



Improved Simulation Through Collaboration

Rensselaer Polytechnic Institute • Stony Brook University • University at Buffalo • Brookhaven National Laboratory • NYSERNet

*Collaborative Research in a Regional Grid
Using HUBzero to Facilitate Collaboration*

Alisa G. Neeman, Steven M. Gallo,
Efstratios Efsthathiadis



Collaborative volcano research and risk mitigation

Outline

- Why a Regional Grid
- Theory of the Community Account
- What is HUBzero
- Security
- Handling Big Data

Motivation: Regional Grid

- Sharing computational resources
- Collaboration
 - Share knowledge and expertise
 - Research
 - Teaching
- Funding agencies encourage this!
- Three institutions in New York with supercomputing resources decide to build a virtual organization

The screenshot displays the homepage of the hpc2 website. The header includes the logo 'hpc²' and the tagline 'Improved Simulation Through Collaboration'. Below this, it lists member institutions: Rensselaer Polytechnic Institute, Stony Brook University, University at Buffalo, Brookhaven National Laboratory, and NYSERNet. A search bar and 'Login'/'Register' links are also present. The main navigation menu includes 'HOME', 'USER INFO', 'ACCOUNTS', 'ABOUT', and 'SOFTWARE'. A central content area features a scrollable box with the title 'The NY State High Performance Computing Consortium' and a description of the partnership. To the right, there is a large, colorful 3D molecular simulation image of Cytochrome P450 binding, with a caption at the bottom: 'Cytochrome P450 binding M. Freindorf, J. Kong, T. Furlani, University at Buffalo'.




The Promise

“HPC² will provide a single, **easy-to-use** entry point (portal) for all users that will provide them with a gateway to HPC² resources, including a description of how to access those resources.”

“End users can potentially be **freed from having to learn about the operating environment** of the various resources, which can vary greatly from institution to institution.”

Challenges

- **Diverse Access Policies**
 - Site specific Terms of Use, account application requirements
 - Site Specific Authentication
 - **One time (crypto card) passwords**
 - challenge-response passwords
 - Public/private key
- **Diverse Hardware and Systems**

Architecture	Batch System	Institution
Dell P4 64-bit Linux cluster	Torque PBS	 CENTER FOR COMPUTATIONAL RESEARCH University at Buffalo <i>The State University of New York</i>
AMD Opteron cluster	SLURM	
Blue Gene L	SLURM	
Blue Gene L	LoadLeveler	
Blue Gene P	LoadLeveler	

A Little Theory Helps Us

- Science gateway based on a **community account**¹
 - One account submits jobs for all community members
 - Shared Responsibilities

Responsibility	Responsible Party	Notes
User authentication	Portal	
Authorization for user actions	Portal	Users transitively receive authorization to submit jobs
Auditing user actions	Compute resource	Resource can't identify individuals
Resource allocation decisions	Portal	
Account for resource use	Portal	
Protect users' processes from each other	Portal	Dynamic individual user workspaces

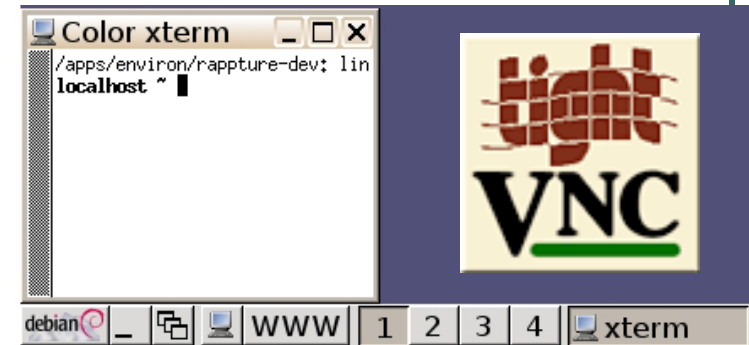
HUBzero roughly follows this model

¹ Von Welch et al, 2006

How? HUBzero Infrastructure

1. Virtual Network Computing

- Lets you see and interact with desktop applications across any network.
- VNC Linux **workspace** included



2. Rappture User Interface

- Rapid Application Infrastructure Toolkit

3. Submit Tool

- submit takes a user command and executes it remotely.

