

# BNL Tier1 Report

- Worker nodes
  - Tier 1: added 88 Dell R430 nodes
  - Tier 3: added 11 Dell R430 nodes
- Easy Condor Quota Gui
- dCache Storage
  - DDN will be retired
  - Hitachi G400 will be added
  - Replication of data have been enabled for certain storage.
    - Transparent to users
  - NFS4.1 enabled using different NFS door
    - Newer kernel on clients
    - New dCache version and change in the configuration
- Tier 3
  - Currently about 70 users with condor queues and storage space.
- Ceph
- HPSS Tape

# Condor Group Quota Web Gui

<https://github.com/fubarwrangler/group-quota>

By William Strecker Kellogg

The screenshot shows the GitHub repository page for 'fubarwrangler / group-quota'. The repository is described as 'HTCondor Accounting Group Software'. It has 337 commits, 2 branches, 0 releases, and 1 contributor. The current branch is 'master'. The repository is being edited by 'fubarwrangler' with the commit message 'Editing per-user group limitations working!'. The latest commit is c879d37, made 8 days ago. The repository contains several files and folders, including 'Documentation', 'atlas-idle', 'bin', 'gq', 'gqweb', 'scripts', '.gitignore', 'LICENSE', 'README.adoc', 'init\_db.py', 'run\_webpage.py', 'setup-gq.py', and 'setup-web.py'. The README.adoc file is open, showing the title 'Accounting-Group Quota management for HTCondor' and the description 'An HTCondor Account Group manager, used by the US ATLAS TeI-1 at BNL.'

GitHub This repository Search Explore Features Enterprise Pricing Sign up Sign in

fubarwrangler / group-quota Watch Star Fork

Code Issues Pull requests Pulse Graphs

HTCondor Accounting Group Software

337 commits 2 branches 0 releases 1 contributor

Branch: master New pull request New file Find file HTTPS https://github.com/fubarwrangler Download ZIP

fubarwrangler Editing per-user group limitations working! Latest commit c879d37 8 days ago

Documentation	Move PANDA-specific stuff out of gq package	24 days ago
atlas-idle	Skeleton user->group relationship and editor for perms	16 days ago
bin	Move PANDA-specific stuff out of gq package	24 days ago
gq	Move PANDA-specific stuff out of gq package	24 days ago
gqweb	Editing per-user group limitations working!	8 days ago
scripts	Add new PANDA queue fix some t3-related scripts	3 months ago
.gitignore	Setup.py for gq library and webpage	7 months ago
LICENSE	Restructure website code a bit	8 months ago
README.adoc	Improve some docs	25 days ago
init_db.py	WIP Tier-3 Group Editor page	5 months ago
run_webpage.py	Re-order folders into group-library and webpage	7 months ago
setup-gq.py	Bump version number and fix name	3 months ago
setup-web.py	Fix instance name and bump version	7 months ago

README.adoc

## Accounting-Group Quota management for HTCondor

An HTCondor Account Group manager, used by the US ATLAS TeI-1 at BNL.

# Condor Group Quota Web Gui

## ATLAS HTCondor Groups

This page allows you to add and remove groups from the group tree (Add/Remove Groups), edit any parameter of those groups (Edit Groups), and edit the quota at one level of the tree without v (Quota EZ-Edit).

Queue

Quota

Overflow

Name	Quota	Prio	Weight	Surplus	Busy
group_atlas	15793	10.0	2.0	×	11610
group_atlas.analysis	1002	10.0	2.0	×	998
group_atlas.analysis.amc	16	10.0	0.0	×	16
group_atlas.analysis.long	493	10.0	1.0	✓	493
group_atlas.analysis.short	493	10.0	1.0	✓	489
group_atlas.prod	14785	10.0	4.0	✓	10612
group_atlas.prod.mp	3407	10.0	8.0	×	2440
group_atlas.prod.production	5151	10.0	1.0	×	920
group_atlas.prod.test	6226	10.0	2.0	✓	7252
group_atlas.prod.xl	1	10.0	16.0	×	0
group_atlas.software	6	5.0	1.0	×	0
group_grid	500	2000.0	1.0	×	500

