Grid Computing & the Open Grid Services Architecture

Ian Foster
Argonne National Laboratory
University of Chicago
Globus Project

Open Group Grid Conference, Boston, July 21, 2003
Is the Grid ...

a) A collaboration & resource sharing infrastructure for scientific applications?
b) A standards-based distributed service integration & management technology?
c) A disruptive technology that enables a virtualized, collaborative, distributed world?
d) An open source technology & community
e) An over-used marketing slogan?
f) All of the above?
Grid Past, Present, Future

• **Past**
  - Origins and broad adoption in eScience, fueled by open source Globus Toolkit

• **Present**
  - Rapidly growing commercial adoption focused on intra-enterprise resource sharing
    - Open Grid Services Architecture (OGSA)

• **Future**
  - Key enabler of new applications & industries based on resource virtualization and distributed service integration
Why You Should Care

1) Grids address pain points now, e.g.
   - Cost of provisioning for peak demand
   - Data federation and integration

2) Grids are a disruptive technology
   - Usher in (and address problems of) a virtualized, collaborative, distributed world

3) An open Grid is to your advantage
   - Insist that your suppliers embrace OGSA, refuse proprietary solutions!
Why the Grid? Origins: Revolution in Science

- **Pre-Internet**
  - Theorize &/or experiment, alone or in small teams; publish paper

- **Post-Internet**
  - Construct and mine large databases of observational or simulation data
  - Develop simulations & analyses
  - Access specialized devices remotely
  - Exchange information within distributed multidisciplinary teams
NEESgrid Earthquake Engineering Collaboratory

Remote Users
(Faculty, Students, Practitioners)

Instrumented Structures and Sites

Laboratory Equipment

High-Performance Network(s)

Field Equipment

Remote Users:
(K-12 Faculty and Students)

Instrumented Structures and Sites

Laboratory Equipment

Curated Data Repository

Global Connections (fully developed FY 2005 – FY 2014)

U. Nevada Reno

www.neesgrid.org
Why the Grid?
New Driver: Revolution in Business

- **Pre-Internet**
  - Central data processing facility

- **Post-Internet**
  - Enterprise computing is highly distributed, heterogeneous, inter-enterprise (B2B)
  - Business processes increasingly computing- & data-rich
  - Outsourcing becomes feasible => service providers of various sorts
  - Growing complexity & need for more efficient management
The New Enterprise Computing Environment

**Application Servers:**
- Meterable Services
- Robust Power
- Content hosting
- Industry applications (ERP, SCM, CRM, e-commerce, data warehouse)

**Edge Servers:**
- Transcoding
- Caching
- Acceleration
- Distribution
- Security
- Directories
- Quality of Service

GTO2000: IBM Research

foster@mcs.anl.gov

ARGONNE  Chicago
Common eScience/eBusiness Requirements

• Dynamically link resources/services
  – From collaborators, customers, eUtilities, ...
  (members of evolving “virtual organization”)

• Into a “virtual computing system”
  – Dynamic, multi-faceted system spanning institutions and industries
  – Configured to meet instantaneous needs, for:

• Multi-faceted QoX for demanding workloads
  – Security, performance, reliability, ...
Grid Technologies Address these Requirements

- Infrastructure ("middleware") for establishing, managing, and evolving multi-organizational federations
  - Dynamic, autonomous, domain independent
  - On-demand, ubiquitous access to computing, data, and services

- Mechanisms for creating and managing workflow within such federations
  - New capabilities constructed dynamically and transparently from distributed services
  - Service-oriented, virtualization
Grids: Resource/Service Virtualization

- Discovery of resources/services with required capabilities and availability
  - Amount of RAM/storage/MFLOPS, # of CPUs, bandwidth, software, ... etc.
- Use of actual resources is “virtualized”
- All part of QoS negotiation ...

foster@mcs.anl.gov
Grids: Resource/Service Integration

- Discovery, composition, management
- Matching of available time-windows for: Data source, bandwidth, input/output, storage allocation, CPU cycles, ..., etc.
- All part of QoS negotiation ...

