OPEN OnDemand & CloudyCluster Collaboration

Agenda

1. Intro
2. Overview of Open OnDemand
3. Overview of CloudyCluster and Omnibond Research Solutions
4. Discussion of the initial collaboration
5. Discussion of the next project
6. General Discussion around collaborations.
7. Demo (Alan has to run to another session)

OPEN OnDemand

Overview for 2023 Internet2 Community Exchange

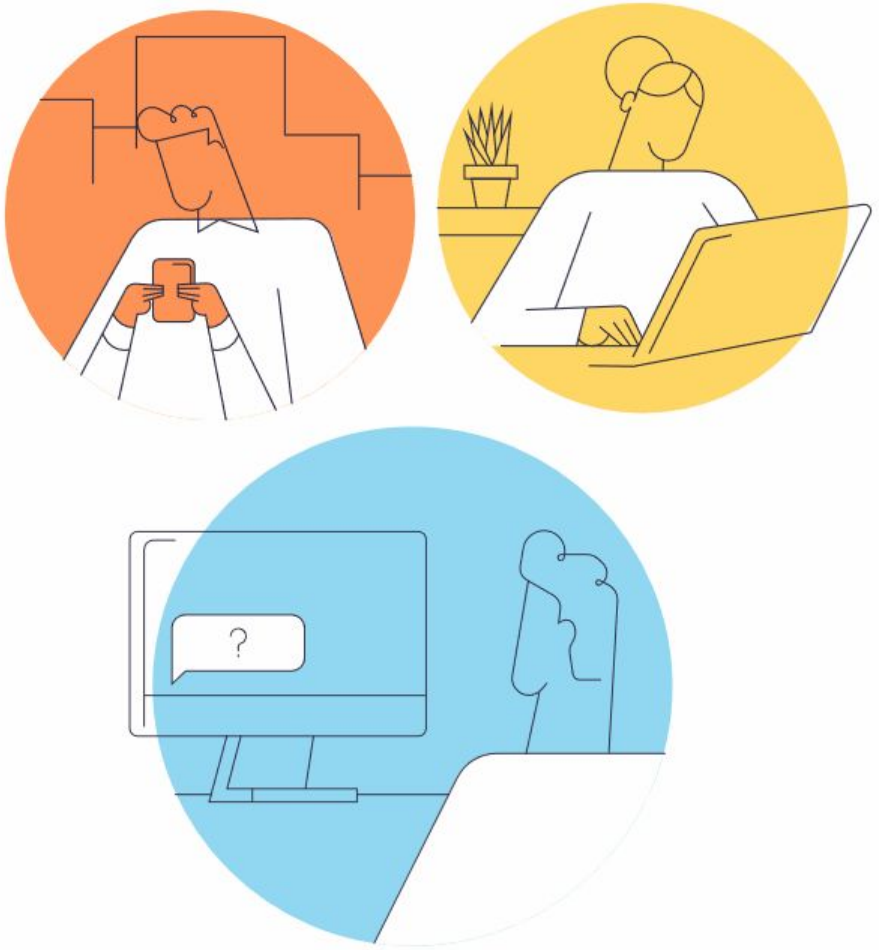
Alan Chalker, Ph.D.

Ohio Supercomputer Center

University at Buffalo Center for Computational Research

VIRGINIA TECH

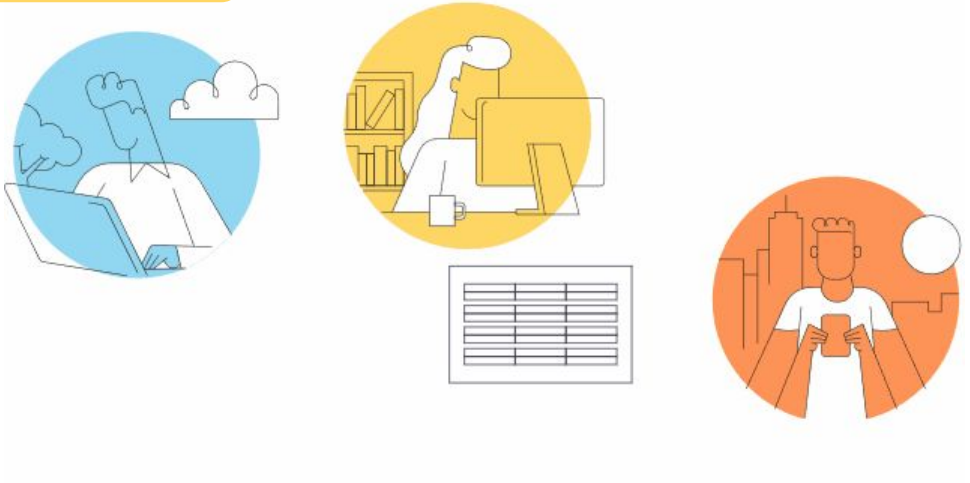
This work is supported by the National Science Foundation of the United States under the awards 1534949, 1835725, and 2138286



openondemand.org

Run Open OnDemand

Access your organization's supercomputers through the web to compute from anywhere, on any device.



Zero installation

Run Open OnDemand entirely in your browser. No client software installation required.

Easy to use

Start computing immediately. A simple interface makes Open OnDemand easy to learn and use.

Compatible with any device

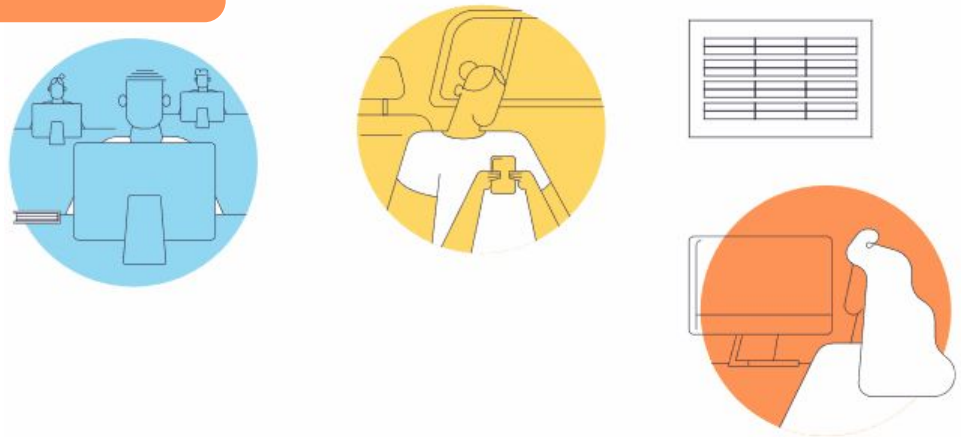
Launch on any device with a browser—even a mobile phone or tablet.

openondemand.org/run

Install Open OnDemand

OPEN OnDemand

Administer remote web access to your supercomputers to transform the way users work and learn.



Low barrier to entry

Empower users of all skill levels by offering an alternative to command-line interface.

Free and open source

Install Open OnDemand for free, and gather knowledge from our large open-source community.