Quotas EZ-Edit

Edit Groups

Add/Remove Groups

Easy Edit

# Condor Group Quota Web Gui

EZEditor for group

Name	Fix	Quota	Slider
group_atlas.prod.test	<input type="checkbox"/>	✎ 6226	0
group_atlas.prod.production	<input type="checkbox"/>	✎ 5151	0
group_atlas.prod.mp	<input type="checkbox"/>	✎ 3407	0
group_atlas.prod.xl	<input type="checkbox"/>	✎ 1	0
		<hr/>	
		14785	

Back Submit

Help

Move the sliders above to adjust what proportion of the total quota each group gets. Select the *Fix* box to fix that group's quota at the current value (you can't leave fewer than 2 free sliders). Click on the quota field or the "edit" pencil next to it to manually input a quota — the up and down arrows increment and decrement by one while hitting enter or clicking off sets the value

# Tier 3 Setup

- Separate Condor queues
  - User belongs to an institution
  - Each institution has own group.
- Separate Storage
  - dCache
    - Part of production dCache
    - Own 5TB space
    - Write once
    - Access by https (browser), Webdav, XRoot, NFS 4.1/PNFS, gridftp, srm.
  - GPFS
    - Own space
- NX Desktop
  - Remote desktop
  - Mac, Widows, Linux, Android supported

# Tier 3 Condor groups

## Tier 3 HTCondor Groups

This page allows you to create, alter, and remove groups from the group tree used by HTCondor in ATLAS. There are three areas of this webpage, one that allows you to create and remove groups (Add/Remove Button), one that lets you edit any parameter of those groups (Edit button), and one that allows easy editing of the quota at one level of the tree without violating other levels (EZ-Edit button).

















Name	Quota	Prio	Weight	Surplus	Busy
group_atlas	973	10.0	0.0	✘	925
group_atlas.albany	42	10.0	0.0	✔	0
group_atlas.anl	42	10.0	0.0	✔	0
group_atlas.arizona	41	10.0	0.0	✔	0
group_atlas.birm	41	10.0	0.0	✔	0
group_atlas.bnl	135	10.0	0.0	✔	925
group_atlas.csu	41	10.0	0.0	✔	0
group_atlas.duke	42	10.0	0.0	✔	0
group_atlas.general	10	10.0	0.0	✘	0
group_atlas.irvine	41	10.0	0.0	✔	0
group_atlas.lbl	41	10.0	0.0	✔	0
group_atlas.louis	41	10.0	0.0	✔	0
group_atlas.nyu	41	10.0	0.0	✔	0
group_atlas.oregon	41	10.0	0.0	✔	0

Each institution has own group with easily adjustable quota.  
The group quota can overflow to the parent quota.



# Tier 3 Condor groups

EZEditor for group\_atlas

Name	Fix	Quota	Slider	
group_atlas.duke	<input type="checkbox"/>	✍ 42	0	 973
group_atlas.anl	<input type="checkbox"/>	✍ 42	0	 973
group_atlas.general	<input type="checkbox"/>	✍ 10	0	 973
group_atlas.sunysb	<input type="checkbox"/>	✍ 42	0	 973
group_atlas.panda	<input type="checkbox"/>	✍ 5	0	 973
group_atlas.albany	<input type="checkbox"/>	✍ 42	0	 973
group_atlas.tsukuba	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.smu	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.nyu	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.birm	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.umich	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.csu	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.umass	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.arizona	<input type="checkbox"/>	✍ 41	0	 973
group_atlas.bnl	<input type="checkbox"/>	✍ 135	0	 973
group_atlas.oregon	<input type="checkbox"/>	✍ 41	0	 973

Group quota can be easily adjusted by Web Gui.

# Storage download tool

- Deficiency of current rucio download.
  - Client host becomes proxy between the source and destination
  - Transfers are not controlled. Many simultaneous requests are DoS to source storage.
  - No catalogs or other management tools.
- New tool
  - Use Rucio and FTS RESTful APIs.
  - Use FTS for transfers
  - Provides local personal catalog in sqlite3 db.
  - Retry many sources. Retry is not automatic. It will use different source by retry.
  - Provides tools to delete
- It can be used for any other FTS aware storage.
  - DDM endpoint is not necessary.
  - It can easily use better Database. (eg. MySQL, Postgres, etc...)
  - Retry can be automated.



