

Resource Deployment Status

Rob Gardner

3-8-10

US ATLAS Facilities Meeting

OSG All Hands - Fermilab





Tracking within Integration Program





























































- Phased program of work for the US ATLAS Facility
- Establishing a baseline set of deliverables that integrates fabric resources, ATLAS software and services, grid level services, operational components, and user tools
- Launched June 2007
 - Now in **Phase 12**
- Focus is primarily on US ATLAS Tier 1 and Tier 2 facilities with some participation from (Tier 2-like or Tier 2-satellite) Tier 3 sites
- <http://www.usatlas.bnl.gov/twiki/bin/view/Admins/IntegrationProgram.html>

Phases by Site

- Upgrades to DQ2 site services, LFC server, equipment procurements, Squid, OSG 1.2, SL5, Throughput

Phase 10





Notation:  completed  work is in progress  defer to next phase  table t









































































Site	DQ2 update	Fabric upgrades	SquidTier2	LFC update	OSG 1.2	NetworkMonitoring	UpgradeSL5
TIER1							
AGLT2							
MWT2_IU							
MWT2_UC							
NET2_BU							
NET2_HU		-					
SWT2_CPB							
SWT2_OU	-		-				
WT2							

Phases by Site

- Upgrades to DQ2 site services, LFC server, equipment procurements, Squid, OSG 1.2, SL5, Throughput

Phase 12

Notation:  completed  work is in progress  not applicable  defer

Site	DQ2 update	Fabric upgrades	SquidTier2	LFC 1.7	OSG 1.2	NetMon	SL5	KitBenchmark
TIER1								
AGLT2								
MWT2_IU								
MWT2_UC								
NET2_BU								
NET2_HU								
SWT2_CPB								
SWT2_OU								
WT2								

2009 Tier 2 WLCG Summary

- Normalized CPU time (HS06-hours) for Jan-Dec 2009
- In terms of WLCG pledge every Tier 2 has exceeded pledged values

Normalised CPU time [units HEPSP06.Hours] by TIER2 and VO

FEDERATION	2008 CPU Pledge (HEPSP06) AVG.		pledge inc. efficiency (HEPSP06-Hrs)	SITE	alice	atlas	cms	lhcb	Total	used as % of pledge
US-AGLT2			AGLT2		101,687,136				101,687,136	
	5183.00	27,273,600	Total		101,687,136				101,687,136	372.84%
			WallClock		112,525,444				112,525,444	
US-MWT2			MWT2_IU		34,513,388				34,513,388	
			MWT2_UC		62,458,812				62,458,812	
	4046.00	21,256,128	Total		96,972,200				96,972,200	456.21%
		WallClock		121,060,088				121,060,088		
US-NET2			BU_ATLAS_Tier2		30,787,264				30,787,264	
	3812.00	20,063,520	Total		30,787,264				30,787,264	153.45%
			WallClock		33,952,876				33,952,876	
US-SWT2			OU_OCHEP_SWT2		12,317,060				12,317,060	
			SWT2_CPB		33,891,776				33,891,776	
			UTA_SWT2		7,229,776				7,229,776	
	6588.00	34,651,584	Total		53,438,612				53,438,612	154.22%
		WallClock		58,953,352		4		58,953,356		
US-WT2			WT2		46,413,896				46,413,896	
	4426.00	23,290,560	Total		46,413,896				46,413,896	199.28%
			WallClock		52,160,744				52,160,744	

Tier 2 Reliability

- Has not been an issue meeting WLCG reliability and availability
- Report from Oct 09 - Dec 09

Federation	Site	Phy. CPU	Log. CPU	KSI2K	Reliability	Availability	Unknown	Reliability History		
								Oct-09	Nov-09	Dec-09
US-AGLT2 (USA, Great Lakes ATLAS T2)										
	AGLT2	N/A	N/A	4,813	100 %	100 %	4 %	100 %	99 %	100 %
US-MWT2 (USA, Midwest ATLAS T2)										
	MWT2_IU	N/A	N/A	3,276	97 %	93 %	3 %	100 %	98 %	93 %
	MWT2_UC	N/A	N/A	3,276	93 %	92 %	14 %	100 %	100 %	99 %
US-NET2 (USA, Northeast ATLAS T2)										
	BU_ATLAS_Tier2	N/A	N/A	1,910	93 %	91 %	4 %	99 %	100 %	100 %
US-SWT2 (USA, Southwest ATLAS T2)										
	OU_OCHEP_SWT2	N/A	N/A	464	100 %	100 %	3 %	100 %	100 %	100 %
	SWT2_CPB	N/A	N/A	1,383	92 %	92 %	17 %	99 %	91 %	100 %
	UTA_SWT2	N/A	N/A	493	97 %	97 %	18 %	100 %	N/A	100 %
US-WT2 (USA, SLAC ATLAS T2)										
	WT2	N/A	N/A	1,202	97 %	88 %	2 %	100 %	100 %	95 %

