State of the HPC Cluster

CDAG Meeting, Spetember 9, 2019

Current Cluster Specs

- 32 Compute systems, all running RHEL Linux
- 7 (or so) administrative or login hosts (hpc2, hpc3, hpc5, hpc7, etc.)
- Isilon storage (2 TB /home, 5TB /projects, plus other project FSs)
- Local storage (/usr/local/ from ha3)
- 10G fiber storage networking
- 1Gb Ethernet data network
- Cluster is housed in both RSPH Data Center and LITS Colocation in NDB

Compute Nodes

- Vary in size/speed, from 8 cores (eagle) to 72 cores (alpha, gene9, gene10)
- Total of 664 cores (1328 threads, 1296 job “slots)
- Typically 3-5 Gb RAM per core, again, depending on architecture of node
- All compute systems upgraded to RHEL Linux 7.6 this summer
- New R base installation of 3.6.0 in /usr/local in SP’19.
- New SAS licensing model available this fall

Admin/Login nodes

- 8 current running “administrative” and/or “login” nodes
- All run as VMs in VMWare on LITS VM server
- All running RHEL 6 - need of upgrade (RH Product retirement 11/2020)
- Several support issues
  - LITS personnel departure; primary emphasis on AWS
  - Configurability
- We have purchased 4 new replacement servers, on-prem hardware
  - Two for replacement Admin node and Login node, running up-to-date software
  - Two for testing (patching, new versions, SLURM)

Storage

- Current “hot topic”: immediate need for large scale, inexpensive storage
- Combined researcher requests for > 500TB in BIOS alone in near term
- LITS Isilon storage:
  - “We’re happy to support you, but we don’t support your model” for local computation - design is primarily for archival and compliant storage.
  - Likely won’t be able to keep up with growing demand
  - Price is currently $560TB/yr
- AWS @ Emory:
    - Storage beyond S3 not approved for local use (e.g., Snowball)
    - Direct connection to RSPH HPC Cluster not currently available.

Storage (cont’d.)

Local storage option appears to be best:

- It is available now
- Estimated costs, depending on options, order-of-magnitude cheaper
- Configurable to our needs and design
- Potentially scalable to Petabytes
- This will require policies (CDAG):
  - Curation of data and archiving by researchers
  - Need a known ongoing funding model
  - Understanding that there would no longer be a 24/7/365 storage team running storage

Storage

Open questions (for CDAG):

- How do we acquire storage? As a whole cluster? By department? By research group? By aggregating current needs?
- How do we retire storage? When funding runs out and disks fail?
- What level of performance is required? (Assumption that we will **NOT** implement a parallel file system, for instance.)

Future Discussions

- Compute nodes:
  - New acquisitions - a current compute node specification exists.
  - Some purchases are imminent.
  - Retirement of old nodes, especially after warranty expires. Re-use for testing.
- Whither GridEngine? SLURM?
- Networking options (10, 40, 100G)?
- And others...