RHIC Computing Facility

Bruce G. Gibbard

DOE/Nuclear Physics Review of RHIC Science and Technology

BNL

24-26 July 2005

RHIC Computing Facility (RCF)

- Organizationally established in 1997
- Staffed as a group in Physics Dept.
- Equipment in raised floor computer area of IT Division, Bldg 515
- Currently co-located and co-operated with ATLAS Computing Facility (ACF) at BNL – US ATLAS Tier 1 Regional Center
  - Long term, 2008+, facilities will be of comparable scale
  - Currently:
    - ACF capacities are ~25% of RCF
    - ACF staff level ~ 65% of RCF

RCF Mission and Scale

- Mission
  - Online recording of Raw data
  - Production reconstruction of Raw data
  - Primary facility for data selection (mining) and analysis
  - Long term archiving and serving of all data
  - … but not originally intended for Monte Carlo generation
- Scale
  - Authorized staff of 20 FTE's
  - Historically ~$2M/year equipment ***refresh*** funding (20-25% annual replacement) – Current year austerity limited to $1.3M
    - Addressing obsolescence
    - Resulting in important collateral capacity growth
  - Growth beyond originally planned scale will require an increase in such refresh funding in the out years

Experiment / RCF Interaction

- Weekly Liaison (Operations) Meeting
  - Review recent performance and problems
  - Plan for near term operations
- Experiments / RCF Annual Series of Meetings to Develop Capital Spending Plan
  - Estimate scale of need for current/coming run
  - Details of distribution of equipment to be procured
  - Most recent in early this Spring for FY ’06 funds
- Periodic Topical Meetings
  - For example: ~Annual Linux farm OS upgrade planning
- Other User Interactions
  - Web site
  - Trouble ticket system (in house) – 3 instances for …
    - RHIC services (1188 tickets for last year)
    - ATLAS services (649 tickets for last year)
    - Grid services (114 tickets for last year)
  - New ticket system (RT – open source) – in trial use

Computing Requirements Estimate

- ***In Response to Recommendation from Last Year's Review***
- A Comprehensive Long Range Estimate Done by PHENIX, RCF & STAR in Fall/Winter 2005
  - Conclusions published as part of ***"Mid-Term Strategic Plan: 2006-2011 For the Relativistic Heavy Ion Collider", February 14, 2006***
- Input is Raw Data Volume for Each Species & Experiment by Year
- Model for Requirements Projection
  - Assume facility resource needs scale with raw data volume
    - Using 2005 Cu-Cu running as a reference
  - With adjustable parameters reflecting expected relative …
    - Richness of data set (density of interesting events)
    - Maturity of processing software
    - Number of reconstruction passes

… for each experiment, species, and year

Computing Cost Estimate

- Requirements Model Output Used as Input to Cost Estimate
- Costing Model is based on
  - Recent procurements
  - Historic trends (Moore's Law and similar technology based trends)
  - Assumed migration from reliance on centralized SAN RAID 5 disk toward use of Intel/Linux processor farm distributed disk for bulk of storage
    - As will be discussed later
  - Assume use of obvious technology evolution (multi-core processors, etc.)
- For running scenarios considered, capacity growth associated with equipment ***refresh*** meets increased capacity requirements in 2007 but increase equipment funding is required in 2008 and beyond
  - Required capacities by year and a funding profile allowing them to be achieved are shown on following slide

Requirements Estimate for a Particular Running Scenario

	FY '06	FY '07	FY '08	FY '09	FY '10	FY '11
Annual Requirement						
Real Data Volume (TB)	1700	2700	3500	4500	7200	8600
Reco CPU (KSI2K)	600	1000	2900	4700	8500	9700
Analys CPU (KSI2K)	310	570	1800	2700	4600	5400
Dist. Disk (TB)	220	480	1500	1700	3000	3600
Cent. Disk (TB)	30	60	190	260	450	560
Annual Tape Volume (TB)	2000	3200	4200	5400	8700	10300
Tape bandwidth (MB/sec)	690	920	920	1700	2100	2300
WAN bandwidth (Mb/sec)	1400	2000	2100	4300	5700	6700
Simulation CPU (KSI2K)	110	200	610	1000	1800	2100
Simulation Data Volume (TB)	330	530	710	900	1400	1700
Installed Requirement						
CPU (KSI2K)	2100	2800	6800	11800	20800	27700
Dist. Disk (TB)	480	720	1900	2600	4300	5700
Cent. Disk (TB)	200	200	290	400	650	880
Tape Volume (TB)	4300	7500	11800	17200	25900	36200
Tape bandwidth (MB/sec)	920	1400	1600	2500	3300	4000
WAN bandwidth (Mb/sec)	1500	2700	3500	6100	8700	11000

Funding Profile ($K)

	FY '06	FY '07	FY '08	FY '09	FY '10	FY '11
CPU + Distributed Disk	270	770	1360	790	1270	1960
Central Disk	150	250	400	330	450	310
Tape Storage System	640	590	250	1060	570	250
LAN	120	190	250	270	320	180
Overhead	130	200	250	270	290	300
Total Annual Cost	1310	2000	2510	2720	2900	3000

Principal RCF Services

- General Collaboration & User Support
- Compute Services (Linux Farms)
  - Programmatic production processing
  - Individual analysis
- Online Storage (Disk)
  - Working data storage (Read / Write)
  - Data serving (mostly Read)
- Mass Storage (HSM Robotic Tape System)
  - Raw data recording and archiving
  - Derived data archiving
- Grid & Network Services

RCF Staff

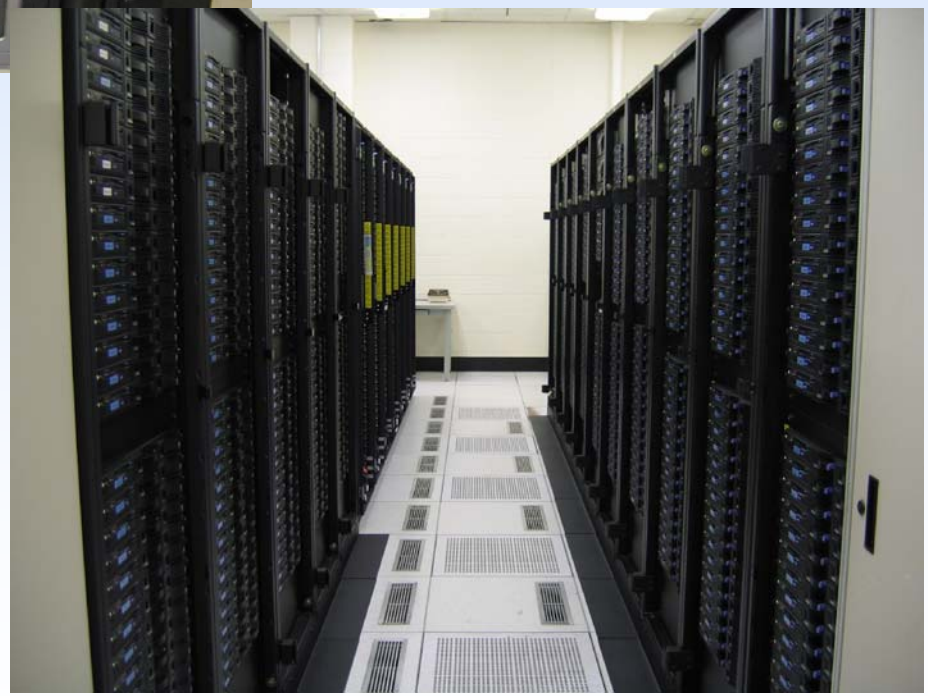
- Current authorized staff level of 20 FTE's is lower than staffs at other sites with comparable missions
- It does work in terms of operations in the context of a co-located ATLAS Tier 1
  - Very high level of commonality and therefore important synergy
  - A dramatic divergence in technical directions could change this but seems very unlikely
- It does not allow for aggressive involvement in developing and/or adopting new technologies
  - Which might in the long or even medium term significantly improve cost effectiveness

