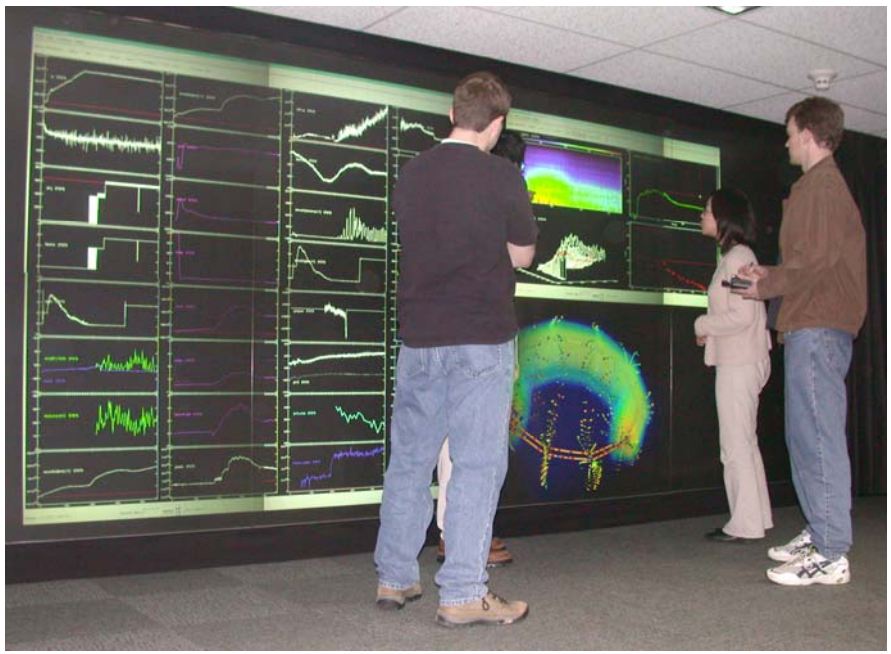
Remote Computing on the National Fusion Grid

Presented by Justin Burruss

4th IAEA Technical Meeting on Control, Data Acquisition, and Remote Participation
San Diego, CA July 2003



FusionGRID
www.fusiongrid.org

Overview

- Grid computing reduces the cost of computing resources
- Grid computing simplifies code maintenance and deployment
- The National Fusion Collaboratory implemented grid computing on the National Fusion Grid using Globus and Akenti
- The TRANSP transport analysis code has been installed as a grid service on the National Fusion Grid
- Other services will be added

FusionGRID
www.fusiongrid.org

Grid computing reduces the cost of computing resources

- In a traditional computing environment, each site has their own set of computers
  - Programs are installed on these local computers
  - If you need more computing power, must buy more computers
- With grid computing, resources can be shared
  - Can share hardware and software
  - Maintenance can be centralized
  - No software porting required
  - Instant software deployment
- Note: not CPU scavenging/SETI@home

FusionGRID
www.fusiongrid.org

A grid environment presents the user with a higher level of abstraction

- Heterogeneous network abstracted into a grid
- User signs on to grid once, not each host in a grid
  - A “single sign-on”
- Applications and other computing resources abstracted into services.
- User uses services without concern to details
- Don’t need username/password for each host
- Don’t even need to know that service may run on a remote host

FusionGRID
www.fusiongrid.org

Single sign-on is accomplished through certificate-based authentication

- Users need to have a single identity on a grid
- This single identity is implemented though digital certificates
- Certificate is a public key plus some other info digitally sited by a Certification Authority (CA) to assure authenticity
- This CA serves as the authority for the authenticity of every certificate on a grid
- All grid hosts must recognize a common CA

