

NASA Technical Memorandum 106397
AIAA-94-1101

Fault Tolerant Onboard Packet Switch Architecture for Communication Satellites: Shared Memory Per Beam Approach

Mary Jo Shalkhauser, Jorge A. Quintana, and Nitin J. Soni
Lewis Research Center
Cleveland, Ohio

Prepared for the
15th International Communications Satellite Systems Conference
sponsored by the American Institute of Aeronautics and Astronautics
San Diego, California, February 28–March 3, 1994



self-routing and efficiently utilizes bandwidth for “bursty” packet traffic. A DDPS was chosen for the ISP implementation, requiring user ground terminals to packetize circuit switched data.

Because actual traffic patterns are not known, a destination-directed packet switch can potentially have both contention and congestion problems. To avoid contention problems, a contention-free switch architecture was selected for the ISP. A contention-free switch is only possible because the total satellite throughput is less than 1 gigabit per second. Contention-free switches for satellite throughput in excess of 1 gigabit per second are beyond current technology. Like contention, however, switch congestion cannot be completely avoided. The ISP must monitor the traffic threshold (before congestion occurs) is reached. Network control can then initiate various congestion control measures to avoid congestion.

To reduce the amount of onboard storage, packets are divided into 16 subpackets of 128 bits each. The first subpacket of a packet contains header information which indicates the packet source, the destination downlink beam and dwell for point-to-point communication traffic, and also signifies multiple destination beams and/or dwells for multicasting. Figure 2 shows the packet and header format. The destination field contains 9 bits to specify the destination beam (1 bit for each of the eight destination beams and 1 bit to route messages to the onboard processor), 3 bits for destination dwell specifying 8 dwells per downlink beam, and 15 bits for destination user addresses. The user addresses are defined by network control and may indicate single users or groups of users. The final destination bit, the multicasting bit, indicates that there are multiple destinations for the packet. To reduce the required number of destination bits in the header and to reduce switch complexity, a restriction has been placed on the network so that when a packet is multicasted it is sent to all the dwells of the destination beam(s).

When a multicasting bit is set, the destination bits are slightly redefined. The destination beams are still indicated by the same 9 bits, however multiple bits may be set for multiple destination

beams. The 3 dwell bits are unused and all 8 dwells within the destination beam(s) will receive the multicasted packet. The 15 user address bits now contain multicasting addresses. The multicast destination users are informed of their multicasting address by network control.

Uplink data is transmitted using a MF-TDMA (multifrequency time division multiple access) technique. Downlink data is transmitted using a TDM format. The uplink and the downlink frames are divided into 16 subframes to take advantage of the subpacket structure. The ISP processes a subframe at a time. The first subframe of a frame contains all the header subpackets, which are used by the switch to set up the routing for the entire frame. Each downlink frame contains 8 dwells which can vary in size according to the needs of the current satellite traffic. The use of subpackets and subframes reduces the amount of onboard storage by a factor of 16, requiring only one subframe of data to be stored onboard the satellite at any time. Reference 1 contains a detailed description and analysis of the packet and frame structures.

A shared memory per beam switch architecture was chosen for the ISP. Reference 2 describes three contention-free architectures and discusses the rationale for choosing the shared memory per beam approach. The modularity of the shared memory per beam approach, and the fact that a good compromise between complexity of processing and memory efficiency is being reached, makes this design the best suitable for a packet switch architecture in the described application.

This paper provides a detailed description of the shared memory per beam architecture, and discusses critical issues such as congestion control and fault tolerant design. Plans for implementation and testing of the switch architecture are also described.

Approach

The shared memory per beam architecture is a destination-directed packet switch which is nonblocking and free of output contention. Switch congestion, however, can cause the beam