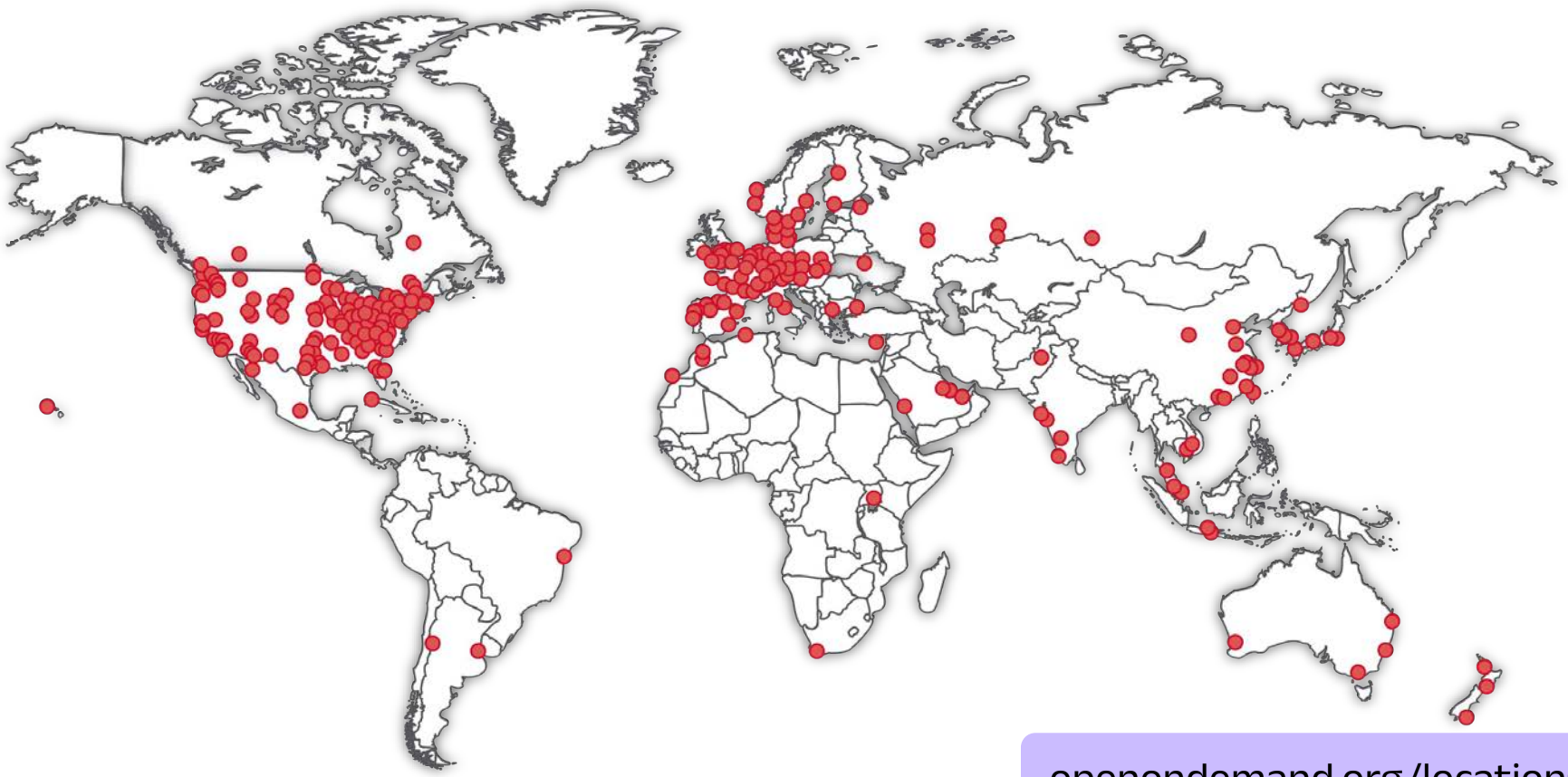
Configurable and flexible

Create and deploy your own applications to meet your users' unique needs.

openondemand.org/install

Deployed Worldwide

OPEN OnDemand



openondemand.org/locations

Example Deployments

Don't see an organization?

The more the merrier! Let us know any that belong on the list

openondemand.org/orgs

Community Events

Tips and tricks calls

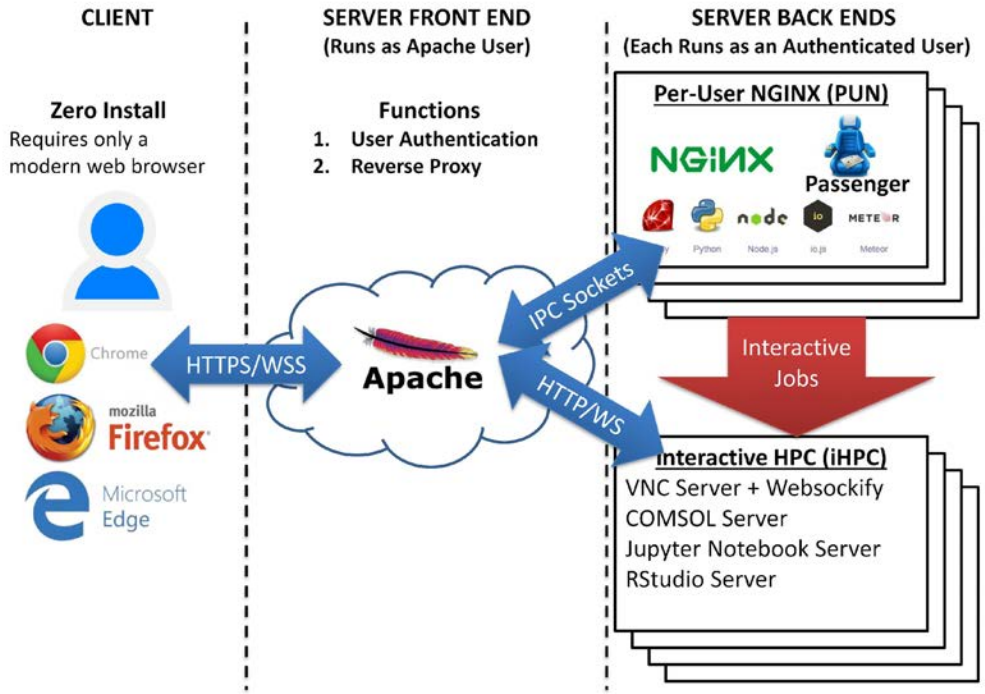
Hosted by the larger Open OnDemand community, tips and tricks webinars share best practices for setting up and using Open OnDemand. They take place on the first Thursday of every month at 1 p.m. ET.

Open office hours

Hosted by our development team, Zoom open office hours are the perfect opportunity to ask questions or make a suggestion. They are held on the second Tuesday of every month from 11:15 a.m. to 12:45 p.m. ET.

openondemand.org/events

Architecture



Requirements

RedHat/CentOS/Rocky/Ubuntu

Software Collections repos

Lsof

Sudo

Nmap-ncat

TurboVNC

websockify

openondemand.org/architecture

Enabled Utilities & Apps

OPEN OnDemand

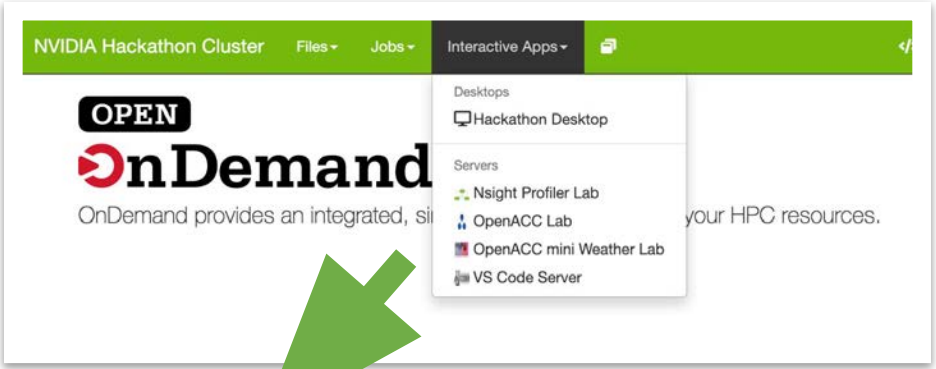
Abaqus
ANSYS
COMSOL
Coot
CSD
Galaxy
Grace
Grafana
Grid Engine
IDL
Jupyter
Kubernetes

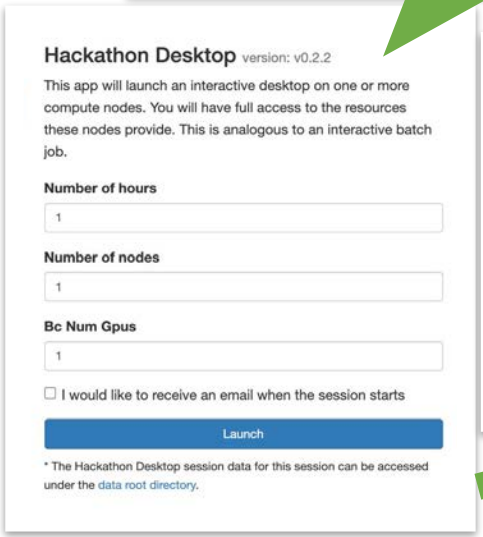
LSF
Lumerical
Mathematica
MATLAB
Meshroom
NAGIOS
Octave
Open XDMoD
Ovito
Paraview
PBS Professional
QGIS

