

65th HPC User Forum, April 2017 – Santa Fe, NM

# Deploying a Software Ecosystem for Science and Engineering Applications at a Regional HPC Center



Karen Tomko

Director of Research Software Applications

Ohio Supercomputer Center

ktomko@osc.edu

# Agenda

The image shows the exterior of the Ohio Supercomputer Center building. The building has a prominent glass facade with a grid pattern, reflecting the sky and trees. The structure is primarily brick with dark accents. A large tree with white blossoms is visible on the right side of the frame. The sky is clear and blue.

- **Overview of Ohio Supercomputer Center**
- Software resources and services
- Looking forward

# About OSC

- 1987: Ohio Legislators Establish OSC
- Statewide resource for all universities in Ohio
  - high performance computing services
  - computational science expertise
  - “ ... propel Ohio's research universities and private industry to the forefront of computational based research.”
- Funded by and reporting line through the Ohio Department of Higher Education
- Fiscal agent is Ohio State University, offices on OSU's west campus

# Client Services CY2016



**26 Ohio-based  
universities**



**41 companies**



**1,358 clients**



**218 trainees**



**8 training  
opportunities**



**218 awards  
made**



**533 projects  
served**



**31 academic  
courses  
used OSC's  
supercomputers**

# Production Capacity CY2016



**115,000,000+**  
**core-hours**  
**consumed**



**3,400,000+**  
**computational**  
**jobs**

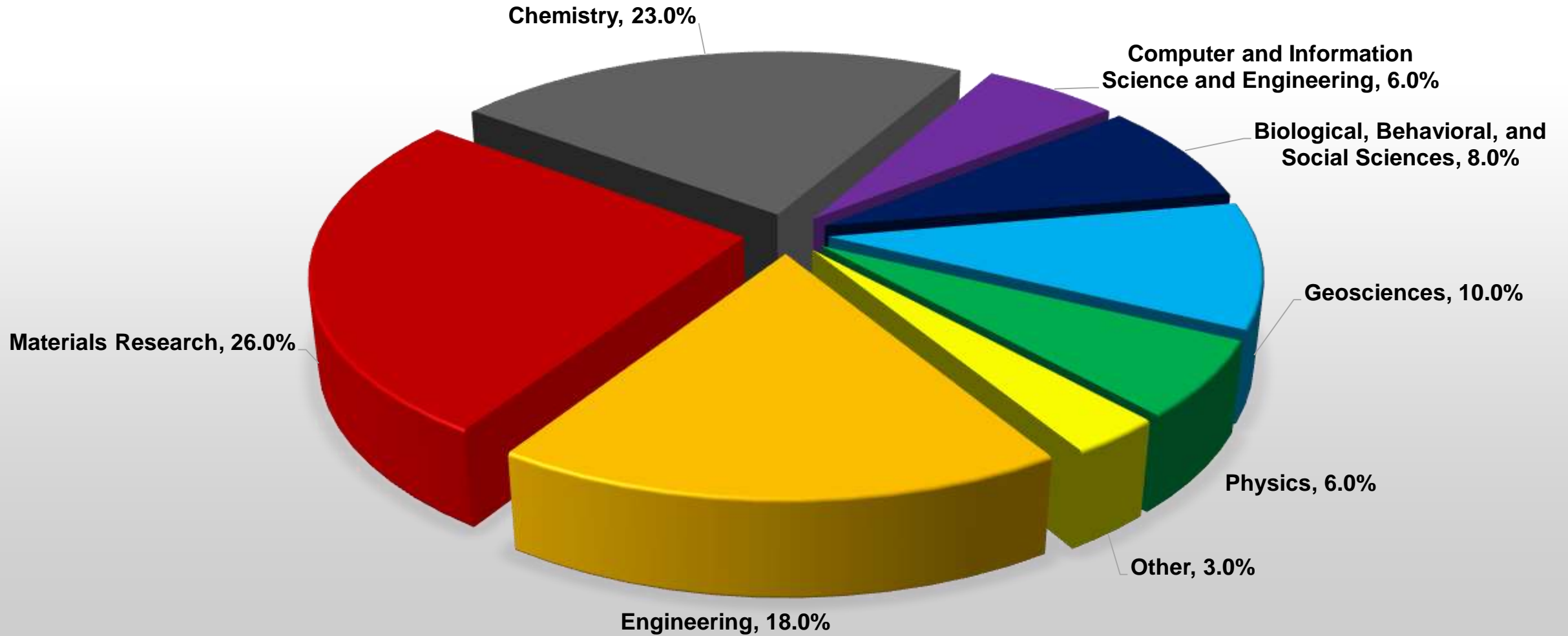


**1,101 TB**  
**data stored**



**97.9% up-time**  
**(target: 96%)**

# Computing Usage by Field of Science CY2016



# OSC Supercomputer and Storage Services



#202 on the Top 500  
w/o GPUs

	Owens (2016)	Ruby (2014)	Oakley (2012)
Theoretical Performance (TF)	~1600	~144	~154
# Nodes	824	240	692
# CPU Cores	23,392	4,800	8,304
Total Memory (TB)	~120	~15.3	~33.4
Memory per Core (GB)	4.5	3.2	4
Interconnect Fabric (IB)	EDR	FDR/EN	QDR

	Capacity (PB)	Bandwidth (GB/s)
Home Storage	0.8	10
Project Storage	3.4	40
Scratch Storage	1.1	100
Tape Library (backup & archive)	5+	3.5

