**SC13 Demonstration**

During the Supercomputing conference 2013 (SC13) in Denver Colorado, Caltech along with international team of researchers designed and demonstrated the first LHC Terabit network Hub in the Caltech booth. The Terabit network hub consisted of four 100G WAN connections and 1Tbps DWDM optical connection between Caltech and Vanderbilt booths. High speed SSD based disk servers with 40GE NICs were used as the end point systems. In addition, for the first time a multipath WAN network controlled by the SDN controller was demonstrated which provided smooth data flows balanced across network paths with varying network speeds. Figure 1 shows the SC13 WAN and show floor network layout.

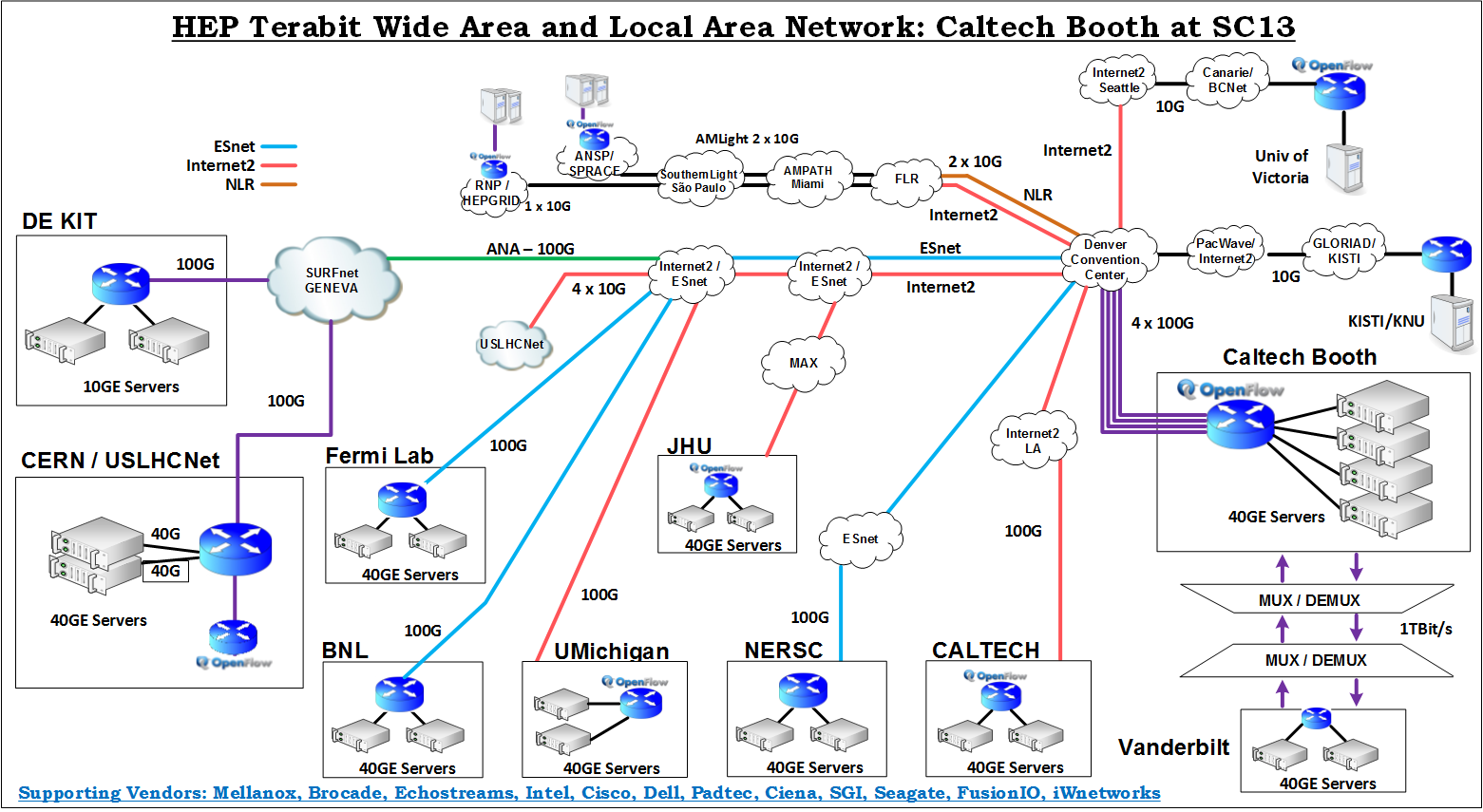


Figure 1: Caltech 2013 - WAN and Inter Booth network layout.