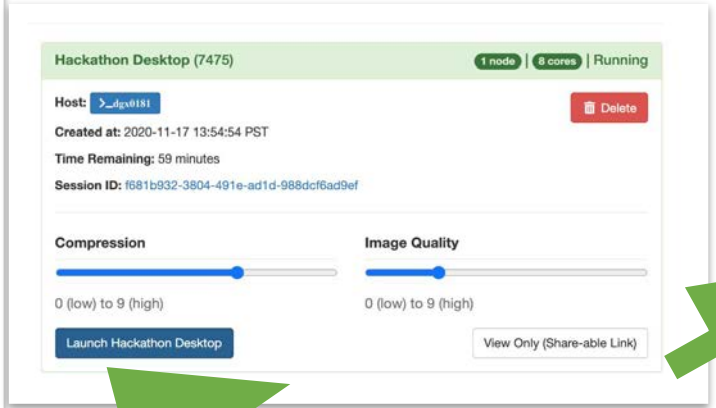
RELION
RStudio
SAS
Shiny
Slurm
Spark
STATA
Tensorboard
Torque
VISIT
Visual Studio Code
VMD

openondemand.org/apps

Desktop Example

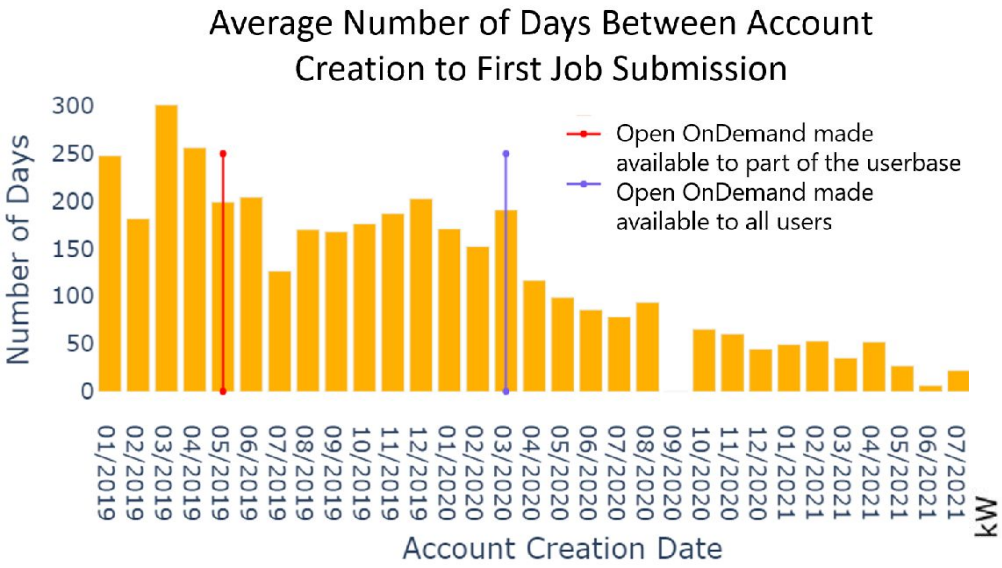






openondemand.org/nvidia

Idaho National Lab

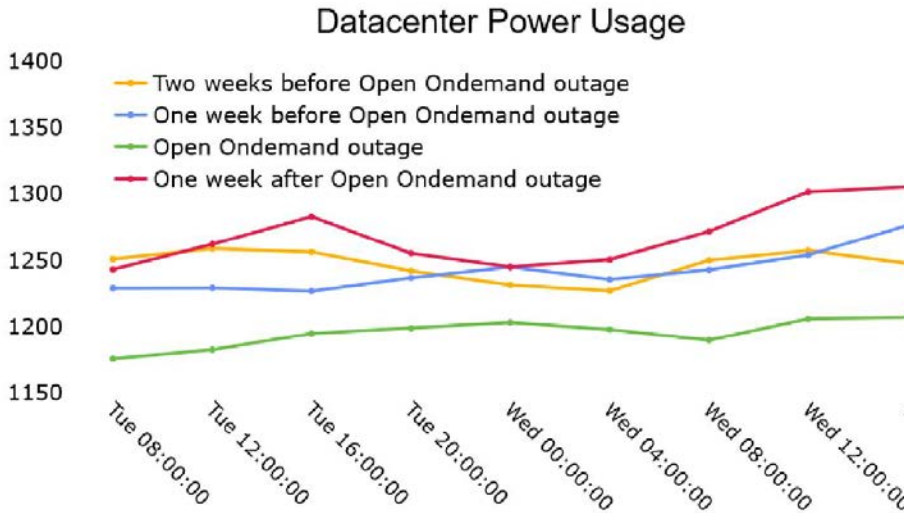


19 times shorter

In days between account creation and first job submission

3.7% drop

In power used during an unscheduled Open OnDemand outage



openondemand.org/idaho

Example Case Studies

WEB PORTALS

Open OnDemand aids engineering school's supercomputing curriculum

■ Research Landscape ▲ Invention/Discovery

In 2019, the Milwaukee School of Engineering (MSOE) unveiled a major addition to its campus: the Dwight and Dian Diercks Computational Science Hall, featuring Rosie the supercomputer. The facility opened in the wake of the launch of MSOE's bachelor's degree in computer science, with a curriculum focused on the growing field of artificial intelligence.

Alumnus Dwight Diercks, who along with his wife Dian donated $34 million for the facility, has worked for the technology company NVIDIA since 1994 and today serves as its senior vice president of software engineering. MSOE worked closely with the company on assembling the necessary hardware and software for the supercomputer, which features NVIDIA graphics processing units (GPUs).

NVIDIA recommended that MSOE adopt Open OnDemand, an open-source high performance computing portal developed by the Ohio Supercomputer Center (OSC). Open OnDemand is used by researchers and college students around the world to access supercomputing resources anywhere from any device.

Derek Riley, a professor and program director for electrical engineering and computer science at MSOE, reports that most of the jobs run on Rosie use the Open OnDemand portal. MSOE established the supercomputer primarily as a teaching tool for undergraduates studying machine learning and data science, and Open OnDemand has made it easy for them to directly access Rosie, Riley said. Students can avoid time-consuming technical setups and instead gain more experience using the supercomputer to answer a variety of scientific and engineering questions.

"It can't be overstated how important it is for students to focus on the problems we want them to learn," he said.

MSOE's approach to supercomputing also reflects the work environment that many students will find themselves in after graduation, as most employers have engineering teams that handle the more technical aspects of supercomputer setup and maintenance while the data scientists conduct analyses, Riley said.

Not only has Open OnDemand been beneficial for undergraduates to use, but MSOE's own system administrators have found it easy to learn and manage, Riley said.

"We've had a really great experience using it—we've been really happy with it," he said. "We've been able to use it primarily out of the box, and it's the main entry point for students and faculty to the cluster."

Read and share online: osc.edu/r77/msoe

Above: Rosie the supercomputer supports the Milwaukee School of Engineering's computer science program.

> "In order to learn from millions and millions of chemical structures, we need a lot of computational power," Ning said.
>
> The Ohio Supercomputer Center (OSC) is integral to the project. Each model requires the use of one graphical processing unit (GPU) and 96 GB of RAM for the data generated, Ning noted.

> Alumnus Dwight Diercks, who along with his wife [illegible] donated $34 million for the facility, has worked for the technology company NVIDIA since 1994 and today serves as its senior vice president of software engineering. MSOE worked closely with the company on assembling the necessary hardware and software for the supercomputer, which features NVIDIA graphics processing units (GPUs).

ARTIFICIAL INTELLIGENCE

Accelerating drug discovery with AI

■ BIOLOGICAL SCIENCES | ▲ INVENTION/DISCOVERY

Xia Ning has a large portfolio of research projects at The Ohio State University that focus on understanding how artificial intelligence can be used to solve issues in health care.

Discovering new drugs to treat disease is one of Ning's goals. Traditional research methods, which call for lengthy trials with animal models, have disadvantages.

"To find a single drug is costly and time consuming," said Ning, an associate professor who holds joint appointments in Ohio State's College of Medicine and College of Engineering.

Ning is drawing on her expertise in computer science and biomedical informatics to create a new path to drug discovery. Her lab examines millions of small molecules and uses that information to create novel models that could be [illegible] candidates for drugs.

[illegible] from millions and millions [illegible] we need a lot of [illegible] said.

[illegible] Computer Center (OSC) is integral [illegible] project. Each model requires the use of [illegible] graphical processing unit (GPU) and 96 GB of RAM for the data generated, Ning noted.

Compared to the conventional central processing units (CPUs) researchers may use in their labs, GPUs offer a significant boost in computing power. OSC features GPUs across its systems to allow clients to efficiently process large amounts of data, and consistently upgrades its hardware to ensure access to the most cutting-edge technologies.

Those computing resources have helped Ning reach important milestones in her drug discovery work. Research findings from the project have been accepted for publication in the prestigious journal Nature Machine Intelligence and also have attracted new grant funding to the lab.

