Funded by the National Science Foundation, Grant #2320373

CyberInfrastructure Training and Education for Synchrotron X-Ray Science (X-CITE)

CyberInfrastructure at CHESS and Beyond

Anirban Mandal, Erik Scott, Sajith Sasidharan (RENCI, UNC Chapel Hill)
Ewa Deelman, **Karan Vahi**, Mats Rynge (ISI, USC)
Matthew Miller, Werner Sun, Peter Ko, Kelly Nygren, Keara Soloway, Rolf Verberg (CHESS, Cornell)
Brandon Sorge (IUPUI)

2024 **User Meeting**
Cultivating Community
2024 CHESS User Meeting

NSF OAC Award Numbers: 2320373, 2320375, 2320374

Funded by the National Science Foundation, Grant #2320373

Introduction

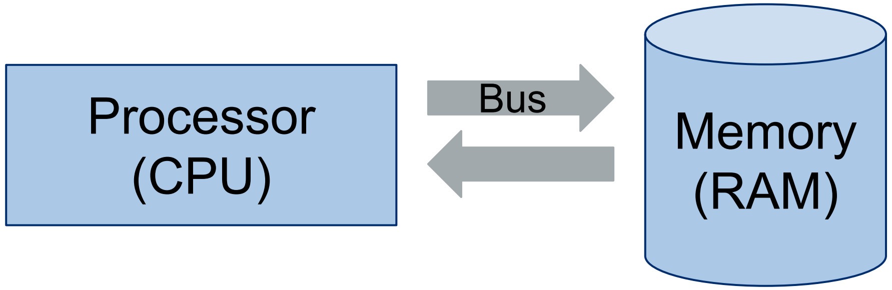
- So far you have had a good overview of the programming fundamentals with Linux and Python
- In this session, we will give you brief insight on
  - Concepts of parallel and distributed computing
  - Submitting jobs to a HPC Cluster at Chess
  - Introduction to CyberInfrastructure (CI)

This presentation is a **forward looking presentation** that gives you an overview of what **CyberInfrastructure(CI) Resources are available** for you at CHESS and nationally.

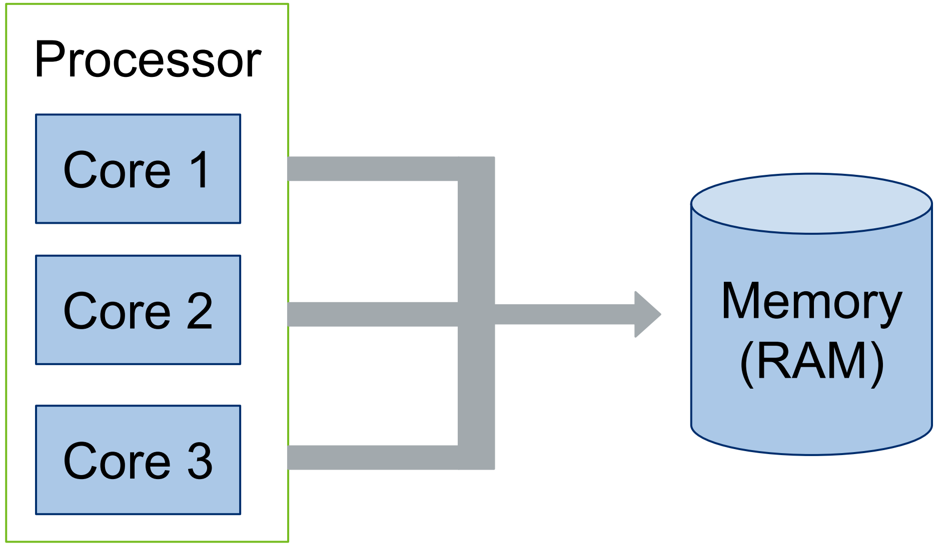
Funded by the National Science Foundation, Grant #2320373

Parallel Computing

- Instead of having a processor with:
  - One unit to add
  - One unit to fetch data from RAM etc
  - Basically one processor



- You actually have multiple units to add, multiply etc within one processor
- Each of these group of units is called a "core" and can act independently
- Now we can execute multiple instructions in parallel!

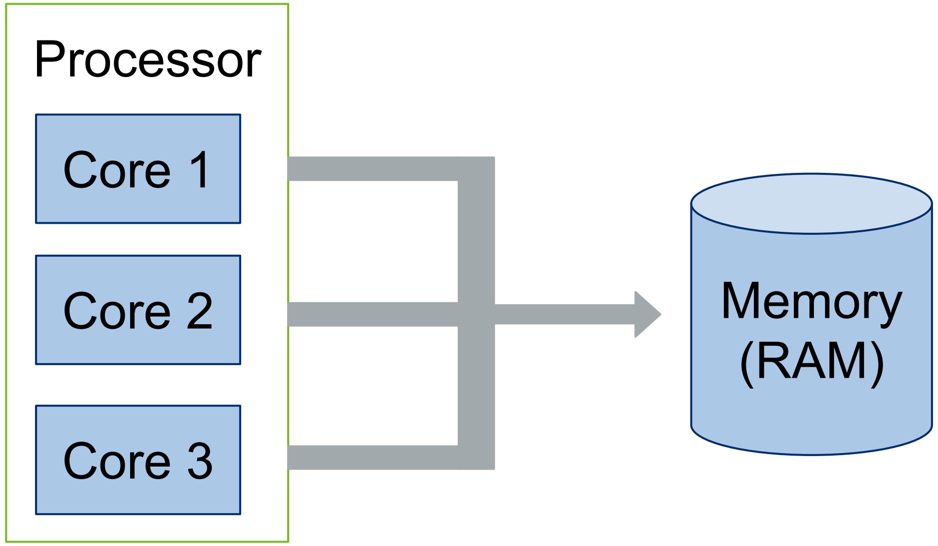


Slide Credit: CICompass CICF

Funded by the National Science Foundation, Grant #2320373

What is Parallel Computing

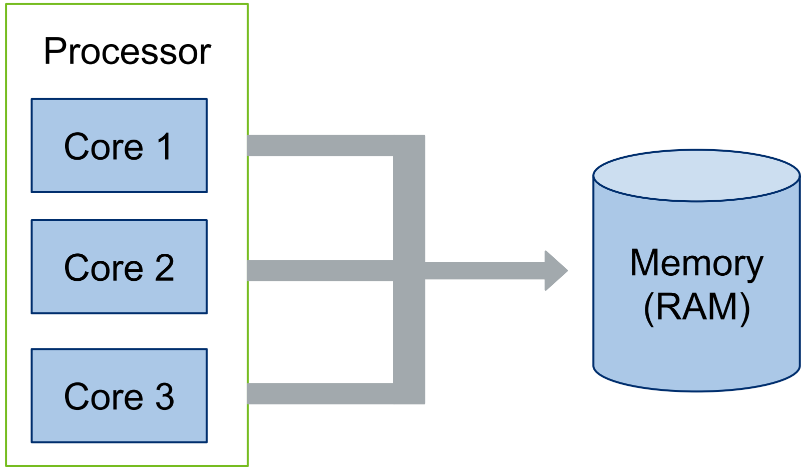
- Your program is the following sequence:
  - ADD 1, 2
  - ADD 3, 4
  - ADD 5, 6
- Now we can execute each ADD on 1 core !



Funded by the National Science Foundation, Grant #2320373

Why parallel computing is complicated?