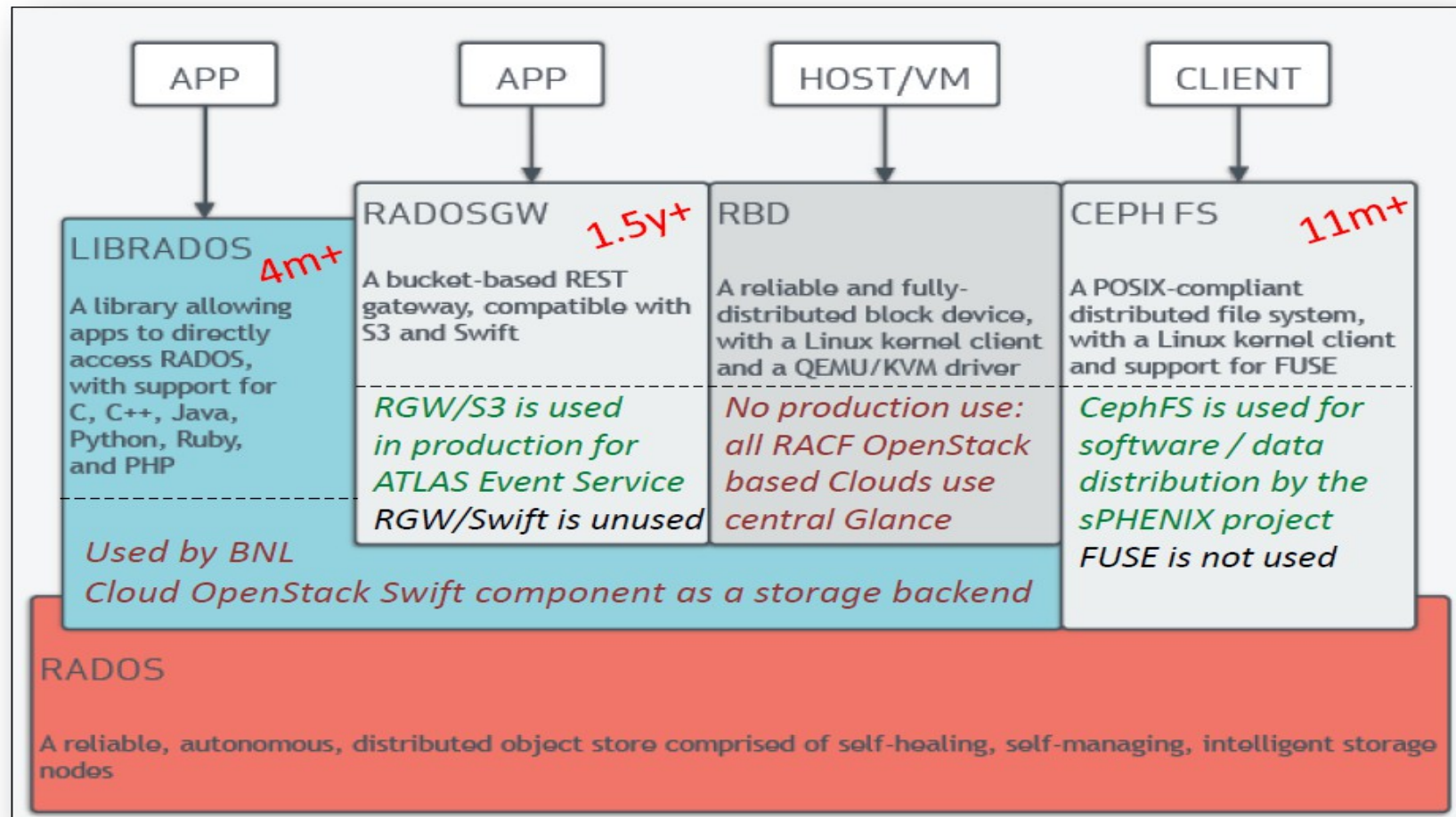
# NX Desktop

- Gnome, KDE, XDM desktop are supported.
- It provides the persistent work environment.
  - Network outage will not terminate your desktop.

The screenshot displays the NX Desktop environment. On the left, a terminal window titled "hito@joje:~" is open, showing a shell prompt "bash-4.1\$". The desktop background is a blue and white abstract pattern. On the right, the ATLAS PanDA monitoring dashboard is visible in a Mozilla Firefox browser window. The dashboard shows the ATLAS PanDA monitor home page with various navigation tabs and a main content area displaying global concurrent running job core counts. The dashboard includes several stacked area charts showing job core counts over time for different sites and activities. The top chart is titled "Global concurrent running job core counts, all sites, all job types, by cloud, last 1, 7, 30 days". The bottom chart is titled "Global concurrent running job core counts, all sites, all job types, by activity, last 1, 7, 30 days". The dashboard also shows a search bar and various menu options like Dash, Tasks, Jobs, Errors, Users, Sites, Incidents, Search, and Admin.

