U.S. ATLAS Computing Facilities

Bruce G. Gibbard

GDB Meeting

16 March 2005

BROOKHAVEN
NATIONAL LABORATORY

US ATLAS Facilities

❄ **Grid3/OSG Connected Resources Including …**

- ❑ Tier 1 Facility at Brookhaven
- ❑ Tier 2 Facilities
    - ⌘ 2 Prototype Tier 2's in operation
        - ⬩ **Indiana-Chicago**
        - ⬩ **Boston**
    - ⌘ 3 (of 5) Permanent Tier 2 sites recently selected
        - ⬩ **Boston-Harvard**
        - ⬩ **Southwest (UTA, OU, UNM, LU)**
        - ⬩ **Midwest (Chicago-Indiana)**
- ❑ Other Institutional (Tier 3) Facilities Active
    - ⬩ **LBNL, Michigan, ANL, etc.**

❄ **Associated Grid/Network Activities**

- ❑ WAN Coordination
- ❑ Program of Grid R&D
    - ⬩ **Based on Work of Grid Projects (**PPDG, GriPhyN, iVDGL, EGEE, etc.**)**
- ❑ Grid Production & Production Support

Tier 1 Facility

❄**Primary Tier 1 Functions**

- ❑ Archive and perform post calibration (and all later) reconstruction of a share of the raw data

- ❑ Group level programmatic analysis passes

- ❑ Store, reprocess, and serve ESD, AOD, TAG & DPD sets

❄**Collocated and Cooperated with RHIC Computing Facility**

- ❑ Combined 2005 capacities …
  - ⌘ CPU – 2.3 MSI2K (3100 CPU's)
    - ◦ RHIC – 1.8 MSI2K (2600 CPU's)
    - ◦ ATLAS – 0.5 MSI2K (500 CPU's)
  - ⌘ Disk – 730 TBytes
    - ◦ RHIC – (220 Central Raid + 285 Linux Dist) TBytes
    - ◦ ATLAS – (25 Central Raid + 200 Linux Dist) TBytes
  - ⌘ Tape – 4.5 PBytes (~2/3 full)
    - ◦ RHIC – 4.40 PBytes
    - ◦ ATLAS – ~100 TBytes

Tier 1 Facility 2005 Evolution

* **Staff Increase**
  - ❑ By 4 FTE's, 3 to be Grid/Network development/support personnel
  - ❑ Staff of 8.5 in 2004 to 12.5 by end of year

* **Equipment Upgrade** - Expected to be operational by end of April
  - ❑ CPU Farm: 200 kSI2k → 500 kSI2k
    - ⌘ 128 x (2 x 3.1 GHz, 2 GB, 1000 GB) … so also ~130 TB local disk
  - ❑ Disk: Focus this year only on disk distributed on Linux procurement
    - ⌘ Total ~225 TB = ~200 distributed + 25 central

Operation Activities at Tier 1

❄ Grid3 based ATLAS Data Challenge 2 (DC2) production
  ❑ Major effort over much of 2004: allocated ~70% of resource

❄ Grid3/OSG based Rome Physics Workshop production
  ❑ Major effort of early 2005

❄ Increasing general US ATLAS use of facility
  ❑ Priority (w/ preemption) scheme: DC2-Rome/US ATLAS/Grid3 jobs.
  ❑ Resource allocation (w/ highest priority): DC2-Rome/US ATLAS : 70/30%