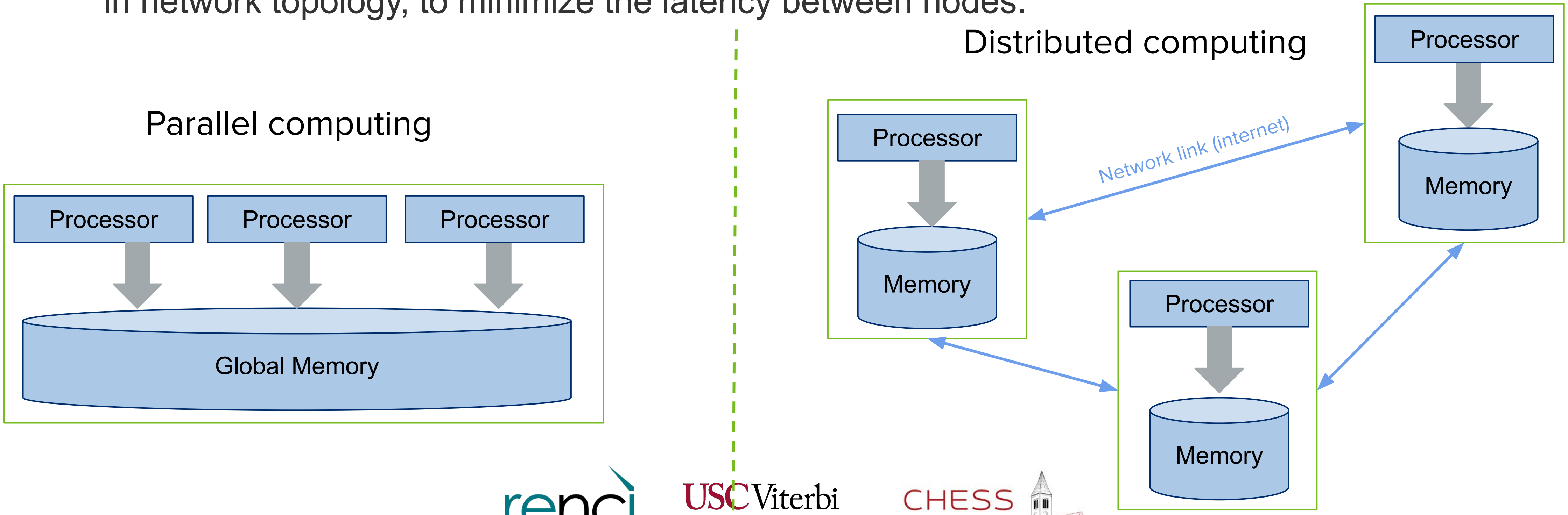
- Thinking in parallel is hard! Humans are better at thinking sequentially
  - We need to adapt sequential algorithm to parallel machines!
  - Example: you want to sort a sequence of numbers: [4, 3, 2, 6, 5, 1]
  - How can you parallelize a sorting algorithm?
- Memory coherency!
  - In your memory you have stored an array: A = [0, 1, 2, 3]
  - If core #1 modifies A and does A[0] = 4 and at the same time core #2 tries to read the value of A[0]? Will core #2 gets 0 or 4?
  - It should get 4! To make sure it does we can use locking mechanisms (mutex etc)
- And many other things.. (costs, cooling, debugging is hard in parallel etc)



Funded by the National Science Foundation, Grant #2320373

Parallel Computing vs Distributed Computing

- What's the differences between parallel computing and distributed computing
  - These two notions are close! But not identical
  - Sometimes, people will use these two terms interchangeably
- High Performance Computing is a type of Distributed Computing
  - In HPC setup; a connected group of servers placed closely to one another both physically and in network topology, to minimize the latency between nodes.



Slide Credit: CICompass CICF

Funded by the National Science Foundation, Grant #2320373

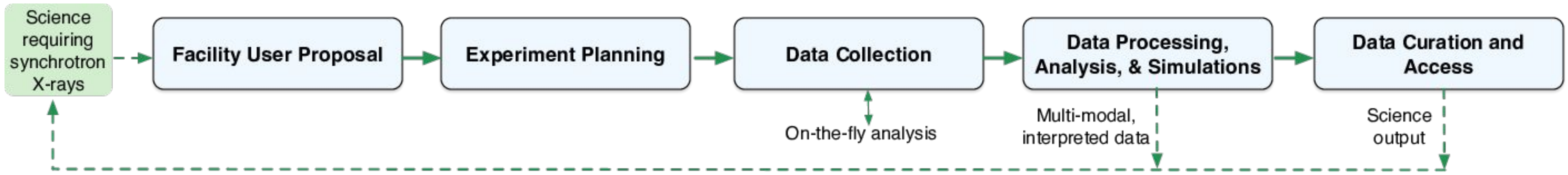
What is CI?

Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and **enable breakthroughs** not otherwise possible [1].

[1] Craig A. Stewart, Stephen Simms, Beth Plale, Matthew Link, David Y. Hancock, and Geoffrey C. Fox. 2010. What is cyberinfrastructure. In Proceedings of the 38th annual ACM SIGUCCS fall conference: navigation and discovery (SIGUCCS '10). Association for Computing Machinery, New York, NY, USA, 37–44. https://doi.org/10.1145/1878335.1878347

Funded by the National Science Foundation, Grant #2320373

Where can CI Help?



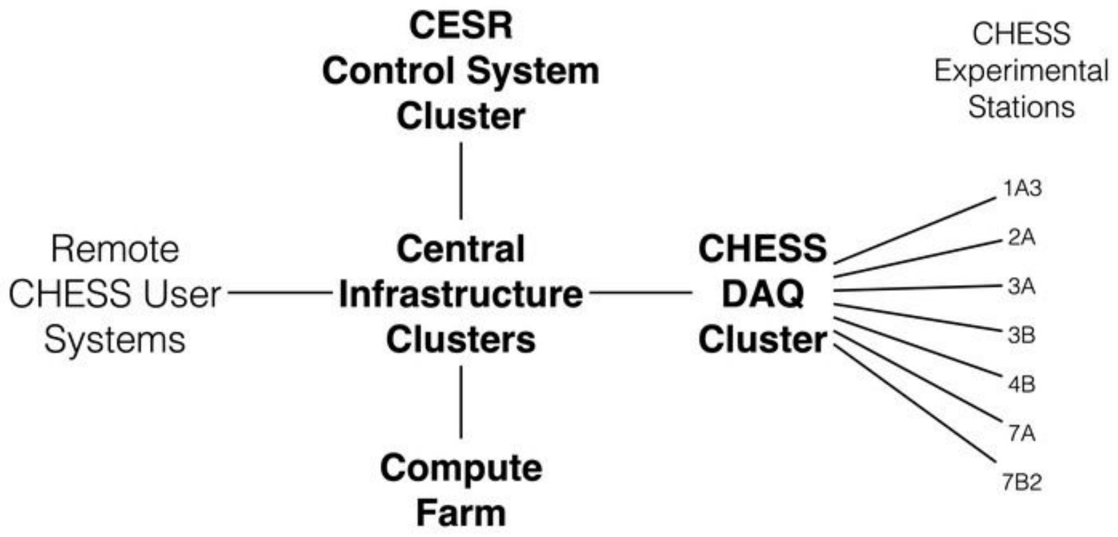
- Challenges at Data Collection
  - Make the data coming off from the beam lines available for analysis in near real-time
- Challenges at data processing stage
  - Processing data (reduction, analysis, simulation, interpretation), handling large data sets, leveraging existing software and CI.
- Challenges at data curation stage
  - Metadata management, Open Science/FAIR and data curation.

