NYSGrid

New York State CyberInfrastructure

Steve Gallo

Center for Computational Research

University at Buffalo

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

What is NYSGrid?

➢ **Consortium of New York institutions formed in July 2006**

➢ **Grassroots collaboration**

➢ **Focused on developing statewide CI capabilities to advance research and education throughout New York State**

➢ **http://www.nysgrid.org/**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

What is NYSGrid?

➢ **19 Member Organizations**

➢ **Diverse cross-section of researchers**

➢ **2 and 4 year schools, public sector institutions**

➢ **Main data centers at Buffalo and Cornell**

➢ **Many small sites (8 – 16 machines)**

➢ **Recently installed 2 Blue Gene machines at RPI and Stony Brook (not available via grid yet)**
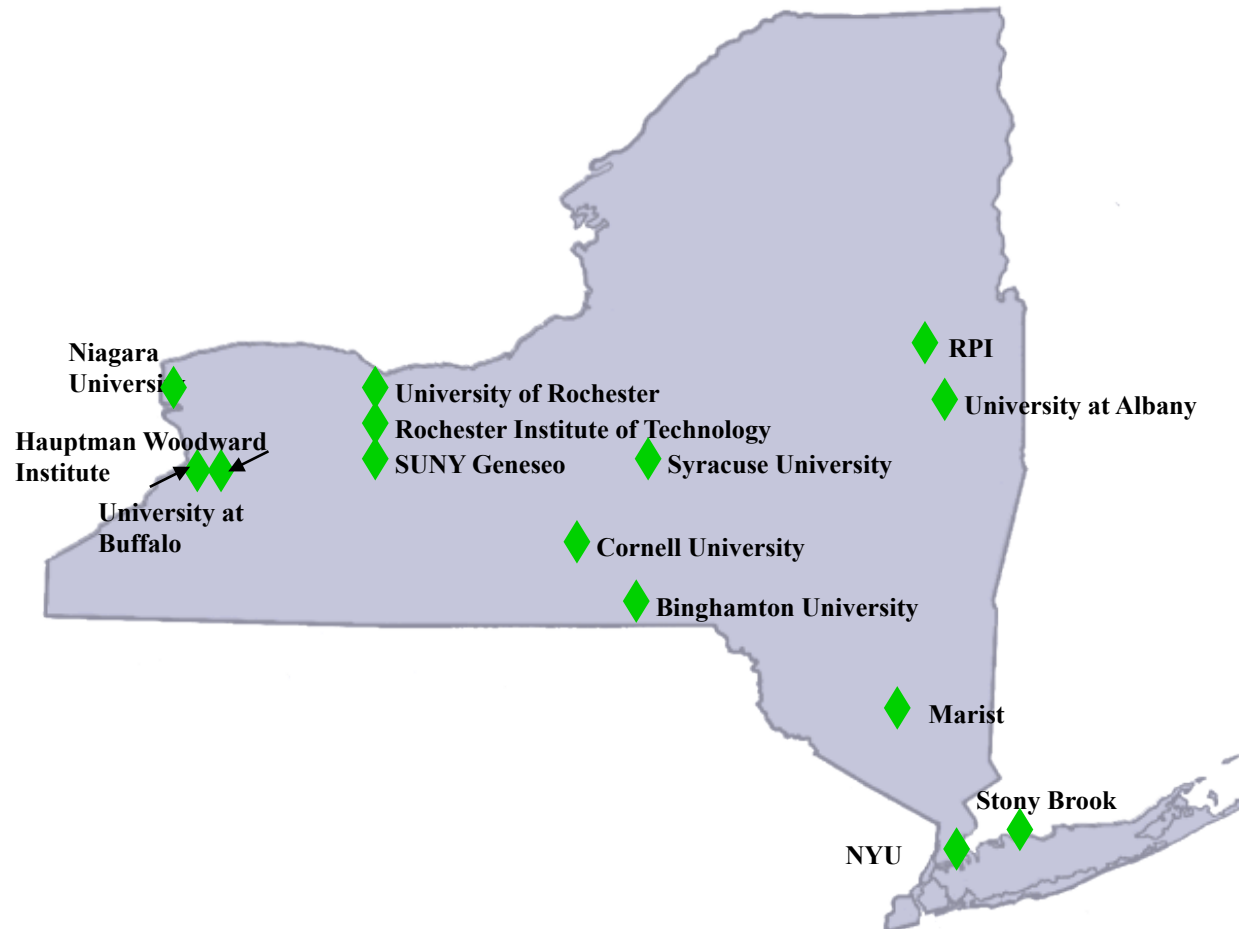
➢ **Campus Condor flocks at RIT and UB**

Open Science Grid

UB CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

What is NYSGrid?

Niagara University

Hauptman Woodward Institute

University at Buffalo

University of Rochester
Rochester Institute of Technology
SUNY Geneseo

Syracuse University

RPI

University at Albany

Cornell University

Binghamton University

Marist

Stony Brook

NYU

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Running NYSGrid

- ➢ **Support center at Buffalo**

- ➢ **Compute resources exist and are online across the state**

- ➢ **Many small sites with little or no committed system administration support**

- ➢ **0.5 support center FTE**

- ➢ **Support center personnel available to assist sites with administering their grid systems, running their applications**

- ➢ **MPI – smaller sites don't have the expertise so they utilize larger sites (Buffalo, Cornell)**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Enabling Science

What Did We Do?

