



December 14, 2017

MMCFTP file-transfer protocol Achieves Transmission Speeds of 231 Gbps

A New World Record for Long-Distance Data Transmission

Japan's National Institute of Informatics (NII; Chiyoda-ku, Tokyo; Dr. Masaru KITSUREGAWA, Director General), one of four organizations that constitute the inter-university research institute corporation, the Research Organization of Information and Systems, has conducted successful data-transfer tests in which MMCFTP (Massively Multi-Connection File Transfer Protocol), a file-transfer protocol developed by NII, was used to achieve stable transmission of 10 terabytes (TB)¹ of data between Japan and the United States at data-transfer speeds of approximately 231 Gbps. In previous tests of data transfer between Japan and the United States, conducted in November of last year, NII recorded a data-transfer speed of 148.7 Gbps, significantly greater than the 80 Gbps speed then reported to be the "world record" for long-distance data transmission. The new test results reported here, which update these previous findings, were announced at the FY2017 annual meeting of the Academic eXchange for Information Environment and Strategy (AXIES) in Hiroshima, Japan, on December 15, 2017.

Test Results

For data-transfer tests, 3 test channels connecting Japan and the United States—each with capacity 100 Gbps, for a total of 300 Gbps—were secured, and data were transferred to Japan from the venue of the **SC17**² international conference (convened in Denver, Colorado, USA from 11/12/2017 to 11/17/2017) in a testing condition known as *memory-to-memory* (M2M).³ In 7 trials each involving a 10 TB data transfer, the effective data transfer speed (or *goodput*⁴) ranged from 224.9 Gbps (transfer time 5 minutes 55 seconds) to 231.3 Gbps (transfer time 5 minutes 45 seconds) (Figure 1). The 10 TB data volume corresponds to 400 typical Blu-ray disks (25 GB each) or approximately 1200 hours of video content at terrestrial digital broadcast quality. The tests announced here successfully transferred this enormous volume of data from the United States to Japan in under 6 minutes.

The transfer tests performed in November of last year, conducted in collaboration with Japan's National Institute of Information and Communications Technology (NICT), involved data transfers to Japan from Salt Lake City in the United States. In these tests, 2 test channels with 100 Gbps capacity each (total capacity 200 Gbps) were used to transfer 10 TB in a memory-to-memory condition, achieving goodput of 148.7 Gbps (total transfer time 8 minutes 58 seconds).

Whereas conventional transmission protocols that use only a single TCP connection are restricted to using just the bandwidth of a single channel, even in the presence of multiple channels, the MMCFTP protocol, which uses multiple TCP connections, can utilize the total bandwidth provided by all available channels; it was by exploiting this advantage of MMCFTP that the tests reported here were able to exceed the record data-transfer speeds achieved last November. The new record set by these tests may be understood as the *world's fastest* data transfer using a pair of servers in the category of intercontinental-class long-distance data transmission.

