



the globus project™
www.globus.org

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

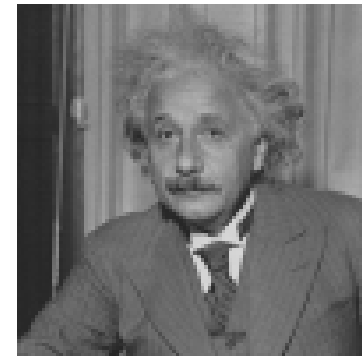
- 1) Grids address pain points now, e.g.
 - Cost of provisioning for peak demand
 - Data federation and integration
- 2) Grids are a disruptive technology
 - Usher in (and address problems of) a virtualized, collaborative, distributed world
- 3) An open Grid is to your advantage
 - Insist that your suppliers embrace OGSA, refuse proprietary solutions!

Why the Grid?

Origins: Revolution in Science

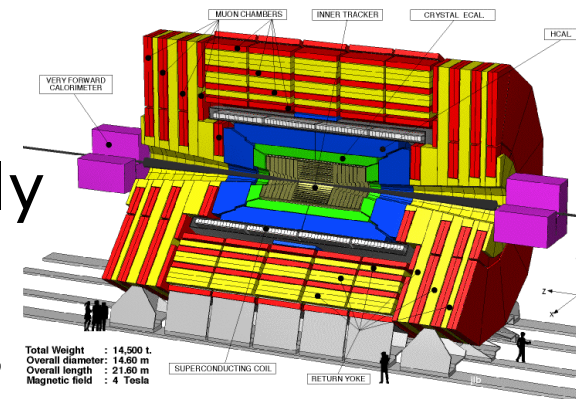
- Pre-Internet

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



- Post-Internet

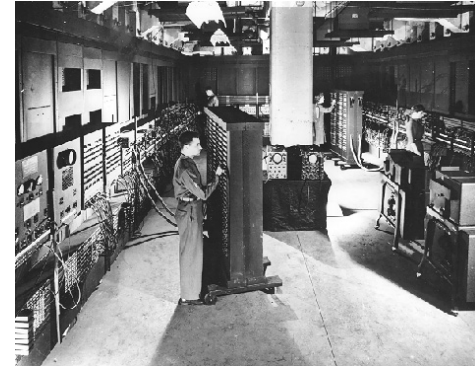
- Construct and mine large databases of observational or simulation data
- Develop simulations & analyses
- Access specialized devices remotely
- Exchange information within distributed multidisciplinary teams



Why the Grid?