Distribution of RHIC Effort

	Current FTE's	Target FTE's
Linux Farms	3.5	3.5
Mass Storage	4.2	4.2
Disk	2.6	2.6
User Support	1.9	2.9
Fabric Infrastr.	2.1	2.6
Wide Area Services	1.8	1.8
Admin	2.5	2.5
TOTAL	18.5	20.0

Mass Storage

Compute Farm

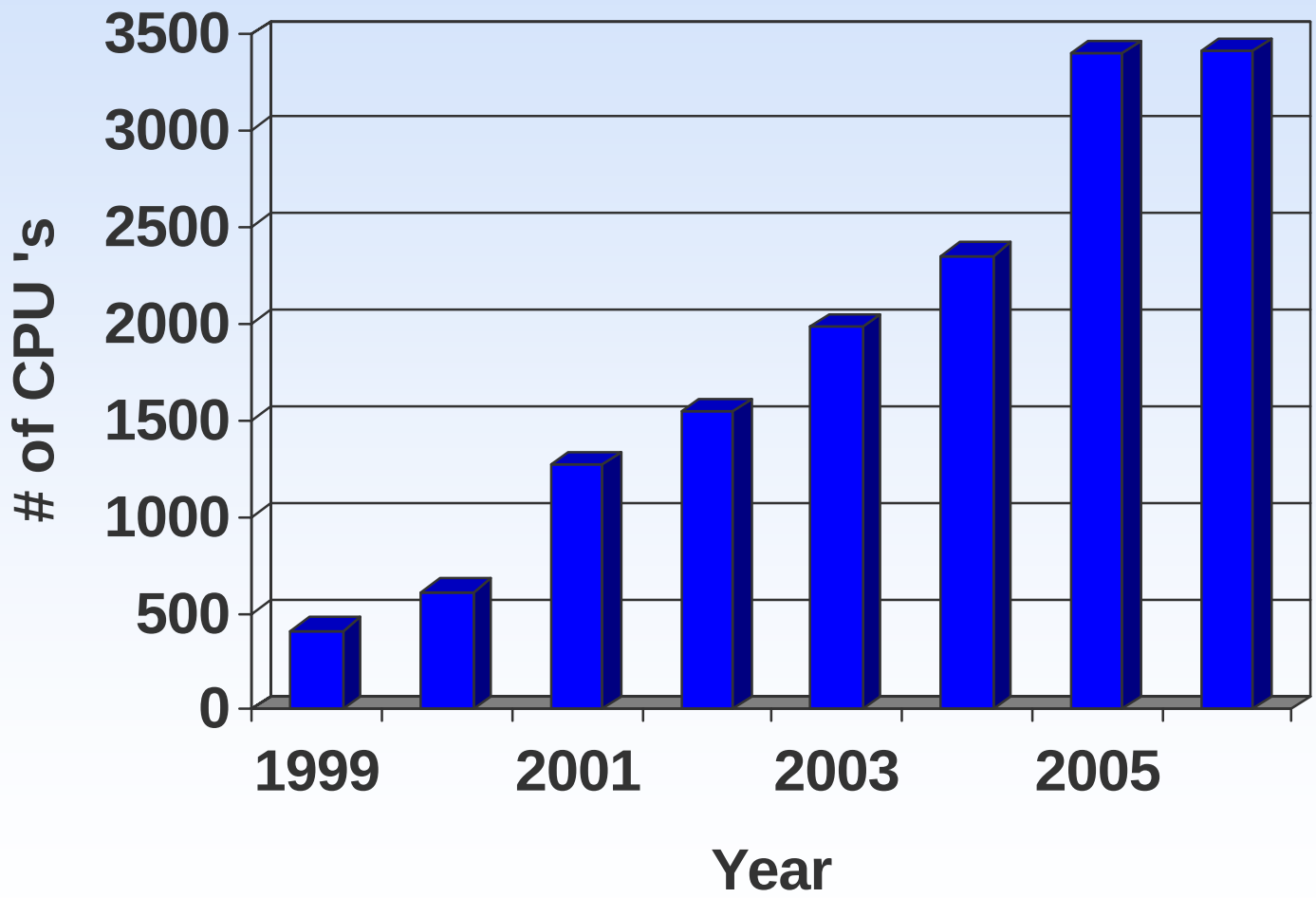


Online Storage

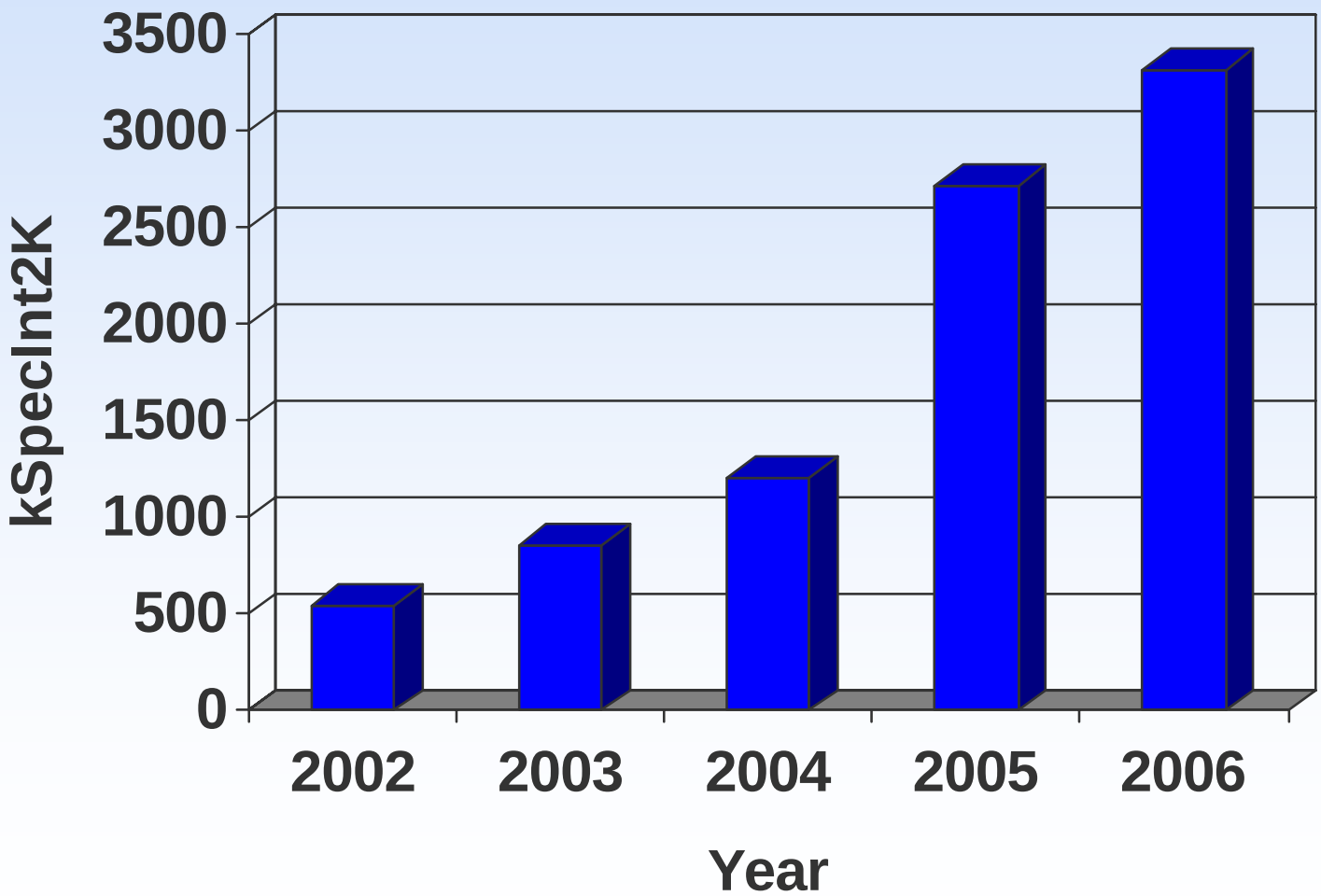
Compute Services

- Intel/Linux dual CPU rack mounted systems
  - Currently 1,700 compute servers with 3,400 CPU's (Intel) for 2,700 kSI2K
  - Delivery is expected in the next week or two of 170 additional machines with 800 kSI2K
    - First local production use of AMD (in place of Intel) and dual core processors (so 680 effective CPU's in 170 servers)
    - Dual core AMD Opteron is optimal for price performance
    - Multi-core also addresses power/cooling barrier by finessing non-linearity of power consumption with clock speed
  - Transition to AMD and to dual core is transparent / non-disruptive
- Expect to address future requirements by continuing to follow Moore's law price/performance in commodity market (including multi-core, & 64 bit advances)

