THE APS CONTROL SYSTEM NETWORK UPGRADE*

K. V. Sidorowicz, D. Leibfritz, W. P. McDowell
Advanced Photon Source, Argonne National Laboratory, Argonne Illinois, USA

Abstract

When it was installed, the Advanced Photon Source (APS) control system network was at the state-of-the-art. Different aspects of the system have been reported at previous meetings [1,2]. As loads on the controls network have increased due to newer and faster workstations and front-end computers, we have found performance of the system declining and have implemented an upgraded network.   There have been dramatic advances in networking hardware in the last several years. The upgraded APS controls network replaces the original FDDI backbone and shared Ethernet hubs with redundant gigabit uplinks and fully switched 10/100 Ethernet switches with backplane fabrics in excess of 20 Gbits/s (Gbps).   The central collapsed backbone FDDI concentrator has been replaced with a Gigabit Ethernet switch with greater than 30 Gbps backplane fabric. Full redundancy of the system has been maintained. This paper will discuss this upgrade and include performance data and performance comparisons with the original network.

1 INTRODUCTION

The APS accelerator control system has been implemented using the Experimental Physics and Industrial Control System (EPICS) software tool kit. At the APS, the control room operator interfaces (OPIs) are Sun Microsystems Unix workstations running Solaris 2.6 and CDE, an X-windows graphical user interface. To facilitate accelerator system troubleshooting, devices capable of running X-windows displays, such as X-terminals, Unix and Linux workstations, or PCs with X-windows emulation, may be placed at any physical location in the facility. A person using any one of these devices has the ability to generate and alter control displays and to access applications, interactive control programs, custom code, and other tools. The front-end computer or input/output controller (IOC) provides direct control to the input/output interfaces for each accelerator subsystem. At APS the standard crate uses the VME or VXI bus standard, a Motorola 68040/60 processor, Ethernet-based network communications, and a variety of signal and subnetwork interfaces. The 68040/60 processor provides the crate with the intelligence to allow it to run its software autonomously with respect to all

other devices in the system. The EPICS core software running in the crate hides hardware dependencies from the high-level software running on the workstation. There are approximately 175 IOCs used in the APS accelerator control system. A real-time operating system, VxWorks, is run in the crate central processing unit (CPU) to provide the basis for the real-time control.

EPICS uses the TCP/IP networking protocol, a commercial standard supported by all network hardware vendors. The TCP/IP implementation is independent of the particular network medium selected to implement the network. APS uses 10 Mb, 100 Mb, and 1 Gb Ethernet.

2 ORIGINAL NETWORK OVERVIEW

A diagram of the original network has been shown at a previous meeting [2]. The network used optical fiber to connect satellite network hubs to a collapsed backbone FDDI concentrator system. All the hubs were dual attached to the concentrator using a star configuration. A router was used to isolate the APS network functions. The network lent itself to being divided along geographical lines, and thus the network was divided into segments that included the control system, the CAT beamlines and laboratory office modules (LOMs), the central laboratory office building (CLO), and the Argonne Guest House (AGH).

When it was designed and installed, the network cabling plant allowed an upgrade path to Fast Ethernet, Gigabit Ethernet, or ATM technology. There were ten remote hubs in the controls network distributed throughout the accelerator facility to provide local Ethernet connections to all network devices. This system also used different physical paths for the fibers between the remote hubs and the concentrators in order to provide protection against common mode physical damage. All of the control system IOCs were redundantly connected to the hubs using fiber Ethernet so that they too could be reconfigured if required. This arrangement also allowed a hub to be serviced or to fail without causing the IOC to lose communication with the network.

3   THE UPGRADE

Figure 1 presents an overview of the controls portion of the upgraded system. There are 12 dual-screen Sun workstations in the main control room. Seven of these are used by operations to control various aspects of the facility and five of the workstations are available for use by engineering groups and physicists. There are two file servers in the control system that are dual attached to