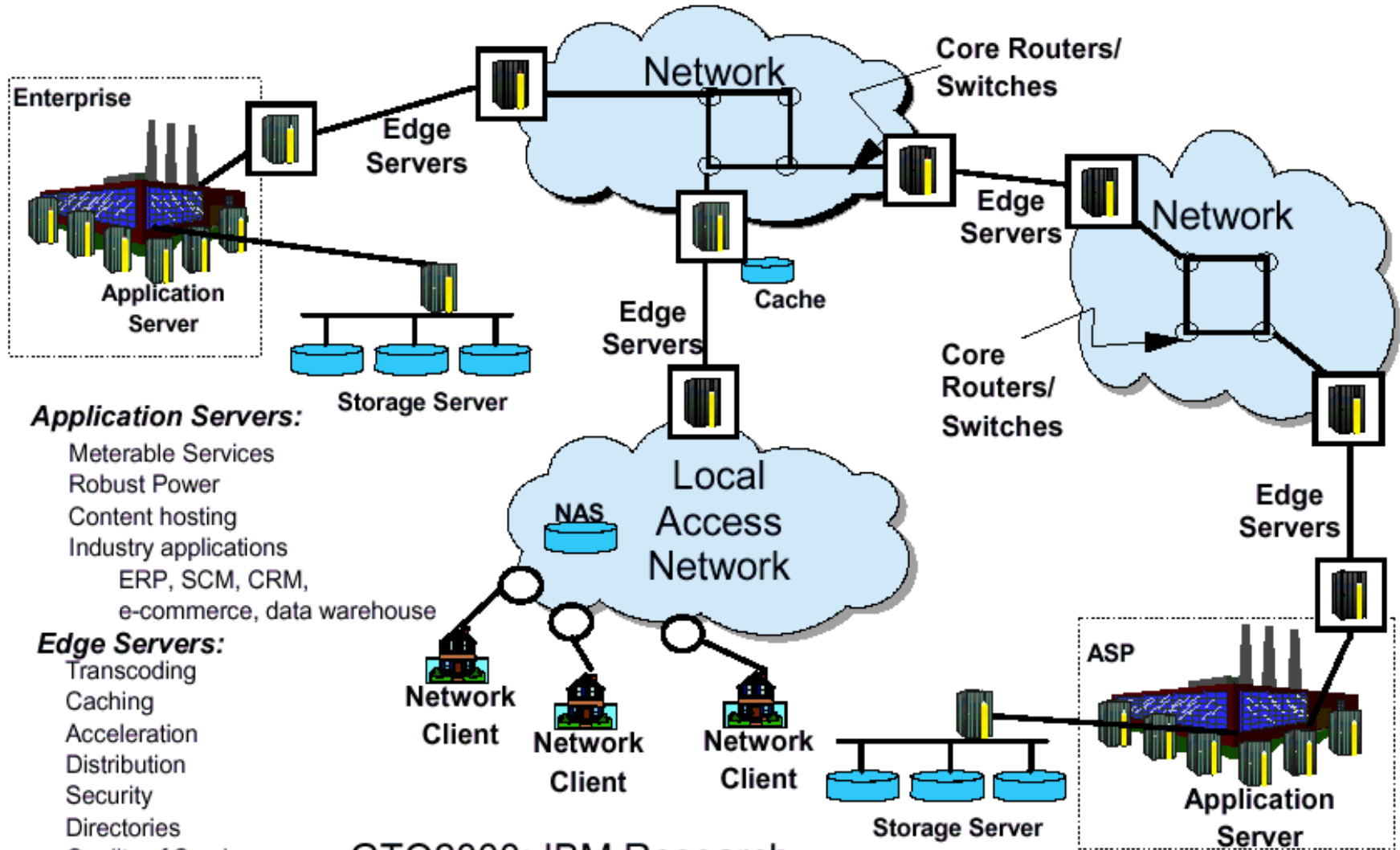
New Driver: Revolution in Business

- 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

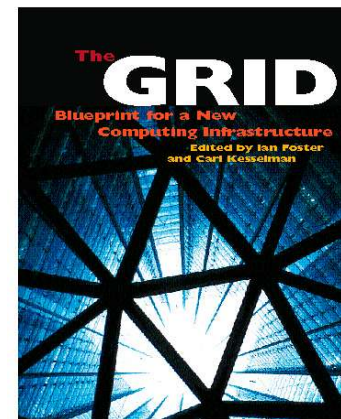
Common eScience/eBusiness Requirements

- 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, ...



