RCF/ACF Status and Plans

Michael Ernst

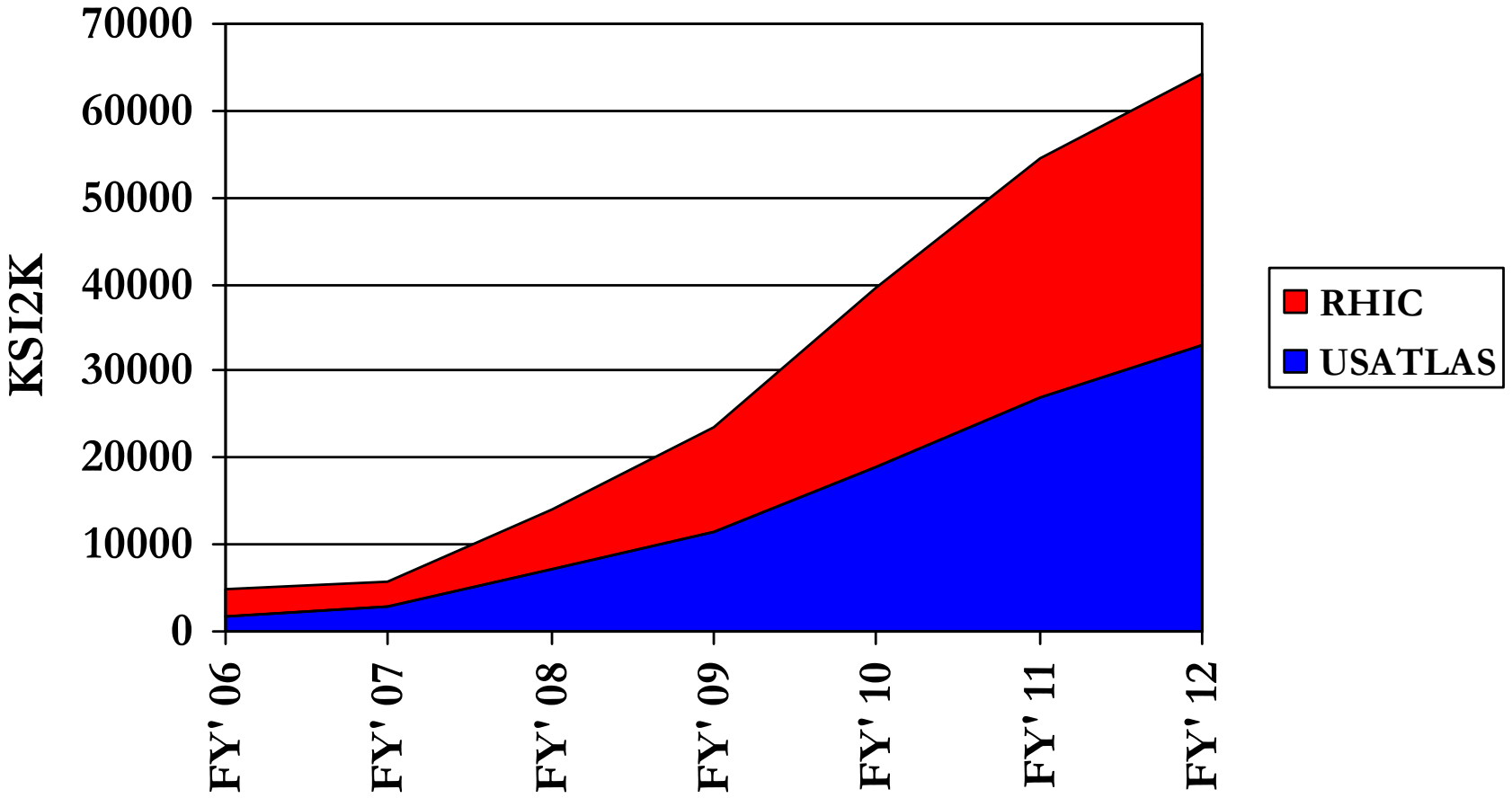
Physics Department Symposium

2 October 2007

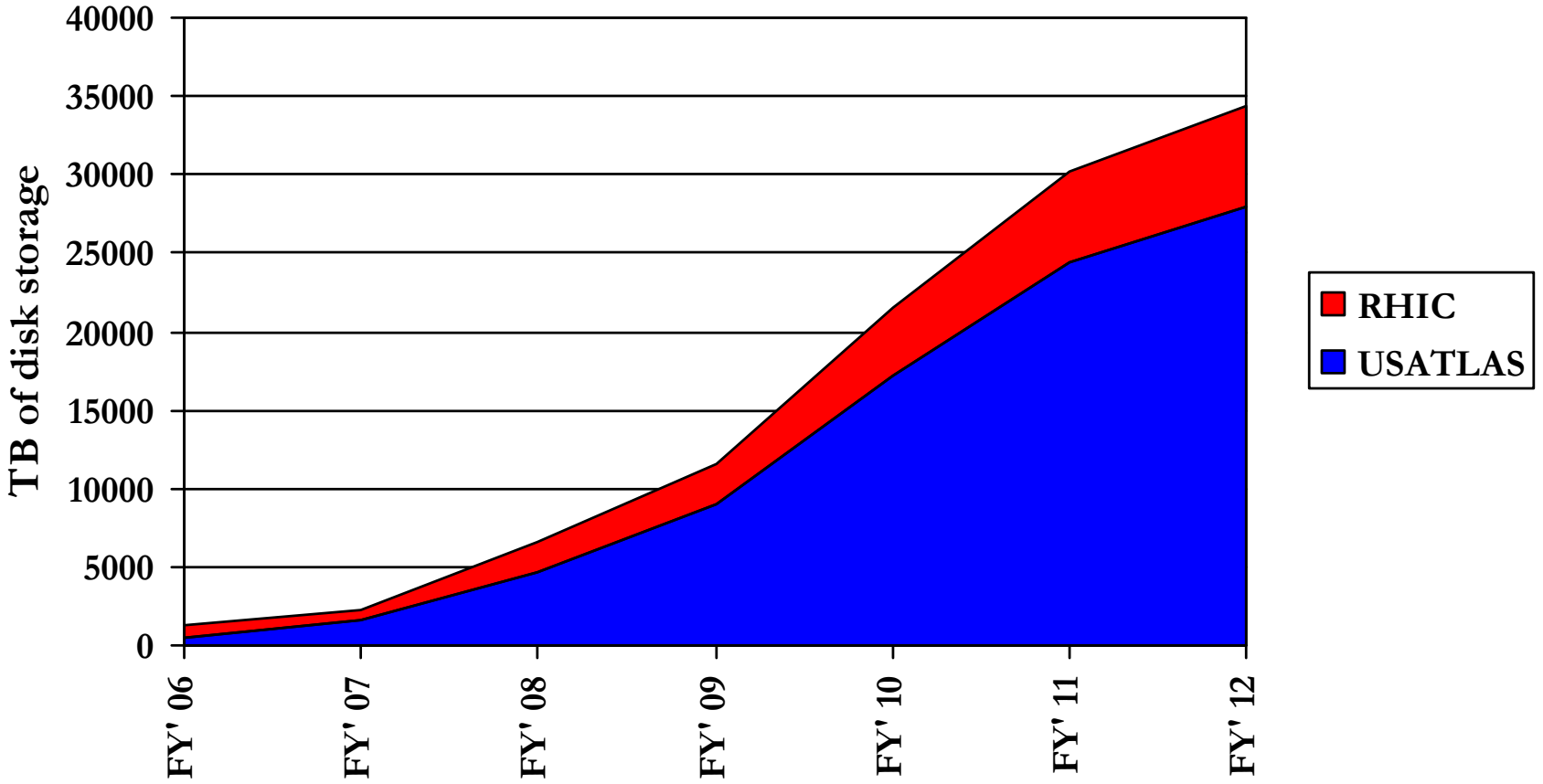
RHIC and ATLAS Computing Facility (RACF)

- Organizationally established in late ’90s
- Staffed as a Group in Physics Department
- Equipment physically located at Brookhaven Computing Facility (BCF)
  - BCF operated by ITD
- Currently co-located and co-operated
  - ACF ramping up quickly, compared to RCF currently
