A Proposal for the Global Remote Participation System

Y. Nagayama, M. Emoto, Y. Kozaki, H. Nakanishi, S. Sudo, T. Yamamoto, K. Hiraki[1)], M. Inaba[1)], N. Tanida[1)] and S. Urushidani[2)]

National Institute for Fusion Science, Toki 509-5292, Japan
[1)] *Grad. School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan*
[2)] *National Institute of Informatics, Tokyo 101-8430, Japan*

(Received: 20 November 2009 / Accepted: 17 May 2010)

A global remote participation system is discussed. From the long experience of SNET that is the Japanese virtual laboratory system for fusion, the global remote participation would meet following problems: (1) conflict between the security and the convenience; (2) data transfer speed. Solutions in our proposal are as follows: (1) a virtual local area network (VLAN) to connect the remote sites and the experimental site using the virtual private network (VPN) technology; (2) a device to accelerator the data transfer in the world area network (WAN). A proof-of-principle test of the WAN acceleration technology has been carried out between ITER site and NIFS. This test demonstrates that the data transfer rate can be more than 1,000 times faster than that by the personal computer.

Keywords: remote participation, data transfer, TCP/IP, ITER, network, VPN, VLAN, fusion

1. Introduction

In modern fusion experiments, as many collaborators in distributed locations use a few large experimental devices, remote communication technologies become more important. Many efforts have been carried out for the remote participation in JET [1, 2] and other fusion experiments [3]. In JET, they have developed infrastructures and tools as follows: the remote computer access, the remote data access, the remote meeting, the remote control overhead display, the remote real time display [1]. The National Institute for Fusion Science (NIFS) has been also developing and operating the Japanese virtual laboratory system for fusion, which is known as "SNET" [4, 5, 6, 7]. The remote participation system of LHD is included in SNET.

A typical example of large fusion experiment is International Thermonuclear Experimental Reactor (ITER), which will be built in Cadarache, France. The world fusion community requires the remote participation to join the ITER experiment. Ideal remote participation system would provide an environment of control room at each remote sites. Shissel has proposed a remote control room for ITER using the GRID technology [8]. As the security level is increased, the convenience is decreased. The highest network security is obtained by no connection. It has been pointed out that future network issues are as follows: security, capacity and virtual private network (VPN) [2]. The conflict between security and convenience (access requirement) limits the expansion of JET remote participation services in a worldwide scalability [3]. The ITER remote participation includes challenging technical issues.

author's e-mail: nagayama.yoshio@nifs.ac.jp

Let us estimate the data size of ITER raw data. The data sizes in present fusion experiments (LHD, JET, Alcator C-mod) are about 10 GB per shot. The reason may be that the diagnostics system is similar in any fusion experimental device. Since the plasma duration time will be 10 times longer than LHD and the diagnostics channels will be 10 times greater than LHD (or other fusion devises), the raw data size may be 1 TB per shot. Assume ITER works as 2 shots/hour, 20 hours/day and 150 days/year, the total data size will be 6,000 TB/year. The original data is so precious that the back up of original data is important. In LHD, the back-up of raw data is stored in Bluray disks (BD). The storage size of BD is 50 GB. If ITER data is backed up into BDs and the data is compressed to 33 % of the original data size, 40,000 pieces of BD will be increased every year. Hard disk (HD) is the most cost effective, but HD is rather fragile. A possible solution is mirror servers to secure the ITER data. If the access of the original ITER data server is limited to a few mirror servers, the security level of the original ITER data will be much increased. In this case, physicists should access mirror servers. Also mirror servers are useful to secure the ITER data in the case of disaster.