BROOKHAVEN
NATIONAL LABORATORY
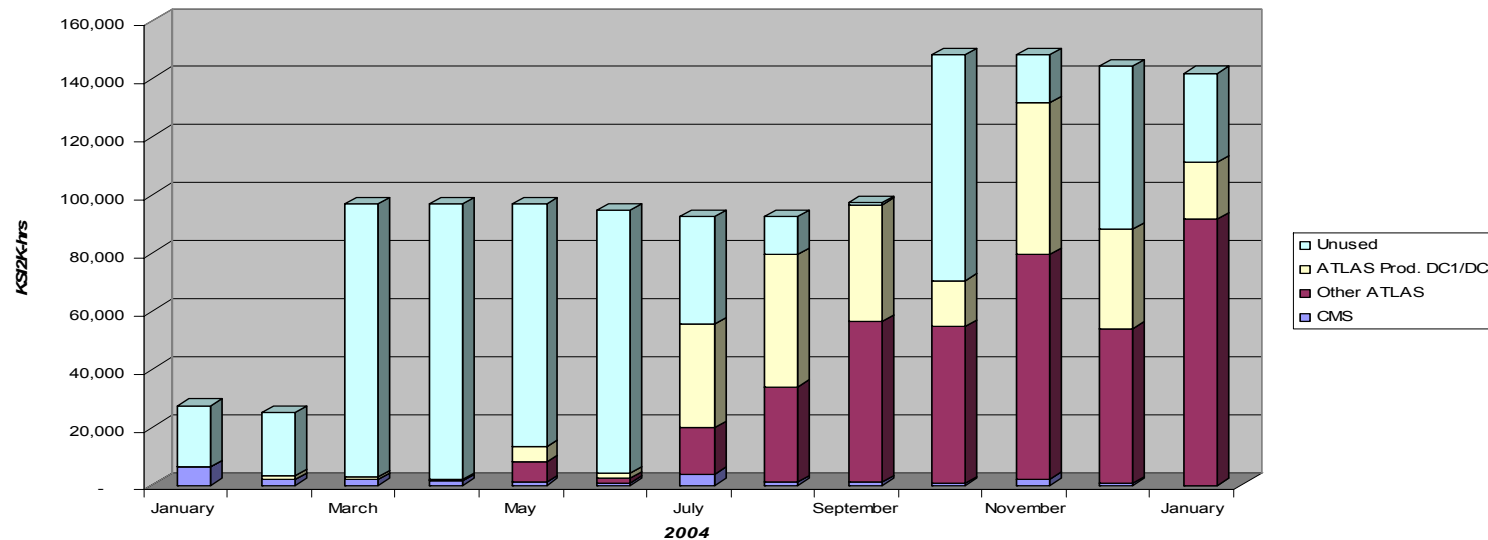
Tier 2 Facilities

❄ **Tier 2 Functions**

- ❑ Primary ATLAS resource for simulation

- ❑ Primary ATLAS location for final analyses

❄ **Expect there to be 5 Permanent Tier 2's**

- ❑ In aggregate will be comparable to the Tier 1 with respect to CPU and disk

- ❑ Defined scale for US Tier 2's
    - ⌘ Standard Tier 2's supported at a level of
        - **~2 FTE's plus MST**
        - **Four year refresh for ~1000 CPU's plus infrastructure**
    - ⌘ 2 Special "Cyber Infrastructure" Tier 2C's will receive additional funding but also additional responsibilities

❄ **Selection of first 3 (of 5) permanent sites recently announced**

- ❑ Boston Univ. & Harvard Univ. – Tier 2C

- ❑ Midwest (Univ. of Chicago & Indiana Univ.) – Tier 2

- ❑ Southwest (Univ. of Texas at Arlington, Oklahoma Univ., Univ. of New Mexico, Langston Univ.) – Tier 2C