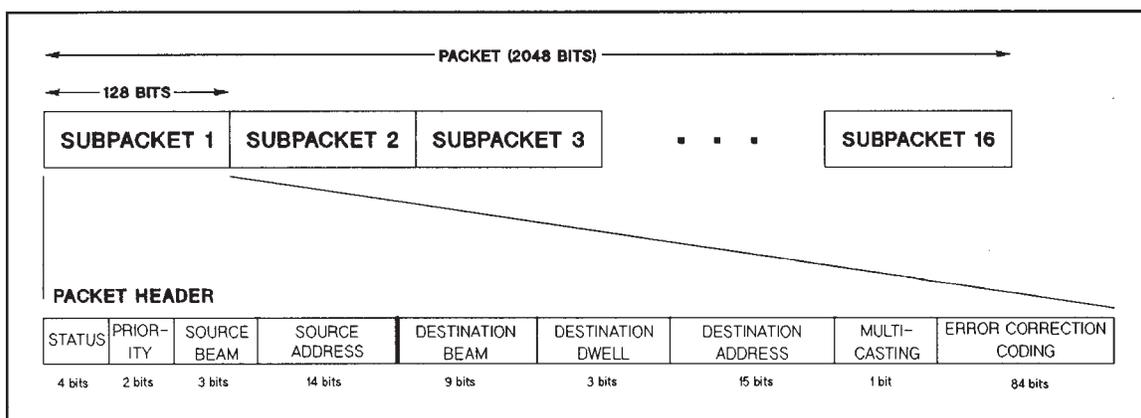


Figure 2 -- Packet and Header Format

control memory first-in, first-out memories (FIFO's) to overflow. Therefore, congestion control algorithms must be considered with this approach. The switch architecture readily supports both point-to-point and multicast communication traffic. The shared memory per beam switch is self-routing. Packet destination information is contained in the packet header and is used to route the packet to the correct downlink beam and dwell. Each downlink beam has a beam memory module which contains a data memory that is shared by each of the eight downlink dwells in the beam. Switching is performed first by routing the subpackets to their correct destination beam memory module and then by reordering the subpackets into their correct dwell locations.

Architecture

The shared memory per beam architecture is a hybrid between the fully shared memory and the shared bus architectures. In the shared memory per beam architecture, each downlink beam has a beam memory module which contains a data memory shared by its 8 downlink dwells. Switching is performed first by routing the subpackets to their destination beam (space switching) and then by reordering the subpackets into their correct dwell locations (temporal switching). Figure 3 contains a block diagram of the shared memory per beam architecture. As shown in the figure, the architecture can be subdivided into three major components: the input module, the switching and processing module, and the output module.

The input module consists of a multiplexer and an input FIFO for each uplink user. The satellite network contains 8 uplink beams, each supporting 32 time division multiplexed (TDM) wideband users (2.048 Mbps), for a total of 256 simultaneous uplink users. Each input FIFO receives input data at 2.048 Mbps, buffers the data and performs a rate conversion. The input FIFO reads are staggered so that the data from all 256 input lines can be

multiplexed onto a single high-speed TDM bus. The output data rate of each input FIFO, 524.288 Mbps, is the composite rate of the input data rate of all 256 users.

The heart of the shared memory per beam architecture is the switching and processing module. This module includes a beam header decoder and multiple beam memory modules, one for each downlink beam. The beam header decoder receives the high-speed TDM bus from the switch input module and determines the destination beam from the packet's header. First the beam header decoder determines if the packet is idle or busy from the status bits in the packet header. Idle packets are empty packets containing no information and are not processed and switched. If the packet is busy, the beam header decoder uses a look-up table or other type of bit mapper to determine the packet destination, and creates an enable signal to the appropriate destination beam memory module. The TDM bus is transmitted to all the beam memory modules but the beam header decoder will enable only the correct destination module. If a multicast signal is received, multiple beam memory modules are enabled.

Packet routing information is contained in the first subframe (header subframe of each frame. Each subsequent subframe in the same frame is routed exactly the same as the header subframe. The decoded output of the beam header decoder is stored in a frame routing memory during the header subframe to preserve the routing information for use in the remainder of the frame.

The shared memory per beam architecture contains a beam memory module for each downlink beam. As shown in Fig. 4, each module contains a dwell header decoder, a shared memory, an address pool FIFO, dwell control memories, and address control memories. Each beam memory module is responsible for reordering incoming subpackets into their correct dwell locations within a destination beam. The data flow is as follows: subpackets