    - ACF capacities are ~ 65% for processing, 95% for disk capacity
    - ACF staff level ~ 85% of RCF

Expected Computing Capacity Evolution



Expected Disk Storage Capacity Evolution



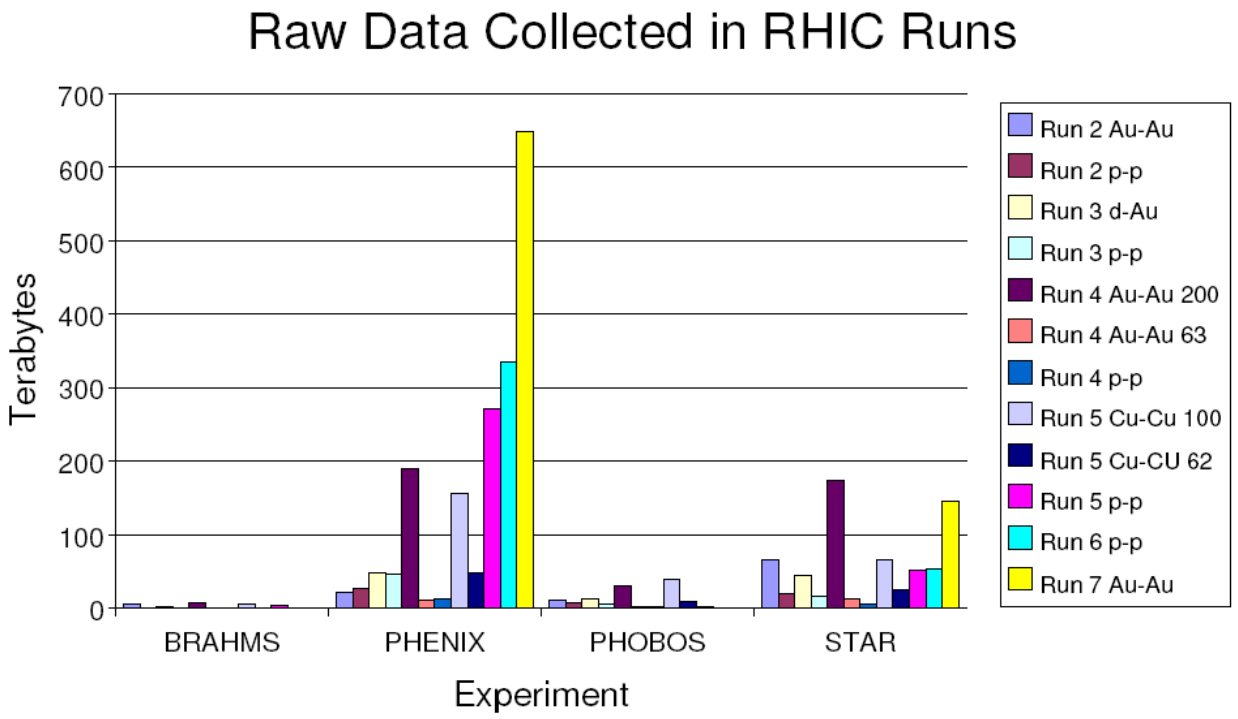
RCF Mission and Scale

- Mission
  - Online Recording of Raw Data
  - Production reconstruction of Raw Data
  - Primary Facility for Data Selection and Analysis
  - Long time Archiving and Serving of all Data
- Scale
  - Authorized staff of 20 FTE's

