Toward a Reliable Gigabit Network

- An Upgrade of the SPring-8 Network

T. Fukui, M. Ishii, R. Tanaka, T. Masuda, T. Ohata

SPring-8
Contents

- Introduction
- Network Overview
- Gigabit Network Performance Test
- Future Plan
- Conclusion
The SPring-8 control system is based on a distributed computing system.

- Using the ONC/RPC with TCP/IP
- Client/Server system
- Any of the equipments on the networks can be controlled from the central control room

- Storage Ring, Booster Synchrotron, Linac, New SUBARU, Insertion Device and Beamline were controlled from CCR
This talk is focus in the accelerator network

BL-USER-LAN with VLAN Technologies was presented on TUAP056
The design policy of the network is:
- Redundancy
- Good performance
- Simple and Flat for ease of management.

FDDI was the best choice at the beginning and we are going to the reliable gigabit ethernet.
Network Overview

At the beginning of the SPring-8

FDDI Backbone

A-D Zone Network Node

Router

Central Control Room Network Node

Linac & Synchrotron Network Node

10BASE-F/

Ethernet & FDDI

Switching HUB

Consoles

VME

Database Servers

FDDI DAS

10BASE-T

Storage Ring

Insertion Device

Beamline
Network Overview -cont.

- In the summer of 1998, the router of injector was replaced to switches.
- In the winter of 2000, the Layer-3 switches were installed due to the rapid increase of beamlines.
- Gigabit ethernet was installed to meet the demands of the higher speed DAQs of the accelerator controls.
  - 1 km long beamline & 27 m undulator.
Network Overview - cont.

- A-D Zone Network Node
  - Fast & Giga Ethernet Switching HUB
  - Layer 3 Switching HUB

- Linac & Synchrotron Network Node
  - Ethernet & FDDI Switching HUB
  - Gigabit Ethernet Switching HUB

- Data Acquisition System
  - PC
  - Database

- VME

- Storage Ring
- Insertion Device
- Beamline

- Central Control Room Network Node
  - New SUBARU
  - Layer 3 Switching HUB

- FDDI Backbone
- 10BASE-F/100BASE-FX
- 100BASE-TX
- 1000BASE-LX

- Consoles
Performance of gigabit ethernet

- Setup of test bench
- The bandwidth of the fast and gigabit ethernet
- Latency of gigabit ethernet switch
- Redundancy of link aggregation
- Latency degradation of link aggregation
- Quality of Service for accelerator control
Setup of test bench

- CoreBuilder 3500
- SuperStackII 9300
- SuperStackII 3900
- Gigabit NIC: 3C985B-SX
- Layer-3 Giga Switch
- Layer-2 Giga Switch
- Layer-2 Giga Switch
- PC with Linux
Performance of gigabit ethernet

- Setup of test bench
  - Netperf developed by Hewlett-Packard
    - Bandwidth, Latency test
    - Generate background traffic
  - ONC/RPC program to simulate SPring-8 control framework
The bandwidth of the fast and gigabit ethernet

- Measured result of gigabit ethernet is 4 times wider bandwidth than fast ethernet.
- There is no degradation of the bandwidth caused by the gigabit ethernet switch.
The bandwidth of the fast and gigabit ethernet

**Bandwidth of the fast and gigabit ethernet**

<table>
<thead>
<tr>
<th>Data Size (KB)</th>
<th>CB3500</th>
<th>CB3500+SS9300+SS3900</th>
<th>100BASE-TX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Transfer Rate (Mbps)**

- CB3500
- CB3500+SS9300+SS3900
- 100BASE-TX
Performance of gigabit ethernet - cont.

- **Latency of gigabit ethernet switch**
  - About 10μsec delay with 65bytes (64bytes header + 1byte data) transmission per hop for the gigabit ethernet switch
  - Fast ethernet is faster than gigabit ethernet at small data size
Latency of gigabit ethernet switch

Latency of the gigabit ethernet switch

Data Size (byte) vs. Response Time (µsec)

- Giga Direct
- Giga 1 Hop
- Giga 2 Hop
- Fast Direct
Latency of gigabit ethernet switch

- About 10\(\mu\)sec delay with 65bytes (64bytes header + 1byte data) transmission per hop for the gigabit ethernet switch
- Fast ethernet is faster than gigabit ethernet at small data size
Setup for link aggregation

CoreBuilder 3500

SuperStackII 9300

SuperStackII 3900

Two ports are aggregate as one logical port
Setup for redundant link test

CoreBuilder

SuperStackII 9300

SuperStackII

Link aggregation is used
As a redundant link
like DAS configuration
of FDDI
Redundancy of link aggregation

- The maximum time for reconnection was about 30 seconds and median value about 3 seconds.
- Much better than STP (about few minutes).

Latency degradation of link aggregation

- About 1μsec with no dependence on data size.
Why we need Quality of Service

- Total Bandwidth
  - High priority
  - Low priority or best effort
  - Drop
Test setup for Quality of Service

ONC/RPC with fast ethernet

Background traffic with gigabit ethernet

CoreBuilder3500 set priority tag for each packet

ONC/RPC with fast ethernet
Quality of Service for accelerator control

- VLAN based priority control on CoreBuilder3500

- High Priority class for ONC/RPC to simulate the accelerator control and Best Effort class for background traffic
Quality of Service for accelerator control

Delayed Packet Rate

- Without Priority Control
- With Priority Control

Percentage of Delayed Packet vs. Background Data Transfer Rate (Mbps)
Future Plan

- In our test, gigabit ethernet with link aggregation showed good performance and good reliability.
- Maintenance cost of the FDDI to increase.
- It will be difficult to buy the FDDI equipment from the market.
- We have a plan to replace the FDDI with gigabit ethernet.
Future Plan – cont’

- Storage Ring
- Insertion Device
- Beamline
- Data Acquisition System
- VME
- Central Control Room

Network Nodes:
- A-D Zone Network Node
- Linac & Synchrotron Network Node
- New SUBARU Network Node

Switching Hubs:
- 100BASE-FX
- Ethernet & FDDI
- VME
- Gigabit Ethernet
- Fast & Giga Ethernet
- 100BASE-TX
- 10BASE-T/100BASE-TX
- 1000BASE-LX
- 1000BASE-LX with Link Aggregation

Connections:
- 100BASE-TX
- Consoles
Conclusions

- Upgrading the SPring-8 network.
- The latency was about 10μsec per hop.
- The link aggregation can be used as a redundant link and 1 μsec increase was measured.
- Priority control showed a good capability to maintain the quality of service.
- We have a plan to move a high performance and reliable network with the gigabit ethernet.