Open OnDemand

HPC for everyone

Robert Settlage, Alan Chalker, Eric Franz, Steve Gallo, Edgar Moore, David Hudak

June 2019

Ohio Supercomputer Center

University at Buffalo
Center for Computational Research

VIRGINIA TECH

Goals and Objectives

Showcase Open OnDemand for HPC

OPEN OnDemand

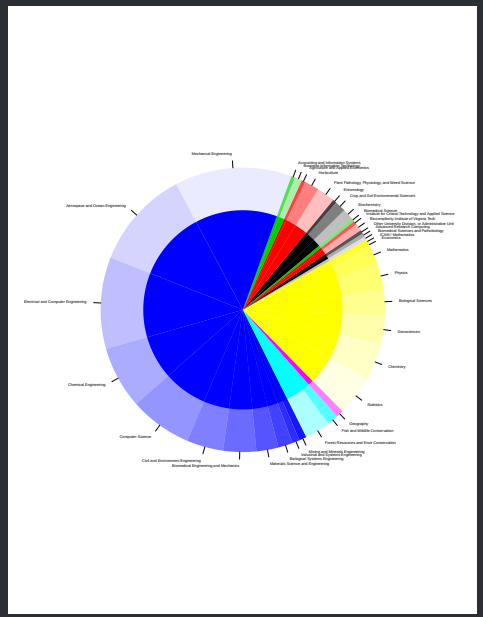
- Introduce ARC at VT
- Discuss HPC barriers
- Introduce OOD
  - features
  - adoption
  - successes
  - roadmap

Advanced Research Computing
Virginia Tech

Unit within the Office of the Vice President of Information Technology.

Goal: Further research by lowering barriers to the use of HPC and Viz

- Centralize resource acquisition, maintenance, and support for research community
- Provide support to facilitate usage of resources and minimize barriers to use
- Enable and participate in research collaborations between departments



Advanced Research Computing

Resources

Heterogeneous clusters supporting many different compute profiles.

934 x86 + 14 Power8 + misc. 7.5 PB BeeGFS, 3 PB GPFS, 275 TB Qumulo

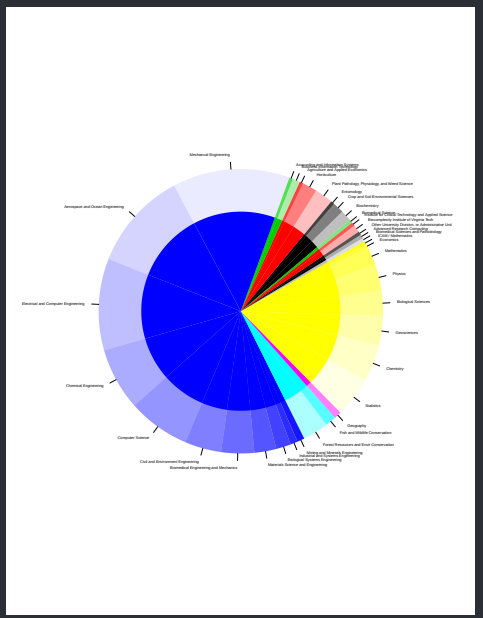
- Ca. 1000 compute nodes split by acquisition generation in 5 clusters
- General X86 compute, x86 + GPU (V100, P100, K80), large mem (3 TB), big data (3 TB local disk + 768 GB RAM), PowerAI (Power8 + 4 P100)
- Visualization resources including 10’ 3D cube, high res wall, more

High Performance Computing
Barriers

Availability of hardware is not an (immediate) issue.

Access and use barriers are largely self-imposed.

- System access: ssh
- Software: no root access, modules
- Data (in/out): ftp, scp, rsync, etc
- Compute configuration, script writing: vi, emacs, etc
- Compute execution: job scheduling



Open OnDemand

Features | Overview

Open, Interactive HPC Via the Web.

Provides easy to use and extend, web-based access to HPC.