Ning has been a heavy user of high performance computing centers since her days as a faculty member at Indiana University. She joined Ohio State in 2018.

"The first thing I did here was look for similar resources," Ning recalled. "I didn't think about having my own cluster—that would have taken me a lot of effort."

In addition to OSC's capabilities and technical support, Ning appreciates its availability to her lab members. OSC offers Open OnDemand, an interface funded by the National Science Foundation that allows users to remotely access the Center resources online from any device.

"(OSC) provides a very nice Open OnDemand tool that we can access through the web," Ning said. "I believe all of my students are using it. It's very easy for us to use, and we don't need to worry about maintenance or software installation."

Ning also employs OSC for teaching Python programming in her biomedical course. The Center creates a class project through which her students can quickly access computational resources and see results immediately through the OSC web portal, she said.

Over the next year, Ning will continue to work with OSC on the drug discovery project, as well as on research on predicting peptides that could be used for vaccine development.

"As long as we're doing research," Ning said, "we'll rely on OSC." ●

Website: u.osu.edu/ning.104

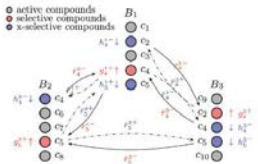
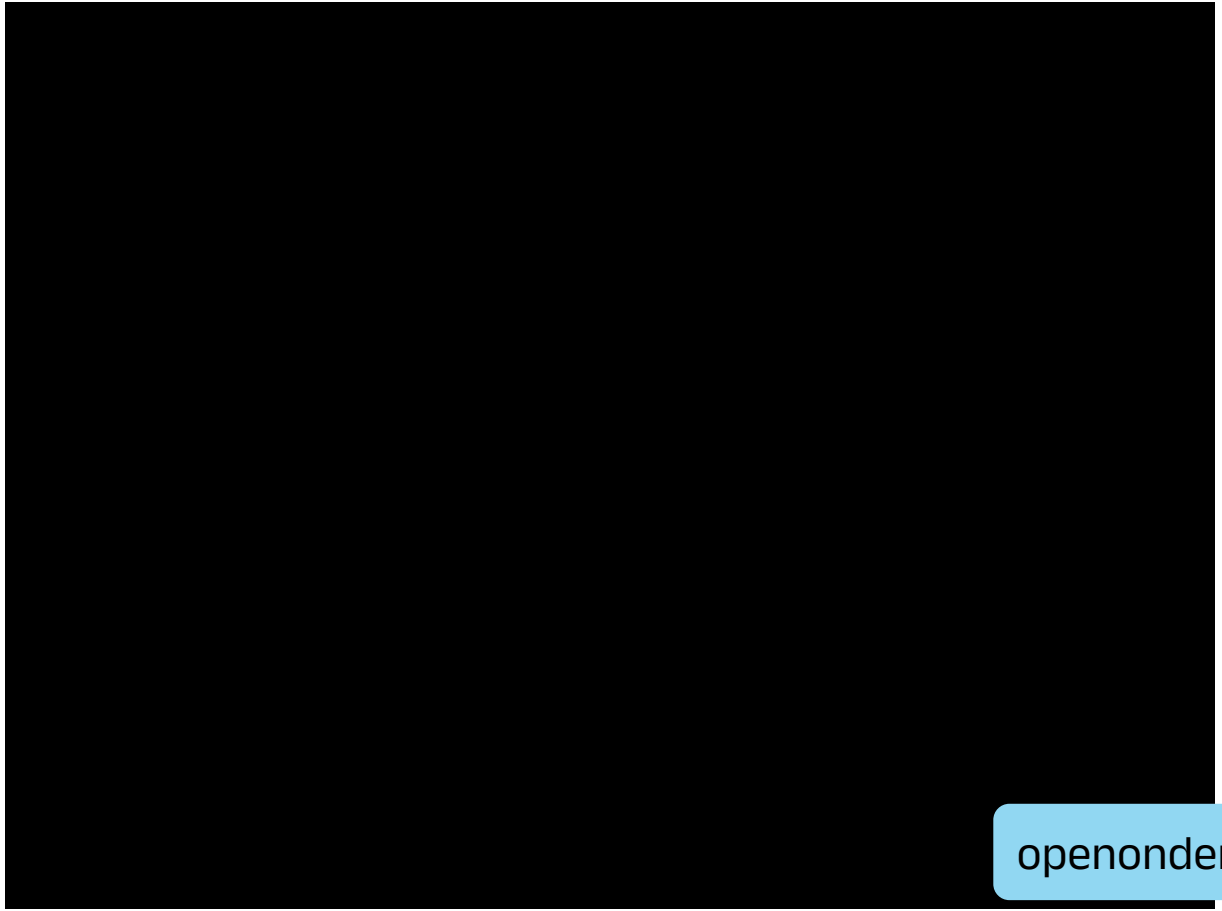


Figure: Advanced AI and large-scale computing enabling multi-purpose drug discovery.

openondemand.org/casestudies

Any Device, Anywhere



openondemand.org/anydevice

omnibond®

Research Computing Solutions Overview for:

Boyd Wilson, CEO

Leadership Team

Omnibond

a customer-focused software engineering and support company

- Over 40 years combined experience in facilitating and supporting academic and corporate research in using the tools and technologies of advanced computing
- Experience at the working and technical, project and executive management levels at Clemson, Purdue and Miami Universities and the National Center for Supercomputing Applications located at UIUC
- Expertise in software development, systems integration, operations, applications support, data transmission, identity and access management, customer relations and research facilitation and engagement
- Founded ACI-REF http://www.aciref.org and CaRCC http://carcc.org
- Over three decades of funded projects from NSF, DoD, DoE, NSA, NIST and DARPA
- INTERNET2 Presidential Fellow & CSTAAC Committee Member

- Identity & Security Management
  - NetIQ Identity Manager Connectors
  - Thousands of customers, sold through Novell/Micro Focus/OpenText, since early 2000’s
  - OmniPasskey Passwordless MFA Shibboleth Plug-in
- Computer Vision & AI
  - TrafficVision - AI based Automated Incident Detection (AID) & Data from existing cameras on roadways
  - BayTracker - Retail Vehicle Tracking and Timing
  - Port Observer - Drayage Queuing, AIS, Dashboard for Ports
- Cloud HPC and Storage Orchestration
  - CloudyCluster
  - OrangeFS
  - Eureka Project
  - Custom Cloud <-> On-Prem Integration

Software Products

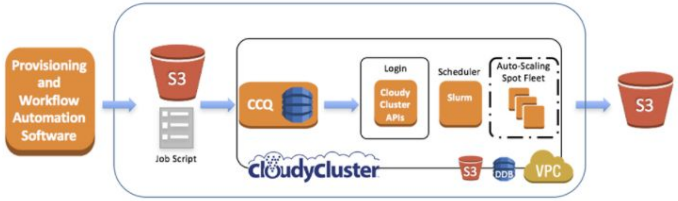
Omnibond

a customer-focused software engineering and support company

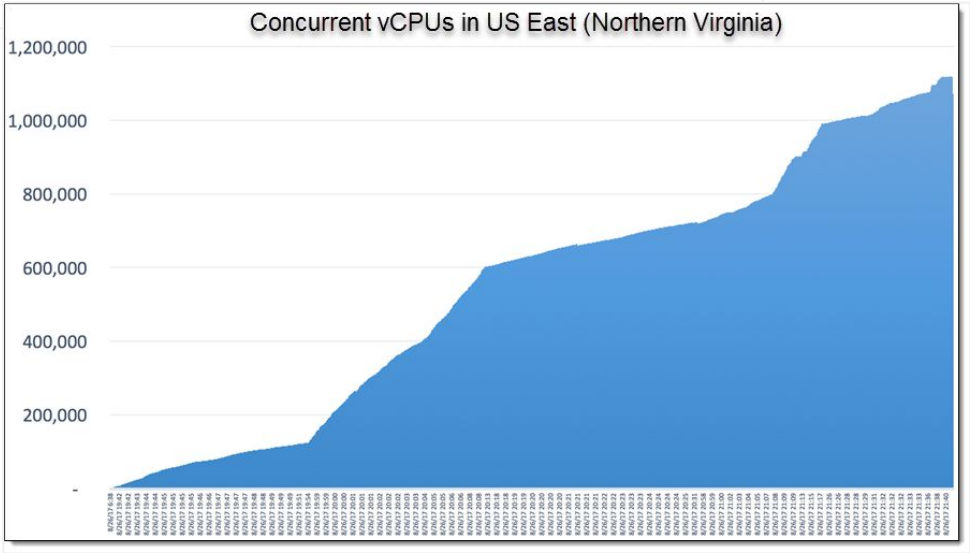
Scaling on AWS

AWS News Blog

Natural Language Processing at Clemson University – 1.1 Million vCPUs & EC2 Spot Instances



