jh190005-MDH

Combination of HPC and high-speed data transfer technologies for big data processing systems

Ken T. Murata (National Institute of Information and Communications Technology)

Abstract
In this research, we combine high performance computing (HPC) and high speed data transfer technologies to improve big data processing systems. Three case studies on large-scale data processing and large-scale information systems are presented which adopt high-performance and flexible protocol (HpFP) as a communication protocol. The HpFP1 is designed for specified networks and puts more emphasis on latency and packet loss tolerances than fairness and friendliness, while the HpFP2 is an improved version of the HpFP1 and is more suitable for real network environments. Based on HpFP2, we implement a file transfer tool, called high-performance copy (HCP), to improve data transfer performance on JHPCN. The performance of our systems is evaluated using real data produced by HPC. The results show that the HCP outperforms the conventional tool for file transfer in JHPCN. We conclude that our systems have significant potential for further researches in JHPCN.

1. Basic Information

(1) Collaborating JHPCN Centers

- Kyoto University
  - Data communication server (server name: XC40, VM hosting): set up high-performance copy (HCP), receive data files from other communication servers at high speed, and save them on large-scale storage.
  - Large-scale storage (1PB): save each research domain data. Researchers at each university access the stored data at high speed from the outside via the data communication server.
  - Data processing server (server name: XC40): perform data processing stored in large-scale storage.
- Nagoya University
  - 8K large-scale visualization system (display server): receive high-speed external storage (especially high-resolution time series of Himawari image data from Chiba University) and display it on an 8K display at high speed (30 fps) in sequence order.
  - Data communication server (server name: UV2000 login node): set up HCP, receive data files from other communication servers at high speed, and save them on large-scale storage. Researchers at each university access stored data at high speed from the outside.
  - Large-scale storage (0.5PB): save each research domain data. Researchers at each university access the stored data at high speed from the outside via the data communication server.
  - Data processing server (server name: UV2000 login node): perform data processing stored in large-scale storage.
- Kyushu University
  - The resources provided by Kyushu University in the interactive environment
  - Other facilities, the resources and methods of use available for

collaborative research: perform data communication experiment using HCP

- Genome data server (Okawa laboratory management, Kyushu University): read the genome data of Kyushu University and transmit the data to the external large-scale storage via the data communication server
- Genome storage (Kyushu University Okawa laboratory management / 150 TB): save the genome data of Kyushu University

□ Tohoku University

- Supercomputer (supercomputer name: SX-ACE): perform Jupiter MHD simulation and then save execution results in supercomputer storage (cache area). The stored data is transmitted to remote large-scale storage via HCP.
- Data communication server (server name: Express 5800): set up HCP and conduct high-speed data transfer experiment of management (communication) server with large-scale storage from other institutions

(2) Research Areas

☑ Very large-scale data processing

☑ Very large-scale information systems

(3) Roles of Project Members

As shown in Section 1(1).

2. Purpose and Significance of Research

This research aims at conducting case studies to demonstrate large-scale data processing and large-scale information systems. We adopt high-performance and flexible protocol (HpFP), which is a communication protocol with high delay tolerance and high packet loss tolerance and is used in various domain science researches. The HpFP1 is designed for specified networks and focuses more on latency and packet loss tolerances than fairness and friendliness, while the HpFP2 is an improved version of the HpFP1 and is more suitable for real network environments. Based on the HpFP2, we implement a file transfer tool, called HCP, to improve data transfer performance on JHPCN. The HCP is used as a tool for the data transmission/reception in the information infrastructure centers of universities via Science Information NETwork 5 (SINET 5).

3. Significance as JHPCN Joint Research Project

SINET5 brings new possibilities with its 100 Gbps full-mesh network connecting more than 800 organizations, clouds, and academic contents. In the viewpoint of high performance computing (HPC), including Grid computing, cloud computing, supercomputers, an inter-university HPC system is one of the goals of high performance computing infrastructure (HPCI). As the high-speed data transfer technology developed and experimented on JHPCN has reached the practical level finally in 2019, it is installed in the information infrastructure system of the application in parallel with the basic experiment to perform file transfer of real data. If this experiment succeeds, the way to future inter-university HPC system will be greatly opened up. By using this way, data output from supercomputers, sensors, measuring instruments, etc. are stored in an arbitrary storage system at high speed and data processing is performed in an arbitrary computing environment. In addition, the processing results are displayed on a large-scale display installed in a specific institution, and it is possible to visualize and analyze collaborative data by multiple researchers as well as demonstrations.