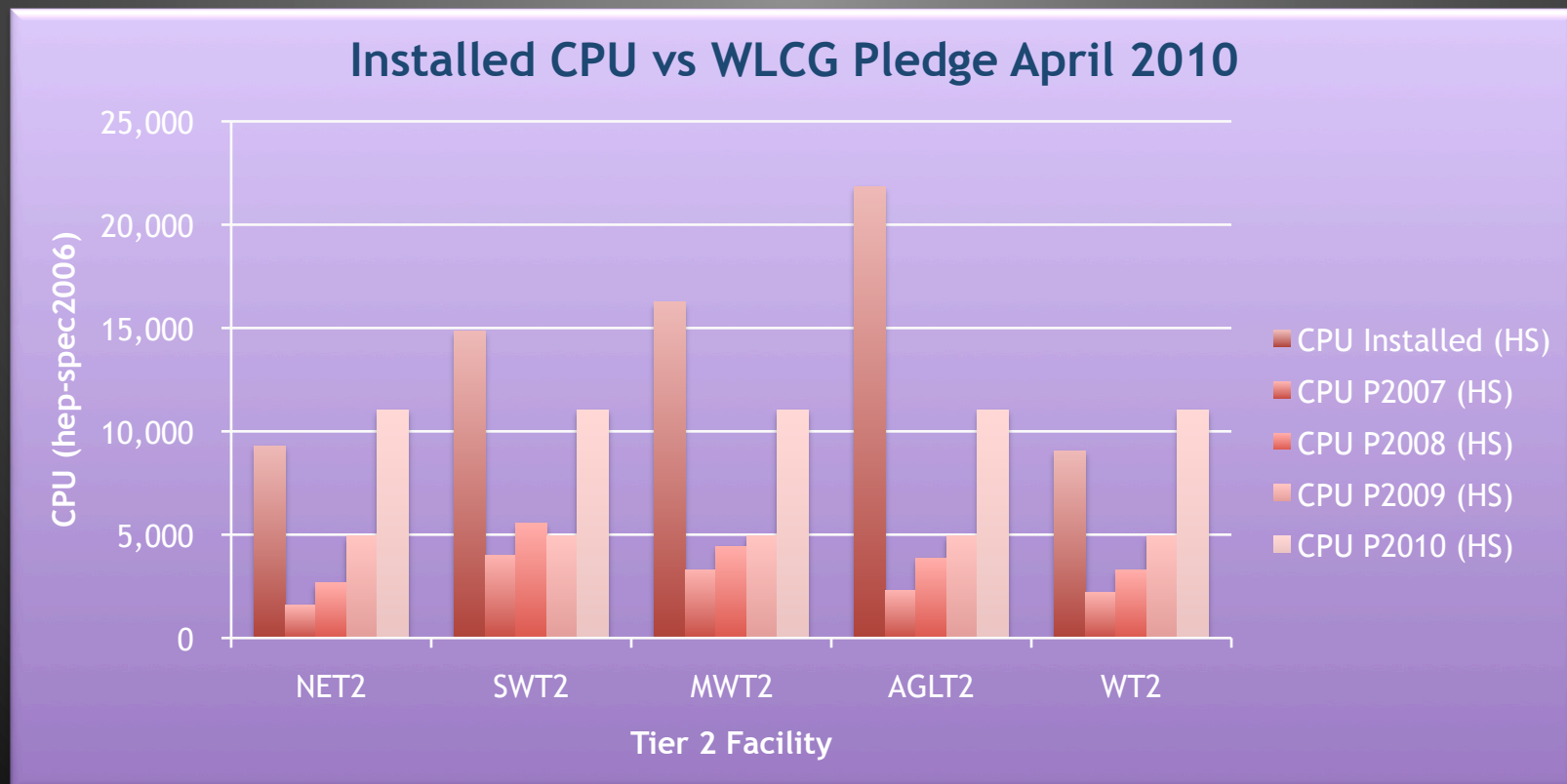
US ATLAS Facility Resources

- CPU resources at Tier 2 and Tier 1 - 2010
- Disk resources at Tier2 and Tier 1 - 2010
- CPU resources at Tier 2 and Tier 1 - up to 2013
- Disk resources at Tier2 and Tier 1 - up to 2013
- Questions about configuration
 - Job slots: analysis versus production
 - Disk: space token allocation
- Capacity summary:
<http://www.usatlas.bnl.gov/twiki/bin/view/Admins/CapacitySummary.html> .

