COMPARATIVE SIMULATIVE ANALYSIS OF WDM LANS FOR AVIONICS PLATFORMS

Casey Reardon, John Profumo, and Alan D. George
High-Performance Computing and Simulation Research Laboratory
University of Florida
Gainesville, Florida

ABSTRACT

With their almost unlimited potential for performance and their decreasing costs, advanced optical components and networks are now being seriously considered for deployment in emerging avionics systems. Towards the goal of developing an advanced avionics network that features wave-division multiplexing (WDM) for performance that is highly scalable, dependable, protocol-independent, and versatile, many disparate architecture strategies need to be evaluated. Due to the high cost of testbed prototyping and integration with existing systems, a simulative approach is used in this study to analyze and compare candidate WDM LAN architectures at a high level. Using discrete-event simulation models developed at the University of Florida, several contrasting approaches are examined for constructing an optical network architecture supportive of future avionics requirements. Each architecture is evaluated in terms of two application scenarios. The results from the simulation experiments enable a high-level comparison of competing architectures and provide insight for aerospace network researchers and designers.

1. INTRODUCTION

Advanced optical networks featuring wavelength division multiplexing (WDM) have been targeted as the technology of choice to realize the desired avionics network of the future. There are numerous advantages that WDM networks can offer, some of which include almost unlimited bandwidth potential, resistance to electromagnetic interference, and the potential for a unified network with protocol independence. Despite these many major advantages, there are just as many challenges when considering optical network technology to realize a high-performance, local-area network (LAN).

Traditionally, optical WDM links have been reserved for long-haul links and high-bandwidth trunks. These cases are drastically different than the avionics LAN environment, where links are relatively short and numerous, and signals need to be routed amongst a large number of nodes. Additionally, the harsh environment of aerospace platforms combined with the mission-critical nature of avionics applications requires the network to be highly reliable and fault-tolerant. Meeting all of these requirements presents a difficult challenge to network designers. Without a widely accepted solution to meet these requirements, numerous ideas for network designs need to be formulated and analyzed.

To investigate and realize such an advanced network, virtual prototyping of potential ideas and architectures to meet these requirements will be necessary. The advantages of a computer modeling approach to network design are obvious, especially when considering the cost and delays associated with the fabrication and testing of physical prototypes. In this paper, we evaluate a set of potential avionics LAN designs using WDM technology. Those designs are compared via simulative experimentation and analytical analysis. The results and analysis presented here provide a high-level comparison of competing architectures, and we believe represents a valuable step towards creating an optimized solution for a WDM avionics network architecture.

The remainder of this paper is organized as follows: Section 2 provides background information describing the requirements of a military avionics network, and the methods used for this research. Section 3 provides details describing the optical network architectures proposed and evaluated in this paper. Section 4 describes the experiments used to compare the proposed architectures. The results of these experiments are presented and analyzed in Section 5. Finally, the conclusions from the work presented are provided in Section 6.

2. BACKGROUND

Before attempting to design a network architecture for any community, it is important to identify the needs and requirements of the network users. We use the basic guidelines proposed by the working session problem at the 2005 AVFOP conference when designing and evaluating network architectures [1]. The requirements described there included high-speed transmission between any two nodes on the network, with the number of nodes scaling up to a maximum of 256. While operation at full scale is important, it will also be desirable that any network architecture can be easily implemented for contemporary avionics platforms, which may often contain far fewer than 256 nodes. Finally, the network should be fully functional under the presence of one or two faults, with graceful degradation after three faults.

All of the network architectures analyzed in this paper were modeled using the Library for Integrated Optical Networks (LION). Developed at the University of Florida, LION provides researchers with an extendable tool to assess both lower- and upper-layer networking issues simultaneously by providing a set of accurate optical components within a powerful network simulation environment. LION was created in a discrete-event simulation environment called MLDesigner from MLDesign Technologies. MLDesigner allows custom models to be defined using C-based code. Numerous low-level models can then be combined in a hierarchical structure to realize complex systems. The components in LION enable accurate modeling of timing and physical effects inside optical devices. New and legacy network protocols can be implemented on top of components to realize and evaluate almost any system design.