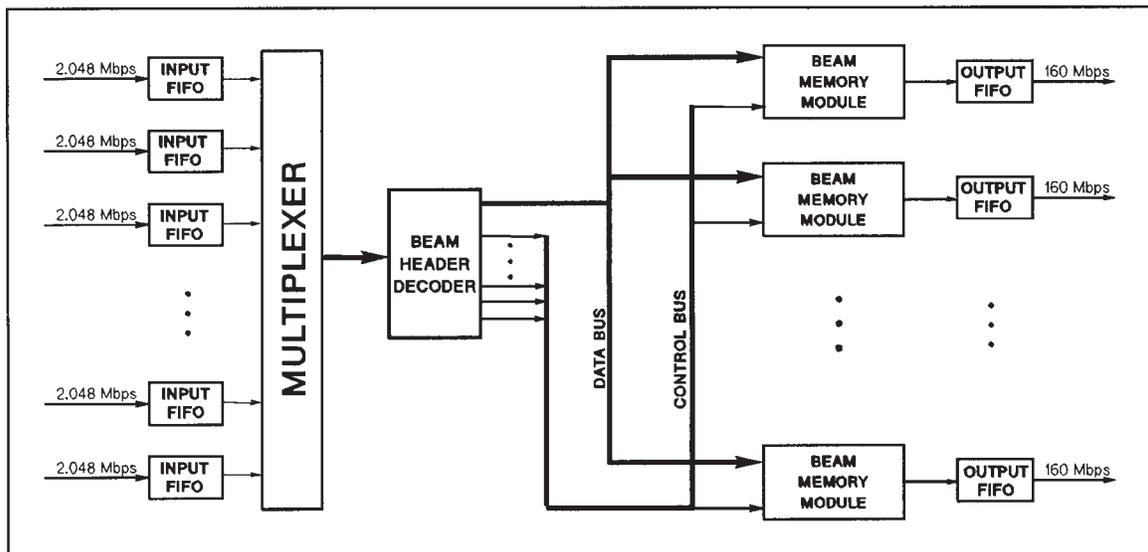


Figure 3 -- Shared Memory Per Beam Architecture

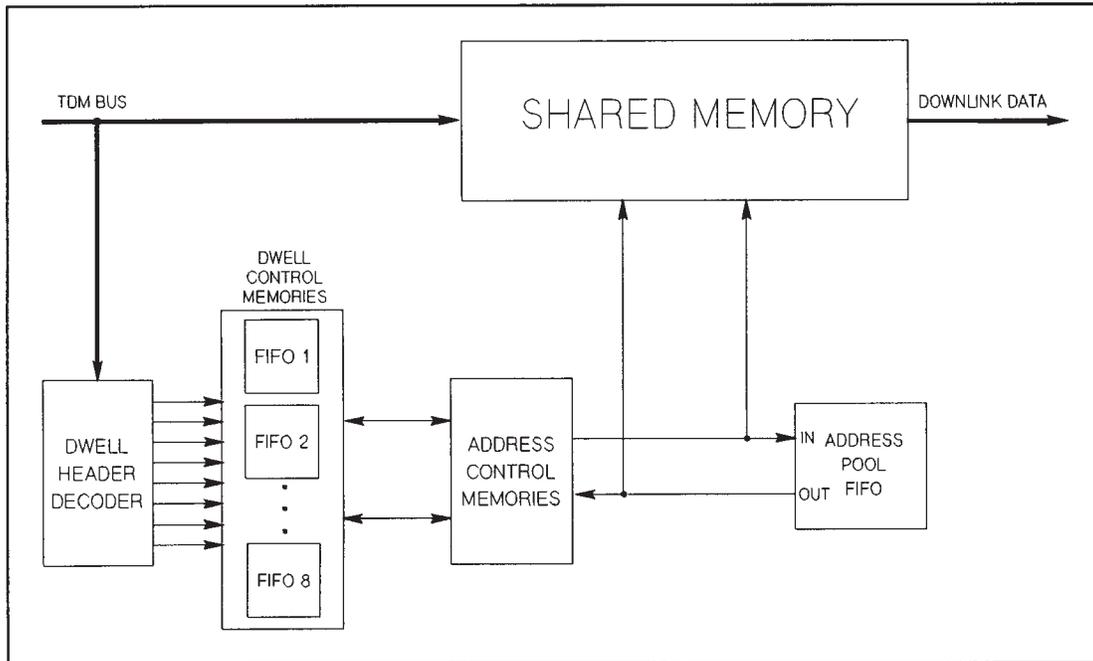


Figure 4 -- Beam Memory Module

received by the beam memory module are written into the shared memory at the first available unused memory location (provided by the address pool FIFO). Simultaneously, a pointer to the shared memory address location is written into an address control memory. Switching of the packet data to the correct destination dwell is achieved by controlling the order in which data is read from the memory.