Tier 2 CPU versus Pledge to 2010**

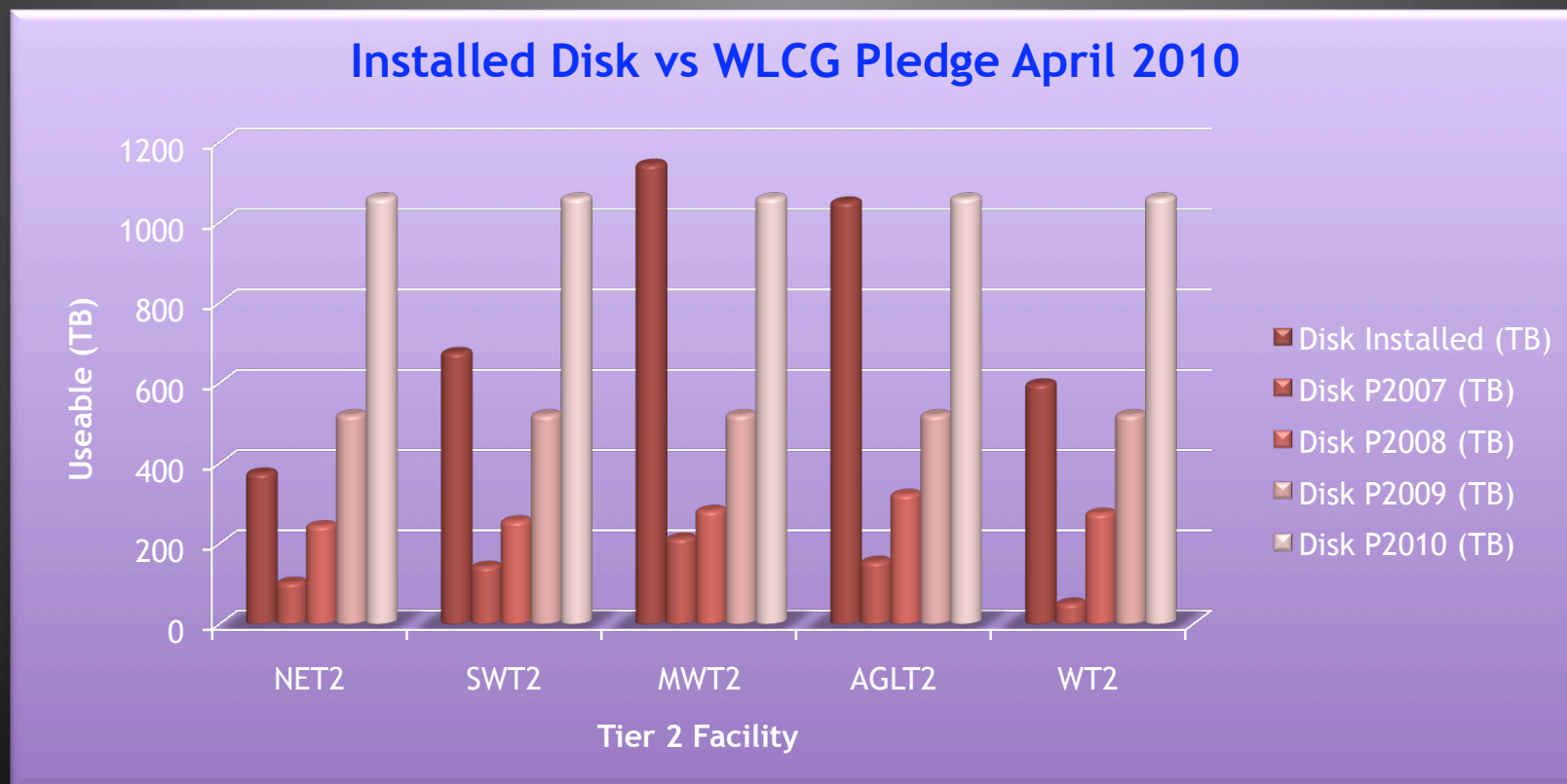
- Each Tier 2 is pledged to deliver ~11K HS06 in '10
 - US ATLAS req. in 2010 is 59.40 kHS06 (pledged 55.20, 72.5 installed)
 - Note HS06 benchmarks are ~7 to ~15 for all processors
 - Site averages are ~8-9
 - T2 min(9000) max(21000)
 - 4/5 Tier 2s well above 2010 pledge

****NET2: +1K cores**



Tier 2 Disk versus Pledge to 2010

- Large disk deployments coming online but deploying adequate storage still a challenge at the Tier 2
 - US ATLAS req. in 2010 is 6.71 PB (pledged 5.70 PB, 3.8 PB installed)
 - Expect larger deployments in later this summer with FY10 funding