4. Outline of Research Achievements up to FY2018

The JHPCN applications that Murata applied (adopted) as a representative applicant are as follows.

 In 2016, (jh150033-IS02) "Data transmission experiment for realizing big data post processing environment using cloud"

 In 2016, sprout (Kyushu University, JPHCN-Q) "Epigenome Big Data Visualization System Technology for SINET"

 In 2016, sprout (Nagoya University, HPC scientific computing collaboration PJ research) "Experiment of large-scale visualization for remote data on cloud"

 In 2017 (jh170034-ISH) "Mash-up of high-performance numerical computing and high-speed data transfer for large-scale data file transfer between universities and their demonstration experiments with real dataset"

 In 2018, (jh180054-ISJ) "Inter-Datacenter File Transfer Examinations for HPC Using Real Datasets"

As for these achievements, the purpose of this research and development is to utilize computer resources (high-performance supercomputers, large-scale storages, large-screen visualization devices/displays) of information infrastructure centers, and IoT sensors (especially video IoT devices) on high-speed infrastructure networks (SINET, JGN, and international networks). By combining these resources, we conduct three case studies of big data processing systems. The results adopted by international conference with peer review in 2019 and presented in domestic conference in 2019 are described in list of publications and presentations as below.

5. Details of FY2019 Research Achievements

We present three case studies on large-scale data processing and large-scale information systems.

(1) Extraction of Meteorological Information Using Artificial Intelligence (AI)

We consider the meteorological information from two websites developed by NICT. The first one is Himawari-8 real-time web (https://himawari.asia/), which provides Himawari-8 satellite sensed images in real time and with full resolution, as shown in Fig. 1. The Himawari-8 real-time web is designed for personal computers and smartphones. The second one is Amaterass web (http://amaterass.nict.go.jp/), which provides ground weather data (e.g., wind speed, solar radiation, air temperature, and rainfall) in real time, as shown in Fig. 2. Both data are analyzed using AI in NICT and Chiba University, and then are processed by supercomputer in Tohoku University and Nagoya University. The visualization is displayed on tiled display wall (TDW) in Chiba University, RIKEN and Kyushu University using STARS by NICT and ChOWDER by Kyushu University. Currently, we transfer up to 371 MB (741 pyramid tile image files) of the Himawari data every 10 minutes and 500 GB (5 files) of the Amaterass data per day. In addition, we also transfer the original PNG files of Himawari data to Kyoto University.

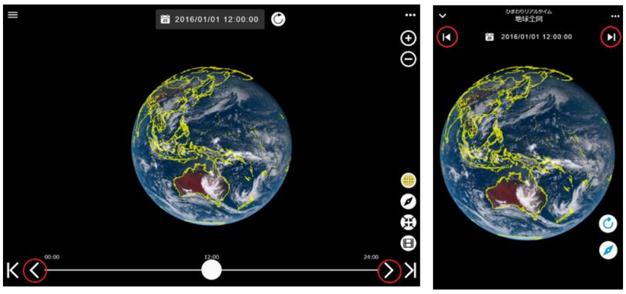


Fig. 1 Himawari-8 real-time web (https://himawari.asia/)

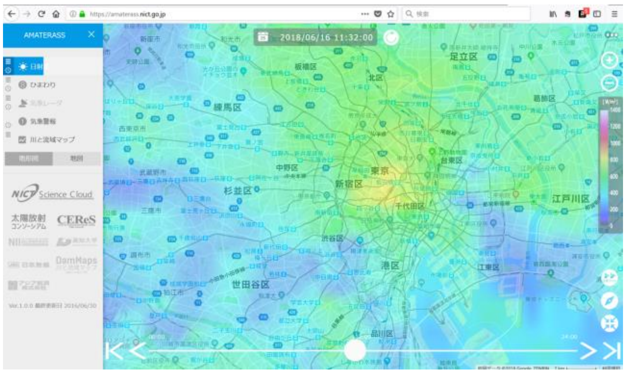


Fig. 2 Amaterass web (http://amaterass.nict.go.jp/)