# Ceph Usage at BNL

By Alexander Zaytsev



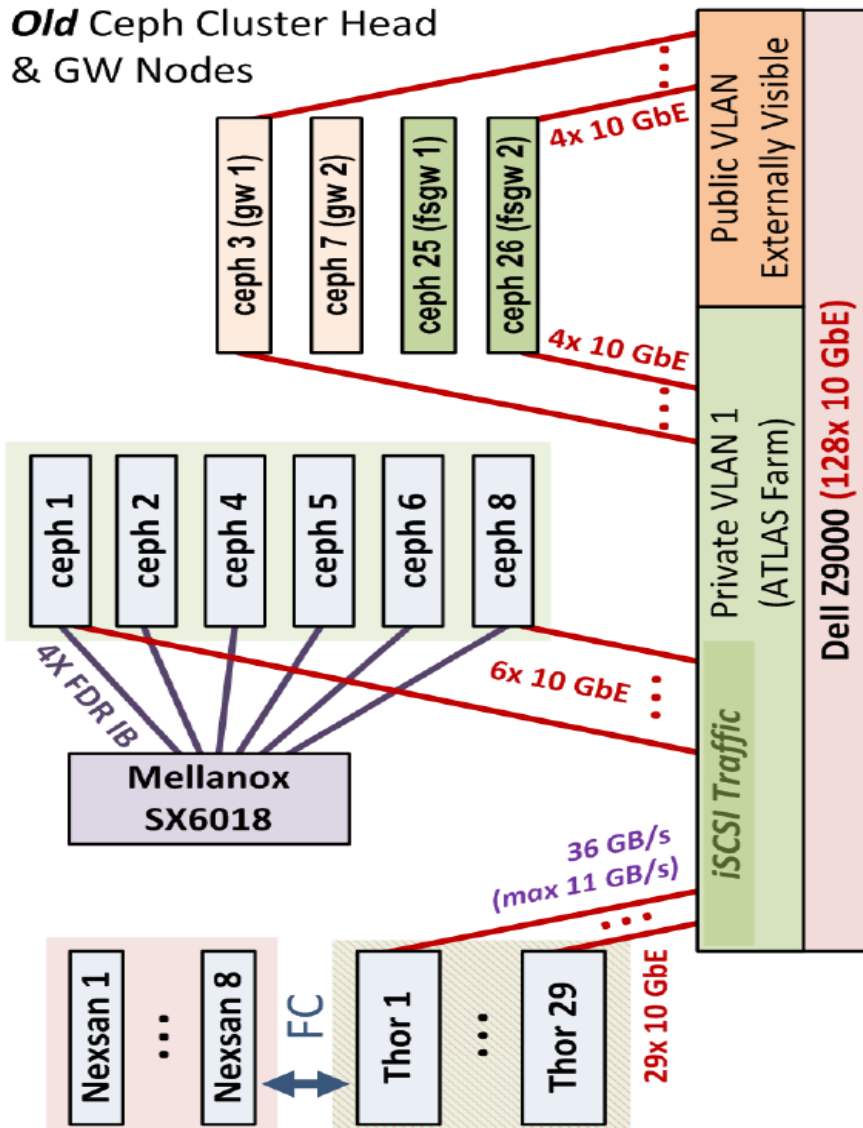
The BNL Cloud instance is the first user of our Ceph installations that utilizes the low level object store API of Ceph directly, and thus benefiting from the lowest API-driven overhead possible (still in pre-production phase).