A lot more Raw Data this Year …

Raw Data (TB)

	Run 2 Au-Au	Run 2 p-p	Run 3 d-Au	Run 3 p-p	Run 4 Au-Au 200	Run 4 Au-Au 63	Run 4 p-p	Run 5 Cu-Cu 100	Run 5 Cu-CU 62	Run 5 p-p	Run 6 p-p	Run 7 Au-Au
BRAHMS	6.4	0.85	3.06	0.62	8.45	0.6	0.58	5.51	1.32	4.23	0.93	0.00
PHENIX	22.01	26.75	48.64	46.15	189.87	11.92	13.5	155.87	48.38	272.37	335.61	648.20
PHOBOS	11.26	7.25	13.53	5.7	30.35	2.23	2.61	39.75	10.5	2	0	0.00
STAR	65.91	19.69	45.6	16.92	174.55	13.79	5.94	66.91	26.43	51.60	54.10	145.29



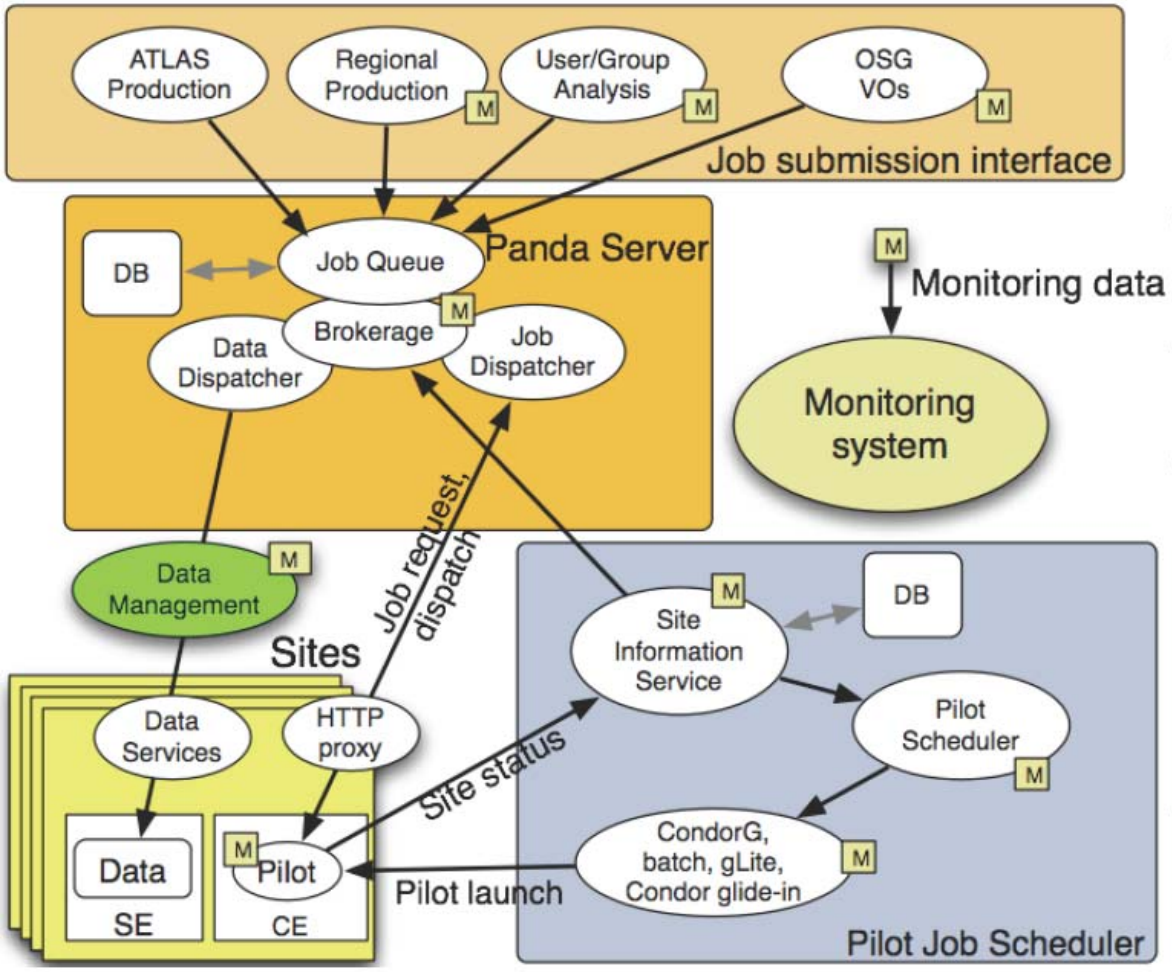
ATLAS Computing at BNL

- Principal objectives
  - Fulfill role as principal U.S. center (Tier-1) in the tiered ATLAS (and LHC) computing model
    - Supply capacity to ATLAS as agreed in the MOU (~23%)
    - Guarantee the capability and capacity needed by US ATLAS physics program
  - Establish a critical mass in computing and software to support ATLAS physics analysis at BNL and elsewhere in the U.S.
  - Contribute to ATLAS software most critical to enabling physics analysis at BNL and the U.S.
  - Provide to U.S. ATLAS physicists expertise and facilities and strengthen U.S. ATLAS physics
  - Leverage projects outside ATLAS, in particular Open Science Grid (OSG), where they can strengthen the BNL and U.S. ATLAS programs