Features:

- Plugin-free web experience
- Easy file management
- Command-line shell environment
- Job Management and monitoring
- Graphical desktop environments and applications

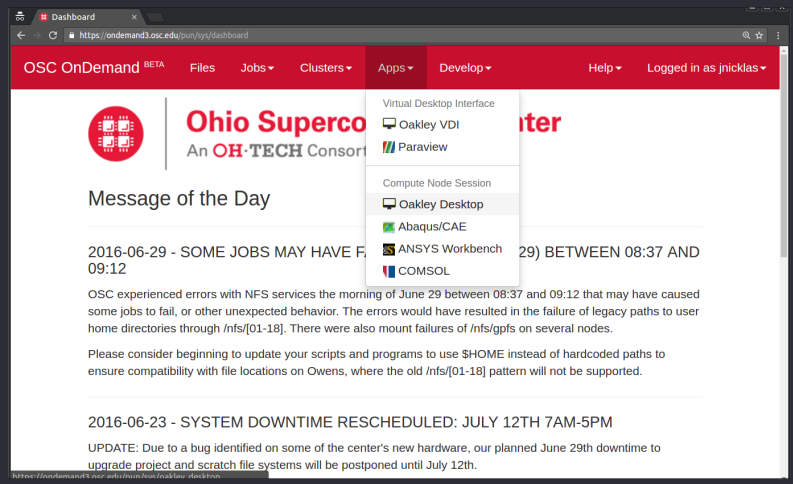
Open OnDemand

Features | Out of the Box

Users come with a modern web browser and HPC credentials.

Open OnDemand provides zero-install and single sign-on solution.

- Landing page
- Files App
- Job Composer App
- Job Monitor

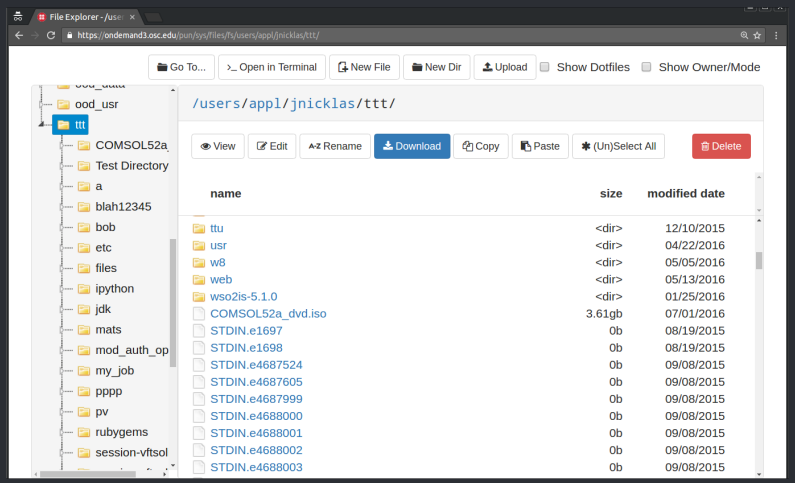


Open OnDemand

Files App

Command line file management is a formidable barrier.

Open OnDemand gives users a familiar tree based file management tool.



- Tree view
- Drag/Drop transfers
- Web viewer
- Web editor

Combined, this reduces inadvertent file errors.

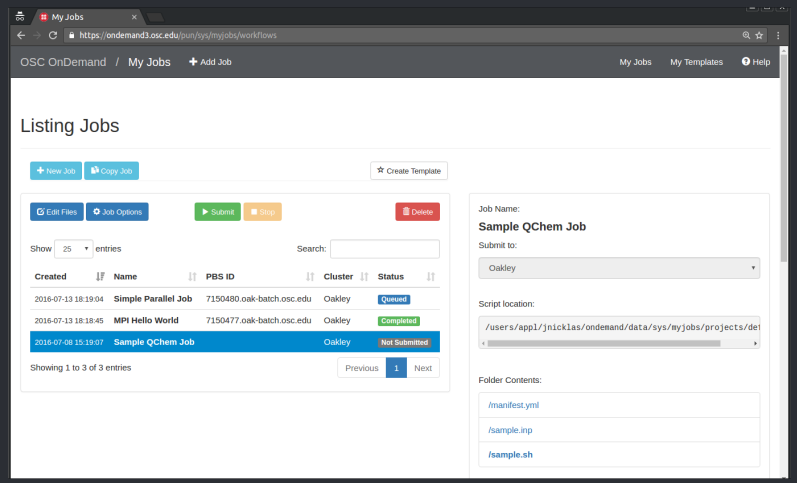
Open OnDemand

Job Composer App

Interaction with schedulers can be both confusing and daunting.

Open OnDemand makes editing and submitting jobs visual.