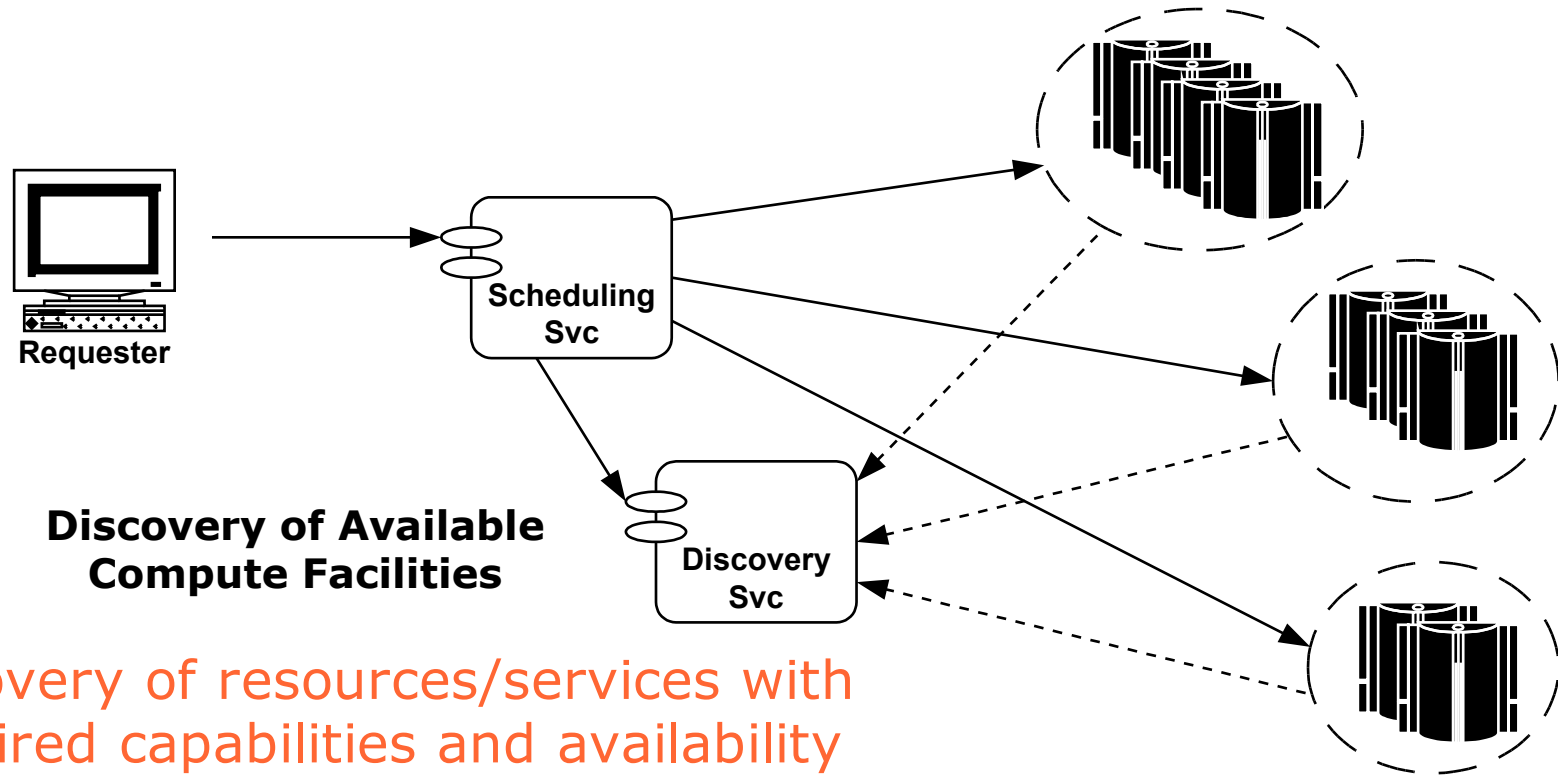
Grid Technologies

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

Resource/Service Virtualization

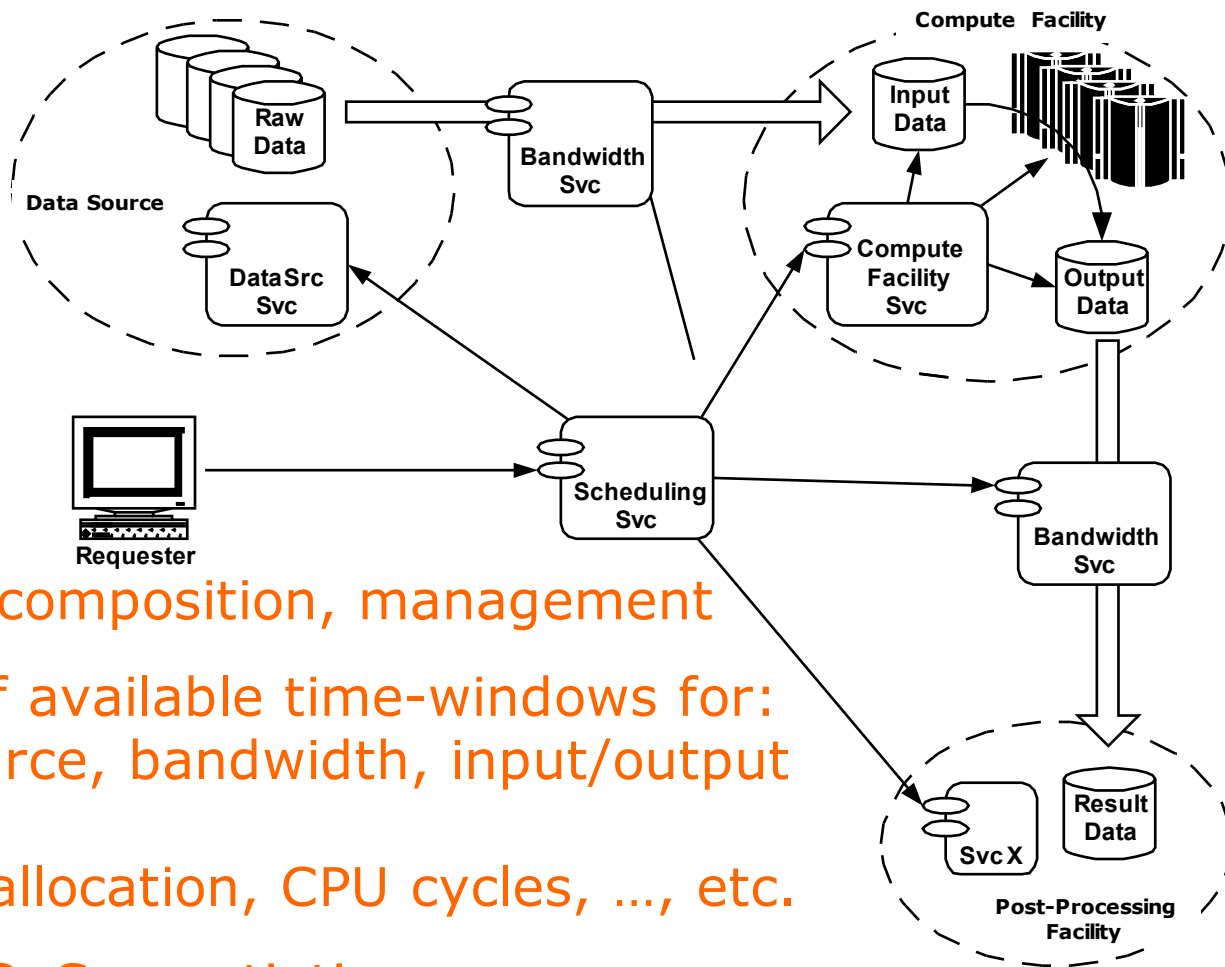


- Discovery of resources/services with required capabilities and availability

Amount of RAM/storage/MFLOPS, # of CPUs, bandwidth, software, ... etc.

- Use of actual resources is "virtualized"
- All part of QoS negotiation ...

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 ...



University of Texas at Austin Grid Computing Portal

Information

[Available Systems](#)

[Grid Status](#)

[Job Status](#)

File Manipulation

[List Remote Files](#)

[List Portal Files](#)

[File Upload](#)

[Transfer to Remote](#)

[Transfer to Portal](#)

[3rd Party Transfer](#)

Scientific Apps

[Seismic Application](#)

Demo Apps

[PI Demo](#)

[Log In](#)

| Deal | System/ Processors | Peak GFLOPs | Memory GBytes | Work Disk GBytes | Name | Grid SW | Network | Status | Load | Jobs |
|------|------------------------|----------------|------------------|------------------------|-------------|------------|---------|--------|------|----------|