Panda Basics

Workload management system for Production ANd Distributed Analysis

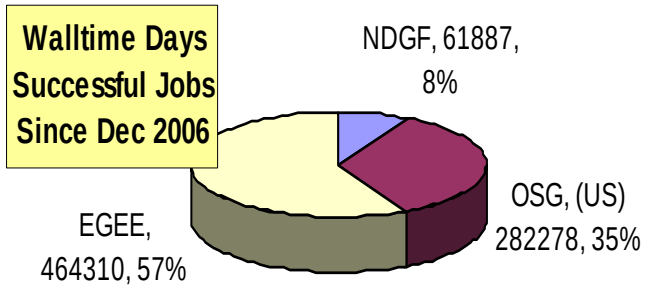
Panda team @ BNL, UT Arlington, U Chicago

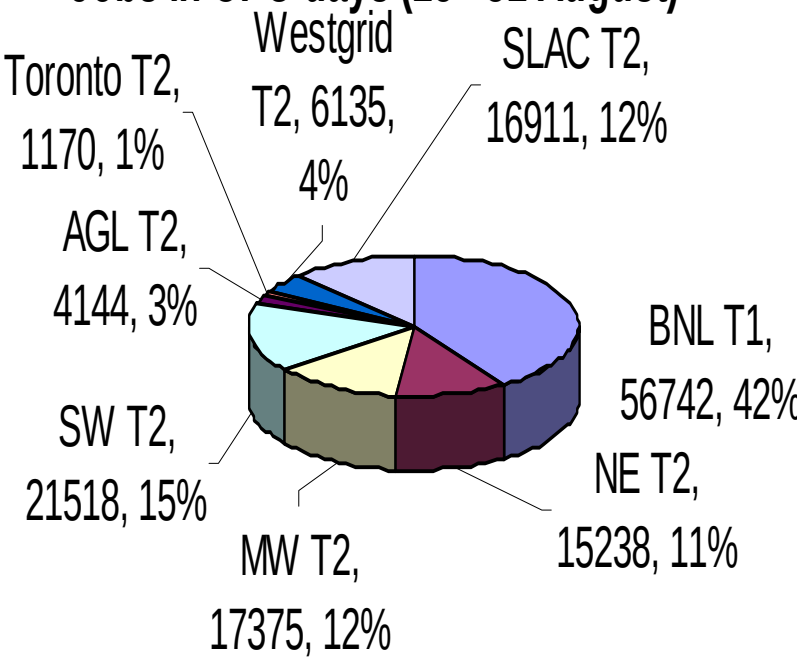


- Launched 8/05 to achieve scalable data-driven workload management
- Tightly integrated with DDM
- Pilot-based 'CPU harvesting'
- *Designed for analysis as well as production*
- Automation, monitoring, low ops manpower
- *Insulate* users from grid complexity, problems
  - Low entry threshold to using the grid
- Cautious in its dependencies
  - Proven components

(US, Canada) Production Statistics (PanDA)

- Since November, all available CPU's occupied (ran out of jobs only for few days, plus few days of service outages)
- About 400 TB of original data stored at U.S ATLAS Tier-1 (includes data generated on other grids)
- Additional ~100 TB of replicas kept at U.S. ATLAS Tier 2 sites





U.S. in the ATLAS Computing Model

- BNL Tier-1 responsibilities (~10 Tier-1's total)
  - Archival shares of raw and reconstructed data, and associated calibration & reprocessing
  - Store and serve 100% of ATLAS reconstruction (ESD), analysis (AOD) and physics tag data (TAG)
  - Physics group level managed production/analysis
  - Resources dedicated to U.S. physicists: additional per-physicist capacity at 50% of the level managed centrally by ATLAS
  - Tier-1 and Tier-2's both support institutional and individual users
    - Primarily end user analysis
- Scale of Tier-1 Center
  - Authorized Staff of 20 FTE's

The Computing Model

- Resources Spread Around the GRID

•Reprocessing of full data with improved calibrations 2 months after data taking.
•Managed Tape Access: RAW, ESD
•Disk Access: AOD, fraction of ESD

•Interactive Analysis
•Plots, Fits, Toy MC, Studies, ...

Tier 3

DPD

30 Sites Worldwide
Tier 2

AOD

•Derive 1st pass calibrations within 24 hours.
•Reconstruct rest of the data keeping up with data taking.

