SC22 (NICT booth No.3247) Network Research Exhibition: Demonstration Preliminary Abstract

Full 400G bps E2E DATA/VIDEO transfer across the Trans-Pacific

Satoshi Matsumoto, Hiroshi Kobayashi,
Akira Nakayama, Hajime Takaki, Kazumasa Kobayashi,
Information-technology Promotion Agency, Japan

Abstract

In this demonstration, we transmit various types of data, 8K / 4K / compressed / uncompressed video, and combinations of them from Japan to the SC22 venue (NICT booth No.3247) using the full 400G bps end-to-end bandwidth. We will aggregate six different 100G bps networks over the Trans-Pacific as a 400Gbps network. The traffic consists of a single or a few streams and is therefore susceptible to different influences than general Internet traffic. Each of the six networks used in the Pacific Rim has different transit locations, distances, delays, equipment, and network characteristics. For E2E 400G bps traffic transmission over long distances, it is now difficult to build a single 400G bps network for experimental purposes, and it is generally a combination of multiple 100G bps networks. Such configurations are common, especially across the Trans-Pacific, even when economic costs are taken into account. The purpose of this demonstration is to clarify the problems that can arise from implementing traffic transmission with a single or a small number of high-capacity streams in such realistic situations.

Goals

We design, implement, and demonstrate long-distance Full 400Gbps E2E DATA/VIDEO transmission in Japan. Through such activities, we have confirmed that the throughput impacts negatively each other when a large-volume traffic (e.g. 8K / 4K / compressed / uncompressed video) is transmitted simultaneously over long-distance networks (the networks include virtual networks such as EVPN/VPLS MPLS). To investigate what happens in the real world, we developed a system that can fully exhaust the 400G bps bandwidth with various packets and traffic patterns (400G bps Packet Gen System). The software runs on commodity hardware and thus the system is portable, and can fully test the 400G bps bandwidth at a low cost. It is suitable for our experiment. The experiment aims to identify problems and technical issues that cannot be clarified in Japan domestically. In this experiment, we utilize international networks across the Trans-Pacific from Japan to the SC22 venue (NICT booth No.3247) and transmit various types of data, 8K / 4K / compressed / uncompressed video, and combinations of them, where the traffic utilizes the full 400G bps of the network bandwidth. We practice long-distance, international data transfer experiments in various environments including the Trans-Pacific and Asia in order to identify undetermined problems of future broadcasting technologies.

1. We use six 100G bps networks that have different transits, distances, latencies, and hardware.
2. Using the combination of the networks,
   A) We transfer Full 400G bps of data.
   B) We also transfer video data (8K/4K/compressed/ uncompressed). Bandwidth varies depends on the types of the data.
   C) We transfer a mixture of the 400Gbps and the video traffic.
3. We monitor traffic received at the terminal device and evaluate stability of the effective bandwidth.

Resources

Full bandwidth of 400G bps is required for SCinet WAN. The networks between Tokyo and US West Coast, six 100G bps networks are used from Tokyo and the West Coast across the Trans-Pacific. [Figure 1] Not all bandwidth is available on each line. Therefore, up to 80 Gbps of bandwidth is reserved for each line, which is combined to achieve Full 400 Gbps.

In this experiment, three PCs are used to generate Full 400 Gbps traffic. Two networks are allocated per PC to ensure efficient use of the six 100Gbps networks.

1. PacificWave/TransPAC (Tokyo-Seattle)
2. SingAREN/NSCC (Tokyo-LosAngeles)
3. NII/SINET (Tokyo- LosAngeles)
4. ARENA-PAC/UoH (Tokyo-Guam-Hawaii-LosAngeles)
5. NII/AARNet(Tokyo-Guam-Sydney-Hawaii-Los Angeles)
6. NICT/NSCC/SingAREN/AARNet(Tokyo-HongKong-Singapore-Perth-Sydney-Hawaii-Seattle)

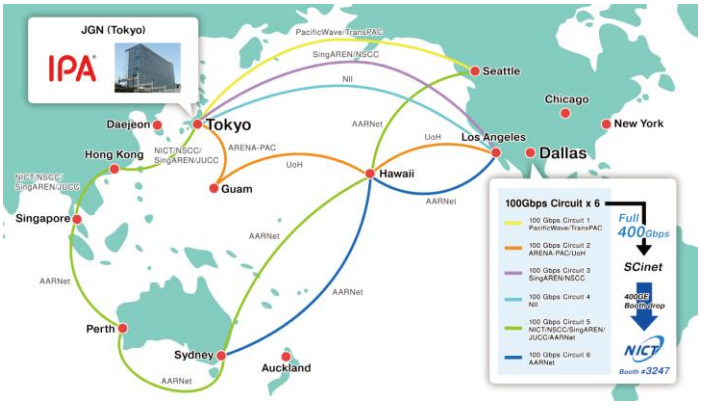

Figure 1 Network overview.