| CS | Linux PC | 1.5 | .1 | 52 | alta | Q | | ↑ | | |
| CS | Linux PC | 1.5 | .1 | 52 | solitude | Q | | ↑ | | |
| TACC | Cray SV1 / 16 | 18 | 16 | 465 | aurora | Q | | ↑ | | 78 |
| TACC | Linux Cluster / 2 | 1 | 6 | 13 | braves | Q | | ↑ | | |
| TACC | Linux PC | 2 | 1 | 10 | cool | Q | Q | ↑ | | |
| TACC | IBM Regatta-HPC / 64 | 303 | 128 | 532 | longhorn | Q | | ↑ | | 4R-4Q |
| TACC | LSF Multi-Cluster 22 | 37 | 14 | 173 | isf | Q | | ↑ | | 0R-2Q-3Q |
| TACC | Linux Cluster / 4 | 2 | 1 | 13 | padre | Q | | ↑ | | |
| TACC | Cray/Dell Cluster / 4 | 19 | 8 | 8 | q | Q | | ↑ | | |
| TACC | Linux PC | 2 | 1 | 10 | sanantonio | Q | Q | ↑ | | |
| TACC | IBM IA-64 Cluster / 40 | 128 | 80 | 140 | santanita | Q | Q | ↑ | | |
| TACC | Sun Workstation | 2 | 1 | 2 | tahoka | Q | | ↑ | | |
| TACC | IBM IA-32 Cluster / 64 | 64 | 32 | 20 | tejas | Q | Q | ↑ | | 6R-4Q-3Q |
| TQAM | Alpha Cluster / 16 | 16 | 8 | 71 | zaphod | Q | | ↑ | | |
| | | Total: | 627 | 299 | 1581 | | | | | |

