A GRID Computing VLBI system for Real-time Monitoring of the Earth

Hiroshi Takeuchi, T. Kondo, Y. Koyama, J. Nakajima, R. Ichikawa, M. Sekido, E. Kawai, H. Osaki, M. Kimura, H. Kuboki
(Kashima Space Research Laboratory, NICT, Japan)

1.Introduction

To achieve real-time monitoring of the earth orientation parameters (EOP), worldwide VLBI stations should be connected to the Internet and both real-time data transmissions and real-time data processings are required. We had developed the first real-time VLBI system using ATM network (K4 system) until 1996 and the system had great contributions to detect the regional crustal deformation of the metropolitan area in Japan. However, the system is not always suitable for worldwide operation since this was based on the conventional VLBI systems which used magnetic tape recorders. Thus we have developed a new VLBI system using conventional PC and TCP/IP technology. We call this K5 VLBI system. Since the K5 system is easily connected to the Internet, we can construct a low cost and high performance VLBI analysis system using GRID computing technology. We are now developing the client/server system like SETI@home. In the system, a large amount of data is divided in small segments and transferred to the client PCs. Sent data are analyzed by the PCs and analytic result are sent back to the server. This type of system generally works well when the network speed is faster than the processing speed in client PCs. At present, K5 software correlator have a capability to process 4 Mbps data in real time when it runs on a PC equipped with a Pentium3 1GHz. Thus, distributed computing method has the potential to increase the software correlation speed if the 10Mbps, 100Mbps or more high speed network environments are equipped.

2. Overview of the system

In the developing system, raw VLBI data are divided into appropriate short-time period data and each segmented data are assigned to each PC. In the case of multi-baseline observations, all the data received at the same moment at different stations are gathered into one client to minimize data transmission costs. As shown in Figure 1, system consists of following components: control server that controls whole system, database server that stores processing conditions of VLBI data and statistics of each client PC, FTP servers at each VLBI station that transmit observed raw data to clients, and a lot of clients by which VLBI data are correlated. When a screensaver-type client software activates, it ask the control server about filenames of data to be processed and IP addresses of observed VLBI stations. The client downloads the data from the FTP servers and correlates the data. Resulting data and related information such as download time and correlation time are reported to the database server via the control server.

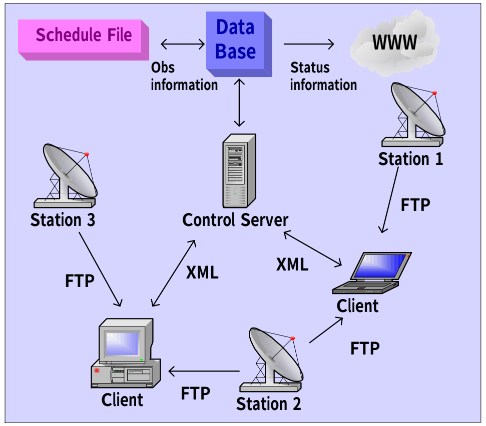

Figure1:Schematic diagram of the system. Radio signals received at world-wide VLBI stations are transmitted via Internet and analyzed in real-time.

3.Detailed structure of control server and client program

A schematic diagram of a control server is shown in figure 3. In the server program, file information of each segmented data are stored as C++ class instances (see the right-hand side of figure 3). Class Instances relevant to unprocessed data segments are listed in a queue (Upper queue in figure 3). Referring the database server, a new class instance is added to this queue in real-time during observations. If a client connects to the server, information in the class instance at the forefront of the queue is transmitted and the instance is transferred to another queue in which information of processing data segments are listed (Lower queue in figure 3). When the correlation process in the client finished, it is reported to the server and relative instance is deleted from the second queue. Finally, reported information is stored in the database server. In client PCs, VLBI-data transfer is realized by a FTP-client function in the screen-saver type client program. A multi-thread simultaneous downloading from different VLBI stations is possible. Correlation process is performed by the external correlator program, cor.exe [Kondo et al., IVS CRL TDC News No.23].

4. Bottleneck of the system

There are three factors which can become a bottleneck of the system. Data processing rates of clients, network speed and data transfer rates of the FTP servers. These three factors can be monitored using statistical data stored in the database server. Data processing rates of clients can be calculated from stored correlation time information. A typical value of the rate using a 1CPU PC is 2 to 12 Mbps. If this factor is the bottleneck of the system, it is necessary to increase the number of clients or to improve the algorithm of correlation programs. The statistics of download time is used to evaluate the network performance. A typical network speed is 10Mbps to several Gbps. The network speeds around FTP servers determine the total performance of the system. If it is not enough compared to a total data processing rate of the clients, it becomes the bottleneck of the system. The data transfer rate of a FTP server is limited by the access speed of hard disks. A typical value is 200Mbps for random accesses and 1Gbps for sequential accesses. When it becomes the bottleneck, FTP mirroring is effective and ram disk can be used for real-time observations. Eliminating these bottlenecks, we can improve the system performance up to 10Gbps, the highest network rate we can use at this time.

Figure 2:GRID VLBI correlator system consists of multiple PCs. In this picture, three standard PCs are correlating VLBI data using screensaver-type client programs like SETI@home.

5. Current status of the system and future plans

The overall performance of the system for a 32-lag XF-type correlation is **70Mbps** using sixteen standard PCs. We are planning to perform a 64-Mbps real-time geodetic experiments. We expect the system will make it possible to estimate the EOP in real-time. The tremendous numbers of PCs all over the world will be able to product high temporal resolution of the EOP time series with high precision. This system is adaptable to increasing or decreasing of the operational stations during observations and significant baselines in terms of EOP estimation can be analyzed preferentially to seek the EOP rapidly. Such EOP information will significantly contribute to a real-time monitoring of vertical crustal deformation, a real-time monitoring of the ionospheric scintillation, a numerical weather prediction using GPS precipitable water vapor, and a real-time precise orbit determination of astrometric satellites and interplanetary spacecrafts.

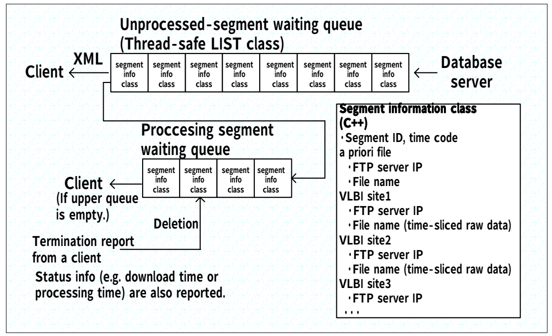

Figure 3:Schematic diagram of control server.