The dwell header decoder consists of a look-up table or bit decoder that determines the destination dwell of the received subpacket from its header. During the first subframe of the frame, the subpacket data is routed to the header decoder which determines the packet destination and creates a write signal to the proper dwell control FIFO. When a multicast bit is present in the packet header, the dwell header decoder creates write signals to multiple dwell control FIFO's. In this manner, multicasting is achieved by multiple pointers to the same shared memory location. Packet duplication within a beam is thus avoided.

The Address Pool FIFO (APF) is used to avoid contention within the shared memory, by allowing packets to be written only into unused shared memory addresses. At startup, the address pool FIFO contains a list of all the shared memory addresses. When an incoming packet arrives at the shared memory, an address is fetched from the APF and the packet is written at that location in memory (address). When the packet is read from the FIFO, the (now available) shared memory address is written back into the APF. To avoid contention, both the APF and the shared memory will be sized slightly larger than one subframe; this guarantees an available shared memory address at all times.

The shared memory is designed in a dual port configuration so reads and writes can occur simultaneously as long as the same

address is not being accessed. Once an available address is determined by the APF, the data is written into the shared memory at that address. At startup, a full frame of data is written into the shared memory before any data is read out and routed to the downlink. Subsequently, frame F is written into the memory while frame $F - 1$ is being read ($F > 1$). The whole transmission takes $F + 1$ cycles compared to $2F$ if a dual-port configuration was not used. The actual size of the shared memory depends upon the chosen bus size and speed, and the amount of coding and redundancy required for fault tolerance.

The dwell control memories manage the temporal switching of the subpackets within a beam by controlling the order in which the data is read out of the shared data memory. These FIFO's consist of a pair of ping-ponged FIFO's for each downlink dwell. During the header decoder which determines the subpacket destination and creates a write signal to the proper dwell control FIFO. Then, the address control memory address (which points to the shared memory address where the packet is stored) is written into that dwell FIFO. This operation acts as a double pointer to the shared memory. For multicasted packets, the same address control memory location is written to multiple dwell FIFO's.

The dwell control FIFO's are also used for congestion detection and control. The fill level of each FIFO is closely monitored. When a fill threshold somewhere below full is reached, potential congestion is detected. Overflow (congestion) can be prevented by adjusting the dwell sizes within a beam, by notifying users to reduce their transmission rate, or by discarding low priority packets.

The dwell control FIFO's are sequentially accessed to read the downlink data out of the data memory and to complete the

switching process. The dwell control FIFO output points to an address control memory location whose output points to the shared memory address location containing the correct data subpacket. With these pointers, the correct subpacket can be read from the shared memory. The output module consists of one output FIFO per downlink beam module. The downlink FIFO's buffer the data and provide rate conversion to accommodate the downlink burst format. Once the outgoing packets are placed into responsible for their transmission.

In addition to these three modules described, the shared-memory per beam architecture will require some onboard processors to update the look-up tables, control the dwell control FIFO read signals, monitor congestion on board the satellite, monitor and report health status of the onboard systems, and communicate with ground based network control.

Memory requirements for the full-scale shared memory per beam switch architecture are shown in Table 1. These figures do not include the memory necessary for the onboard processors or for fault tolerance purposes. The total architecture was estimated to consume 300 W of power and to consist of 14 million equivalent gates, assuming a TTL and CMOS implementation.

TABLE 1.-ISP MEMORY REQUIREMENTS

Memory	Depth	Width	Number
Shared memory	20,000 bytes (15 address bits)	16 bits	8
Address pool FIFO	2500 bytes (12 address bits)	12 bits	8
Frame routing memory	8192 bytes	8 bits	1
Dwell control memories	2500	12 bits	64
Address control memories	2500 bytes (12 address bits)	12 bits	2
Uplink FIFO's	256 bits	1	256
Downlink FIFO's	40,000 bytes	8	8

Fault Tolerance

Onboard satellite systems are susceptible to failures due to the harsh space environment. Therefore, fault tolerant design plays a crucial role in maintaining reliable operation of the onboard systems. The shared memory per beam packet switch will be implemented using fault tolerance techniques such as fault avoidance, fault detection, fault correction, redundancy, and reconfiguration.³ Critical portions of the design, those most susceptible to failure and those whose failure would cause the most problems, will first be identified. Fault tolerant design and special manufacturing techniques will then be used. Critical subsystems will be designed with techniques of redundancy, error detection and correction, and reconfiguration to maintain reliable operation and to gracefully degrade the systems in the event of failures.