3. PROPOSED NETWORK ARCHITECTURES

In order to design and realize an optimized network architecture to meet the current and future needs of the military avionics community, a number of system designs need to be evaluated. Thus, we have identified an initial set of promising optical network designs, which have been modeled and analyzed. These systems represent a wide range of design approaches, featuring numerous topologies and control protocols. Additionally, the extent to which the networks rely on optics to function varies, as will be clearly illustrated. Most of the proposed designs also represent examples of network architectures that have been previously developed for alternate networking applications. The remainder of this section is spent detailing each of the network architectures.

The first network design, the *ring-ring* architecture, is illustrated in Figure 1. The basic design of the ring-ring comes from the ROBUS network architecture [2]. This network consists of multiple rings of local nodes, connected by a separate master ring. At the local level, network nodes are grouped together and connected using redundant rings. The use of redundant rings allows this network to maintain full connectivity under the presence of multiple faults, a key advantage of this architecture. Each node places data on the ring traveling in both directions, while being able to receive data from both directions on the ring. The use of bidirectional transmission allows each ring to suffer one cut and still operate at full functionality. Each local ring is equipped with a ring-leader node, which provides the interface between the local ring and master ring. This node is responsible for routing wavelengths between rings, and removing local traffic from the local ring. The ring-leader will also perform additional functions, depending on the control protocol. We allow each local ring to host up to 16 network nodes. Using 16 of these rings will provide scalability to 256 nodes. A drawback of this approach is the difficulty in maintaining acceptable optical signal powers, as each optical signal must reach all local nodes and ring-leaders. This limitation requires carefully tuned optical amplifiers throughout the rings.

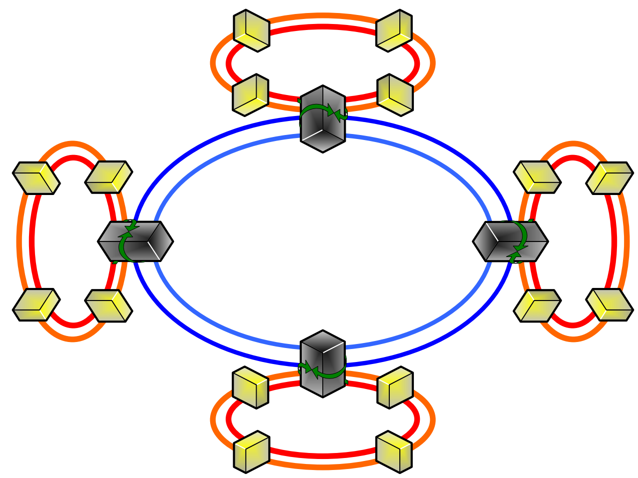

Figure 1: Ring-Ring Architecture

For the ring-ring architecture, two different control protocols are employed in this study. The first is a static time-division, multiple-access (TDMA) protocol. In this protocol, each destination on the ring is assigned a wavelength. The same wavelengths can be reused within each group, so only 16 wavelengths are required for all local traffic. Longer time slots are used for inter-group traffic, since this traffic will travel longer distances thus additional time is required to allow traffic to clear any fiber links it was using. For smaller-sized systems, such as a 16-node ring, a simple TDMA system can work well. Unfortunately, pure static TDMA will break down when the network scales, and we have to consider traffic traveling between rings. For inter-ring traffic, we use a compromise between a TDMA and reservation protocol, while introducing an optical-electrical conversion. Each ring is assigned one or two wavelengths to receive traffic from remote nodes. The groups that must share a wavelength to reach a remote group use TDMA to eliminate contention. Within each ring, nodes request the ring-leader for access to transmit on the desired non-local wavelength. Once the ring-leader grants access, the sending node can send during their group's time slot. The transmitted data is collected and retransmitted by the remote ring-leader.