# OSC Service Catalog

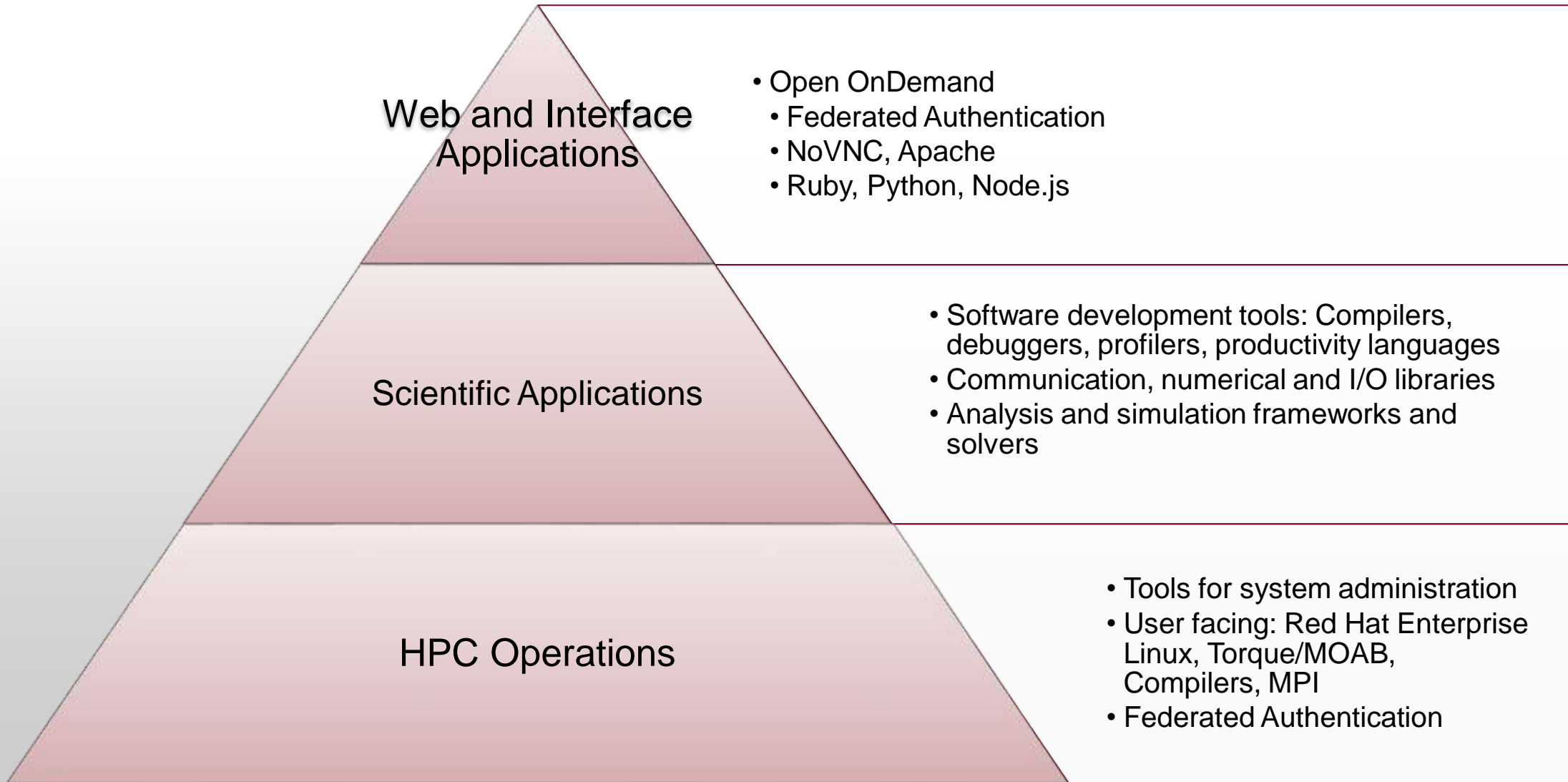
<b>Cluster Computing</b> <ul style="list-style-type: none"><li>• High Performance Computing</li><li>• High Throughput Computing</li><li>• Data-intensive Computing</li></ul>	<b>Research Data Storage</b> <ul style="list-style-type: none"><li>• Project Storage</li><li>• Archival Storage</li></ul>
<b>Client Services</b> <ul style="list-style-type: none"><li>• 24x7 Call Center</li><li>• Level 2 Engineering Support</li></ul>	<b>Client Facilitation</b> <ul style="list-style-type: none"><li>• Consultation (in-person and online)</li><li>• Training and Education</li><li>• Classroom accounts</li></ul>
<b>Scientific Software Development</b> <ul style="list-style-type: none"><li>• Software Development</li><li>• Software Parallelization</li></ul>	<b>Web Software Development</b> <ul style="list-style-type: none"><li>• Software Development</li><li>• Software Consulting</li></ul>
<b>Partner on Proposals</b> <ul style="list-style-type: none"><li>• Cyberinfrastructure solutions</li><li>• Modeling &amp; simulation for industry</li></ul>	<b>Visualization &amp; Virtual Environments</b> <ul style="list-style-type: none"><li>• Visualization Services</li><li>• Virtual environments (DSL)</li></ul>

