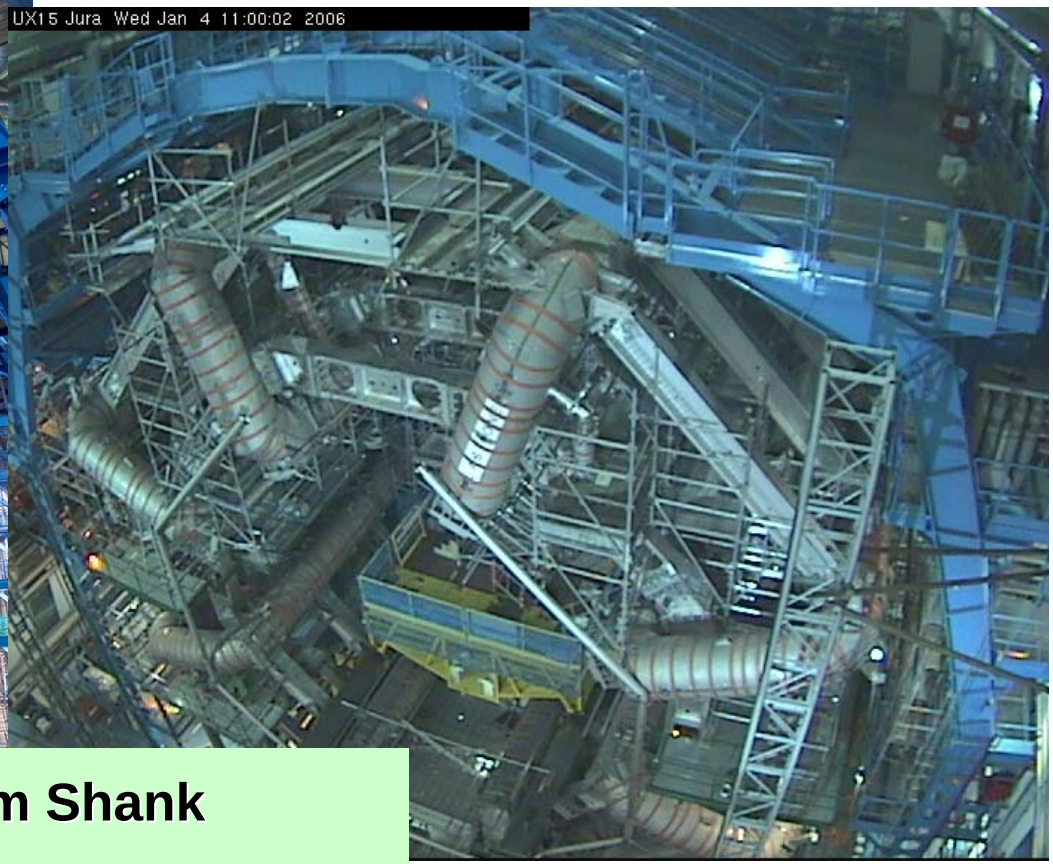
US Tier 2 issues



Jim Shank

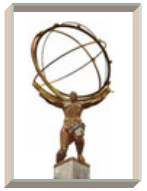
Boston University

U.S. ATLAS Tier 2 Meeting

Harvard University

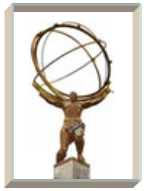
17-18 August, 2006

Overview



- LHC Schedule news and implications
- Latest funding news
  - Upcoming reviews:
    - Computing Advisory Panel (CAP) 18-19 Dec @ BNL
    - DOE/NSF review of LHC S&C 17-19 Jan, 2007 @ UTA
- Goals for this workshop
  - T2 Planning Wiki
  - Action Items from last T2 meeting in Boston
  - Update hardware ramp
  - Reconcile with the ATLAS Computing Model