The second control protocol we consider is a reservation-based control protocol (RSVP). Again, each destination on the ring is assigned a wavelength, and the same wavelengths can be reused among groups. In this scenario, nodes that wish to send data issue a request to the ring-leader on a reserved wavelength. The controller responds when the requested wavelength is available, allowing the sender to transmit. When transmission is complete, the sender notifies the controller again, releasing its token for that wavelength. Inter-group traffic is handled in the same way as described with TDMA protocol. The one exception is that the group controller does not wait for its assigned time slot to send remote data to local nodes. The controller instead waits for all pending requests to the destination to be satisfied, before transmitting the received data to the destination.

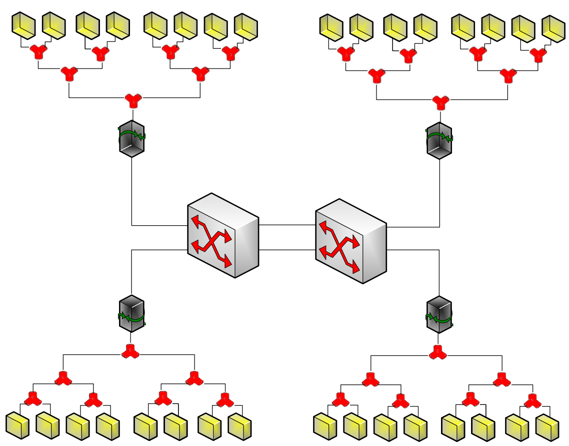

Figure 2: Optical Tree Architecture

Our second candidate architecture uses optical trees to connect groups of nodes, while using a star coupler to connect the distributed trees [3]. Similar to the ring-ring system previously described, this architecture can be viewed as a two-level design. Figure 2 provides an illustration of the optical tree architecture. We allow 16 nodes to be connected on each tree, with 16 trees leading to a network that can scale to 256 nodes. The tree-leaders, which are placed at the root of each local tree, will be asked to perform many of the same functions as the ring-leader in the previous system. Thus, the tree-leaders interface each tree with the central coupler, and perform the necessary wavelength routing. Data produced by each node is sent up the tree to the root, or tree-leader. The tree-leader is responsible for routing the wavelength to the appropriate tree. The tree architecture is not as inherently fault-tolerant as rings, but there are two advantages the tree architecture provides. Nodes on the trees will be simpler, as they only need one interface with the tree. Second, the lengths between each node and the root are consistent, unlike in rings, thus time slots are used more efficiently. Both control protocols used with the ring-ring architecture are considered here as well. Each local tree functions the same as a local ring, and tree-leaders now replace the ring-leader nodes from before.

The third candidate system represents a hybrid optical-electrical network architecture. This network is based largely off of the LAN architecture and solutions provided by commercial companies such as Matisse Networks [5]. In this architecture, nodes are connected to switches with electronic links, while optical links are reserved for communication between switches. The switches in the network are connected in a ring. By using common components such as Ethernet of Fibre Channel devices in this architecture, the costs of implementing this network are very low. Costs are also minimized by using a very minimal amount of optical devices. One drawback is the difficulty of fault-tolerance, as each switch must be duplicated to overcome most faults. Another drawback is the limited bandwidth provided to each individual node where electronic links are used, but this can be overcome by allocating multiple ports to a single node or device. Figure 3 provides an illustration of the hybrid architecture.