Funded by the National Science Foundation, Grant #2320373

CHESS Cyber Infrastructure

- DAQ Cluster is the Data Acquisition System that runs on a dedicated server cluster. Makes available 2PB of storage for raw data collection, analysis, and simulation.

- Data collected at the stations written directly to the DAQ and is immediately available for processing/analysis by users on the Compute Farm.



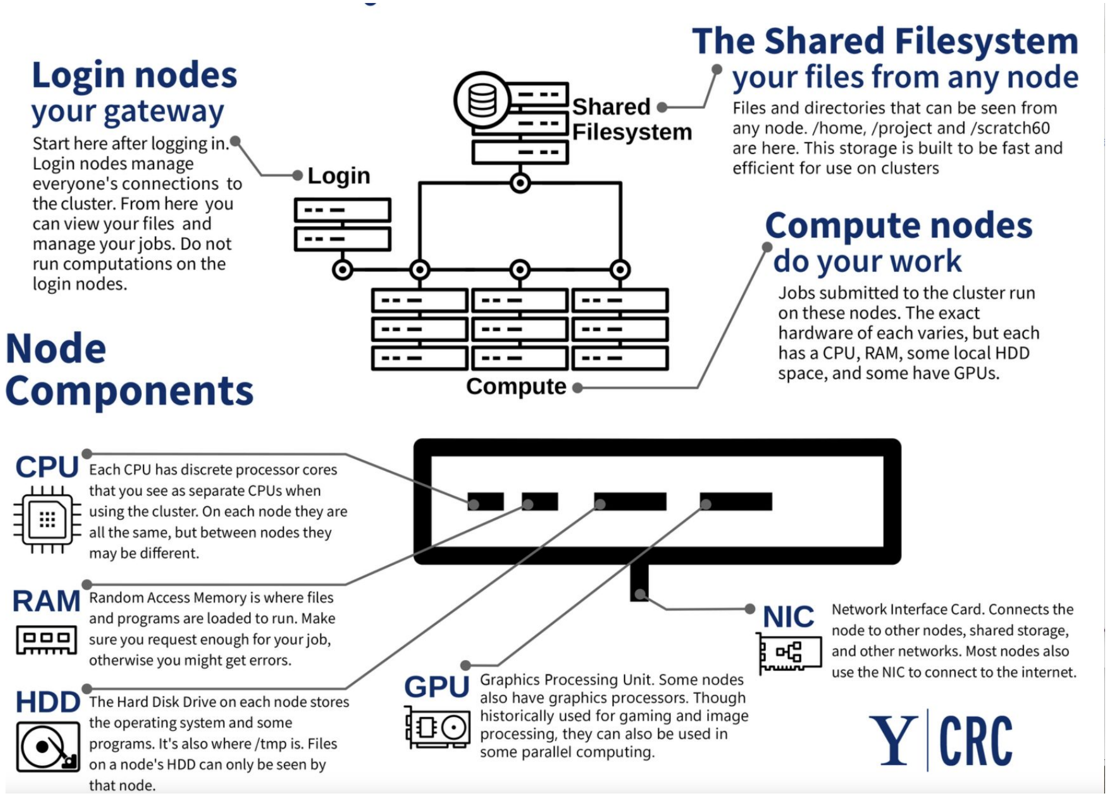
CHESS CI: Interconnected immediately available for processing/analysis by users on the Compute Farm.

Funded by the National Science Foundation, Grant #2320373

CI Resources at CHESS - CLASSE Compute Farm

- HPC Cluster at CHESS/CLASSE
  - central resource of 60+ enterprise-class Linux nodes (with around 400 cores)
  - uses SGE as a front-end queueing system
  - supports interactive, batch, parallel, and GPU jobs
  - ensures equal access to the Compute Farm for all users.

General HPC Cluster Layout



Credit: https://docs.ycrc.yale.edu/clusters-at-yale/

renci USC Viterbi School of Engineering Information Sciences Institute CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

Job Submission to Chess Compute Farm

Two basic steps to submitting a job

1. Create a shell script containing the commands to be run
2. Submit this script to the Compute Farm using the **qsub** command.uses SGE as a front-end queueing system

```
$ cat myscript.sh

#!/bin/bash
echo Running on host: `hostname`.
echo Starting on: `date`.
sleep 10
echo Ending on: `date`.
#$ -q all.q
#$ -S /bin/bash
#$ -l mem_free=8G
/path/to/program.exe
```

```
$ qsub -q all.q myscript.sh
```

