Constructing High Availability Network Systems

Yasushi Inoguchi, Mikifumi Shikida, Yasuo Tan and Teruo Matsuzawa
Dec. 6, 2001
IS-RR-2001-025

Center for Information Science
Japan Advanced Institute of Science and Technology
Asahidai 1–1, Tatsunokuchi
Nomi, Ishikawa, 923-1292, JAPAN
inoguchi@jaist.ac.jp

ISSN 0918-7553

Abstract

Recently scale of LAN systems is rapidly increasing, and integrated network systems are required to improve usability and reduce maintenance cost. However there are many problems to construct the integrated networks such as network design, network connection for file servers and designing application servers.

This paper addresses schemes for integrated network systems based on a real integrated network in our university. Network designing policy, routing schemes for central huge file servers and high availability schemes for application servers are discussed.

Our proposal for network design is construction with two different generation network facilities to avoid hardware and software failure. For file servers which has many network interface to improve bandwidth, many routing algorithms are standardized to select network interface for packets from file servers, but there are no effective implementation for client which throw packets toward file servers. We propose network interface selecting method that looks up optimal interface in NIS maps and set as "automount" arguments.

We implement these approaches on our network and evaluate its efficacy under real network system operation. In consequence, routing schemes for file servers can reduce roundabout routing.

Introduction

Recently a number of servers and workstations on a network are rapidly increasing, and many sites have a large number of client machines and a lot of large scale high performance servers. However there are no effective implementations to manage these large scale systems.

This paper shows network management operation method for large scale and integrated systems based on experience on JAIST (Japan Advanced Institute of Science and Technology) network. The network in JAIST is one of a great scale network system and it include a number of workstations for each students or faculties, large file servers, some massively parallel computers, and other network servers.

Both availability and performance are required for network design because all users in JAIST access all workstations or servers through the network. Each client workstations have to select optimal routing to handle a great number of accesses in the network.

In this paper we propose approaches and techniques to control these system efficiently, and discuss its effectiveness. We implement these approaches in our site and shows result of operation with these techniques. We show detail of methods that are adopted by JAIST network about network design, routing for file servers.

Network Design Policy

Network Design Policy

Usually, many large scale networks is constructed by connecting small scale (laboratory scale) networks. However, network scale is rapidly increasing and it is required to maintain these networks as an integrated system. In this subsection we discuss advantages of the integrated network design policy compared to conventional one.

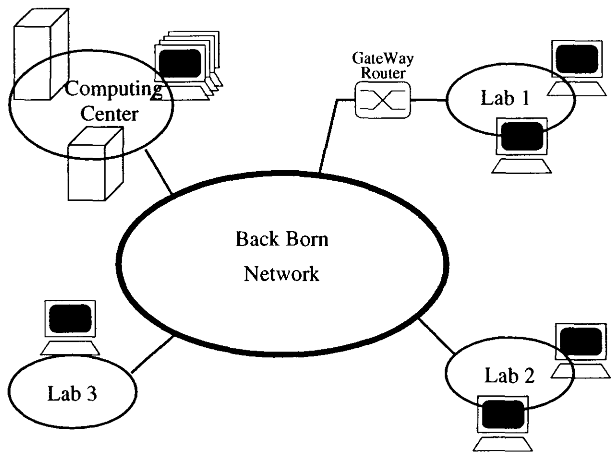

Figure 1: Concept of Conventional Network Design.

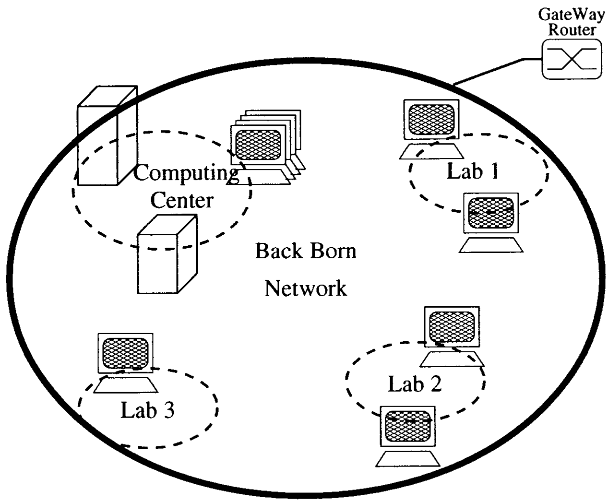

Figure 2: Concept of Integrated Network Design.

Conventional Network Design

Figure 1 shows a typical conventional network design. The conventional networks consist of a set of laboratory networks. Each laboratory has it's own workstations and servers in their local network, and university scale backbone network connects these laboratory networks each other. This network design has an advantage that each local network is adapted to each laboratory environment. For example, user can register new machine without sending a paper for requesting an IP address.