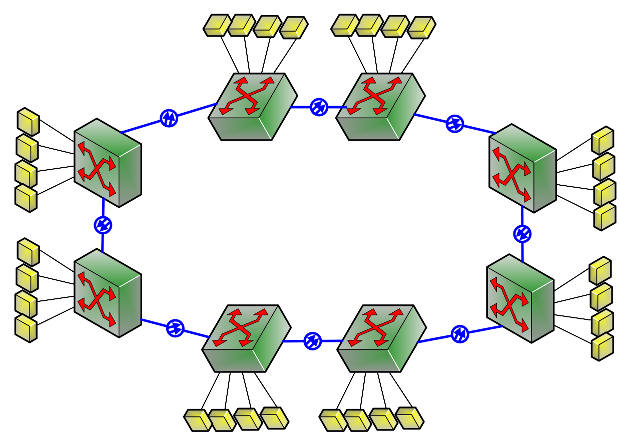

Figure 3: Hybrid Architecture

The switches in this case-study contain 32 local ports. By using 32-port switches, only 8 switches are required for the network to scale to 256 nodes. Each switch is also outfitted with two optical ports, each with a laser transmitter and set of optical receivers. By using multiple receivers at each optical port, an entire wavelength can be dedicated for transmission between each pair of switches. Since optical components are limited to the switches only, this will not be very costly. Each switch on the optical ring will filter locally destined wavelengths, while allowing all other wavelengths to pass through. For each node, we have assumed a low-latency NIC which requires 4 μs of processing time per packet at each end, not including any queueing delays within the NIC. While this setting is faster than seen by most typical commercial NICs, these values can be viewed as exclusive of application and transport layer delays that may eventually increase packet latencies (in all systems).

Our last system architecture uses optical switches to form a Clos network, as illustrated in Figure 4. First proposed by Charles Clos in 1953, a Clos network is a highly-connected multi-stage switched network that provides multiple paths between end ports, which minimizes or eliminates blocking in the network [7]. In this system, we use switches modeled after the time-slotted OSMOSIS switch architecture developed at IBM, as part of the DARPA HPCS program [4]. A simplified version of their architecture is used and presented here.

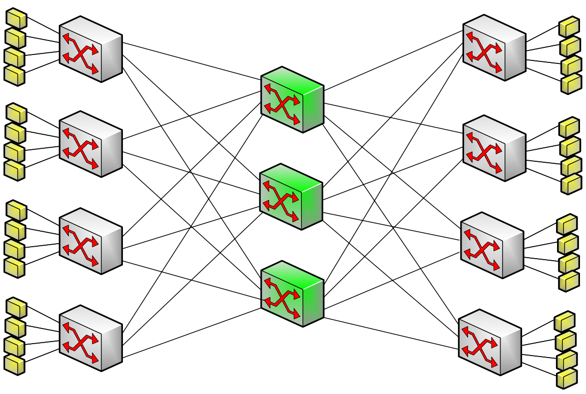

Figure 4: Optical Clos Architecture

Our Clos network is comprised of eight 32-node perimeter switches, plus three 8-port backbone switches. Each connection between two ports includes both electronic and optical links. The electronic link is used to inform the switch of transfer requests from the nodes. These requests include the desired destination and the number of required time slots needed for the data. An arbiter within each switch quickly processes these requests, and schedules transfers in a simple round-robin format in our implementation. When a requested transfer is scheduled to take place, the arbiter informs the participating nodes of this, including how many consecutive time slots the sender is allotted. If the time slots allotted to the transmitter are fewer than the number it requested, the node waits until informed again by the arbiter it is allowed to transfer. For each new message transmission, nodes must submit a new request to the arbiter before sending again. The optical switching is performed using the broadcast and select approach detailed in [4]. The use of purely optical switches, combined with a Clos topology that provides multiple paths between perimeter switches, provides the highest potential bandwidth of any architecture proposed here. These advantages come at a cost though, since the Clos would be the most costly proposed architecture to implement. Currently the optical switches themselves are complex and unproven, although this is likely to change in the coming years as the technology matures. The cost increases even more when considering fault-tolerance in a Clos network, which requires redundant switches all around the perimeter. Furthermore, the Clos network requires a high degree of cabling, which may be difficult to achieve in the tight confines of many aircraft.

4. EXPERIMENTAL SETUP

In an effort to generate useful results and analysis predicting network performance, we attempt to create experiments that will mimic real conditions on an avionics platform. Thus, two experimental configurations representative of actual avionics platforms have been constructed to stimulate our simulation experiments. The traffic sources in these configurations generate traffic that is bursty, periodic, or random in nature. These two configurations are briefly described in this section.

