Jet User Meeting

05 October 2011

1-2 PM MT, room 2A-305

Dial in info: 877-852-8961, 6145942#

Topics

- Quotas on /lfs0,/lfs1
- HFIP phase 4 (2011 Upgrade)
- HSMS migration to NESCC (RDHPS only)
- Zeus Update- NESCC

Quotas on lfs0/lfs1

- Small bug in group quotas affecting /lfs0 and /lfs1 has been fixed
- What does this mean for the users?
  - It means we will stop sending you emails to clean up your directories.
  - A single user or group will no longer be able to negatively impact all other users
- Steps to implementation
  - We are going to change the group ownership of all files so that /lfsX/project/$ACCT is owned by $ACCT
  - Set all directories so newly created directories will be owned by $ACCT
    - chmod g+s $DIR $ACCT

Quotas – What does it mean for the users?

- We will be changing group ownership for all projects
  - This will not affect running jobs
  - We have tested this
- Do not use –a or –p when transferring files **with cp** from the HSMS because of group ownership.
  - Use cp "—preserve=timestamps" if you want to preserve the original times
  - Using –a or –p will fail
- If you are using MssGet, the –p will work.

Checking Quotas

- Use jet_quota_report.pl

```
# jet_quota_report.pl -h

jet_quota_report <-h> (-f) <FileSystem> -g <Group> -f <User> -b

Options:
      -b -- Print usage without unites (disk usage in GB)

Output:
     Project  -- Name of the allocated project
     ResCtlQ  -- Quotas as set in the official resource control file
     DuUsage  -- The Space used as reported by the old quota tools
     DataUsed -- Data used by this project as reported by the filesystem [*]
     DQuota   -- Soft quota for data
     DLimit   -- Hard quota for data
     DGrace   -- Grace time (in seconds)
     NumUsed  -- Number of files used by this project
     FQuota   -- Soft quota for total files
     FLimit   -- Hard limit for total files
     FGrace     -- Grace time for file limit (in seconds)

     [*] The important column is the DataUsed column.  This is the enforced quota limit.
```

Checking Quotas

- Use jet_quota_report.pl

```
# jet_quota_report.pl -f /lfs0

Quota Report for FS: /lfs0

  Project ResCtlQ DuUsage | DataUsed  DQuota  DLimit   DGrace NumUsed DQuota  DLimit
DGrace
-------------------------------------------------------------------------------
-
naos-ruc   1000G    677G  |    673G       0        0        -     44K       0        0         -
wrfruc    30000G  27974G  |  24999G       0        0        -      5M       0        0         -
Rucsref     750G   1024M  |     8K       0        0        -       2       0        0         -
fim       15000G  15953G  |  17210G       0        0        -      3M       0        0         -
Rtfim      3250G   8358G  |   5944G       0        0        -    995K       0        0         -
rtrr      12000G   4966G  |   4453G       0        0        -    846K       0        0         -
Jetmgt     5000G   4421G  |   4420G    5000G    5100G       -    684K       0        0         -

  Project ResCtlQ DuUsage | DataUsed  DQuota  DLimit   DGrace NumUsed DQuota  DLimit
DGrace
-------------------------------------------------------------------------------
-
frd             -       -  |    353G       0        0        -    181K       0        0         -
nrtrr           -       -  |     48G       0        0        -     24       0        0         -
```

How quotas will be enabled

- Your soft quota will be set to your allocation
- Your hard quota will be set to 105% of your allocation
- The grace period will be set to 3 days
  - “grace period” length of time of when you can be over your soft quota but not have writes blocked
- For group accounts which are not allocated projects
  - Soft quota will be 1024 bytes (effectively 0)
  - Hard quota will be 100 GB
    - Eventually this will be tuned down to zero

Quotas - Schedule

- Oct 12th, 8am – Enable quotas for /lfs0
- Oct 19th, 8am – Enable quotas for /lfs1