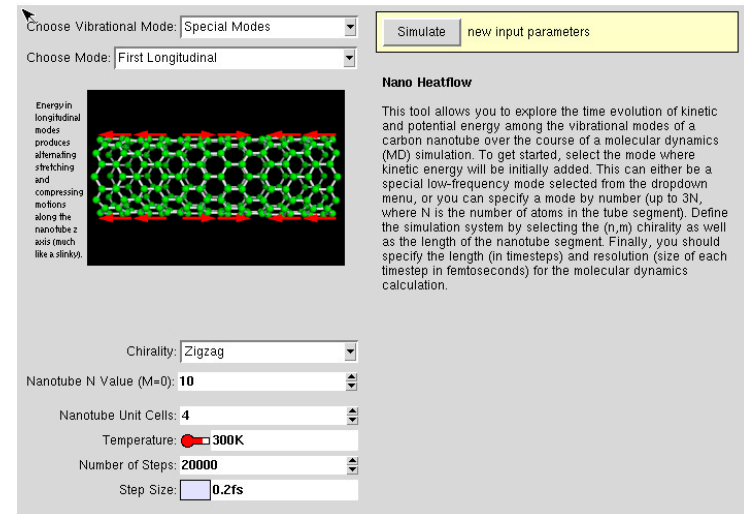


4. Joomla Web Content Management System

Rappture Introduction

GUI toolkit to wrap simulation software

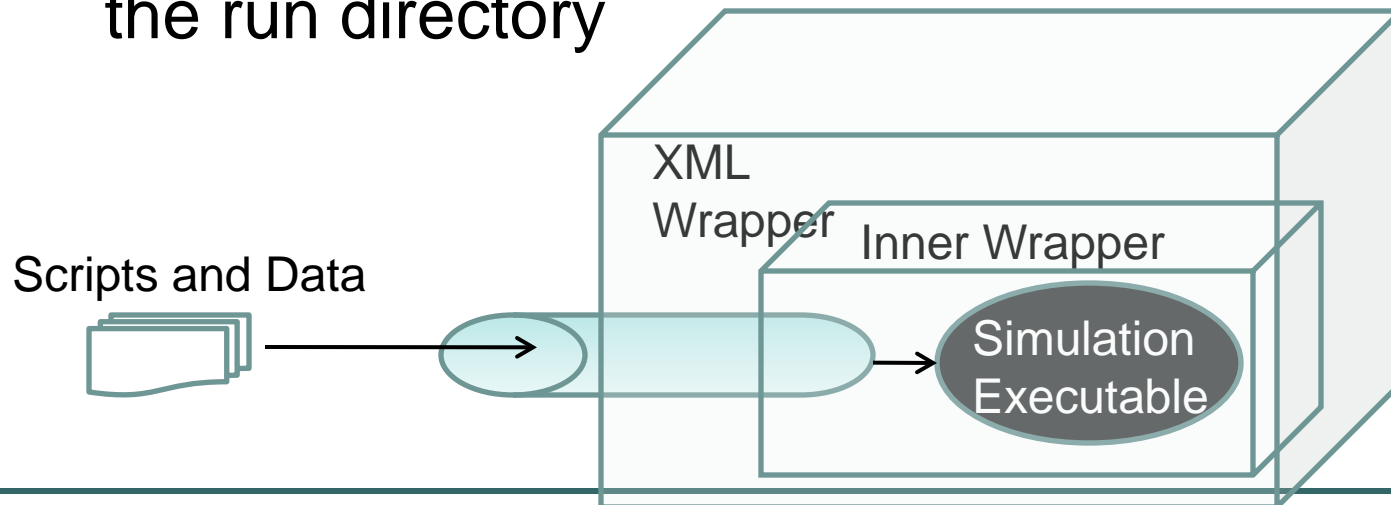
- Designed principally for input **values**
- Simulation code **modified** to read/write values using Rappture library
- Can do file-based I/O, but clunky



The screenshot shows the Rappture GUI for the Nano Heatflow simulation. At the top, there are two dropdown menus: "Choose Vibrational Mode: Special Modes" and "Choose Mode: First Longitudinal". To the right of these is a yellow "Simulate" button with the text "new input parameters" next to it. Below the dropdowns is a 3D visualization of a carbon nanotube segment, colored green and red, with red arrows indicating vibrational modes. To the left of the visualization is a text box that reads: "Energy in longitudinal modes produces alternating stretching and compressing motions along the nanotube z axis (much like a slinky)". Below the visualization are several input fields: "Chirality: Zigzag", "Nanotube N Value (M=0): 10", "Nanotube Unit Cells: 4", "Temperature: 300K", "Number of Steps: 20000", and "Step Size: 0.2fs". To the right of the input fields is a text box titled "Nano Heatflow" that contains the following text: "This tool allows you to explore the time evolution of kinetic and potential energy among the vibrational modes of a carbon nanotube over the course of a molecular dynamics (MD) simulation. To get started, select the mode where kinetic energy will be initially added. This can either be a special low-frequency mode selected from the dropdown menu, or you can specify a mode by number (up to 3N, where N is the number of atoms in the tube segment). Define the simulation system by selecting the (n,m) chirality as well as the length of the nanotube segment. Finally, you should specify the length (in timesteps) and resolution (size of each timestep in femtoseconds) for the molecular dynamics calculation."