LHC Schedule



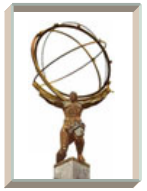
year	*energy*	*luminosity*	*physics beam time*
2007	**450+450 GeV**	**$5x10^{30}$**	**protons - 26 days at 30% overall efficiency ➔ $0.7*10^6$ seconds**
2008	**7+7 TeV**	**$0.5x10^{33}$**	**protons - starting beginning July ➔ $4*10^6$ seconds ions - end of run - 5 days at 50% overall efficiency ➔ $0.2*10^6$ seconds**
2009	**7+7 TeV**	**$1x10^{33}$**	**protons:50% better than 2008 ➔ $6*10^6$ seconds ions: 20 days of beam at 50% efficiency ➔10^6 seconds**
2010	**7+7 TeV**	**$1x10^{34}$**	**TDR targets: protons: ➔ 10^7 seconds ions: ➔ $2*10^6$ seconds**

Funding

U.S. ATLAS Research Program Target Chart (AYk$s)

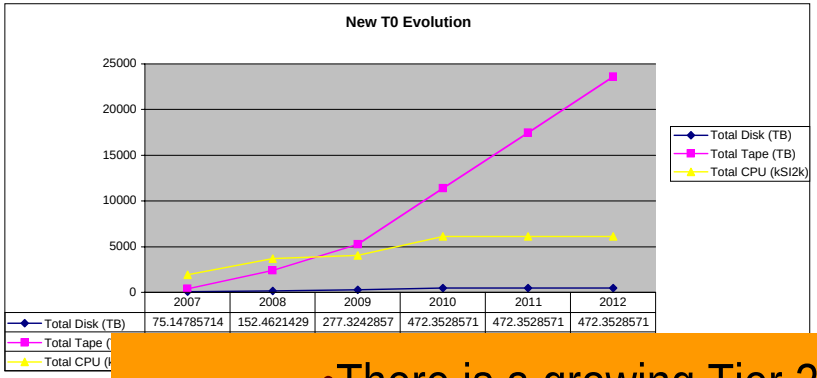
Category	WBS	Description	FY07	FY08	FY09	FY10	FY11
Computing	2.2	Software	5,268	5,199	5,439	5,535	5,790
	2.3	Facilities	9,965	11,928	12,611	14,504	12,587
	2.0	**Total Computing**	**15,233**	**17,127**	**18,050**	**20,039**	**18,377**
M&O	3.1	Silicon	2,095	1,642	1,642	1,642	1,638
	3.2	TRT	547	445	440	465	440
	3.3	Liquid Argon	2,716	2,093	2,161	2,162	2,227
	3.4	Tile	1,531	768	785	804	825
	3.5	Endcap Muon	2,714	1,032	880	895	919
	3.6	Trigger/DAQ	1,814	765	556	571	586
		**Common Funds Cat. B (included in subsystems above)	1,066	1,515	1,609	1,659	1,589
	3.7	Common Funds Cat. A	1,630	1,700	1,877	2,141	2,198
	3.7.2	C&I	1,256	-	-	-	-
	3.8	Outreach	50	53	54	55	57
	3.9	Program Management	600	837	866	895	924
	3.10	Technical Coordination	1,137	463	488	514	540
	3.0	**Total M&O**	**16,090**	**9,799**	**9,749**	**10,145**	**10,354**
Upgrade R&D	4.1	Silicon Upgrade R&D	2,015	2,095	1,878	1,885	1,859
	4.2	Liquid Argon Upgrade R&D	1,053	1,142	1,360	1,353	1,378
	4.5	Muon Upgrade R&D	10	-	-	-	-
	4.0	**Total Upgrades**	**3,078**	**3,237**	**3,238**	**3,238**	**3,237**
Subtotal (Comput.+ M&O + Upgrades)		**Subtotal U.S. ATLAS Research Program**	**34,401**	**30,163**	**31,036**	**33,422**	**31,968**
Management Reserve		Management Reserve (%)	9%	10%	10%	8%	15%
		Management Reserve	**3,589**	**3,437**	**3,464**	**3,078**	**5,532**
Total U.S. ATLAS Research Program		**Total U.S. ATLAS Research Program**	**37,991**	**33,600**	**34,500**	**36,500**	**37,500**

