OPEN ONDEMAND 101

Ryan DeRue, Senior Computational Scientist
What to Expect from Open OnDemand 101?

Objectives

- Understand how science gateways like Open OnDemand address the challenges faced in typical high-performance computing (HPC) work.
- Ensure you know how to login to the gateway associated with your HPC resource here at Purdue.
- Become comfortable with the purpose and location of different features in Open OnDemand.
- Feel confident running your own workflows in Open OnDemand.
OOD 101

Lowering the Barrier to HPC
Layout of a Cluster

- HPC systems are often organized into clusters
  - We call them clusters because they are a multitude of independent nodes that can communicate with one another
  - Because each node is independent, we can have different types of nodes with specialized hardware

- When you login to a cluster, typically you will land on login node along with all the other users on the cluster
  - In order to carve out dedicated resources for your work, you need to ask the scheduler to allocate you resources on the compute nodes
HPC Presents a Steep Learning Curve

- Working within HPC in this traditional manner presents many challenges to new users
  - Typically use an SSH client/Linux terminal
  - Learning new commands for interacting with the scheduler
  - Software runs without a graphical user interface (GUI)

- This can lead to new domain scientists spending most of their time learning how to use the system as opposed to doing science
Science Gateways

- To make these resources easier to use, there has been a push for the creation of “science gateways”

- Science Gateways are typically web based interfaces that allow users to run science/engineering applications on HPC resources from their web browser

- Open OnDemand is an example of such a science gateway
OOD 101

Using Open OnDemand: Dashboard
Logging into Open OnDemand

- We offer Open OnDemand on all Purdue community clusters, however each cluster has its own access gateway
  - Each of these can be found by navigating to: gateway.<cluster>rcac.purdue.edu
  - For example, for the Gilbreth cluster,
- Upon navigating to the gateway for your cluster, you will need to login using BoilerKey
- Once you are logged in, you will land on your dashboard page
  - My dashboard for gateway.gilbreth.rcac.purdue.edu is on the right
Getting Familiar with The Dashboard

1. Current Quota Usage
   - Shows the remaining space in each of your different filesystem quotas
     - Includes your file count quota as well!
   - To learn more about storage quotas see our user guide page on quotas
Getting Familiar with The Dashboard

2. Directory Shortcuts
   - Allows you to quickly navigate between your important directories
   - Clicking these will open a file navigator allowing you to view/edit your files or upload/download files to/from the cluster
Getting Familiar with The Dashboard

3. Jobs Menu
   • Provides shortcuts to widgets for creating new Slurm jobs as well as viewing the status of your current ones
Using Open OnDemand: Dashboard

Getting Familiar with The Dashboard

4. Interactive Application Menu
   - This dropdown lists all the applications we support on Open OnDemand
   - Clicking one of the applications will take you to a form through which you can request the resources you want to use
Getting Familiar with The Dashboard

5. Documentation

- Provides a shortcut to our user guide for the cluster whose Open OnDemand instance you are using.
Getting Familiar with The Dashboard

6. Cluster

• Provides options for interacting with the cluster directly

• Gilbreth Shell Access will provide you with a Unix shell on one of the login nodes

• Open XDMoD provides a tool for viewing metrics about the cluster
Getting Familiar with The Dashboard

7. My Interactive Sessions
   - This will take you to a page wherein you can access all your currently running application sessions
OOD 101

Using Open OnDemand: Files
A Web-Based File Explorer

1. To access the File Explorer, click on one of the mounted locations listed in the Files dropdown
   • This will open a new tab containing the File Explorer
   • You can also click the blue hyperlink text within the quota widget
Using Open OnDemand: Files

A Web-Based File Explorer

2. The menu along the left-hand side of the screen is for navigating to the different subdirectories contained within your home directory

- Even if you change the directory you are currently examining, this menu will continue to display subdirectories relative to your home directory
A Web-Based File Explorer

3. The central part of this webpage displays the contents of your present working directory as well along with some metadata like the last modified timestamps.
A Web-Based File Explorer

4. The toolbar located above the contents of the directory displays your current working directory as well as tools for affecting the files contained here.
Using Open OnDemand: Files

A Web-Based File Explorer

5. The toolbar in the upper right corner provides tools for changing your current directory, opening your current directory in a terminal to apply shell commands to the files, and creating new files/directories that are either empty or uploaded from your local workstation.
OOD 101

Using Open OnDemand: Applications
Using Open OnDemand: Applications

**Application Forms**

1. To access the App Form for Gilbreth Compute Desktop, select it from the Interactive Apps dropdown on the Dashboard
   - The form will be opened in your current tab
   - You can return to the Dashboard by clicking on “Gilbreth–Gateway” in the top left corner
Application Forms

