WE-P19

UPGRADE OF THE HERA/PETRA/DORIS ACCELERATOR CONTROL NETWORK AT DESY

Steve Herb, Rüdiger Schmitz, Rolf Schröder, Harald Schulz , DESY, Hamburg, Germany

Abstract

The control system at DESY for the HERA, PETRA, and DORIS accelerators, and the associated pre-accelerators, includes about 500 computers (mostly PCs) using the IP and IPX protocols over Ethernet for real-time communications. The control system network has recently undergone a major upgrade, with the BNC-based cabling and ~10 year old Router/repeater technology having been replaced by a structured cabling infrastructure using modern switches and routers from Cisco[1]. The new system is built in conformance with, and is fully integrated into the DESY site network, but can also operate in stand-alone mode. Using the same structure as the site network has brought major benefits, as well as some understanding of the how the requirements for controls are different. We discuss this, as well as some of the tools which were (or would have been) useful for the commissioning and maintenance

1 INTRODUCTION

When starting the upgrade process we confronted a network that had grown over a period of about 10 years. It reflected changing technologies, our increasing understanding of the Ethernet and the different approaches of the control systems of on the one hand, the HERA accelerators and on the other hand, the pre-accelerators LINAC2 , LINAC3, DESY2, DESY3 and the storage rings PETRA2 and DORIS3 [2]. One difference is that most of the computers for the smaller rings are centrally located in dedicated racks, with the remote hardware connected by fieldbus, while many of the computers for the large HERA rings are located in service buildings several kilometers from the control room. A second difference is that the systems for the smaller machines make extensive use of IPX broadcasting within machine-specific subnets, while the HERA controls use mostly UDP for machine to machine communications.

In either case, the interruption of network connections results in loss of ability to change machine parameters, but not in threats to the safety of the machine components. Under stationary conditions, the beams continue to circulate, but no diagnostic data is available.

2 SUBNETS AND OLD TOPOLOGY

The ~500 hosts are distributed over 12 subnets, each representing a class of tasks in the total control system (Table 1). Each subnet used 10Base2 BNC cabling with repeaters as necessary to permit direct connections between all members. This structure provided modularity and some independence from central devices.

Table 1: Subnets in the Control System

Task	No of hosts
Hosts for LINAC2, DESY2	47
Hosts for LINAC3, DESY3	47
Hosts for DORIS3 and beam lines	51
Hosts for PETRA2 and beam lines	54
HERA-Consoles	34
Device Servers for HERA I	125
Device Servers for HERA II	23
Hosts for the timing/ trigger system	15
Common consoles, services	21
Central services, Archive Serves	22
Developers, central servers I	65
Developers, central servers II	102

Routing between the pre-accelerator subnets was accomplished using Netware file servers, and at a higher level a dedicated 12 port router (ALANTEC) serviced these together with the local HERA subnets and the fiber optics connections from the HERA halls. One port was used to connect to the DESY site network, and filters were used to control traffic into the control network. This topology proved its worth on several occasions when the control system network continued to function during massive disruptions of the site network.

3 UPGRADE ISSUES

There were a number of motivations for an upgrade:

- Provide more bandwidth to applications
- Replace the aging ALANTEC router by up-to-date technology
- Eliminate the router functionality of the Netware file servers
- Improve transparency and documentation for the network
- Transfer part of the administration job to the DESY network group (NOC)

There were also some constraints which influenced the decisions of what to do and buy:

- Conform to the NOC equipment standards
- Retain the modular structure and some sort of independence for the subnets
- Minimize reconfiguration work for the hosts

- The initial switch from the old to the new network structure should not take longer than one day
- Costs should be reasonable

3.1 The new components

The components we chose were:

- As core a Cisco Catalyst 6509 Switch with
  - Supervisor engine and routing capability
  - 24 port 10Mb FO
  - 24 port 100Mb FX
  - 96 10/100Mb RJE
  - 2 Gigabit uplinks to the DESY core routers
- 15 Cisco 2924XL switches at remote locations (such as the HERA halls) connected by optical cables to the core
- Inexpensive non-managed switches (e.g. AT 708) which connect clusters such as the console PCs to one 100baseT port of the core switch

3.2 VLANs

The managed switches support the VLAN (virtual LAN) technology by which subnets may be defined logically in the routers and switches, so that groups of devices can communicate as if they were on a single wire when in fact they are located on a number of different LAN segments. It was therefore natural to implement our subnets as VLANs. This can simplify cabling, since a single trunk line to a distant hall can serve several subnets. The disadvantage is that stable communications on the VLAN depend on the correct operation of several routers and many switches.

4 MIGRATION

We decided for reasons of manpower and minimizing disruption of service on a strategy of migration, first replacing the central router and switch, and then gradually converting the BNC strings to structured cabling. The steps of the migration are listed in Table 2. The router replacement was performed in July 2001 and the conversion to structured cabling is still in progress.

Table 2: Migration steps

Step	Description
1	Prepare cabling infrastructure and site preparation for the new core switch
2	Replace the ALANTEC by moving the subnets at the 12 ports one by one to the Cisco core switch, resulting in a half day of network interruptions
3	Directly connect important servers to the switch, remove 10BaseT repeaters during normal machine operation
4	Provide an administrative framework for the management of the 2924XL switches
5	Install the 2924XL switches to connect the distant locations (ongoing)..
6	Replace local BNC strings with non-managed switches (ongoing).