Members of Japanese fusion community are intrinsically ITER collaborators. The remote participation system is crucial for Japanese to join the ITER experiment. As the ITER remote participation center is to be built at Rokkasho site in Japan, we assume so in this paper. In this case, the data transfer becomes a big issue. Due to the long distance between Japan and France, which is 15,000 km in fiber length, the data transfer rate may be extremely reduced. In a previous paper [9], we describe how to reflect physicist's requirement to the ITER remote participation

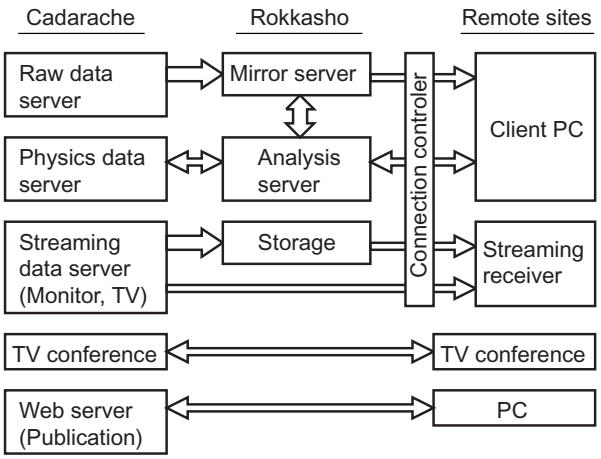

Fig. 1 Proposed connection of servers and clients in the ITER remote participation system.

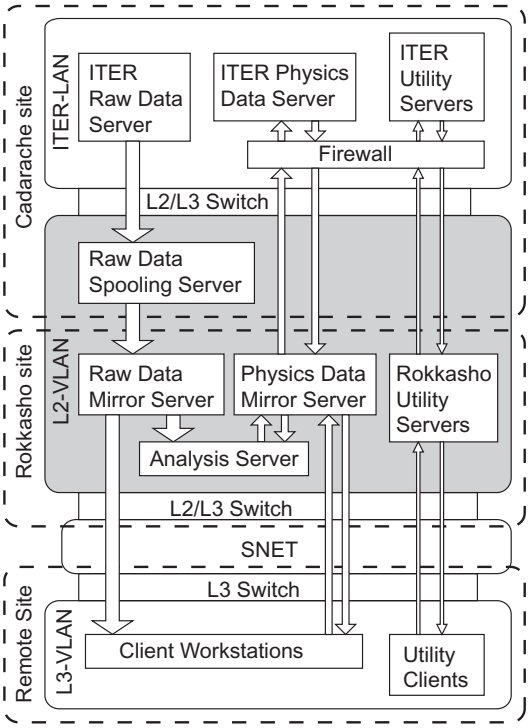

Fig. 2 Proposed virtual local area network (VLAN) system to connect remote sites and ITER site.

system. This paper will present different aspect, such as the VPN and the fast data transfer. These are key technologies to establish ITER mirror sites. In this paper we discuss the case of the Japanese ITER remote participation system, but technologies discussed here can be useful in any remote participation system. In sec. 2, we discuss conflict between security and convenience. It takes a long experience to obtain a good solution. In sec. 3, the network system is discussed. In sec. 4, we present a proof-of-principle experiment of the fast data transfer technology.

2. Security and Convenience

In order to maximize the user's convenience, let us consider physicists' behavior. Typical physicists behaves as follows: (1) preparation, (2) experiment, (3) analysis, (4) publication. In the preparation of the experiment, they discuss their plan with co-workers, propose the experiment, and get machine time. In order to discuss with co-workers in different places, they use a video conference system. It would be better the video conference system has high resolution such as High Definition TeleVision (HDTV) and a viewgraph viewer. In order to join the experiment from the remote site, monitoring systems for their own device and plasma (real time TV and parameters) are required. Both video image and sound in the control room and the machine would provide the reality. Streaming servers may provide broadcasting service of the real time monitoring and HDTV. Video conference would be also useful, when the co-workers are separated. Operation of own diagnostics would be preferable. After the experiment, they analyze the data at the remote site. When carrying out the data analysis, they use the data server and sometimes they would use the data analysis near by the data servers. Once they obtain new result, they would publish it in the conference or in the academic journal. They discuss with co-workers using the video conference system, they circulate the manuscript (or presentation), then they would have approval of the publication.

Therefore, a physicist staying at the remote site requires, (1) the video conference system; (2) the broadcasting system of experiment (HDTV, monitoring data, machine status); (3) the data and analysis servers; (4) the publication system. Remote participation system should provide the above four items in a secure network system.