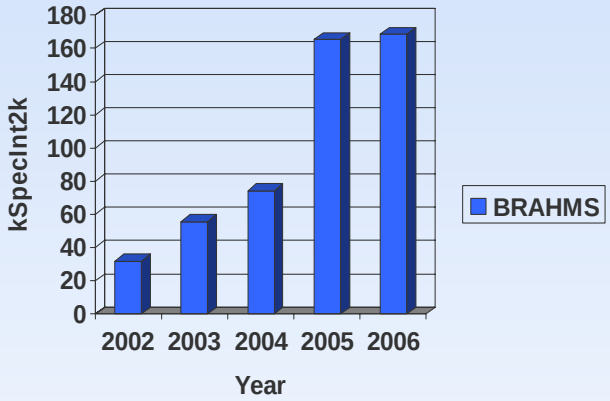
Growth of RCF Linux Farm

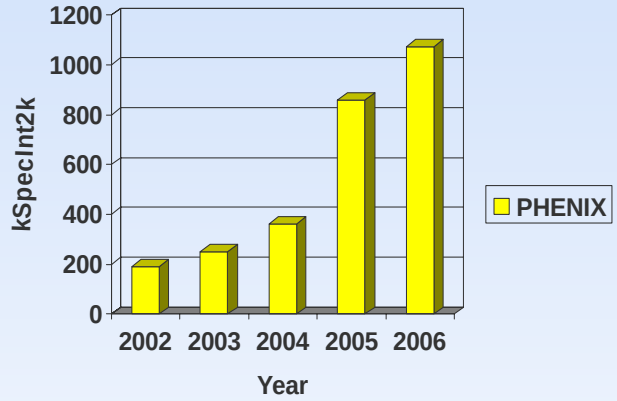


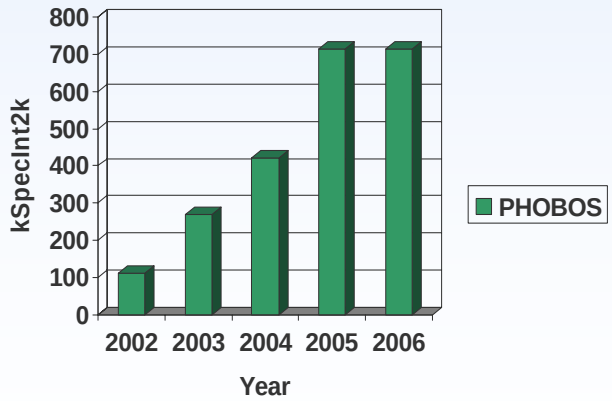
Growth in RCF Computational Power

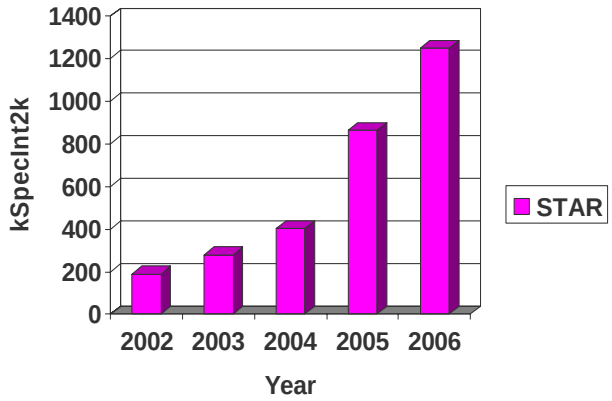


Growth in RCF Computational Power by experiment



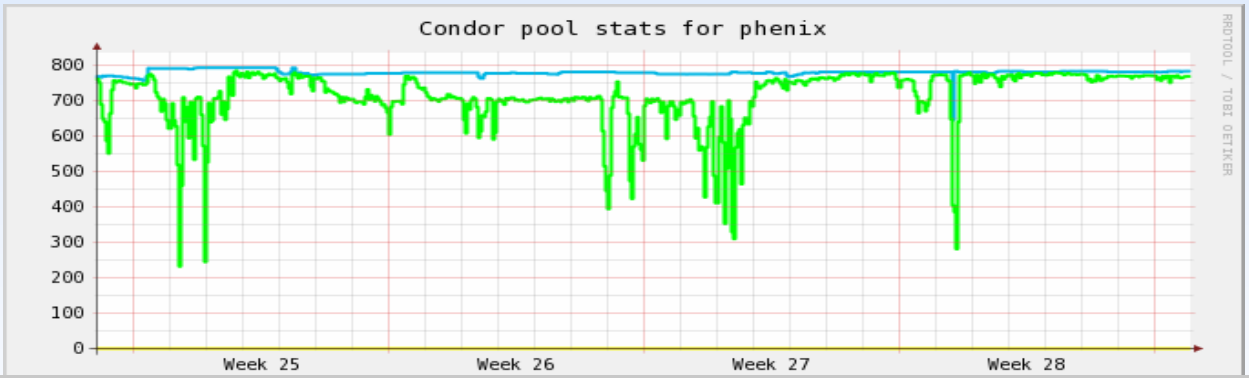


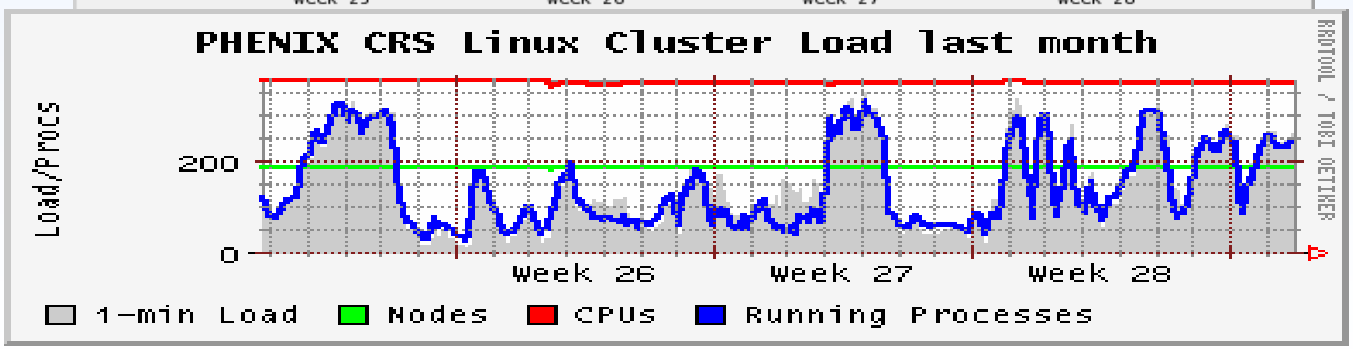




Resource Utilization Issues

- ***In Response to Recommendation from Last Year's Review***
- Ganglia Monitoring by Load or CPU show significant time with no apparent active job in a large fraction of the farm
- But Condor monitoring shows that in general there are active jobs
- Issue is one of not properly accounted I/O waits
  - In particular network I/O's (dCache, NFS write, possibly others)
  - But not in general HPSS





Resource Sharing Queue

- Goal is to make available idle cycles in processor farms to other experiment users without impacting primary "owner"
- Mechanism is to evict "guest" user jobs when owner jobs starts

Statistics from 06-29-2006 to 07-06-2006

no. jobs completed

source \ destination	phenix	phobos	star	brahms	atlas	rcf	total
phenix			69	289	1		359
phobos							
star	560			2	148		710
brahms							
atlas							
rcf							
total							

no. jobs evicted before completion

source \ destination	phenix	phobos	star	brahms	atlas	rcf	total
phenix			23	113			136
phobos							
star	85				2		87
brahms							
atlas							
rcf							
total							

- System needs further optimization
  - Works well for short jobs but not widely used
    - See case of long PHENIX jobs versus short STAR jobs above
  - Its use needs to be more transparent (short jobs default to it)
  - Consider increasing rights of guests
    - Only suspend for owner or …
    - Give guest a grace period to complete

Online (Disk) Storage

- Changes in disk storage model
- As discussed last year, historic RCF model was Sun/NFS served RAID 5 SAN connected central disk for all storage
  - … augmented by local scratch disk on CPU boxes
- Growth in demand drove disk costs to match and exceed CPU costs
- New strategy: Differentiate disk by function
  - Central NFS "full function" disk for Read/Write not yet backed up data, of which there is a limited amount
  - Less expensive distributed disk for "mostly Read" already backed up data, of which there is a great deal
  - Committing to support two online (disk) storage paths

