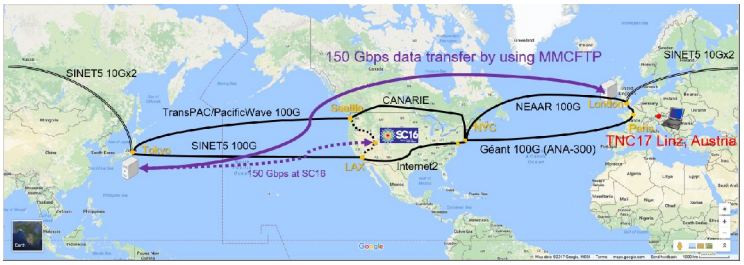
TNC17 Networking Conference

150 Gbps data transfer between Tokyo and London by MMCFTP

K.YAMANAKA[1], V.CAPONE[2], S.ABE[1], S.URUSHIDANI[1] [1]NII [2]GÉANT

Introduction

- Massively Multi-Connection File Transfer Protocol (MMCFTP [1]) is a new file transfer protocol designed for int'l sharing of big data.
- MMCFTP uses several thousands of TCP connections to sustain the specified target speed, and controls the number of TCP connections dynamically according to the network conditions.
- MMCFTP achieved 150 Gbps data transfer speed between Tokyo and Salt Lake City in SC16. At TNC17, we will try data transfers between Tokyo and London at the same speed, as a collaboration activity of between NII and GÉANT.

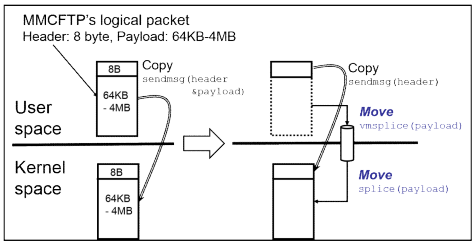


Problem

- By analyzing the results gathered at SC16, we found that system memory bandwidth (BW) is a limiting factor for data transfer speed.
- The zero copy technique seems to be useful to reduce BW usage. However, major Operating Systems (OSs) support sender-zero-copy, but not receiver-zero-copy.

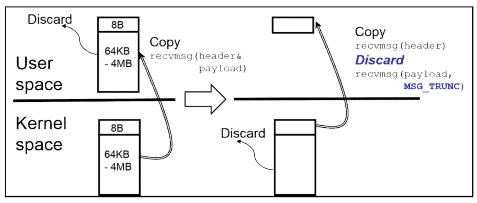
Solution

- Two techniques were introduced to MMCFTP.
- Zero copy (ZC) for sender appl.



- Payload discard (PD) for receiver appl.

(Applicable only to "reception to memory")



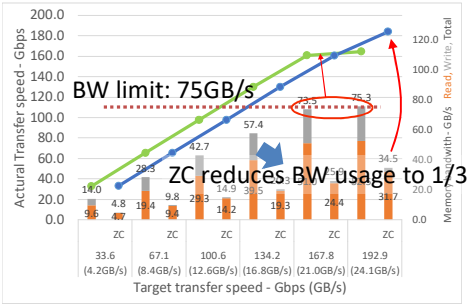
Experiment

Specifications of Sender & Receiver PCs

CPU	Xeon E5-2687W v3 (3.1GHz, 10C/20T) x 2
Memory	512 GB (DDR4-2133) Ideal bandwidth: 17 GB/s x 4 x 2 = 136 GB/s
OS	CentOS 7.3 (1611)
Disk	System: SATA SSD Data: M.2 NVMe SSD 1TB x 8 (RAID0): 8 TB
NIC 100GbE	Mellanox ConnextX-4 VPI x 2 MTU: 9000

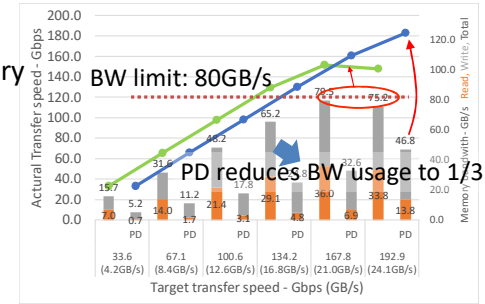
- Sender appl.

Transmission from memory



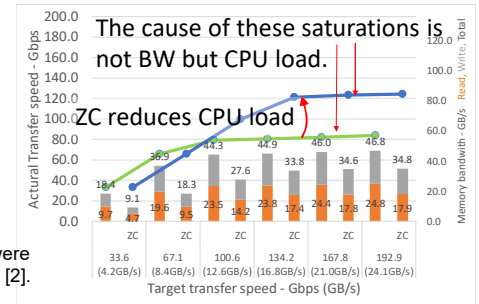
- Receiver appl.

Reception to memory



- Sender appl.

Transmission from disk (SSD array)



Memory bandwidth usages were measured by Intel PCM tools [2].

Conclusion

- System memory bandwidth is a limiting factor for data transfer speed.
- The zero copy technique is effective to reduce memory bandwidth usage and to increase data transfer speed.
- Major OSs don't support a receiver-zero-copy technique, currently. The effect of it can be estimated by the payload discard technique.
- The result of our experiments showed a necessity of receiver-zero-copy for 200 Gbps file transfers.

Contact:
Kenjiro Yamanaka (yamanaka@nii.ac.jp)
Vincenzo Capone (vincenzo.capone@geant.org)

Reference
[1] K. Yamanaka, et al., "A TCP/IP-based constant-bit-rate file transfer protocol and its extension to multipoint data delivery," Fusion Eng. Des., 89 (5), 2014.
[2] Intel Performance Counter Monitor, https://software.intel.com/en-us/articles/intel-performance-counter-monitor