 140.97 MB/s total, which included 0.51

MB/s for incoming communication. The total average peak value is 11.01% of the 10-Gb/s speed.

For UDP, more information was available. For the server-to- client bandwidth, Iperf provides information on the total amount of data that was transmitted and the throughput for the commu- nication. These statistics are shown in Table I.

These statistics show that time to failure is not affected by whether the client or the server were in the beam. When compared to the baseline, though, the total data transferred and the throughput are signifi- cantly lower. In baseline testing, the controller was able to com- plete an hour-long test without crashing, and the throughput was an order of magnitude faster than all of the server-to-client communication in the neutron environment.

For client-to-server communication, Iperf provides the same information as the server side, as well as extra information on lost datagrams, variation in arrival time (jitter), and out-of-order datagram arrival. The client-to-server communication statistics for UDP are shown in Table II. With these statistics, we can see that time to failure, total data transferred, throughput, jitter, and percentage of lost datagrams did not vary significantly whether the client or server was in the beam. In terms of the client-to- server communication, the throughput was nearly an order of magnitude slower than the server-to-client communication for all the cause of the throughput issues for both UDP data sets stems from lost datagrams. From Table II, we can see that there is a very high percentage of lost datagrams in the beam and in the baseline. While on average 79% of all datagrams are lost in the neutron radiation environment, the percentage in the base- line was 18%. We believe that the lost datagrams in the base- line test are due to

an undersized UDP buffer in the operating system's settings.

Assuming that caused the 18% lost datagrams in the baseline test,

that only accounts for one quarter of the lost packets in the radiation environment.

The high percentage of data loss could be either the limi-

tation of the Iperf software, or it could be the observable re- sult of a SEU. The limitations of the Iperf test software is that the code for determining lost or out-of-order datagrams is very simple. The software expects the datagrams to arrive in sequen- tial order based on the packet ID. If datagrams do not arrive ordered, the code assumes all of the datagrams between the last received packet ID and the current packet ID are lost or out of order. Therefore, a SEU in the datagram that changes the packet ID from 3371 to 7467 causes 4096 lost datagrams. Similarly, a SEU that switched the packet ID from 7467 to 3371 causes 4096 out-of-order datagrams to be recorded. Further testing is necessary to determine whether the error rate can be lowered to a reasonable rate. There might also be a tradeoff that can be made such that slower transmission speeds can be used to allow for the use of a smaller buffer size.

It should be noted that we found no correlation between jitter and lost datagrams or jitter and throughput. While it seems like the radiation increased the average jitter, we do not know if the increase in jitter affects anything other than the arrival time of packets.

Finally, as UDP can only communicate 20% of the packets reliably, UDP is not that different than TCP/IP, which trans- mits data reliably at 6%-11% of the peak bandwidth. UDP has the disadvantage, however, of needing to retransmit the 80% of the lost packets. Many of these packets will need to be retrans- mitted multiple times, though, as the retransmitted packets will be subject to the same problems as the original transmission. Therefore, for many applications TCP/IP will be simpler than UDP.

V. RESULT

There could be temporary communication failures that cause the network to not pass data. Finally, we found that there is a degra- dation in performance in the neutron environment. Under these circumstances, we saw a 10–40-times decrease

in throughput for UDP and a 9–16-times decrease in throughput for TCP/IP.

ACKNOWLEDGMENT

Mrs. Anubhuti Khare one of the authors is indebted to Director UIT RGPV Bhopal for giving permission for sending the paper to the journal. Manish Saxena (HOD) is also thankful to the Chairmen, Bansal Institute of Science & Technology Bhopal for giving permission to send the paper for publication. Last but not least, I would also like to thank our HOD and colleagues for supporting us.

REFERENCES

1. A. Komornicki, G. Mullen-Schultz, and D. Landon,―Roadrunner: Hardware and software overview,‖ 2009 [Online].Available: http://www.redbooks.ibm.com/abstracts/REDP4477.html
2. Broadcom BCM57710 product brief,‖ 2008 [Online].Available: http://ko.broadcom.com/collateral/pb/57710 -PB05-R.pdf
3. A. Silburt, A. Evans, I. Perryman, S.-J. Wen, and D. Alexandrescu, [5] ―Design for soft error resiliency in internet core routers,‖ IEEETrans.Nucl. Sci., vol. 56, no. 6, pp. 3551 –3555, Dec. 2009.
4. T.-Y. Chuang, E. Schmidt, and S.-J. Wen, ―Platform neutron testing for single event upset (SEU),‖ in Proc. IEEE SELSE, Mar. 2009 [Online].Available: http://selse5.selse.org/program.html
5. A. Silburt, A. Evans, I. Perryman, S.-J. Wen, and D. Alexandrescu,―Specification and verification of soft error performance in reliable internet core routers,‖ IEEE Trans. Nucl. Sci., vol. 55, no. 4, pp.2389–2398, Aug. 2008.
6. S. Buchner, ―Evaluation of commercial communication network pro- tocols for space application,‖ 2003.
7. BCM57710/BCM57711 programmer's guide: Highly integrated media access controller programmer's guide,‖ BroadCom, Irvine, CA, Tech. Rep. 57710_57711-PG200R, 2009.
8. Iperf,‖ 2010 [Online]. Available: http://sourceforge.net/projects/iperf/
9. F. Halsall, Data Communications, Computer Networks, and Open Sys- [11] tems. Reading, MA: Addison-Wesley, 1996.
10. E. Normad and T. Baker, ―Altitude and latitude variations in avionicsSEU and atmospheric nuetron flux,‖ IEEE Trans. Nucl. Sci., vol. 40, no. 6, pp. 1484–1490, Dec. 1993.
11. B. Smith, ―An approach to graphs of linear forms (Unpublished work style),‖ unpublished.
12. E. H. Miller, ―A note on reflector arrays (Periodical style—Accepted for publication),‖ IEEE Trans. Antennas Propagat., to be published.
13. J. Wang, ―Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication),‖ IEEE J. Quantum Electron., submitted for publication.
14. C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.
15. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, ―Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces(Translation Journals style),‖ IEEE Transl. J. Magn.Jpn., vol. 2, Aug. 1987, pp. 740–741 [Dig. 9th Annu. Conf. Magnetics Japan, 1982, p. 301].
16. M. Young, The Techincal Writers Handbook. Mill Valley, CA: University Science, 1989.
17. (Basic Book/Monograph Online Sources) J. K. Author. (year, month, day). Title (edition) [Type of medium]. Volume(issue). Available: http://www.(URL)
18. J. Jones. (1991, May 10). Networks (2nd ed.) [Online]. Available: http://www.atm.com
19. (Journal Online Sources style) K. Author. (year, month). Title. Journal [Type of medium]. Volume(issue), paging if given.
20. R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. IEEE Trans. Plasma Sci. [Online]. 21(3).