Detailed Instructions

https://wiki.classe.cornell.edu/Computing/GridEngine

Funded by the National Science Foundation, Grant #2320373

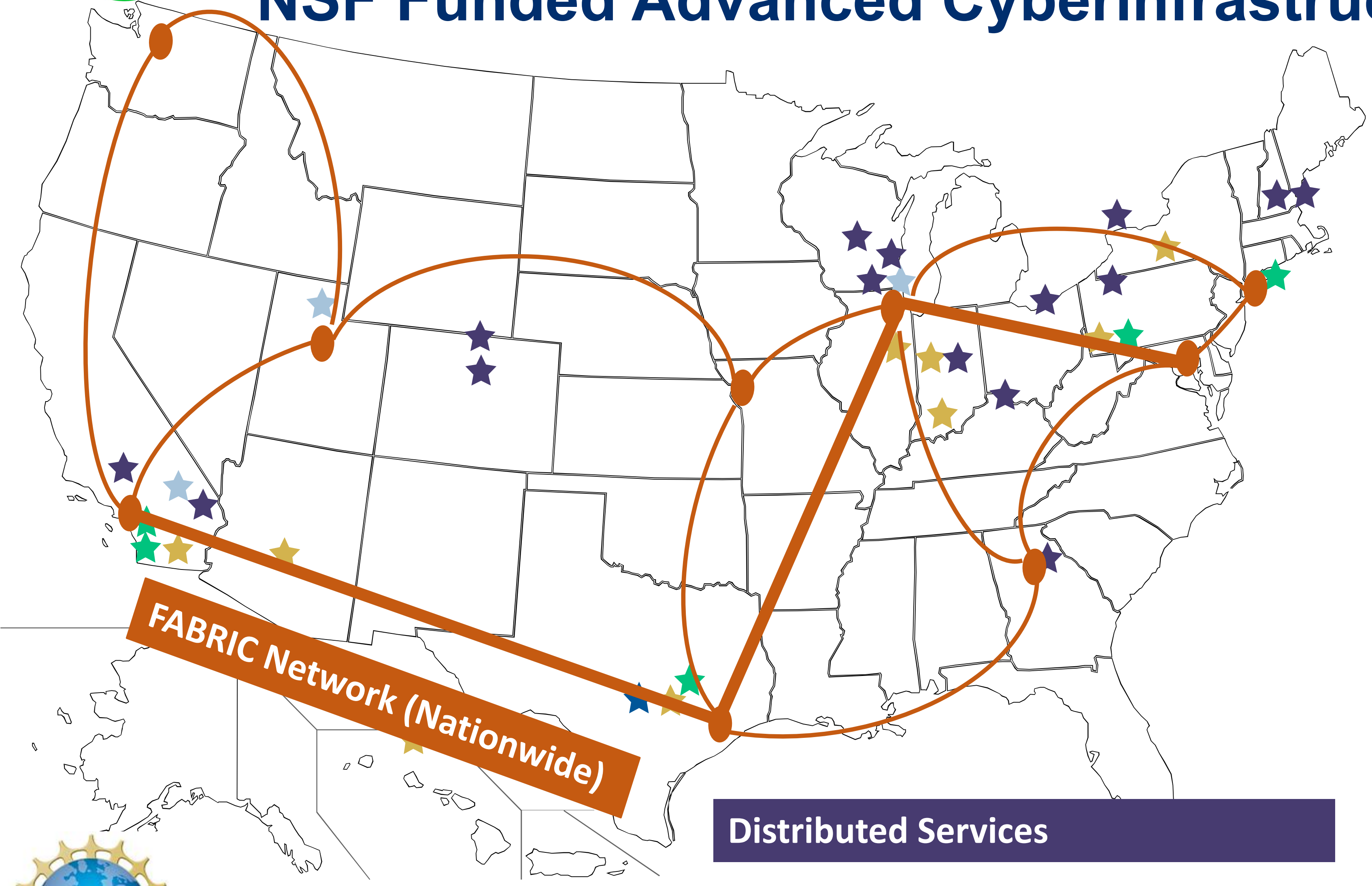
Beyond CHESS

- What if the CHESS Cluster is not sufficient for your needs?
- Are there options?
  - YES! In fact lots in terms of access to Compute and Storage

- NSF makes available two distributed CI available for US researchers to use.
  - ACCESS
    - Collection of both large and experimental HPC resources.
    - Similar to CHESS cluster, multiple and on steroids!
  - PATh/OSG
    - High-throughput computing environment leveraging fair-share access to contributed compute capacity
    - Makes spare resources available from various Campus Clusters

Funded by the National Science Foundation, Grant #2320373

NSF Funded Advanced Cyberinfrastructure Program



Distributed Services

PATh/OSG	University of Wisconsin, Madison
ACCESS	Fourteen Partners

Leadership-class

Frontera	U of Texas, Austin

Innovative Production Systems

Anvil	Purdue University
Bridges 2	Carnegie-Mellon University
Delta/DeltaAI	U of Illinois Urbana-Champaign
Expanse	U of California, San Diego
Jetstream 2	University of Indiana + Partners
Stampede 2	U of Texas, Austin

Prototypes/Testbeds

Neocortex	Carnegie-Mellon University
Voyager	U of California, San Diego
Ookami	Stonybrook University
NRP	U of California, San Diego
ACES	Texas A&M University

Cloud technology resources

Cloudbank	U of California, San Diego
CloudLab	University of Utah
Chameleon	University of Chicago

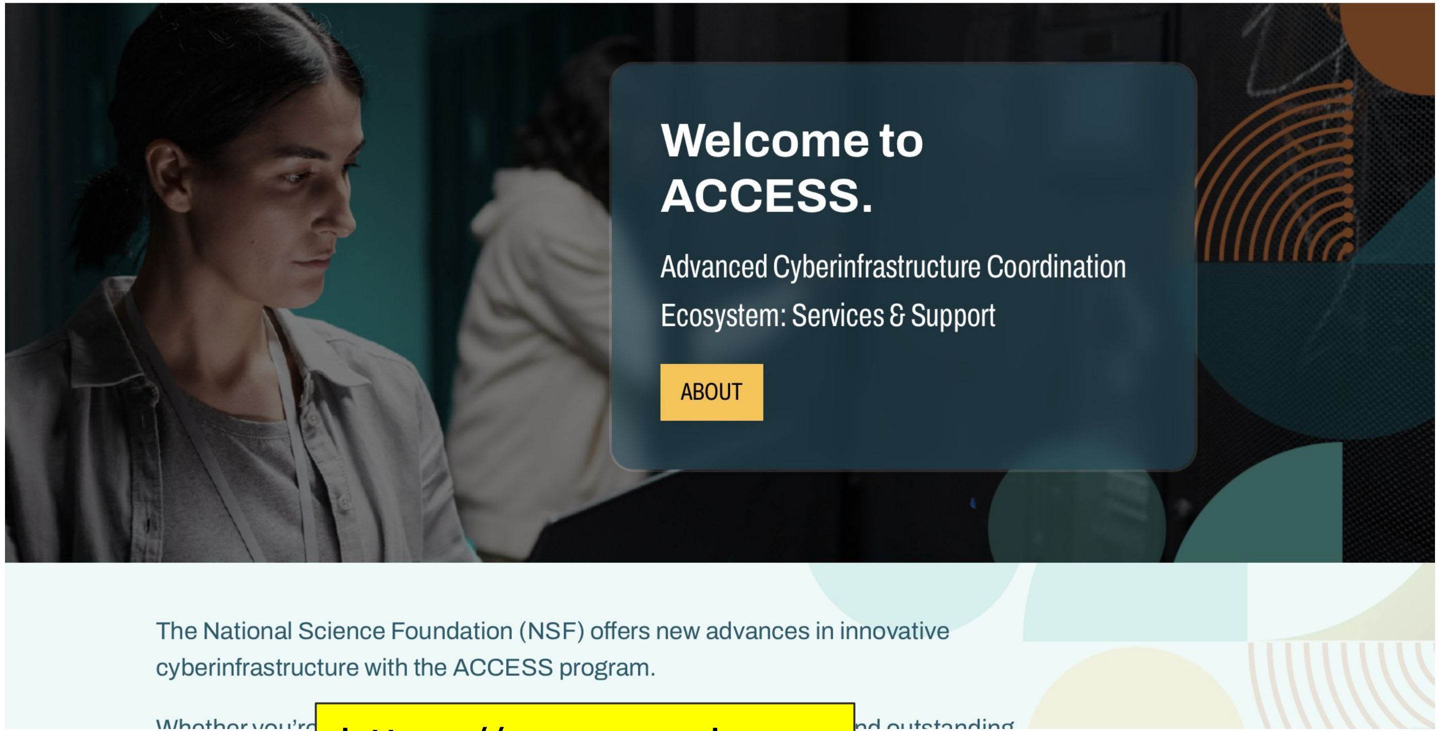
renci USC Viterbi School of Engineering Information Sciences Institute CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

ACCESS - Advanced Cyberinfrastructure Coordination Ecosystem: Services and Support

ACCESS | Advancing Innovation

- Collection of advanced compute, data, networking, and software
- Combines resources and expertise to form an ecosystem to support scientific research and discovery
- Established as a set of independently funded, but coordinated awards



https://access-ci.org

renci | USC Viterbi School of Engineering Information Sciences Institute | CHESS Cornell High Energy Synchrotron Source

Funded by the National Science Foundation, Grant #2320373

Evolution in national cyberinfrastructure informed by community input

TeraGrid™

2001 - 2011

https://en.wikipedia.org/wiki/TeraGrid

XSEDE

Extreme Science and Engineering Discovery Environment

2011 - 2022

https://www.nsf.gov/news/news_images.jsp?cntn_id=121181&org=NSF

ACCESS

Advancing Innovation

2022 -

https://access-ci.org

renci

USC Viterbi School of Engineering Information Sciences Institute

CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

ACCESS Available Resources

- *Multi-core Compute*
- *GPU Computing*
- *Novel Computing*
- *Storage Resources*
- *AI-Focused*