Various fault tolerance objectives have been identified for the information switching processor. Minimizing the probability of packet loss, misrouting, and corruption of utmost importance to achieve fault-free switching. Faulty paths through the switch will be detected and corrected by the onboard fault tolerant circuits. Self-checking circuits, which continuously monitor subsystems during normal operation and detect these faults,⁴ will be used. Off-line testing cannot detect the transient or intermittent faults that are typically found in VLSI circuits. Ensuring availability, reliability, and performance of the packet switch over a long period of time is crucial.

When designing for fault tolerance and fault avoidance, identifying which component should be protected and how much protection is enough, but not too much, are two main concerns that arise. One critical subsystem in the shared memory per beam architecture is the TDM bus. Redundant paths through the TDM bus are essential because a failure can result in total loss of switching capability. Another critical subsystem is the beam header decoder. If the beam header decoder fails, data may be routed to an incorrect destination beam. Having a "hot spare" will help remedy this situation. Other critical subsystems are identified in Table 2.

Fault tolerant design and manufacturing techniques to minimize interconnect-related, environment-related and contamination-related failures are being considered. The single most significant factor in the failure of electronics circuits is the interconnect complexity. Minimizing these failures can be accomplished by minimizing quantity of interconnects and by controlling the quality of interconnects. Radiation phenomenon, thermal problems, and electrostatic discharge can be avoided by hardening, temperature control, and electrical shielding respectively.⁵

As circuits become more complex, determining that a failure has occurred becomes more difficult. In the shared memory per beam structure, the beam memory module has several memory components that are susceptible to single event upsets (SEU) and this can lead to misrouted or lost data. Typically, memories that are used for switching will be protected from SEU effects by error correcting or majority voting schemes. For hard failures, detection will be done by on-line and off-line testing techniques.

In-circuit error correction resides at the gate or latch hardware level, and monitors data errors by syndrome or check bits appended to the base data. This type of compensation is attractive because the correction is automatic and transparent, and can potentially eliminate both hard and soft errors.⁵ As shown in Table 2, the correction mechanism suggested for each subsystem failure is predominantly error correction coding. Another mechanism suggested is redundancy, either on-chip or system level. Subsystems such as beam header decoder and dwell header decoders are replicated one-for-one, but beam memory modules are replicated one-for-eight. Therefore, if one beam memory module experiences failure, the spare one can be switched into service

TABLE 2.--FAULT TOLERANCE CHART

Critical	Subsystem	Failure effects	Detection mechanism	Correction mechanism	Reconfigure technique
	Input FIFO's	Lost data Lost headers	Error correction coding (ECC)	ECC	Spares
*	TDM Bus	Total loss of switching capability	Majority voting (MV)	MV	Redundant paths, fault tolerant topology
	Shared memory	Lost data	ECC	ECC	1 for 1 spare, partitioning, APF
*	Beam header decoder	Misrouted data	ECC,MV	ECC,MV	Hot spares
*	Dwell header decoder	Misrouted data	ECC,MV	ECC,MV	Hot spares
*	Dwell control FIFO's	Misrouted data	ECC, self-checking circuits	ECC	1 for 8 spares, time-sharing
*	Address control memories	Misrouted data	ECC	ECC	1 for 1 spare, partitioning
	Address pool FIFO	Lost SM locations, possible data overwrite	ECC	ECC	1 for 1 hot spare
	Beam memory modules	Loss of downlink beam	ECC,MV	ECC,MV	1 for 8 spare, time-sharing
	Output FIFO's	Lost data	ECC	ECC	1 for 1 spare
	Processors	Loss of switch control	Watchdog timers, software	Software techniques	Spares

without losing a downlink beam. In many cases, redundancy is the only alternative to guarantee a predictable performance baseline over the spacecraft lifetime. Off-line testing techniques such as boundary-scan and self-test also need to be investigated.