# Agenda

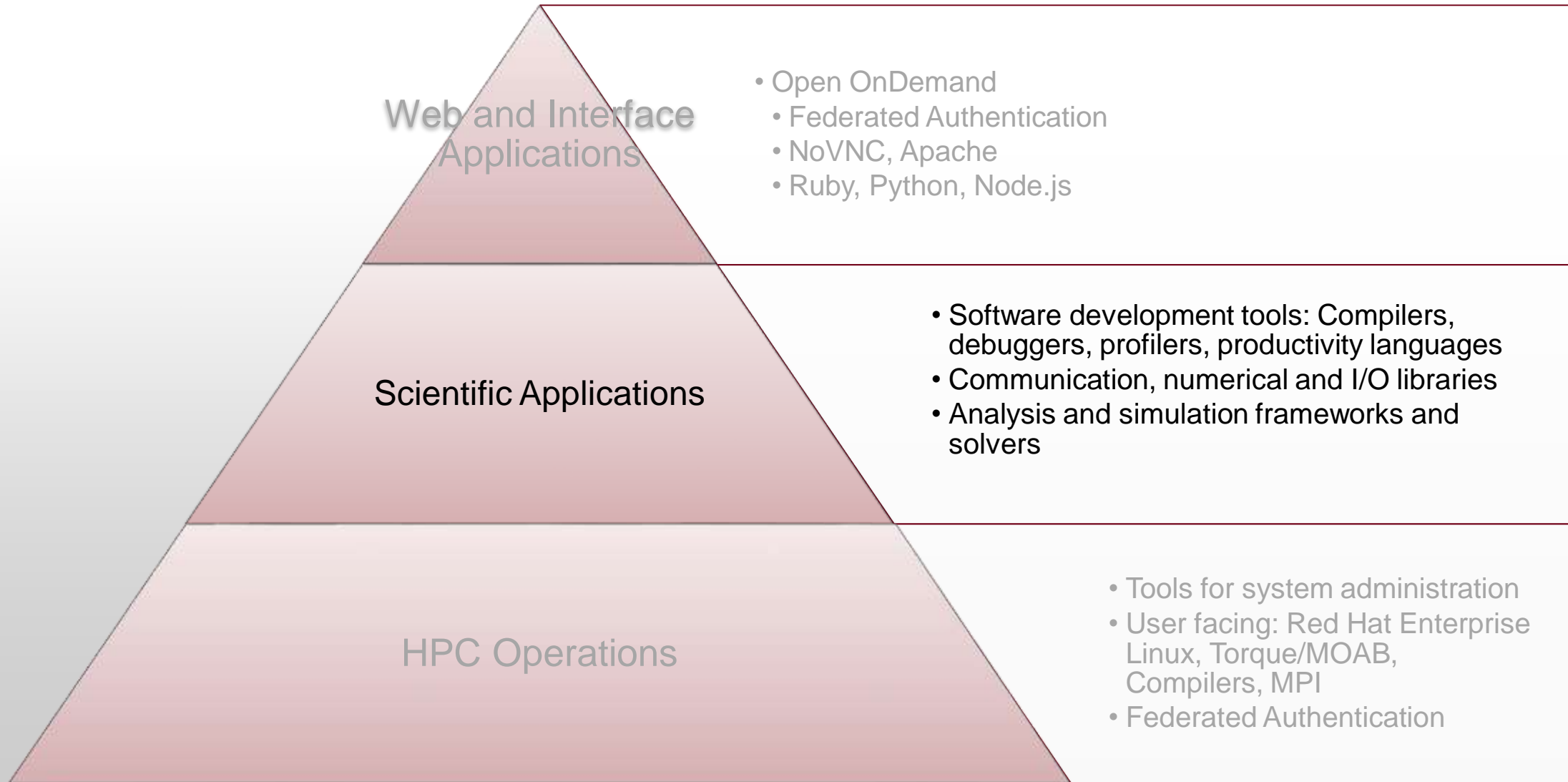
The image shows the exterior of the Ohio Supercomputer Center building. The building has a prominent glass facade with a grid pattern, reflecting the sky and surrounding trees. The structure is primarily brick with dark accents around the windows and roofline. A large tree with white blossoms is visible on the right side of the frame.

- Overview of Ohio Supercomputer Center
- **Software resources and services**
- Looking forward

# Deploying Software at OSC



# Scientific Applications Examples

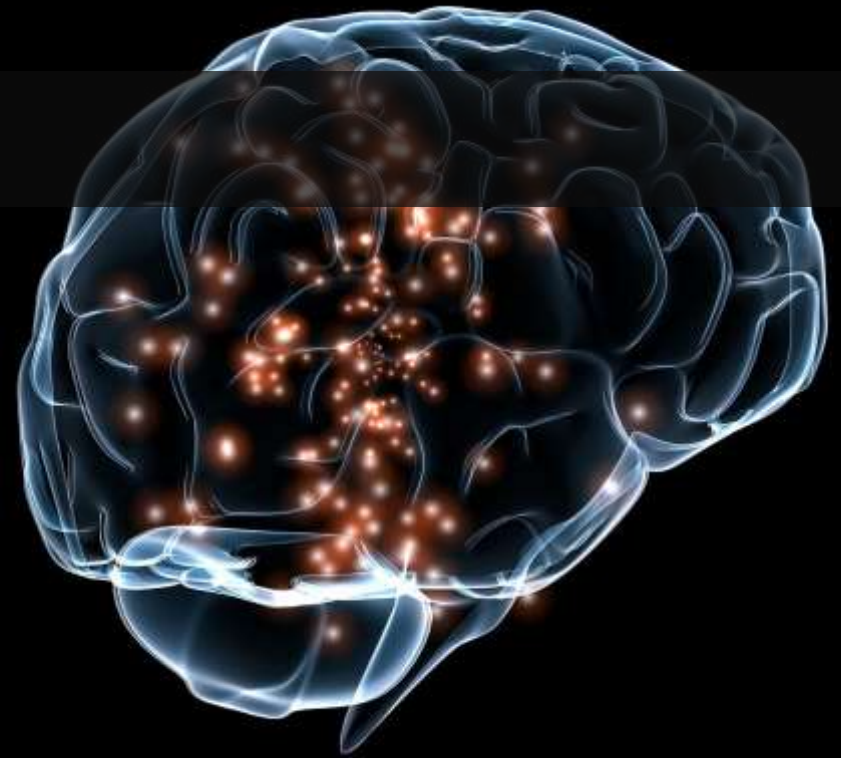


# Software Resources for OSC Clients



- More than 150 software packages,
  - [https://www.osc.edu/resources/available\\_software](https://www.osc.edu/resources/available_software)
- Types of Software include compilers, development tools, numerical libraries, applications
- Mix of ISV, free and open source packages
- Agreements with ISVs are often challenging given our multi-institution and mixed academic/non-academic client base