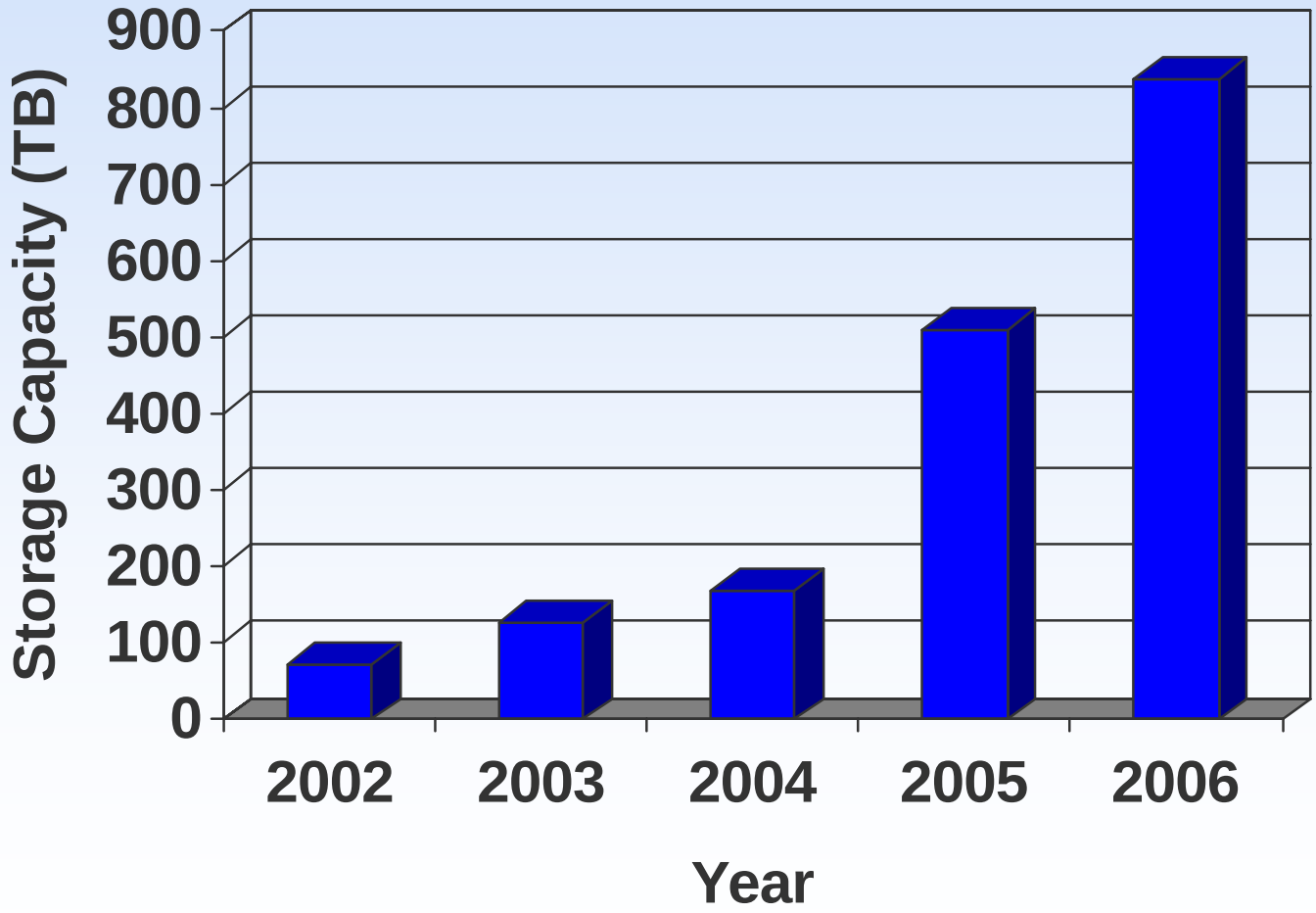
“Full Function” Disk Service

- Read/Write, reliable, high performance and high availability – NFS served RAID systems
  - Historically
    - ~150 TB of SUN served RAID 5 disk
    - ~70 TB of Panasas (appliance) served RAID 5 disk
  - Current acquisitions
    - ~100 TB of Nexsan & Aberdeen Linux served RAID 5/6 disk
  - Moved to lower Tier of RAID disk vendors this year
    - Hoped for very high reliability/performance from somewhat more expensive vendors not fully realized and offset by other technical problems

"Mostly Read" Disk

- Disk deployed on Linux processor farm nodes
  - IDE & SCSI → SATA
  - ~3 x less expensive then full function disk
- Requires additional disk/data management software and yields less than a truly full function solution
- Two such management systems currently in use at RCF
  - dCache - DESY/Fermilab developed Grid aware product
    - Scalable, robust, unified pool of storage with integral mass store backing, posix-like access, ROOT support
    - ATLAS is major BNL user with 150 → 500 TB
      - Heavily used in production for ~18 months
  - Xrootd – SLAC, CERN, BNL + other community developers
    - Similar scalable product currently under very active development
    - STAR is major BNL user with 130 → 330 TB
      - Heavily used for more than a year

Growth in RCF Distributed Storage Capacity



Some Detail

- PHENIX is RHIC dCache user
  - Currently 26 TB of disk on 120 nodes
    - Read rates on occasion exceed 50 TB/day
  - Expansion to 110 TB expected soon with further expansion to 200 TB likely this year
  - Local dCache development work in conjunction with US ATLAS Tier 1
    - Currently supporting one of the largest dCache instances
    - Officially accepted as a developer by primary product developers
    - Actively investigating
      - Issues of high performance scalability as a disk system
      - Issues of Grid performance and reliability (in particular its SRM)
- STAR and xrootd
  - Development cycle
    - Accurate usage monitoring (at user application level)
    - Integrate SRM (Storage Resource Manager) technology on the back-end
  - Proceed with development to make Xrootd next-generation Grid-aware
    - Alter model to an object based model
    - Built-in capability to scan large datasets in search for rare event
    - Plan strongly relies on
      - SRM technology - a US-leading activity to date
      - Additional support beyond existing ~ 2 FTE x 3 years

PHENIX dCache Configuration

HPSS Tape Silo

PHENIX Compute Farm
(107 pool servers, SL3)

26 TB disk in total
All Pool Servers are also
Clients (Workers)

External Write Pools
(9 pool servers, RHEL[3,4])

Per node: VA
800Mhz, 400GB disk,
Fast Ethernet, EXT
and XFS file system

Core Server
(PNFS, Admin, etc.)