(2) Concurrent Processing System

We consider the Genomic information collected from Genetic Analyzer (Illumina Novaseq 6000) in Kyushu University. There are only five Genetic Analyzers in Japan. The running time is about half a day to two days. Therefore, it is necessary to transfer the Genomic information collected from Genetic Analyzers to the large-scale storages at high speed. The data transmission is performed from Kyushu University to Kyoto University using HCP, as shown in Fig. 3. The visualization is displayed in real time on the web application, as shown in Fig. 4. Currently, we transfer up to 1.1 TB by demand.

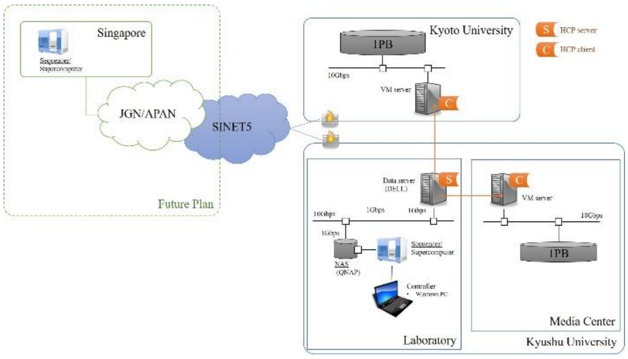


Fig. 3 Data transmission model in JHPCN from Kyushu University to Kyoto University

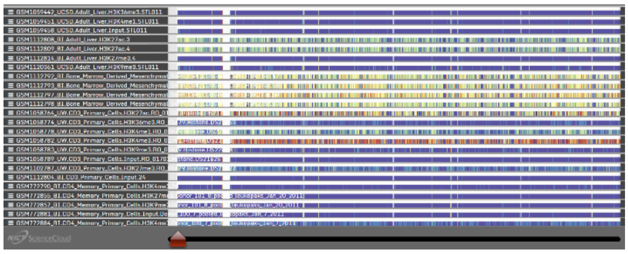
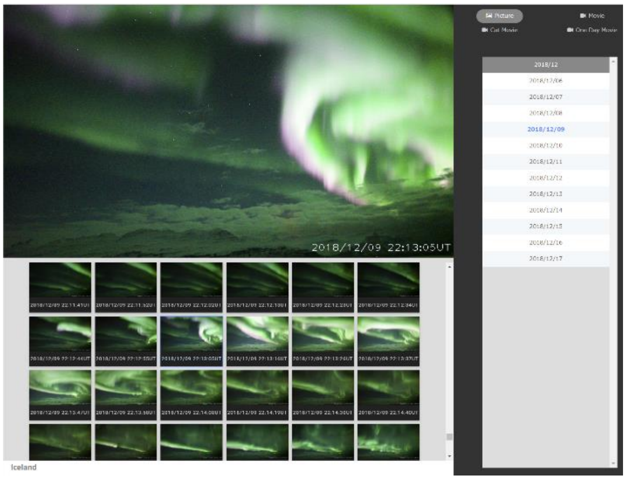


Fig. 4 Example of web application for viewing epigenomic data

(3) Extraction of Visual Internet of Things (IoT) Data using Deep Learning

We consider the visual IoT data collected from SmartSight camera, as shown in Fig. 5(a). The image/video transmission is fixed with 30 fps based on high resolution (full HD). The SmartSight cameras are set up in 20 locations of Kochi, Arctic region, Hirosaki University, Chikuma, and Tokyo, and Chiang Mai, Thailand. All of the data are transferred to Kyoto University, and then are analyzed and processed using deep learning with TensorFlow on supercomputer. The visualization is displayed on TDW, as shown in Fig. 5(b). Moreover, the additional visual IoT data from other locations are transferred to Chiba University. Currently, we transfer images every 10 seconds and videos every 10 minutes.

(a) Visual IoT device



(b) Example of real-time monitoring system

Fig. 5 SmartSight camera

6. Progress during FY2019 and Future Prospects

We evaluate the performance of our file transfer tool for high-speed data transmission on JHPCN in three case studies. The HCP is operated in fair mode to avoid stressing transmission control protocol (TCP) traffic. The results show that we succeed the data transmission in three case studies.