Tier 1

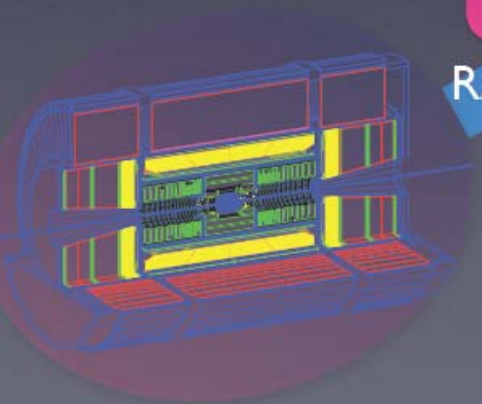
10 Sites Worldwide

RAW/
AOD/
ESD

Tier 0

RAW

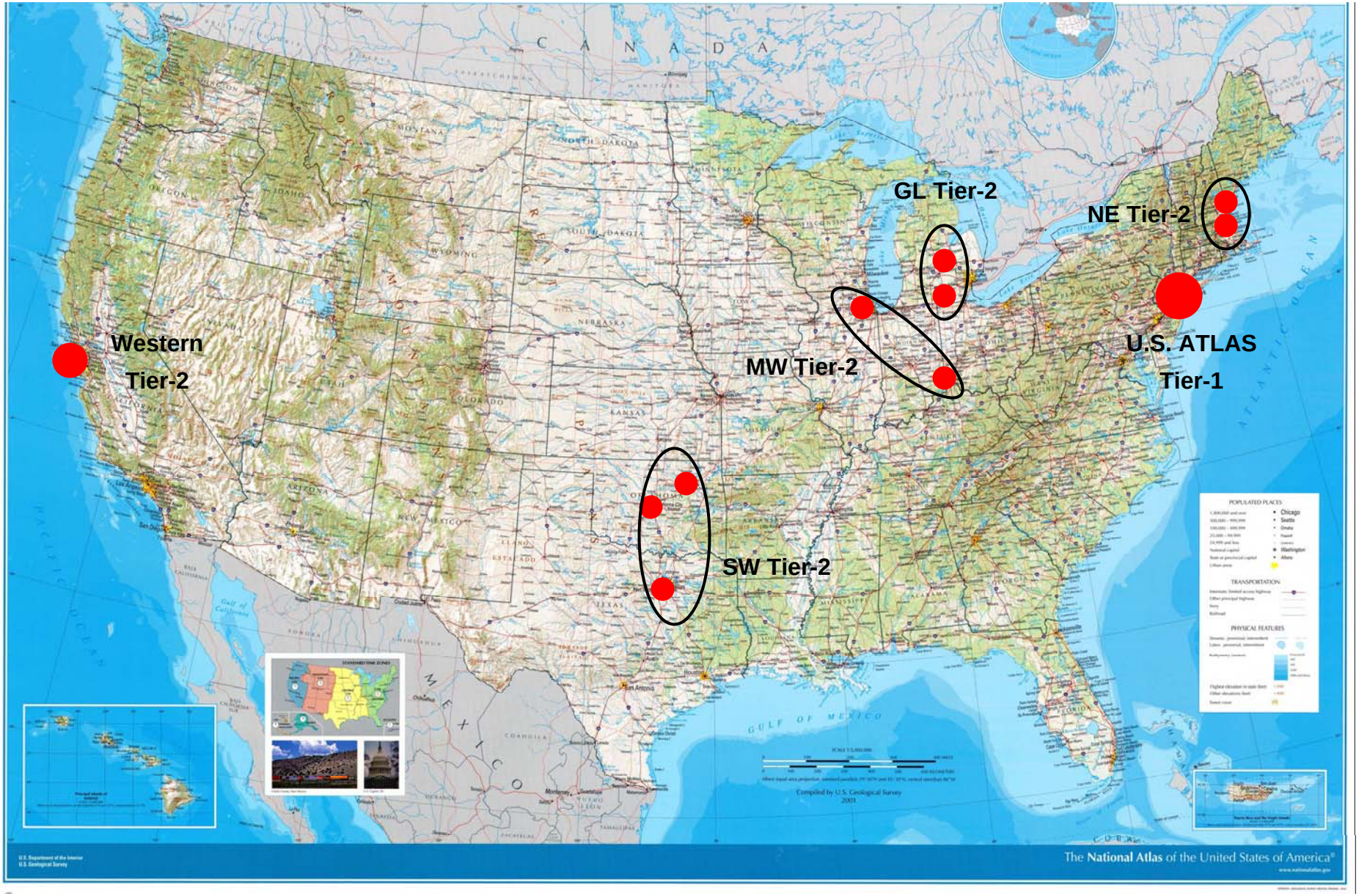
•Production of simulated events.
•User Analysis: 12 CPU/ Analyzer
•Disk Store: AOD