Click on column headers to sort.

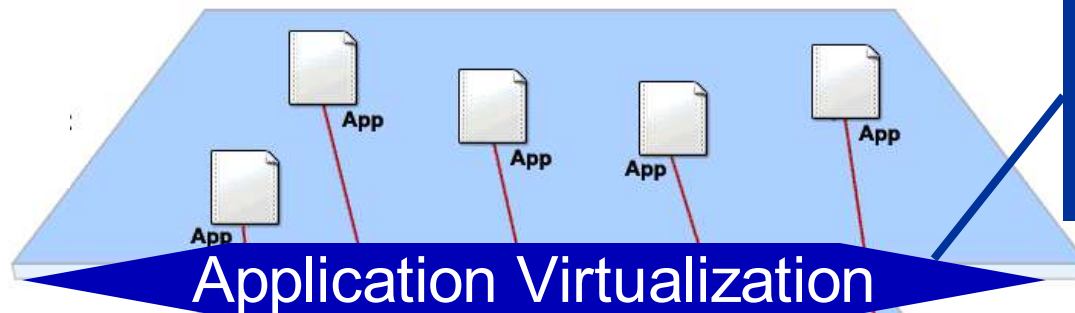
Click the magnifying glass icon for more information about grid software status or network connectivity





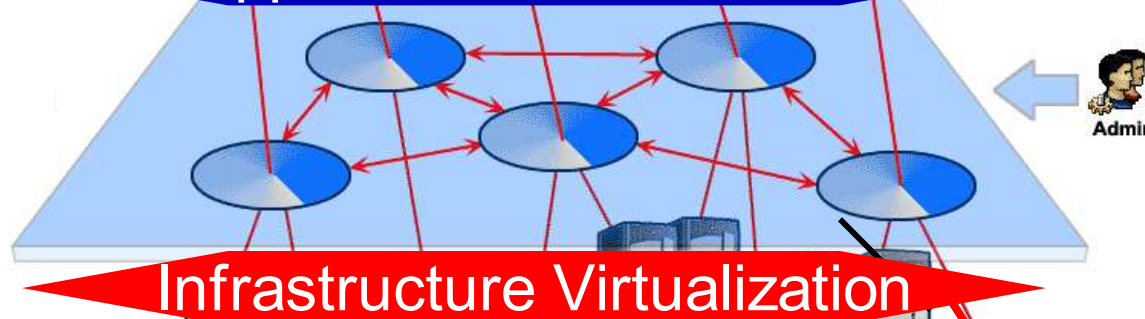
Platform Symphony: Real-Time Online Processing

**Applications:
Delivery**

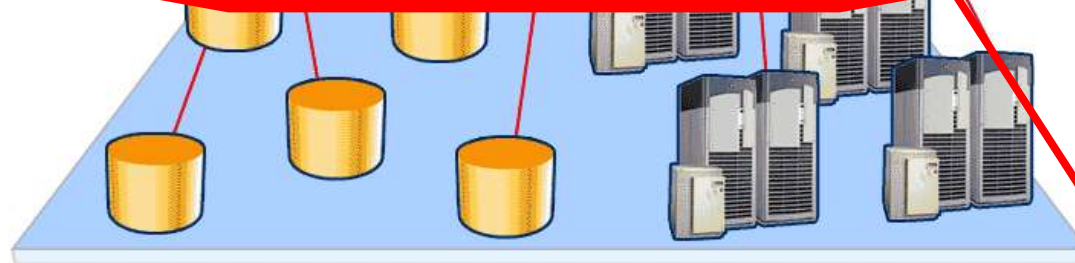


- Automatically connect applications to services
- Dynamic & intelligent provisioning