2. You can navigate to other Interactive Apps by selecting them from the left-hand menu.
Application Forms

3. In order to request access to dedicated resources, you must select the queue to which you wish to submit your job

   - Different queues serve different purposes and information about each can be found in our user guide
Application Forms

4. You must also specify the number/type of resources you wish to have allocated and for how long you intend to use them

- You will be given memory proportional to the number of CPUs/GPUs requested. For GPUs, typically 1 GPU is sufficient
- It's important to remember that the scheduler prioritizes more modest requests. You should not ask for the maximum duration walltime just because it's possible
Application Forms

4. Because there can be a wait time dependent on the queue selected as well as the size/duration of the allocation requested, you can check this box to receive an e-mail when your session is ready.
Application Forms

5. After selecting the options you would like for your session, click launch to send your request to the scheduler
Application Forms

6. This will create a "card" for this session which will assign it two IDs

   • The first ID is the Slurms scheduler's job ID and is parenthesized next to the application name
   • The second ID is the Session ID and is the name of the directory where Open OnDemand will track information about the session
   • Clicking this session ID will open this directory in the File Explorer
   • When the scheduler finds resources for your job, the card will transform into the second image, and you can access the application by clicking the blue button labelled "Launch..."
Specific Applications: Compute Desktop

Every Purdue RCAC cluster has an application named “Compute Desktop.”

- This application will launch a lightweight graphical desktop that you can use to interact with the HPC cluster as you would your own local workstation
- It includes a native file explorer, Firefox as an internet browser, a terminal, and a typical Desktop that you can save files to
- This application can be used to launch software that requires a GUI
  - This can be used on system-wide installations by RCAC as well as with software you personally install
Specific Applications: Compute Desktop

1. When you first launch the Compute Desktop, you will be presented with a desktop that appears as the one on the right.
   
   - You can access your files by selecting the folder icon in the taskbar along the bottom.
Specific Applications: Compute Desktop

1. When you first launch the Compute Desktop, you will be presented with a desktop that appears as the one on the right.
   - You can access your files by selecting the file-cabinet icon in the taskbar along the bottom.
Using Open OnDemand: Applications

Specific Applications: Compute Desktop

2. From this file menu, you can access any file on the cluster, but by default it will open to your home directory

- You can change the directory you are viewing by clicking on other directories or by typing the absolute path into the search bar at the top
Specific Applications: Compute Desktop

2. You can launch a terminal on the desktop by selecting the terminal icon from the bottom taskbar
   
   • This is extremely useful for launching GUI applications from the command-line
Using Open OnDemand: Applications

Specific Applications: Compute Desktop

3. Notice that the prompt says that host is gilbreth-f002. This is the name of the compute node that we are on
   • If we were to try to do the same thing through ThinLinc, which provides a similar interface, the terminal would indicate that we are on a front-end (not a compute node) with a label of gilbreth-feXX
Specific Applications: Compute Desktop

4. By loading a module (ansys) for software that I know provides a GUI, I can launch that GUI by typing the name of the command to do so.

   • Usually, you can find such commands in the user manual for software you personally install
   • This allows me to run scientific software the same way many users do on ThinLinc, but I get all the resources dedicated to my application which improves performance!
Specific Applications: Jupyter

1. Jupyter is by far the most popular application for Open OnDemand on our clusters
   • Very useful application for interactive work with languages like Python, R, Julia, etc.
   • App form is very similar to the previous one, but it does have an option to select whether you want to run “Jupyter Lab” or “Jupyter Notebook”.
   • Notebook is the original Jupyter application and is a more bare-bones method of using Jupyter.
Specific Applications: Jupyter

2. We have installed several system wide Jupyter kernels for various languages.
   - Some of these kernels, the “learning kernels” contain ready to use installations of the most common machine learning libraries.
   - You cannot install additional packages into these kernels, but you can create your own anaconda environment containing the packages you want.
     - If your anaconda environment contains the packages “IPyKernel” and “IPython”, you will be able to select them as a Jupyter kernel when using Open OnDemand.