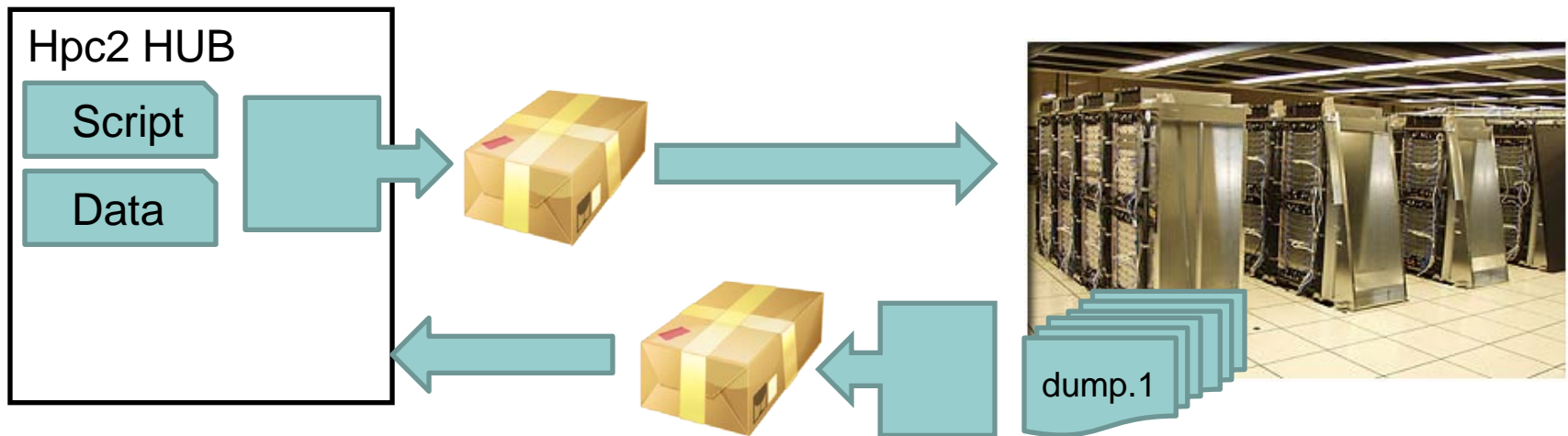
Rappture and File I/O

- GUI Interface specified with XML document
 - Input data gets inserted into XML document
 - From there, inner wrapper script writes file to the run directory



Submit Acts Like a Cluster User

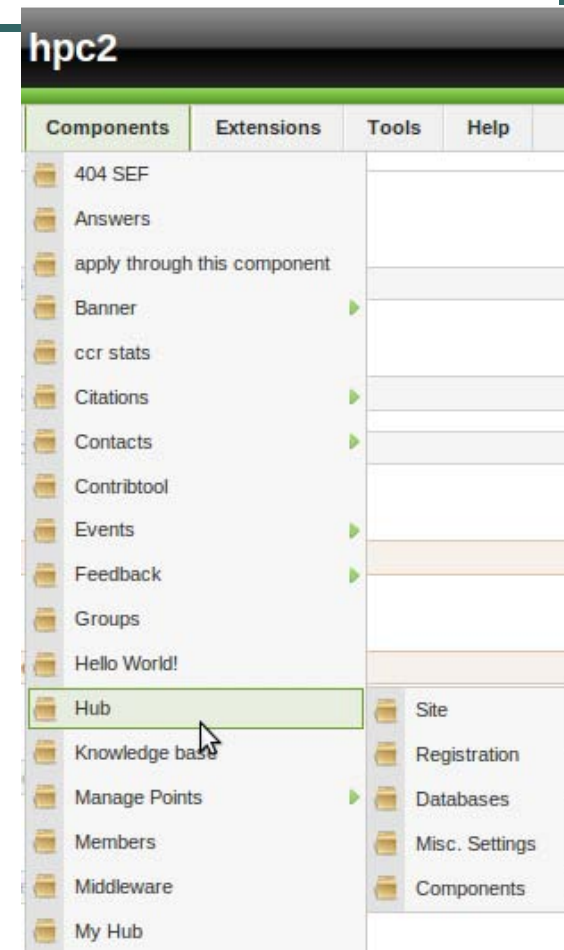
1. **Startup:** Tars up input directory, scp to cluster
2. **Run:** Job runs in temporary directory
 - Status messages piped back through rapture to user
3. **Finish:** Tars up output files and scp back to user
4. **Cleanup:** Removes temporary directory and contents



Details hidden from the user, job runs in name of the community account.