ATLAS Requirements start 2008, 2010

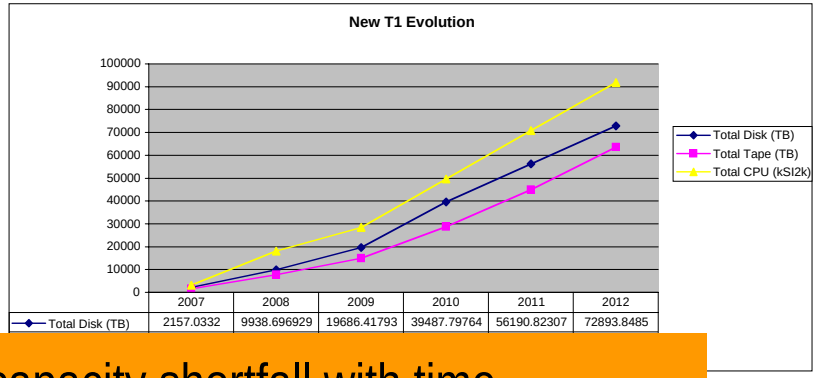


	CPU (MSi2k)		Disk (PB)		Tape (PB)	
	2008	2010	2008	2010	2008	2010
Tier-0	3.7	6.1	0.15	0.5	2.4	11.4
CERN Analysis Facility	2.1	4.6	1.0	2.8	0.4	1.0
Sum of Tier-1s	18.1	50	10	40	7.7	28.7
Sum of Tier-2s	17.5	51.5	7.7	22.1		
Total	41.4	112.2	18.9	65.4	10.5	41.1

Evolution

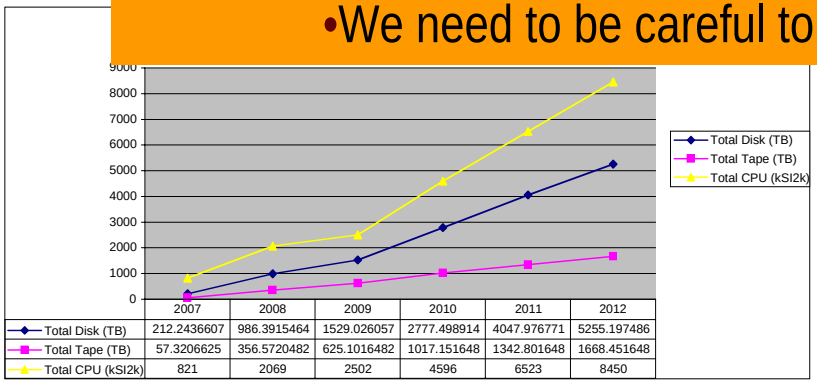


	2007	2008	2009	2010	2011	2012
Total Disk (TB)	75.14785714	152.4621429	277.3242857	472.3528571	472.3528571	472.3528571
Total Tape (						
Total CPU (						

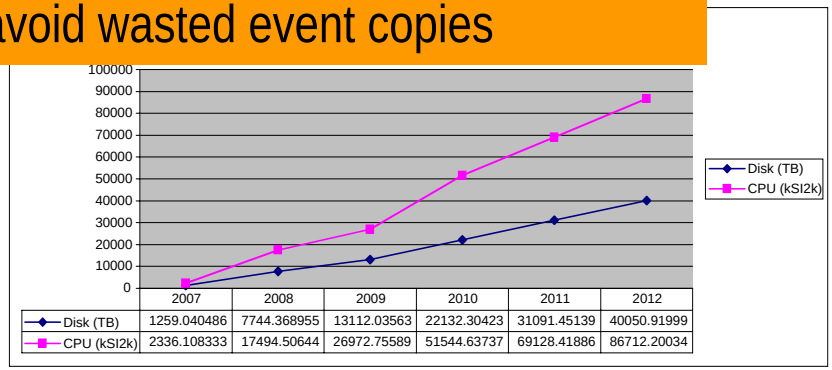


	2007	2008	2009	2010	2011	2012
Total Disk (TB)	2157.0332	9938.696929	19686.41793	39487.79764	56190.82307	72893.8485

- There is a growing Tier 2 capacity shortfall with time
- We need to be careful to avoid wasted event copies