Once fault tolerance techniques have been chosen for implementation, demonstration and performance verification of the techniques becomes an issue. The control logic needed to demonstrate the capabilities of fault tolerant schemes can become complex. These circuits must be capable of simulating failures, such as SEU's and hard failures, in the ISP hardware. Additional circuits and techniques must be used to verify the success of the fault tolerant hardware. The bit error rate is one parameter that is often observed during performance verification.

NASA Lewis plans to identify and document the fault tolerance design techniques needed for the entire switch design. Components will be chosen that are manufactured using techniques such

as radiation hardness and shielding, and good fault avoidance design practices will be used. Identification of fault susceptible and critical performance subsystems is already underway. For the demonstration hardware model, two or three areas of interest will be selected for protection, and a different fault tolerance technique will be applied to each one.

Implementation and Testing

The shared memory per beam packet switch architecture is complete. NASA plans to simulate the entire packet switch and to implement critical portions in hardware and demonstrate them in a laboratory environment. A scaled-down model of the shared memory per beam architecture has been identified and is shown in Figure 5. This scaled-down switch will be modeled using VHDL (VHSIC Hardware Description Language) to demonstrate the feasibility of the entire switch architecture.

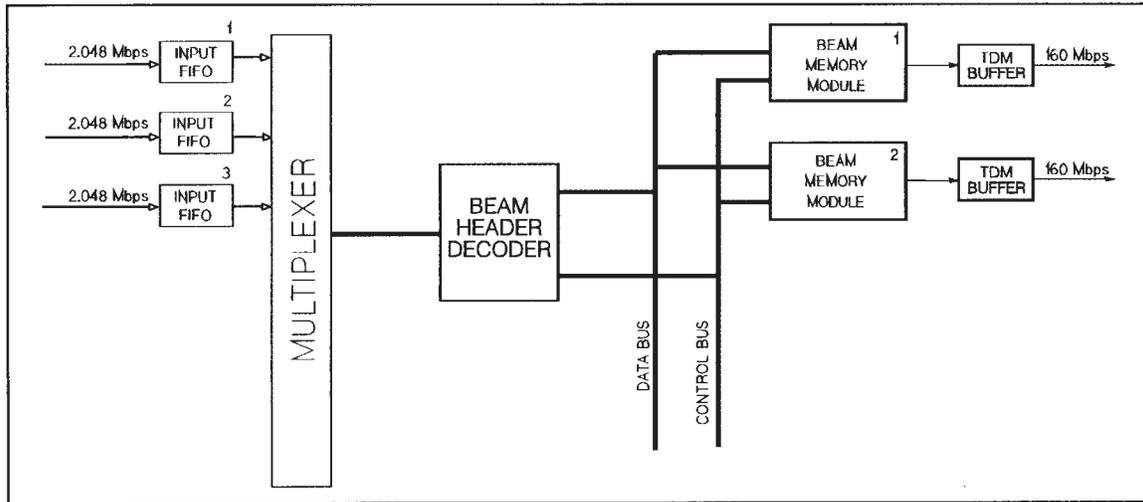


Figure 5 -- Scaled Down Information Switching Processor Architecture

To simulate fault tolerance, NASA Lewis is using a software package, called Ptolemy. This package enables a user to study the interaction between call-processing software and hardware switch elements. The Ptolemy software can also simulate fault detection and monitor the overall health of the subsystem during normal operation, without taking the unit off-line.⁶

Two portions of the switch architecture have been identified as critical subsystems—the high-speed TDM bus and the beam memory modules. The high-speed TDM bus, which includes the beam header decoder, will be studied to determine the best method for implementation so that the bus is fault tolerant in a space environment. Alternatives to an electronic bus, such as fiber optics, are under investigation. The chosen architecture will then be implemented in hardware for proof of feasibility and demonstration.

The other critical subsystem is the beam memory module which performs the bulk of the packet switching. A scaled-down version of the beam memory module will be demonstrated in hardware. Portions of the beam memory module will be designed using VHDL models and logic synthesis software to take advantage of the models in the VHDL simulation of the entire switch. Within the beam memory module, fault tolerance techniques will be implemented where necessary and feasible.

A significant amount of special test equipment (STE) will be needed to fully test the features of the high-speed TDM bus and the beam memory module. The STE responsibilities will include the creation of user traffic patterns and data to test the functionality of the switch and to demonstrate congestion control, control and monitor of all tests, verification of proper routing through the switch, and the simulation of faults and verification of fault recovery.