# Current Ceph Setup

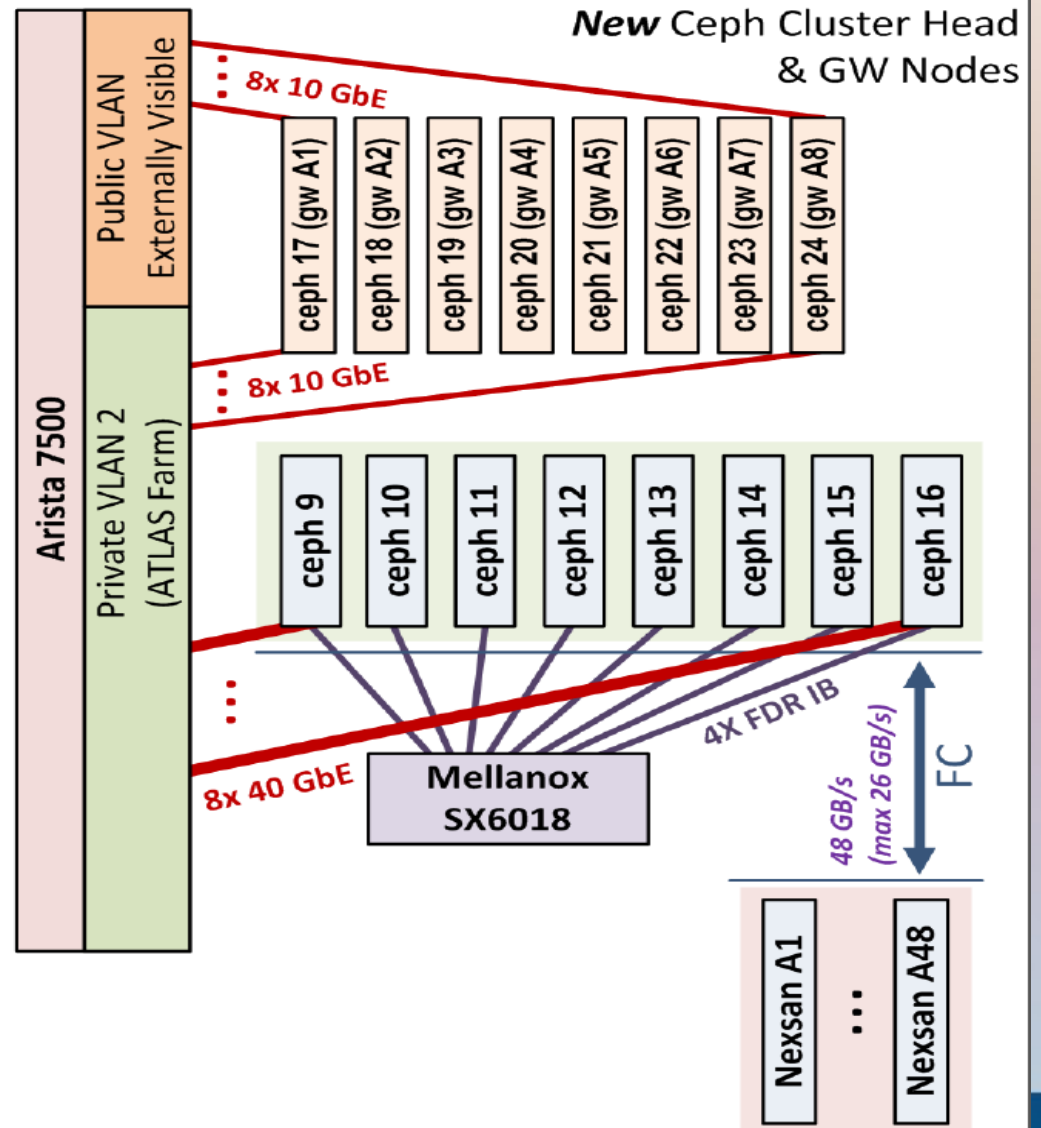
Two Ceph clusters deployed in RACF as of 2015Q4