FusionGRID
www.fusiongrid.org

Akenti security engine takes care of authorization

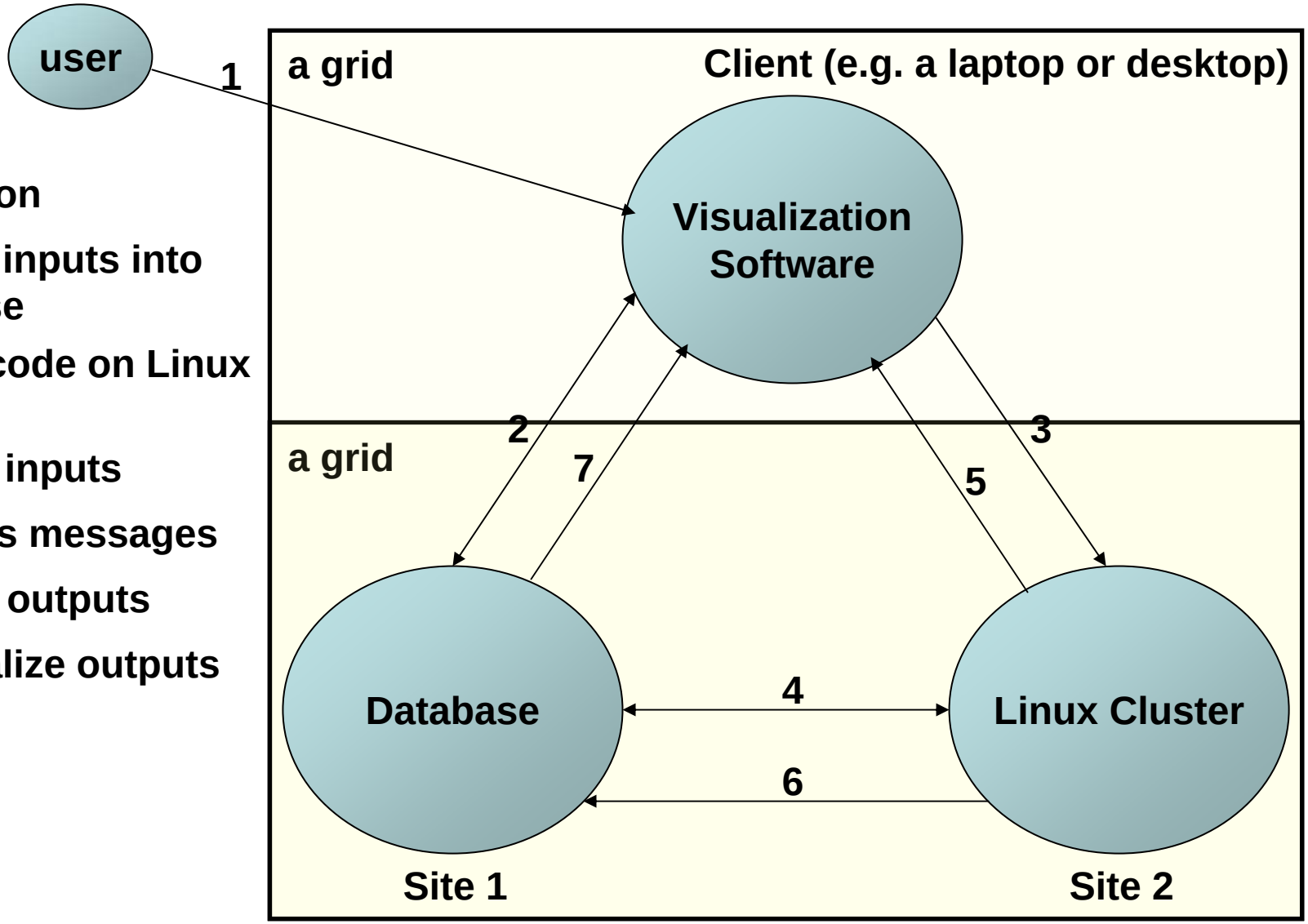
- **Akenti understands digital certificates**
- **Allows for very detailed resource usage requirements**
- **Supports the kind of distributed access control essential for a grid environment.**
- **Note: authorization is not the same thing as authentication**

	Answers the question	**Analogy**
Authentication	Who are you?	Passport
Authorization	Do you have permission?	Visa

FusionGRID
www.fusiongrid.org

Example outline of running a grid code

1. Sign on
2. Load inputs into database
3. Run code on Linux Cluster
4. Read inputs
5. Status messages
6. Write outputs
7. Visualize outputs



The National Fusion Collaboratory used the Globus Toolkit and Akenti to build the National Fusion Grid

