OSC Fall Update

- Budget Trends
- SOCC Charges
- Business Model Proposal
- Discussion
Budget Trends – Revenue Analysis

OSC Revenue Categories:

- State Subsidy
- Sponsored Research Funding
- User Fees
  - Commercial Usage
  - Condo (Dedicated) Usage

2018 Breakdown

- State Line Item Appropriations: 10%
- Sponsored Research Funding: 79%
- User Fees: 11%

Revenue Trend

2013 2014 2015 2016 2017 2018 (Budget)
OSC Expense Categories:
- Direct Costs
- Indirect Costs
- Personnel

2018 Breakdown

- Personnel: 54%
- Space Rent and Power: 11%
- Equipment Maint. and Purchased Services: 3%
- Equipment: 2%
- Other Direct Expenditures: 5%
- Overhead / Indirect F&A: 2%
- Cost Distributions and Transfers: 0%

Expense Trend

- 2013
- 2014
- 2015
- 2016
- 2017
- 2018 (Budget)
Budget Trends – Overall

Projected funding gap is a result of:
- Decreasing Funding
- Increasing Expenses
Data center costs fully transferred to OSC by 2017
- Permanent, substantial addition to OSC’s expenses
- From 2008 to 2018, OSC lost nearly $1.4M for core operations

Impact of SOCC Charges

State Subsidy Trend

- Millions
- Core Operations
- Space and Power

10% Decrease in FY18
OSC Business Model Proposal

• OSC is an initiative of the Ohio Department of Higher Education (ODHE)

• Decision was made to ask universities for support
  – Need a single policy to apply to all Ohio universities

• ODHE has experience with fee models for services (e.g. OARnet, OhioLink), wants to do the same thing with OSC
  – Faculty chargeback policies up to each university
OSC Business Model Proposal Features

• Maintain subsidized access model
  – Subject to peer review by OSC Allocations Committee
  – Fully subsidize a portion of ALL awards

• Charge fees based on compute (RU) usage
  – Options for institutions to pay centrally to lower per-RU fees
  – Establish predictable costs over multiple years

• Create an advisory model similar to OARnet
  – User/Institutional-based Finance Committee
  – Each fall, Committee will develop pricing for coming fiscal year

We are doing everything we can to constrain costs and will continue to offer heavily subsidized services that are well below market rates.
Step 1: Project resource usage

Available RUs (Millions)

2013 2014 2015 2016 2017

10
20
30
40

Step 2: Quantify recovery target

Expenses

Revenue

Target

Step 3: Subsidize 1st 10K RU for all clients

Usage

10% 90%

90% 10%

Step 4: Divide target by usage

<table>
<thead>
<tr>
<th>RUs</th>
<th>Unit Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10,000</td>
<td>No Cost</td>
</tr>
<tr>
<td>&gt;10,000</td>
<td>$0.17/RU</td>
</tr>
</tbody>
</table>

Step 5: Incentivize central payment

<table>
<thead>
<tr>
<th>Payment</th>
<th>Discount</th>
<th>Unit Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10,000</td>
<td>20%</td>
<td>$0.136/RU</td>
</tr>
<tr>
<td>$30,000</td>
<td>40%</td>
<td>$0.102/RU</td>
</tr>
<tr>
<td>$60,000</td>
<td>60%</td>
<td>$0.068/RU</td>
</tr>
<tr>
<td>$350,000</td>
<td>80%</td>
<td>$0.034/RU</td>
</tr>
</tbody>
</table>

Step 6: Calculate cost/university

Institutional Costs ($Ks)

DRAFT COSTS
Local Cluster

- 1 Owens node = ~22K RUs / year
- 100K RUs = ~5 nodes
- Typical node upfront cost = ~$5,500
- Typical node annual operating cost = ~$1,900
- 5 node cluster cost = ~$28K upfront, ~$10K/year
- 5 node cluster five-year lifetime cost = $78K
- Additional costs for:
  - system administration services
  - software licenses
  - data storage and backup services

Amazon Cloud Services

- 1 Owens node =~ Amazon r4.8xlarge instance
- 100K RUs =~5 instances
- Typical instance upfront annual cost = ~$12K
- 5 instance annual cost (paid upfront) = ~$60K
- 5 instance five-year lifetime cost = ~$300K
- Additional costs for:
  - data transfer
  - system administration services
  - software licenses
  - data storage and backup services

These costs are hard to quantify generically since each client has specific unique needs, although they are covered by OSC today.
Comparison to OSC for Example 100K RUs / Year Client

<table>
<thead>
<tr>
<th>Institution</th>
<th>Annual Central Payment</th>
<th>RU Rate</th>
<th>Annual Cost for 100K RUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local Cluster</td>
<td>$0</td>
<td>$0</td>
<td>$15,600</td>
</tr>
<tr>
<td>Amazon</td>
<td>$0</td>
<td>$0</td>
<td>$60,000</td>
</tr>
<tr>
<td>Ohio State University</td>
<td>$350,000</td>
<td>$0.034</td>
<td>$3,060</td>
</tr>
<tr>
<td>Case Western Reserve University</td>
<td>$60,000</td>
<td>$0.068</td>
<td>$6,120</td>
</tr>
<tr>
<td>Ohio University</td>
<td>$30,000</td>
<td>$0.102</td>
<td>$9,180</td>
</tr>
<tr>
<td>University of Cincinnati</td>
<td>$30,000</td>
<td>$0.102</td>
<td>$9,180</td>
</tr>
<tr>
<td>Bowling Green State University</td>
<td>$10,000</td>
<td>$0.136</td>
<td>$12,240</td>
</tr>
<tr>
<td>All Other Institutions</td>
<td>$0</td>
<td>$0.170</td>
<td>$15,300</td>
</tr>
</tbody>
</table>

Does NOT include:
- System administration
- Software licenses
- Data storage
- Backup services

These are based upon projected institutional usage and lowest overall impact to each institution.
Client Feedback to Date

Representative quotes:

• "I would have to downgrade my research program to fit within the computation that I can afford, which is basically nothing at the rates that you are proposing."