**Network Infrastructure**

The network was designed using high speed optical and Ethernet switching devices. Key hardware components used during the demonstration are described below:

* Mellanox MLXe-16 Ethernet switch with 4 x 100GE, 40 x 40GE ports and 8 x 10GE ports.
* Dell-Force10 Z9000 Ethernet switches (OpenFlow capable).
* Mellanox SX6036 Ethernet switches.
* 40GE Network cards from Mellanox along with active optical cables.
* Padtec optical DWDM equipment for inter-booth data transfer at 1Tbps

**Demonstration Results**

Data was transferred from the show floor to several LHC end sites around the globe. Figure 2 shows both the inter booth and the WAN data transfers. In total, average data transfer rates of 750Gbps with peaks at 850Gbps were achieved.

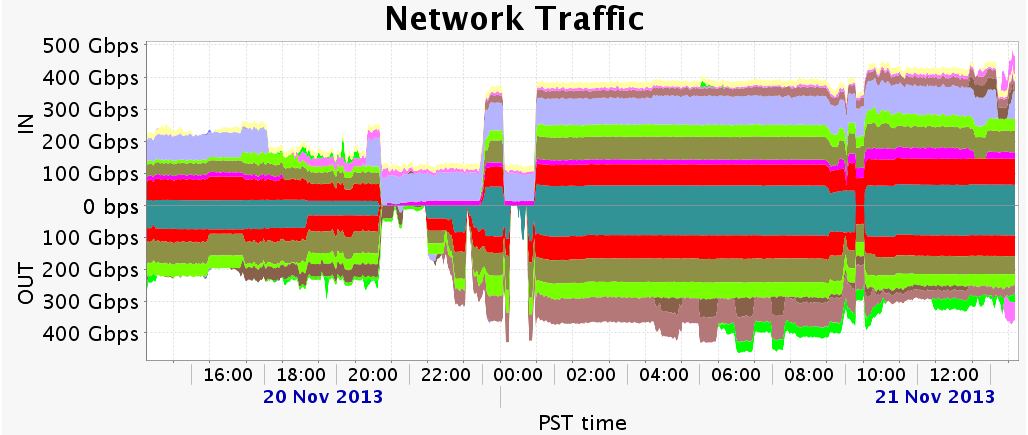


Figure 2: Total traffic flow from Caltech Booth

Following points provide a summary of data transfer results achieved between Caltech booth on the show floor and the various LHC end sites. This summary also includes challenges faced on each of the network path and what techniques were used to resolve them.

* SC13 – DE-KIT (Germany, via ANA transatlantic link)

75Gbps from disk to disk was achieved. DE-KIT used multiple 10GE servers while two servers were used at the show floor.

* SC13 – BNL over ESnet:

80Gbps achieved over two pair of hosts at each end site. Only memory to memory tests were performed due to non-availability of disk based servers

* SC13 – NERSC over ESnet:

Packet loss was encountered initially due to the usage of data center grade Ethernet switches having low buffers in the WAN path. However the path became clean once those switches were removed from the picture. A consistent 90Gbps throughput was achieved by reading from two SSD hosts at NERSC facility sending to a single host at the booth with multiple 40GE network cards.

* SC13 – FNAL over ESnet:

The wide area path showed packet loss. It was not clearly identified which network, router, end hosts or the NIC firmware had issues. A single stream TCP session can reach up to 5Gbps. However a single UDP stream could go up to 15Gbps per flow. Later on, Linux traffic shaper tools 'tc' was used to pace the TCP flows which helped achieving the single stream to reach up to 15Gbps.

However multiple streams was still a problem to FNAL. This seems to indicate that something in the path most probably a router or a switch device has small buffers and thus it is dropping the packets.

* SC13 – Pasadena over Internet2 AL2S:

80Gbps transfer rates by reading from the disks at the show floor and writing on the servers at Caltech Tier2 center. This was a disk to memory transfer because the link was lossy in the other direction.

* SC13 – CERN over ESnet (ANA transatlantic link):

A maximum of 75Gbps memory to memory was achieved by using two servers at CERN and two servers on the show floor. Disk to disk data throughput was 40Gbps.