The first configuration used in these experiments is our military configuration. This system is largely based off of the architecture of the F-22 Raptor, as documented in [6]. The Raptor uses a centralized architecture, where data acquired from remote sensors and actuators are gathered and processed at the core processing units. The core processing system includes two common integrated processors, or CIPs. Each CIP includes a network of data and signal processors, memory and other units. Our military configuration includes two CIP sets in our core processing subsystem. In total, our military configuration includes 97 nodes across seven subsystems. In the baseline case, this system generates about 200 megabits of traffic per second. Figure 5 gives a top-level diagram of the military configuration design.

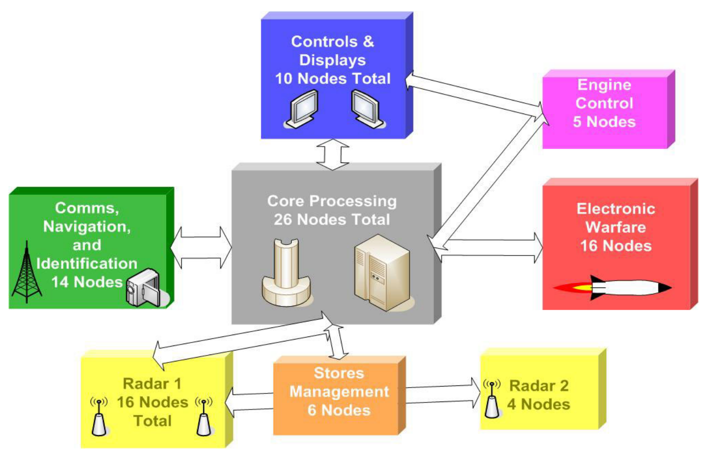

Figure 5: Military Configuration

The second configuration used in these experiments is our commercial configuration. The primary difference between the two configurations is that the commercial system is a more distributed architecture than the military system, which was a centralized one. There is no central unit in the commercial system used to process critical data. Each subsystem instead takes on more responsibility for processing local information. Subsystems are now free to directly communicate with each other, without going through central processing. Thus, the network will see an increased variety of communication between subsystems. The commercial configuration includes a total of 84 nodes across eight subsystems. In the baseline case, this system generates about 300 megabits of traffic per second. Figure 6 gives a top-level diagram of the commercial configuration design.

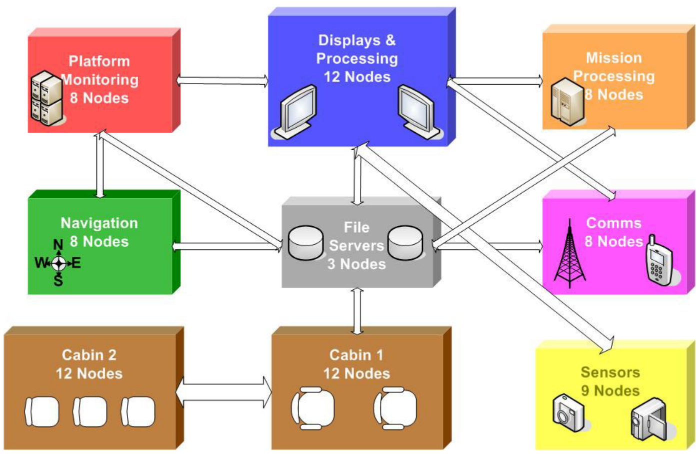

Figure 6: Commercial Configuration

5. RESULTS

In this section the structure and results of the simulation experiments are presented. The results for each of the architectures are then discussed, along with comparative analyses between the candidate architectures. Each of the six candidate architectures was tested and evaluated using four different experiments. The first two experiments consisted of subjecting the architecture to one second of traffic from the baseline military and commercial configurations previously described. For the final two tests, the traffic rates from all of the nodes were scaled by a factor of 10 to represent demanding traffic loads on future platforms. As before, one second of network traffic was again used to stimulate the systems. In all of the experiments, the payload size of each message was uniformly distributed between 100 and 2,500 bytes.