foster@mcs.anl.gov
### Available Systems

<table>
<thead>
<tr>
<th>Dept</th>
<th>System/Processors</th>
<th>Peak GFLOPs</th>
<th>Memory Gigabytes</th>
<th>Work Disk Gigabytes</th>
<th>Name</th>
<th>Grid</th>
<th>SW</th>
<th>Network Status</th>
<th>Load</th>
<th>Jobs</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>Linux PC</td>
<td>1.5</td>
<td>1</td>
<td>62</td>
<td>alta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>Linux PC</td>
<td>1.5</td>
<td>1</td>
<td>52</td>
<td>solitude</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>Cray SV1 / 16</td>
<td>19</td>
<td>16</td>
<td>480</td>
<td>aurora</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>T1</td>
</tr>
<tr>
<td>TACC</td>
<td>Linux Cluster / 2</td>
<td>1</td>
<td>6</td>
<td>13</td>
<td>brazos</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>Linux PC</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>cool</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>IBM Regatta-HPC / 64</td>
<td>333</td>
<td>128</td>
<td>532</td>
<td>longhorn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>LSF Multi-Cluster / 22</td>
<td>37</td>
<td>14</td>
<td>173</td>
<td>laf</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4R-4Q</td>
</tr>
<tr>
<td>TACC</td>
<td>Linux Cluster / 4</td>
<td>2</td>
<td>1</td>
<td>13</td>
<td>padre</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>Cray/Dell Cluster / 4</td>
<td>19</td>
<td>8</td>
<td>8</td>
<td>q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>Linux PC</td>
<td>2</td>
<td>1</td>
<td>10</td>
<td>sanantonio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>IBM IA-64 Cluster / 40</td>
<td>128</td>
<td>80</td>
<td>140</td>
<td>sanantonio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>Sun Workstation</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>tahoka</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TACC</td>
<td>IBM IA-32 Cluster / 64</td>
<td>64</td>
<td>32</td>
<td>20</td>
<td>texas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6R-4Q-3Q</td>
</tr>
</tbody>
</table>

Total: 627 280 1581

Click on column headers to sort.

Click the magnifying glass icon for more information about grid software status or network connectivity.
Platform Symphony: Real-Time Online Processing

Applications: Delivery

Application Services: Distribution

Servers: Execution

Application Virtualization

• Automatically connect applications to services
• Dynamic & intelligent provisioning

Infrastructure Virtualization

• Dynamic & intelligent provisioning
• Automatic failover
Grids: Multiple Independent Organizations

- Each organization enforces its own access policy
- Identity federation + authorization assertions
- Trusted third parties
- All part of QoS/QoP negotiation...

foster@mcs.anl.gov
Extending Symphony to a Cross-Institutional Grid

NY
Financial Institution – Insurance Group

UK
Financial Institution

NY
Financial Institution – Capital Markets Group
NASA: Aviation Safety

**Crew Capabilities**
- accuracy
- perception
- stamina
- re-action times
- SOPs

**Wing Models**
- Lift Capabilities
- Drag Capabilities
- Responsiveness

**Airframe Models**

**Stabilizer Models**
- Deflection capabilities
- Responsiveness

**Engine Models**
- Thrust performance
- Reverse Thrust performance
- Responsiveness
- Fuel Consumption

**Human Models**

**Landing Gear Models**
- Braking performance
- Steering capabilities
- Traction
- Dampening capabilities

**NASA: Aviation Safety**
Grids: A Possible (Likely?) End State
What is a Grid?

• Three key criteria:
  – Coordinates distributed resources ...
  – using standard, open, general-purpose protocols and interfaces ...
  – to deliver non-trivial qualities of service.