Summary and Future Directions

The NASA Lewis Research Center is developing a flexible, high-throughput, fault tolerant onboard information switching processor for a multichannel communication signal processing satellite. This information switching processor (ISP) is a baseband switch used for routing user information among user ground terminals. A shared memory per beam switch architecture was chosen for the ISP destination directed packet switch. This approach performs switching by first routing the subpackets to their correct destination beam memory module and then by reordering the subpackets into their correct dwell locations.

NASA plans to simulate the entire shared memory per beam packet switch, and then to implement critical portions in hardware and demonstrate them in a laboratory environment. A scaled-down switch will be modeled using VHDL to demonstrate the feasibility of the entire switch architecture. Two critical portions of the switch architecture, the high-speed TDM bus and the beam memory modules, will be implemented in proof-of-concept hardware.

Because onboard satellite systems are susceptible to failures in a harsh space environment, fault tolerant design is required to maintain reliable operation. The critical portions of the shared memory per beam packet switch will be implemented using fault tolerance techniques. A significant amount of special test equipment will be needed to fully test the features of the high-speed TDM bus and the beam memory module.

Upon successful completion of the switch simulation and demonstration of the identified critical subsystems of the switch,

NASA will develop a full-scale ISP, either in-house or through industry contracts. In addition to the switch development, efforts will focus on the creation of realistic user data and traffic loading to fully test the switch functionality, congestion control, and network control.

References

1. Ivancic, W.D., et al., "Destination Directed Packet Switch Architecture for a Geostationary Communication Satellite Network," World Space Congress, 43rd Congress of the International Astronautical Federation, Aug. 28–Sept. 5, 1992.
2. Shalkhauser, M.J., and Quintana, J.A., "On-Board Packet Switch Architectures for Communication Satellites," NASA TM-106328, Sept. 1993.
3. Raghunandan, K., Caokley, F.P., and Evans, B.G., "Fault Tolerance of On-Board DSP Circuits," John Wiley & Sons, Ltd., Aug. 1991.
4. Lala, P.K., "Self-Checking Logic Flags Errors as They Happen," EDN-Design Feature, Jan. 7, 1993, pp. 73–78.
5. "Identification and Analysis of Failure Mechanisms for Space-Based Electronics," by Space Systems/LORAL, Final Report of Task 3 of Contract NAS3-25934, June 1993, pp. 4.1–4.15.
6. Buck, J., et al., "Ptolemy: A Framework for Simulating and Prototyping Heterogeneous Systems," by Department of EECS at University of California, Berkeley, Aug. 31, 1992.

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> <i>OMB No. 0704-0188</i>	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE January 1994	3. REPORT TYPE AND DATES COVERED Technical Memorandum		
4. TITLE AND SUBTITLE Fault Tolerant Onboard Packet Switch Architecture for Communication Satellites: Shared Memory Per Beam Approach			5. FUNDING NUMBERS WU-506-72-21	
6. AUTHOR(S) Mary Jo Shalkhauser, Jorge A. Quintana, and Nitin J. Soni				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Lewis Research Center Cleveland, Ohio 44135-3191			8. PERFORMING ORGANIZATION REPORT NUMBER E-8224	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, D.C. 20546-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA TM-106397 AIAA-94-1101	
11. SUPPLEMENTARY NOTES Prepared for the 15th International Communications Satellite Systems Conference, sponsored by the American Institute of Aeronautics and Astronautics, San Diego, California, February 28-March 3, 1994. Responsible person, Mary Jo Shalkhauser, (216) 433-3455.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 17			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The NASA Lewis Research Center is developing a multichannel communication signal processing satellite (MCSPS) system which will provide low data rate, direct to user, commercial communications services. The focus of current space segment developments is a flexible, high-throughput, fault tolerant onboard information switching processor. This information switching processor (ISP) is a destination-directed packet switch which performs both space and time switching to route user information among numerous user ground terminals. Through both industry study contracts and in-house investigations, several packet switching architectures were examined. A contention-free approach, the shared memory per beam architecture, was selected for implementation. This paper describes the shared memory per beam architecture, fault tolerance insertion, implementation, and demonstration plans.				
14. SUBJECT TERMS Onboard processing; Packet switch; Communications satellite; Fault tolerance			15. NUMBER OF PAGES 10	
			16. PRICE CODE A02	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	