Networking and HEP Data Collection

Pekka K. Sinervo
University of Toronto / IPP

1. The Anatomy of a FNAL HEP Experiment
2. The CDF Experiment
3. HEP Dataflow
4. Local Area Networks at FNAL
5. ESNET and INTERNET
6. The Future for HEP Networking
7. Conclusions

Fermilab Affiliates
September 1994

Anatomy of an FNAL HEP Experiment

High Energy Physics experiments can only be performed at a small number of international facilities

- experiments require large groups
- generate large amounts of data
- require massive software development efforts
- typically take years to analyze

A typical experiment at FNAL

- performed by a several hundred physicists from several dozen institutions
- relies heavily on physicists working remotely
- requires diverse computing facilities
  - large amounts of CPU and data storage
  - high-speed data access
  - support for several operating systems

Networking has become essential infrastructure

The CDF Experiment

To give a concrete example, the CDF Collaboration

- consists of 430 Ph.D. physicists and grad students
- drawn from 34 institutions
- spread across US, Italy, Japan, Canada and Taiwan

Collaboration communicates/interacts in several ways

- bi-monthly standing meetings
- collaboration meetings every 3 months
- computer bulletin boards, NCSA Mosaic, WWW
- e-mail

Almost all communication has an electronic component

How We Use Networks

Can break down uses into a few categories:

- Data and Software Distribution
- Information distribution
- Communication within collaboration
- Teleconferencing

Look at each in the CDF context to give some flavour

Data Flow in CDF

The scale of the dataflow problem:

- CDF detector records events at ~ 5 Hz
- each event consists of ~ 150 Kbytes of data
- recorded 16 million events (~ 2 Tbytes) in 1992
- subsamples of data analyzed at ~ 15 sites

Large software base needed to manipulate and examine data

- Code base is ~ 2 million lines of FORTRAN 77
- Developed to run on VMS and UNIX platforms
- Maintained by ~ 20 "distributed" physicists

Online Data Collection and Processing

Our use of high-speed networks starts at detector

- Use 100 Mbyte/s Ultranet hub to collect data
- Written to Exabyte tape

Data is processed using a "farm" of parallel processors

- Data is then split into smaller datasets
- Various physics groups work on these
- Often requires large amounts of reprocessing

Local Data Access

Data is accessed locally at FNAL using

1. VAX/VMS cluster ($\sim$ 500 MIPs and 150 Gb disk)
2. SGI UNIX multi-processor ($\sim$ 240 MIPs and 75 Gb disk)
3. VMS workstation cluster ($\sim$ 1000 MIPs and 200 Gb)
4. UNIX workstation cluster ($\sim$ 500 MIPs and 25 Gb)
5. Heirarchical storage system ($\sim$ 0.8 Tb)
6. About 140 Exabyte drives

Components are connected via an FDDI ring

- Try to keep “hot” datasets on disk
- Access with local servers
- Workstations are still on copper ethernet

Networking Offsite

FNAL is connected to Energy Sciences Network (ES-NET)

- Provides T1/T3 connections to other DOE facilities
- Connects us to rest of INTERNET
- Provides international links to other labs

Implications to University Users

Network topology is fairly convoluted For example, from Toronto ⇒ FNAL is 18 hops

```
 1  irides (128.100.78.6)  10 ms  0 ms  0 ms
 2  mcl3.gw (128.100.75.224)  10 ms  10 ms  0 ms
 3  utorgw.gw (128.100.96.19)  10 ms  10 ms  0 ms
 4  psp.on.canet.ca (192.68.55.1)  10 ms  10 ms  0 ms
 5  psp.ny.canet.ca (192.70.164.142)  20 ms  20 ms  10 ms
 6  ENSS133.CIT.CORNELL.EDU (192.35.82.101)  20 ms  20 ms  20
 7  t3-2.cnss48.Hartford.t3.ans.net (140.222.48.3)  20 ms  20
 8  t3-2.cnss32.New-York.t3.ans.net (140.222.32.3)  20 ms  30
 9  t3-1.cnss56.Washington-DC.t3.ans.net (140.222.56.2)  30 ms
10  mf-0.cnss58.Washington-DC.t3.ans.net (140.222.56.194)  30
11  t3-0.enss145.t3.ans.net (140.222.145.1)  30 ms  40 ms  30
12  umd-rt1.es.net (192.203.229.252)  130 ms  120 ms  120 ms
13  pppl-umd.es.net (134.55.6.34)  140 ms  1010 ms *
14  pppl3-e-pppl1.es.net (134.55.6.132)  40 ms  40 ms  100 ms
15  fnal-pppl-t3.es.net (134.55.21.5)  80 ms  80 ms  70 ms
16  r-s-frw.fnal.gov (198.151.133.1)  310 ms  70 ms  60 ms
17  r-s-hub.fnal.gov (131.225.57.200)  60 ms  120 ms  60 ms
18  fnalp.fnal.gov (131.225.109.2)  140 ms  140 ms  70 ms
```

Other Network Uses

Teleconferencing is now heavily used

- Make use of two dedicated facilities
- Experimenting with teleconferences over network

Document distribution

- Los Alamos has on-line preprint library
- Many groups distribute their own preprints via WWW

$\boxed{\text{Future Advances in HEP Networking}}$