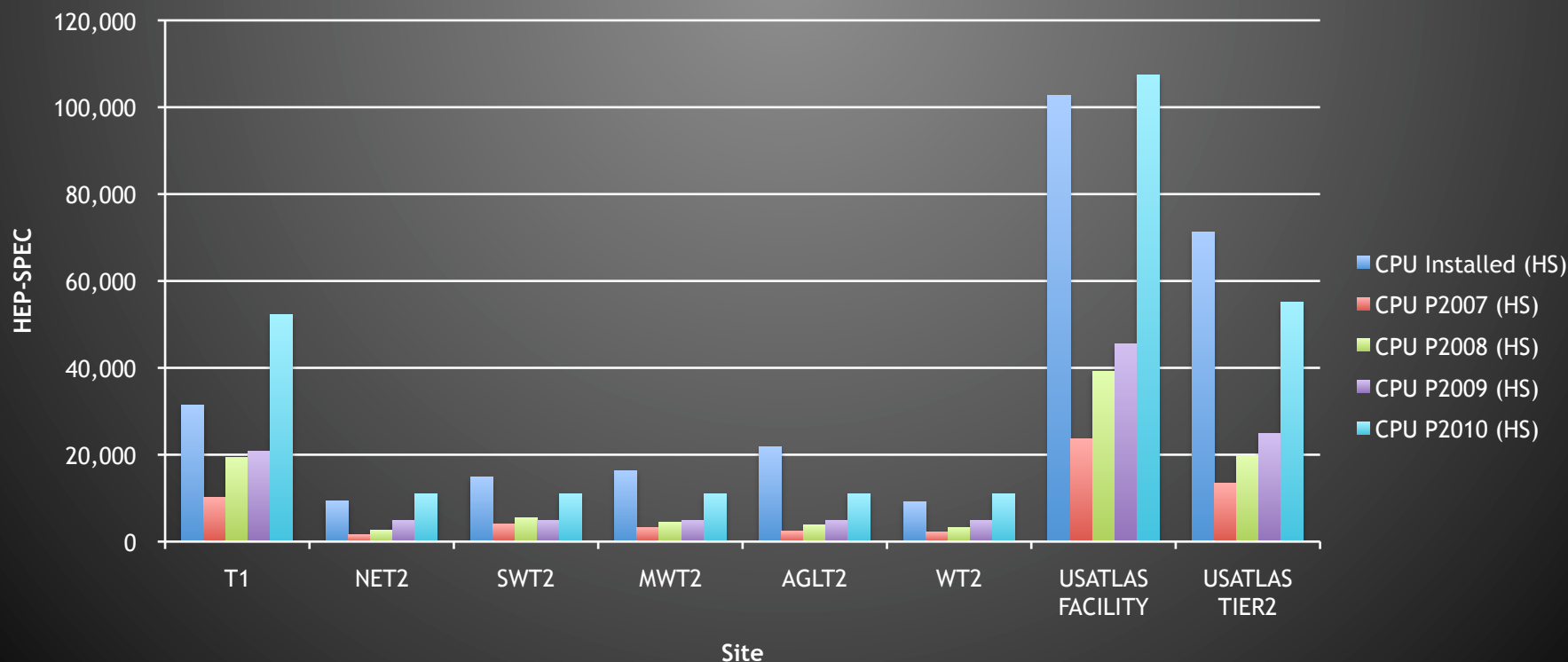


US ATLAS Total (T1+T2) CPU vs Pledge 2010**

- Installed capacity has now exceeded 100 kHS06
 - Tier-1 CPU: US share in 2010 is 43.47 kHS06 (pledged 49.68), 53.59 kHS06 needed in 2011
 - Tier-2 CPU: US share in 2010 is 59.40 kHS06 (pledged 55.20), 70.40 kHS06 needed in 2011
 - Very close to 2010 pledge overall