In the first case study, we demonstrate that the HpFP2 with fair mode outperforms the traditional TCP. Figure 6 shows the comparison of TCP and HpFP2 with fair mode in JHPCN from NICT to Kyoto University. The throughput of HpFP2 with fair mode is higher than that of TCP. Figure 7 shows the Himawari data transferred to Kyoto University. The data size changes depending on the local time of day, since the PNG compression works particularly well for black area. It is obvious that the data size in the daytime is larger than that in the nighttime. Figure 8 shows the daily data transmission from NICT to Kyoto University. The total size of the Himawari data is up to 27 GB.

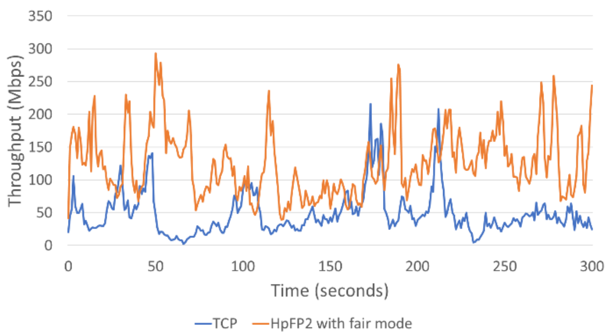


Fig. 6 Comparison of TCP and HpFP2 with fair mode in JHPCN from NICT to Kyoto University

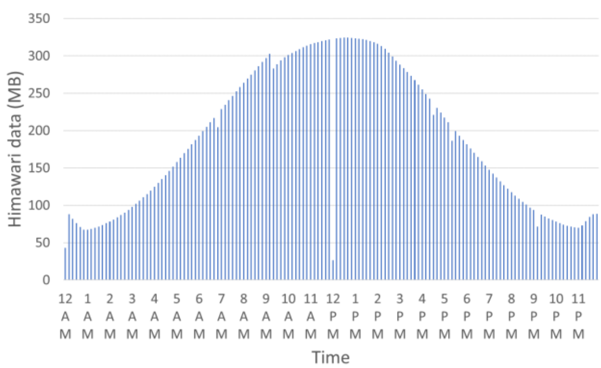


Fig. 7 Himawari data on April 26, 2020

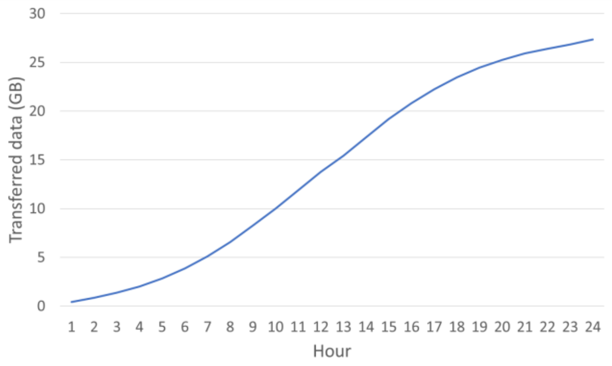


Fig. 8 Daily data transmission from NICT to Kyoto University

In the second case study, we investigate the improvement of data transmission over JHPCN using HCP, compared to the traditional tool, called rsync. The rsync command is run with the -a (archive mode) and -v (verbose) options. In this experiment, the performance of the rsync and the HCP with fair mode is evaluated over JHPCN from Kyushu University to Kyoto University using epigenomic data collected from supercomputer resources, as shown in Fig. 4. The epigenomic data size in different data series are 20 MB, 16 MB, 327 MB, 328 MB, 1.1 TB, 850 GB, 75 GB, and 54 GB, respectively. It is obvious that the throughput of HCP with fair mode is higher than that of rsync in every file transfer, as shown in Fig. 9. In other words, the HCP with fair mode achieves data transmission improvement, compared to the traditional tool.