Joomla Content Management

- Components integrate VNC plugin, windowing kit, and submit tool, providing full scientific gateway capabilities.
- Custom HUBzero components
 - For tool installation
 - User Registration
 - Groups for access control
 - Publishing and sharing materials



HUBzero Joomla components provide authentication, authorization, resource allocation

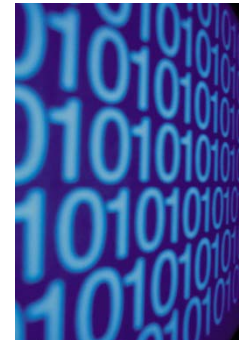
HUBzero Security Infrastructure

- Authentication
 - secure registration and authentication in Joomla
 - HTTPS for communication and LDAP for lookup
- Authorization: Data Storage
 - Home directory read, write access available to each user (traditional Unix privileges)
 - Users can upload a “resource” for sharing through a Joomla HUB component
 - Public, group, private

Authorization: Remote Execution

- Submit runs on a remote cluster as a user with full user privileges accorded by the local cluster environment.
 - create and delete files in its home directory
 - execute files from /usr/bin, /usr/apps, and /usr/local/bin.

In other words, submit can run arbitrary executables built by users on an arbitrary number of processors.



Authorization: Execution

- Access control list for registered tools, by users or groups

Tools.dat
entry

```
[lammps]  
destinations = u2-grid  
executablePath = lmp_linux  
remoteManager = u2-grid_lammps  
restrictedToGroups = ccr
```

- Arbitrary executables can run via **workspace**
 - on HUB server
 - submitted to a cluster

Tool builder can restrict users to members of a group she creates.

Security Modifications for a Regional Grid

```
[u2-grid]
venues =
u2-grid.ccr.buffalo.edu
remotePpn = 2
remoteBatchSystem = PBS
remoteUser = xyz
remoteManager = mpi
venueMechanism = ssh
remoteScratchDirectory =
/san/scratch/hpc2
jobssiteMonitorDesignator =
u2-grid
arbitraryExecutableAllowed =
False
checkProbeResult = False
restrictedToGroups = ccr
```

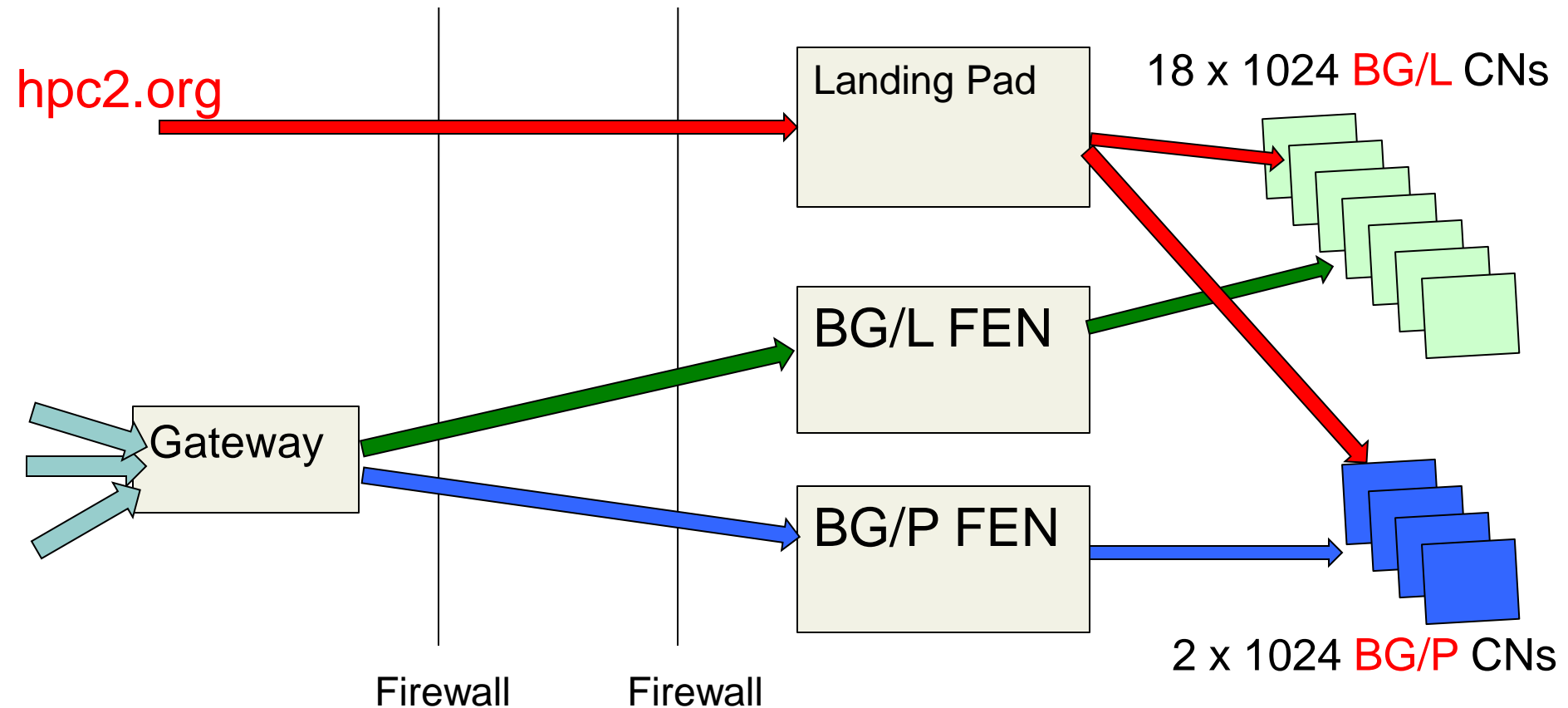
Authorization: Apply principle of least privilege