****NET2: +1K cores**

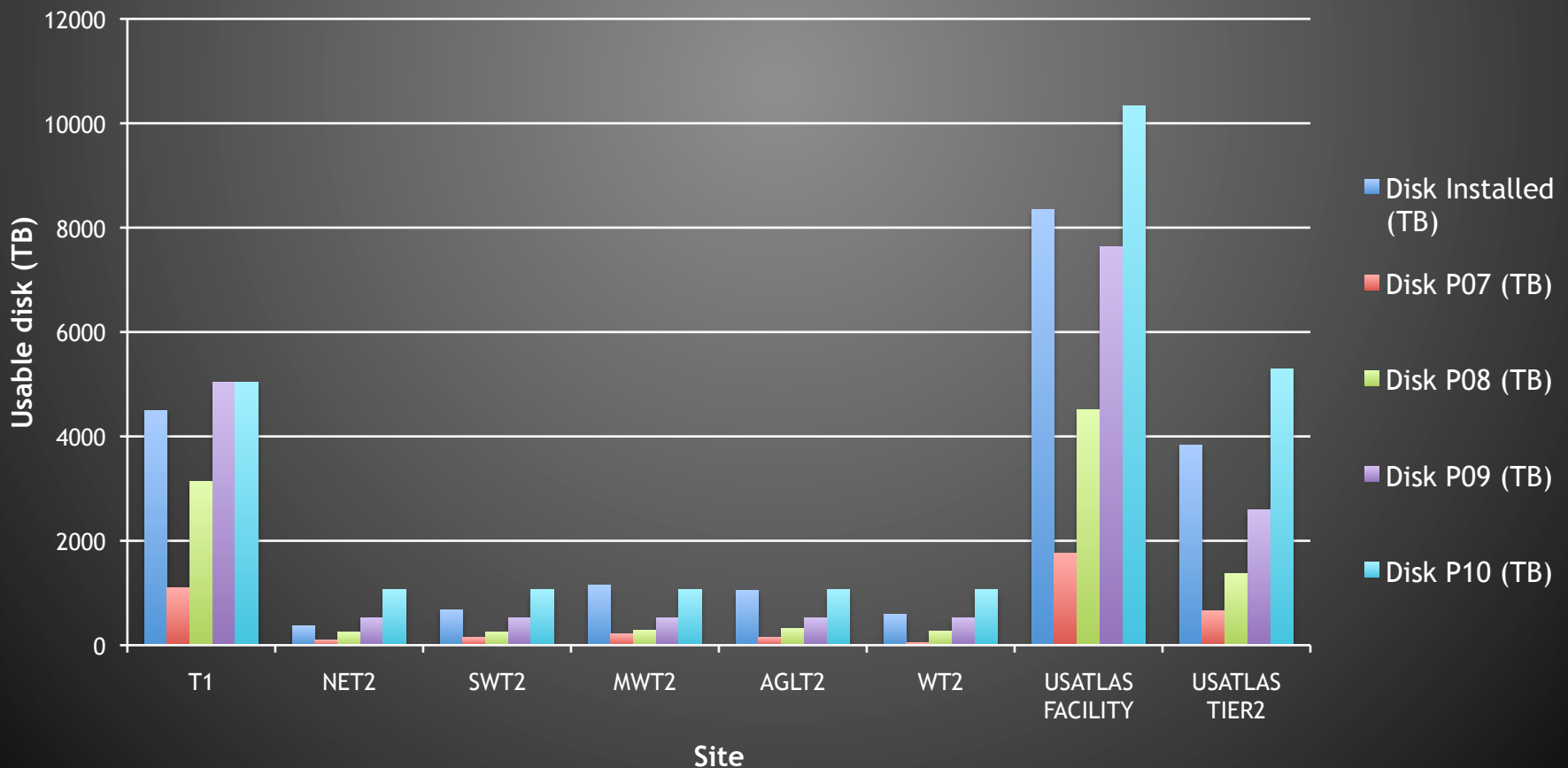
Installed CPU vs WLCG Pledge



US ATLAS Total (T1+T2) Disk vs Pledge 2010

- 8.3 PB total (T1: 4.5 PB T2: 3.8 PB)
 - Tier-1 Disk: US share in 2010 is 6.72 PB (pledged 5.04 PB), 10.07 PB needed in 2011
 - Tier-2 Disk: US share in 2010 is 6.71 PB (pledged 5.70 PB), 10.89 PB needed in 2011

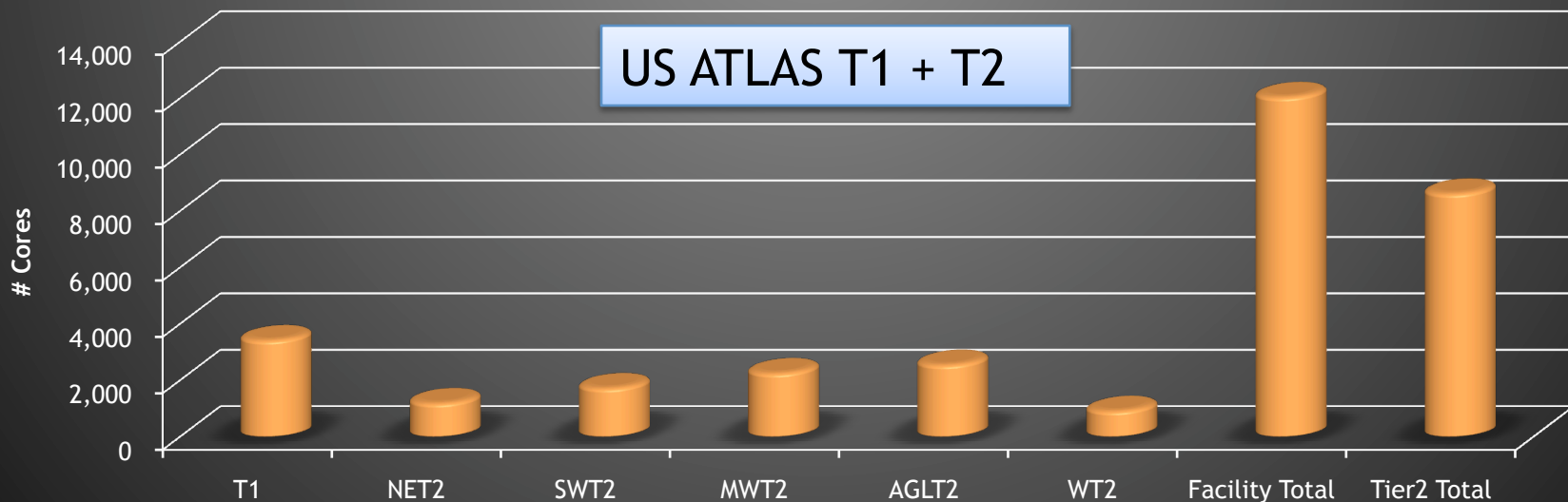
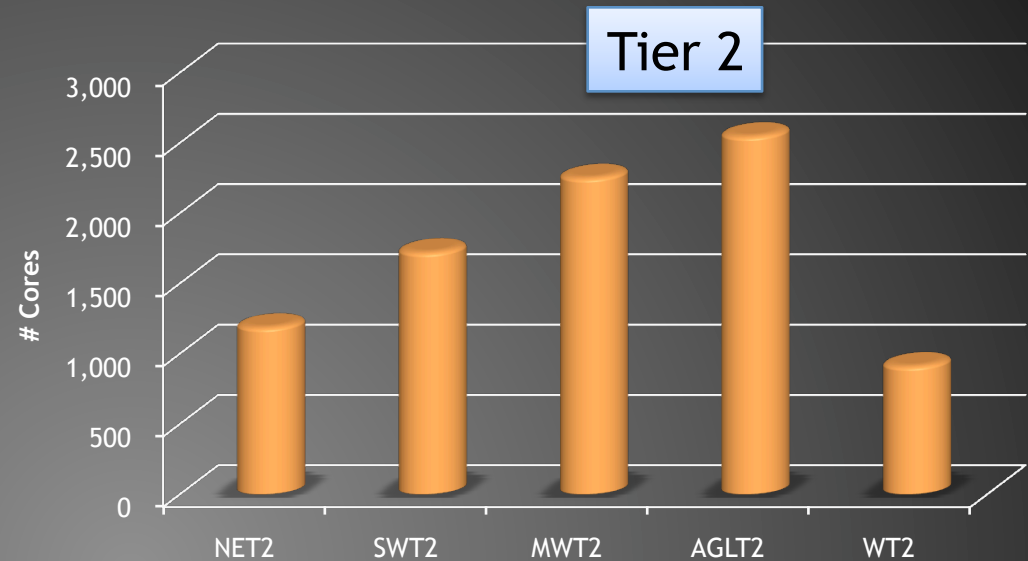
Installed Disk vs WLCG Pledge