On the other hand, the conventional network design has some disadvantages. One is that the user environment such as a ".cshrc" and his home directory are different between user's workstation and other systems such as central computational servers. And if a laboratory doesn't have a good network manager, its network will be wasted. Moreover, each laboratory manages their network individually, thus total cost for network management is very expensive.

Integrated Network Design

To reduce these problems about network management, integrated network design is required. The figure 2 is an image of integrated network design.

It has no barriers between laboratories and a computational center. Integrated network design has many advantages. Most significant advantage is seamless information environment. All computers access a common user's home directory, thus users can access same data in their home directory without using ftp to transfer data between workstation and other servers. The sum of network management cost of this network design is cheap because a few staffs manage entire system efficiently using NIS, SNMP, or other network management tools. Thus even if a laboratory do not have a capable network manager, all users can use well-maintained systems. However, this network design requires a huge and high availability network and server systems. If important server or network is broken, no one can use entire system.

Constructing an Integrated Network

Overview of FRONTIER

In this subsection, we introduce our network system as a sample of a target network and show a construction strategy to keep both performance and availability.

Information environment of our university is called FRONTIER, and network system is called FRONTNET. FRONTNET is one of integrated network systems. In FRONTIER, anybody can use any system at anytime and anywhere in the university. Users in FRONTIER are not only professional of information science. Students belong to other department and office staffs also use same information environment. FRONTIER offers total information environment in our university. It contains networks, file servers, some kind of application servers, computational servers, and client workstations on user's desks.

Table 1 shows scale of FRONTIER. It includes 120 sub-nets, 5,000 hosts entry, and more than 1,000 accounts. All of them belong to only one NIS domain and one mail domain.

FRONTIER Network System

Figure 3 shows overview of FRONTNET. The network consist of two independent backbone networks drawn both sides in that figure. Center part of the figure is floor (laboratory) networks. These two backbone networks are completely independent systems. "In-House backbone" ("IP Switch" and "Router" in the figure) are high performance routers for each buildings. "Super backbone" ("HiPPI Switch" and "Ultra Network") connect these high performance routers with high bandwidth. Each file server has many network interfaces, and these interfaces are connected high performance routers directory. Thus most of traffics for file servers don't path through the super backbone.

Figure 4 shows details of floor networks. Each floor network consists of two segments. The floor network has some local servers for network management. These local servers are used as a NIS slave server, a printer spooling server, or a DNS slave server. Local servers have two network interfaces and connected each segment directory.

To support laboratory-scale NFS services or local computational services, some other local servers are disposed at floor networks. Floor networks also has laser printers, color scanner or some kind of tape drives for local backup.

Two Generation Network

Now we discuss about schemes to realize high performance and high availability networks. Most of high performance network products lack availability, on the other hand high availability routers can not exchange network packets quickly.

If we don't consider cost for the network, duplicate high performance network system can be constructed to increase network availability. The duplicated network system may avoid most hardware trouble. However, this approach is not effective for software or firmware trouble because both network facilities execute same software or firmware that has same bugs.

Table 1: Scale of FRONTIER (1998.)

Sub-nets	120
Hosts (NIS entries)	5,000
IS-Center offering hosts	1,650
Users	1,090
NIS domain	1
Mail domain	1
Staffs	faculty 3, technical officials 4

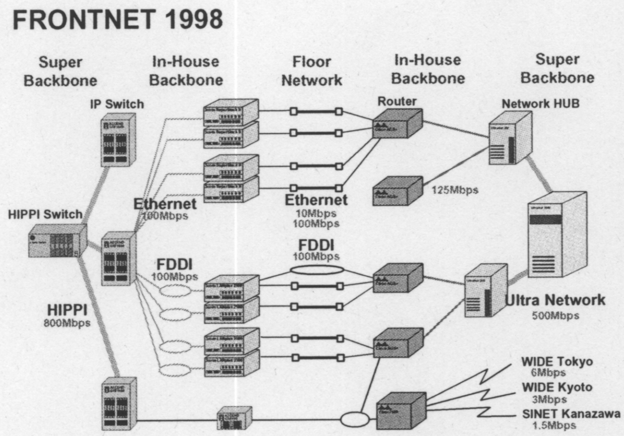

Figure 3: Overview of FRONTNET.

To improve both performance and availability, FRONTNET consist of two different generations of network facilities. Initially, we construct one network system. The first network is completed, we can start to supply network service, and at the same time we start to construct second generation network system. After the second network is accomplished, FRONTNET can provide duplicated network.

Generally, transferring speed of new routers is faster than old generation routers. Thus FRONTNET pass most of network traffics through new generation network at normal state, and the old generation network is used as a backup network. If the new generation facilities get into troubled, routings are automatically switched to the old network, and users can continue to use network without disconnection.