<table>
<thead>
<tr>
<th>Jupyter</th>
<th>This app will launch a Jupyter Notebook or Jupyter Lab session on the Gilbreth cluster.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Queue</strong></td>
<td>standby (Max 4.0 hours)</td>
</tr>
<tr>
<td>Please select a queue from the drop-down and enter the number of hours below (up to the max listed above).</td>
<td></td>
</tr>
<tr>
<td><strong>Number of hours</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Number of GPUs</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>Type of GPU</strong></td>
<td>First Available</td>
</tr>
<tr>
<td>I would like to receive an email when the session starts</td>
<td></td>
</tr>
<tr>
<td><strong>Launch Jupyter Lab or Jupyter Notebook</strong></td>
<td>Jupyter Lab</td>
</tr>
<tr>
<td>In order to use your existing Anaconda environments as Jupyter kernels within Open OnDemand, they must contain IPyKernel and IPython</td>
<td></td>
</tr>
</tbody>
</table>
Specific Applications: Jupyter

3. Regardless of whether you choose to use Jupyter Notebook or JupyterLab, it will open to your home directory.

   • To create a new notebook in Jupyter Lab click the blue “+” button in the top left corner (Shown in the top-most picture)

   • To create a new notebook in Jupyter Notebook, select the “New” button in the top right corner. (Shown in the bottom-most picture)
Using Open OnDemand: Applications

Specific Applications: Jupyter

4. After we have created a new notebook, we will be presented with our new notebook in an interface shown to the right

   • Jupyter Lab is shown in top-most picture and Jupyter Notebook shown below that

   • To select the kernel you wish to use, click on the name of the kernel shown to the right in a yellow box, and an option to select a new kernel will appear
OOD 101

Using Open OnDemand: Job Composer
A Widget For Talking to the Scheduler

Often, we may also want to run non-interactive jobs. In such cases Open OnDemand provides an interface for creating such jobs as well as monitoring your interactive and non-interactive jobs.

We can access these tools using the Jobs menu
A Widget For Talking to the Scheduler

1. The Active Jobs widget will allow you to query certain usage statistics from the scheduler that you would otherwise need to do through the command line interface (CLI)
   - When a job fails due to an error, this is a useful tool to discover what may have happened
   - You can look at the amount of memory that was requested to see if perhaps there was an out of memory error
   - You can also use the “Open in File Manager” button to view the output.log file that may contain information about the error that occurred
A Widget For Talking to the Scheduler

2. The Job Composer widget is used for submitting non-interactive jobs to the scheduler. This is very useful when you have some work to accomplish as part of your workflow that doesn’t need to be done interactively.

- An example of this might be post processing the output of some work that was generated interactively
- This tool provides a simple default template for submitting a “Hello World” job script, but allows you to save template job scripts for workflows you may regularly repeat
A Widget For Talking to the Scheduler

3. Selecting the “+ New Job” button will allow you to begin working on creating a new job either from one of your existing templates, or from the default hello world template.

• This will create a new entry in your table of jobs with a status of “Not Submitted”
A Widget For Talking to the Scheduler

3. If we select the newly created job in the table, we can then click on the “Job Options” button to bring up configurable options for that job.
   
   • The most important fields here are the job name, which is for your own reference, and the “Account”
   • “Account” is another name for queue. You can enter the name of one of the queues which you are able to select from the queue dropdown on any of the app forms
A Widget For Talking to the Scheduler

4. After saving the job options, we can scroll along the boxes on the right side of the table to do two things

   • We can see some details about the job
   • We can edit the job script that this job will run
     • This is the most important part of this workflow because this is defining what we want our non-interactive job to do
   • Select “Open Editor” to change the contents of this script
A Widget For Talking to the Scheduler

5. Within your job script, you will want to put all the commands which you want to be ran as part of the job, as well as any options to Slurm, and then hit “Save”

- For more information about how to do this, see the user guide page for example Slurm job scripts
A Widget For Talking to the Scheduler

6. After we have made our edits to the job script, we can close this tab and return to the job composer page. Now all that remains is to select this job and hit submit!

- After this point the status of the job updates from "Not Submitted" to "Running", and when the job terminates, this will update finally to "Completed"
A Widget For Talking to the Scheduler

7. After we have made our edits to the job script, we can close this tab and return to the job composer page. Now all that remains is to select this job and hit submit!

• After this point the status of the job updates from “Not Submitted” to “Running”, and when the job terminates, this will update finally to “Completed”
A Widget For Talking to the Scheduler

8. After the job completes, we can view the job details along the right-hand side of the page

   • We see that the “Folder Contents” has been updated with a new output file: “slurm-2677561.out”

   • If we click the link to this file, we will see that our job successfully generated the expected output:

   ```
   1  Going to sleep for 30 seconds, but when I woke up, I’ll greet you!
   2  Hello from Open OnDemand!
   3
   ```
Conclusions
Conclusions

- Open OnDemand helps to lower the barrier of entry to HPC for domain scientists seeking to do computational work

- We can use Open OnDemand to manage our directories as well as transfer files to/from the cluster

- Open OnDemand is especially useful for running interactive style jobs that require a graphical user interface

- We can use Open OnDemand to manage our non-interactive jobs as well
OOD 101

Questions?
THANK YOU

Feel free to reach out to rderue@purdue.edu with questions.

Slides are posted at: https://www.rcac.purdue.edu/training/ood101