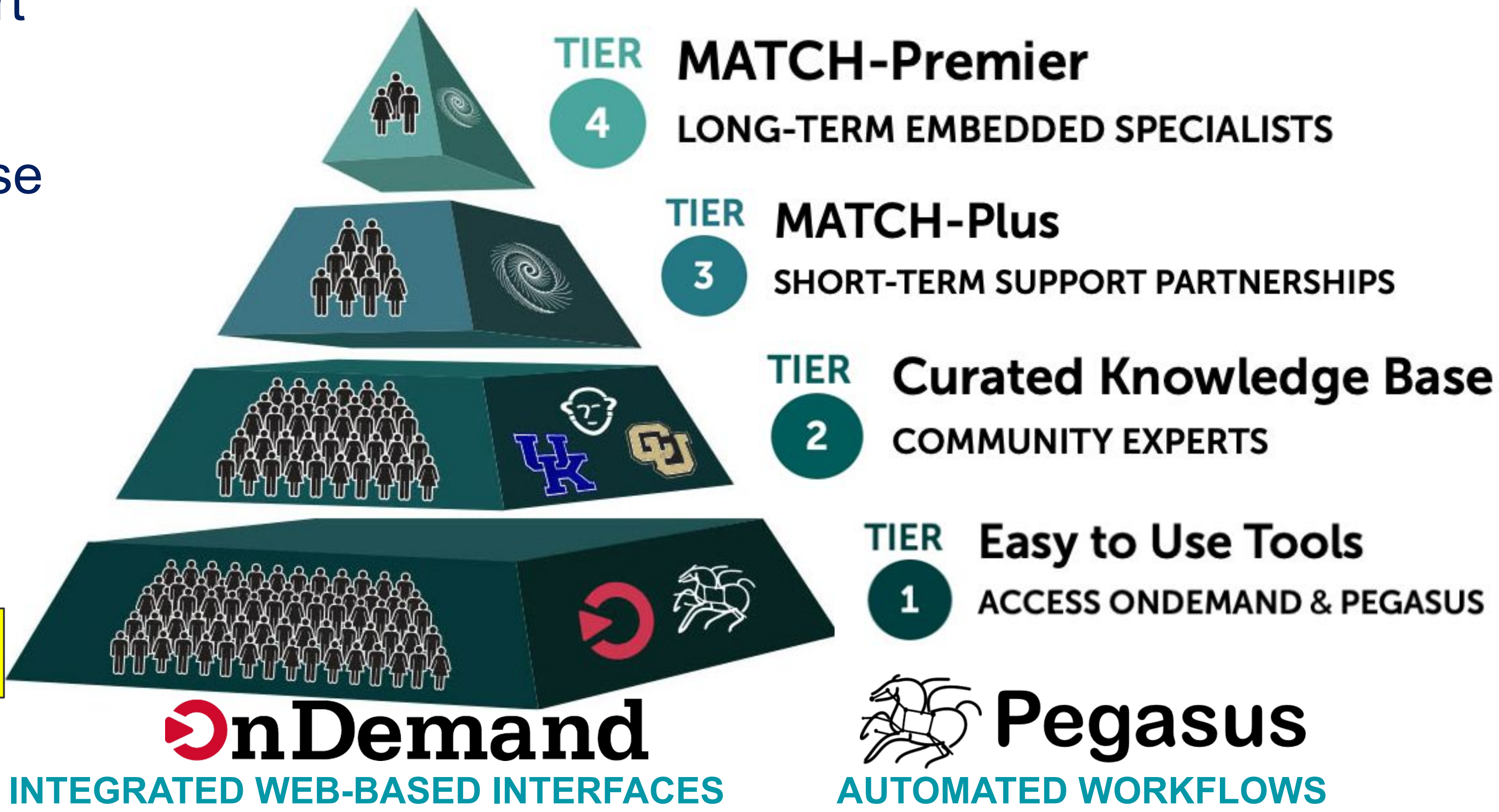
renci
USC Viterbi School of Engineering Information Sciences Institute
CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

ACCESS Researchers Support Services

- Enable innovative research through equitable and scalable support
- Four tiers of support
- Tools, growing knowledge base
- Match-making with experts
- Student engagement
- Engagement from community
- CSSN incentives

https://support.access-ci.org



renci USC Viterbi School of Engineering Information Sciences Institute CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

What is the OSG Consortium

The **OSG Consortium** builds and operates a **set of pools of shared computing and data capacity for distributed high-throughput computing (dHTC).**

https://display.opensciencegrid.org

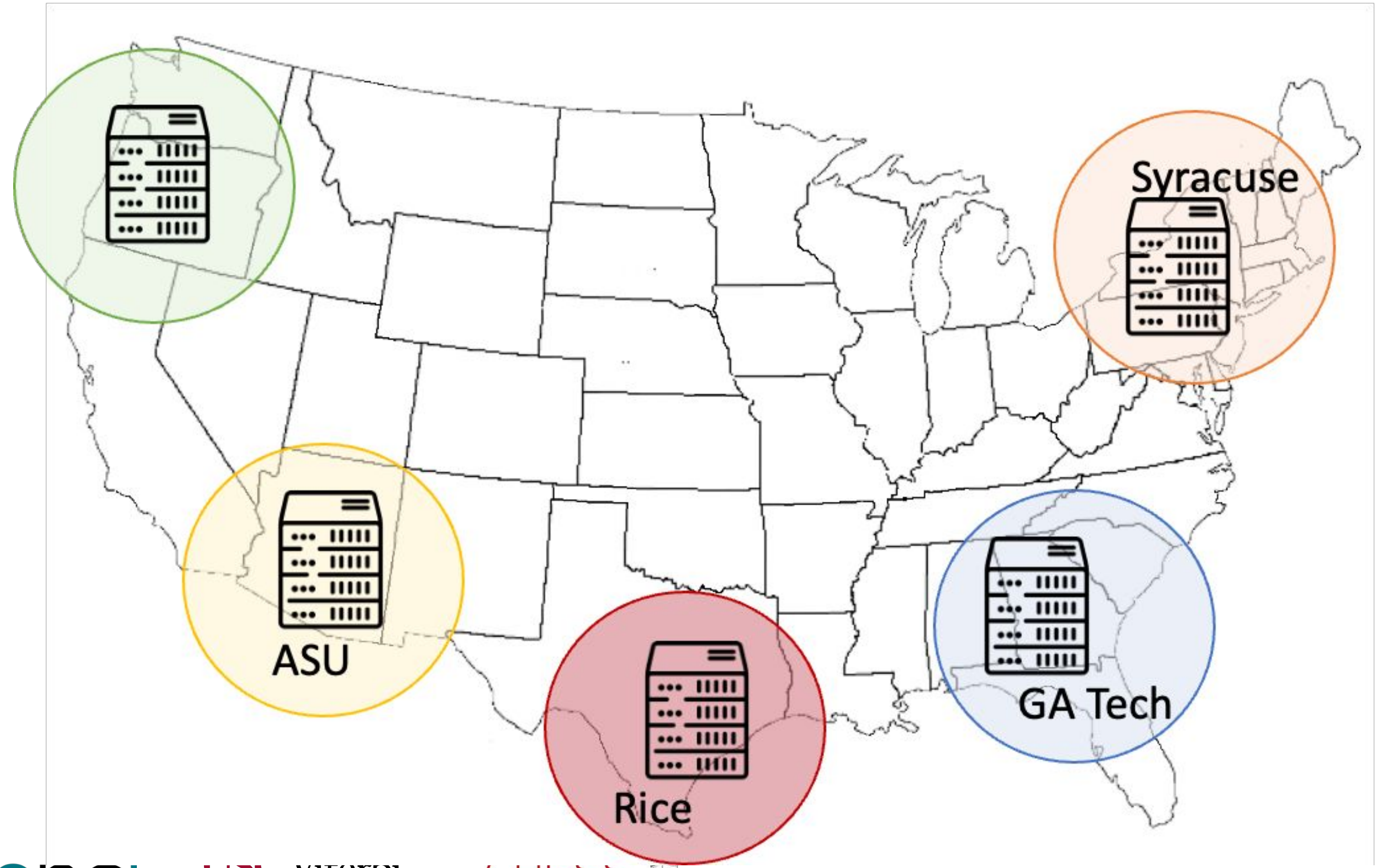
Slide Credit: OSG Research Facilitation Team