Provisioning and Workflow Automation Framework



Data Intensive Computing Ecosystems (DICE)
School of Computing, Clemson University

https://aws.amazon.com/blogs/aws/natural-language-processing-at-clemson-university-1-1-million-vcpus-ec2-spot-instances/

Scaling

Google HPC Blog Post

Cloud against the storm: Clemson's 2.1 million VCPU experiment

https://cloud.google.com/blog/topics/hpc/clemson-experiment-uses-2-1-million-vcpus-on-google-cloud

THENEXTPLATFORM

Urgent HPC can Burst Affordably to the Cloud

https://www.nextplatform.com/2020/01/08/urgent-hpc-can-burst-affordably-to-the-cloud/

- 133,573 GCP Instances at peak
- 2,138,000 vCPUs at peak
- 6,022,964 vCPU hours

Processed 2,479,396 hours (~256TB) of video data

- ~4 hours of runtime
- ~1M vCPU within an hour
- ~1.5M vCPU within 1.5 hours
- 2.13M vCPU within 3 hours

Total Cost: $52,598.64 USD
Average cost of $0.008 USD per vCPU hour

CLEMSON UNIVERSITY

Google Cloud

intel

Cloudy Cluster by Omnibond

TrafficVision

Real-time HPC & AI TrafficVision

Features

Incident Detection

- Stopped Vehicle
- Slowed Traffic
- Debris in Roadway
- Low Visibility
- Pedestrians
- Wrong-Way Drivers*

Alerts, Incident Images & Clips

Real-Time Data Collection

- Speeds Per Lane / Per Direction
- Vehicle Counts Per Lane / Direction
- Lane Occupancy and Density
- Congestion Index
- Classification:
  **Motorcycle / Car / Truck / Large Truck

Data CSV export & JSON API

Customers

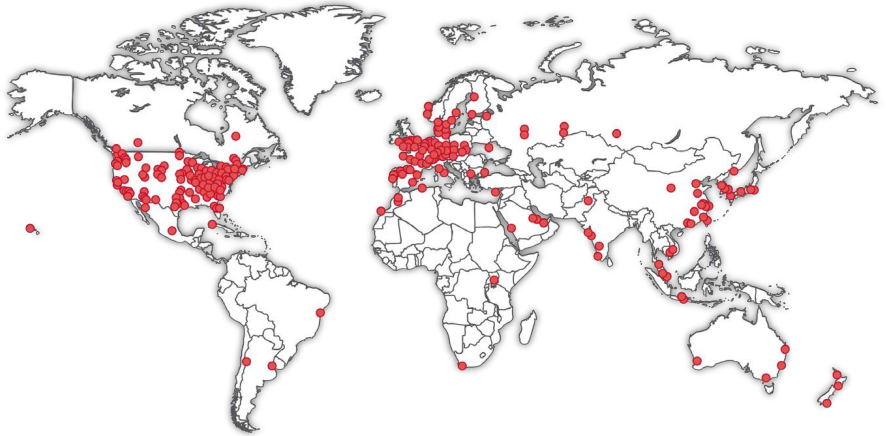
Caltrans · CDOT · Calgary · FDOT · 360ns · GDOT · SCOUT · SKYLINE · kapsch · WSP · Ontario · Michigan DOT · Minnesota DOT · MoDOT · NMDOT · NTTA · PENNA TURN PIKE · Peraton · Port of LONG BEACH · RCTC · SCDOT · TRIMARC

HPC User Experience

OPEN OnDemand

Integration of OOD into CloudyCluster

- As part of the Cloudify Gateways program (Omnibond, OSC, U Buffalo, Va Tech)
- Full deployment automated as part of the CloudyCluster launch
- Working on project based UI initially to support data scientist project teams.
- The Collaboration Continues



Open OnDemand Deployments

"Your browser is the supercomputer: On Demand is a no-tears shortcut to research-computing" from Matt Windsor of University of Alabama at Birmingham. A key phrase: "No experience necessary"

This material is based upon work supported by the National Science Foundation under grant numbers 1534949 and 1835725. and under active development by a team from the Ohio Supercomputer Center, U. of Buffalo CCR, and Virginia Tech

Relion & CRYO-EM

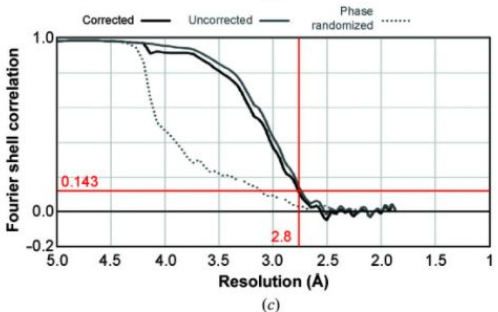
Working with the Cianfrocco Lab at University of Michigan to test data transfer and scalability to the Cloud for CRYO-EM projects using RELION

Also working with related Startups in the space that don't have access to large on prem clusters.

Initially Introduced as part of the Cloudify Program through the SGCI.

Cianfrocco is willing to consult with customers in conjunction with Omnibond and CloudyCluster

Another CRYO-EM project is in progress with the ERN ECOSYSTEM for RESEARCH NETWORKING



High-resolution cryo-EM using beam-image shift at 200 keV

Jennifer N Cash [1], Sarah Kearns [1], Yilai Li [1], Michael A Cianfrocco [1]

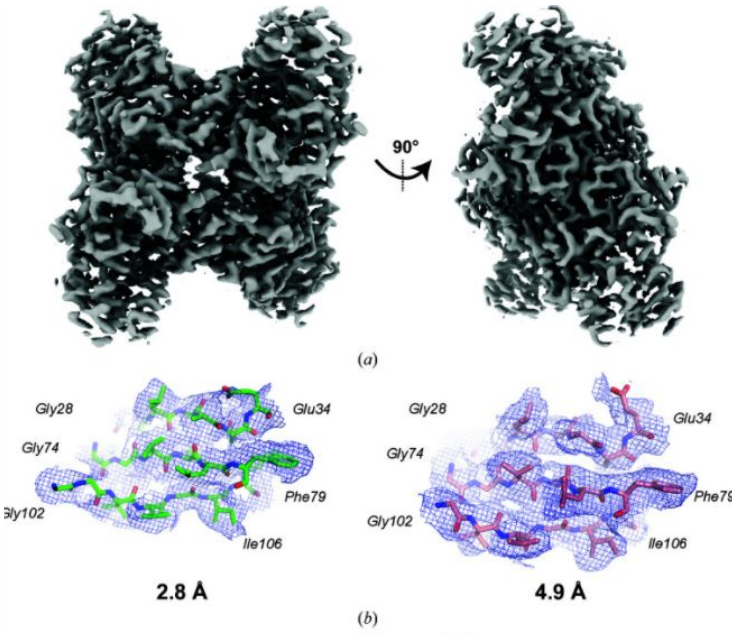
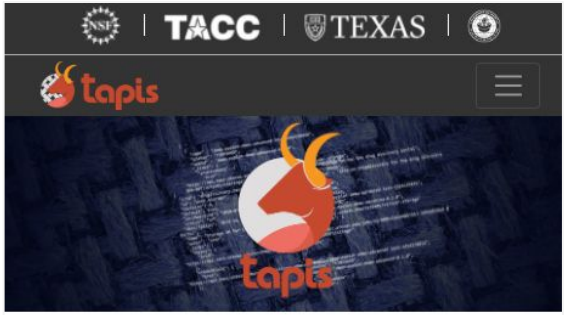


Figure 5 Final aldolase reconstruction at 2.8 Å resolution. (*a*) Sharpened aldolase reconstruction at 2.8 Å resolution. (*b*) Example densities and models for aldolase at 2.8 and 4.9 Å resolution. (*c*) FSC curve for the final reconstruction.

TAPIS - PEARC ‘22 Poster, Gateways ‘22 Paper



Tapis

Unify, simplify, automate and secure your interactions with advanced computing infrastructure.

The Tapis Framework provides a hosted, unified web-based API for securely managing computational workloads across institutions so that experts can focus on their research instead of the technology needed to accomplish it.