- Globus Toolkit (GT) offers programs and protocols for security
- GT uses X.509 certificates for authentication
- GT also takes care of remote invocation
- GT resource discovery not needed
- The Collaboratory selected Akenti for authorization
- Implemented custom resource monitoring system

FusionGRID
www.fusiongrid.org

The first service on the National Fusion Grid is TRANSP

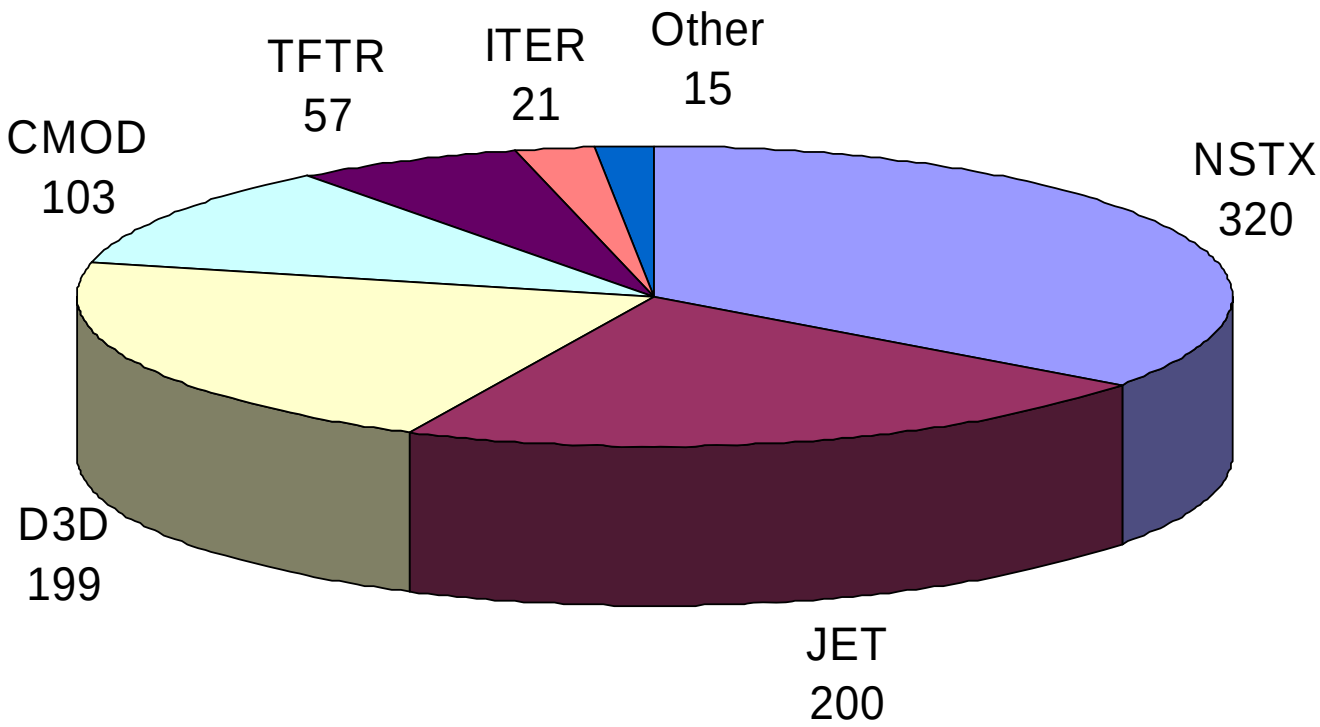
- TRANSP is a one million line transport code from PPPL
- GA used to spend 3 programmer months a year maintaining a local copy
- Now runs as a grid service on the National Fusion Grid
- Administration centralized at PPPL
- No porting
- Zero deployment time

Researchers are more productive with grid TRANSP

- **TRANSP runs on a Linux cluster at PPPL**
- **Each run executes 4-5 times faster**
- **There are many nodes, so runs can execute in parallel**
- **As a result, researchers are no longer limited by computing power and can be much more productive**