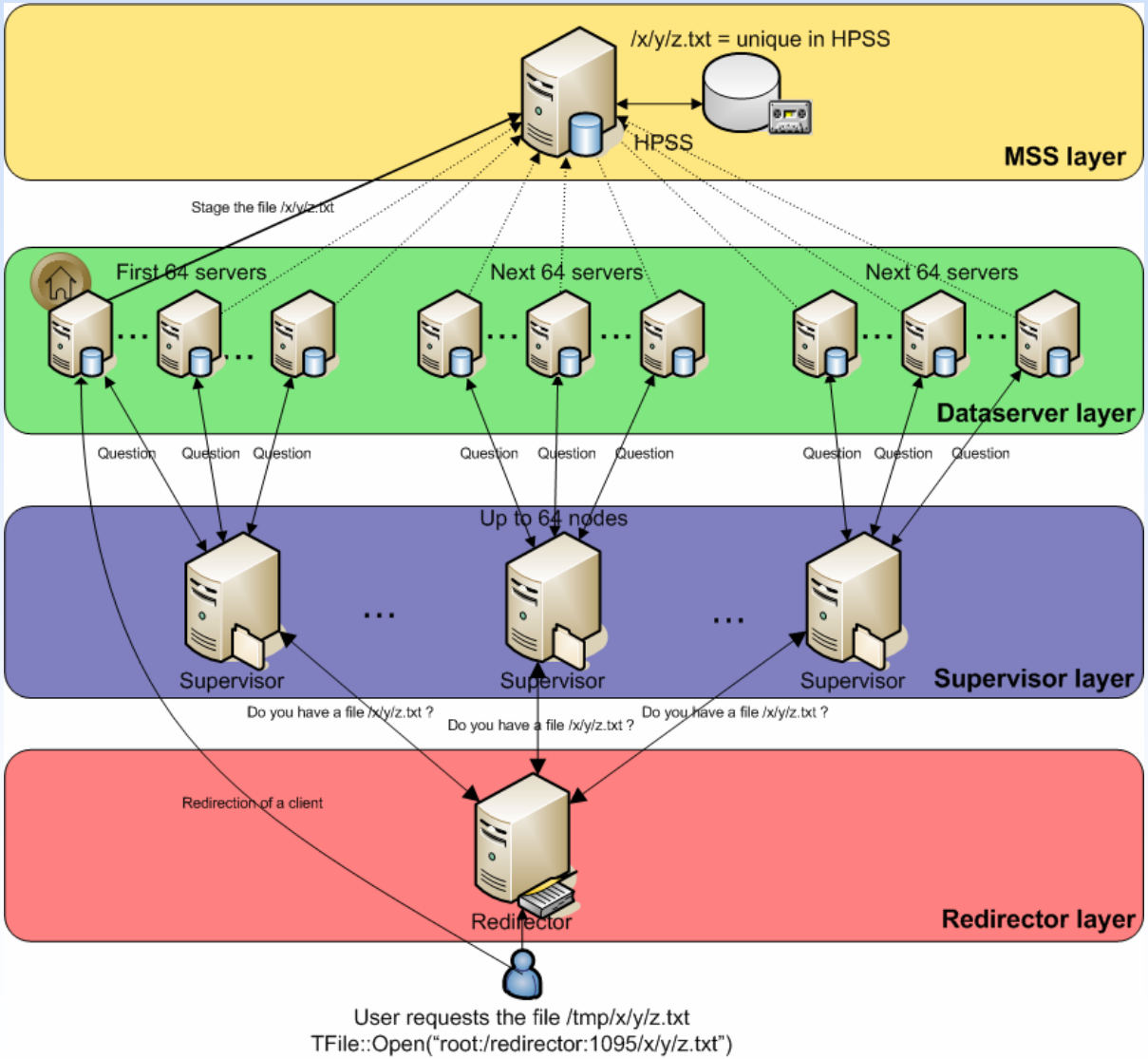
4 SRM/GFTP Doors with
fast ethernet

3 dCap Doors with
fast Ethernet

Outbound
Only

Firewall

Xrootd architecture



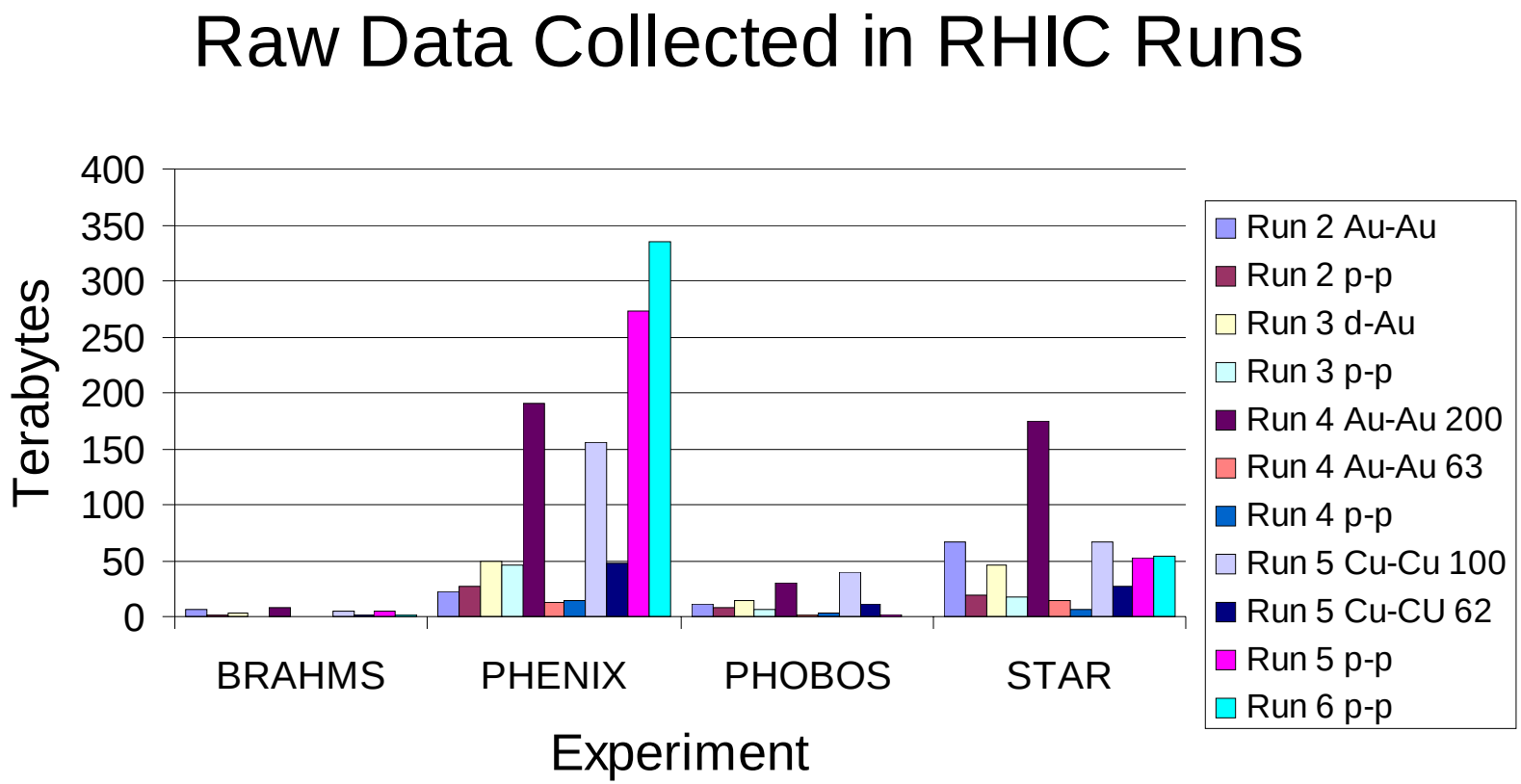
Coordination with MSS required integrating with the DataCarousel tool.