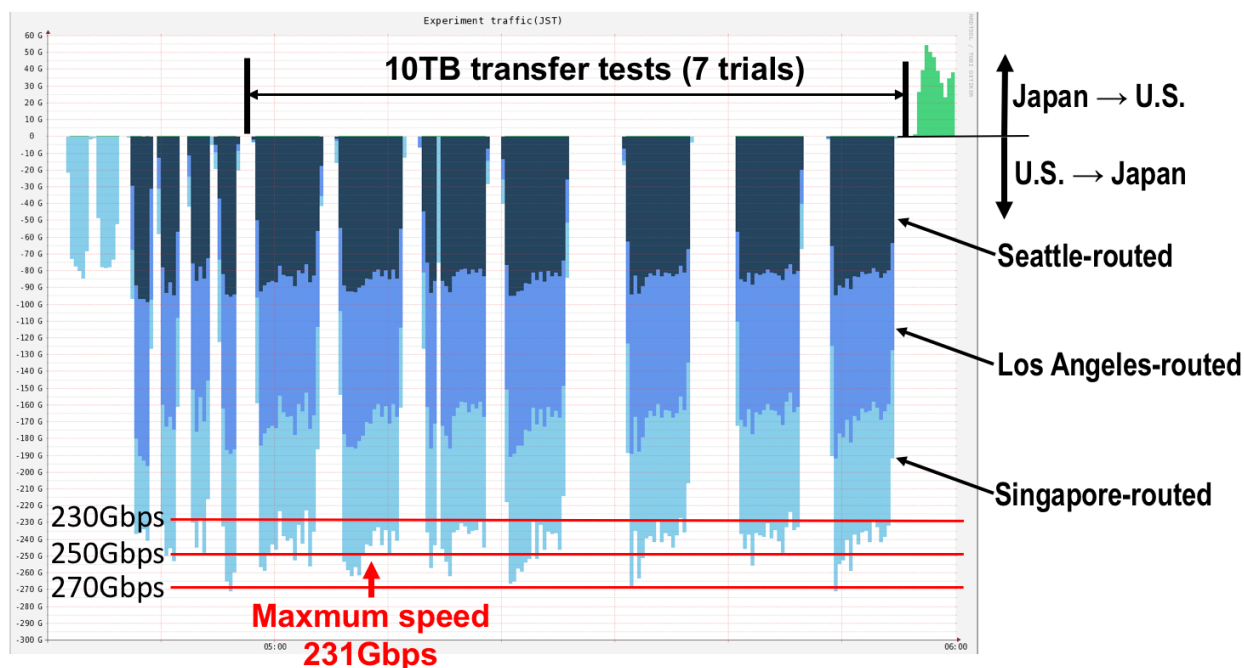


Figure 1: Experimental results (Bandwidth utilization status for equipment installed within the SC17 convention center; measurement interval 20 seconds. Provided by JP-NOC.)

The connection allocation control mechanism in MMCFTP

The MMCFTP protocol, which was developed to facilitate transfers of large quantities of experimental data among international collaborators in advanced scientific and technological fields, has the advantage of allowing multiple TCP connections to be utilized simultaneously when transferring big data. The protocol monitors network status (including the magnitude of delays and packet-loss ratios) and dynamically adjusts the number of TCP connections appropriately to achieve stable data transmission at ultra-high speeds. The tests reported here simultaneously used multiple channels with different transmission distances; in conventional protocols, large volumes of traffic would flow through short-distance channels, resulting in congestion that degrades transmission quality and goodput. This problem is solved by MMCFTP's connection allocation control mechanism.

The tests reported here used **SINET5**, a 100 Gbps academic information network built and operated by NII that connects the United States to Japan via Los Angeles, together with a 100 Gbps channel via the U.S. city of Seattle and a 100 Gbps channel routed through Singapore. The Tokyo-Singapore portion of the Singapore-routed channel is a new channel that began operation in November 2017 as a joint project of NICT, the National Supercomputing Centre Singapore, and the Singapore Advanced Research and Education Network (SingAREN). The round-trip delay time of the channels, a benchmark that measures transmission distance, was

125 milliseconds for the Seattle-routed channel, 122 milliseconds for the Los Angeles-routed channel, and 288 for the Singapore-routed channel; thus, the Singapore-routed channel is more than 2 times longer than the other two channels.

When using MMCFTP to transmit data via multiple channels, it is possible to specify the relative proportions of the numbers of TCP connections assigned to each channel. In the tests reported here, the ratio of connections allocated to the Seattle-routed, Los Angeles-routed, and Singapore-routed channels was set at 1:1:4. In practice, the actual numbers of TCP connections used to achieve a data-transfer speed of 231.3 Gbps were 3,713, 3,713, and 14,851 (total 22,277), or a 1:1:3.9997 ratio, in agreement with the specified ratio. As a result, the average goodputs for the Seattle-routed, Los Angeles-routed, and Singapore-routed channels were 80.5 Gbps, 80.8 Gbps, and 70.6 Gbps, respectively, confirming the suppression of congestion on the various channels and enabling the high data-transfer speed of 231.3 Gbps to be attained.

Structure of the test network

For the tests reported here, the test-channel environment of the **JGN** R&D test-bed network operated by NICT was extended to the NICT booth in the SC17 convention center, and general-purpose servers were installed as transmitters and receivers at the NICT booth and at the JGN Tokyo node. In addition to SINET5 and JGN, the test network was constructed with assistance from the U.S. academic network **Internet2**, SingAREN, the **TransPAC** network operated by the University of Indiana in the United States, the **PacificWave** interconnection point for U.S. academic networks, Japan's **WIDE Project**, and the **SCinet** network inside the SC17 convention center (Figure 2). Coordination of the overall test network was handled by NICT, while the JP-NOC team, with KDDI at its core, was responsible for configuration and tuning of network devices.

