NESCC, zeus, and the HSMS

About NESCC

The NOAA Environmental Security Computing Center (NESCC), located in Fairmont, WV, is the location of NOAA's newest High Performance Computing Data Center...

There are currently 2 major systems at NESCC:

- Zeus -- A 382 Tflop SGI-based high performance computing system
- HPSS -- A 50 Petabyte IBM/Oracle hierarchical storage management system (HSMS)

from https://nesccdocs.rdhpcs.noaa.gov/

(Tentative) Schedule

- 1 Mar 2012 -- jet /arch2 and /arch3 mount points removed
- 1 Mar 2012 -- transition of targeted projects begins
- 15 Mar 2012 -- full transition begins
- 22 Mar 2012 -- data migration begins
- 30 Sep 2012 -- data migration ends

What 'Transition' Means

- 'Transition' refers to two major components: getting users/projects to use the zeus HSMS, and copying all data from the current HSMS to the zeus HSMS
- Heavy HSMS projects (those who are routinely saving more than a small amount of data) will need to start saving to the zeus HSMS right away!
- While this is going on, RDHPCS staff will copy data in the background
- Light HSMS projects will begin transitioning later

/mss, /arch2, /arch3 Mount Point Removal

- HSMS wrapper scripts will utilize more efficient communications technology (note that they will not be based on ssh/scp (although for a short time authentication may be done through ssh))
- Users should begin to convert their codes to use the wrapper scripts at their earliest convenience (mkdir -> mssMkdir, cp -> mssPut/Get, etc.)
- Other techniques will be made available (e.g., hsi, htar) for power users

General Wrapper Script Philosophy

- The wrapper scripts (mssMkdir, mssPut, etc.) will be updated for the zeus HSMS. We will attempt to make the wrapper scripts behave similarly to how they currently behave, in particular:
  - -u <user> arguments or <user>@ prefixes supported
  - <host>: prefixes supported ('mss2:', 'mss3:', 'zeus:')
  - root paths supported ('/mss', '/mss2', '/mss3', '/arch2', '/arch3')
- Some of these to be phased out eventually

General Wrapper Script Philosophy (cont)

- **Recommended: use the /mss root directory whenever possible**
  - **this was always meant as a user convenience**
  - **i.e., users were never supposed to care about '/arch2' or '/arch3' -- these got out into the wild as output and error messages weren't reverse-mapped**
  - **zeus root directory structure completely different. No one should care (unless you wish to use hsi/htar, say) what this really is -- '/mss' will hide this**
  - **will attempt to reverse map in output and error messages (may not be ready in initial versions).**
  - **likewise for the zeus HSMS host -- 'zeus:' is adequate**

General Wrapper Script Philosphy (cont)

- There will be some differences. In particular:
  - not all option arguments supported
  - exit codes, error messages may differ
  - current mssGet recursive destination directory incorrectly handled; zeus mssGet will be more standard
  - analogues to --ssh-opts arguments not yet supported
- Goal -- at least during the initial transition period, users (who use the wrapper scripts in a vanilla fashion) shouldn't have to modify their scripts. If you need to do more than the wrapper scripts provide, please start a help ticket (rdhpcs.jet.help@noaa.gov)

What about rsync and direct tar?

- No mssRsync currently supported (or envisioned at this point)
  - use mssPut/mssGet with -u and -a (or -p (--preserve=mode,timestamps))
- Poor-man's mssTar written long ago (see Glen for details), but
  - not a direct tar (uses temporary space)
  - not yet tested on zeus

jet Pre-transition Wrapper Scripts and Systems

- If your jet project is in pre-transition mode, all HSMS activity will be to the current ("Old", "legacy") HSMS
- If your project is in transition mode, all writes will go to the zeus ("New") HSMS; all other operations will check both the zeus and the current HSMSes

What 'and' Means

- When we say the zeus HSMS 'and' the current HSMS, we mean just that -- we attempt the operation first on the zeus HSMS, then on the current HSMS
- This may be confusing for mssLs -- listings are not merged
- This may be confusing for mssGet (and mssStage) -- -a is your friend

zeus Wrapper Scripts and Systems

- On zeus, the wrapper scripts will refer solely to the zeus HSMS (mssMkdir, mssPut, etc.).
- No access to the current HSMS (this may change in the future)
- From the zeus production side, the wrapper scripts will access the production HPSS.

More on Wrapper Scripts

- Uses different underlying utilities: uberftp and globus-url-copy. There are many ways for these to fail that my limited testing hasn't yet found. If you've found a new one (or have other issues), contact us at rdhpcs.jet.help@noaa.gov

- Authentication may take a good few seconds -- if you make lots of calls to the wrapper scripts, your processing may be slower

Copying Data

- Behind the scenes, we'll be copying data from the current to the new HSMS
- A number of gotchas:
  - timestamps might not be preserved (we hope to have this fixed, or make a special touch pass someday), and we don't know if all uids/gids will be preserved. (And what about character encodings, etc.?)
  - symbolic links ignored for now
  - mssLs, mssGet confusion, especially mssGet -- not yet smart enough to not copy data twice (current policy: not delete files from the current HSMS after they've been copied). Safer to list/copy files from each HSMS independently

Data Migration Priority

- When data are migrated we will ordering transfers from newest to oldest
- Data retention policy is after 2.5 years we can remove the data.
  https://jetdocs.rdhpcs.noaa.gov/wiki/index.php/HSMS_Data_Retention_Policy
- Goal is to migrate all data, but if we run into issues, data older that 2.5 years may not be migrated.

Every Datum is Sacred

Q: Should I delete old, unwanted data from the current HSMS?

A: Yes and No

Pros: less data to copy!

Cons: deleted data are gone and CANNOT be restored. We cannot fix a user mistake.

So, if you have data you can delete, please delete by 4/1/12 (i.e., please don't delete files when we're in the midst of copying them).

Reminder of HSMS Best Practices

- **Don’t store files that can be recreated:**
  - **complier *.o, *.a files**
  - **run-time *.e, *.o files; run-time links**
  - **sockets, pipes, character/block special files**
  - **browser history/cache files**
- ***PLEASE*** **tar up (small) files**