Redirector could be set as round robin

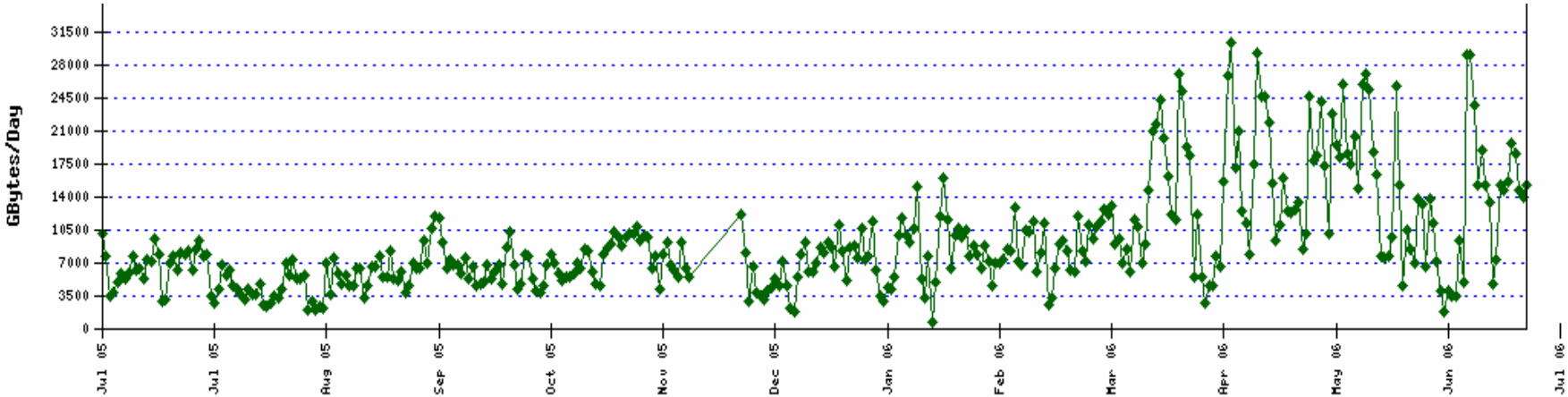
Mass Storage System

- Hierarchical Storage Manager is ***HPSS*** from IBM
  - Version 5.1 moving soon to 6.2
- StorageTek Robotic Tape Libraries
  - 4 classic cylindrical silos
  - 1 new SL8500 linear library
  - Total capacity ~7 PBytes now containing ~4 PBytes of data
- Tape Drives
  - 37 StorageTek 9940b (30 MB/sec)
  - 20 LT0 Gen 3 (80 MB/sec)
- ~30 TB of disk cache
- New data movers are Linux/Intel instead of AIX/IBM
- In house developed tape access optimization software

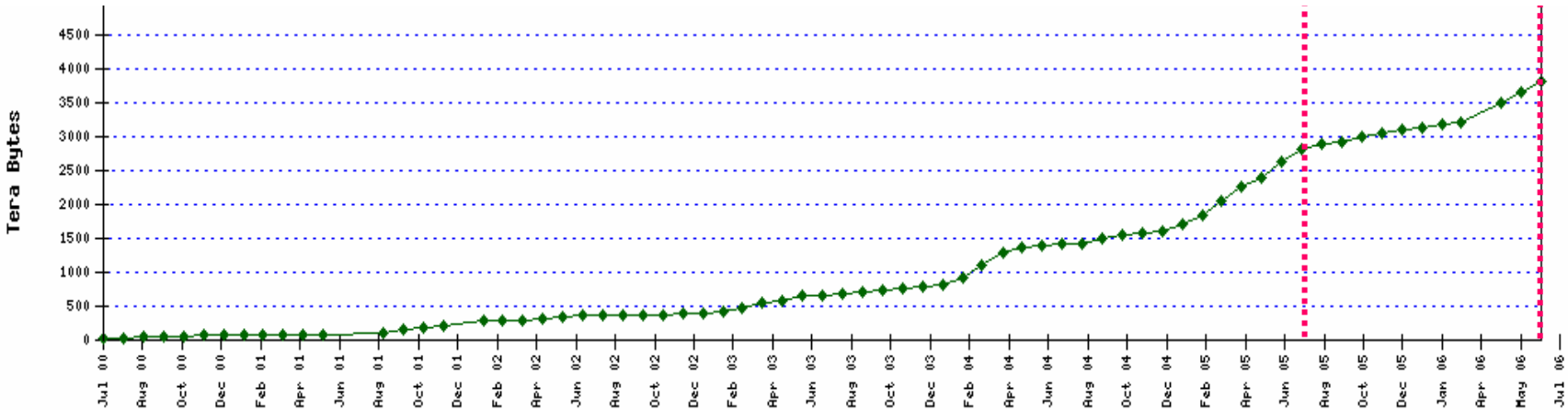
History of Raw Data Collection



Tape I/O Rate (GB/day) over Past Year



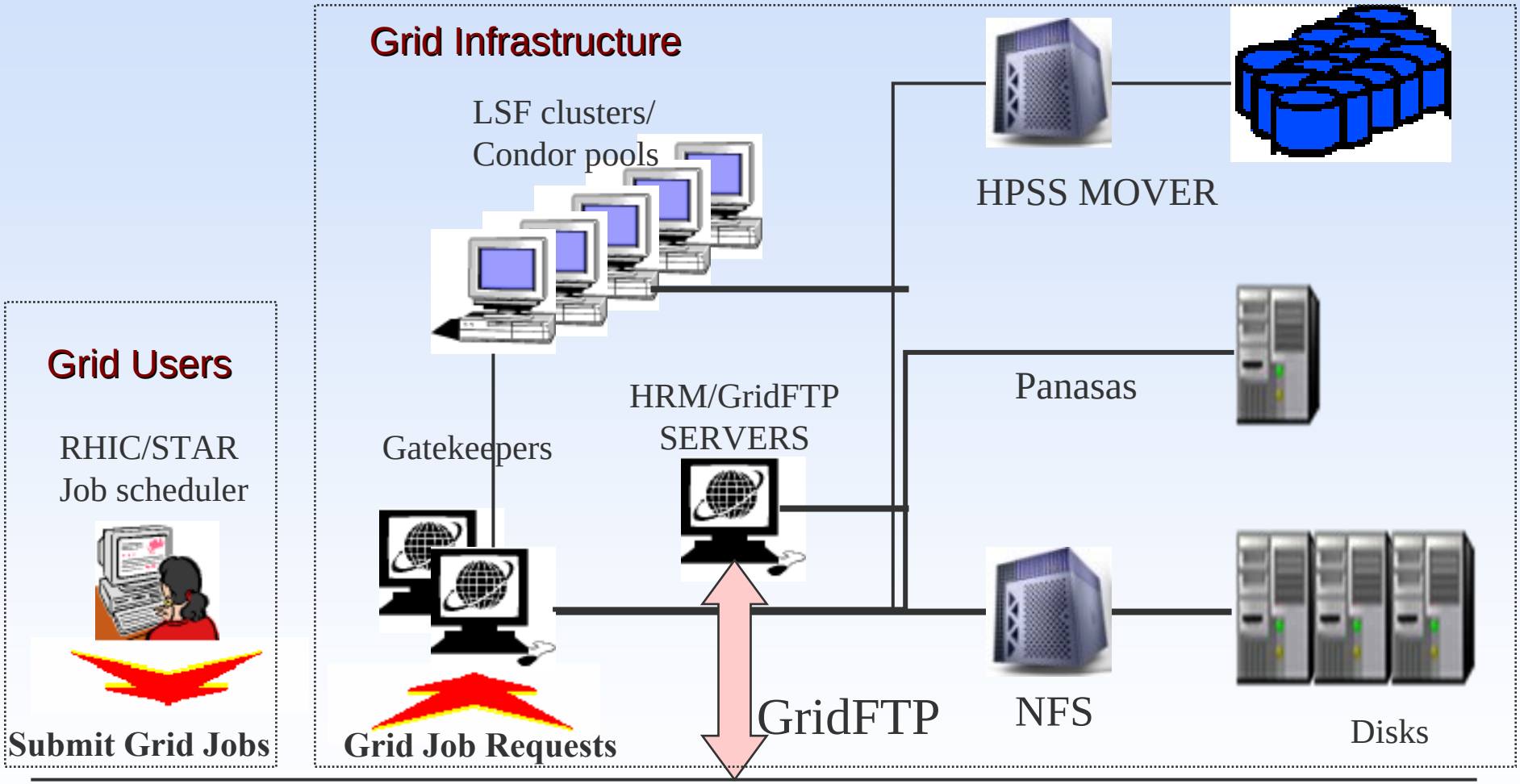
Volume of Data (TB) Stored in HPSS over Six Years