Migrating User's Default Accounts

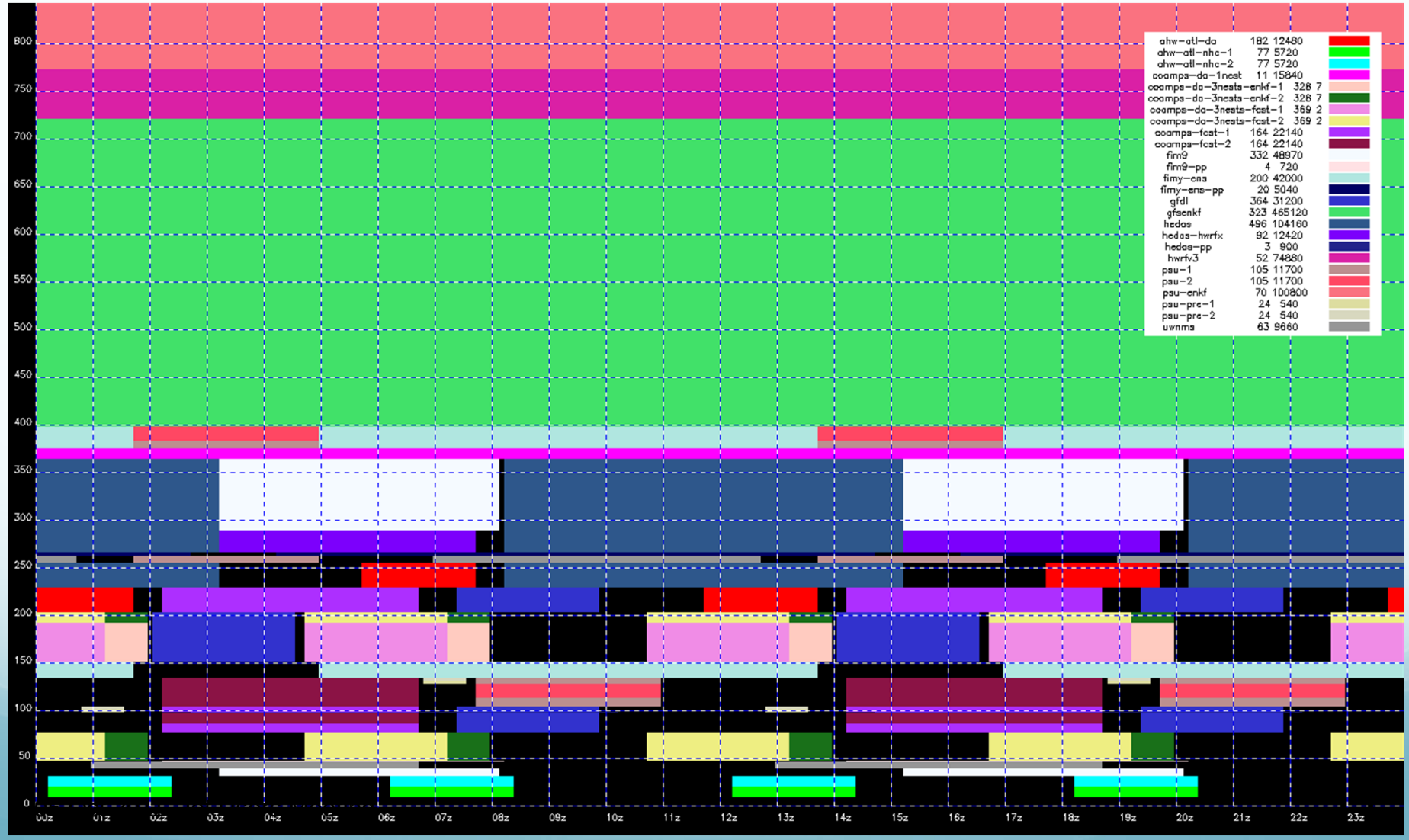
- Many users' default groups (frd, al, fsl, etc.) are not allocated accounts on Jet
  - This was due to historical reasons
- Over the next few months, we will be changing users' default projects so they are an allocated project
- We are still working on a plan that minimizes impact to the users, but it will probably work something like this:
  - Default groups will be changed to the account from which the user consumes the most of the cycles on Jet
  - Users could request an alternate group if desired
  - The change over would happen during a downtime
- You will be hearing more soon about this