The optical transmitters used in every scenario operate at 2.5 Gbps. Additionally, the electronic links used in the Clos and Hybrid architectures operated at 1 Gbps. The timeslot period used in both TDMA architectures was 5 us for local traffic, and 6 us for traffic between rings or trees. A 500 ns timeslot period was used with the OSMOSIS switches in the Clos architecture.

The results from all of these experiments are presented in the following tables and figures. Table 1 summarizes the overall average packet latencies from each system under each of the four experimental configurations. Tables 1 and 2 illustrate the overall packet latency for each system in the baseline military and baseline commercial experiments, respectively. In addition, Table 2 shows the worst-case packet latencies for each experiment, which are determined by taking the average latency of the ten slowest packets in each test. This ten-packet average was used so that the result of a single packet would not skew the perception of the system's overall performance. At the same time, with each system providing data from tens or hundreds of thousands of packets, the average of the ten slowest will generally not be far off from the slowest packet overall.

Table 1: Average Packet Latency (us)

	Ring-TDM	Ring-RSVP	Tree-TDM	Tree-RSVP	Hybrid	Clos
Mil-1x	1,109	223	1,122	230	52	21
Mil-10x	93,075	65,880	92,973	60,955	78	24
Com-1x	4,264	140	4,188	139	40	22
Com-10x	113,249	113,640	116,841	113,072	44	23

Table 2: Worst-Case Latency (us)

	Ring-TDM	Ring-RSVP	Tree-TDM	Tree-RSVP	Hybrid	Clos
Mil-1x	20,045	3,167	20,045	3,453	365	124
Mil-10x	100,279	646,698	988,777	703,736	973	264
Com-1x	291,607	2,121	289,717	2,203	143	94
Com-10x	1,053,385	943,428	1,051,196	943,727	352	109

The results tabulated from all of our simulative experiments reveal several clear trends. The first major trend is that the Clos network exhibits the highest performances in all four experimental setups. The average packet latency for the Clos networks is less than half that of the hybrid network in each setup, while performing several orders of magnitude better than both RSVP and TDMA cases. The results should not be surprising for several reasons. The OSMOSIS architecture uses switches with an optical backplane that can handle up to 60 Gbps of optical data in this configuration. As importantly, the scheduling in OSMOSIS is very fast and highly efficient, since each node has its own dedicated link to the scheduler. This method means that bandwidth is inefficiently used in most traffic scenarios, which is illustrated by the low worst-case latency values. Moreover, a Clos network provides high connectivity between switches, which increases the total bandwidth the system can support and provides multiple paths to allow the network flexibility in how data is routed across it. All of these factors combined lead to the high performance observed. Even with these expectations, the performance of the Clos network is still impressive, as 20 us latencies for messages averaging 1300 bytes in size is very fast.

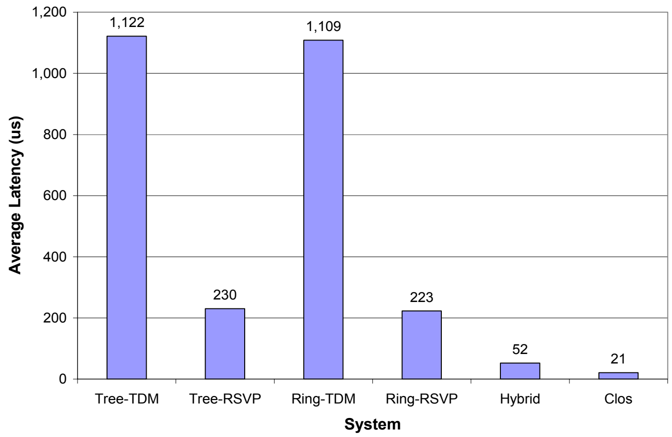

Figure 7: Average Packet Latency (Baseline Military Configuration)