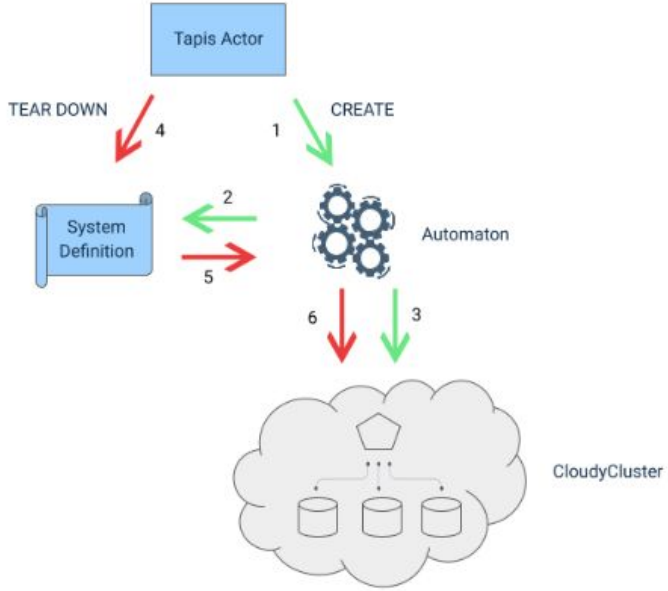
Projects Using Tapis

- A2CPS
- Bridging Barriers
- CyVerse
- DesignSafe
- Drug Discovery Portal
- ECCO
- Ike 'Wai (at the University of Hawaii Manoa)
- iReceptor
- iR+
- NeuroNex 3DEM
- Planet Texas 2050
- Science Gateways Community Institute
- Synergistic Discovery and Design Environment (SD2E)
- VDJServer
- UTRC

https://github.com/omnibond/automaton

Extending Tapis Workflow Management Framework with Elastic Google Cloud Distributed System using CloudyCluster by Omnibond

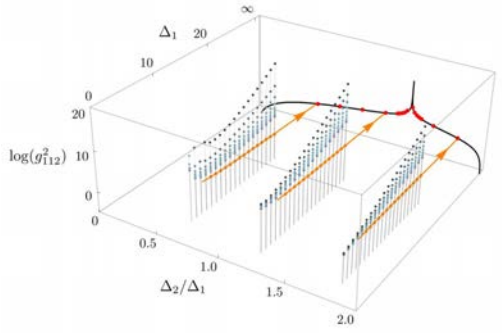
ERIC LAM, University of Hawaii, Manoa, USA
SEAN CLEVELAND, University of Hawaii - Information Technology Services - Cyberinfrastructure, USA
COLE MCKNIGHT, Omnibond Systems, USA
BOYD WILSON, Omnibond Systems, USA
RICHARD CARDONE, Texas Advanced Computing Center, USA
JOE STUBBS, Texas Advanced Computing Center, USA



SDPB Solver MPI Tuning

The Problem of Strong Coupling

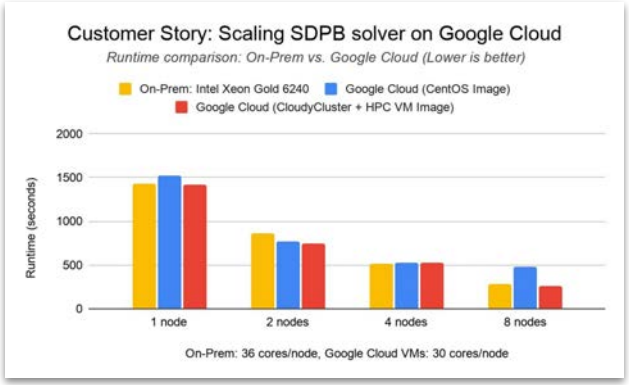
Quantum field theory (QFT) is a universal language for theoretical physics, describing the Standard Model of particle physics, early universe inflation, and condensed matter phenomena such as phase transitions, superconductors, and quantum Hall fluids. A triumph of 20th century physics was to understand weakly coupled QFTs. However, weakly interacting systems represent a tiny island in theory space and cannot capture many of the most interesting physical phenomena.



https://bootstrapcollaboration.com/

Our Collaboration

Recently, members of our collaboration discovered new bootstrap techniques that apply in general dimensions. In the past few years we have applied these techniques to a wide variety of seemingly unrelated problems: to perform the world's most precise analysis of the 3d Ising model, to constrain strongly coupled theories of physics beyond the Standard Model, to aid in classifying superconformal field theories, to derive locality and black hole thermality in models of quantum gravity, and to prove irreversibility of renormalization group flows. We believe this is the beginning of a much larger enterprise, crossing traditional boundaries between string theory, condensed matter physics, and phenomenology, and making strong connections to modern mathematics and computer science.



https://cloud.google.com/blog/topics/hpc/introducing-hpc-vm-images

To expand the collaboration's computation capabilities, Walter Landry wanted to see how SDPB would scale on Google Cloud. Working with Omnibond's **CloudyCluster** and leveraging the HPC VM image, Landry achieved comparable performance and scaling to an on-premises cluster at Yale, based on Intel Xeon Gold 6240 processors and Infiniband FDR.

Turn-Key Elastic HPC, HTC, & Storage in GCP

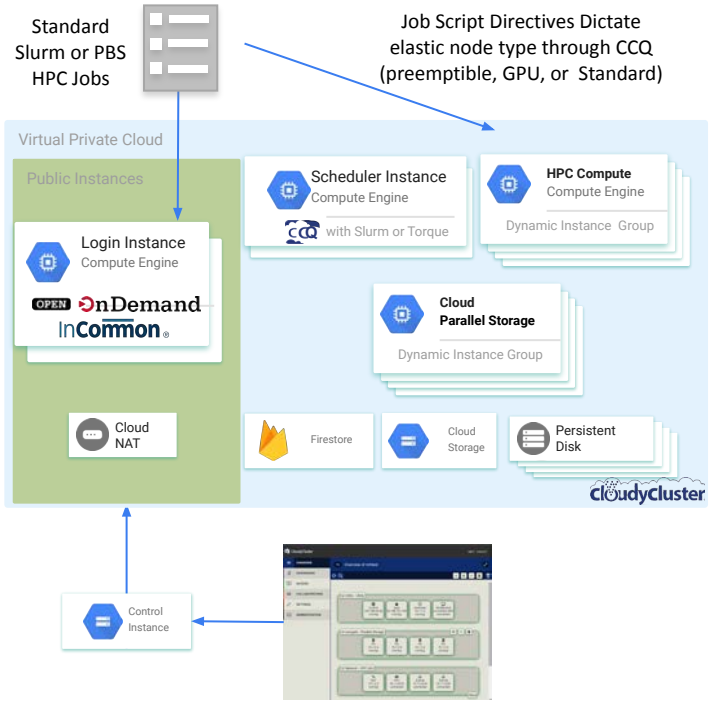
Create a familiar, secure & fully operational computational cluster in minutes, complete with:

Encrypted Storage: GCS, OrangeFS on PD

Compute: Job Driven Elastic Compute through CCQ (Preemptable, On-Demand, GPU)

Schedulers: Torque & SLURM with the CCQ Meta-Scheduler
Supports Billing Labels per job/user
Placement Groups, BulkAPI, Tier1 Networking
Supports Multiple Custom Images

End User Friendly UI: Open OnDemand



Includes Familiar Software

HPC Libraries: Boost, Cuda Toolkit, Docker, FFTW, FLTK, GCC, Gengetopt, GRIB2, GSL, HDF5, Intel MPI, Intel Runtimes, ImageMagick, JasPer, mpich, NetCDF, NumPy, Octave, OpenCV, OpenMPI, PROJ, R, Rmpi, SciPy, SWIG, WGRIB, UDUNITS, .NET Core, Singularity, Queue, Picard, xrootd, etc…

HPC Software: Ambertools, ANN, ATLAS, BLAS, Blast, Blender, Burrows-Wheeler Aligner, CESM, GROMACS, JupyterLab, LAMMPS, NCAR, NCL, NCO, nwchem, OpenFoam, papi, paraview, Quantum Espresso, SAMtools, WRF, Galaxy, Vtk, Su2, Dakota, Gatk, etc…

ML Software: MIpack, NuPIC, Octave, OpenCV, PICARD, Queue, Scikit-learn, Tensorflow