# Software that leverages GPU



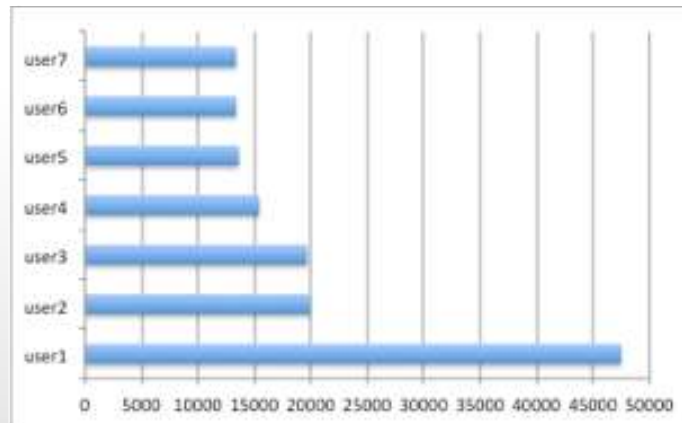
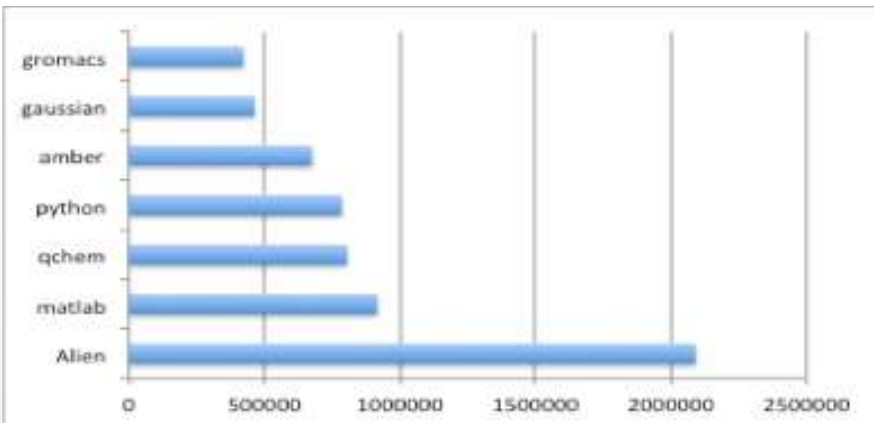
- Molecular Dynamics (MD) Simulations
  - 3X - 7X faster than CPU
  - Materials Science, Biochemistry, Chemistry, Biophysics
  - Software: NAMD, LAMMPS, AMBER, GROMACS
- Machine Learning/Deep Learning
  - 4X - 10x faster for “training” than CPU
  - Wide range of disciplines
  - Software : Caffe, TensorFlow, Torch

