RDHPCS Computing for HFIP and HAFS

Vijay Tallapragada
Chief, Modeling and Data Assimilation Branch
NOAA/NWS/NCEP/EMC

HFIP Annual Meeting
Miami, FL, Nov 4-6, 2019
System Status for WCOSS

Reston: IBM System – Tide / Cray System – Luna

Orlando: IBM System – Gyre / Cray System - Surge

The Cray Systems will include graphics on the front panels of the systems as shown in these two images.
System Status for WCOSS

Current WCOSS:
- IBM P1/IBM P2 (Tide/Gyre): 1098 nodes (24 cores per node)
- Cray (Luna/Surge): 2048 nodes (24 cores per node)
- Dell P3 (Mars/Venus): 1200 nodes (28 cores per node)

P1/P2 Retirement
- Old IBM P1/P2 transitioning to new Dell
- Current moratorium to end on January 22, 2020

New Addition for WCOSS in 2020
- Dell Phase 3.5: 625 nodes (40 cores per node)

Next WCOSS: TBD
- All systems to migrate to new WCOSS during Feb 2021 – Jan 2022
- Moratorium kicks in on Feb 1, 2021 (No model upgrades for 2021)
System Utilization: Phase 1 for CY2019

CFS

GWENS

Prod

Prod

System retired from production in January 2020
System Utilization: Phase 2 for CY2019

System retired from production in January 2020
System Utilization: Cray for CY2019

- HWRF
- NWM
- GFS
- FV3GFS
- Prod
- HRRR
- HWRF

Missing Data
RDHPCS Program Wide

- Hera is the replacement for Theia
  - being turned on incrementally with same amount of Theia turned off – complete by November 25, 2019
- MSU (Orion) is the new HSUP computer
  - Early users are onboarding
  - Facility, Interconnect challenges
  - Mass port of users planned from Theia/Hera
- Gaea – mostly used by GFDL
- Jet is the primary HFIP machine
# Theia Allocation with HS for Hera and Orion

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Hera From Theia</th>
<th>Hera HS</th>
<th>Hera Total</th>
<th>Hera Only Req</th>
<th>Orion HS</th>
<th>Total</th>
<th>Percentage of Allocation</th>
<th>Percentage of Orion Allocation Alloc</th>
<th>Portfolio</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCEP</td>
<td>7,488,321</td>
<td>6,583,168</td>
<td>14,071,489</td>
<td>1,890,000</td>
<td>17,610,193</td>
<td>31,681,682</td>
<td>37.0%</td>
<td>33.50% NCEP</td>
<td></td>
</tr>
<tr>
<td>BMC</td>
<td>8,976,699</td>
<td>4,699,470</td>
<td>13,676,169</td>
<td>3,330,000</td>
<td>12,571,237</td>
<td>26,247,405</td>
<td>30.6%</td>
<td>23.92% BMC</td>
<td></td>
</tr>
<tr>
<td>GFDL</td>
<td>1,705,337</td>
<td>1,129,785</td>
<td>2,835,122</td>
<td>0</td>
<td>3,022,211</td>
<td>5,857,333</td>
<td>6.8%</td>
<td>5.75% GFDL</td>
<td></td>
</tr>
<tr>
<td>NAGAPE</td>
<td>764,164</td>
<td>4,182,376</td>
<td>4,946,540</td>
<td>0</td>
<td>11,187,994</td>
<td>16,134,534</td>
<td>18.8%</td>
<td>21.29% NAGAPE</td>
<td></td>
</tr>
<tr>
<td>NESDIS</td>
<td>381,070</td>
<td>1,086,331</td>
<td>1,467,401</td>
<td>0</td>
<td>2,905,972</td>
<td>4,373,374</td>
<td>5.1%</td>
<td>5.53% NESDIS</td>
<td></td>
</tr>
<tr>
<td>NOS</td>
<td>0</td>
<td>2,390</td>
<td>2,390</td>
<td>11,000</td>
<td>6,393</td>
<td>8,783</td>
<td>0.0%</td>
<td>0.01% NOS</td>
<td></td>
</tr>
<tr>
<td>NMFS</td>
<td>10,697</td>
<td>0</td>
<td>10,697</td>
<td>0</td>
<td>0</td>
<td>10,697</td>
<td>0.0%</td>
<td>0.00% NMFS</td>
<td></td>
</tr>
<tr>
<td>INCUBATOR</td>
<td>80,230</td>
<td>0</td>
<td>80,230</td>
<td>0</td>
<td>0</td>
<td>80,230</td>
<td>0.1%</td>
<td>0.00% INCUBATOR</td>
<td></td>
</tr>
<tr>
<td>Coastal</td>
<td>1,337,162</td>
<td>0</td>
<td>1,337,162</td>
<td>0</td>
<td>0</td>
<td>1,337,162</td>
<td>1.6%</td>
<td>0.00% Coastal</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>20,743,680</strong></td>
<td><strong>17,683,520</strong></td>
<td><strong>38,427,200</strong></td>
<td><strong>5,231,000</strong></td>
<td><strong>47,304,000</strong></td>
<td><strong>85,731,200</strong></td>
<td><strong>100.00%</strong></td>
<td><strong>100.00%</strong></td>
<td></td>
</tr>
</tbody>
</table>
HFIP 2019 Reservation