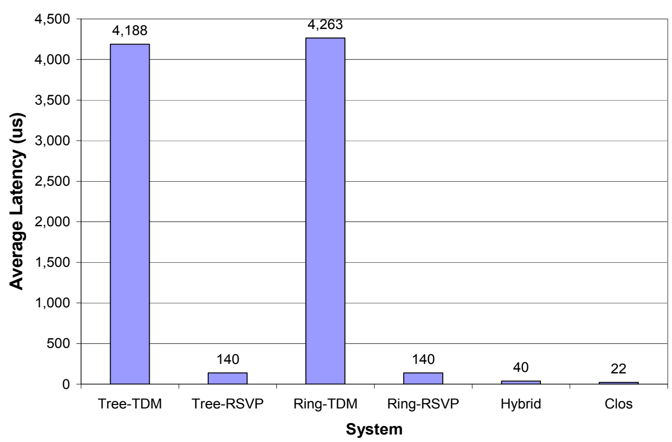

Figure 8: Average Packet Latency (Baseline Commercial Configuration)

The performance of the Clos network was constant across all four experiments. This behavior was not observed with any other candidate architecture, whose average packet latencies saw very significant increases as the generated traffic rates were elevated. The results suggest that the Clos network not only offers the highest performance at current traffic levels, but will provide superior performance with even higher traffic demands. It should also be noted that a Clos network can easily accommodate additional nodes by simply connecting an additional switch to the perimeter. Also, if performance was affected due to contention through backbone switches, a fourth or fifth switch can be added to the backbone to accommodate additional bandwidth requirements. However, the high performance and fault-tolerance do come at a significant price, as described in Section 3.

The hybrid system consistently showed the second best performance results across all systems. While the average latencies of the hybrid system were over twice those exhibited by the Clos network, it still performed orders of magnitude better than the other four systems. The primary reason for the large disparity is the efficiency of the control protocols. Since traffic in data networks is often bursty in nature, it is important that the network architecture can handle traffic bursts. Architectures that can quickly allocate resources these bursts are highly desirable, such as switched networks which are common and proven in high-performance networks. The Clos network takes advantage of the fast arbitration the OSMOSIS architecture was designed around. In the hybrid system, the only contention occurs from buffering packets at the switch and NIC ports.

Despite their efficiencies, the hybrid system lags behind the Clos system for two major reasons. First, the Ethernet switches do not provide nearly as much throughput as the optical switches are capable of providing. Meanwhile, the Ethernet switches modeled here do not switch as fast as the optical switches. Despite this limitation, the operation is highly efficient, as bandwidth in the switches is not wasted when nodes are inactive. The hybrid system does degrade in performance when traffic levels were increased, especially on the military system where the latency average increased 25 us. This result is largely due to the strain put on the core processing subsystem's switch, which sees traffic from all other subsystems. Since the switching in the hybrid architecture is electronic, the bandwidth of the switching backplane is limited compared to optical switches. The strain on the core processing is further illustrated by the high worst-case latency in the 10x military experiment. This strain could be alleviated by using smaller switch modules and spreading the load at high traffic areas such as the core processing system. By contrast, expanding the number of switches significantly would increase inter-switch communication, and would likely not be desirable for a network implementation.

As described in Section 3, another major advantage of the hybrid network is the potential low costs of deployment on current and future platforms. Since Ethernet is used as the primary transport in our experiments, this approach allows the use of commodity components that are very mature and cost-efficient.