renci | USC Viterbi School of Engineering Information Sciences Institute | CHESS Cornell High Energy Synchrotron Source

Funded by the National Science Foundation, Grant #2320373

For Example:

Often, if a campus wants to provide more computing to its researchers, it will invest in a local cluster.



Slide Credit: OSG Research Facilitation Team

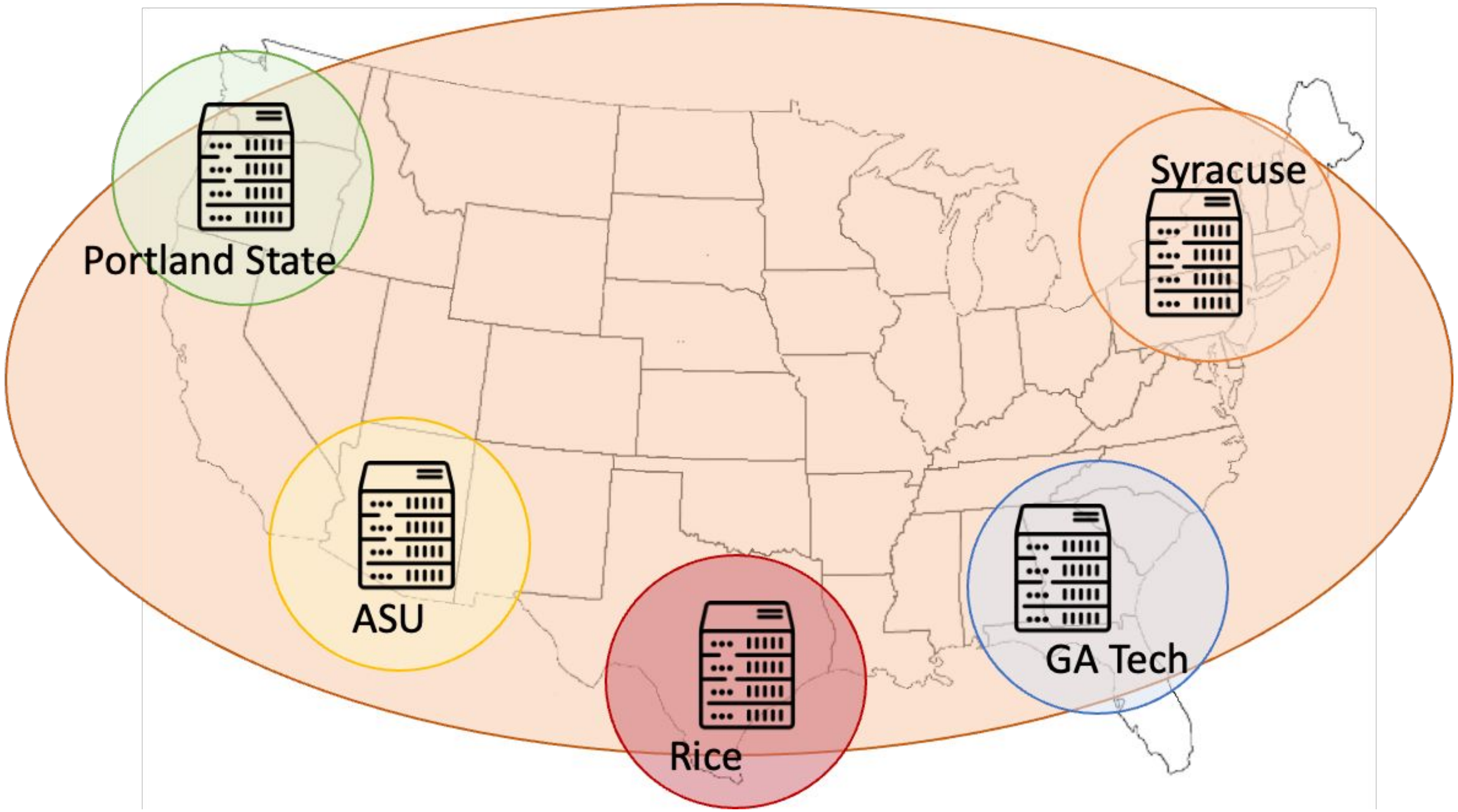
renci
USC Viterbi School of Engineering Information Sciences Institute
CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

Sharing Capacity

But what if institutions pooled their “extra” resources together in a virtual cluster?

(and provided access to researchers without a local cluster?)



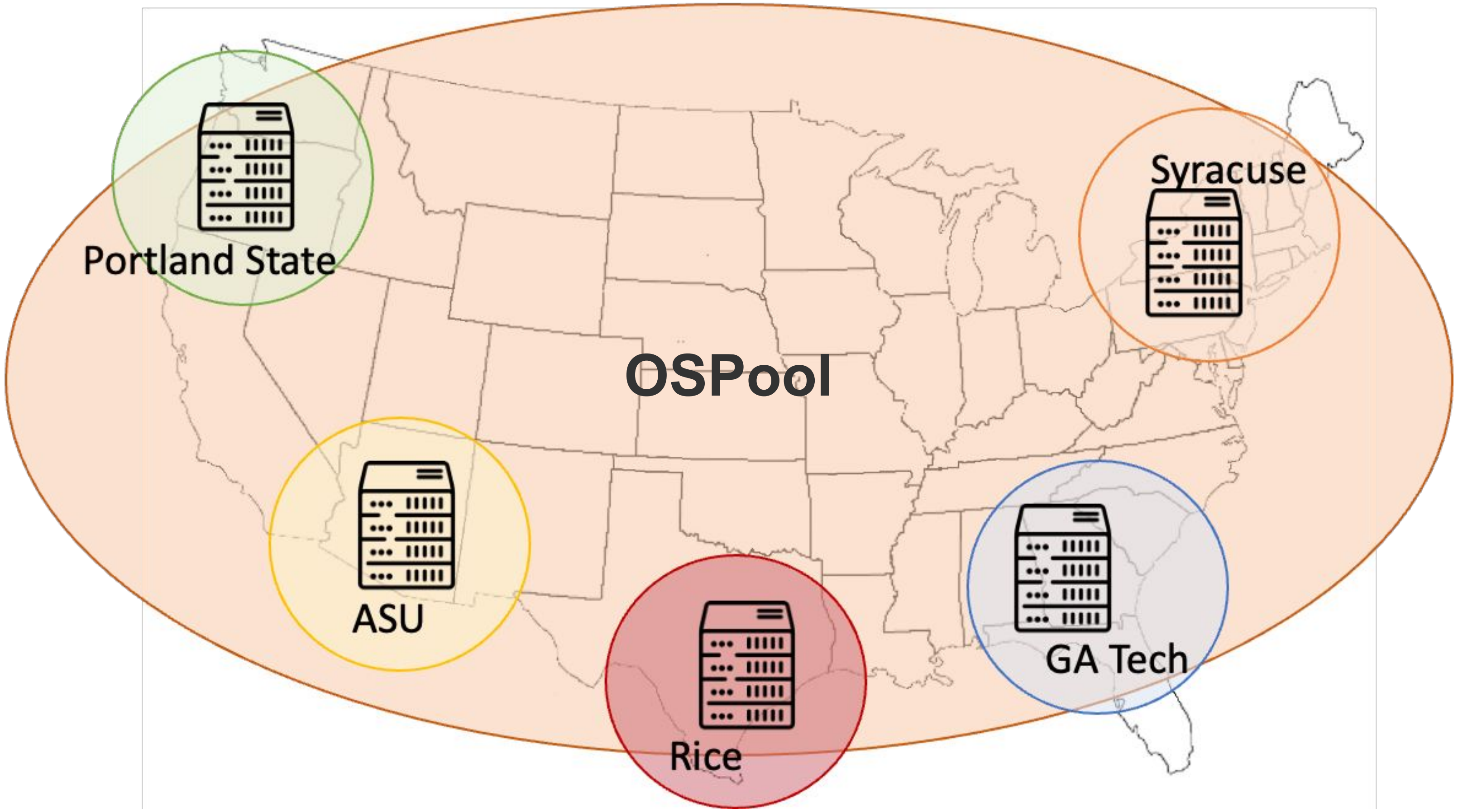
Slide Credit: OSG Research Facilitation Team