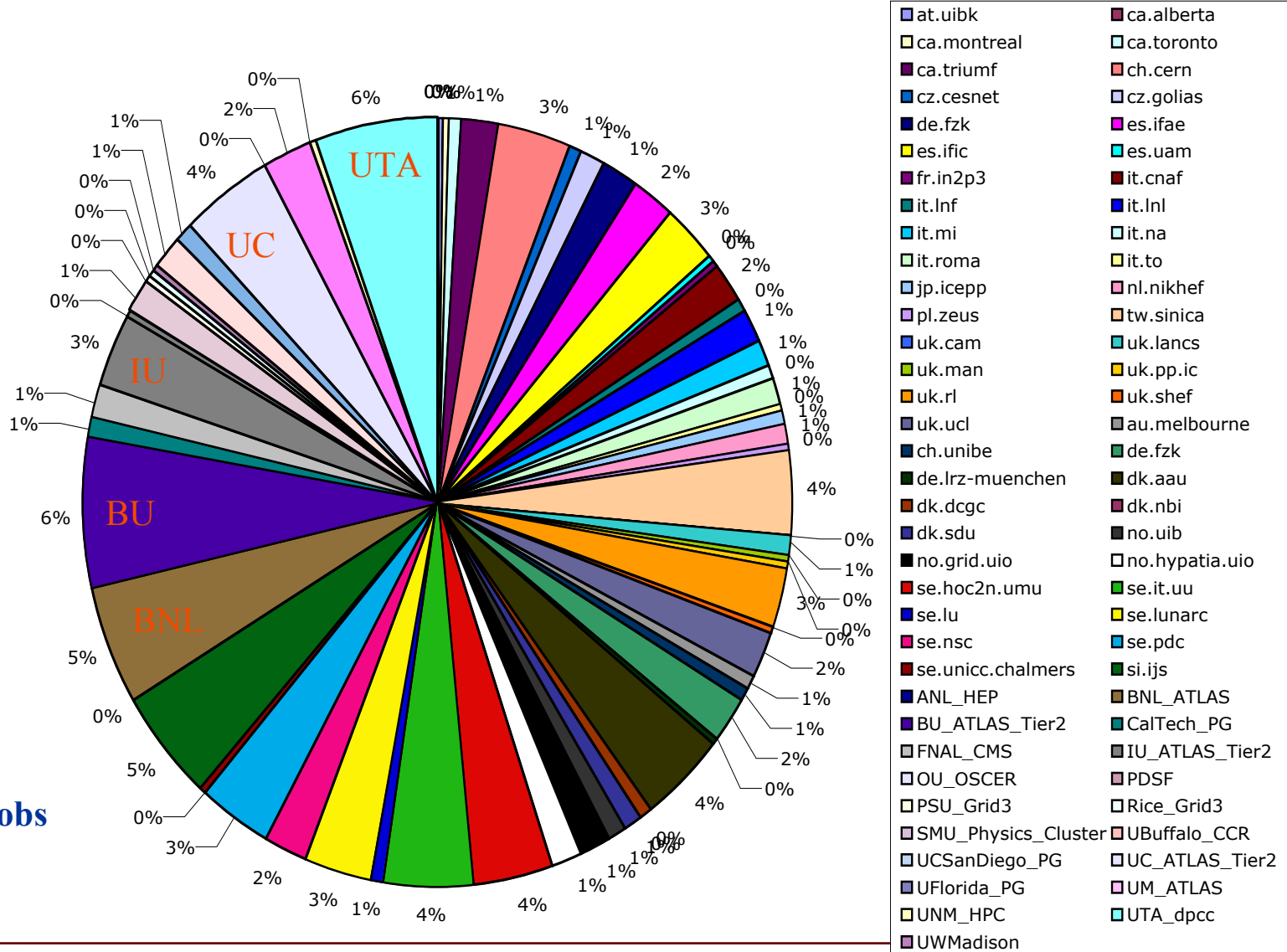❄ **Remaining 2 sites to be selected in 2006**

Prototype Tier 2 Centers

❋**Prototype Tier 2's in operation as part of Grid3/OSG**

❑ Principle US contributors to ATLAS DC1 and DC2

❑ Currently major contributors to production in support of Rome Physics Workshop

❋**Aggregate capacities similar to Tier 1**

	2004		2005	
	CPU (KSI2K)	Disk (TB)	CPU (KSI2K)	Disk (TB)
Boston Univ.	91	14	128	30
Univ. of Chicago	141	16	141	16
Indiana Univ.	78	10	78	10
TOTAL Tier 2's	310	39	347	55
Tier 1	*200*	*75*	*500*	*225*