After the Clos and Hybrid architectures, the RSVP control protocol on both the ring and tree networks showed the next best results. The RSVP protocol provided average latencies around 140 and 230 us for the military and commercial baseline configurations respectively. The performance of the RSVP protocol suffered greatly when the traffic rates were increased. These numbers are well above the Clos and hybrid averages, but much better than the TDMA averages. Even though the topologies used with the RSVP protocol do not offer the bandwidth capabilities of the Clos, they do provide bandwidth potentials comparable to the Hybrid architecture, and thus this limit is not the primary cause for decreased performance. Instead, the long amount of time required to complete arbitration is the major drawback. Arbitration is significantly slowed by the TDMA used within the control wavelength. The slow arbitration process can cause major delays when several nodes need to send to the same node at once, causing packets to back up when bursts occur. Using a faster form of arbitration, such as using electronic control channels as in OSMOSIS, would greatly increase the performance potential of those architectures. This method is almost certainly the most attractive way to attempt to achieve the performance of the Clos architecture with the fault-tolerance of rings.

The TDMA protocol showed the worst performance, and the latencies grew super-linearly as the traffic rates were increased. The worst-case latencies exceeded a full second when both configurations' traffic levels were increased. This outcome can be largely attributed to the inefficiency of a static TDM-based protocol for a packet-switched network. Each node has access to only a fraction of the bandwidth that each node can receive. Thus, even when there is only one node sending to a destination, it only utilizes a fraction of the full bandwidth the destination node can receive. As the number of nodes increases, smaller fractions of bandwidth must be pre-allocated, leading to highly inefficient performance, especially in the presence of bursty traffic. Using more complex strategies such as statistical TDMA can increase the efficiency of this protocol, but not the several orders of magnitude required to be comparable with the Hybrid or Clos architectures. Thus TDMA is not an attractive basis for designing a powerful and flexible avionics network.

6. CONCLUSIONS

To arrive at an optimized design for the "irresistible network" that the military avionics community demands, numerous designs from disparate paradigms must be analyzed. A preliminary set of potential designs have been presented and described in this paper. Those designs were evaluated and compared via simulative experiments. The results of those experiments revealed important insights for designing a WDM avionics network.

While WDM optical networking technology provides a powerful and effective solution for some networking applications, a purely passive optical approach to a large packet-switched LAN for avionics may not be an ideal approach. Purely optical approaches for dynamically establishing lightpaths and avoiding collisions between large numbers of nodes on shared optical mediums thus far leads to highly inefficient control protocols, and disappointing performances. This outcome was evidenced by the poor performance results observed with our ring-ring and optical tree architectures, each using two different control protocols. For such architectures to be used in avionics LANs, significantly more efficient control protocols will be required.

Without the functionality of buffering and active switching that electronic networks can provide, a scalable and flexible LAN with high-performance is difficult to achieve. In contrast, architectures that use electronic components to provide additional network functionality, and thus leverage the strengths of both electronic and optical network components, seem to offer more promise for providing the flexible and powerful network solution the avionics community is currently seeking. The Clos and Hybrid switched architectures proposed here both fall into this category, whose performance results were very promising. Unfortunately, these implementations lead to complex wiring demands, and require costly solutions to achieve the desired levels of fault-tolerance.

7. REFERENCES

[1] Avionics Fiber-Optics and Photonics (AVFOP) Conference, MIT Lincoln Lab, Minneapolis, MN, Sep. 20-22, 2005.

[2] Gardner, R., et al., "High-Performance Photonic Avionics Networking Using WDM," MILCOM 1999, Volume: 2, 31 Oct.- 3 Nov. 1999, pp. 958-962.

[3] Gerla, M., Kova, M., and Bannister, J., "Optical Tree Topologies: Access Control and Wavelength Assignment," Computer Networks and ISDN Systems Journal, Vol. 26, No. 6-8, pp. 965-983, 1994.

[4] Hemenway, R., and Grzybowski, R., "Optical-Packet-Switched Interconnect for Supercomputer Applications," Journal of Optical Networking, Vol. 3, No. 12, pp. 900, Dec. 2004.

[5] Matisse Networks, www.matissenetworks.com

[6] Spitzer, Carl, "The Avionics Handbook," CRC Press LLC, Boca Raton, Florida, 2001.

[7] Clos, Charles, "A Study of Non-Blocking Switching Networks," *Bell System Technical Journal*, March 1953, pp.406-424.