Cores**

**NET2: +1K cores

- Number of cores in facility:
- Determining allocation production vs. analysis an issues



Installed Facility Summary

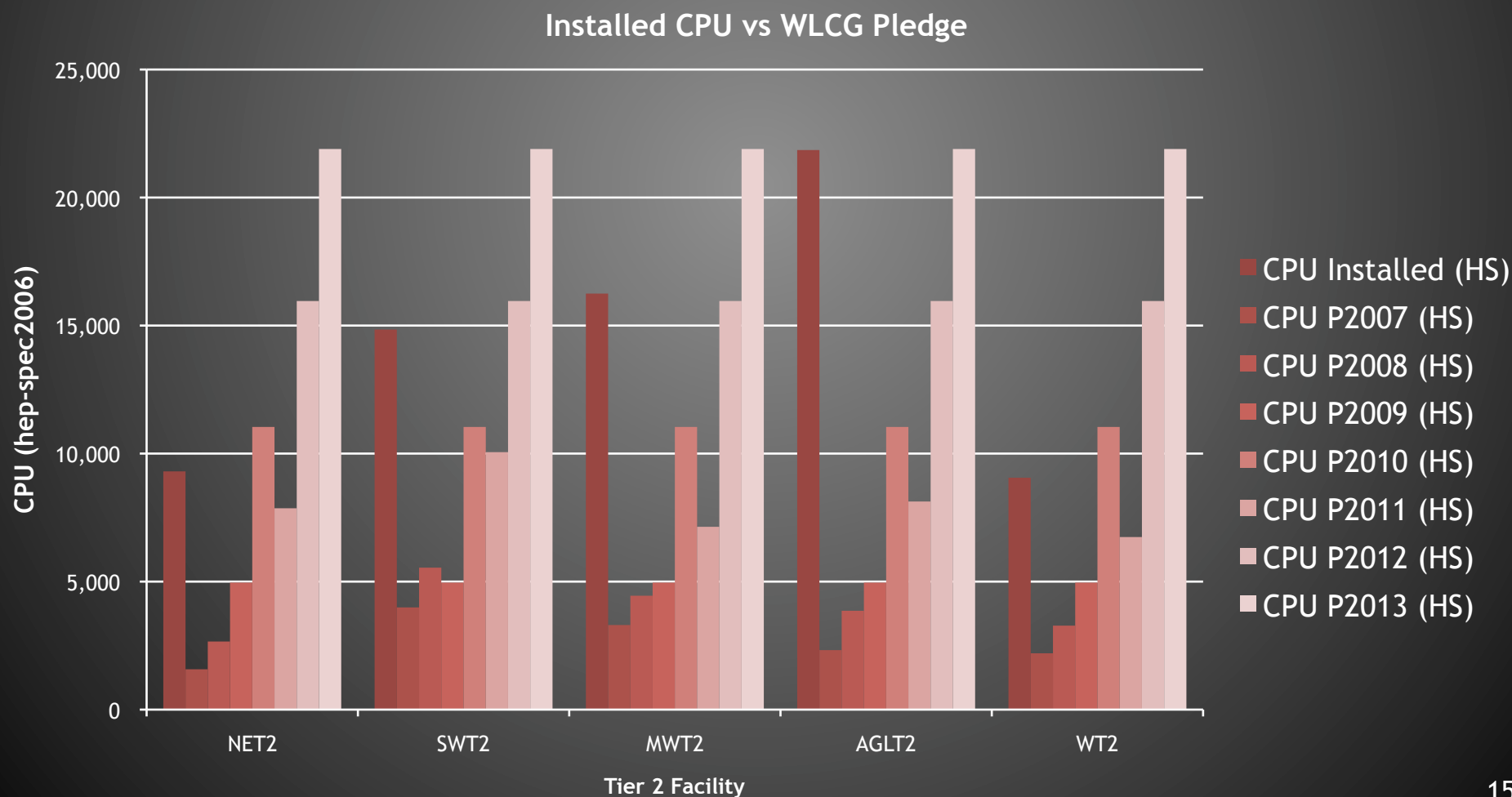
(ending 3/31/10)**

- **US ATLAS Facility**
 - 11,642 cores 100,268 HS06
 - 8.3 PB
- **Tier 1 Center**
 - 3,420 cores 31,496 HS06
 - 4.5 PB
- **Tier2 Centers**
 - 8,622 cores 72,569 HS06
 - 3.84 PB