Figure 1 shows a schematic diagram of proposed connection of servers and clients in the ITER remote participation system. ITER collaborator may directly access the utility servers, such as web servers, which provide useful informations and receiving system of applications of experimental proposal. User may access the mirror server, to which the ITER raw data is sent by the unidirectional transfer to protect the ITER data from infection. At the Rokkasho site, a data analysis server (cluster) is to be installed. Users can analyze huge data by using the analysis server and get the results through the network, of which speed may not be critical. Physicists will be able to analyze ITER data using standard analysis code in the analysis server [10]. The problem is as follows: the Internet is too dangerous to connect the ITER severs; the local area network (LAN) is limited to the each local site.

In order to secure the above items for convenience,

we propose a VPN system, as shown in Fig. 2. The Japanese ITER remote participation system is enclosed in a closed VPN. This must be safe because it can be accessed from the outside network through only few well controlled connection points. A layer 2 (L2) virtual local network (VLAN) covers both Cadarache site and Rokkasho site with the VPN technology. The VLAN works as a highly secured virtual dedicated line. Inside the L2-VLAN, most communication protocols can be freely transferred. For example, the most layer 3 (L3) VLAN limits the multicast but the L2-VLAN approve it. The multicast is very useful, as it transfers the same information (the ITER timing) to many clients at the same time. The L2-LAN provides an environment that the analysis server in the Rokkasho site starts the calculation of the ITER data analysis at the right timing. The L2-VLAN including the ITER site, the University of Tokyo and NIFS was temporally established during the 4 Gbps data transfer experiment, and it worked well. In Fig. 2, utility servers mean streaming data server, TV conference and Web server in Fig. 1. The detail of the network system is described in Sec. 3.

3. SNET

SNET has been developed using the "SINET3" high-speed network operated by the National Institute of Informatics (NII) [11]. SNET consists of 3 categories as follows: (1) the LHD remote participation, (2) the bi-directional remote participation, (3) the remote use of supercomputer. These categories have own VPN. At each remote site, a LAN is made and is connected to SINET3 by using a L3 switch. The LHD remote participation system has been in service since 2002. In this system, a secured closed network is made by using the L3-VPN. This system provides the same environment of the LHD control room. Physicists can operate their diagnostics and analyze LHD experimental data at the remote site. The security policy of SNET is as follows [6]: (1) an outside network should be connected through a firewall, which limits protocols and other items; (2) a device with a registered media access control (MAC) address can be connected by using a function of a L3 switch; (3) a SNET user should install a vaccine software delivered by the SNET administration and update the operation system of own computer properly. The L3 switch can check the MAC address at every interface port.

The bi-directional remote participation system has been in service since 2008. This system delivers the experimental data of the remote site to Japanese fusion community. In 2008, this system has been developed to deliver experimental data of the Q-shu University Experiment with Steady-state spherical Tokamak (QUEST), which is built at Kyushu University. In this case, the QUEST data is transfered to the LABCOM system [12, 13] with the L2-VPN between Kyushu University and NIFS. The QUEST data is delivered to remote site, which is in a L3-