- Common job workflow:
  - copy previous job
  - edit
  - submit
- Monitor status



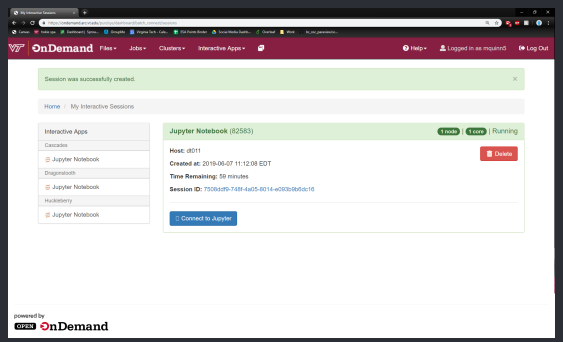
Open OnDemand

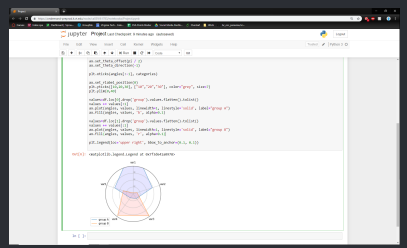
Features | Extensibility I

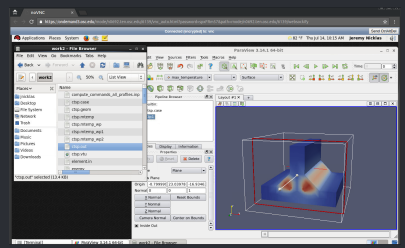
OnDemand uses a plug-in style wrapper to facilitate app development.

Users and sites can develop and share custom apps.

- Jupyter Notebooks
- Matlab
- Rstudio
- ParaView, Comsol, etc





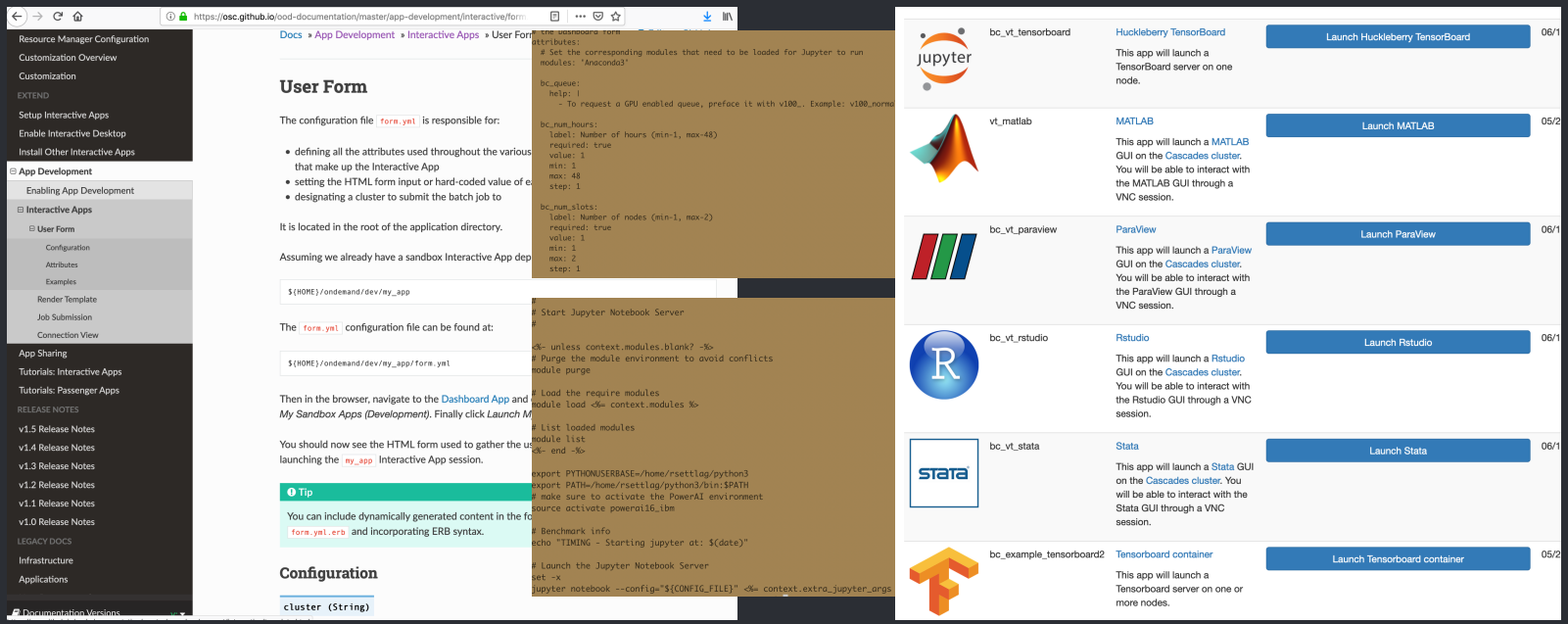


Open OnDemand

Features | Extensibility II

OnDemand has rich documentation.