**NET2: +1K cores

Forward Pledges Tier 2 CPU

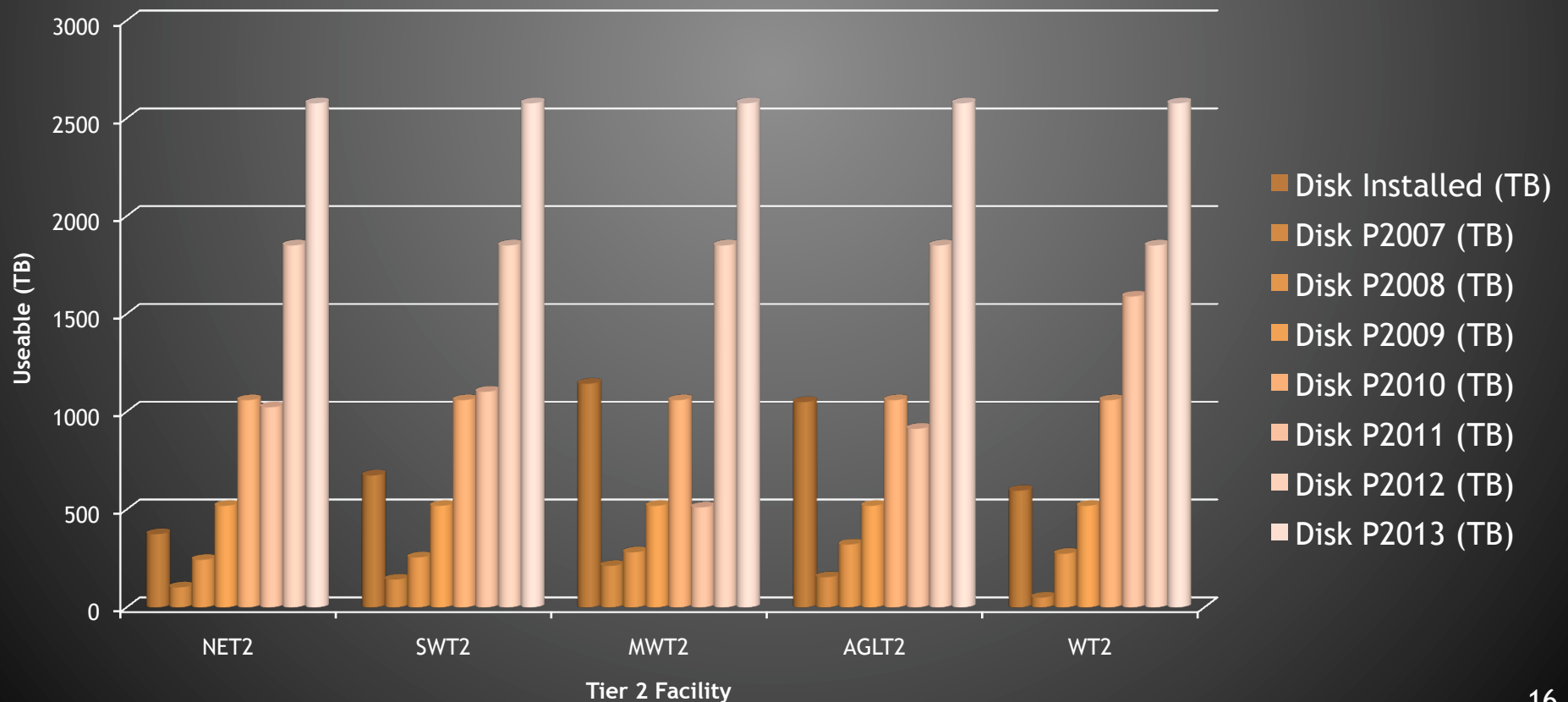
- Tier 2's normalize in 2012 in CPU (2011 pledge is low)
- Tier 2 CPU: US share in 2010 is 59.40 kHS06 (pledged 55.20), 70.40 kHS06 needed in 2011 (14 kHS06 per T2)



Forward Pledges Tier 2 Disk

- Pledges among Tier 2's normalize in 2012 (2011 pledge low)
- Tier-2 Disk: US share in 2010 is 6.71 PB (pledged 5.70 PB), 10.89 PB needed in 2011 (2.2 PB / T2)

