Reconfigurable “Platform” for Transdisciplinary Research

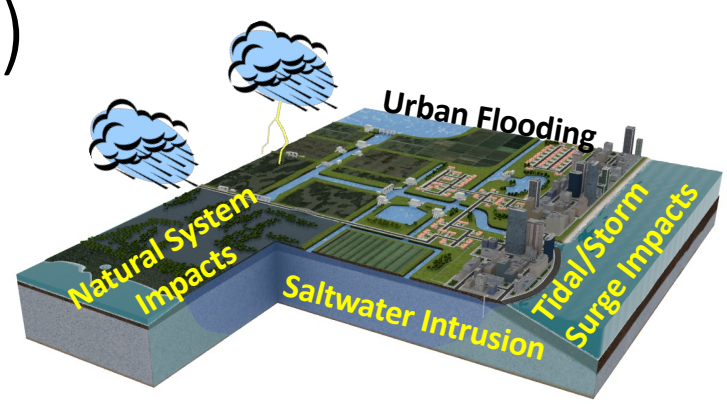
Jason Liu

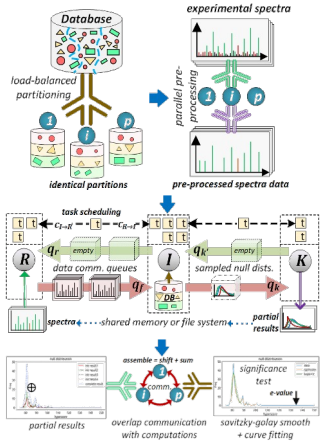
Knight Foundation School of Computing and Information Sciences
Florida International University

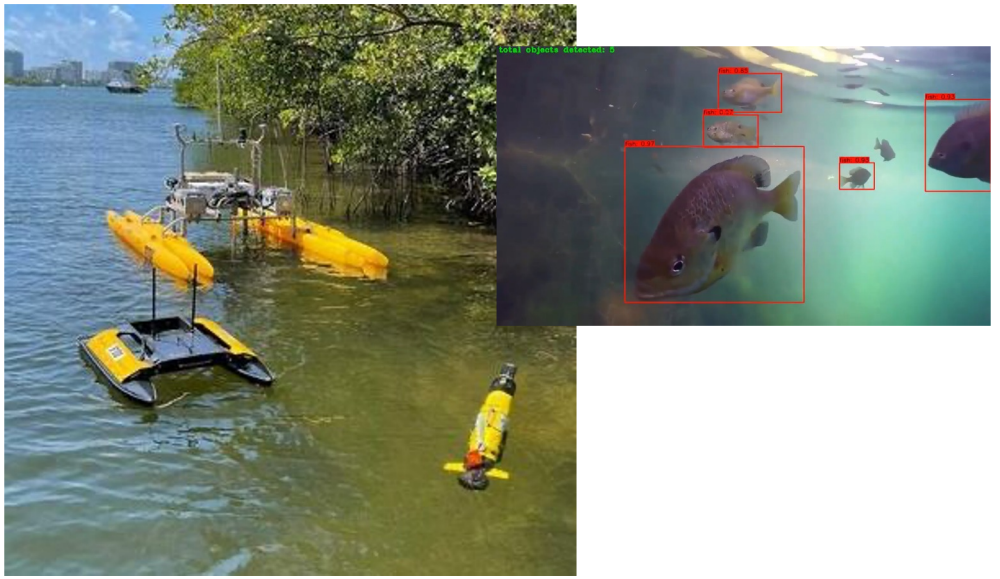
MERIF Workshop, June 2, 2022, Madison, Wisconsin, USA

Example Domain Drivers (at FIU)

- Environmental Science and Engineering
  - Coastal environmental monitoring
  - Freshwater quality analyses and ecotoxicological studies
  - Everglades restoration
- Extreme Events
  - Disaster risk and resiliency analysis
  - Hurricane loss and storm surge modeling
- Computational and Systems Biology
  - Proteomics
  - Genomics
  - Connectomics