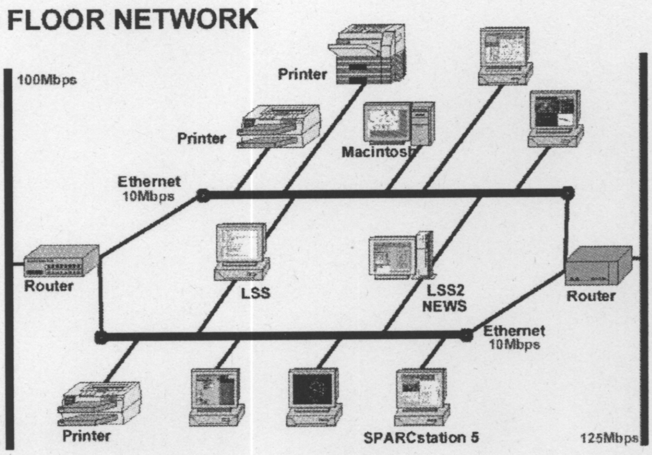

Figure 4: Floor network of FRONTNET.

Routing for File Servers

File Servers in Integrated Networks

This subsection discusses about network design and routing control for file servers. For file servers in large scale network systems, availability, network performance, and of course file serving performance are required.

As described in the previous section, FRONTNET has In-House backbone routers for each building and all users' home directories are stored in a file server. To improve network performance, every file server has network interfaces connected each backbone router directly. Although these routers are connected by high-speed network, it is better to reduce routing hops. Thus user's client workstation select an optimal network interface of the file server and throw packets toward the network interface. Problem is how to select the optimal network interface of the file server for each client workstations.

Figure 5 shows network connection for a file server. If a packet goes through optimal routing, a packet pass only one router (shown in figure 5(a)) and network delay is minimized. On the other hand, if routing is not optimized, a packet pass through roundabout routers and transfer delay will be increased (shown in figure 5(b)).

Routing Scheme for File Servers

There are good solutions for output packet from file servers. Many general routing protocols like RIP (Routing Information Protocol)[RFC1058, RFC2453], IGRP (Interior Gateway Routing Protocol), and OSPF (Open Shortest Path First)[RFC2328] are useful to select network interface of the file server. However, routing for input packets is difficult, because clients recognize different interfaces on same file server as different hosts.

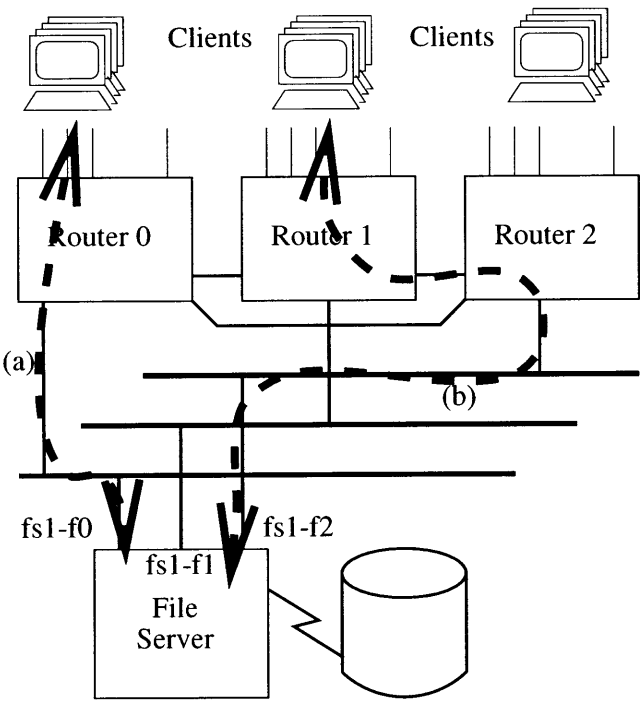

Figure 5: Routing for File Servers.

Thus we have to consider a scheme to select the optimal interface of file server for packets toward file servers.

Our goal is simple. We have to select the best network interface of the file server. If the routing is optimal, there is only one "In-House backbone router" between a client and the interface of the file server. User's home directories are stored in file servers and every file server has network interfaces connected to each backbone router directly. Thus the optimal interface is given as an argument of "mount" command for client workstations.

Figure 6 shows the scheme to select the optimal network interface for each client. As arrangements, we make some NIS database used by client workstations along with network connection map by manually, as shown in figure 6 - 1. "Servermap" NIS map relates Ethernet segments to optimal network interfaces of a file server, and "auto.home" NIS map relates home directories to network interface of file server. Thus client workstation can determine the optimal interface for user's home directory access by looking up in these NIS maps.

When user access one's home directory in a file server, client workstation tries to mount the home directory with optimal interface of the file server. This procedure is shown in figure 6 - 2. Client workstation looks up optimal network interface in "servermap" NIS database by its Ethernet segment.