0.6 PB + 0.4 PB usable capacity split

**Old** Ceph Cluster Head & GW Nodes

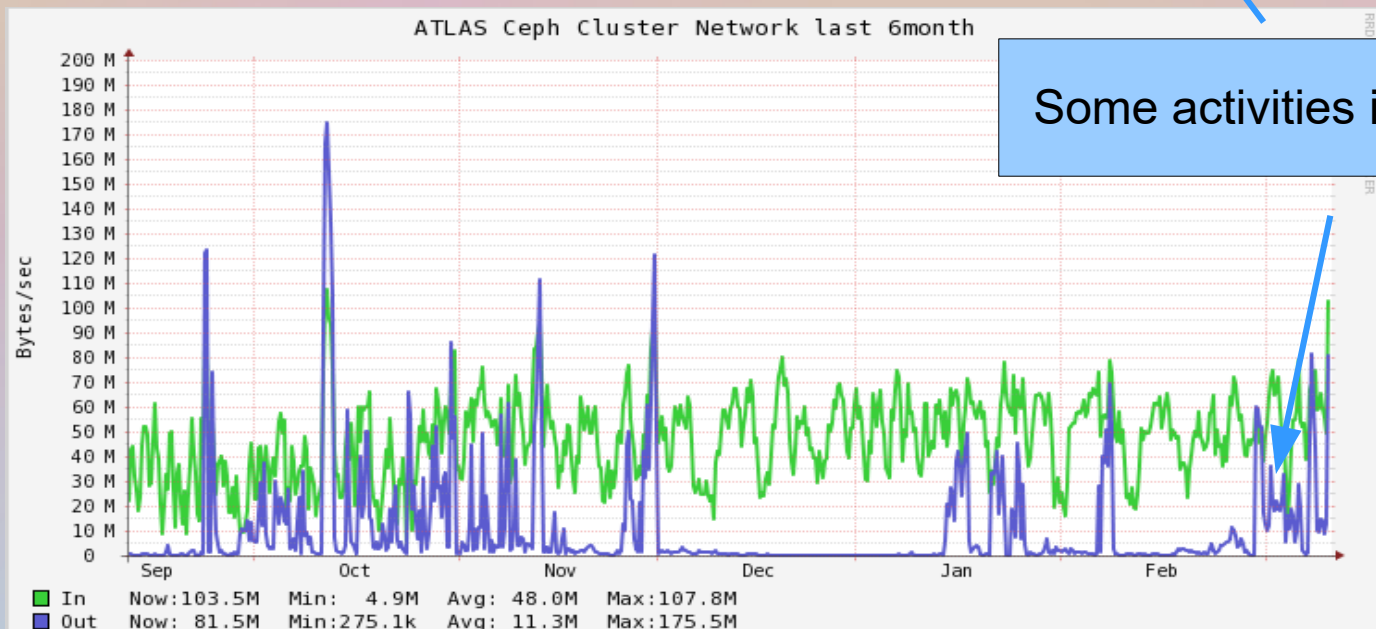
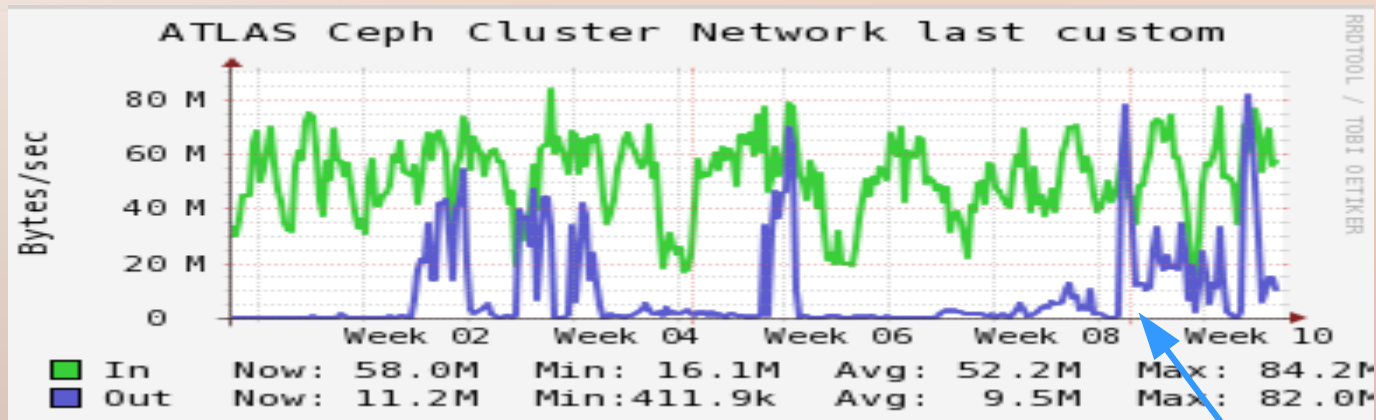


**New** Ceph Cluster Head & GW Nodes





# Ceph S3 network traffic



Some activities in the recent weeks

# Ceph Project

- Monthly Ceph Meeting
  - Share working knowledge.
  - The first meeting among CERN, BNL and RAL
  - Plan to expand a bit for other sites in WLCG
- Ceph Day during the next ATLAS Computing and Software workshop
  - <http://www.eventbrite.com/e/ceph-day-switzerland-tickets-17342577115>
- BNL
  - XRootd Cache over Ceph FS.
    - Data cache by ATLAS
    - Use of FRM
  - Ceph as front end to HPSS tape system.