In the last case study, we demonstrate the daily data transmission from the SmartSight cameras to the large-scale storages. Figures 10 and 11 show the daily data transmission from the SmartSight cameras to Kyoto University and Chiba University, respectively. The total sizes of the visual IoT data transferred to Kyoto University and Chiba University are up to 96 GB and 82 GB per day, respectively. This implies that the monthly data transmission from the SmartSight cameras to the large-scale storages is more than 5TB.

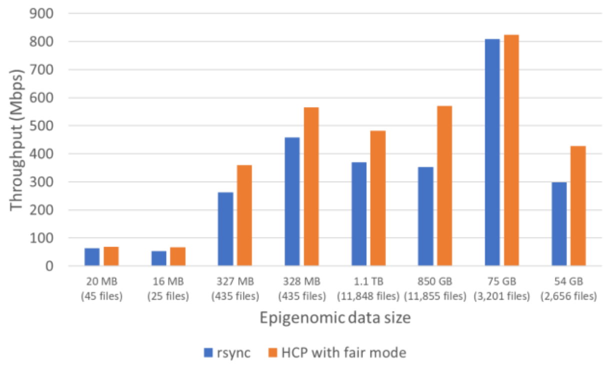


Fig. 9 Data transmission from Kyushu University to Kyoto University

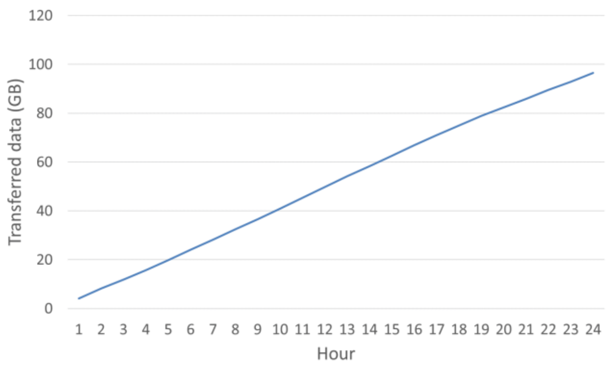


Fig. 10 Daily data transmission from SmartSight cameras to Kyoto University

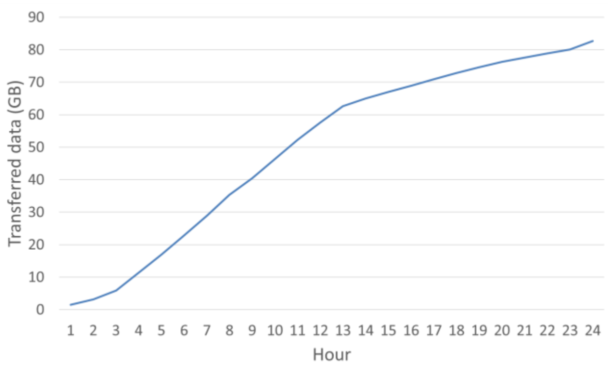


Fig. 11 Daily data transmission from SmartSight cameras to Chiba University

In future prospects, we will evaluate the performance of TCP and HpFP2 in a variety of data sizes on Gfarm environment over JHPCN. We will also investigate the performance of the HCP in each mode and fine-tune the parameters of the HCP to achieve the best performance. Furthermore, we will also investigate the performance of the HCP over collaborative international networks, e.g., SINET5, Japan Gigabit Network (JGN) and Asia Pacific Advanced Network (APAN), as shown in Fig. 3. The results of this research are expected to be standardized for file transfer over JHPCN.

7. List of Publications and Presentations

(1) Proceedings of International Conferences (Refereed)

P. Pavarangkoon, K. T. Murata, K. Yamamoto, K. Muranaga, T. Mizuhara, K. Fukazawa, R. Egawa, T. Katagiri, M. Ogino, and T. Nanri, "Performance Improvement of High-Speed File Transfer over JHPCN," 5th IEEE International Conference on Cloud and Big Data Computing (CBDCom 2019), Aug. 2019.

(2) Presentations at domestic conference (Non-refereed)

P. Pavarangkoon, K. T. Murata, K. Yamamoto, T. Mizuhara, Y. Kagebayashi, A. Takaki, K. Muranaga, and K. Fukazawa, "Experimental Evaluation of High-Speed File Transfer over JHPCN," Japan Geoscience Union Meeting 2019 (JpGU 2019), May 2019.