`http://openondemand.org/`



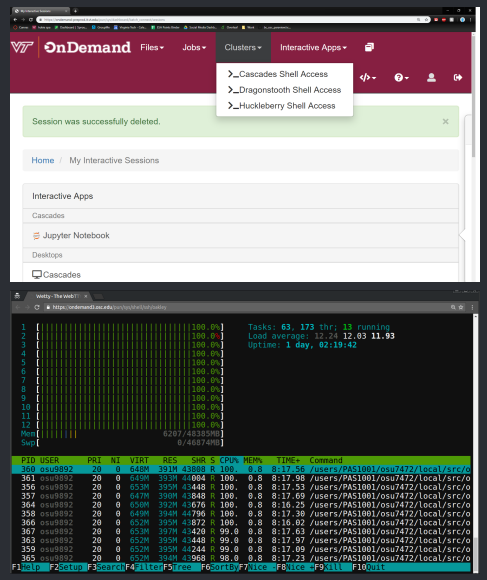
Open OnDemand

Successes | Teaching

Many class settings benefit from HPC as a computing platform.

Simplifying access helps students and instructors alike.

- Platform variability reduced
- Unified view of clusters
- Shell App
- Reduced time to compute
  - pre-OOD full class introduction
  - post-OOD less than 15 min



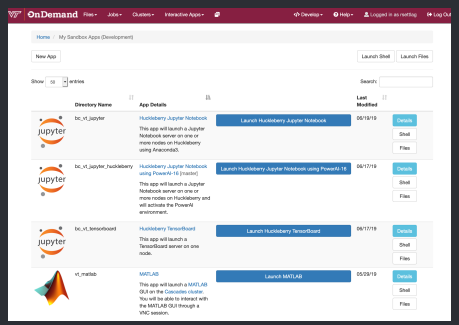
Open OnDemand

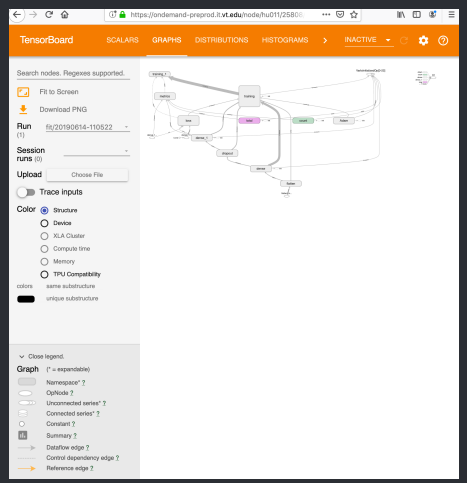
Successes | Hackathon

VT-OpenPOWER Hackathon Spring 2019.

Goal: bring a model and accelerate using PowerAI.

- >50 participants, 2 week
- Many had zero HPC experience
- OOD
  - Shell App
  - Jupyter Notebook with PowerAI
  - TensorBoard via Jupyter





Winning teams showed acceleration and scaling in diverse applications from GANs for CFD, RNNs in game AI, Siamese NN in cell type classification.

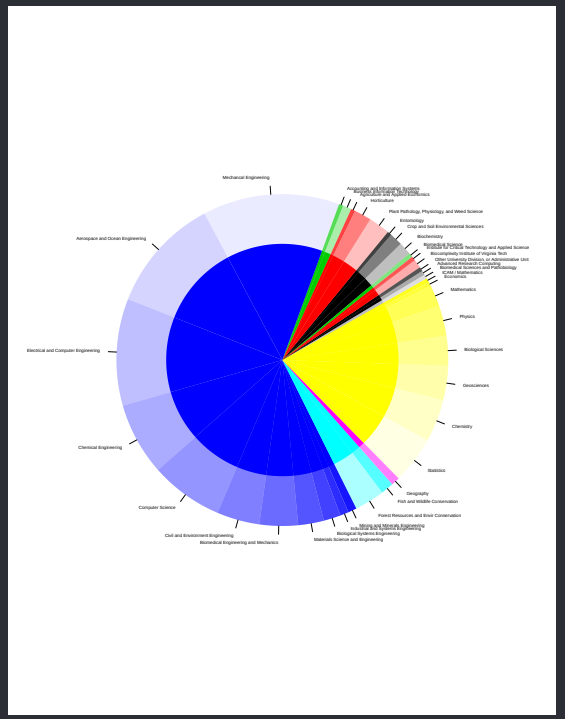
Open OnDemand

Successes | Research

Open OnDemand: HPC for everyone.