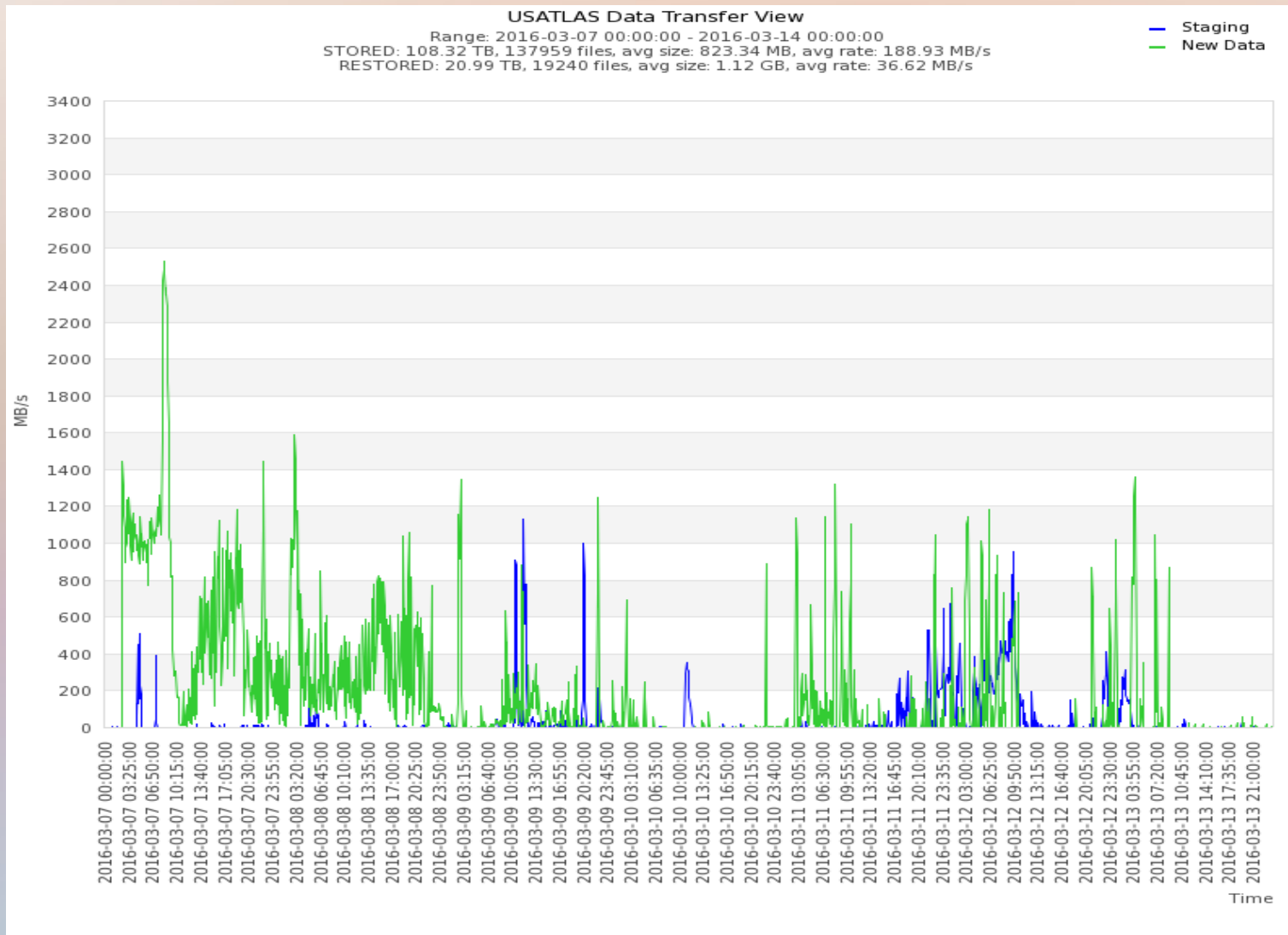
# Ceph Project

- Monthly Ceph Meeting
  - Share working knowledge.
  - The first meeting among CERN, BNL and RAL
  - Plan to expand a bit for other sites in WLCG
- Ceph Day during the next ATLAS Computing and Software workshop
  - <http://www.eventbrite.com/e/ceph-day-switzerland-tickets-17342577115>
- BNL
  - XRootd Cache over Ceph FS.
    - Data cache by ATLAS
    - Use of FRM
  - Ceph as front end to HPSS tape system.