each concentrator. Six workstations boot from each server to prevent the control system from becoming disabled in the event of a server crash. The original Cisco 7513 router has been replaced by a Cisco 8540. The Catalyst 8540 switch router is a modular Layer 2 and Layer 3 switch router that provides wire-speed Ethernet routing and switching services. The system has a 13-slot chassis that supports 48 port modules of 10/100 Fast Ethernet connectivity, or 2 port modules of Gigabit Ethernet. Eight user slots are available for user interface cards. The system can be deployed as a high-speed switch router for campus or enterprise backbones.
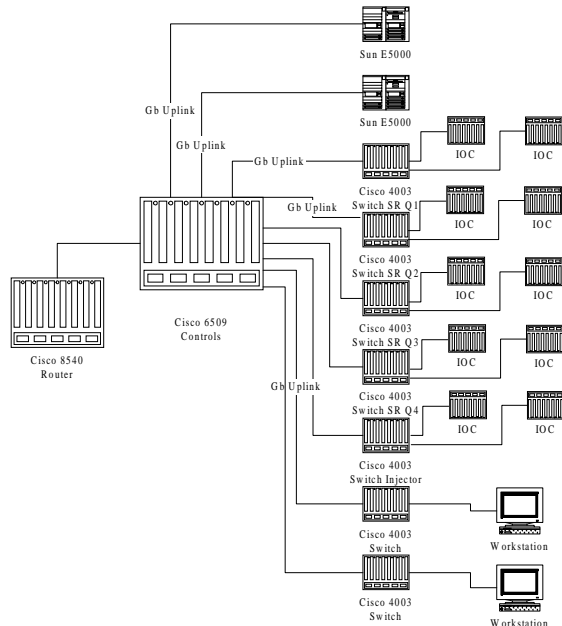

Figure 1: The APS Controls Network

The key features of the Catalyst 8540 router include wire-speed Layer 3 IP, IP multicast routing, and forwarding across Ethernet and Fast EtherChannel (FEC) interfaces. The switch router also provides high quality of service (QoS) capabilities, including support for four queues per port and flow classification based on IP precedence bits.

The major improvements over the 7513 include the performance of the 40-Gbps nonblocking switching fabric scaling to 24 million pps compared with the two 1-Gbps data buses of the 7513 backplane. In addition, replacing the FDDI interfaces on the 7513 with the gigabit interfaces on the 8540 will greatly improve routing performance of the entire network.

APS has selected the Cisco 6509 nine-slot switch for internal distribution of gigabit network links. These high-performance, modular, frame-based switches support high-density Fast Ethernet and Gigabit Ethernet and have a 32-Gbps switching capacity. The Gigabit Ethernet switching module provides a high-performance gigabit switching backbone while the supervisor engine gigabit interfaces serve as uplinks to the Cisco 8540 router.

Sixteen port Gigabit Ethernet modules aggregate traffic from the high-density Ethernet 10/100-Mbits/s (Mbps) wiring closets.

The Cisco 6509 switches support fault tolerance and redundancy through the use of two hot-swappable supervisor engines and two fully redundant, hot-swappable, AC-input or DC-input, load-sharing power supplies. Each power supply has a separate power input. In addition, the fan assembly is hot-swappable and the backplane-mounted clock modules are redundant. All modules (including the supervisor engine if you have redundant supervisors), fans, and dual power supplies support hot swapping, so that modules can be added, replaced, or removed without interrupting the system power or causing other software or interfaces to shut down.

The gigabit uplinks from the 4003 switches are connected to the Cisco 6509. The 4003 switch provides intelligent Layer 2 services leveraging a 24-Gbps bandwidth architecture for 10/100/1000-Mbps Ethernet switching. Redundant features include two load-sharing, fault-tolerant AC power supplies, and a hot-swappable fan tray. Up to 96 10/100 Fast Ethernet ports or up to 36 Gigabit Ethernet ports can be installed into one chassis.

We have now replaced the FDDI and the network hubs with switched technology. All connections and equipment still allow fail-over to redundant paths and equipment. The same equipment and strategy have been followed for the network hub equipment in the CLO, the CAT areas, and the AGH buildings. Thus the computers installed in offices, labs, and residence rooms will use Category 5 wiring at 10- or 100-Mbps Ethernet rates, while the network equipment has been upgraded to Gigabit Ethernet on an incremental basis.

4 PROBLEMS

The initial task of removing the shared Ethernet hubs and replacing these hubs with switches required significant effort. All the network fiber cables from each hub, approximately 80, had to be removed and labeled before the new switch was physically installed. In addition, media converters and chassis for these converters had to be installed because the ports on the Cisco 4003 only support copper interfaces. Even though all possible preliminary work was done before a network shutdown, it still required nearly a day to install one switch. Copper patch cables were added between the switch port and media converter for every network connection that required fiber. After the switch was connected, the switch itself was checked to ensure all the interfaces were up and operating properly without any errors.

The switches ran correctly for nearly a week until they started rebooting randomly. Cisco personnel suggested we upgrade the firmware on the switches to resolve this problem. While waiting for the next maintenance period,

the switches were upgraded with new firmware. This resolved the problem of the switches rebooting randomly, but then another problem appeared. Two switches in one location were having channeling problems with the uplinks. Initially it appeared to be a Cisco problem or perhaps a problem with the distance to this location. After duplicating this configuration with spare switches, we determined the problem originated from the media converter chassis itself, which was not operating properly. After replacing this bad media converter chassis and replacing some bad media converter modules, our channeling problems were resolved. Our Cisco switches and media converters have not faltered since these initial problems were resolved.