Statistics	MYSQL	SPARK
------------	-------	-------

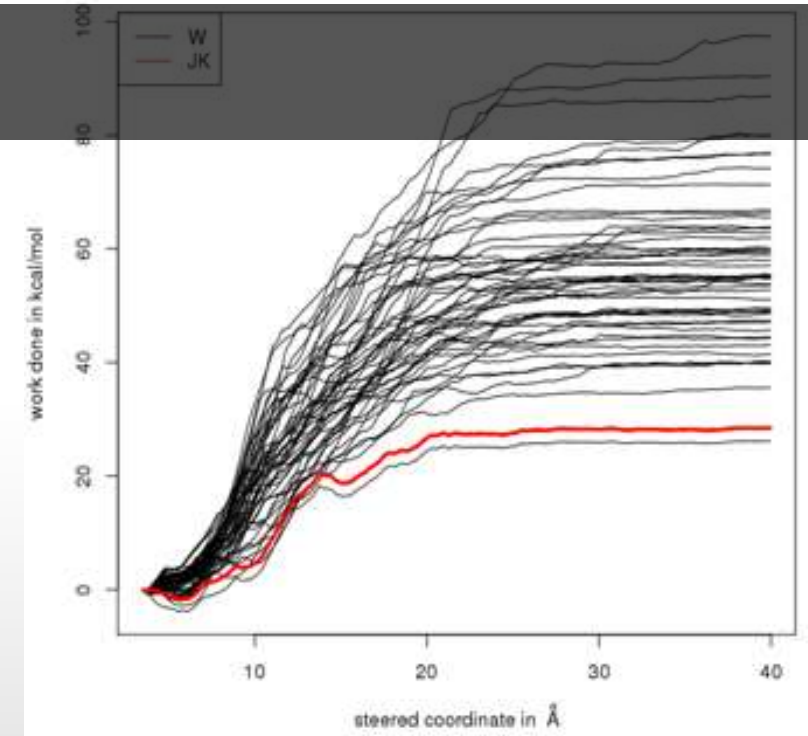
# Data Analytics Use Cases and Software

CPU vs Account	1.25 hour	5 sec
----------------	-----------	-------

Walltime vs user	1.40 hour	5 sec
------------------	-----------	-------

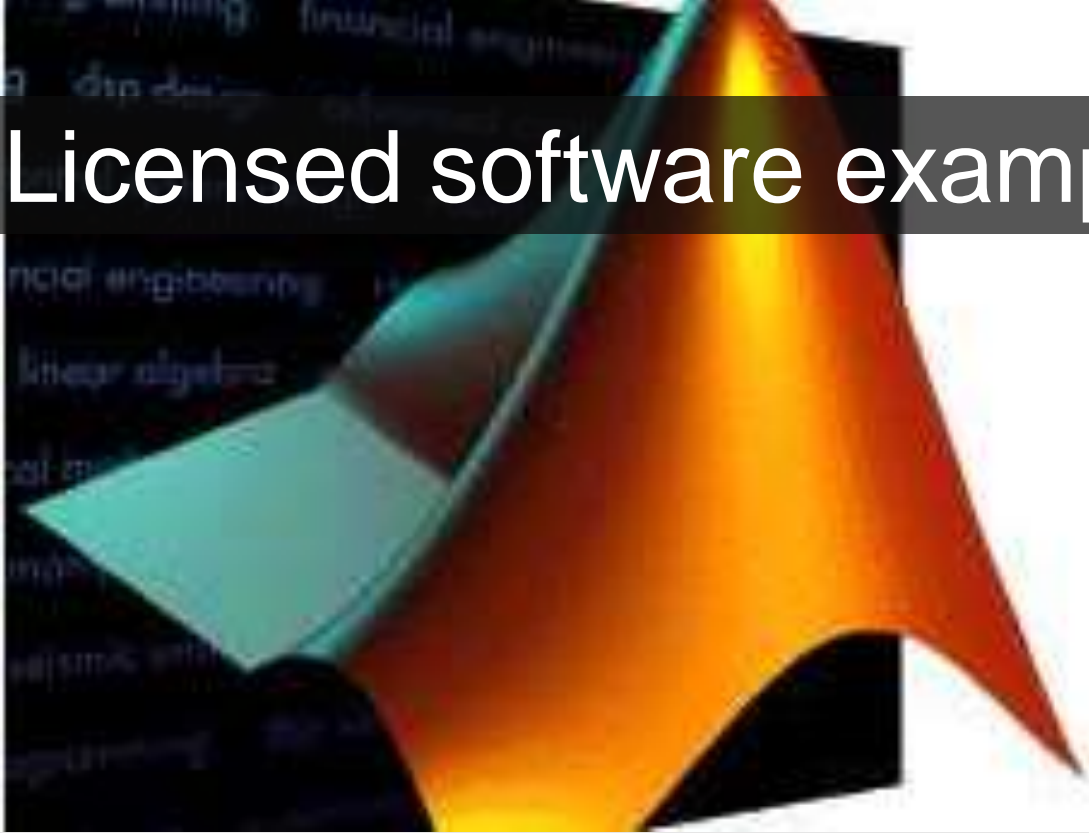


2-PAM in  $S_p$  SARIN inhibited AChE,  $k=100\text{kcal/mol}/\text{\AA}^2$   $v=0.004\text{\AA}/\text{ps}$



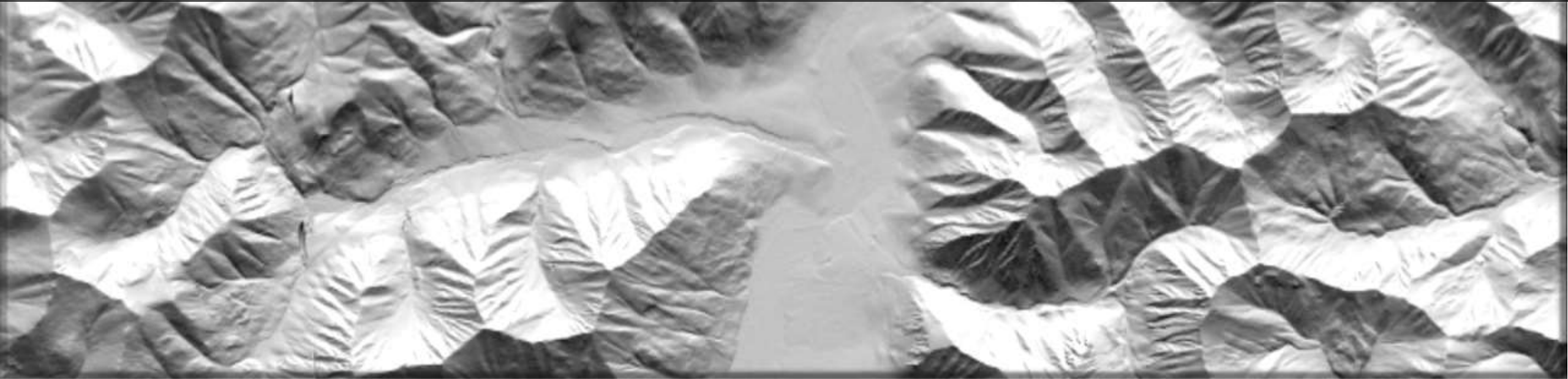
- Analytics on OSC Job data
  - Complex queries on historical job data
  - More than 700x faster than MYSQL query of same data
  - Software: Apache SPARK, PySpark
- Analysis of Simulation results
  - Large data sets from suite of simulation runs
  - Biochemistry/Bioinformatics
  - Software : VMD, R
- Other services: Hadoop, Statistical and mathematical software, high performance storage

# Licensed software example: Matlab Access at OSC



- OSC included in OSU site license
- History: non-OSU clients using local licenses from campus servers when running at OSC
- New terms allows for academic use of MATLAB on OSC systems by any OSC clients from educational institutions
- Additionally, OSC has licenses for MATLAB DCS for multi-node parallel MATLAB