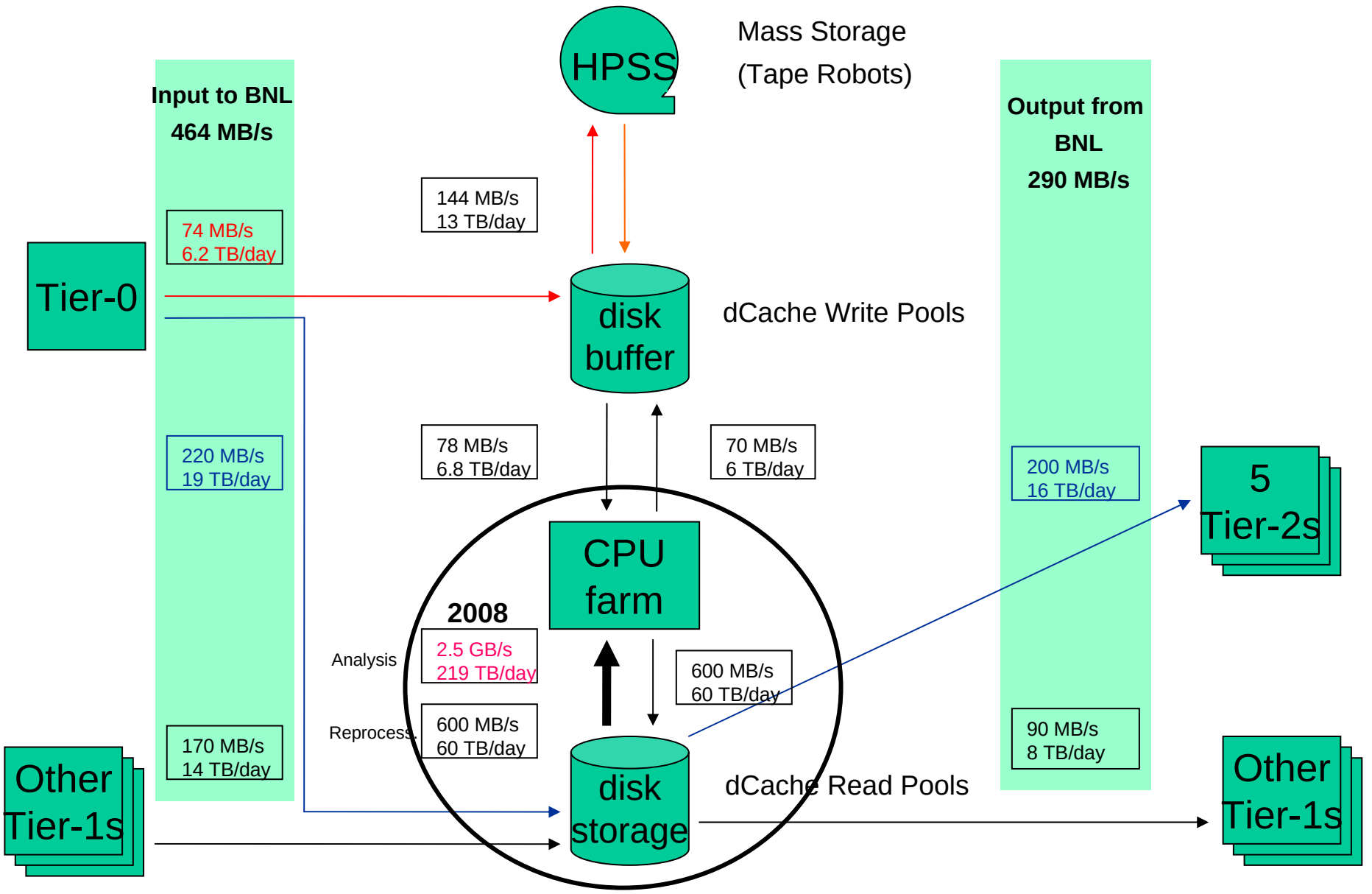
CERN Analysis Facility

•Primary purpose: calibrations
•Small subset of collaboration will have access to full ESD.
•Limited Access to RAW Data.

A. Farbin


GL Tier-2
NE Tier-2
Western
Tier-2
MW Tier-2
U.S. ATLAS
Tier-1
SW Tier-2
CANADA
MEXICO
GULF OF MEXICO
The National Atlas of the United States of America

U.S. ATLAS Tier-1 Networking Needs



ATLAS Distributed Data Management

Overview | Dataset Info | Page Help | User Guide | Feedback

OVERVIEW

OVERVIEW Activity
Activity in Last Hour
Activity in Last 4 Hours
Activity in Last 24 Hours
Activity in Last 7 Days
Activity in Last 30 Days
Activity in ...

Cloud Activity
ASGC Cloud
BNL Cloud
CERN Cloud
CNAF Cloud
FZK Cloud
LYON Cloud
NDGF Cloud
PIC Cloud
RAL Cloud
SARA Cloud
TRIUMF Cloud

Throughput (MB/s)

Throughput from 2007-08-01 12:30 to 2007-09-24 16:30

Data Transferred (GBytes)

Bytes transferred from 2007-08-01 12:30 to 2007-09-24 16:30

Completed File Transfers

Number of errors from 2007-08-01 12:30 to 2007-09-24 16:30

Total Number Errors

Number of errors from 2007-08-01 12:30 to 2007-09-24 16:30

Activity Summary ('*2007-08-01 12:30*' to '*2007-09-24 16:30*')

Click on the cloud name to view list of sites

	Transfers				Services		Errors		
Cloud	Efficiency	Throughput	Files Done	Datasets Done	DQ	Grid	Transfer	Local	Remote
ASGC	31%	2 MB/s	172848	1406			381250	29679	
BNL	98%	29 MB/s	1537722	403223			32188	0	
CERN	38%	4 MB/s	259721	126			424652	90843	
CNAF	54%	7 MB/s	551138	32938			469933	258895	
FZK	61%	5 MB/s	350385	13918			220385	51686	
LYON	67%	5 MB/s	319163	13623			156663	30974	
NDGF	57%	4 MB/s	275382	234			207272	53909	
PIC	49%	5 MB/s	374701	3367			390448	90970	
RAL	47%	3 MB/s	178798	513			199090	63599	
SARA	35%	3 MB/s	225193	10874			412684	196002	
TRIUMF	57%	5 MB/s	266430	10059			197602	106156	

CRITICAL | WARNING | NORMAL | GOOD | NO_ACTIVITY

Done | My Computer

Data Replication Performance in August and September as measured by the ATLAS Distributed Data Management Monitoring (ARDA Dashboard)