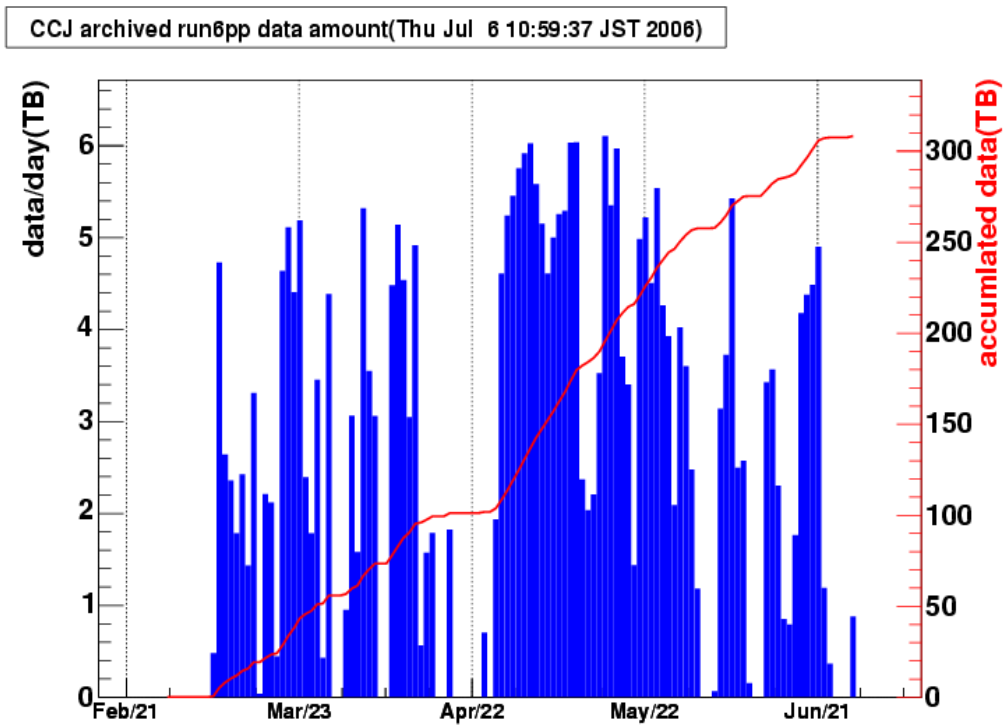
Grid and Network Services

- Computing models of RHIC experiments predate the Grid
  - So, unlike ATLAS, they were not originally based on a Grid computing model
  - However the desire to utilize substantial distributed resources is driving evolution of these models toward Grid computing
    - Originally for simulation but now more generally for analysis
      - LBNL, etc. for STAR
      - Riken, etc. for PHENIX
- Same staff engaged in major US ATLAS Grid effort also supports RHIC wide area distributed computing with …
  - Support for Grid tools and services as well as network expertise
    - GridFTP, SRM, etc
    - High volume network transfer optimization
  - Support for involvement (of STAR) in Open Science Grid
    - OSG software deployment and integration of resource into OSG
    - OSG administration: registration, authorization, etc.

RHIC Grid Configuration at BNL



PHENIX Data Transfer to CCJ



Courtesy of Y. Watanabe

- Total data transferred to CCJ (Computer Center in Japan) was 309 TB of Raw Data
- Average transfer rate was ***~3TB/day! And 6TB/day max speed.***
- RCF supplied Network/Grid expertise includeing:
  - Installation of Grid transfer and monitoring tools
  - Trouble shooting and tuning of network path
  - Integration dCache/SRM (Storage Resource Manager) for future transfers

Wide Area Network

- Jan '06 WAN upgrade of BNL connectivity

  OC48 (2.5 Gb/sec) → 2 x λ (20 Gb/sec)
- Motivated by heavy network use by both RHIC and ATLAS (potential contention)
- Funded in equal part by ESnet, DOE NP, DOE HEP and BNL
- Connection still lacks desired redundancy and diversity
  - Will require significant additional funding not yet identified

Physical Infrastructure

- Last year the limit of capacities on existing floor space were reached for
  - … chilled water, cooled air, UPS power, and power distribution
- Therefore this year for first time major physical infrastructure improvements were needed
  - 250KW of local UPS / PDU systems to support new procurements
  - New chilled water feed
  - Local rack top cooling for new procurements
  - These are now installed and operational with costs covered by Lab GPP funds
- Approaching limit of available floor space itself
  - Raised floor with fire detection & suppression, physical security
  - ATLAS expansion, with limited equipment retirement, is driving space growth (while both programs continue power/cooling growth)
  - Reallocation of space to RCF/ACF will allow 2007 and 2008 expansion
    - Additional power & cooling will be needed each year
- Need for expansion of space itself in 2009 and beyond is fully recognized by Lab management but detailed plan for addressing this need is not yet in place

Cyber Security

- Facility is a firewall protected enclave within the BNL firewall protect site
- Good working relationship with BNL ITD cyber security
- Facility has a ~single sign-on Kerberos base authentication infrastructure
- Cyber security lapses of the last year
  - All a result of externally sniffed passwords (network or keystroke)
  - One incident resulted in a root level comprise of a single machine but none result in the utilization of facility resources against other sites
- Major current efforts
  - Contributing to new BNL Cyber Security Program Plan
  - Deploying, facility wide, ***Ordo*** (BNL developed host based configuration tracking/auditing tool for Unix-like systems)
  - Preparing for expect move to non-reusable (one time) passwords in next two months (addressing sniffed password vulnerability)
- Concern remains of potential future conflicts between Grid requirements, regulatory requirements, and a cyber security policy/architecture which does not disrupt effective facility use

Conclusions

- Plans to evolve and expand facility services to meet expected needs
  - Are based on so far successful adjustments of technical directions
  - Remain within the mainstream of the HENP computing
  - Will require agreed and planned for increases in 2008 and beyond
- Physical infrastructure needs are recognized but long term plans are not yet fully in place
  - Specifically the facility will needs new space with appropriate characteristics and services for 2009 and beyond
  - Immediate annual increases in power/cooling capacity and costs are ongoing and are being addressed
- The Grid is likely to significantly change future computing
  - Based on ATLAS pathfinder experience at BNL, RCF should be well positioned to exploit its advantages
- Cyber Security remains a looming presence