Building on Open OnDemand at Texas A&M
Richard Lawrence

ECSS Symposium
March 15, 2022

High Performance Research Computing
DIVISION OF RESEARCH
Outline

1. Introduction
2. Open OnDemand Framework
3. Quality of Life Improvements at Texas A&M
4. Vision for the Future
5. Conclusions
1. Introduction
State of the Field

- Cyber Infrastructure (CI) is challenging
- Researchers by and large prefer interactive, graphical interfaces
  - IDE (Integrated Development Environment)
- Science requires increasingly sophisticated software practices
  - FAIR (Findability, Accessibility, Interoperability, and Reusability)
  - FEAT (Fairness, Ethics, Accountability and Transparency).
A popular Python IDE

The Jupyter Notebook

```python
In [ ]:
1. import tensorflow as tf
2. print("Num GPUs Available: ", len(tf.config.list_physical_devices('GPU'))
```
Containers

- Containers make software portable
- Popular container runtimes
  - Docker, Singularity, Charliecloud
- Many scientific packages have already been containerized
  - Bioconda, LAMMPS, and Tensorflow

Containers seem mysterious, but they’re actually pretty easy
Typical Research Workflow

Can we support all of that?

- **Learning**: 😊📖
- **Development**: 😞✍
- **Analysis**: 😞📈
- **Production**: 😎☕
2. Open OnDemand
Web Interface for HPC

openondemand.org/

open source community
HPRC Portal

HPRC Portal is an instance of OOD portal.hprc.tamu.edu

OnDemand provides an integrated, single access point for all of your HPC resources.
HPRC Portal

HPRC Portal is an instance of OOD portal.hprc.tamu.edu

Spotlight: Interactive Apps menu
Job Submit (default)
Sometimes you want to teach a class. Here's how we use the View-only link to provide support for remote participants.
Jupyter Notebook
(default)

Jupyter Notebook version: v1.0.1-3-g94d29b4

This app will launch a Jupyter Notebook server on one or more nodes.

Account

Queue

Number of hours

1

Number of nodes

1

I would like to receive an email when the session starts

Launch

* The Jupyter Notebook session data for this session can be accessed under the data root directory.
3. Quality of Life
Improvements in Texas
A&M's OOD Portal
The dashboard shows information about the cluster, user accounts, and disk quotas, that would be difficult to obtain on the command line.
Job Composer

This job composer helps the researcher write their batch files.
<table>
<thead>
<tr>
<th>Current Scratch Quota</th>
<th>1 TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Scratch Quota</td>
<td>TB</td>
</tr>
<tr>
<td>Current File Limit</td>
<td>250000</td>
</tr>
<tr>
<td>New File Limit</td>
<td></td>
</tr>
<tr>
<td>Justification (Required)</td>
<td></td>
</tr>
<tr>
<td>What data is stored with requested quota?</td>
<td></td>
</tr>
<tr>
<td>What job requires this quota increase?</td>
<td></td>
</tr>
<tr>
<td>What is the input/output size of the job?</td>
<td></td>
</tr>
<tr>
<td>What is your long-term plan for this data?</td>
<td></td>
</tr>
</tbody>
</table>

**Error Message**

- e.g. file `my_script.py` cannot be found

**Path to Your Job File**

- e.g. `$SCRATCH/myawesome_research/my_sim_job.sh`  

**Job ID**

**Path to Your Executable/Script file**

- e.g. `$SCRATCH/myawesome_research/hello_world.out`  

**Additional Information**

- e.g. The data is located at `$SCRATCH/my_data`. Also, when I run the job, I use my Python virtual environment located in `$SCRATCH/venv`  

**Comment (Optional)**

**I verify that I will remove any unnecessary data and compress files/folders to save shared resources.**

**Submit Request**

Either Disk Quota or File Limit must be filled.

