Efficient "BIG DATA" Transfers among LHC Sites

Azher Mughal, Artur Barczyk

Caltech / USLHCNet

GLIF-2012, Chicago

http://supercomputing.caltech.edu

The Motivation

- **The LHC experiments, with their distributed Computing Models and global program of LHC physics, have a renewed focus on networks, and correspondingly a renewed emphasis on "capacity" and "reliability" of the networks**
- **Networks have seen an exponential growth in capacity**
  - 10X in usage every 47 months in ESnet over 18 years
  - About 6M times capacity growth over 25 years across the Atlantic (LEP3Net in 1985 to USLHCNet as of today)
  - LHC experiments (CMS / ATLAS) are generating large data sets which need to be efficiently transferred to end sites, anywhere in the world
- **A sustained ability to use ever-larger continental and transoceanic networks effectively: high throughput transfers**
- **HEP as a driver of R&E and mission-oriented networks**
- **Testing latest innovations both in terms of software and hardware**

Harvey Newman, Caltech

40GE Server Design Kit

- ✓ **SandyBridge E5 Based Servers:**
  (e.g. SuperMicro X9DRi-F)
  Intel E5-2670 with C2 Stepping
  64 GB of DDR3 RAM

- ✓ **Mellanox VPI CX-3 PCIe Gen3 NIC**

- ✓ **Dell / Mellanox QSFP Active Fiber Cables**

- ✓ **LSI 9266-8i or 9271-8i, 8 port SATA 6G RAID Controller with CacheCade**

- ✓ **OCZ Vertex 4 SSD, 6Gb/s**
  (preferably choose enterprise disks

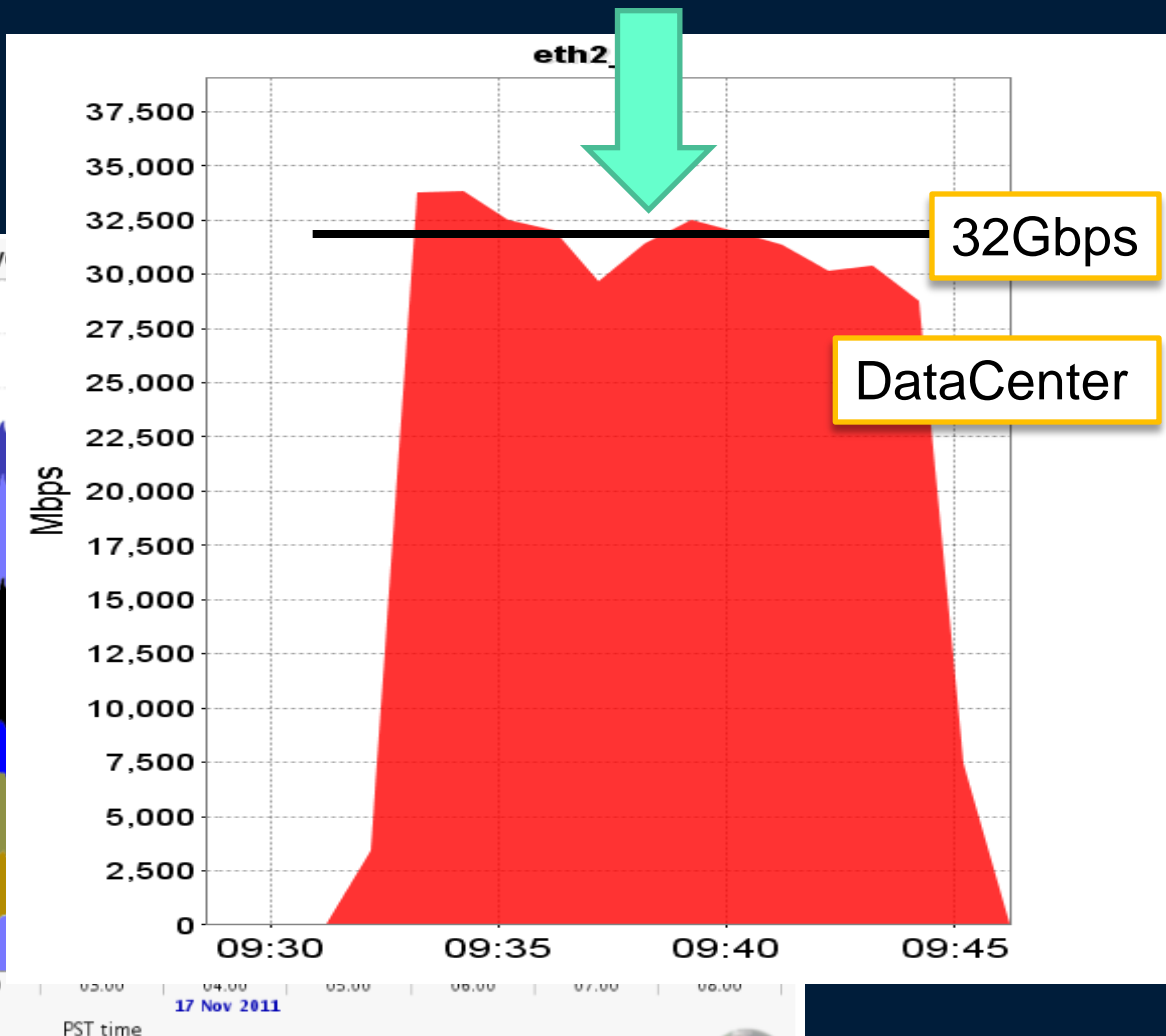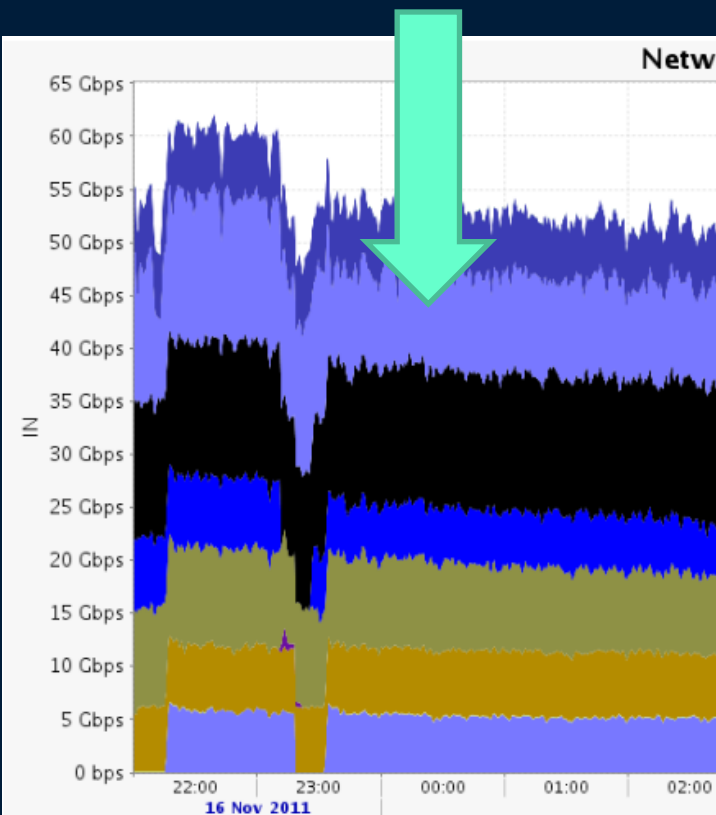- ✓ **Dell – Force10; Z9000 40GE Switch**