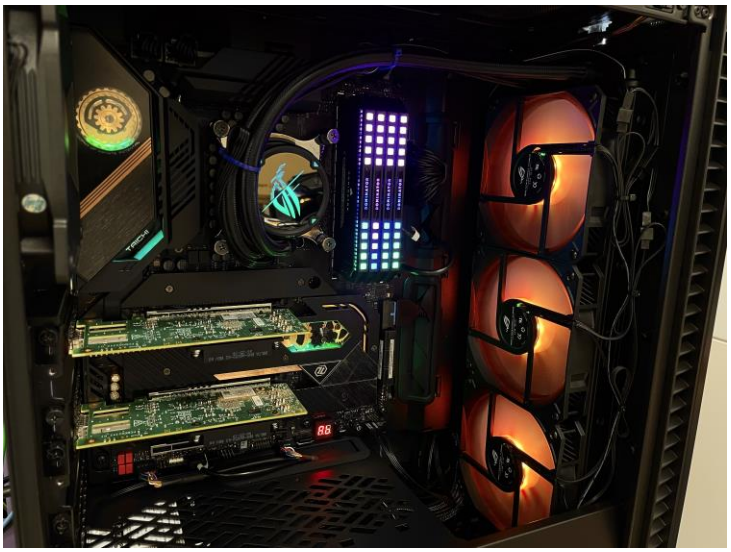

Figure 2 picture of 400G bps Packet Gen System

About 400G bps Packet Gen System

The 400G bps Packet Gen System consists of commodity hardware. [Figure 2] No special ASICs are used. All made from commonly available parts. This makes it relatively easy to create similar systems. We have already made 6 systems (PCs). The specs of the system are as follows. It can be easily scaled up and can be placed at various points on the network. The ability to create more test points allows for flexible measurements and traffic generation.

CPU	Core i9 12900K or 12900KS
RAM	DDR5 5600 16GB * 4
M/B	Intel Z690 Chipset motherboard
NIC	Nvidia ConnectX-6 Dx * 2 (PCIe Gen4 x8 * 2) (200GbE model, MCX623105AN-VDAT)
SSD	512GB NVMe SSD Storage
OS	ubuntu Server 22.04 LTS
Software	DPDK 21.11.1

We use Nvidia ConnectX-6 Network Interface Card. This is because it can use DPDK. It is also the only NIC that supports 200GbE (such as 200G-SR4 transceiver). Currently, the maximum performance is 200Gbps in a single system (due to PCI Express Gen4 bandwidth limitation). The system can generate 200 Gbps stream. By combining two units can fully test the 400G bps bandwidth. Benchmarks currently show that the system can generate 280 M Packet Per Second by a single PC.

Due to NW equipment limitations, we will use two 100 Gbps transceivers (100G-SR4) per system, so the performance of a single PC will be 200Gbps/280Mpps. In this demonstration, three systems were deployed at the senders (at Tokyo Japan) and three at the receivers (at SC22 venue).

In the Future, We plan to replace Nvidia ConnectX-6 Network Interface Card with Nvidia ConnectX-7 Network Interface Card. In that case, it will have the ability to generate 400Gbps in a single stream by a single PC.

Demonstration

In this demonstration, We will transfer Full 400Gbps Traffic across the Trans-Pacifc by using 400Gbps Packet Gen Systems. Six 100 Gbps networks will be combined in various patterns. This allows us to generate and measure traffic with various characteristics.

In addition, as a demonstration at the venue, traffic will be passed through to the NGFW equipment. We will monitor the behavior of the devices, especially when large amounts of traffic are forwarded. Fortigate FG-3700F and Paloalto PA-5450 will be used as NGFW devices. We will also benchmark the different characteristics of these two devices.

We will demonstrate that flexible traffic generation techniques can be used for a variety of purposes, from network measurement to device benchmarking.

Involved Parties

- Satoshi Matsumoto, IPA, st-matsu@ipa.go.jp
- Hiroshi Kobayashi, IPA, hr-koba@ipa.go.jp
- Akira Nakayama, IPA, a-nakaya@ipa.go.jp
- Hajime Takaki, IPA, h-takaki@ipa.go.jp
- Kazumasa Kobayashi, IPA/NICT, kazu-k@udx.icscoe.jp / kazu-k@jgn-x.jp