DC2 Jobs Per Site

69 sites
~276000 Jobs

Legend:
- at.uibk
- ca.alberta
- ca.montreal
- ca.toronto
- ca.triumf
- ch.cern
- cz.cesnet
- cz.golias
- de.fzk
- es.ifae
- es.ific
- es.uam
- fr.in2p3
- it.cnaf
- it.lnf
- it.lnl
- it.mi
- it.na
- it.roma
- it.to
- jp.icepp
- nl.nikhef
- pl.zeus
- tw.sinica
- uk.cam
- uk.lancs
- uk.man
- uk.pp.ic
- uk.rl
- uk.shef
- uk.ucl
- au.melbourne
- ch.unibe
- de.fzk
- de.lrz-muenchen
- dk.aau
- dk.dcgc
- dk.nbi
- dk.sdu
- no.uib
- no.grid.uio
- no.hypatia.uio
- se.hoc2n.umu
- se.it.uu
- se.lu
- se.lunarc
- se.nsc
- se.pdc
- se.unicc.chalmers
- si.ijs
- ANL_HEP
- BNL_ATLAS
- BU_ATLAS_Tier2
- CalTech_PG
- FNAL_CMS
- IU_ATLAS_Tier2
- OU_OSCER
- PDSF
- PSU_Grid3
- Rice_Grid3
- SMU_Physics_Cluster
- UBuffalo_CCR
- UCSanDiego_PG
- UC_ATLAS_Tier2
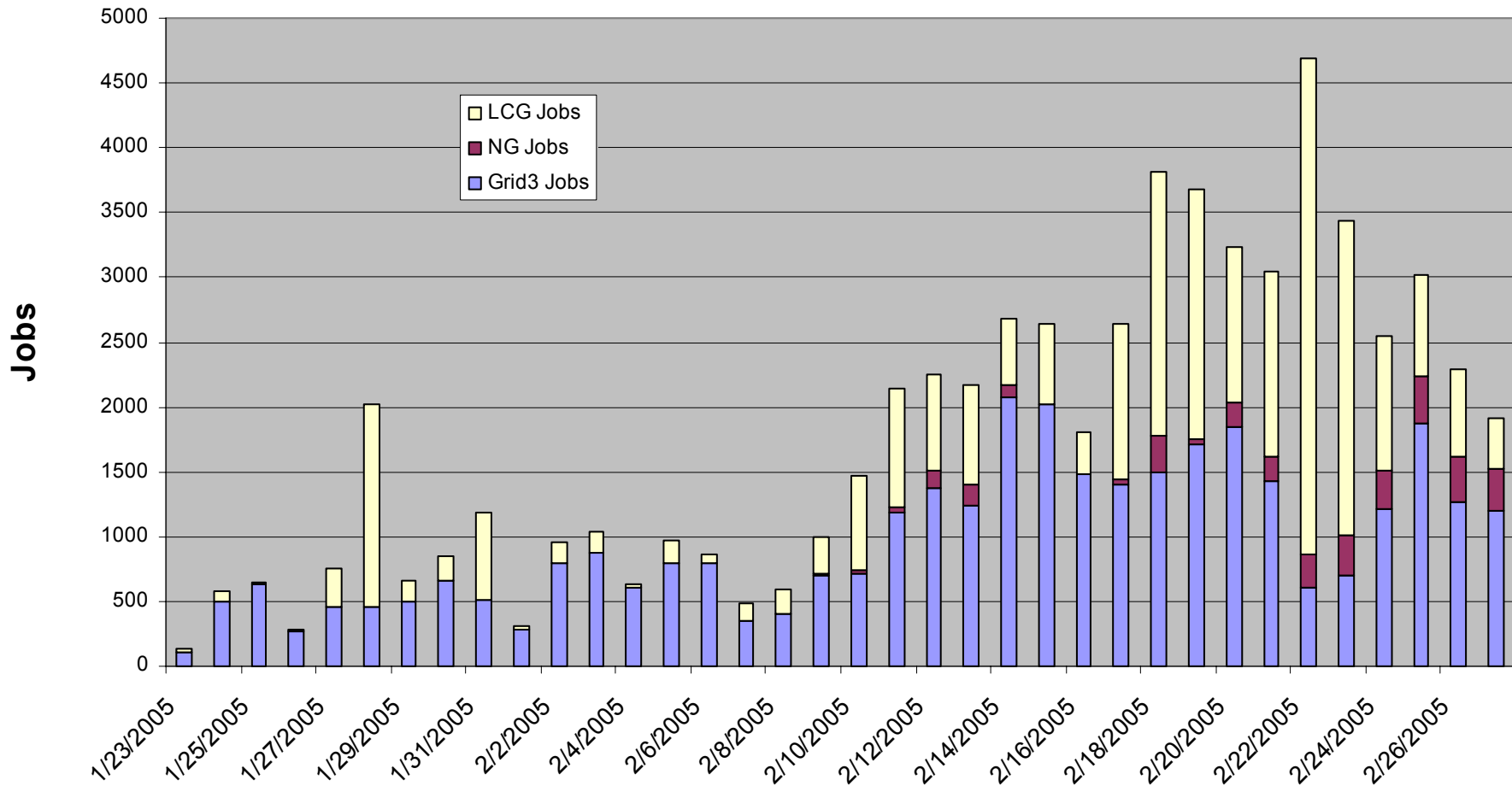- UFlorida_PG
- UM_ATLAS
- UNM_HPC
- UTA_dpcc
- UWMadison