- **Create infrastructure within the state (hardware and software)**
- **Provide education**
  - **Train users at individual sites statewide**
  - **Multiple workshops**
  - **Submission examples**
- **Bring on the users!**
- **Where are the users?**
- **Identify Champions**
  - **Convince champions that they will benefit**
  - **Work closely to get their applications running over the grid**
  - **Wow – I can get more science done!**
  - **Leverage their successes to bring on new users**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Enabling Science

➢ **Identifying willing users can be difficult**

- **"Why should I use the grid?"**

- **"I'm happy with running locally."**

- **"I'm too busy right now."**

➢ **Champions are important**

- **Dave Kofke (UB, Chemistry) - Molecular simulations**

- **Barbara Poliks (Binghamton, Physics) – Molecular dynamics (MPI)**

- **Abani Patra (UB, Engineering) & Mike Sheridan (UB, Geology) – Geophysical mass flow**

➢ **Leverage success stories to attract new users**

Working With OSG

- ➢ **Utilize OSG software stack**

- ➢ **Leverage the experience of the OSG community**

- ➢ **Learn best practices**

- ➢ **Identify tools that others are using**

- ➢ **Use educational materials**

- ➢ **Utilize Engage VO expertise**

  - ▪ **Grid submission using Condor (Kofke)**

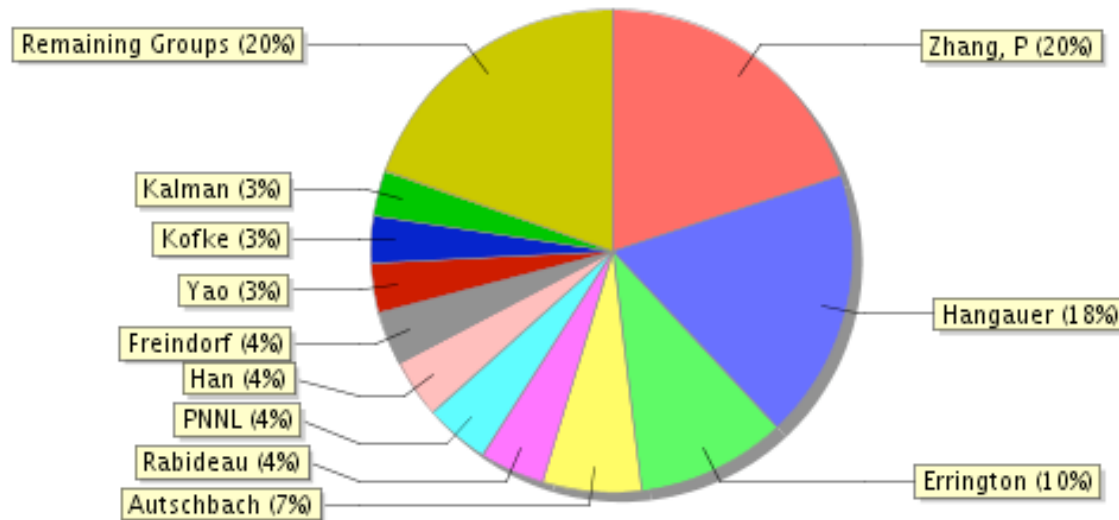 - ▪ **Modifying job manager to run site-specific MPI (Poliks)**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Campus Grids

- ➢ **Institutions want to support research computing**
- ➢ **Condor is very appealing**
- ➢ **CIO support is important!**
- ➢ **Buffalo condor flock (public sites initially ~400 cores)**
- ➢ **RIT condor flock (~1000 cores)**
- ➢ **Campus-specific usage monitoring is important (UBMod)**

Pie chart labels:
- Remaining Groups (20%)
- Zhang, P (20%)
- Kalman (3%)
- Kofke (3%)
- Yao (3%)
- Freindorf (4%)
- Han (4%)
- PNNL (4%)
- Rabideau (4%)
- Autschbach (7%)
- Hangauer (18%)
- Errington (10%)

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Challenges

- ➢ **Education**
  - ▪ **Why should researchers look at grid computing?**
  - ▪ **Is the benefit worth the time investment?**

- ➢ **The Human Element**
  - ▪ **Bringing on new users takes time and effort**
  - ▪ **Many institutions do not have dedicated FTE**

- ➢ **MPI**
  - ▪ **MPI applications are special**
  - ▪ **MPI task launchers all behave slightly differently**
  - ▪ **Applications may not deal well with variances**

Challenges

➢ **Smaller institutions pose their own challenges**

  ▪ **2 year colleges & smaller 4 year institutions**

  ▪ **Little or no systems administration support**

  ▪ **May not have any compute power locally**

  ▪ **Limited network bandwidth**

  ▪ **Need more support from systems administrators and computational scientists**

➢ **But they also present opportunities**

  ▪ **Researchers more willing to use the grid because of these restrictions**

  ▪ **Eager to get their applications running**

  ▪ **Willing to provide feedback**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Future Activities

➢ **Make campus Condor pools available on the grid (RIT, UB)**

➢ **Attracting more users!**

➢ **Continued workshops and education**

➢ **Seeking funding from NY state for computational scientists**

  ▪ **Regional training & support**

  ▪ **Based in different regions across the state**

  ▪ **Travel to institutions to educate and assist users**

➢ **Support specific disciplines using science portals**

➢ **Support education using virtualized clusters to allow students to experiment with parallel & grid computing**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Future Activities

- ➢ **Focus on smaller institutions with little or now local compute resources**

- ➢ **Attract new users by offering modest competitive grants to grid-enable their research**

- ➢ **Provide access to submit nodes and data repositories**

- ➢ **Develop procedures to ship pre-installed gatekeepers to small sites**

Open Science Grid

CENTER FOR COMPUTATIONAL RESEARCH
University at Buffalo *The State University of New York*

NYSGrid

Conclusion

Questions?