CONSUMMATION OF AN OBSERVABLE NETWORK SYSTEM

T. Ohata, M. Ishii, T. Sakamoto, T. Sugimoto, JASRI/SPring-8, Hyogo, Japan

Abstract

Recent network system becomes more complex and larger because of virtual LAN and other virtualization technologies. Proliferation of a variety of network switches and routers makes a network system gigantic; hence, management of the misty network faces problems. This is the largest factor that deteriorates stable operation of a network system that should be robust and reliable. One of the promising solutions to keep a network system simple and understandable is introduction of the monitor tools that makes a network system visual and observable. We introduced the sFlow technology in addition to the traditional SNMP-based network node management (NNM) system. We could take statuses of network nodes by NNM such as hardware failure, and also we could grasp long perspective of network traffic at one view by the sFlow. In addition, an integrated log management system was introduced to collect all events on the whole network system. As a result, we could detect a trouble outbreak in real time even if a trouble occurred on the end point of the network, and could solve the problem promptly. We describe a way to achieve an observable network system to maintain stable network operation.

INTRODUCTION

The network system has progressed rapidly in resent years. Now days, the network system is one of the most important infrastructures that enable the continuous operation of large experimental facilities. However, it is true that the network system becomes very complex and the prospect worsened.

SPring-8 was opened to the public use in 1997, and currently we have an accelerator complex and about 50 beamlines for synchrotron radiation experiments. Indeed, SPring-8 has grown together with the progress of the network technology. The total number of experiment users exceeded 10,000 in a year. We have three networks [1] categorized by the purpose: a Control-LAN for accelerators and beamlines operation, a BL-USER-LAN for beamline experiments, and a public network for office work. The Control-LAN is the network that is strictly protected by firewall systems to realize stable operation. There is no route to access the Control-LAN directly from external networks. Recent years, several industrial networks that uses physical layer of Ethernet are proposed around the world. EtherNet/IP, PROFINET, and FL-net, etc grew popular. We selected the FL-net [2] as a device control network. The FL-net uses UPD broadcast on IP frame with token passing scheme. It is necessary to segregate it from other networks because network traffic of FL-net is hum. In the BL-USER-LAN, various experimental control and data acquisition systems are operated. The BL-USER-LAN is isolated from external networks to keep secure by using VLAN (virtual local area network) and NAT (network address translation) technologies. Additionally, we installed an IPS (intrusion protection system) that protects the network from threats such as computer worms. The public network provides services such as mail, web and wireless network system. As shown in figure 1, the networks are segregated by each other to avoid disturbance from other network troubles. It is a huge and sophisticated network where of each exceeds 1,000 nodes. And all of them are mission-critical

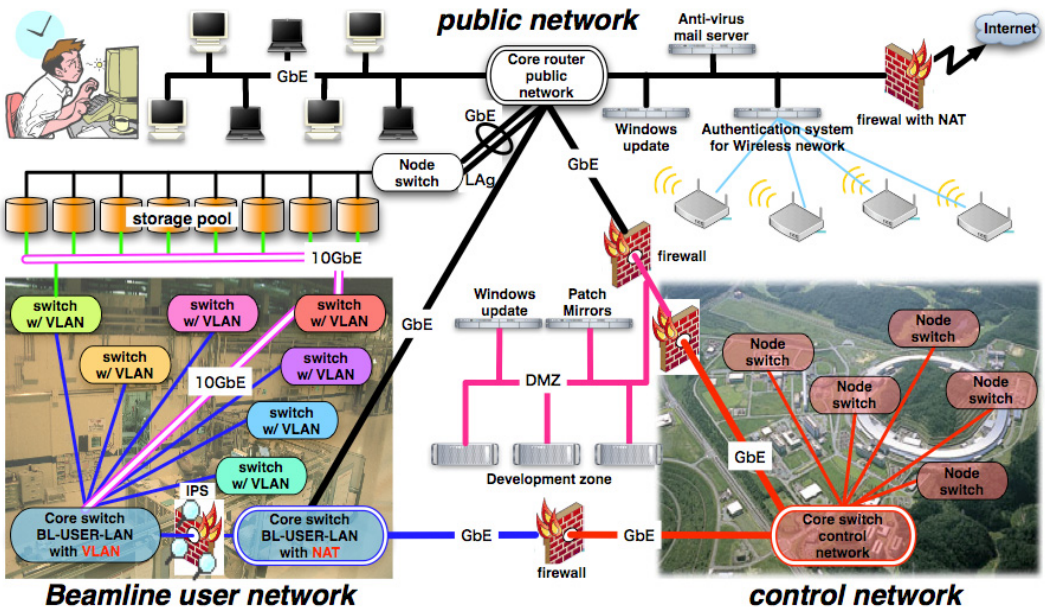

Figure 1: Overview of segregated networks in SPring-8.