- Users no longer automatically granted workspace
- Disallow running arbitrary executables at clusters
- Submit modified; **cluster access control by group**
 - ccr, ccni, sbbnl

Cluster Side Security

- SSH for communication/data transfer
- Connection ONLY from HUB domain
- Connection ONLY by HUB community account user

Brookhaven National Lab "Landing Pad" Node For Cluster Security



FEN: Front End Node CN: Cluster Node

Cluster Security: Auditing

- Submit modified to add logging at cluster sites
- Logs user name so **cluster admin can identify user**
- Logs executable path so admin can test whether unlicensed users run licensed software

Final format

Date	User Name	User ID	HUB Job ID	Executable path
Wed Aug 25 12:55:33 PDT 2010,	slatest,	92,	513,	/bgl/apps/VASP-5.2-Parallel/multiple-kpoint-version/vasp.5.2, llsubmit: The job "fengpfs.134776" has been submitted.

← Batch job ID

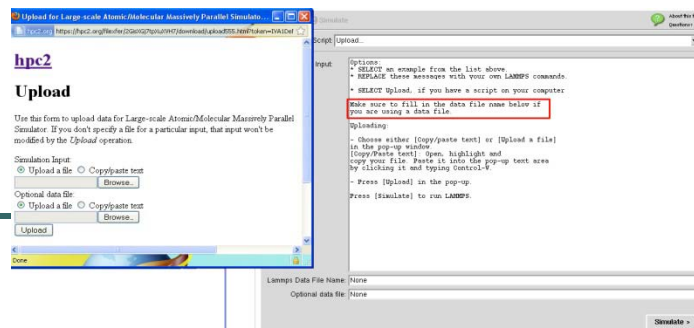
Big Data: Rappture Limitations

Rappture Upload Max Data Size: **400 MB**

Rappture Download Max File Size: **121 MB**

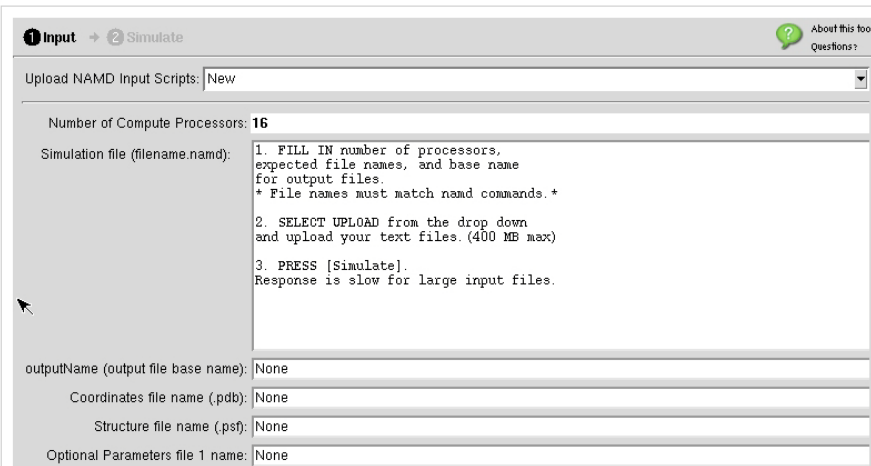
Protocol: **POST**

1. For binary input files, necessary for user to find a MIME encoder and preprocess files.
 - File gets inserted into XML wrapper at upload
2. User must type input file name(s)
 - inner wrapper copies file into run directory and has to create with the file name simulation input script expects.