# Software Collaboration Example: SETSM Project



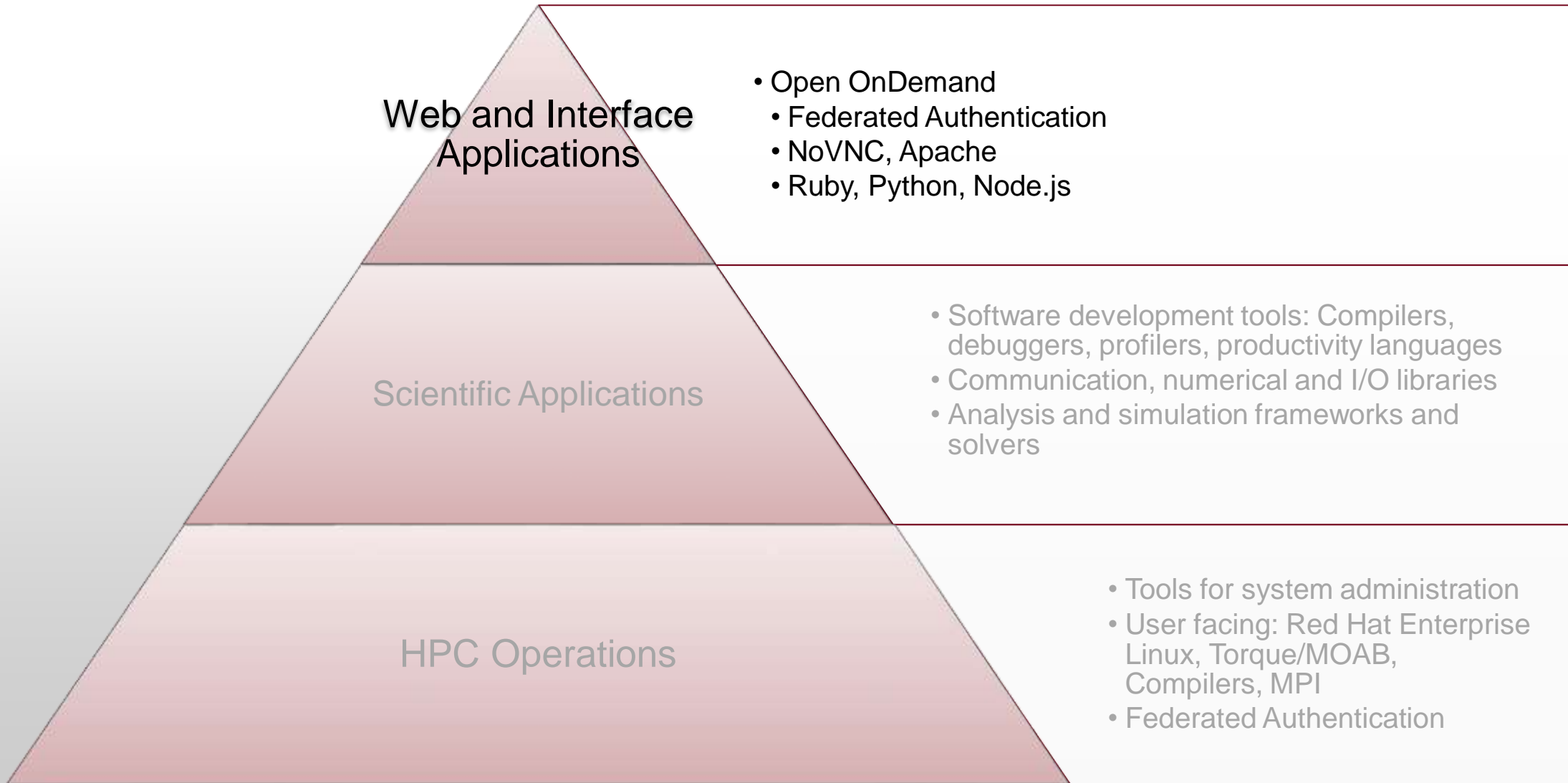
NSF funded “Automated, High Resolution Terrain Generation for XSEDE”

**Howat**, Noh – OSU, Tomko – OSC

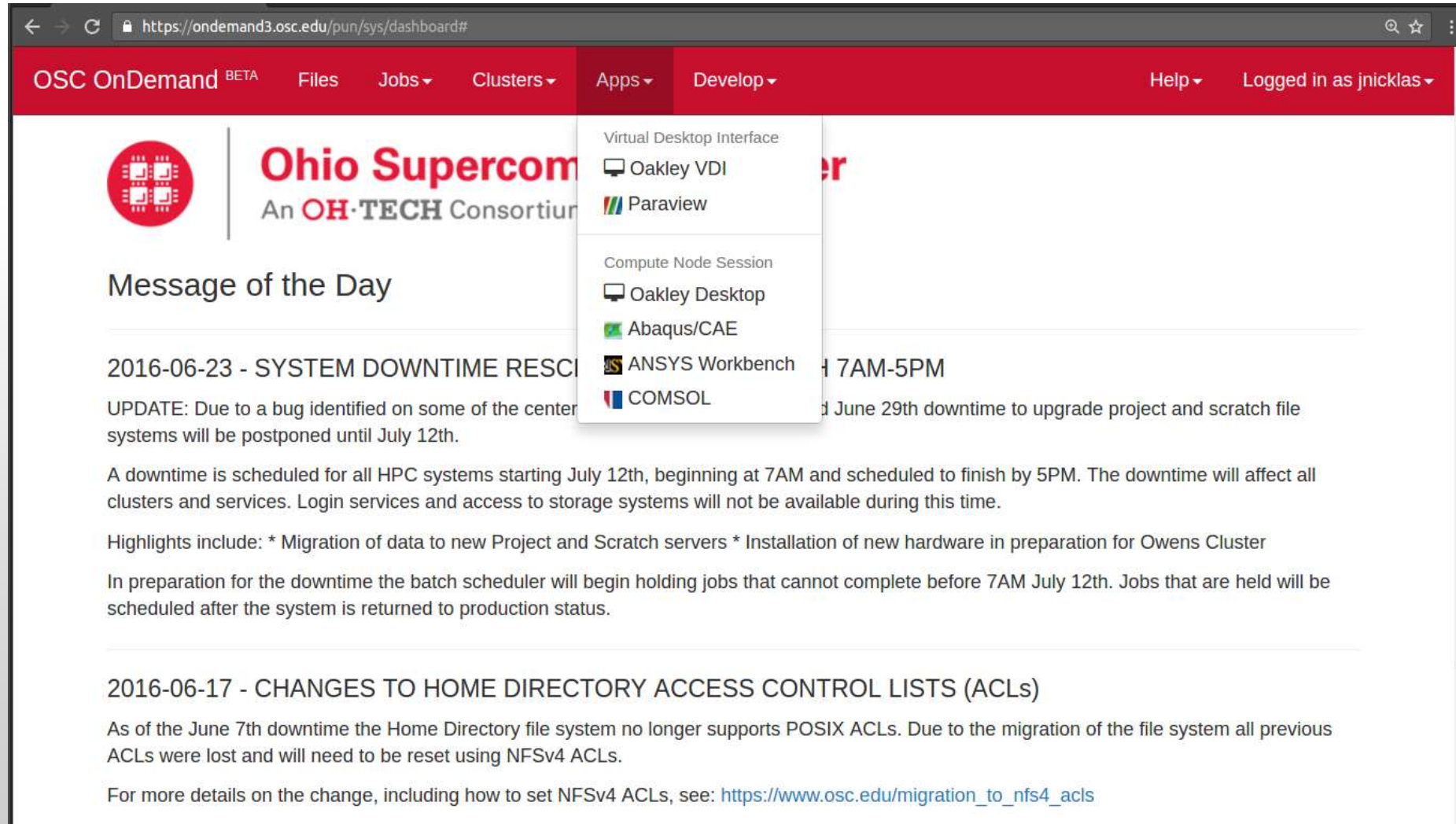
Funding Source: National Science Foundation (NSF)