Rome Production Ramp-up

Rome Production

Grid / OSG Related Activities

❋ **Two SRM's Tested, Deployed, and Operational at Tier 1**

- ❑ HRM/SRM from LBNL: HPSS capable out of the box

- ❑ dCache/SRM from Fermilab/DESY: dCache compatible out of the box

- ❑ Interoperability issues between above addressed and basic level of compatibility with CASTOR/SRM verified

- ❑ Evaluation of appropriate SRM choice for Tier 2's underway

❋ **Cyber Security and AAA for VO Management Development at Tier 1**

- ❑ GUMS: a Grid identity mapping service working with suite including VOM/VOMS/VOX/VOMRS

- ❑ Consolidation of ATLAS VO registry with US ATLAS as a subgroup.

- ❑ Privilege management project underway in collaboration with Fermilab/CMS
  - ⌘ Dynamic assignment of local access.
  - ⌘ Role based authorization.
  - ⌘ Faster policy implementation.

Grid / OSG Related Activities (2)

❄ **OSG Integration and Deployment Activities**

- ❏ US ATLAS Tier 1 and Tier 2's are part of Integration Test Bed
- ❏ Where OSG components are tested and certified for deployment

❄ **LCG Deployment**

- ❏ Using very limited resources and only modest effort
  - ⌘ Nov '04: LCG 2.2.0
  - ⌘ Recently completed upgrade to LCG 2.3.0
  - ⌘ Now extending upgrade to use SLC
- ❏ Primarily function is to facilitate comparisons and interoperability studies between LCG and OSG

Grid / OSG Related Activities (3)

❈ "Terapaths" Project

- ❑ Addressing issues of contention between projects and activities with in projects on shared IP network infrastructure

- ❑ Integrat LAN QoS with WAN MPLS for end-to-end network resource management

- ❑ Initial tests on both LAN and WAN have been performed

❈ Effectiveness of LAN QoS has been demonstrated on BNL backbone

Network QoS with Three Classes: Best Effort, Class 4 and EF

Legend: Best Effort — Class 4 — Express Forwarding — TOTAL — Wire S
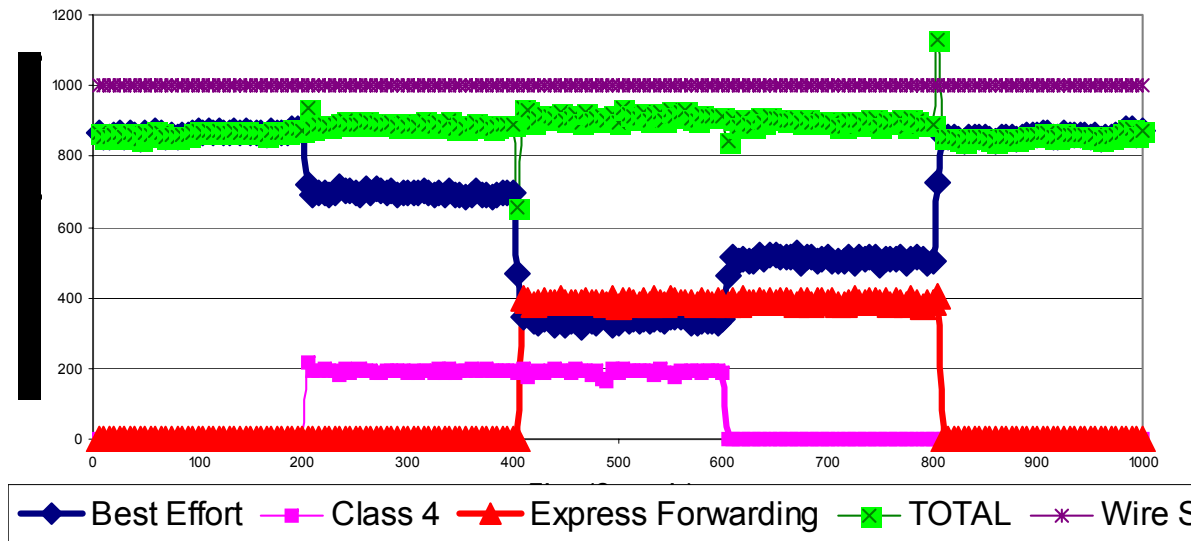
Service Challenge 1