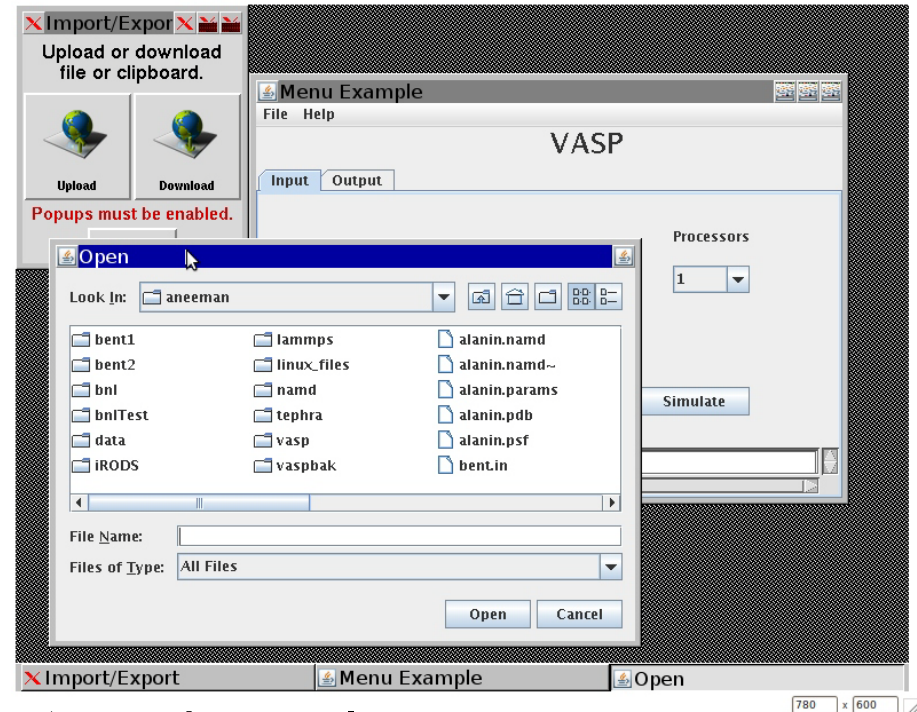


Workaround: sftp/webdav + Java Interface

Rappture



- ◆ User has to type in file names/paths
 - ✧ More work
 - ✧ Easy to make errors
 - ✧ No feedback until run aborts



- ◆ Traditional Menu
 - ◆ File browser prevents errors
 - ◆ Quick feedback
 - ✧ Displays imported files
 - ✧ Popup dialog guides the user
- ◆ Input file check before job run

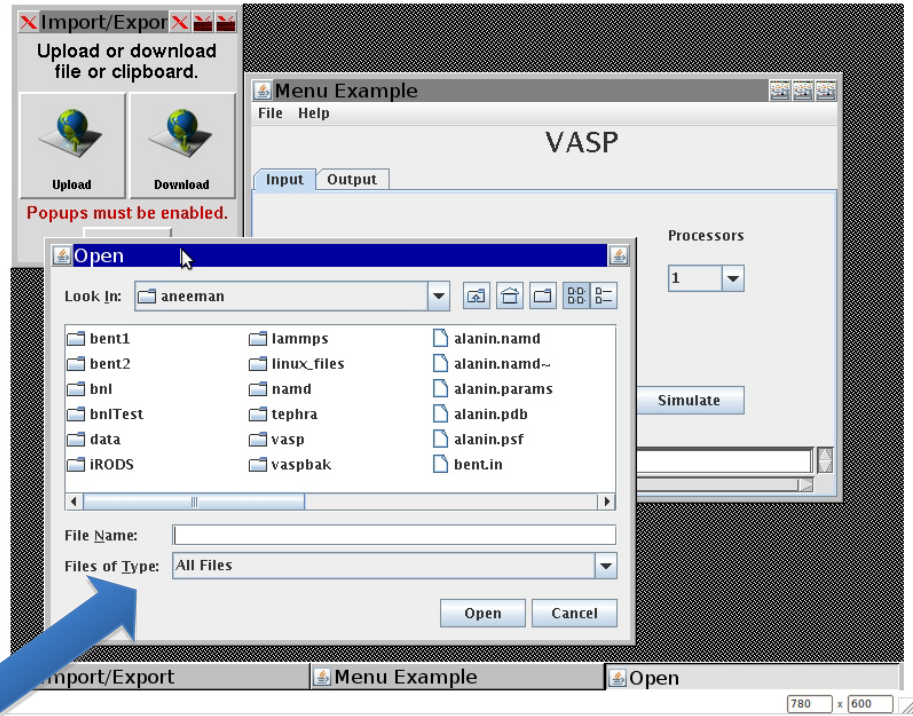
Three ways to get files into simulator

1. Uploader

- Nice file browser
- Limited file size

2. SFTP

- Any size file, any format (binary or text)
- Robust
- After upload, browse for file name to import to run directory