• What is not a Grid?
  – A cluster, a network attached storage device, a scientific instrument, a network, etc.
  – Each is an important component of a Grid, but by itself does not constitute a Grid.
The Grid World: Current Status

• Large number of Grid success stories
  – Many major projects in science
  – Growing number of commercial deployments

• Open source Globus Toolkit® a de facto standard for major protocols & services
  – Simple protocols & APIs for authentication, discovery, access, etc.: infrastructure
  – Information-centric design
  – Large user and developer base
  – Multiple commercial support providers

• Global Grid Forum: community & standards

• Emerging Open Grid Services Architecture
Grid Evolution: OGSA (Open Grid Services Architecture)

• Goals
  – Refactor Globus protocol suite to enable common base and expose key capabilities
  – Service orientation to virtualize resources and unify resources/services/information
  – Embrace key Web services standards, leverage commercial efforts

• Result = standard interfaces & behaviors for distributed system mgmt: the Grid Service
  – Standardization within Global Grid Forum
  – GT3 open source implementation

• OGSA = Web services on steroids!
Open Grid Services Infrastructure (OGSI)

- Service registry
- Service requestor (e.g. user application)
- Service factory
- Grid Service Handle
- Service data
- Keep-alives
- Notifications
- Service invocation
- Service discovery
- Service instances
- Service registration

Resource allocation

Interactions standardized using WSDL and SOAP

Authentication & Authorization are applied to all requests
OGSA Standardization & Implementation

• **OGSI** defines core interfaces and behaviors for manageable services

• Efforts are underway to define standards for
  - Agreement negotiation
  - Common management model
  - Data access and integration
  - Security and policy
  - Etc.

• Supported by strong open source technology & major commercial vendors
Hurdles for WS & OGSA to Overcome

• Industry and customer focus
  – Pass maturity point before next silver bullet...

• Standardization, standardization, standardization...
  – Interoperability, pluggability, replaceability, ...
  – Protocols and infrastructure services
  – Global Grid Forum

• “Unfriendly” licenses
  – IP issues can spoil everything
  – There is no money in middleware!
    (ubiquity is key to make money on the added value!)

• Learn from previous efforts
  – We are reinventing some wheels...
Why Grids will Succeed

• WS standards/interoperability issues
  – Too many options, too little time...

• Grid requirements >> WS requirements
  – OGSA is WS++; addresses key operational issues

• Global Grid Forum & Globus are very pragmatic

• Globus provides working, open source toolkit
  – Growing, global, demanding user community
  – Vendors can, do, and will use Globus Toolkit

Nothing “sells” better than working free code
Silver Bullet Hype-Curve...

Time

Success/Maturity/Acceptance

DCE

CORBA

WebService

OGSA + WebServices + Globus Toolkit
Recap: The Grid Is ...

a) A collaboration & resource sharing infrastructure for scientific applications

b) A standards-based distributed service integration & management technology (OGSA)

c) A disruptive technology that enables a virtualized, collaborative, distributed world

d) An open source technology & community (Globus Toolkit: “Linux for the Grid”)

e) An over-used marketing slogan
Grid Past, Present, Future

• Past
  – Origins and broad adoption in eScience, fueled by open source Globus Toolkit

• Present
  – Rapidly growing commercial adoption focused on intra-enterprise resource sharing
     – Open Grid Services Architecture (OGSA)

• Future
  – Key enabler of new applications & industries based on resource virtualization and distributed service integration
Why You Should Care

1) Grids address pain points now, e.g.
   - Cost of provisioning for peak demand
   - Data federation and integration

2) Grids are a disruptive technology
   - Usher in (and address problems of) a virtualized, collaborative, distributed world

3) An open Grid is to your advantage
   - Insist that your suppliers embrace OGSA, refuse proprietary solutions!
Summary

• Look beyond “The Grid” hype ...
  – A lot of good stuff—including working software
• Web Services are pretty basic
  – Look at the added value of OGSA
• Grid features: sophisticated plumbing + services
  – Great framework for your apps
  – Benefit from service and utility abstractions
  – Address challenging cross-domain issues
• Vendors are commercializing “The Grid” now
  – The “cool” ones (-:
• Get involved with the Global Grid Forum
For More Information

• The Globus Project™
  – www.globus.org

• Global Grid Forum
  – www.ggf.org

• Background information
  – www.mcs.anl.gov/~foster

• GlobusWORLD 2004
  – www.globusworld.org
  – Jan 20–23, San Fran

2nd Edition: November 2003