Installed Disk vs WLCG Pledge



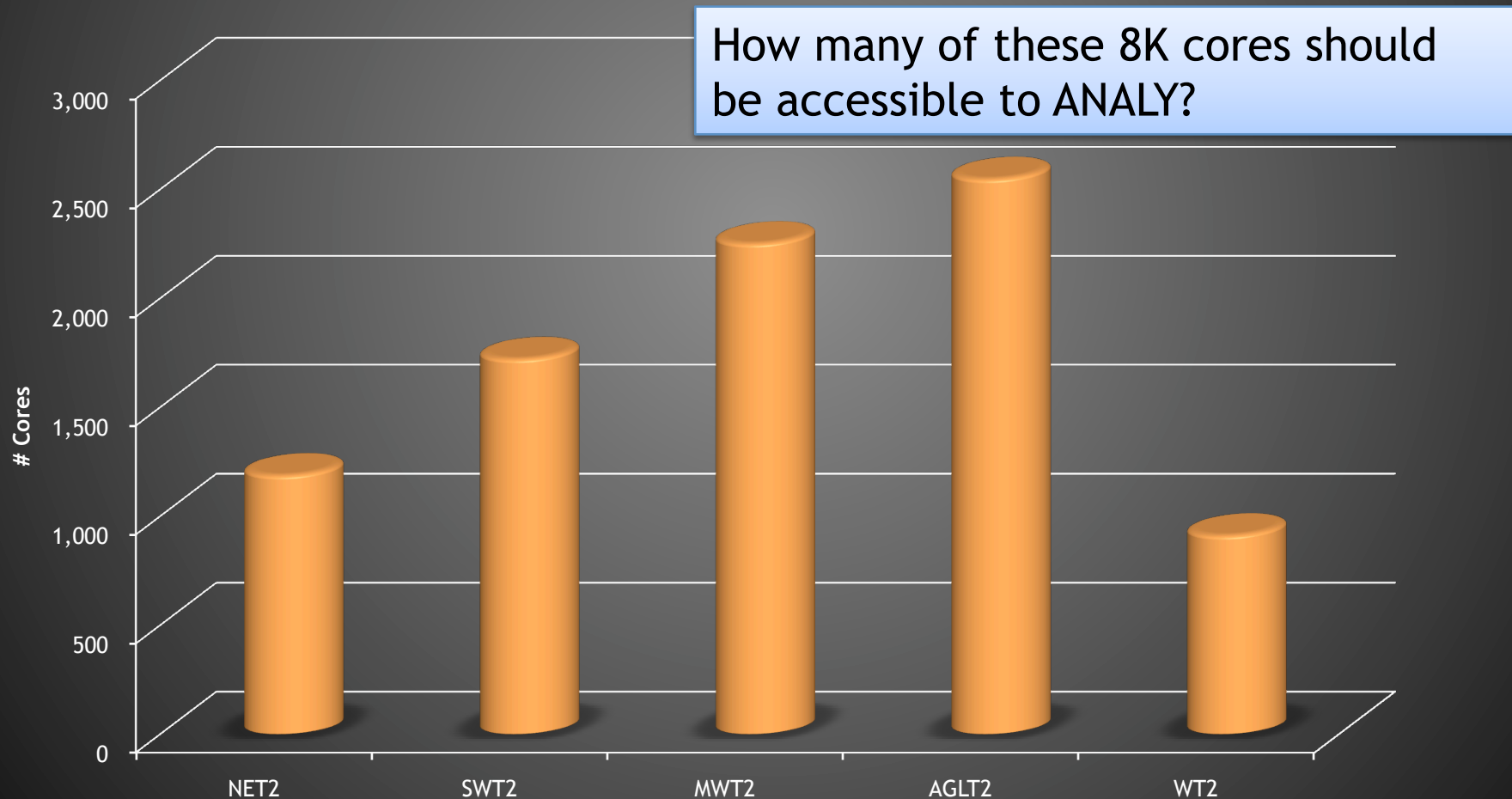
Analysis Resources ...

Analy queue profiles

Queue	Storage system	Access/Mover	Min job slots	Max job slots	Queue policy
ANALY_BNL_ATLAS_1	dCache	dcap, dccp, lcg-cp	250	420	90 minutes walltime limit, highest priority
ANALY_LONG_BNL_ATLAS	dCache	dcap, dccp, lcg-cp	250	420	no walltime limit, highest priority, no production jobs run here
ANALY_AGLT2	dCache	dcap, dccp, lcg-cp	32	528	Analysis scheduled higher than production; 3-day wall time limit; 4GB max image size (50/50 RAM/Swap); 25GB max disk usage
ANALY_MWT2	dCache	dcap, dccp, lcg-cp	(0) 400	1200	Analysis scheduled higher than production; 400 cap when production; 1200 when no production jobs. 3 day wall time limit.
ANALY_NET2	GPFS	lsm	280	600	no limit
ANALY_SWT2_CPB	xrootd	xcp	20	200	Analysis jobs have higher priority than production; 41 Hour walltime limit
ANALY_OU_OCHEP_SWT2 (soon)	Lustre	cp, lcg-cp	16	256	no limit
ANALY_SLAC	xrootd	xcp	0	~1000	48% fairshare for analysis, 32% for production

Analysis Queues

- Analysis queue capacities?
- Historically site managers determine based on performance and reliability



Summary, Concerns, Questions

- In Phase 12 of Integration Program we have made many upgrades to the facility
 - SL5, OSG, new equipment, I2 tools releases, ...
- Some equipment nearly deployed, online imminently to support 2010-2011 run
- Getting all sites deployed with pledged storage for remainder of 2010 and 2011 (looking forward) will be a big challenge
- Supporting analysis at scale is biggest concern
 - Performance (storage access)
 - Size of ANALY queue