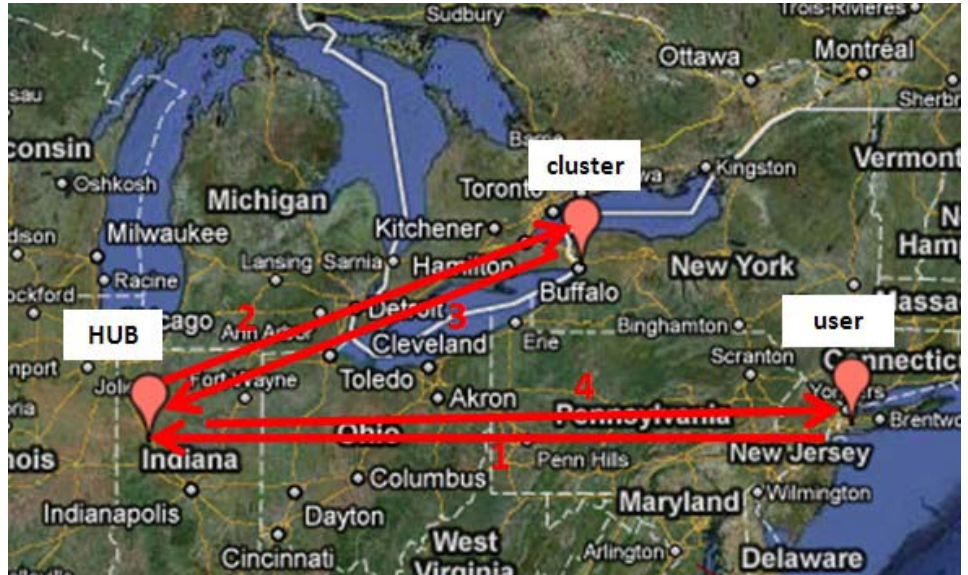
3. Use results from previous run

- Browse for file name to import to run directory

Data Grid Motivation

- Provide more storage at less cost
- Reduce wait time for run to start
- Share large datasets easily

1. User uploads input files to HUB.
2. Submit transfers input to cluster.
3. Submit transfers output to HUB.
4. User downloads files from HUB.



File Transfers for Portal Job



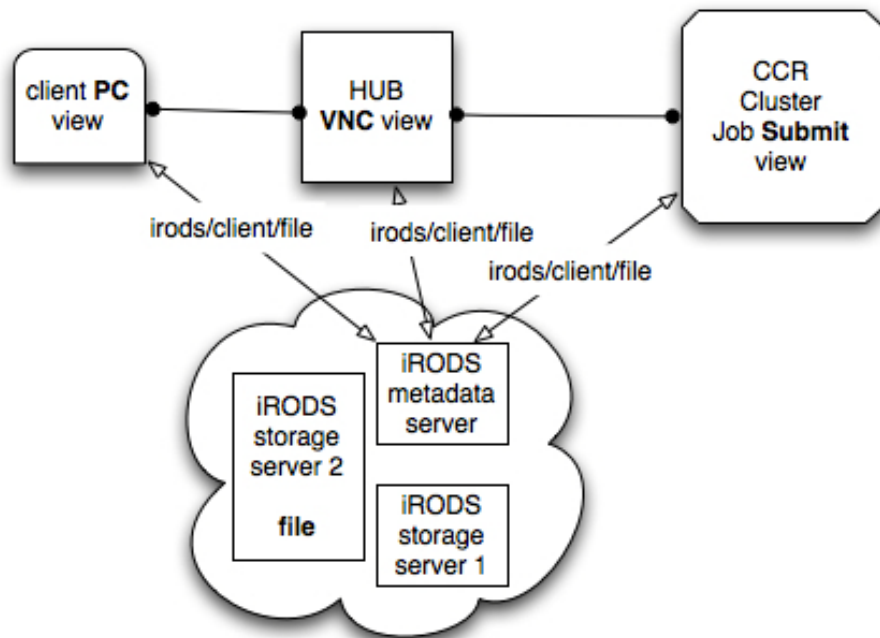
iRule Oriented
Data Systems

iRODS

FUSE

-
- Scalable, virtualized, distributed data system
 - Data Intensive Cyber Environments (DICE) research group headed by Dr. Reagan Moore, UNC Chapel Hill
 - Data Servers, metadata server
 - Emphasis: high performance data transfer
 - FUSE (file system in user space) iRODS client
 - Linux kernel module captures file system calls
 - Connects over network to iRODS servers to fetch/send data, but appears as local mount
 - FUSE Clients deployed on desktop, HUB, and cluster