Cosmic data replication within clouds (new replication policy)

History Bookmarks Tools Help

http://gridui03.usatlas.bnl.gov:25880/server/pandamon/query?mode=listM4#LYON

Free Hotmail indico Windows Media Windows BeOnLine: связь, ко... Windows

Tier2	Datasets	Total Files in datasets	Total CpFiles in datasets	Completed	Transfer	Subscribed
BEIJING	50	370	30	4	10	36
CPPM	11	73	7	1	2	8
LAL	98	857	60	1	39	58
LAPP	11	184	25	0	5	6
LPC	21	159	23	3	7	11
LPNHE	98	1205	101	0	36	62
NIPNE_02	11	65	16	1	5	5
NIPNE_07	11	112	18	5	3	3
SACLAY	98	547	28	0	18	80
TOKYO	291	3768	156	0	65	206

Lyon Cloud

BNL Cloud

100% of data requested by 5 T2s

Tier-3 at UWM

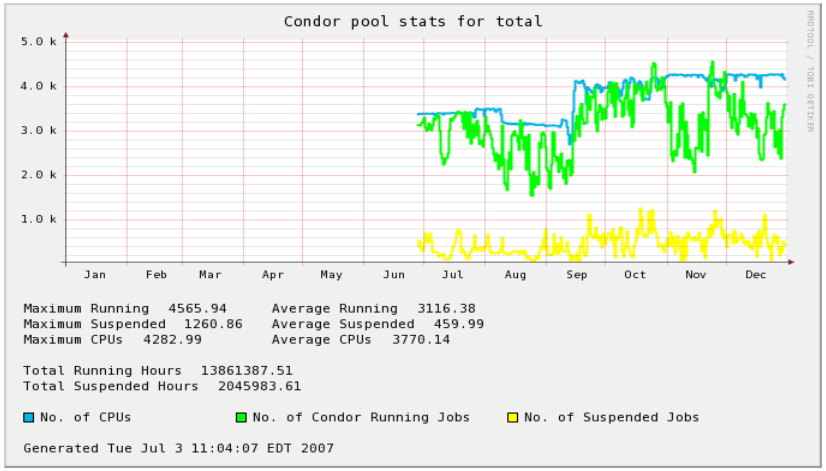
Dataset	Files on Source Host	BEIJING	CPPM	LAL	LAPP	LPC	LPNHE	NIPNE_02	NIPNE_07	SACLAY	TOKYO
...0019354.Default.L1TT-b10000000.ESD.v13002503.part0001	3		1								
...0019391.Default.L1TT-b00000010.ESD.v13002503.part0001	1										
...0019391.Default.L1TT-b00000010.ESD.v13002503.part0002	3										
...0019391.Default.L1TT-b00000010.ESD.v13002503.part0003	2										
...0019536.Default.L1TT-b00000001.ESD.v13002502.part0001	49				8						
...0019544.Default.L1TT-b00000001.ESD.v13002502.part0001	93										
...0019562.Default.L1TT-b00000001.ESD.v13002503.part0001	39										
...0019580.Default.L1TT-b00000001.ESD.v13002503.part0001	4										
...0019585.Default.L1TT-b00000001.ESD.v13002503.part0001	16										
...0019587.Default.L1TT-b00000001.ESD.v13002503.part0001	8										
...0019619.Default.L1TT-b00000010.ESD.v13002503.part0001	1										
...0019626.Default.L1TT-b00000010.ESD.v13002503.part0001	25										
...0019638.Default.L1TT-b00000010.ESD.v13002503.part0001	8										
...0019654.Default.L1TT-b00000010.ESD.v13002503.part0001	1										
...0019656.Default.L1TT-b00000010.ESD.v13002503.part0001	10			1							
...0019659.Default.L1TT-b00000010.ESD.v13002503.part0001	3			2		2			2		

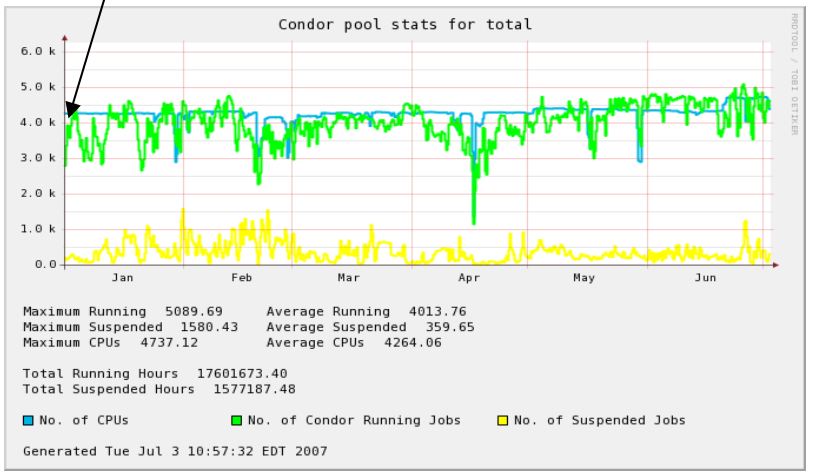
A. Klimentov at WLCG Collaboration Meeting
Victoria, 1 September 2007
http://indico.cern.ch/conferenceTimeTable.py?confId=3578

100% of data shared by T2s within the cloud

Processor Farm Occupancy (RACF)