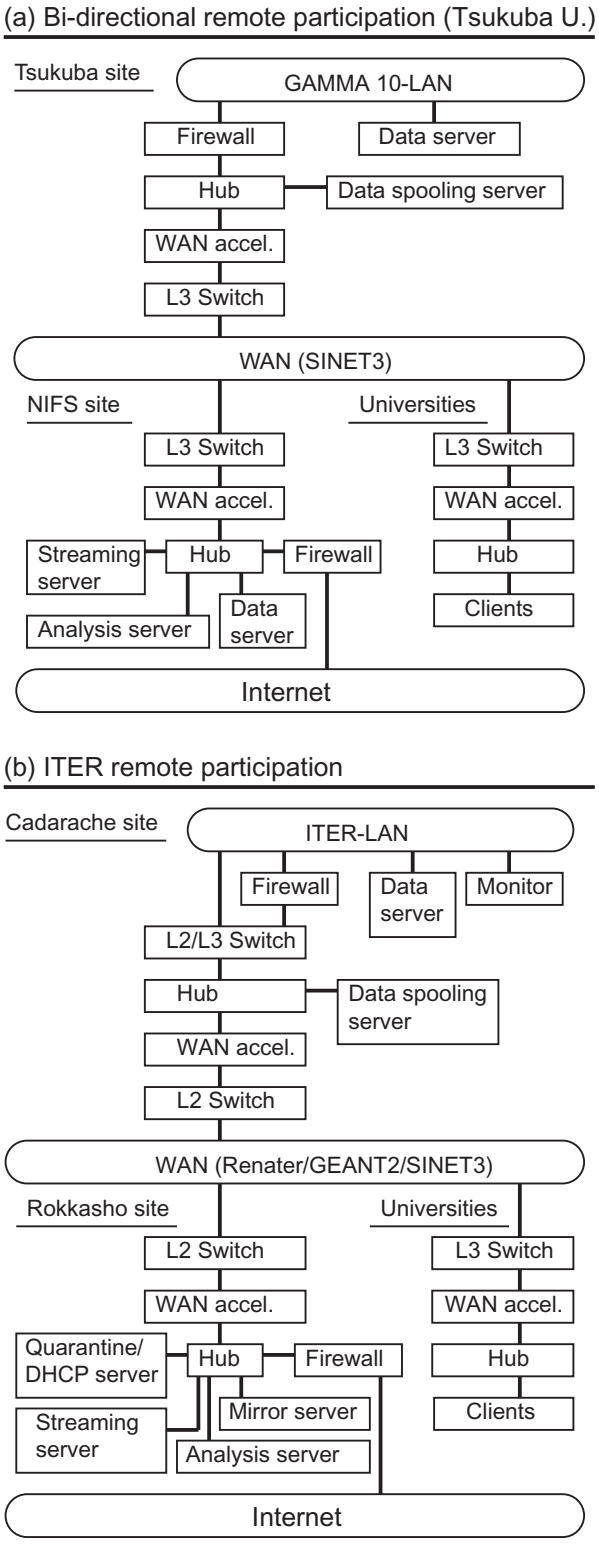

Fig. 3 (a) Network in the Bi-directional remote participation system for Tsukuba University. (b) Proposed network in the ITER remote participation system.

VPN. The GAMMA 10 experimental data has been delivered by using the bi-directional remote participation system since 2009. The GAMMA 10 is a tandem mirror device at Tsukuba University. Figure 3(a) shows schematic diagram of the network connection in the bi-directional remote participation system for the GAMMA 10 experiment in Tsukuba University. The experimental data in the GAMMA 10 data server is transfered to the data spooling server through a firewall. Data in the data spooling server is sent to the data server at NIFS through the SINET3. As the data server at NIFS is shared in the SNET, collaborators at remote sites can use the GAMMA 10 data. NIFS also provide an analysis server and a Web data viewer [10]. Since the experimental data is huge, fast data transfer is required. The obtained data transfer rate is 300 Mbps by using a wide area network (WAN) accelerator, which is installed at the front end of LAN.

Differences between the QUEST system and the GAMMA 10 systems come from the fact that the QUEST data is taken by the LABCOM system, while the GAMMA 10 data is taken by the original data server. The reason why the client cannot access the data server of GAMMA 10 directly is as follows: The original GAMMA 10 data server stays inside the GAMMA-10 control network, which the GAMMA-10 control computer belongs to. If the GAMMA 10 data server is open to the public, the GAMMA 10 control computer could be also open. This is very dangerous. The data is transferred through the firewall from the original GAMMA 10 data server to the public data server. As the ITER CODAS is for the control of devices and for the data acquisition, the public data server should be outside the ITER CODAS in order to improve the security. So, the ITER case is similar to the GAMMA 10 case.