We had hoped to report the performance of the gigabit uplinks at this meeting. Cisco scheduled delivery of the gigabit cards for June 1999. The last communication from Cisco indicated that the cards were due to be shipped on October 8, 1999, the day this meeting ends.

5 PERFORMANCE

Since we were unable to completed the network upgrade to Gigabit Ethernet due to a delay in receiving the products from the vendor, we decided to perform the upgrade in two phases. The phase I upgrade consists of 10/100 switches in the accelerator and a Fast Ethernet switch for the backbone. In phase II we will upgrade the backbone switch to Gigabit Ethernet. The phase I backbone Fast Ethernet switch has been installed in the computer room with Fast Ethernet switches installed in the linac, PAR, booster, rf, and one quadrant of the storage ring. For uplinks we channeled two Fast Ethernet full-duplex ports from the remote accelerator switches back to the computer room. This provides a maximum theoretical speed of 400 Mbps. We ran the same software tests as were run when we installed the first-generation control system network. To perform this test, a database is loaded into an IOC and an MEDM application, which allows selection of the number of events per second, and is executed on a workstation. To monitor CPU utilization, top, a freeware performance monitor is run on the workstation and spy, a VxWorks utility, is run on the IOC. In 1995, the hardware consisted of Ethernet hubs, FDDI backbone, a Sun Sparcstation 20, and Motorola MV167 (68040) IOCs. Table 1 represents data from those tests. Maximum events/second were 5600 with 2% idle time on the IOC.

The hardware today consists of Ethernet/Fast Ethernet remote switches, a Fast Ethernet backbone switch, a Sun Ultra 2 workstation for OPI, and Motorola MV167 (68040) and MV172 (68060) IOCs. Using MEDM version 2.3.5a and EPICS version 3.13.1.1, Tables 2 and 3 show an improvement of 200 in events/second for the MV167. Maximum events/second for the MV172 is 9762 and maximum network utilization was 41%.

Since the second-generation phase II network is designed for the latest Ultra workstations and Motorola Power PC IOCs, a Power PC test stand was setup for testing. The Power PC IOCs support full-duplex Fast Ethernet. Data from these tests is given in Table 4. The limiting factor here appears to be the Ultra 2 workstations. Future testing will include the latest in hardware from Sun.

Table 1: 167 Shared 10 Mb

MV167		
Request Events/S	Actual Events/S	IOC Idle
1000	1000	
2000	2000	
3000	2950	
4000	3800	
5000	4800	
6000	5300	24%
7000	5500	2%
8000	5600	

Table 2: 167 Switched 10 Mb

MV167			
Request Events/S	Actual Events/S	IOC Idle	WS Idle
1000	1000	91%	91%
2000	2000	79%	85%
3000	3000	76%	72%
4000	4000	70%	65%
5000	5000	59%	60%
6000	5800	46%	80%
7000	5600	43%	81%
8000	5300	30%	80%

Table 3: 172 Switched 10Mb

MV172			
Request Events/S	Actual Events/S	IOC Idle	WS Idle
1000	1000	87%	87%
2000	2000	85%	81%
3000	3000	82%	70%
4000	4000	81%	62%
5000	5000	79%	50%
6000	6000	67%	43%
7000	7000	64%	38%
8000	8000	61%	39%
9000	9000	54%	39%
10000	9400	28%	58%
11000	9762	27%	60%

Table 4: 2700 Switched 100Mb

MV2700			
Request Events/S	Actual Events/S	IOC Idle	WS Idle
1000	1000	99%	99%
2000	2000	97%	93%
3000	3000	96%	89%
4000	4000	96%	73%
5000	5000	96%	63%
6000	6000	96%	58%
7000	7000	96%	50%
8000	8000	96%	44%
9000	9000	95%	35%
10000	10000	95%	34%
11000	10240	88%	32%
12000	10871	87%	16%

6 CONCLUSIONS

The network upgrade has improved the performance of the APS controls network by a factor of at least 100. Although we are not able to report the final performance figures because of vendor delivery problems, we will post final figures on our Web site.

7 REFERENCES

[1] K. V. Sidorowicz and W. P. McDowell, "The APS Control System Network," ICALEPCS'95, Chicago, October 1995 (CD-ROM).

[2] W. P. McDowell and K. V. Sidorowicz, "An Accelerator Controls Networks Designed for Reliability and Flexibility," ICALEPCS'97 Beijing, November 1997, pp. 302-304.