Goal: find users with HPC use cases and enable using OnDemand.

- New users
- English
- History
- Statistics
- Biomedicine/Health Care
- Artists



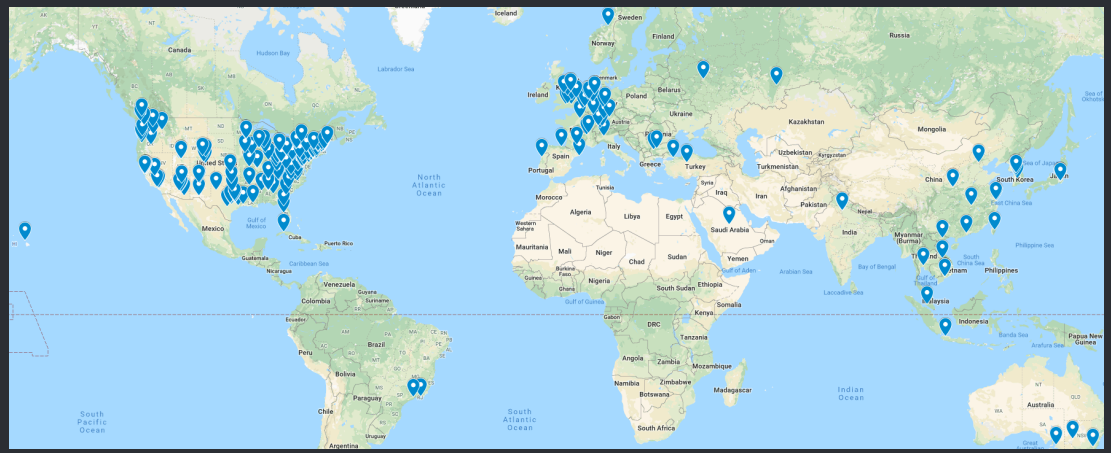
Open OnDemand

Adoption

Open OnDemand is a community driven open source project.

Our current user base is pretty broadly distributed. Unique installations:

- 136 US
- 70 International



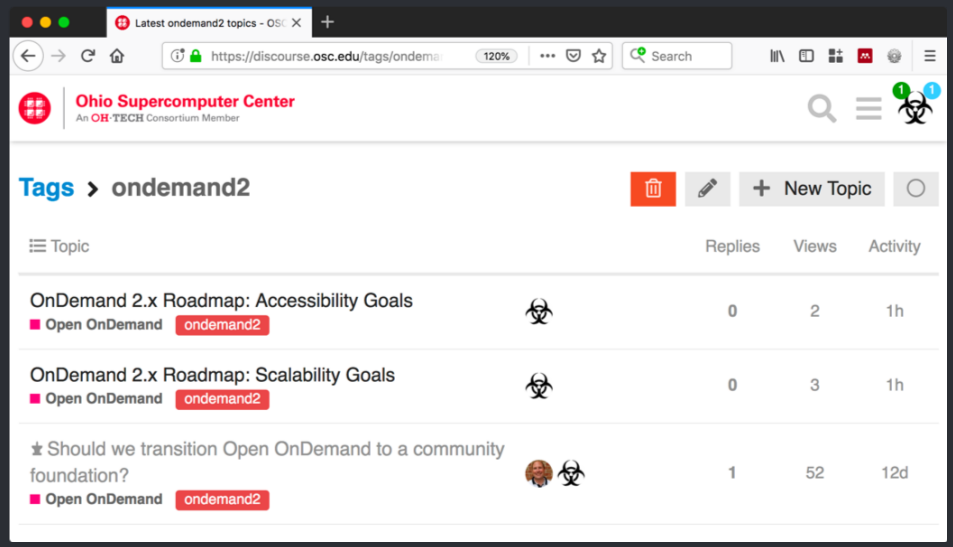
Open OnDemand2

Roadmap

Open OnDemand 2.x, NSF award #1835725

Four focus areas:

- Visibility
- Scalability
- Accessibility
- Engagement



Pinned topics on Discourse.