• “Funding rates are plummeting, and the amount per grant is dropping in relative terms…If costs are pushed onto the grants that mainly support the salaries of young talent, the obvious outcome is that the workforce must be reduced.”

Summary of comments:

• Paying for compute from direct grant funds hurts ability to hire students
• Concern over cost containment
• Concern over long-term cost predictability
• Desired flexibility to spend funding when available
• Explore charging for other services (e.g. storage, GPUs, big-data, HIPPA)
Status and Next Steps

• Discussions with provosts from 6 institutions this summer

• DHE and OSC will continue to engage universities to discuss the proposed model and consider adjustments

• OSC and Finance Committee will be looking at budget and usage projections

• Begin charging in FY19 (July 1, 2018) with the goal of reaching $1.5M in funding by FY20
Questions?
Campus Visits

- Bowling Green State University (January)
- University of Cincinnati (April)
- Ohio State University (March, June)
  - Consultation hours at Research Commons every other Tuesday
Upcoming Training and Events

- Introduction to Supercomputing / Big Data at UC: October 10th
- Big Data at OSC: October 26th

See https://www.osc.edu/events
## OSC Supercomputers + Storage

### Systems

<table>
<thead>
<tr>
<th>Systems</th>
<th>Oakley</th>
<th>Ruby</th>
<th>Owens</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>2012</td>
<td>2014</td>
<td>2016</td>
</tr>
<tr>
<td>Cost</td>
<td>$4 million</td>
<td>$1.5 million</td>
<td>$7 million</td>
</tr>
<tr>
<td>Theoretical Performance</td>
<td>~154 TF</td>
<td>~144 TF</td>
<td>~1600 TF</td>
</tr>
<tr>
<td>Nodes</td>
<td>692</td>
<td>240</td>
<td>824</td>
</tr>
<tr>
<td>CPU Cores</td>
<td>8304</td>
<td>4800</td>
<td>23392</td>
</tr>
<tr>
<td>RAM</td>
<td>~33.4 TB</td>
<td>~15.3 TB</td>
<td>~120 TB</td>
</tr>
<tr>
<td>GPUs</td>
<td>128 NVIDIA Tesla M2070</td>
<td>20 NVIDIA Tesla K40</td>
<td>160 NVIDIA Pascal P100</td>
</tr>
</tbody>
</table>

**Total compute: ~1900 TF**

### Storage

<table>
<thead>
<tr>
<th>Storage</th>
<th>Home</th>
<th>Project</th>
<th>Scratch</th>
<th>Tape Library (backup &amp; archive)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>.8 PB</td>
<td>3.4 PB</td>
<td>1.1 PB</td>
<td>5+ PB</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>10 Gbps</td>
<td>40 Gbps</td>
<td>100 Gbps</td>
<td>3.5 Gbps</td>
</tr>
</tbody>
</table>
## Owens Node Configurations “side-by-side” Comparison

<table>
<thead>
<tr>
<th>Node Type</th>
<th>Compute</th>
<th>GPGPU</th>
<th>Data Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node Count</td>
<td>648</td>
<td>160</td>
<td>16</td>
</tr>
<tr>
<td>Core Count</td>
<td>28</td>
<td>28</td>
<td>48</td>
</tr>
<tr>
<td>Core Type</td>
<td>Broadwell</td>
<td>Broadwell</td>
<td>Haswell</td>
</tr>
<tr>
<td>Memory</td>
<td>128 GB</td>
<td>128 GB</td>
<td>1500 GB</td>
</tr>
<tr>
<td>Disk</td>
<td>1 TB</td>
<td>1 TB</td>
<td>20 TB</td>
</tr>
<tr>
<td>GPU</td>
<td>N/A</td>
<td>P100</td>
<td>None</td>
</tr>
</tbody>
</table>
Owens GPU Rollout and Adoption

- 160 Nvidia P100 GPUs were made available to clients in the beginning of April 2017
- Usage:
  - ~50% Molecular dynamics (e.g. Amber, Gromacs, Namd, Lammmps)
  - ~50% Machine learning / neural networks (e.g. Tensorflow, Caffe, Torch)
Client Portal Project

• A user friendly client portal that is a “one-stop” shop for registration, project management, profile updates, etc.

• Status:
  – Internal user testing underway, external client testing beginning shortly
  – Development timeline expected to finish first week of October
  – AweSim website in process of being transferred to internal hosting
New Cluster (C18) Vision / Timeline

• Vision
  • Complement Owens
  • Dense compute component
  • GPU computer component
  • Big data component

• Timeline
  • RFI issued Sept 25
  • RFI due Oct 22
  • RFP issued Nov 2
  • RFP due Dec 15
  • Facilities updates May
  • System delivery June