1. Arrangements.
   (a) Let name of network interfaces of a file servers as "fs1-f0," "fs1-f1," "fs1-f2."
   (b) Make NIS map for optimal interface for each segment.
   ex) 150.65.108 fs1-f2
   (This map called as "servermap.")
   (c) Make NIS map for mounting directory and insert a variable.
   ex)home01 $FS1,fs1-f0:/home/home01
   (This map called as "auto.home.")
2. Clients behavior.
   (a) Get segment number where the client is in.
   ex) 150.65.108
   (b) Look up optimal interface name in the "servermap" NIS map by the segment number.
   ex) fs1-f2
   (c) Set the interface name to the variable in "auto.home" NIS map.
   ex) automount -D FS1=fs1-f2 -D FS2=...
   (d) Every accesses to home directories uses the optimal interface determined above.
   ex) Access for "/home/home01" uses "fs1-f2" interface.

Figure 6: Routing Scheme for Packet toward a File Server

Then the network interface is set to a variable which is in "automount" NIS database. The mounting operation is achieved automatically by using this optimal network interface like example in the figure.

Evaluation

We use this routing scheme in FRONTNET. Figure 7 shows traffic at a network interface of a file server, and figure 8 shows traffic between In-House backbone routers. These graphs shows that the average traffic between backbone routers is about half Mbit/s although the average traffic coming from or going to the file server is about 1.5MBit/s. Actual traffic for a file server between the backbone routers is less than 500Kbit/s because traffic between backbone routers include packets toward the file server, packets between clients, and packets to other servers such as WWW or mail server. Traffic between backbone routers might be much more if the scheme does not work, like figure 5 (b), thus it can be said that the selecting optimal interface scheme works effectively.

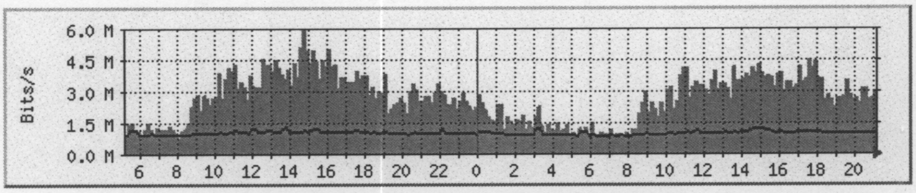

Figure 7: Traffic between a File Server and a Router.

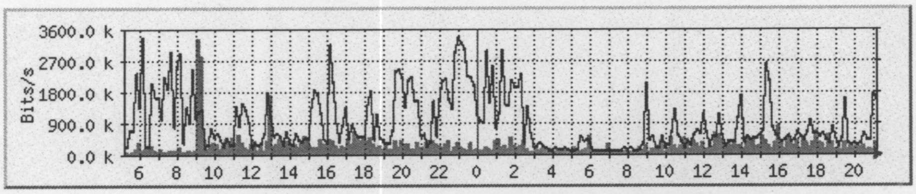

Figure 8: Network Traffic between Routers.

There are some competitive schemes for selecting optimal network interface of file servers. "Rotation of IP address by DNS[DNS]" is one candidate for traffic balancing. However it is difficult to use in In-House network, because rotation speed is not so faster and it might be bottleneck of access speed for file servers. "MAC address rewriting" is also a candidate but rewriting speed determine access speed for file servers too.

Conclusion

This paper addressed schemes to construct integrated network. Comparing integrated network design and conventional network design, it was shown that the integrated network design has many advantages. However it is difficult to realize a large scale integrated network.

Constructing network policy, a routing method for file servers in a large scale network are discussed. Duplicated network schemes are installed and evaluated in FRONTIER which is a sample of large scale integrated networks. The routing method for file servers effectively reduced traffic between In-House backbone routers. As the result, proposed schemes are helpful to construct the integrated network. Evaluation under controlled conditions is future work.

References

[DNS] Paul Albits and Cricket Liu, *DNS and BIND,* O'Reilly & Associates, Inc., 1992.

[MAIL] Bryan Costales and Eric Allman, *sendmail, Second Edition,* O'Reilly & Associates, Inc., 1997.

[NETB] Brendon W. Mills, *Where To Turn For Network Bandwidth,* Unix Review, Jluy 1995.

[NS] Charles Kaufman, Radia Perlman, and Michael Speciner, *Network Security: Private Communication in a Public World,* Prentice Hall, 1995.

[RFC1058] Hedrick, C., *Routing Information Protocol,* STD 34, RFC 1058, 1988

[RFC2328] John Moy, *OSPF Version 2,* RFC2328, 1998.

[RFC2453] Gary S. Malkin, *RIP Version 2,* RFC2453, 1998.

Affiliation

Yasushi Inoguchi, Mikio Shikida, Yasuo Tan and Teruo Matsuzawa:
Center for Information Science,
Japan Advanced Institute of Science and Technology.