Sample Jobs: Intel Cluster Checker, Mpi_prime, wrf, etc…

cloudycluster

Google Cloud Marketplace

Turn-Key Elastic HPC, HTC, & Storage in AWS

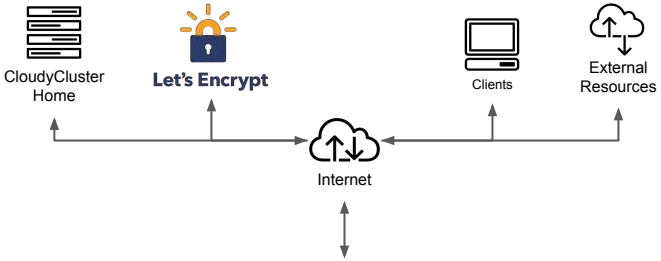
Create a familiar, secure & fully operational computational cluster in minutes, complete with:

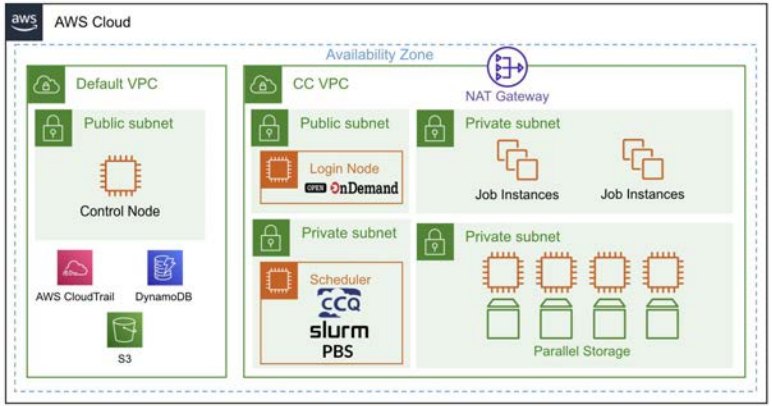
Encrypted Storage: EBS, OrangeFS on PD

Compute: Job Driven Elastic Compute through CCQ (Spot, On-Demand, GPU)

Schedulers: Torque & SLURM with the CCQ Meta-Scheduler
- Supports Billing Tags per job/user
- Placement Policies, Advanced Networking
- Supports Multiple Custom Images

End User Friendly UI: Open OnDemand





Includes Familiar Software

HPC Libraries: Boost, Cuda Toolkit, Docker, FFTW, FLTK, GCC, Gengetopt, GRIB2, GSL, HDF5, Intel MPI, Intel Runtimes, ImageMagick, JasPer, mpich, NetCDF, NumPy, Octave, OpenCV, OpenMPI, PROJ, R, Rmpi, SciPy, SWIG, WGRIB, UDUNITS, .NET Core, Singularity, Queue, Picard, xrootd, etc…

HPC Software: Ambertools, ANN, ATLAS, BLAS, Blast, Blender, Burrows-Wheeler Aligner, CESM, GROMACS, JupyterLab, LAMMPS, NCAR, NCL, NCO, nwchem, OpenFoam, papi, paraview, Quantum Espresso, SAMtools, WRF, Galaxy, Vtk, Su2, Dakota, Gatk, etc…

ML Software: Mlpack, NuPIC, Octave, OpenCV, PICARD, Queue, Scikit-learn, Tensorflow

Sample Jobs: Intel Cluster Checker, Mpi_prime, wrf, etc…

cloudycluster

The Meta-Scheduler Approach

Scheduler Independent

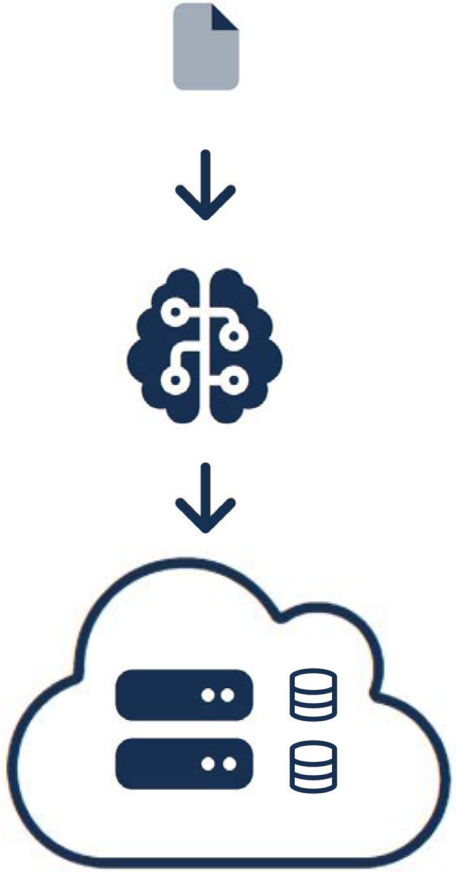
- Torque
- Slurm

Allow for Meta-Scheduler Directives

- Instance Type, each job can have a different instance type and billing model.
- GPU
- Spot, Preemptible
- Billing Tags/Labels
- Volume Type
- Custom Images for Different Jobs

Turns “Scheduler” into Dispatcher Handling:

- Instance and appropriate subnet provisioning
- Instance deletion



Self Paced Learning (Ask for Free Credits for Lab)

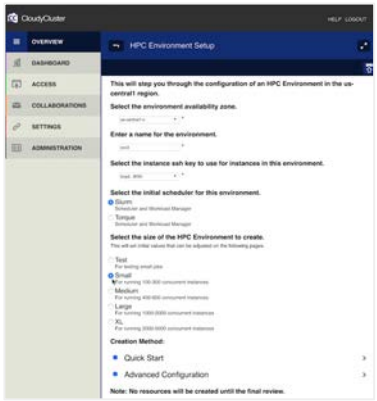
← Omnibond: Creating an HPC Environment in Google Cloud with CloudyCluster

Start Lab 01:30:00

Student Resources

- Job Scripts
- Using the CCQ Meta-scheduler
- Blog Articles
- Behind the Scenes Architecture of CloudyCluster v3 on GCP

Omnibond: Creating an HPC Environment in Google Cloud with CloudyCluster



1 hour 30 minutes 5 Credits ☆☆☆☆☆

This lab was developed with our partner, Omnibond. Your personal information may be shared with Omnibond, the lab sponsor, if you have opted-in to receive product updates, announcements, and offers in your Account Profile.

GSP862

Google Cloud Self-Paced Labs

CloudyCluster Online Training

Learn How To:

- Setup GCP Account, Check Quotas, Configure IAM
- Launch a CloudyCluster instance from the GCP Marketplace
- Setup Admin user and credentials
- Choose the size of your environment (Test, SM, M, L, XL)
- Customize options for the environment -> Select Create
- CloudyCluster automatically creates a Private Network, Login, Scheduler (Slurm or PBS), and Parallel Storage all ready to run HPC jobs.
- Upload your data / Tweak your HPC job script
- Run your job and Download Results.

Initial Collaboration Story
In the Style of a Fireside Chat

Cloudify Gateways

Cloudify Program:

- **Solicitation**: Program from Science Gateways Community Institute (SGCI) to give out Google Cloud Credits for research projects to be “Cloudified”
- **Submissions**: Submissions over 3 years a batch per year
- **Total**: 15 Awardees over the 3 years
- **Awards:** Google Cloud Credits and Cloudify technical assistance from Omnibond

Cloudify Gateways Proposal

“Scientists accessing HPC resources today, either at the campus or national level typically **need advanced knowledge of Linux**, familiarity with command-line interfaces and installation and configuration of custom client software. These additional requirements **create an accessibility gap** that for many scientists serves as **an impediment to adopting HPC** for their research. This accessibility gap is a long standing and recognized issue in HPC. **Removing, or even reducing this impediment to HPC** use, **will immediately increase the productiveness** of computationally intensive research. **Lowering the barrier** for access to HPC systems has long been the “**Holy Grail**” in the HPC community.”

…

“Open OnDemand (OOD) is an **innovative**, open-source, **web-based portal for accessing HPC services** that removes these intricacies, and in so doing reduces time to science for researchers by **facilitating their access to HPC** resources.”

…

“However, the **CloudyCluster team** is explicitly focused on providing an easy to use **Cloud-based HPC infrastructure**”

OPEN OnDemand

Supercomputing. Seamlessly. Open, Interactive HPC Via the Web.

About Open OnDemand

Provides an easy to install and use, web-based access to supercomputers, resulting in intuitive, innovative support for interactive supercomputing.

Features include:

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

Cloudify Project

- Need: Many HPC centers are evaluating the usage of cloud-based HPC resources
- Problem: OOD architecture based on traditional HPC resource managers and intra-node communication doesn't work in the cloud
- Solution: Combine the expertise of the OOD and CloudyCluster teams, with CCQ and Slurm running on the Google Cloud Platform
- Outcomes: A demo of OOD running on GCP and supporting an example HPC client workflow..