Scheduling and Backlog on Jet

- It has been taking longer or jobs to get through the system.
  - This is not a problem with the system scheduler.
- The HFIP portion of the system has over 90% of it reserved for stream 1.5 experiments.
  - Experiments end Nov 1st
- More user's continue to use there allocations on the RDHPCS portion (ncomp, hcomp).
  - Utilization is around 88%.

Solution? - Acquire more equipment (uJet and Zeus)

Reservations on tJet



HFIP p4 - Storage Upgrade

- HPS upgrade: /pan2
  - ~650 TB usable, (840TB raw)
    - ~520 TB: HFIP
    - ~130 TB: DOE
  - 10-12.5 GB/s aggregate bandwidth
  - No down time anticipated for current resources
  - Available: on or before 10/28/11

HFIP p4 – Compute Upgrade

- Compute upgrade: uJet
  - 6048 cores (504 nodes)- Westmere 2.66GHz
    - Identical hardware and environment to tJet
    - No change in user environment
  - No down time anticipated for current resources
  - System documentation:

    https://jetdocs.rdhpcs.noaa.gov

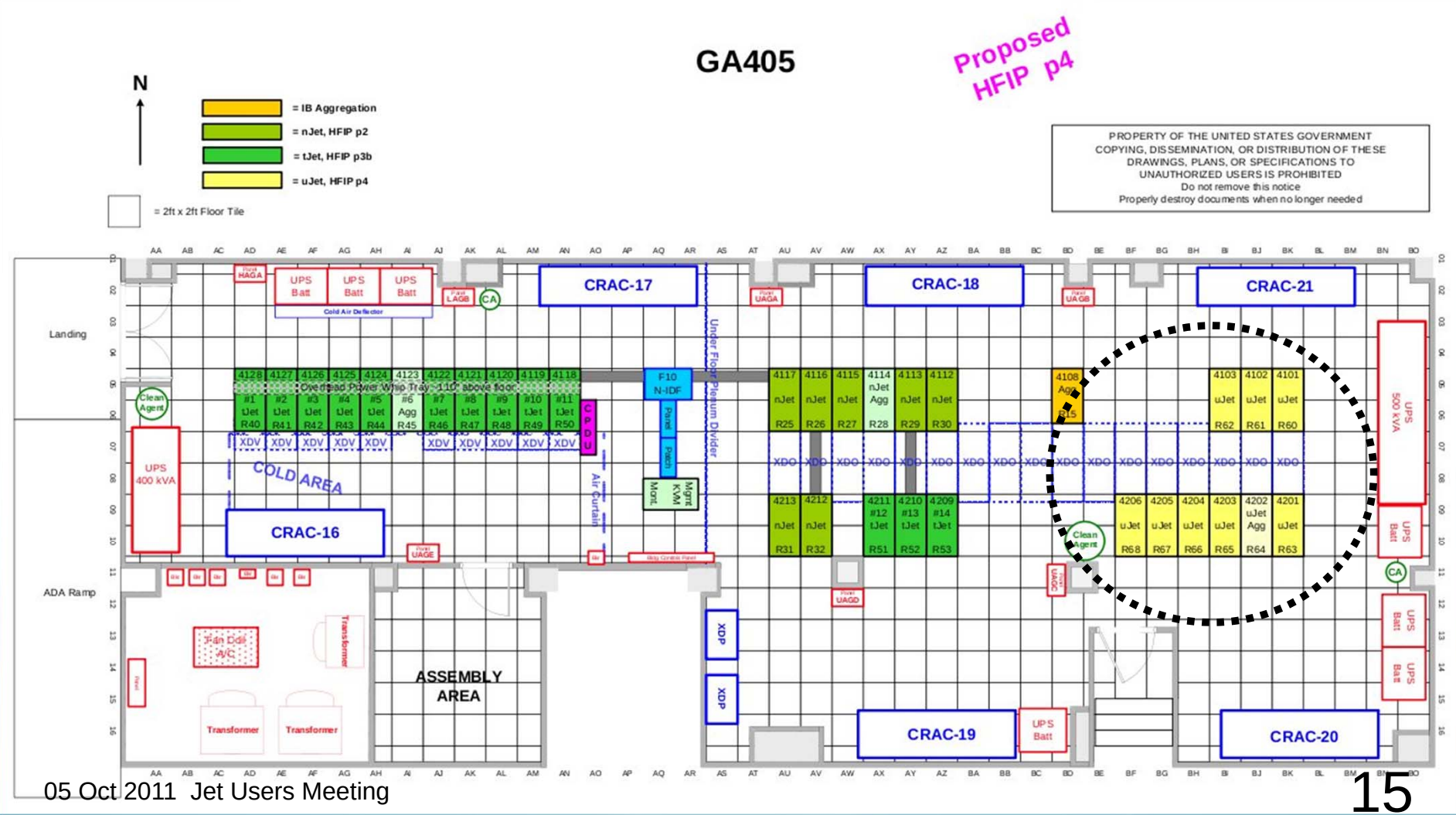
  - Available: on or before 12/09/11

System Resources
Allocations by projects after upgrade

System	RDHPCS	HFIP	FAA	Total
hJet	2016			2016
nJet	1376	472	1736	3584
tJet		10128		10128
uJet		6048		6048
Total	3392	16648	1736	21776

Number of cores available to each project per system.
Total system performance: 235 Tflops

HFIP p4 – Compute Upgrade



HSMS Transition to NESCC (RDHPCS only)

- Oracle SL8500 Tape Library
  - 8 active robots
  - 50 PB (uncompressed) capacity
  - 32 tape drives
- HSMS Software – HPSS
- ~ 1 PB first tier disk

Transition to HSMS at NESCC

- The HSMS data will be migrated to NESCC
- New data will be also written to HSMS
  - We are still working on the plan for this transition. We will have more details in the next month
- The data transition does not affect HFIP data at this time.