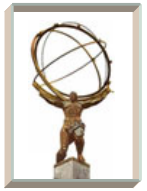


	2007	2008	2009	2010	2011	2012
Total Disk (TB)	212.2436607	986.3915464	1529.026057	2777.498914	4047.976771	5255.197486
Total Tape (TB)	57.3206625	356.5720482	625.1016482	1017.151648	1342.801648	1668.451648
Total CPU (kSI2k)	821	2069	2502	4596	6523	8450



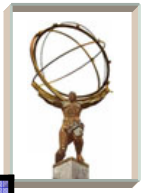
	2007	2008	2009	2010	2011	2012
Disk (TB)	1259.040486	7744.368955	13112.03563	22132.30423	31091.45139	40050.91999
CPU (kSI2k)	2336.108333	17494.50644	26972.75589	51544.63737	69128.41886	86712.20034

Resource Allocation Committee



- Newly formed committee to coordinate U.S. resources (CPU/disk)
  - Prioritize: Physics simulation, analysis, calibration
- Committee members:
  - Physics Adviser (I. Hinchliffe)
  - Chair of the ASG (S. Willocq)
  - Facilities Manager (B. Gibbard)
  - Tier 2 (R. Gardner)
  - The US Production manager (K. De)
  - The EPM (J. Shank)
  - The Software Manager (S. Rajagopalan)
  - Physics users (B. Mellado)
  - Calibration users (B. Zhou)
- First meeting 25 Jan, 2006
  - http://www.usatlas.bnl.gov/twiki/bin/view/AtlasSoftware/ResourceAllocationCommittee

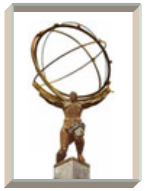
Projected T2 Hardware Growth (dedicated to ATLAS)



Tier 2 Center	2005	2006	2007	2008	2009
Boston/Harvard					
CPU (kSi2k)	210	350	730	1,090	1,600
Disk (TB)	40	170	370	480	630
Southwest					
CPU (kSi2k)	500	900	1,500	1,700	2,100
Disk (TB)	60	200	380	540	700
Midwest					
CPU (kSi2k)	360	510	900	1,100	1,300
Disk (TB)	50	130	260	465	790

- Assumes Moore's law doubling of CPU (3 yrs) and disk capacity (1.5 yrs) at constant cost
- Assumes replacement of hardware every 3 years

Reconciling Requests with Target



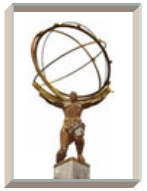
- As shown in Feb, we still have about a $3M difference in requests over our targets in 2008 and beyond.
  - As shown then, we put about $1M for software and $2M for Facilities as requests to the Management Reserve (MR).
- We are still working on what our actual 2006 spending will be
  - Late hires at T1, late purchases at T2
  - Should allow us to get where we want on 2007 with minimal call on MR (< $500k ?)
- 2008 and beyond still a problem
  - We are working closely with the ATLAS Computing Model Group to understand our T1/T2 needs out to 2012
  - New LHC running assumptions COULD lead to some savings in a later ramp of hardware
  - Software: Emphasis on user support, if not enough MR, we will have to cut some Core effort

Need to Ramp up the T2 hardware

- The Research Program has a large FY06 Rollover
- We are begging for more money for FY07, but the agencies are saying use our rollover.
- Use it or loose it

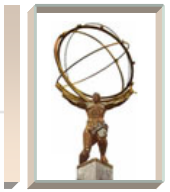
- Goal:
  - Detailed schedule (not just rolled up into 1 year) of hardware ramp.

The T2 Meeting Last May



- Tier 2 Planning Wiki:
  - http://www.usatlas.bnl.gov/twiki/bin/view/Admins/TierTwoPlanning

Tier2 Planning



- Tasks and Milestones
- Questions / Issues for facilities
  - NET2
  - SWT2
  - MWT2
- OSG Software and Panda validation
- References

Tasks and Milestones

- TierTwoFacilities and TierTwoNetworking
- TierTwoStorageDataServices
- TierTwoPolicyAccounting
- TierTwoOperationsUserSupport

Questions / Issues for facilities

NET2

- **Current set of action items:** NortheastTier2ActionItems
- Notes taken during the May 2006 Tier2 workshop:
  - Main set of questions concern storage system to use.