renci USC Viterbi School of Engineering Information Sciences Institute CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE

Funded by the National Science Foundation, Grant #2320373

The OSPool

This is exactly the idea behind the OSPool, one of the OSG Services available to researchers.



Slide Credit: OSG Research Facilitation Team

Funded by the National Science Foundation, Grant #2320373

OSG Services: Compute and Storage for Researchers

Open Science Pool (OSPool):
National resource of computing capacity for high *throughput* workloads

OSG-operated Access Points:
OSG-operated service to submit jobs to the OSPool**

Open Science Data Federation (OSDF):
Network of data origins (file servers) and caches for data accessibility

Slide Credit: OSG Research Facilitation Team

Funded by the National Science Foundation, Grant #2320373

What workloads are good for OSG

	Ideal Jobs! (up to 10,000 cores across Jobs, per user!)	Still Very Advantageous!	Less-so, but maybe
Cores (GPUs)	**1** (1; non-specific type)	**<8** (1; specific GPU type)	**>8 (or MPI)** (multiple)
Walltime	**<10 hrs*** *or checkpointable	**<20 hrs*** *or checkpointable	**>20 hrs**
RAM	**<few GB**	**<10s GB**	**>10s GB**
Input	**<500 MB**	**<10 GB**	**>10 GB**
Output	**<1 GB**	**<10 GB**	**>10 GB**
Software	*'portable' (pre-compiled binaries, transferable, containerizable, etc.)*	*most other than* →	*Licensed software; non-Linux*

** the "less-so, but maybe" column could still be an HTC workload, but one that would run more effectively on a local, dedicated HTC system instead of the OSPool*

Slide Credit: OSG Research Facilitation Team

Funded by the National Science Foundation, Grant #2320373

OSG Size and Scale

Last 12 months: 122M jobs - 3B compute hours

Status Map | Jobs | CPU Hours

In the last 24 Hours	
317,000	Jobs
9,081,000	CPU Hours
In the last 30 Days	
9,342,000	Jobs
263,087,000	CPU Hours
In the last 12 Months	
122,417,000	Jobs
3,091,109,000	CPU Hours

