

Enabling AI/DL via Containers and Open OnDemand

.....

HPC Knowledge Meeting'20

Robert Settlage, Srijith Rajamohan

June 2020



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



Goals and Objectives

Showcase Open OnDemand for HPC, AI/DL



- Introduce ARC at VT
- Discuss HPC barriers
- Introduce OOD
 - features
 - adoption
- Create an app – BYOC
- AI/DL workload example



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



VIRGINIA
TECH. 1/12

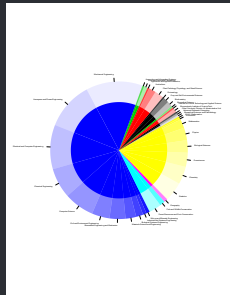
Advanced Research Computing

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

Basically, we are asking all our users to be computer scientists.

Think "tickets".



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



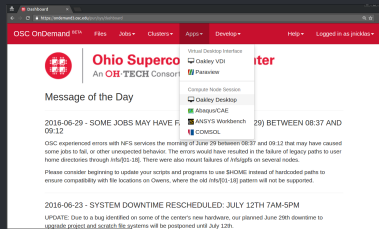
Open OnDemand

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



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



VIRGINIA
TECH. 6/12

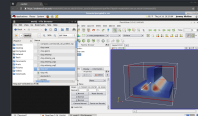
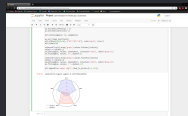
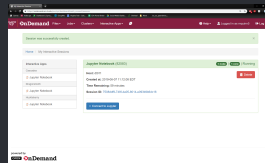
Open OnDemand

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

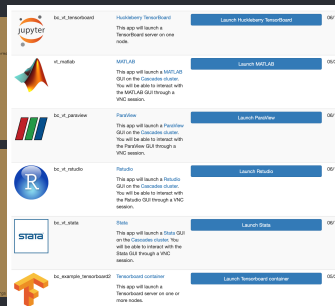
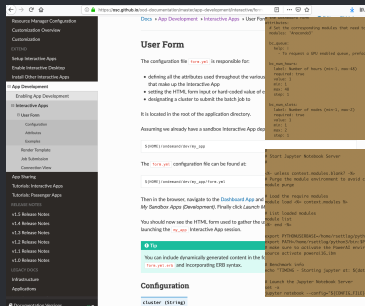


Open OnDemand

Features | Extensibility II

OnDemand has rich documentation.

<http://openondemand.org/>



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



VIRGINIA
TECH. 8/12

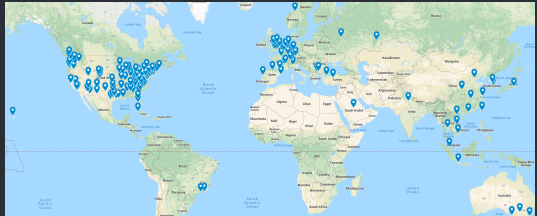
Open OnDemand

Adoption

Open OnDemand is a community driven open source project.

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

- 136 US
- 70 International



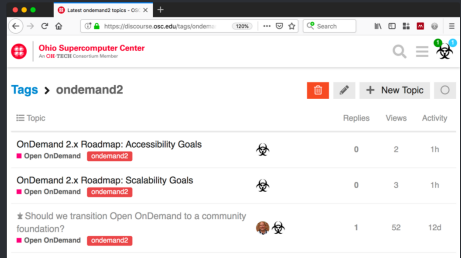
Open OnDemand2

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>



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



VIRGINIA
TECH. 10/12

Open OnDemand2

Create an app demo

Bring Your Own Container

Create an app to launch a user singularity container

- edit form.yml
- edit submit.yml
- edit before.sh
- edit script.sh
- test

<https://osc.github.io/ood-documentation/master/app-development/tutorials-interactive-apps.html>



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



VIRGINIA
TECH. 11/12

Containers in AI/DL?

Thank you.



Ohio Supercomputer Center



University at Buffalo

Center for Computational Research



VIRGINIA
TECH. 12/12