About RDI²

• RDI² is focused on Advanced Computing and Data Cyberinfrastructure (ACI):
  – Drive innovation in both compute science and computational science
    • New architectures
    • New algorithms, software methodologies/stacks
    • New problem formulations
  – Research collaboration, path-finding and IP generation

• Research instruments with specialized functionality

• Access to RDI² resources
  – Based on state-of-the-art allocation-based policy
  – Complements other modes such as condo model (in development at Rutgers) which provides general-purpose compute cycles
RDI² Research Instruments

- Ocean Observatories Initiative Cyber-infrastructure
  - Funded by NSF, 25 years operation

- Caliburn: Omnipath, NVMe, Lustre
- ELF: Infiniband and GPFS
  - Funded by State of NJ (ELF funds)

- Data Infrastructure Building Blocks (DIBBs)
  - Virtual Data Collaboratory (NJ and Pennsylvania)
  - Funded by NSF

- CAPER: Platform for energy-efficiency research (RDI² members)
- Spring Cloud: Openstack-based platform (RDI² members)
- Other RDI² internal instruments
ELF Early Adopters Program

- Access to ELF has been done ramping up progressively
- Most of early adopters on speed by March 2016

<table>
<thead>
<tr>
<th>Number of Active Projects</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Active Users</td>
<td>134</td>
</tr>
<tr>
<td>Number of departments</td>
<td>20</td>
</tr>
<tr>
<td>Number of Schools</td>
<td>10</td>
</tr>
<tr>
<td>Campus coverage</td>
<td>New Brunswick, Newark, Camden, NJCU</td>
</tr>
</tbody>
</table>
ELF Early Adopters Program – User Distribution

Count of User Department

- RD/2: 12.8%
- Computer Science: 9.6%
- Electrical and Computer Engineering: 7.4%
- Chemical and Biochemical Engineering: 7.4%
- Physics and Astronomy: 5.3%
- Physics: 5.3%
- Pathology and Laboratory Medicine: 5.3%
- Biology: 5.3%
- Biochemistry and Microbiology: 4.3%
- Center for Integrative Proteomics Research: 4.3%
- Research Computing: 3.2%
- Waksman Institute of Microbiology: 3.2%
- Genetics: 3.2%
- Ecology, Evolution and Natural Resources: 3.2%
- Theoretical Chemistry: 3.2%
- other: 19.1%

Count of User School

- ORED: 43.7%
- SAS: 18.4%
- Engineering: 12.8%
- SAS-Camden: 9.7%
- RWJMS: 9.7%
- SEBS: 9.7%
- OIT: 9.7%
- NCAS-Newark: 5.6%
- NJCU: 7.4%
- other: 5.6%

Count of User Campus

- New Brunswick: 85.1%
- Camden: 5.6%
- Newark: 7.4%
- NJCU: 5.6%
ELF Early Adopters Program

- ELF utilization statistics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Jobs Run</td>
<td>58,551</td>
</tr>
<tr>
<td>CPU Hours Used</td>
<td>10,309,800</td>
</tr>
<tr>
<td>Storage Allocated (% of 1PB Total)</td>
<td>~110TB (~11%)</td>
</tr>
<tr>
<td>Average CPU hours per job</td>
<td>176.08</td>
</tr>
<tr>
<td>Average job execution time (wall time)</td>
<td>1.08 hrs</td>
</tr>
<tr>
<td>Overall resource utilization (steady phase)</td>
<td>~95%-98%</td>
</tr>
</tbody>
</table>
ELF Early Adopters Program

- ELF utilization statistics – Service Units (SUs)

**Utilization by Department**
- Physics and Astronomy: 18.2%
- Physics (Camden): 19.5%
- Waksman Institute of Microbiology: 13.3%
- Theoretical Chemistry (Newark): 12.1%
- Chemical and Biochemical Engineering: 9.1%
- Center for Integrative Proteomics Research: 21.9%
- Chemistry and Chemical Biology: 13.3%
- Other: 4%

**Utilization by School**
- SAS: 43.5%
- SAS-Camden: 14.3%
- NCAS-Newark: 14.3%
- SOE: 21.9%
- Other: 12.1%

**Utilization by Campus**
- New Brunswick: 71.4%
- Camden: 14.3%
- Newark: 14.3%
ELF Early Adopters Program

- ELF job size statistics (most jobs are parallel 129-512 cores)
  - There is a large demand for HPC resources

~40% ELF down due to cooling constrain in Hill Center
Transition from ELF Early Adopters to Production

• Reminder: three different access modes

1. General Users
   • Start-up allocations, limited resources
   • Open to Rutgers community

2. Awarded Users
   • Proposals required (2 call for proposals per year)
   • Allocation based on merit
   • Project-based (PI + users associated with the project)
   • Open software, tools and libraries

3. RDI² Partners
   • User support services, software by arrangement
   • Allocated compute resources
   • Allocated data/backup services
Transition from ELF Early Adopters to Production

• Access in 2016

  – Early adopters
    • Extended until end of 2016
    • New requests to elf@rdi2.rutgers.edu

  – Open access (start-up allocations)
    • Ad-hoc requests to elf@rdi2.rutgers.edu
    • Web-based requests being implemented

  – RDI² membership
    • elf@rdi2.rutgers.edu
    • Contact RDI² Director, Manish Parashar
Transition from ELF Early Adopters to Production

• Access in 2017

  – Call for proposals

    • Proposals by Principal Investigator (PI)
    • Call for proposals: 10/24/2016
    • Proposals submission: 11/14/2016
    • Notification of awards: 12/16/2016
      – Committee reviews requests based on merit and need for resources
    • Awarded allocations period: 1/1/2017 – 6/30/2017
Transition to Caliburn

• Phase 1: Starting 10/17/2016
  – Main goal is to avoid disruption/downtime
  – Early adopter users migrated to Caliburn
    • Users will be contacted by RDI² personnel (based on current usage at ELF)
    • Requests for preferred timelines are welcome
  – Data migrated to (temporary) Lustre filesystem (http://lustre.org)
    • Downtime will be needed for final synchronization
  – Approx. 60% of ELF running in current location
    • GPFS will stay but not synchronized with Lustre (can be requested afterwards)
  – Remaining ~40% of ELF being set up in final destination
Transition to Caliburn

- GPFS vs. Lustre
  - Reference using Naïve IOR benchmark (medium size job)
  - GPFS better at write and Lustre better at read
Transition to Caliburn

• Phase 2: Starting 11/07/2016
  – Early adopter users working primarily on Caliburn
  – Lustre filesystem as only filesystem
  – Approx. ~40% of ELF running in final location
    • Temporary filesystem
  – Remaining ~60% of ELF transition
  – GPFS filesystem transition
Transition to Caliburn

- Phase 3: end of 2016
  - Both Caliburn and ELF available in final location
  - GPFS available in final location
Contact

• Information: elf@rdi2.rutgers.edu

Manish Parashar
RDI² Director, Distinguished Professor

Ivan Rodero
RDI2 Associate Director for Technical Operations, Associate Research Professor
Thank you!

http://rdi2.rutgers.edu/