The SINET3 is connected to international academic networks, ESnet in USA and GEANT2 in EU. The Japanese fusion community can easily utilize the remote communication technologies by adding to SNET a new VLAN for ITER. Experience and technologies in SNET will be useful to establish the ITER remote participation system. An example of possible network connection of the ITER remote participation is shown in Fig. 3(b). As Rokkasho site provides a mirror data server that is in the outside of the ITER-LAN, the ITER system may be similar to the GAMMA 10 remote participation system. In this case, the ITER-LAN is connected to SNET with a switch because the fast data transfer is required. As the direction of data is one way from ITER to SNET, the network switch can protect the ITER-LAN. In order to protect servers, user's computers should get a security check before connection. Here, we propose to use a quarantine server, which checks security level of user's PC and approve the network connection using a dynamic host configuration protocol (DHCP) server. A removable storage should be also checked because it may have a chance to be infected.

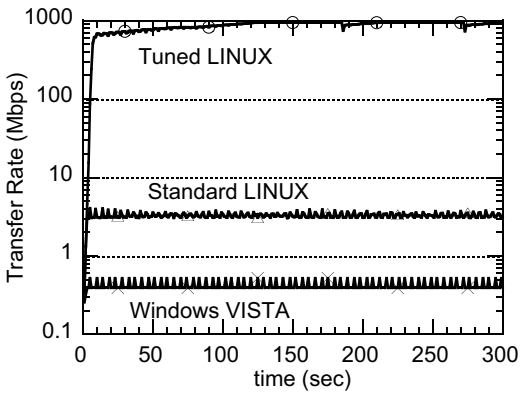

Fig. 4 Data transfer rate between ITER site and NIFS in the case of Windows VISTA, the standard LINUX and the tuned LINUX.

4. Experiment of Data Transfer

Present computer network technology uses the TCP/IP protocol. Since data packet is sent after receiving the acknowledgment in the TCP/IP protocol, the data transfer rate is decreased as the distance increased. So, the effective throughput is much less than the network bandwidth [14]. We are developing an acceleration technology [15, 16]. Between the ITER head quarter building (ITER-HQ) and NIFS, we carried out two proof-of-principle tests, as follows: (1) 1 Gbps test, (2) 4 Gbps test.

In the case of 1 Gbps test, ITER-HQ is connected to the 1 Gbps port of Cadarache node of Renater, and NIFS is connected to the 1 Gbps port of NIFS node of SINET3. The network path between the ITER-HQ and NIFS is as follows: ITER-HQ, Renater, GEANT2, SINET3, NIFS. The band width of the network lines between Renater and SINET3 is 10 Gbps, but the total bandwidth is 1 Gbps due to 1 Gbps ports. The total network length is 15,000 km, and it takes 0.3 sec from sending data to receiving acknowledgment. At NIFS, a rack mounted server, of which operation system (OS) is LINUX, is connected to the SNET router. At ITER-HQ, a note PC is connected to a port of the network switch. The data transfer rate is measured with a software "iperf" [17]. Here, the data is transfered from the memory of note PC at ITER site to the memory of the LINUX server at NIFS.

The result is shown in Fig. 4. When the OS of note PC is Windows VISTA, the measured data transfer rate from ITER to NIFS is 0.4 Mbps. When the OS of note PC is standard LINUX, the measured data transfer rate from ITER to NIFS is 3 Mbps. In order to accelerate the data transfer, the driver of network

interface is modified to optimize the data transfer. By optimizing the window size and the interval of packet, packet loss was very few in this test. The data of 1.18 TB was sent within 3 hours, so that the averaged data transfer rate is 881 Mbps. In this case, the peak data transfer rate is 899 Mbps.

In the case of 4 Gbps test, ITER-HQ is connected to the 10 Gbps port of Renater, and the NIFS is connected to the 10 Gbps port of SINET3. The network path between the ITER-HQ and NIFS is similar to the case of the 1 Gbps port test. Since the network is a shared line, the data transfer rate is limited to 4 Gbps in order to reduce disturbance to the network traffic. Computers at ITER-HQ and at NIFS are faster LINUX computer with solid state drive (SSD). The data of 86 GB from the disk of ITER-HQ was sent to the disk of NIFS within 205 seconds, so that the averaged data transfer rate is 3.3 Gbps. Beside the shared line, a L2-VLAN was temporally established to connect ITER-HQ, NIFS and the University of Tokyo. However, the the data transfer rate was similar to that in the case of shared line. These tests indicate that the data transfer rate of about 90 % of the band width is available.