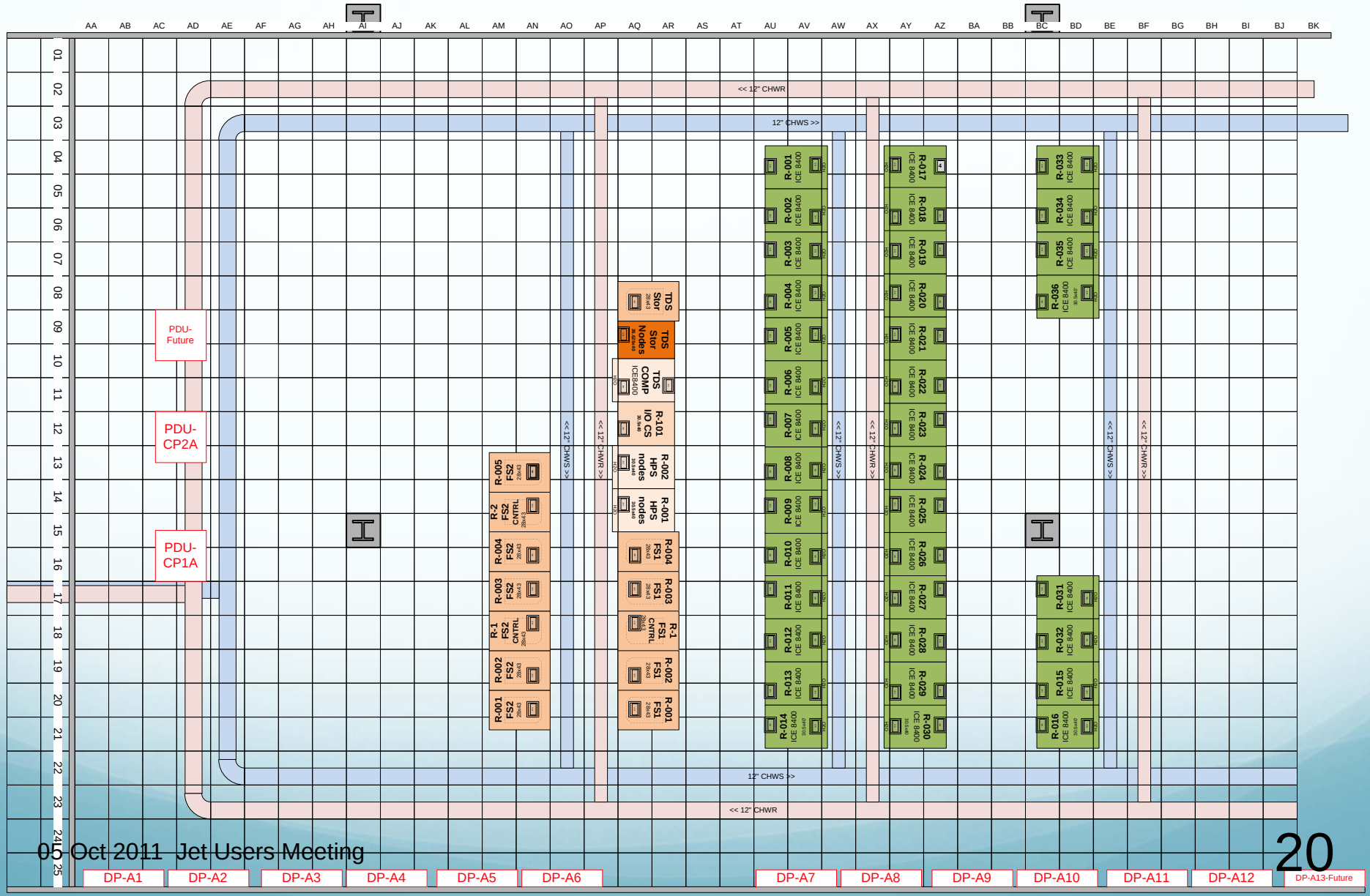
National Environmental Security Computing Center

- Compute Platform – SGI Ice - Zeus
- Tape Library – Oracle SL8500
- HSMS Software – IBM HPSS

Compute System

- SGI ICE
  - 27648 cores, 3.46 Ghz
  - 2GB/core
  - QDR dual-rail Infiniband
- High Performance Storage, Lustre
  - 5.6 PB
  - 50 GB/s
  - File system will be split into two parts

NESCC - Room L121: 1st Floor – HDDS Compute/HPS



Software Environment

- Intel Compilers
  - PGI will be supported
  - Looking into supporting Lahey
- SGI and Intel MPI
- 3rd Party software
  - Everything on Jet should be installed on Zeus
- Modules will be used to select different software packages

Support

- We hope to have the test and development system available to some users in the November/December timeframe
- User support will be similar to current process
- Website is being created, will be similar to current Boulder website
  - http://rdhpcs.noaa.gov/nescc
- Help system will still be email based.
  - help.nescc@noaa.gov
  - We will start with the current package for support but will likely transition to a new package (should be transparent to email-only users)

Schedule

- HSMS
  - Sept 23 – Acceptance Testing Complete
  - Late October/Early November– Start ESRL data migration
- Compute
  - Oct 11: Production and Test System Delivery
  - ~Feb 1: Start Acceptance

Questions?