University at Buffalo Center for Computational Research

Ohio Supercomputer Center

VIRGINIA TECH

openondemand.org

Open OnDemand & CloudyCluster

Extended Outcomes:

- **Open OnDemand - CloudyCluster Full Integration**: Because of this collaboration OOD is now an integral part of CloudyCluster and automatically launches when you bring up CloudyCluster
- **Expanded Coverage**: This full integration is now across Google Cloud and AWS
- **Expanded Use Cases**: Biomedical Startups

Open Discussion

Cloudify Gateways from the SGCI & Google

Open Discussion Starter Questions

- **How can you see other areas that would benefit from similar collaborations?**
- **Are there examples of similar collaborations that you have seen?**
- **Are there ways to help foster these?**

Next Phase of Collaboration

Project Eureka - Unify Data Mgt & Computation

EUREKA

see the show

Eureka Goals

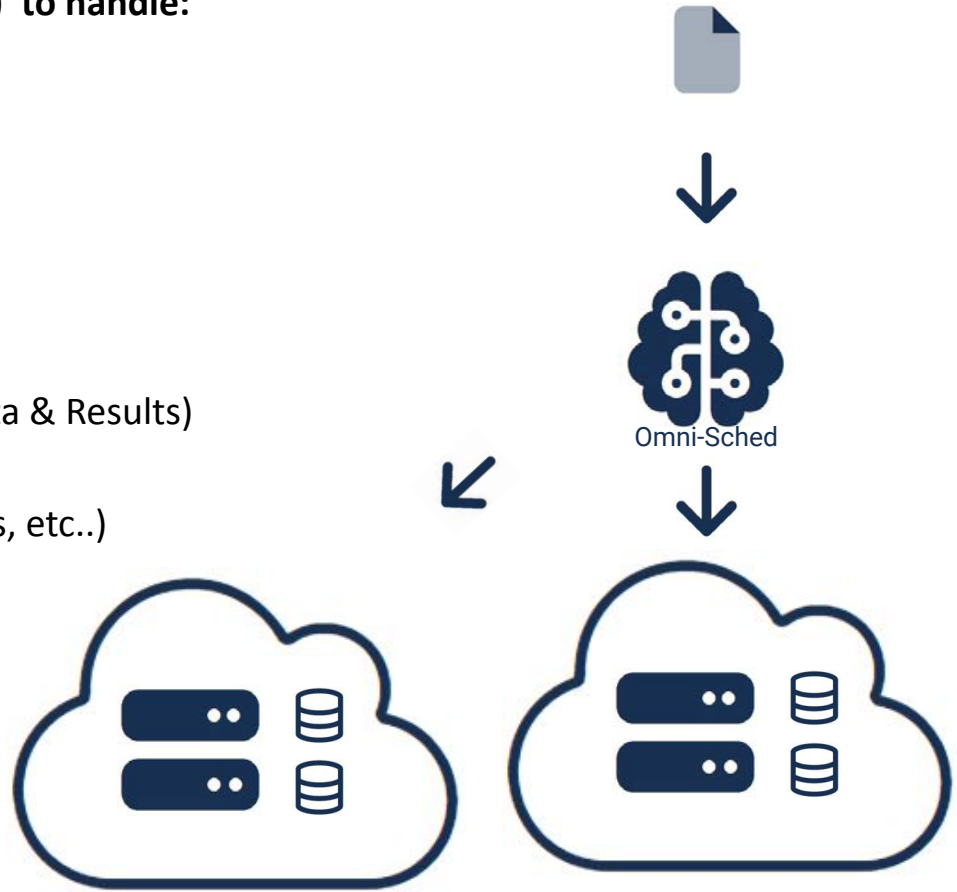
- Data Science, HPC, and HPC Adjacent
  - K8s, spark, etc…
  - Data Science Apps
  - Job Routing
- Storage
  - Staging
  - Dynamic scratch
  - Publish
- Project based UI in Open OnDemand
- Bridge to on-prem
- Easier for end users (cloud stuff still gets in the way for researchers)
- More modular
- Leverage more -> Provide more

The Omni-Scheduler Approach

Expand meta-scheduler concept (OmniSched) to handle:

- Cross Cloud - leverage TerraForm
- Enable On-Prem (TF -> kvm)
- Job routing between locations
- Dynamic Scratch
  - Driven by job Directives
- Data Staging
  - Driven by Job Directives (Input Data & Results)
- Beyond HPC
  - HPC adjacent (publishing, websites, etc..)
  - Enable Apache Spark
  - K8s
  - Etc…
- Scheduling Data Movements

With iRODS integration with OmniSched



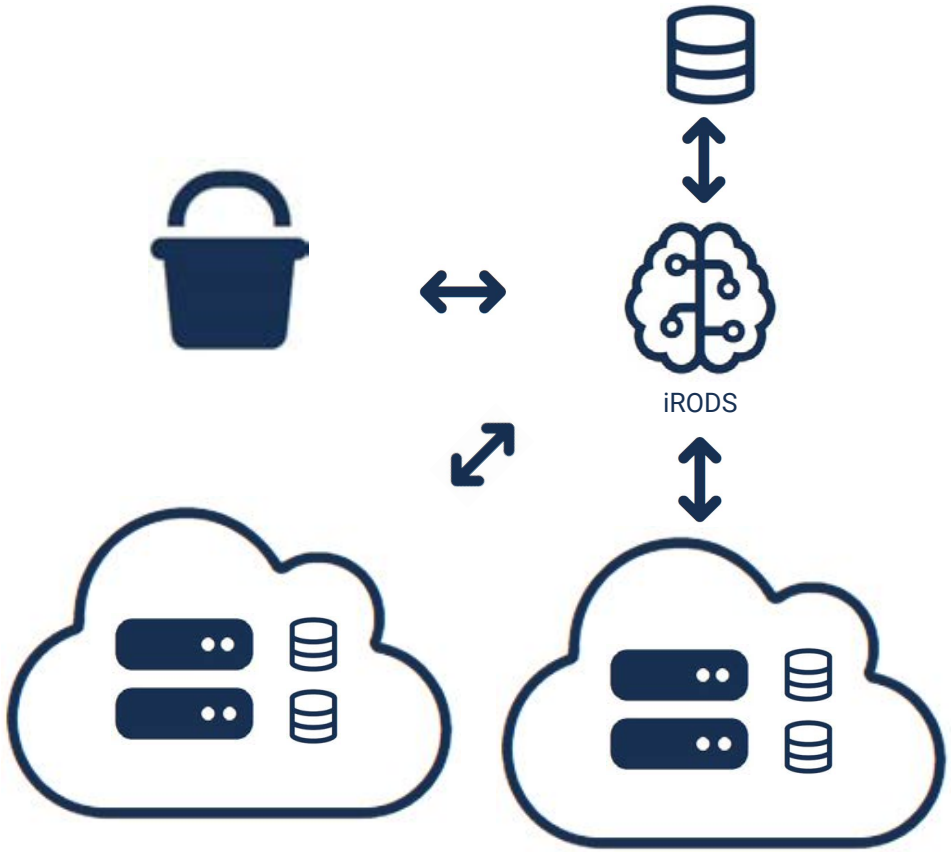
The Integrated Data Management

In addition to Job based directives

Data Management Capabilities

- Replicate
- Archive
- Ingest
- Tier
- Publish

All configured in OOD with iRODS



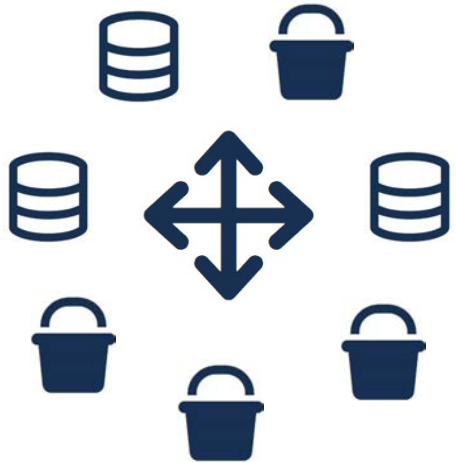
Design Thinking UI

Project

- **Data Science (and other) App Launcher**
- **Research Project Based Collaboration**

Data

- **Cross System File Manager in OOD using iRODS**
- **Storage Resource Integration configured in OOD with iRODS**
- **Data Action Scheduling, configured in OOD with iRODS**



Demo