Characterizing Diverse (Computing) Demand

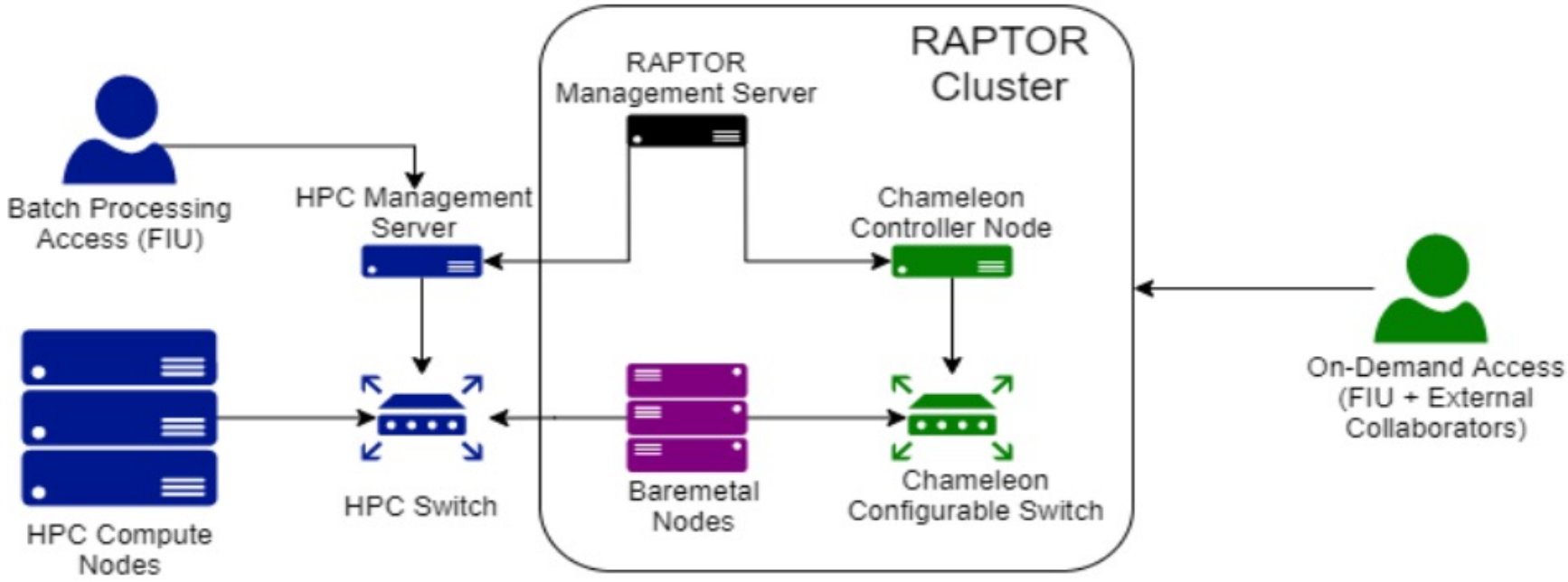
- **High-performance computing**
- **High-throughput computing**
- **Data-intensive computing**
- **Machine learning**
- **Distributed computing**
- **Federated learning**
- **On-demand computing**
- **Cloud computing**
- **Edge computing**
- **…**

- Multiplexing
- Sharing
- Reconfigurable
- Extensible

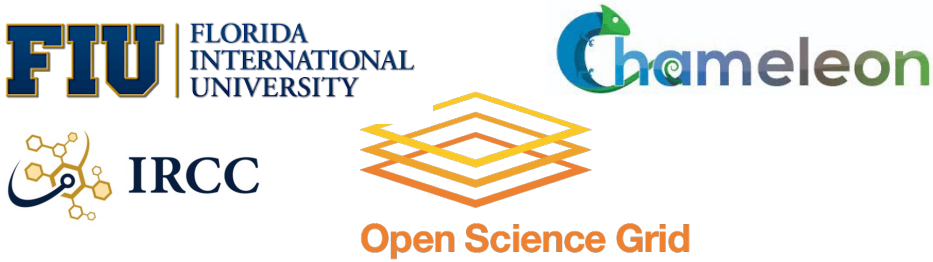
RAPTOR – Reconfigurable Advanced Platform for Transdisciplinary Open Research (a CC* project)

The major goal of RAPTOR is to increase FIU's research production by enhancing its computing capabilities both at the campus level and through participation in a resource-sharing federated distributed computing community. One important aspect of RAPTOR is that it's a reconfigurable platform to address the diverse computing needs of science drivers