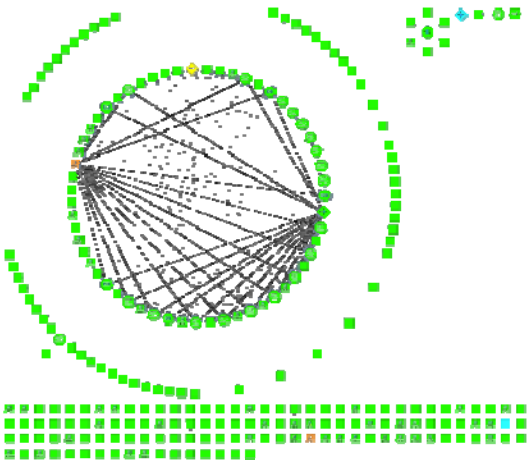

Figure 2: Logical network map of the Control-LAN.

network systems for facility.

These huge and complicated networks bring on a lot of petty problems such as slow network communication. These might sometimes growth serious network failure. The cause is a broadcast storm come from breakdown of the network equipment, or loops due to the operational error. Recent years, the risk of the loop connection increases extremely by the spread of Auto-MDIX (automatic medium-dependent interface crossover) that does the crossing connection in the network switch. For that case, we can only recognize a dysfunction of the network system. Although each network component is operated normally, the system of network doesn't function. It is like a brain death. If there is no appropriate monitoring system we cannot tackle network faults.

Therefore, we constructed network-monitoring system to solve network failures promptly. We describe the method of effectively observing the network at next section.

MIERUKA

MIERUKA is a visualization method known in the TOYOTA production system. It aims identifying problems and bringing them visible. We constructed the observable network system in accordance with the concept of MIERUKA. Following monitoring systems are especially important.

- Status monitoring system of whole network nodes by using network node manager (NNM)
- Integrated log server system for all network equipments (Syslog server)
- Flow monitoring system of network traffic (Flow monitor)

Details of each monitoring system are described as follows.

Network Node Manager

Network node manager (NNM) is a traditional network management system. NNM collects information from any network nodes that supported SNMP (simple network management protocol). By using NNM, overall status of the network system can be recognized such as number of nodes, node alive, connections between nodes, statistics of traffics, and so on. It is the best solution for all-round network management. However, because NNM doesn't distinguish between applications, sometimes it is difficult to specify the cause of the trouble.

We installed HP OpenVIEW NNM to the Control-LAN and the BL-USER-LAN. Figure 2 shows the network map of the logical network connection on the Control-LAN. There is a node shown by a red, yellow icon. It is understood that the node failure occurs in several nodes at one view.

Syslog Server

In the Control-LAN, we have a lot of high-functional network equipments, such as firewall, database, file server, and so on. The monitor of an internal state of these equipments is extremely important besides the state observed from the outside on the network by NNM. In general, syslog records detail information of the equipment to local storage or memory. We introduced the syslog server into the Control-LAN to integrate all log information from the network equipments. The syslog server stored log information that reaches 1TB in a day. Then we can trace behaviors of network and control system at time that trouble occurs by using log data on the syslog server. Figure 3 shows overview of the syslog server. To prevent the packet fall from a large amount of data, we adopted syslog-ng and a load-balancing scheme.

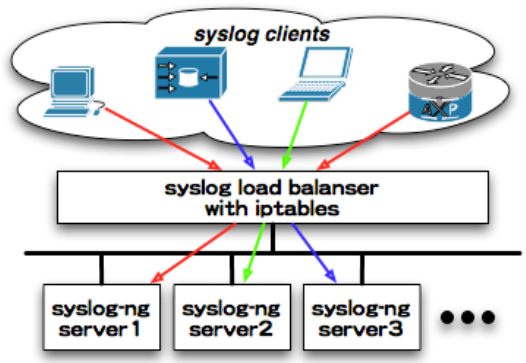

Figure 3: Overview of the Syslog environment.

Flow Monitor

The flow monitor concretely monitors and records what kind of network traffics where the packets are sent and where they are received. It makes network completely visible and gives information that enables rapid troubleshooting. The flow monitor supplements the weak point of NNM. For example NNM cannot distinguish a breaking of network cable and a loop failure. On the other hand, the flow monitor can clearly distinguish these. Moreover, it is also important to be able to identify FL-net traffic at the Control-LAN. Additionally, the flow monitor can be applied for the monitor of network security to defense against security threats at the public network.