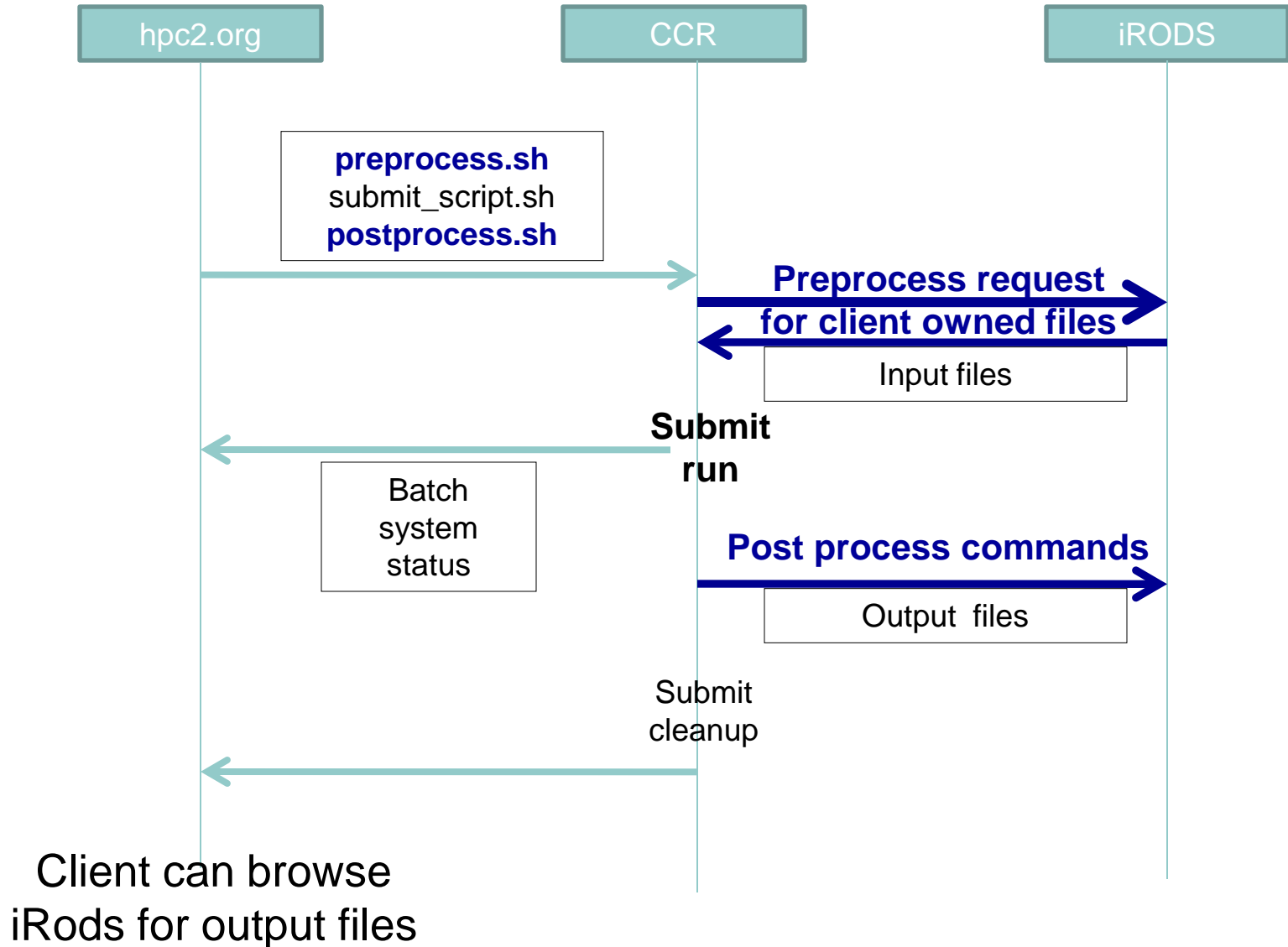
Submit and Data Grid Interaction



File Transfers using Data Grid

1. Client drag-and-drops file into iRODS
 2. Client can browse iRODS via HUB GUI for files (path saved)
 3. Submit tool, **running at cluster**, accesses user's files in iRODS (on behalf of the user)
- **Identical path views of a file needed in iRODS from client desktop, HUB, and compute cluster**

Submit Modifications



HUB side modifications

- Install FUSE kernel module and iRODS commands into path
- Automate mount of iRODS for users
 - VNC container caches mount state!!
- Coming this summer: Automate
 - iRODS user registration
 - creation of user home directories
 - User's irods home directory co-owned by vhub community account

Conclusions

- HUBzero computing IS grid computing
- With the correct setup and new modifications, the security model works well for regional grid
- HUB has limited capacity for data handling but possible to extend. More research needed.
- Interface frustrating for sophisticated users
- But with effort, shortcomings can be surmounted



Improved Simulation Through Collaboration

Rensselaer Polytechnic Institute • Stony Brook University • University at Buffalo • Brookhaven National Laboratory • NYSERNet

Thanks!

Please check out <http://hpc2.org> and <http://vhub.org>

- This work was funded by the New York State Foundation for Science, Technology and Innovation (NYSTAR).
- Thanks to Matt Jones for many useful discussions
- Special thanks to Steven Clark, Nick Kisseberth, Derrick Kearney, Mike McLennan and the HUBzero staff.



Submit Modifications