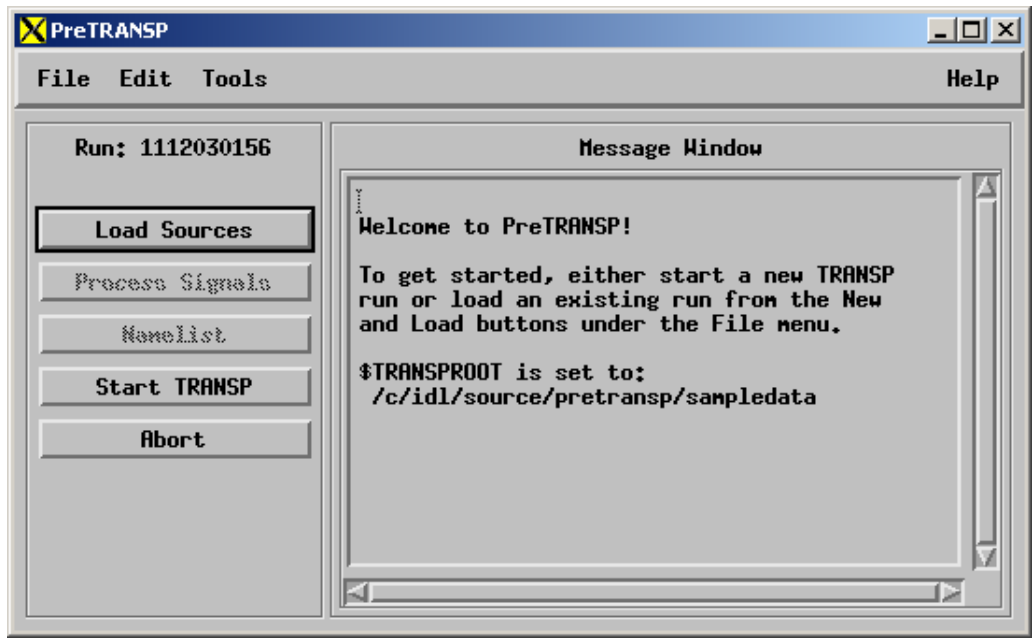
FusionGRID
www.fusiongrid.org

Over 900 Grid-enabled TRANSP runs (Oct 02 – Jun 03)



PreTRANSP simplifies the process of starting TRANSP

- IDL-based GUI application
- Manages code runs using a code run database
- Loads inputs
- Simple start button launches TRANSP