- Establish a service for on-demand polar Digital Elevation Model (DEM) production and distribution utilizing the XSEDE High Performance Computing framework and the NSF-funded Polar Geospatial Center (PGC) data services
- OSC is providing software engineering, code optimization and parallel software development for the project

# Web and Interface Applications Examples




# Improving Client Access with Open OnDemand



The screenshot shows the OSC OnDemand dashboard interface. The top navigation bar is red and contains the following items: 'OSC OnDemand BETA', 'Files', 'Jobs', 'Clusters', 'Apps', 'Develop', 'Help', and 'Logged in as jnicklas'. The 'Apps' dropdown menu is open, displaying a list of applications: 'Virtual Desktop Interface', 'Oakley VDI', 'Paraview', 'Compute Node Session', 'Oakley Desktop', 'Abaqus/CAE', 'ANSYS Workbench', and 'COMSOL'. The main content area features the Ohio Supercomputer Center logo and the text 'Ohio Supercomputer Center An OH-TECH Consortium'. Below this is a 'Message of the Day' section with two entries: one dated 2016-06-23 regarding system downtime and another dated 2016-06-17 regarding changes to home directory ACLs.

OSC OnDemand <sup>BETA</sup> Files Jobs Clusters Apps Develop Help Logged in as jnicklas

 **Ohio Supercomputer Center**  
An OH-TECH Consortium

### Message of the Day

**2016-06-23 - SYSTEM DOWNTIME RESCHEDULED** 7AM-5PM

UPDATE: Due to a bug identified on some of the center's storage systems will be postponed until July 12th.

A downtime is scheduled for all HPC systems starting July 12th, beginning at 7AM and scheduled to finish by 5PM. The downtime will affect all clusters and services. Login services and access to storage systems will not be available during this time.

Highlights include: \* Migration of data to new Project and Scratch servers \* Installation of new hardware in preparation for Owens Cluster

In preparation for the downtime the batch scheduler will begin holding jobs that cannot complete before 7AM July 12th. Jobs that are held will be scheduled after the system is returned to production status.

**2016-06-17 - CHANGES TO HOME DIRECTORY ACCESS CONTROL LISTS (ACLs)**

As of the June 7th downtime the Home Directory file system no longer supports POSIX ACLs. Due to the migration of the file system all previous ACLs were lost and will need to be reset using NFSv4 ACLs.

For more details on the change, including how to set NFSv4 ACLs, see: [https://www.osc.edu/migration\\_to\\_nfs4\\_acls](https://www.osc.edu/migration_to_nfs4_acls)

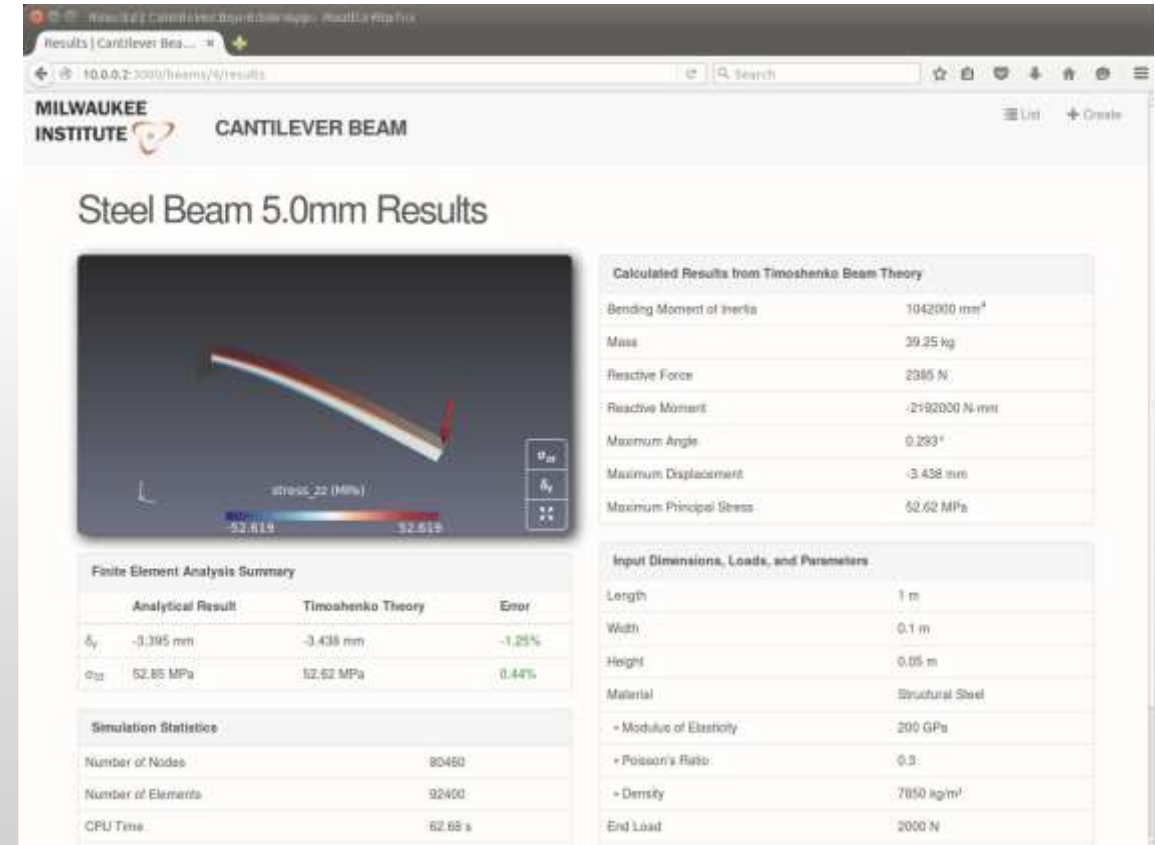
Visit our website: <https://osc.github.io/Open-OnDemand/>