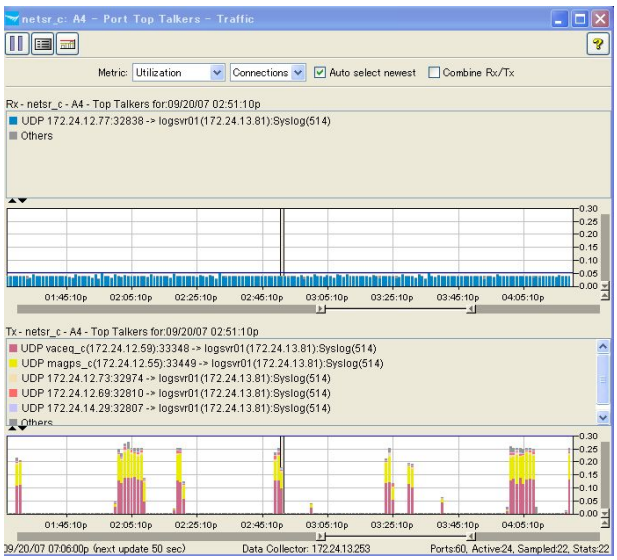

Figure 4: Example of the flow monitor. Bursty syslog packets are observed.

There are two major implementations of flow technology NetFlow and sFlow. NetFlow was developed by CISCO Systems. And sFlow was proposed by InMon and was supported by a lot of network vendors.

In the Control-LAN of SPring-8, HP ProCurve switch occupies large majority, we adopted ProCurve Manager Plus as an sFlow monitor. The monitor discriminates variety of ProCurve switch and integrates switch management function. Figure 4 shows an example of the flow monitor. Bursty syslog packets are observed.

EFFECTIVENESS OF MIERUKA

Two examples of effectiveness of MIERUKA are shown as follows.

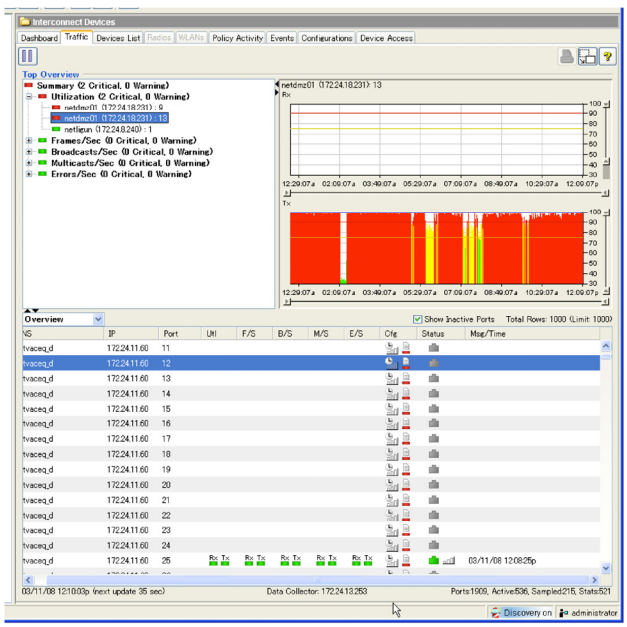

Figure 5: The overflow of the network bandwidth observed by ProCurve Manager Plus.

Investigation of Insufficient Network Speed

As shown in figure 5, the sFlow monitor has shown the limitation of the network bandwidth at the 100M Ethernet. A red graph shows that the bandwidth is used by 100% completely. We could quickly find and fix the problem, and the upgrade of the interfaces of the firewall was moved up. MIERUKA clarifies the problem of the network design, and becomes an important indicator that decides the solution.

Syslog Flooding

The network fault by the flooding packets that had been generated when the syslog server system was installed in the Control-LAN was found. The flooding was kicked up by UDP syslog communication that is connectionless protocol. Because the UDP packet doesn't leave ARP information of the transmission source in the network switch, it becomes a reason why an unnecessary flooding is generated. MIERUKA clarified that the flooding attacks the embedded equipment on the Control-LAN, and causes the trouble. And an important clue to the problem solving is shown.

CONCLUSION

We constructed the observable network system with the concept of MIERUKA to realize stable network operation. To identify problems and bringing them visible are most important to maintain huge and complex network system. In SPring-8, the networks for XFEL facilities will be newly added, and the network system becomes a huger, more complex. We will upgrade the system of MIERUKA, and construct a better network environment.

REFERENCES

[1] M. Ishii et al., "Upgrade of SPring-8 Beamline Network with VLAN Technology over Gigabit Ethernet", Proc. of ICALEPCS'01, SAN JOSE, USA.

[2] Specifications of Open PLC Network (OPCN), http://www.jema-net.or.jp/Japanese/hyojun/opcn_e/top-opcn.htm.