

# Western Tier 2 Site Report

by Wei Yang

# CPU Capacity

WT2	CPU Type	GHz	Batch Slots	HS06 / core	Total HS06	HT	Slots / machine
Resource = WT	Intel Xeon(R) X5355	2.66	320	7.99	2,557		8
	Intel Xeon® X5570	2.93	312	15.01	4,683		8
	Intel Xeon® X5650	2.67	456	13.77	6,279		12
	Intel Xeon® X5650	2.67	1,360	9.97	13,559	HT	20
	Intel Xeon® E5-2660	2.2	960	16.05	15,408		16
	Intel(R) Xeon E5-26	2.3	1,056	10	10,528	HT	48
		<b>Totals</b>	<b>4,464</b>		<b>53,015</b>		
			<b>HS06 average</b>	<b>11.88</b>			

- ◆ 22 newly purchased CPU nodes. Each 24 physics cores / 48 logical cores
- ◆ Will run OpenStack VMs
  - OpenStack RDO Kilo on Cent OS 7
  - Bare metal host: Dell M630 48-core/128GB, 10Gbps
  - 8-core 20GB VMs
  - 16-core 40GB VMs

# CPU Capacity, Benchmark on VMs

Host name	cores	RAM (GB)	Hypervisor	HS06	HS06/core	Average	Deviation	Overhead
deft-239-10	8	20	deft-0003	78.22	9.78			
deft-239-11	8	20	deft-0003	77.73	9.72			
deft-239-12	8	20	deft-0003	77.49	9.69			
deft-239-13	8	20	deft-0003	77.02	9.63			
deft-239-14	8	20	deft-0003	86.38	10.80			
deft-239-15	8	20	deft-0003	78.76	9.85			
Summary			Total =	475.60		9.91	0.44	7.25%
deft-239-16	16	40	deft-0005	169.67	10.60			
deft-239-17	16	40	deft-0005	160.80	10.05			
deft-239-18	16	40	deft-0005	161.63	10.10			
Summary			Total =	492.10		10.25	0.31	4.04%
deft-239-21	16	40	deft-0009	264.37	16.52			
deft0001	48	128	N/A	512.80	10.68	10.68		

HEP Spec 2006 measurement

Will also use real ATLAS jobs to compare performance/overhead

# Storage Capacity

- Multi-Tier Storage Managed by Xrootd
- Total Capacity ~ 3980TB
- Of which ~1000TB Shinglet Magnetic Recording Drives
  - » 168x 8TB Seagate SMR drives in 2 trays (and 2 hosts)
  - » Self Managed Drive. Appear as normal drives to HBA but with very different performance characteristics
    - Normal reading performance
    - Good for burst sequence writing.
    - File systems that know the internal of SMR are still under development
- 12TB SSD array for Cache performance studying
- Xrootd file block level cache:
  - » 2x 1TB on borrowed host.
    - Good for software test. Not big enough to function as a cache.
  - » Ordering new hardware: See Andy's talk for hardware detail
  - » Analysis jobs on half of the batch nodes will use this cache. The other half will not.
    - Will use Analytics to studying the performance difference

# Storage Capacity, SMR storage

- SMR storage:
  - » Do as good as we can based on our understanding of SMR technology:
    - Store only large, old (untouched) data, single writing stream
  - » 8+2 RAID 6 (x16)
  - » XFS File system
  - » allocation size = 512MB
    - $512/8=64\text{MB} >$  Seagate SMR drive's band/zone size



# Infrastructure:

- Shared Batch Farm
  - » Use LSF fair share to manage priority. ATLAS ~ 30%
  - » Batch nodes older than 5 yr will be “confiscated”
    - Policy is under discussion. Fair share removed.
    - Computer Center will run it as long possible
- Free power/cooling/rack space/basic network
  - » Each node (of any type) get 1x 10Gbps free.
  - » Addition switch port cost money
- Also free OS installation/maintenance/cyber security
- Multiple 40Gbps backbone LAN
- 100Gbps WAN
  - » Dedicated DTN for ATLAS GridFTP (20 Gbps)
  - » Shared DTN for FAX (20 Gbps)
  - » Batch nodes with outbound TCP connection
- OSG 3.3 with HTCondor CE
  - » Dynamic allocation for S-core, M-core, High Memory, Analysis jobs