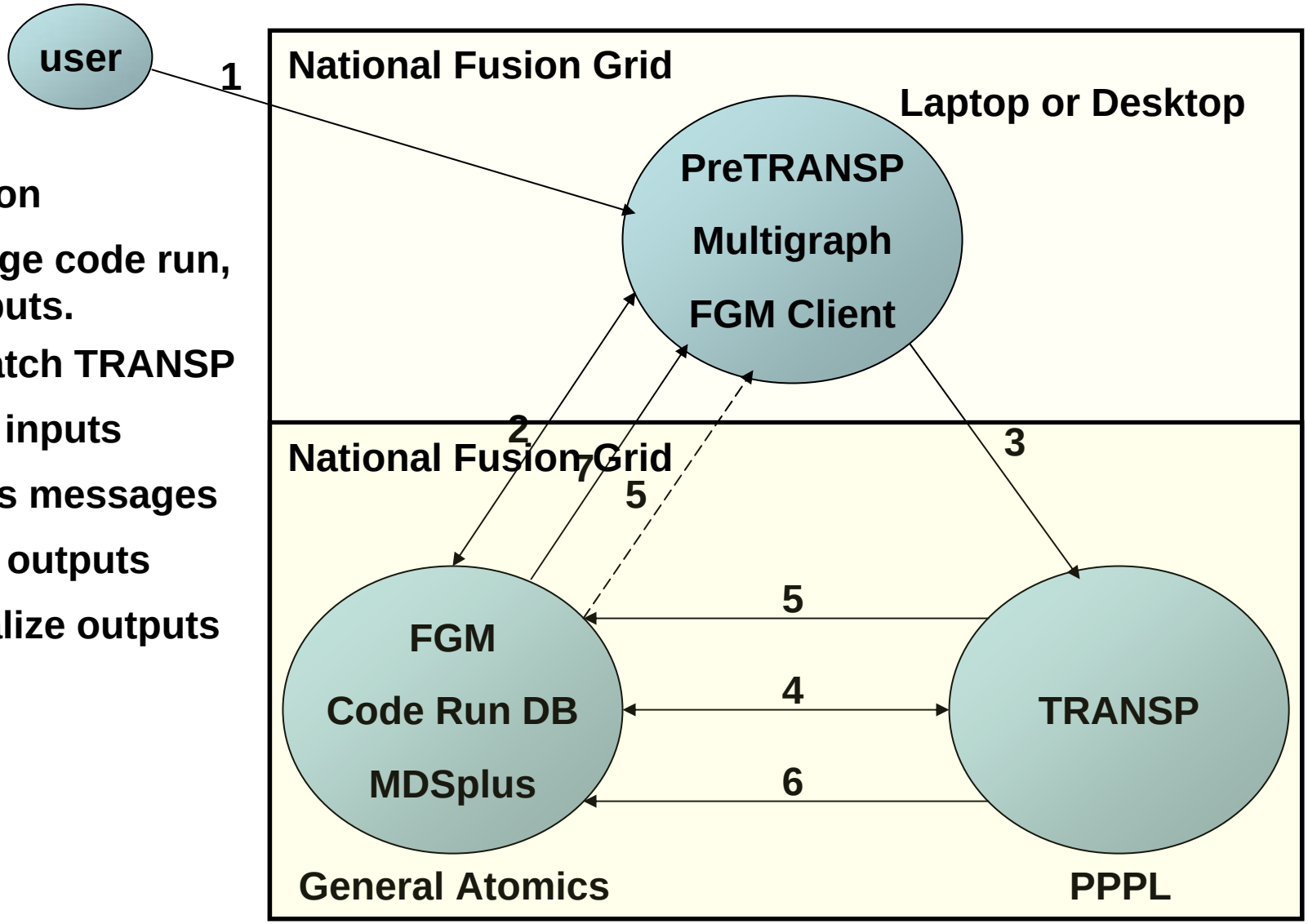
The Fusion Grid Monitor (FGM) is used to monitor TRANSP

- Monitoring info posted to FGM server and written to database
- Users view info through simple web interface
- Client web browsers updated automatically through server-push
- Will be used to monitor other services when they come online
- See Sean Flanagan's poster for details

FusionGRID
www.fusiongrid.org

How everything fits together to run TRANSP

1. Sign on
2. Manage code run, load inputs.
3. Dispatch TRANSP
4. Read inputs
5. Status messages
6. Write outputs
7. Visualize outputs



Lessons learned

- Certificate management is a hassle
  - Need to export from browser, convert to right format, install manually on each client machine
  - No buffer between time old certificate expires and renewed certificate becomes valid

- Non-routable IP address cause problems (e.g. private IP, NAT)

Firewalls and Globus don’t mix

- Underlying problem is that Globus and Firewalls do not work together
  - Globus may need hundreds of open ports
  - No way to change firewall configuration on the fly
  - If firewall was dynamic, might open ports without human intervention
- What happens is that Globus tries to open stdout/stdin through ports that may be blocked by a firewall
- TRANSP workaround was to redirect stdout/stdin to a file, then copy file to remote host.
- A general solution is needed

FusionGRID
www.fusiongrid.org

New developments & future work

- GS2 is being tested on the National Fusion Grid
  - GS2 is a microturbulence code
  - Next service to be added to the National Fusion Grid
  - Runs on 24-node Linux cluster at University of Maryland
  - Already 10 active GS2 users

- DOE Grids CA and European Data Grid (EDG) agreement
  - Mutual agreement to recognize each other's certificates
  - Necessary step towards international collaborations

FusionGRID
www.fusiongrid.org

Summary slide

- **Grid computing reduces the cost of computing resources**
- **Grid computing simplifies code maintenance and deployment**
- **The National Fusion Collaboratory implemented grid computing on the National Fusion Grid**
- **The TRANSP transport analysis code has been installed as a grid service on the National Fusion Grid**
- **Other services will be added**

FusionGRID
www.fusiongrid.org