Server Cost = ~ $15k
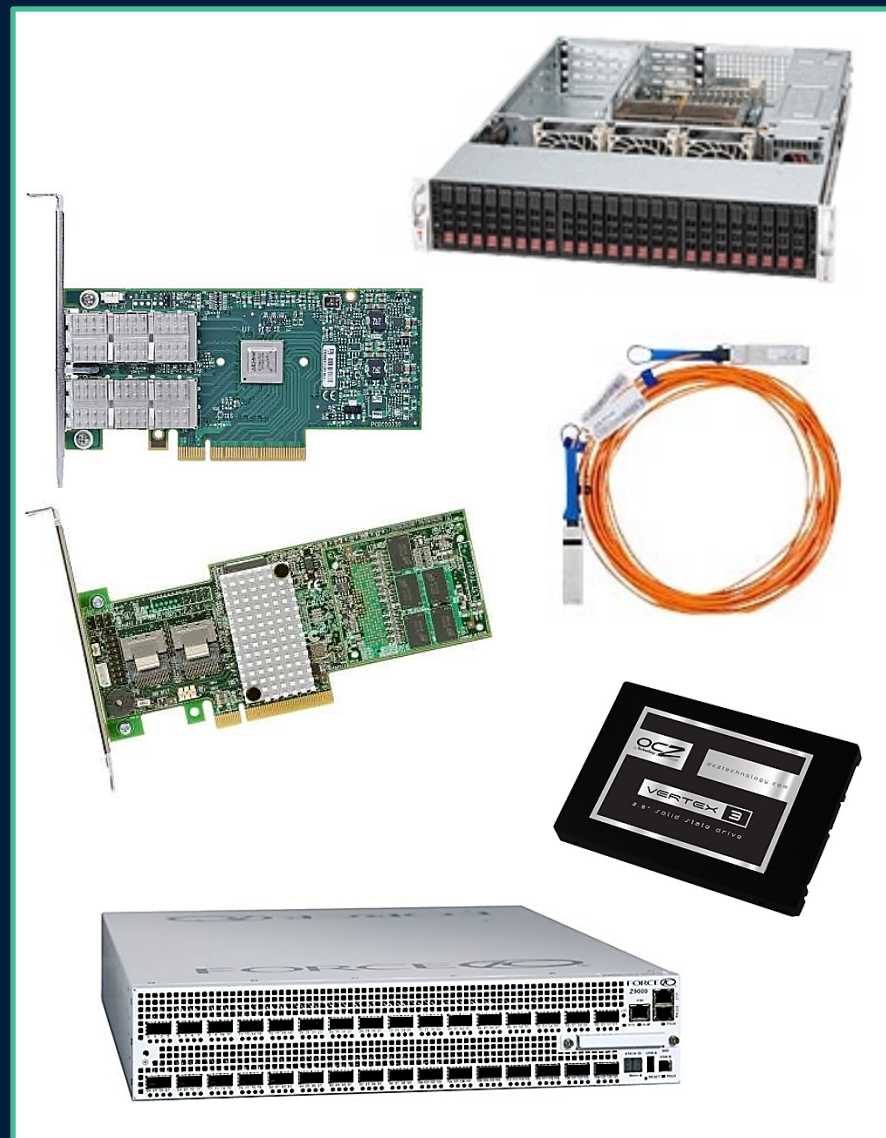
http://supercomputing.caltech.edu/40gekit.html