4.1 Areas of responsibility

The operation of the Cisco Catalyst including maintenance and supplying spares is the responsibility of the NOC of DESY. The proper function of the network connected below the Catalyst is in the responsibility of the controls group, but with assistance from NOC.

4.2 Security and reliability aspects

The access-lists of the ALANTEC router have not been re-established in the new environment. The security concept must be reviewed.

The hardware in the central switch can be replaced easily, but this produces connection problems in all control VLANs.

We still make heavy use of the IPX-Protocol, but the NOC has limited experience with IPX, compared to IP.

For the IP-routing there exist several fall-back routers at NOC, but this is not the case for IPX.

4.3 Nomenclature and Documentation

The new devices provide a description field for each port. We use it to document the device connected and its location. To do this effectively and unambiguously we need for example a room number scheme, rack naming, etc. for the whole site, a DESY-wide convention to fill in the field. There is an intrinsic limitation because the field length is limited and not the same on all devices.

4.4 Integration into the DESY network

After the migration we can offer our customers within the control system whichever VLAN they need, and we can make our VLANs available on any switch of the DESY site network, as required.

The maintenance is NOCs responsibility, but it affects all control VLANs. Network maintenance must now be coordinated with the accelerator schedules as well as the requirements of the NOC and the DESY site users.

5 EXPERIENCE

The migration to the new network structure has certainly been successful, but we have also spent more time than we would have liked on related problems. We try here to summarize some of these experiences, especially as they relate to accelerator operations.

5.1 Router/Switch Hardware Failures

On two occasions the supervisor engine failed, resulting in loss of connectivity between machine subnets for a total of about 1/2 day. Because most of the machine subnets were *not* structured, connectivity within these subnets was maintained.

5.2 Switch Port Shutdown

The Cisco switches shut down a port for 30 seconds when too many collisions or other packet errors occur. If the port is servicing a BNC string with many computers, the data from all computers on the string will be lost. We lost perhaps a total of one shift of HERA beam time due to the disappearance of crucial diagnostic information during several accelerator ramps. The problem went away when the structured cabling was installed at this level.

5.3 Site Router Reconfiguration

After a power outage at another location on the DESY site, we experienced connectivity problems for about 15 minutes, apparently caused by high loads in routers which were recalculating routes for rebooted switches.

5.4 IP Address Confusion

For reasons not understood, PCs running the old systems DOS and Win3.11 sometimes fail to come on net, with the message 'IP already in use', although that is not the case. This has had considerable nuisance value.

5.5 Loss of Connections to 2924XLs

On several occasions, the last for 3 hours, the subnets including hosts on the 2924XL switches were lamed by unexplained heavy traffic on the trunks to the 2924XL switches. This stopped operation of HERA and DORIS.

6 APPRAISAL

We present our appraisal after one year of experience.
Benefits:

- 100Mb/sec available at each host if necessary
- No work with spares and updates of core switch
- VLAN technology works
- More transparency of the network, online description via WWW available
- The switches support the location of devices by their hardware address (MAC address). Find out how many devices are active on the ports is easily done. In combination with the IP-Database (QIP from Lucent Technologies) one can determine their names etc.

Problems and Drawbacks:

- The core switch represents a single point of failure (and it has failed twice)
- No direct access of configuring the core switch due to the responsibility regulations. This causes little delay and needs more administration.
- Integration into the DESY site network has side-effects (we are not so well isolated as before)
- Mixing of structured and BNC cabling has resulted in unexpected problems during the migration.
- We chose the 2924XL switches because they seemed suited to our needs and are standard Cisco products, although they were not being used by the NOC group. They do not integrate well into the framework of the other switches (they use different management and operating systems), and NOC has less ability to help us in the trouble-shooting. We should have purchased models identical to those already in use
- The network technology has become more complex, and we often feel somewhat helpless when trouble-shooting is required. Tools such as LANalyzer [3] which were effective in the old topology cannot see very much at a switched port, and we do not have adequate replacements
- Maintenance slots are rare since even when the accelerators are down, the network may be required for vacuum work and the personnel access system

What we like to have:

- One graphical user interface (GUI) which supplies:
  - A living picture of the network:
    - The utilisation of the switches
    - Error conditions at single ports
    - The active addresses
  - Listing of the actual configuration of port and VLAN mappings
  - Search for MAC-Addresses and VLANs
- Integration of the different administration tools, and accompanying automatic processes to check the integrity of the network databases

7 CONCLUSION

- Modern network technology is powerful but complex, and requires professional attention and tools. We underestimated these requirements
- Saving money on hardware may be a mistake if it increases the system complexity
- Mixing old and new technology brings surprises; migration is an option, but it may not be painless
- Hardware can fail: don't build a network for several accelerators with a single point of failure
- Introducing new technology and having no trained people is a mistake as well.
- Future (as present) accelerators should not assume 100% reliability for network connections

8 REFERENCES

[1] © Cisco Systems Inc.. http://www.cisco.com

[2] Rüdiger Schmitz, "A Control System for the DESY Accelerator Chains", proceedings of PCaPAC'99, http://conference.kek.jp/PCaPAC99

[3] Network Analyzer Program for Windows by Novell Corporation, Provo, Utah, USA.