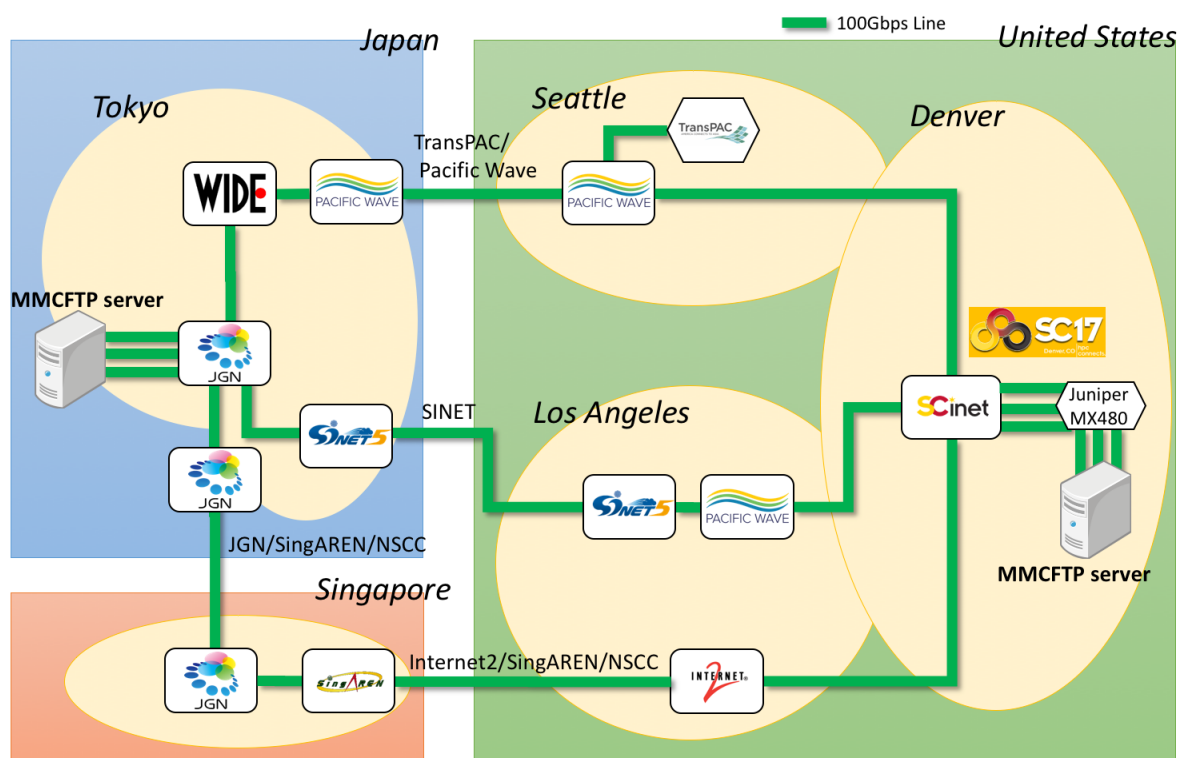


Figure 2: Structure of the test network⁵

Background on the historical motivation for MMCFTP and recent efforts to improve it

In advanced disciplines of scientific research such as elementary particle physics, nuclear fusion, and astronomy, large volumes of experimental data acquired by enormous experimental instruments constructed by teams of international collaborators are transmitted around the world for analysis in participating countries. To facilitate this, the construction of 100 Gbps-grade ultra-high-speed networks is currently in progress, with the **SINET5** network now extended to all regions of Japan and connected to the United States through a 100-Gbps link. As networks themselves become increasingly high-speed, the limitations of data-transfer protocols have prevented long-distance data-transfer speeds from increasing accordingly.

NII developed MMCFTP to address this issue and has worked to improve its performance through tests in Japan and abroad (Table 1). Under *disk-to-disk (D2D)* conditions, in which a file is read from a disk and transmitted to a disk on the remote machine, tests conducted in May 2017 recorded a data rate of 97 Gbps.

Date	Transfer Between	Round-trip delay time	Goodput	
			M2M	D2D
March 2015 ⁶	Koganei (Tokyo, Japan) – Nomi (Ishikawa Prefecture, Japan)(round trip)	26ms	84Gbps	-
August 2016	Saint-Paul-les-Durance (France) – Rokkasho Village (Japan)	200ms	-	7.9Gbps
November 2016	Salt Lake City – Tokyo	113ms/115ms	150Gbps	-
May 2017	London – Tokyo	240ms/242ms	131Gbps	97Gbps
November 2017	Denver – Tokyo	122ms/125ms/288ms	231Gbps	-

Table: Results of data-transfer tests using MMCFTP

NII provides MMCFTP for the purposes of advancing cutting-edge science and technology, and in the future we will continue striving for better stability and higher speeds through testing and real-world utilization.

(End of document)

Media Contact:

**Research Organization of Information and Systems
National Institute of Informatics**

Communications

Direct:+81(0)3-4212-2164; E-mail : media@nii.ac.jp

Footnotes

1. TB: A unit of data volume. One terabyte is 1 trillion bytes (1,000 gigabytes).
2. SC17: An international conference and exhibition in the field of high-performance computing. <http://sc17.supercomputing.org/>
3. Memory-to-memory (M2M): A testing condition that measures performance in writing data from memory on the transmitting machine to memory on the receiving machine. Because M2M data-transfers involve no reading or

writing of data to or from disks, they are not affected by disk-performance limitations and are thus optimal for measuring the performance of transmission protocols.

4. Goodput: Throughput (data volume transferred per unit time) calculated based only on the actual data to be exchanged between applications, neglecting overhead used for communication control such as retransmission and protocol headers.
5. The network equipment installed within the SC17 convention center was provided by Juniper Networks Inc.
6. See the NII News Release of May 13, 2015 titled "NII Succeeds in Achieving One of World's Fastest Long-Distance Transmission Speeds: Records transmission speed of 84 Gbps using new MMCFTP protocol" (http://www.nii.ac.jp/userimg/press_20150513_E.pdf)