

Tier1 Hardware Review

Martin Bly

HEPSysMan - RAL, June 2013

Capacity Storage

- Storage in a box
 - Variations on 4U chassis with 24 or 36 bays
 - Single controller:
 - Areca 1280, 3ware 9650SE, Adaptec 5405,52445, LSI 9260-8i
 - 24 or 34 drives:
 - 2 system RAID1, 22 or 32 data RAID6, no hot spare (*)
 - Data drive sizes: 1TB, 2TB, 3TB – WD or Hitachi
 - Mostly 7.2k SATA, one batch SAS
 - Useable sizes: 20, 38, 40, 90TB
 - 2 CPUs or more recently only 1 CPU
 - RAM: 8, 12, 24, 64GB (swap typically 8GB)
 - NIC: 1GbE to 2010, 10GbE from 2011

Capacity Summary

- **2012:** 4.0PB in 46 servers, Viglen and OCF 4U 36-bay chassis, LSI MegaRAID SAS 9260-8i, 9261-8i, 3TB HDDs
- **2011:** 2.6PB in 68 servers, Viglen and Clustervision 4U 24-bay chassis, Adaptec 5404 and LSI MegaRAID SAS 9260-8i, 2TB HDDs
- **2010:** 1.3PB in 36 servers, Streamline and Viglen 4U 24-bay chassis, Adaptec 5404 and 3ware 9650, 2TB HDDs
- **2009:** 3.64PB in 98 servers, Streamline and Viglen 4U 24-bay chassis, Adaptec 5405 and Adaptec 52445, 2TB HDDs
- **2008:** 2.2PB in 110 servers, Streamline + Viglen 4U 24-bay chassis, Areca 1280 and dual 3ware 9650, 1TB HDDs
- **2007:** 1.64PB in 182 servers, Viglen 3U 16-bay chassis, dual 3ware 9650, 750GB HDDs

Other Storage

- Dell R510, PERC H7xx, X520 10GbE NIC
 - 12 x 1TB SATA
 - 12 x 600GB 10k SAS
- iSCSI:
 - 3 x Equallogic PE6510 arrays
 - 2 x Infortrend
- FC
 - 3 x MD3620f
 - 4 x EMC Clariion
 - Several Infortrend + other OEM

Capacity Compute

- Multi-system chassis:
 - SM Twin and Twin²
 - Dell C6100
 - HP Proliant z6000
- CPUs: E56xx, E5-26xx
- RAM: 3 or 4GB/core (*2012 – per thread)
- NIC: 1GbE (2GbE pair in 2012)
- Disk: at least two
 - ~100GB per core or thread
 - SW RAID0

Non-capacity

- Virtualise it!
- R410, R510, R710, R620 + odds and sods
- R710: 96GB RAM, 2 x 6-core CPUs
- R620: 96GB RAM, 2 x 8-core CPUs
- R[5-7][12]0 with 2 or 4 x 10GbE NICs
- Shared storage on Equallogic arrays

Switches

- Extreme routers:
 - 2 x x670V: 48 x 10Gb/s SFP+ + 4 x 40Gb/s QSFP
- Force10:
 - 2 x Z9000: 32 x 40Gb/s QSFP
 - 13 x S4810P: 48 x 10Gb/s SFP+ + 4 x 40Gb/s QSFP
 - 9 x S60: 44 x 1Gb/s + 2 x 10Gb/s SFP+
- Arista: 2 x 7124 24 x 10Gb/s SFP+
- Fujitsu: 1 x 2600G: 26 x 10Gb/s
- Nortel/Avaya:
 - x 56xx, xx=50, 98
 - x 55xx, xx=10, 30
- Netgear
 - 50+ x FS750T2, 2 x FS726T
 - 4 x GS724T, 1 x GS748T
- 3Com: 25+ old 10/100Mb/s

Thoughts...

- No plan to change type of capacity storage this year
 - 4TB drives will make RAID6 lump size ~120TB useable in a 36-bay chassis (single controller)
 - 1 CPU seems to be enough, more RAM good
 - 10GbE NIC OK. May look to double up links for resilience rather than capacity
 -

Thoughts (2)

- CPU requirements continue to be defined by HS06 – unlikely to change any time soon
- Quad-system chassis fine with 2-way systems
 - Have not tested 4-way systems
- Maybe go for 10GbE on-board
- Disk I/O bandwidth a known issue as job counts per server rise
 - More spindles or faster disks
- Power supply spreading – what happens if a rail drops?
 - 3 or 4 x PDUs per rack for proper resilience

Switches

- Variety is the spice of life – until you have to make them work together
- Mesh with aggregation layer and access layer
 - all Force10
 - About to test an Arista 7050 as a possible access layer switch for our mesh
- Big layer 2 domains are a problem (arp, spanning tree...)