Inputs provided by Wei Yu
Summary

• Workflow review
  • The user workflow review started on June 21
  • 6 projects selected. Total 10,753,000 Cr-Hrs/month
  • 2,079,000 Cr-Hrs/month on tjet, and the rest are on xjet
  • There are five CM tickets for reservation changes related to user’s workflow adjustment after July 11

• Reservations
  • Pre-production Reservation Deployment on July 11
  • There are 81 reservations for HFIP project
  • All reservations are monitored every 6 hours (missing or not)
  • Reservation usages have been monitored for each epoch and reviewed every week

• Problems
  • There are 31 help tickets for system and reservation issues.
Release Reservation

- Slurm transition issues
  - No existed tools to monitor reservation usage
  - No existed tools to release and create the daily reservation
  - New developed tool may have potential bugs

- Developed tool during this season
  - Monitor reservation usage
  - Basic reservation release script
  - Automatic release unused reservation
    - For unused reservations only (tested and worked fine)

- Other issues
  - Lack of the tool to release the reservation when the job is finished.
Reservation with Auto Release

The difference between red and green colors is related to the reservations which are not released after the job finished. This will be improved next season.
Summary of Issues

• Data transfer
  • NCEP data delayed some time (a random issue)
  • In one case, we saw that other users’ job piling up to use front node and caused HFIP data delaying. After we made the reservation for HFIP data transferring on front node, the data delaying issue was not reported any more.

• Slurm Bugs
  • Failed to load jobs into reservation nodes
    • Solution: slurm provided patches and restarted the reservations after every downtime
  • Failed to replace the bad nodes when reservation started/to be started
    • Temporary solution: added one more node to each reservation

• Hardware issues
  • It is normal in distributed computer system that the bad nodes appeared from time to time. The batch job system will automatically avoid to use those bad nodes. It caused more attention this year, since it was mixed together with slurm bugs.

• Current Status
  • During last three weeks, all reservations worked fine. It looks like that we got the right solutions.
Lessons from 2019 Demo

- **Offset** - t is the time to start the workflow. Need to consider delay of the data.
- **Duration** - It is the job running time. Suggested to be 10 minutes longer than the benchmark of the running job; could be impacted by disk and network.
- **Nodes and Cores** - If the job is going to take whole node, request 1 node, even when using one core. If the job is going to run on shared node with one or more cores, need to state it.
- **Service nodes** - If the job is going to use front node, need to state and not count in computing nodes.
- **Front End nodes** - Make reservation on FE nodes for real time data transfer. Other user’s behavior could mess up the real time job.
- **Compute Nodes** – Request extra node for each reservation to prevent job failures due to occasional bad nodes.
Questions?

Thank You