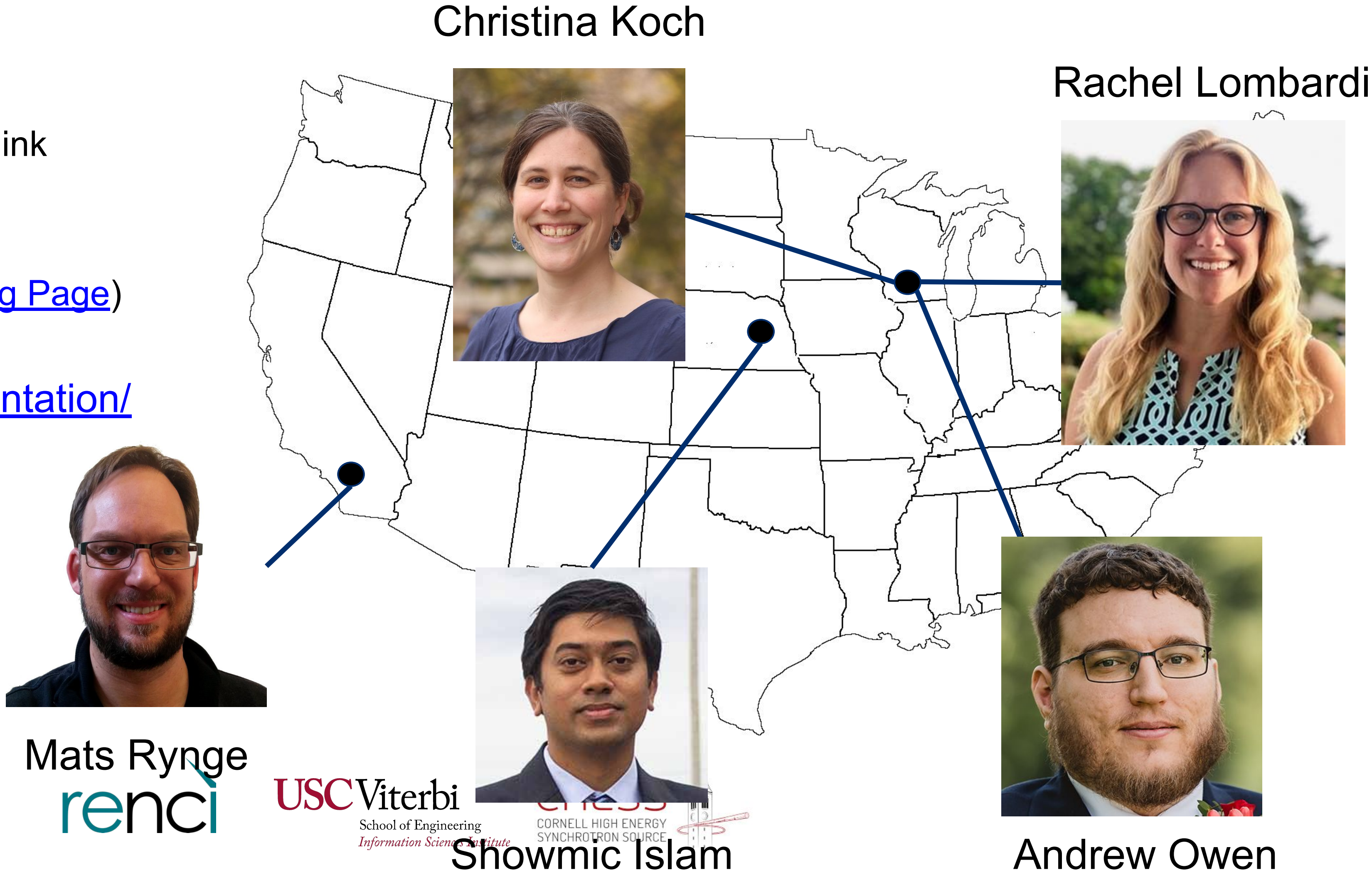
OSG services enabled across 155 sites

Slide Credit: OSG Research Facilitation Team

Funded by the National Science Foundation, Grant #2320373

Support: OSG Research Facilitator Network

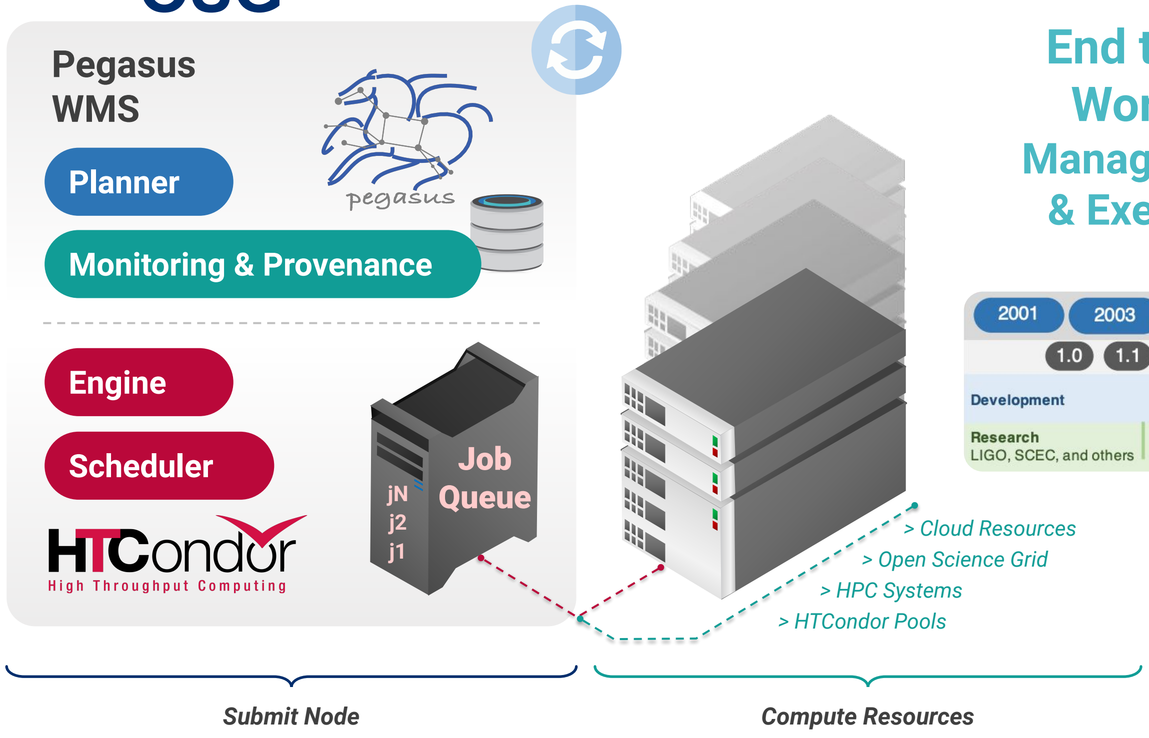
- **Email**: support@osg-htc.org
- **Office Hours**:
  - Tues/Thurs, contact us for zoom link (Help Page)
- **Training Opportunities**:
  - 3rd Tuesday of the month (Training Page)
- **Guides**: https://portal.osg-htc.org/documentation/



Slide Credit: OSG Research Facilitation Team

Funded by the National Science Foundation, Grant #2320373

Pegasus Workflows: Common tool that supports both ACCESS and OSG



End to End Workflow Management & Execution

- Develop portable scientific workflows in Python, Java, and R
- Compile workflows to be run on heterogeneous resources
- Monitor and debug workflow execution via CLI and web-based tools
- Recover from failures with built-in fault tolerance mechanisms
- Regular release schedule incorporating latest research and development

2001 · 2003 · 2005 · 2007 · 2009 · 2011 · 2013 · 2015 · 2017 · 2018 · 2020

1.0 · 1.1 · 1.2 · 1.3 · 1.4 · 2.0 · 2.1 · 2.2 · 2.3 · 2.4 · 3.0 · 3.1 · 4.0 · 4.1 · 4.2 · 4.3 · 4.4 · 4.5 · 4.6 · 4.7 · 4.8 · 4.9 · 5.0

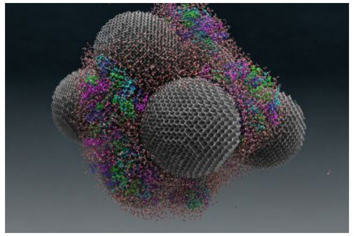
Development: support for GT4 · task clustering · support for AWS · hierarchical workflows · pegasus-lite engine · monitoring dashboard · ensemble manager · support for containers · redesign of APIs

Research LIGO, SCEC, and others: data cleanup algorithms · data footprint · cloud computing evaluation · MPI-based workflow engine design · Real time performance data capture · metadata capture · data integrity assurance

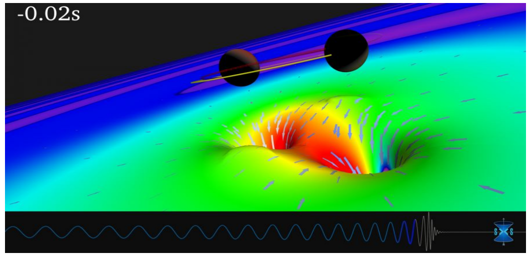
Pegasus in practice

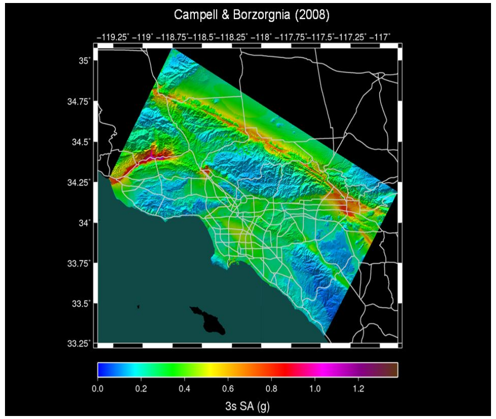
- Laser Interferometer Gravitational Wave Observatory (LIGO) develops large scale analysis pipelines used for gravitational wave detection.
- Southern California Earthquake Center (SCEC) CyberShake project generates hazard maps using hierarchical workflows .
- Oak Ridge National Lab (ORNL) conducted studies on tRNA and nanodiamonds to improve drug delivery design principles.

Visualization of water on nanodiamond spheres from ORNL



LIGO observation of colliding black holes





Hazard map indicating maximum amount of shaking at a particular geographic location generated from SCEC's CyberShake Pegasus workflow

renci · USC Viterbi School of Engineering Information Sciences Institute · CHESS Cornell High Energy Synchrotron Source

Thanks

National Science Foundation

NSF Cybertraining Award # 2320373

renci

USC Viterbi School of Engineering Information Sciences Institute

CHESS CORNELL HIGH ENERGY SYNCHROTRON SOURCE