5. Summary

Technical issues of the remote communication system are as follows: (1) conflict between security and convenience; (2) data transfer rate, because the huge ITER data should be sent to Japan; (3) a global distributed data mirroring system to store the ITER data (raw data and physics data). Our solution to solve the conflict is enclosing the Cadarache site, the Rokkasho site and other remote sites by a closed VPN. This is natural expansion of SNET. The solution of data transfer is the WAN accelerator. The proof-of-principle experiment to obtain 90 % of the network bandwidth was carried out between the ITER-HQ and the NIFS. We learned that the data transfer rate may be limited to 40 % of the bandwidth when using shared line in order to minimize the disturbance to the network traffic. In order to send 1 TB within 200 sec, the data transfer rate should be 40 Gbps. When the ITER experiment begins, the network bandwidth of 100 Gbps will be available. Further development will be required in the data transfer technology and the data mirroring. One of the most important technologies is the 40 Gbps WAN accelerator, which is 100 times faster than the present one that is installed in SNET. The relevant technology is applicable for international large-scale research activities in various fields. Also SNET will provide a long experience to obtain a good solution to relax the conflict between security and convenience.

Acknowledgments

SNET is partly supported by Cyber Science Infrastructure development project of the National Institute of Informatics, and NIFS (NIFS08USNX001 and NIFS08USNN002). The data transfer experiments are supported by ITER HQ, SINET3, JGN2plus, WIDE, GEANT2, RENATER and other related NOC teams.

[1] W. Suttrop, D. Kinna, J. Farthing, O. Hemming, J. How, et al., Fusion Eng. Des. **60**, 459-465 (2002).
[2] J. How, V. Schmidt, Fusion Eng. Des. **60**, 449-457 (2002).
[3] D.P. Schissel, J.W. Farthing, V. Schmidt, Fusion Eng. Des. **74**, 803-808 (2005).
[4] M. Emoto, T. Yamamoto, S. Komada, Y. Nagayama, Fusion Eng. Des. **81**, 2051-2055 (2006).
[5] Y. Nagayama, M. Emoto, H. Nakanishi, S. Sudo, S. Imazu, et al., Fusion Eng. Des. **83**, 170-175 (2008).
[6] K. Tsuda, Y. Nagayama, T. Yamamoto, R. Horiuchi, S. Ishiguro, et al., Fusion Eng. Des. **83**, 471-475 (2008).
[7] T. Yamamoto, Y. Nagayama, H. Nakanishi, S. Ishiguro, S. Takami, et al., submitted to Fusion Eng. Des. (2010).
[8] D.P. Schissel, Fusion Eng. Des. **83**, 539-544 (2008).
[9] Y. Nagayama, M. Emoto, Y. Kozaki, H. Nakanishi, S. Sudo, et al., submitted to Fusion Eng. Des. (2010).
[10] M. Emoto, S. Murakami, M. Yoshida, H. Funaba, Y. Nagayama, Fusion Eng. Des. **83**, 453-457 (2008).
[11] S. Urushidani, J. Matsukata, Fusion Eng. Des. **83**, 498-503 (2008).
[12] H. Nakanishi, M. Ohsuna, M. Kojima, S. Imazu, M. Nonomura, et al., Fusion Eng. Des. **82**, 1203-1209 (2007).
[13] H. Nakanishi, M. Ohsuna, M. Kojima, S. Imazu, M. Nonomura, et al., Fusion Eng. Des. **83**, 397-401 (2008).
[14] V. Jacobson, R. Braden, D. Borman, IETF RFC1323 (1992) (http://www.ietf.org/rfc/rfc1323.txt).
[15] H. Kamezawa, M. Nakamura, J. Tamatsukuri, N. Aoshima, et al., ACM/IEEE Proc. SC04 (2004) p.24.
[16] N. Tanida, K. Hiraki, M. Inaba, submitted to Fusion Eng. Des. (2010).
[17] National Laboratory for Applied Network Research, (http://sourceforge.net/projects/iperf)