SWT2

- **Current set of action items:** SouthwestTier2ActionItems
- Notes during May 2006 Tier2 workshop:
  - Deployed 160 nodes
  - Number of headnodes - 8 TB
  - 16 TB SAN storage: IBRIX filesystem. To study
  - Additional xx TB available in distributed storage (single SATA drive)
    - Q: use dcache, or resilient dcache
  - Platform Rocks 3.3.0 (RHEL X86_64) - additional Dell rolls
  - Platform freeware scheduler Lava (looking into)
  - Networking: oc12 from ntg to houston peering site. future: LEARN and NLR? Need to look into peering.
  - OU: same as UTA, 1/4 scale. Plan on putting SRM/dCache on top of IBRIX filesystem.
  - OU: network connectivity option is through NLR.

Southwest Tier2 Action Items

Applies to ALL T2!

- Introduction
- Facilities
- Networking
- Storage and Data Services
- Policy and Accounting
- Operations and User Support

Introduction

These are some notes to follow-up on TierTwoPlanning tasks and deliverables as discussed at the May Tier2 workshop in Chicago. This is being updated for the Tier2 workshop at Harvard, August 17-18, 2006.

Facilities

Follow-up from the TierTwoFacilities working group:

- ...

Networking

Follow-up from the TierTwoNetworking working group:

- ...

Storage and Data Services

Follow-up from the TierTwoStorageDataServices working group:

- ...

Policy and Accounting

Follow-up from the TierTwoPolicyAccounting working group:

- ...

Operations and User Support

Follow-up from the TierTwoOperationsUserSupport working group:

- ...

Goal:

Fill-in/Update these Wikis

to top

Progress on T2 Services?

Tasks and Deliverables

To the following deliverables has to be added what is specified within ATLAS Computing TDR or other documents. Total space requirement (in TB) have to be added and crossreferenced with the content of the TierTwoFacilities document.

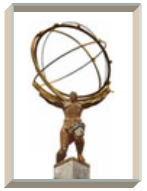
Each T2 has to:

- be able to connect to Tier1 through FTS channel (no work necessary, the channel is open by the Tier1)
- provide a SRM managed storage: backend storage can be dCache (SRM/dCache) or other efficient sotrage (high cost solutions) A suggested SRM/dCache server would require at least 2, better 3 or more machines.
  - a CORE machine providing a file server serving at least 1 TB using PNFS. Possibly this machine (the DB and PNFS manager) should be separate on a powerful machine (leave all the other daemons of the core and dors on a the other machine): BNL has 2 3.0 Xeon, 4 GB
  - one or more machines with DOORS (gsiftp, dcap, SRM). You may like a separate machine for the SRM door.
  - provide some reader/writer nodes (at least one reader/writer) An alternative would be a SRM/DRM solution on top of a high performance shared FS, like Ibrix, Panases, Luster, GPFS
- provide a DQ2 server

Here is a list of tasks/deliverables. Most will have also related milestones in the next section:

1. Provide a dCache server suitable for SC4 (v1.6.6.5) and production:
   - At least 2 machines (see above) with 1 TB (min to have a realistic use)
   - dCache installation (Easy to bring up! - BNL will offer support)
   - Tune-up and improvement can be done later
   - update of dCache when needed
2. Provide a DQ2 server
   - have a DQ2 server functional (v.?)
   - update DQ2 server when needed
3. Prototype and recipe of T2 DB infrastructure
   - Test DB service for Condition and Tag
   - start with current model
   - investigate differen models
   - provide requirements
   - provide a deployment recipe (Pacman?) and documentation
4. Readiness for the Calibration Allignment challenge (CAC will be this fall, after release 13, at the end of September)
   - Have a DB service (one machine with DB server hosting Tag and Condition DB - will be finalized by prototyping):
     - Tag: probably dedicated mysql server
     - Condition: probably file based DB like SQLite, but may be mysql
     - access will be only from within T2 site (all clusters part of it)

Tier 2 Documentation



- Uniform Web pages
- Up to date snapshots of hardware configuration
- Kill/Erase/destroy old web pages!