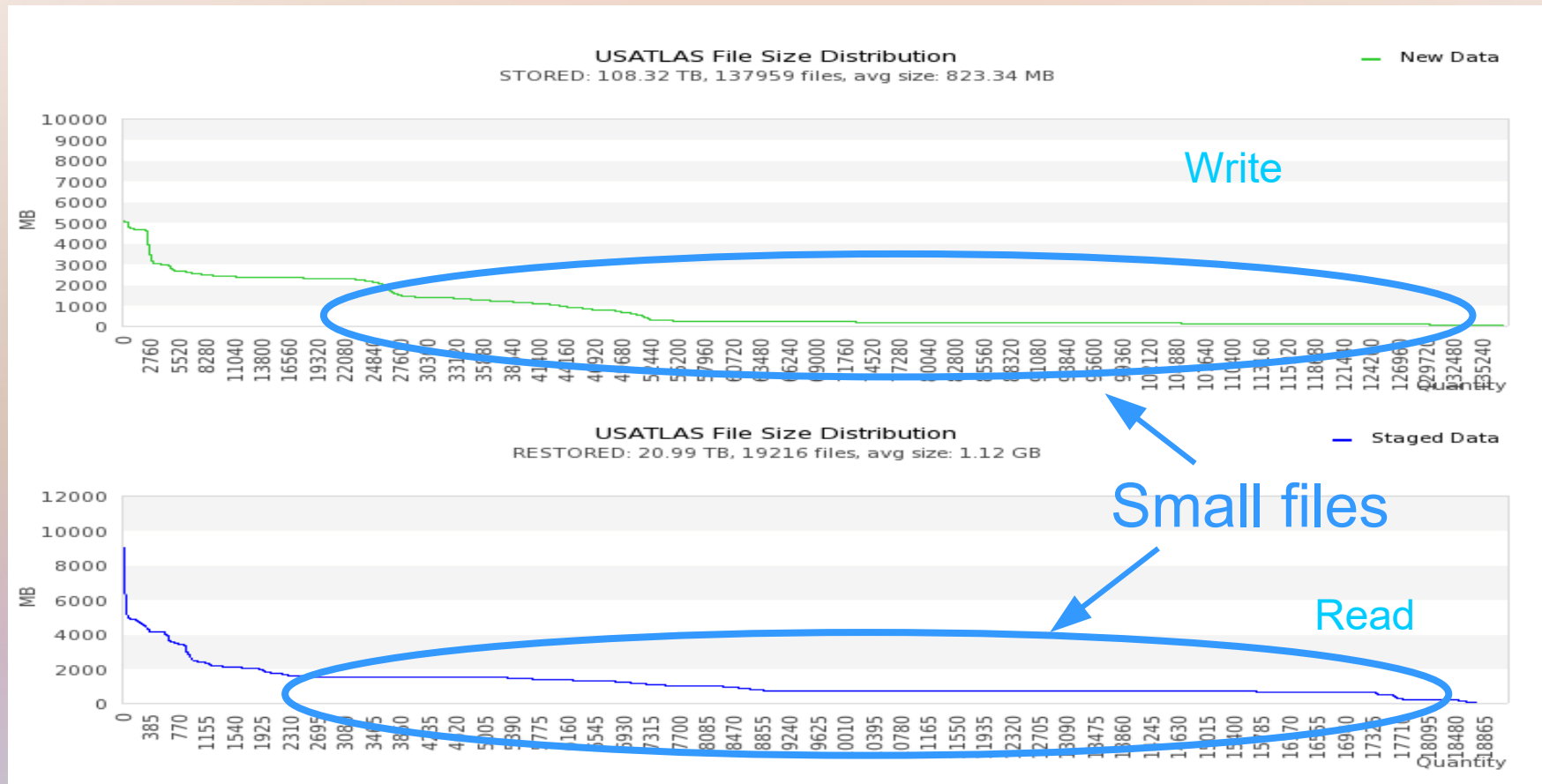
# HPSS Performance Monitor

By David Yu



- The fraction of tape drives used for reading from and writing to tape can be adjusted.
- Over 20 tape drives
- Performance varies with size of files.

# Data Size Distribution in HPSS TAPE



ATLAS use of TAPE is dominated by small files.  
The small files are not as efficient as the large ones in TAPE system.