RAPTOR Architecture



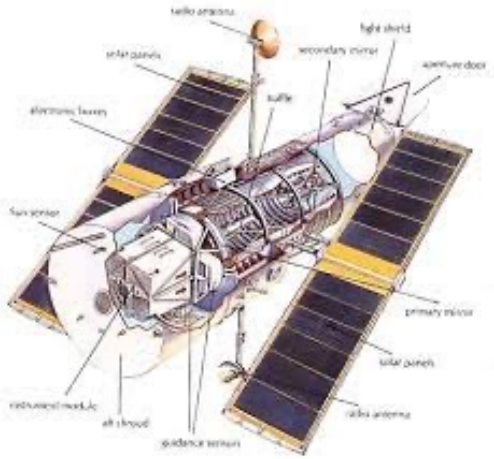
RAPTOR architecture

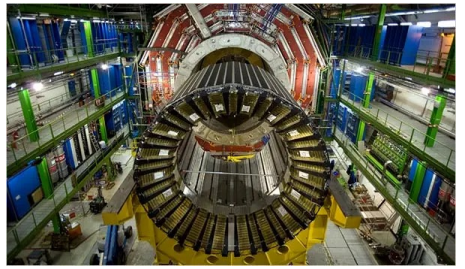


- The RAPTOR architecture will consist of a Chameleon cloud platform and will integrate with the production HPC cluster at FIU
- RAPTOR will allow dynamic allocation of hardware resources between the Chameleon Platform and the HPC cluster as managed and controlled by the RAPTOR Management Server
- The production HPC will also contribute to the Open Science Grid's (OSG) Open Science Pool
- FIU researchers will be able to harness the capacity of the OSPool via OSG-operated access points by going to OSG Connect and getting an OSG Account.
- In addition to providing resources to FIU researchers, RAPTOR will contribute to a national research community through Chameleon and OSG.

Some Observations

- Overloaded terms: instrument, testbed, platform, infrastructure

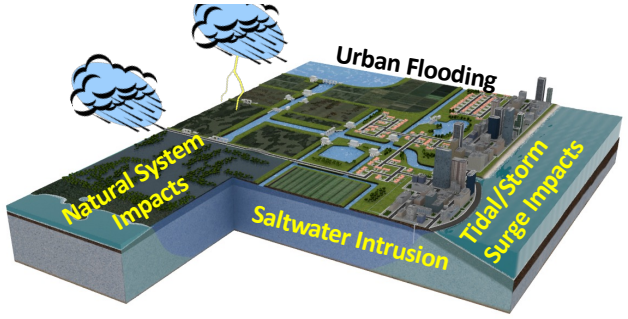




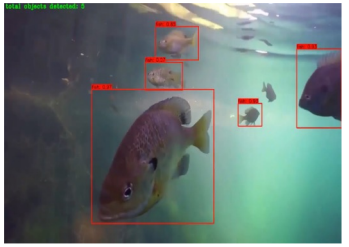
Note: pictures not to scale!

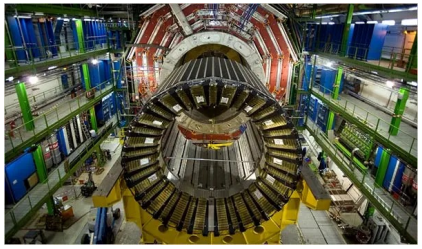
- Telescope is not designed to improve the design of telescopes
- So much so that Internet not designed to improve the design of Internet
- What's the target clientele (beneficiary, stakeholder)?
  - In NSF terms: What's the scientific question? What's community buy-in?

Target Clientele



Astronomers
Astrophysicists

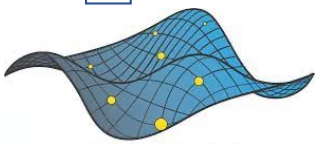




Physicists

(Some) Computer Scientists
Network Researchers

Scientific Computing
HPC Users



Chameleon FABRIC

Current Cloud Computing/Networking Instruments/Testbeds/Platforms/Infrastructures

- Platform centric:
  - "Build the platform; let users come" (in that order!)
- Build the platform:
  - Use state-of-the-art software/hardware building blocks
  - Make a "Swiss army knife" with all built-in functions: support for experiment configurations, user management, security, measurement system, data collection, data curation, and data management, etc.
- Let users come:
  - Democratize access to the platform
  - Bring your own device (BYOD)
- Metrics of success: #users, #experiments, #papers
- **It will not scale (to the level of a telescope or LHC)!**
  - **3 orders of magnitude difference in investment**

To Scale, Need to Be Application Centric

- Build the platform for the specific type(s) of applications, on site
- “Create” your own infrastructure instance
- Plunk it in and make to work: instantiate, customize, combine, compose
- Metrics of success: #platforms, #sites, #applications
  - Make it easy or easier to instantiate, customize, combine, compose
- Case in point: ARA – Wireless living lab for rural communities (from Hongwei Zhang’s presentation yesterday)
  - Deployment of advanced wireless platforms in Central Iowa
  - Adoption of open-source software platforms for living lab management and experimentation: OpenStack, CHI-in-a-Box, ONF (SD-RAN, SD-CORE, ONOS), srsRAN, OpenAirInterface, etc.
- If one can replicate this a thousand times, we have our “telescope” project

Which approach to take for next mid-scale or even billion $$ infrastructure?

Tree

VS.

Tree Nursery