`https://discourse.osc.edu/tags/ondemand2`

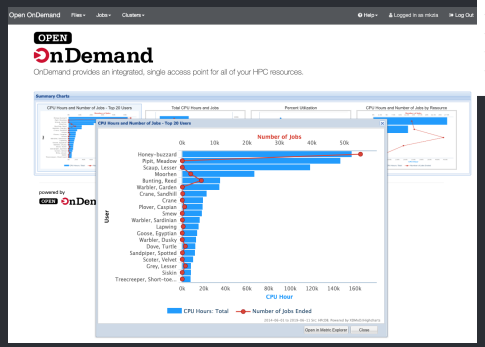
Open OnDemand2

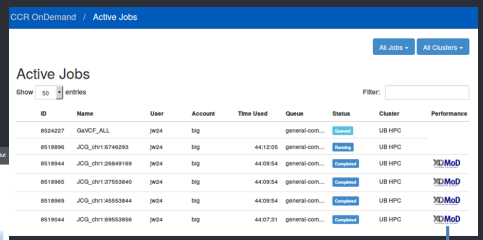
Visibility I

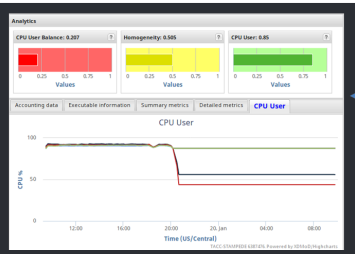
Enhance resource utilization visibility by integrating Open XDMoD.

Providing both novice and seasoned users with more resource utilization metrics will lead to more efficient computes.

- Overall cluster utilization metrics
- System performance
- Individual job performance
- Add GPU utilization





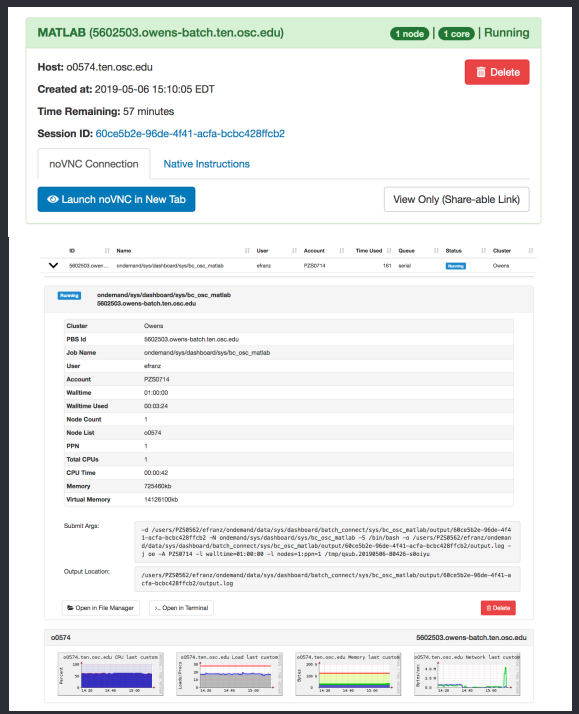


Open OnDemand2

Visibility II

Enhance resource utilization visibility by integrating Open XDMoD.

Real time metrics should also be visible.



- Active job performance
- Add button to connect to job via shell

Open OnDemand2

Scalability

Support more types of computing resources and software.

Enable less sophisticated users and enhance the veteran power user.

- Enable Git
- Enable pipelines/parameter sweeps
- Extend Files App
- Support spawning VMs in Cloud
- Bring your use case …

Open OnDemand2

Accessibility

Present HPC in a way that makes the computing resources more accessible to more users.

Often this means provide a more familiar interface.

- Further simplification of the Job Composer
- Further increase power of the Job Composer
- Build out more domain specific apps
- Desktop metaphore – completely automate job submission from users desktop
- Can we simplify the app creation process?
- Can we automate software switch discovery?

Open OnDemand2

Engagement

Open OnDemand is a community project.

We will actively discover new HPC use cases, advocate for the novice user, and ensure the community is engaged.

- Establish community of HPC users
- Establish community of administrators
- Continuously poll the community for development direction
- Establish Science and Client Advisory Group

Questions?

Thank you.