**Create Help Ticket**

- What do you need help with?  
  - Other: Software Installation

**Please be as specific as possible. This will help us resolve your issue faster.**
Rich Jupyter Environments Menu

**Jupyter Notebook (TESTING)** version: 3205f76

This app will launch a Jupyter Notebook server on the Grace cluster.

**Notice:** This form has changed. Please pay attention to what options you select and what the defaults are.

**Type of environment**

- Module load + Python virtualenv

Select the type of environment in which Jupyter is installed.

Help me choose

**Module selected**

SciPy_tamu/2021.10-Python-3.8.6

Select a module to load. All modules listed will also load the
In the back-end:

\texttt{jupyter notebook <args>}

\textbf{Becomes:}
\begin{verbatim}
<%singularity-statement=
singularity <path-to-image-file>%%>
<%=singularity-statement%%> jupyter notebook <args>
\end{verbatim}
Case study: Teaching Containers

Course Agenda
● Overview of Containers
● Singularity on HPRC
● Getting a Container Image
● Container Usage Basics
● More Container Examples
  ○ Tensorflow with GPU
    OnDemand Interactive showcased
  ○ LAMMPS with MPI
  ○ Biocontainers
● Build your own Containers
Feedback overwhelmingly positive
Case study: Teaching Containers

Course Agenda

- Overview of Containers
- Singularity on HPRC
- Getting a Container Image
- Container Usage Basics
- More Container Examples
  - Tensorflow with GPU
  - OnDemand Interactive showcased
  - LAMMPS with MPI
  - Biocontainers
- Build your own Containers

Feedback overwhelmingly positive
Case study: Teaching Containers

Course Agenda
- Overview of Containers
- Singularity on HPRC
- Getting a Container Image
- Container Usage Basics
- More Container Examples
  - Tensorflow with GPU
    OnDemand Interactive showcased
  - LAMMPS with MPI
  - Biocontainers
- Build your own Containers

Feedback overwhelmingly positive

Attendance/Registration
- 4/11 (2021-04-30)
- 2/20 (2021-10-15)
- 7/13 (2022-03-11)
4. Vision for the Future
Unify Interactive and Batch Job Composers

Disjoint forms with redundant elements
*Unsustainable*

- Job composer
- Quick job
- Job template
- Batch jobs
- Interactive app form
- Interactive job

Common form elements centralized
*Sustainable*

- Job composer
- Quick job
- Job template
- Batch jobs
- Interactive job
Graphical Environment Manager
for Python and other package-based environments

- “Create a virtual environment” is a task we want researchers to be utilizing, but the command line is clunky
- Personal computers already manage their Python packages graphically, so researchers on HPC can too.
Graphical Environment Manager
for Containers

- “Pull or build a container image” is a task we want researchers to be utilizing, but the command line is clunky
- Resource-intense container tasks push towards Batch scheduling and accounting
- Menus can help researchers with unfamiliar container runtime syntax
Classroom Features for Instructors

- HPC skills can be taught via the Texas A&M portal in a variety of ways - apps, Jupyter notebook, CLI
- Distribute course materials; Student-teacher interaction, feedback, grading
- Course enrollment dynamically linked to batch scheduling
5. Conclusions
Researchers Want This

- Education and Research have a natural synergy if the interface is good.
- Graphical interfaces for tools such as Python or Containers lower the barrier to entry for new users.
- Improves the quality of research by enabling new technologies.
- Important to provide user-friendly interfaces for the technologies to foster researcher adoption.
Acknowledgements

This work was supported by the

- National Science Foundation (NSF) award number 1925764, "CC* Cyberteam: SWEETER"
- NSF award number 2019129, "MRI: FASTER"
- NSF award number 1730695, "CyberTraining: CIP: CiSE-ProS: Cyberinfrastructure Security Education for Professionals and Students"
- NSF award number 2019136, "CC* BRICCs: Building Research Innovation at Community Colleges"
- NSF award number 1829799, "Cybertraining: CMS³"
- NSF award number 2112356, "ACES - Accelerating Computing for Emerging Sciences"