❊ US ATLAS is very interested in Service Challenge participation

- ❑ Very interest, technically capable and committed staff at Tier 1
- ❑ Fully prepared in terms of hardware and software
- ❑ … though no 10 Gb/sec connectivity at this time

❊ Proposed schedules and statements of readiness continue to be "effectively" ignored by Service Challenge coordinators leading to substantial frustration

❊ However, in conjunction with LCG Service Challenge 1, 2 RRD GridFTP servers doing bulk disk-to-disk data transfers achieved:

- ❑ BNL / CERN: ~100 MByte/sec. Limited by site backbone bandwidth
- ❑ Testing limited at 12 hours (night time) to not impact other BNL projects
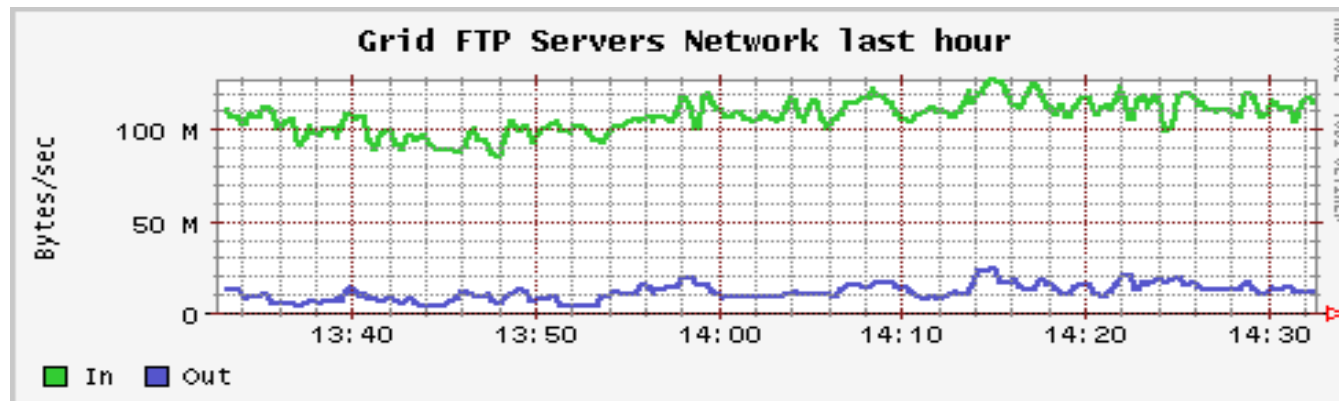
Service Challenge 2 ...

❄ **BNL SC2 Configuration**

   ❑ OC48 (2.5Gb/sec) site connection … but current internal 1000 Mb/sec limit

   ❑ dCache head node and 4 pool nodes (1.2 TB of disk) & HPSS backend store

   ❑ Using dCache-SRM

❄ **Proposed BNL SC2 Goals**

   ❑ Week of 14 March: Demonstrate functionality

      ⌘ Verify full connectivity, interoperability and performance - Completed 14 March !

   ❑ Week of 21 March: Demonstrate extended period performance

      ⌘ SRM/SRM transfers for 12 hours at 80% of available bandwidth, 80-100 MB/sec

   ❑ Week of 28 March: Demonstrate robustness

      ⌘ SRM/SRM transfers for full week at 60% of available bandwidth, 75 MB/sec

❄ **SC3 ...**

      ⌘ Full OC48 capacity (2.5 Gb/sec) will be available at Tier 1

      ⌘ Limited time 100 MB/sec tape capacity if coordinated with RHIC experiments

      ⌘ Strong interest in Tier 2 participation