# What can clients do with Open OnDemand?

- Supercomputer access
  - File Browser, File Editor, Terminal
  - Job Constructor, Job Status, System Status
- Virtual Desktop Interface (VDI)
- Interactive HPC (iHPC)
- Web apps (AweSim AppKit)
  - Open OnDemand apps (Jobs, VDI, iHPC)
  - Workflow apps (Cantilever Beam example)

# AppKit Example: Beam

- Classic cantilever beam problem
  - User specifies geometry, materials and load
  - App calculates displacement and stress
- 100% open-source back end
  - gmsh for geometry and mesh generation
  - Elmer for the analysis
  - Paraview to generate a WebGL file for results visualization



# Example Non-academic Client: Testing



Engineers and non-expert designers use OSC simulation apps to run fully automated, on-demand and accurate crush tests of new bucket designs inexpensively and quickly

Dan Meyer, Comet Solutions

# Agenda

The image shows the exterior of the Ohio Supercomputer Center building. The building has a prominent glass facade with a grid pattern, reflecting the sky and trees. The structure is made of brick and has a modern architectural style. A large tree with white blossoms is visible on the right side of the building. The sky is clear and blue.

- Overview of Ohio Supercomputer Center
- Software resources and services
- **Looking forward**

# New Technology Example: DDN Infinite Memory Engine (IME)

- “Burst Buffer” for scratch file system
- NVMe SSD based storage, acts as write-back/read cache, or temporary storage
- Test and document use cases for client rollout
  - POSIX interface, /ime/scratch instead of /fs/scratch
  - Native API
  - MPI-IO (NetCDF, HDF5, etc)
- Still in testing, friendly user availability soon



# Interactive Applications via Web Browser

Demos in package 'base':

`error.catching` More examples on catching and handling errors  
`is.things` Explore some properties of R objects and `is.FOO()` functions. Not for newbies!  
`recursion` Using recursion for adaptive integration  
`scoping` An illustration of lexical scoping.

Demos in package 'graphics':

`Hershey` Tables of the characters in the Hershey vector fonts  
`Japanese` Tables of the Japanese characters in the Hershey vector fonts  
`graphics` A show of some of R's graphics capabilities  
`image` The image-like graphics builtins of R  
`persp` Extended `persp()` examples  
`plotmath` Examples of the use of mathematics annotation

Demos in package 'grDevices':

`colors` A show of R's predefined colors()  
`hclColors` Exploration of `hcl()` space

Demos in package 'stats':

`glm.vr` Some `glm()` examples from V&R with several predictors  
`lm.glm` Some linear and generalized linear modelling examples from 'An Introduction to Statistical Modelling' by Annette Dobson  
`nlm` Nonlinear least-squares using `nlm()`  
`smooth` 'Visualize' steps in Tukey's smoothers

The screenshot shows the RStudio web interface. The top navigation bar includes 'File', 'Plots', 'Packages', 'Help', and 'Viewer'. The 'Environment' pane is active, displaying a list of objects in the Global Environment:

Object	Type	Value
opar	List of 1	
pie.sales	Named num [1:6]	0.12 0.3 0.26 0.16 0.04 0.12
pin	num [1:2]	7.72 4.38
scale	num	0.00729791666666667
Soft	Factor w/ 3 levels	"Hard", "Medium", ...: 1 1 1 1 1 1 1 1 2 2 ...
Temp	Factor w/ 2 levels	"Low", "High": 1 1 2 2 1 1 2 2 1 1 ...
usr	num [1:4]	-131 1011 -14 634
x	num [1:27]	-12.57 -11.6 -10.63 -9.67 -8.7 ...
x.at	num [1:8]	100 200 300 400 500 600 700 800
xadd	num	98.8038823865256
xdelta	num	860
xscale	num	0.00897480620155039
xx	int [1:202]	0 1 2 3 4 5 6 7 8 9 ...
y	num [1:27]	-12.57 -11.6 -10.63 -9.67 -8.7 ...
y.at	num [1:6]	100 200 300 400 500 600

- New capability not available at other supercomputer centers
- Accessible via a web browser with a few clicks through OSC OnDemand
- High performance computing live via a dedicated HPC node(s) (vs. local laptop)
- Currently in Beta testing: Rstudio, Jupyter Notebook for python, MATLAB

# Client-Centric View