4,200 Job Slots

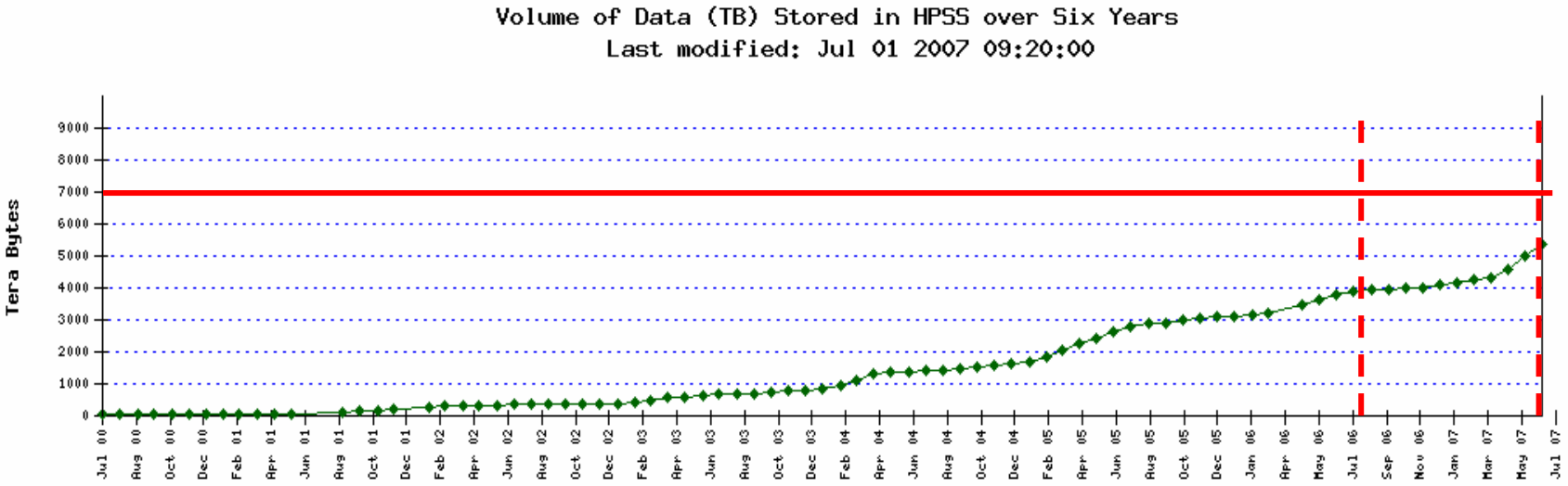




- Left-hand plot is for late June'06 to 12/31/06
- Right-hand plot is for 01/01/07 to 07/03/07
- Occupancy rose from 83% to 94% between the two periods
- Created general queue in 2006 to increase occupancy
- NP experiments providing resources to ATLAS and vice versa
  - STAR getting cycles from U.S. ATLAS Tier-1 – ATLAS jobs running on RHIC CPUs

Mass Storage System

- Tertiary Storage Data Volume
  - Automated Magnetic Tape repositories managed by HPSS

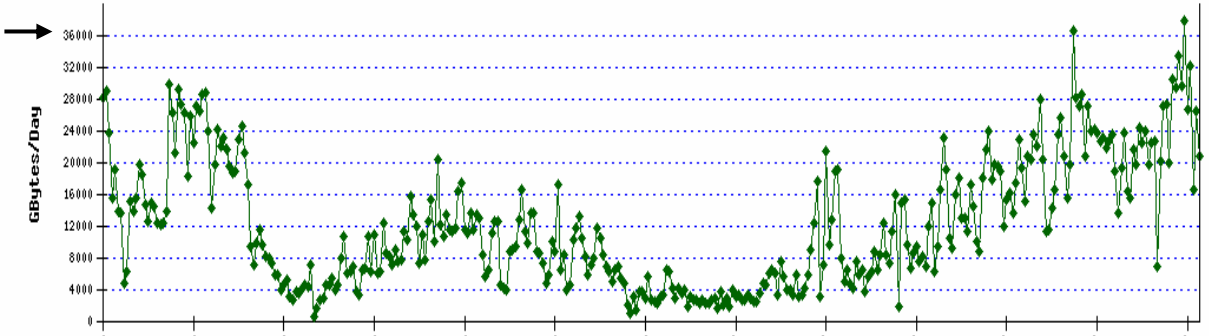


7 PB

5.5 PB

  - Data handling rate by Tape Drives over the past year

36 TB/Day



Physical Infrastructure

- Have reached limits in all areas
  - Without additional space RACF will not be able to accommodate the next robot (due in early spring 2008) and the upgrade to processing power and disk storage
  - Reallocation of space to RCF/ACF allows 2008 expansion
    - Additional power & cooling is needed each year
  - Need expansion of space in 2009 and beyond
    - Working with ITD, BNL Plant Engineering and BNL Management on a plan
      - **Very tight schedule**
      - **Progress is not as good as we had hoped for**
        - **Technical and organizational problems**
- This is our top concern at the moment

Summary

- The RHIC and the ATLAS Computing Facility (RACF) serves as the hub and principal center for the user community, with scale-up in proportion to data taking
- The RACF at BNL is on track to meet the performance and capacity requirements of the RHIC and the ATLAS research programs
- Plans in place to evolve and expand facility services to meet expected needs
  - Are based on successful adjustments of technical directions
  - Remain within the mainstream of NP and HEP computing
  - Requires agreed and planned for increases in 2008 and beyond
- Space, Power & Cooling at the RACF is on the critical path