3

Watch Live Demo this evening @ EVL

40GE line rate at 39.6Gbs

40Gbps

Network Traffic

Write ~ 17.5 Gbps
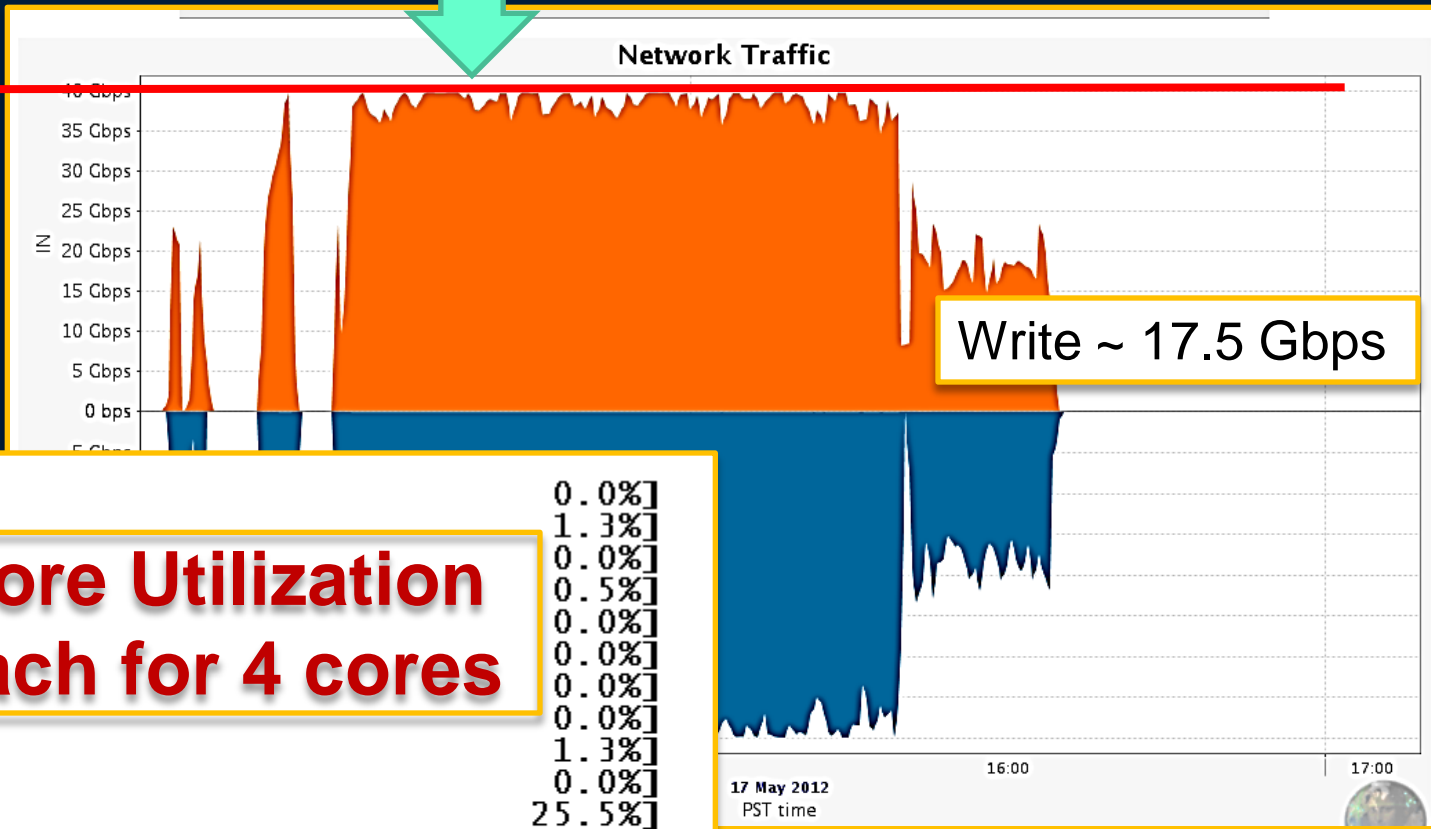
CPU Core Utilization 25% each for 4 cores

```
1  [                           0.0%]
2  [*                          1.3%]
3  [                           0.0%]
4  [*                          0.5%]
5  [                           0.0%]
6  [                           0.0%]
7  [                           0.0%]
8  [                           0.0%]
9  [#*                         1.3%]
10 [                           0.0%]
11 [#************             25.5%]
12 [                           0.0%]
13 [************              25.0%]
14 [#***********              24.7%]
15 [                           0.0%]
16 [************              24.4%]
Mem[||#*            2763/64399MB]
Swp[                   0/66591MB]
```