**Application
Services:
Distribution**

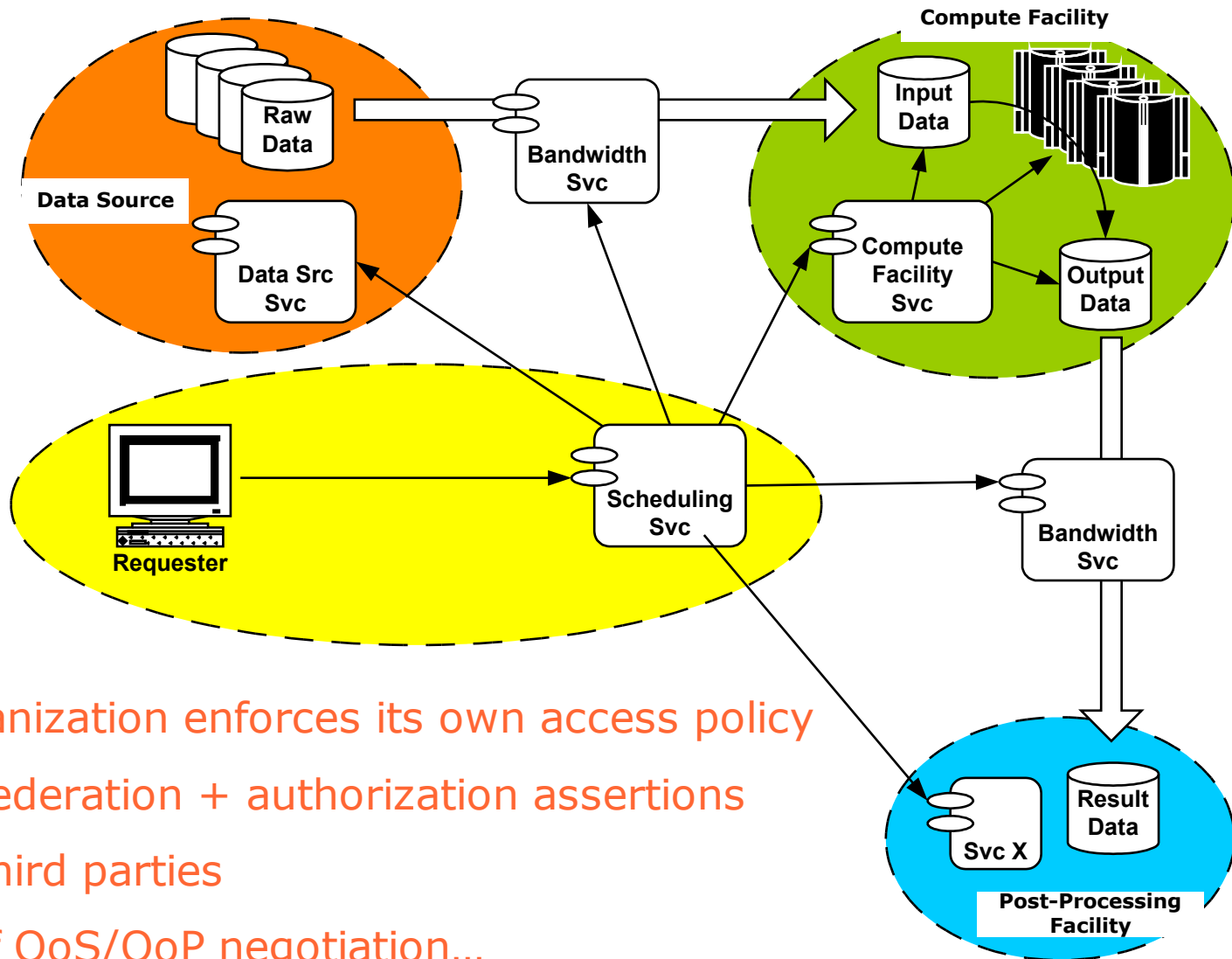


**Servers:
Execution**



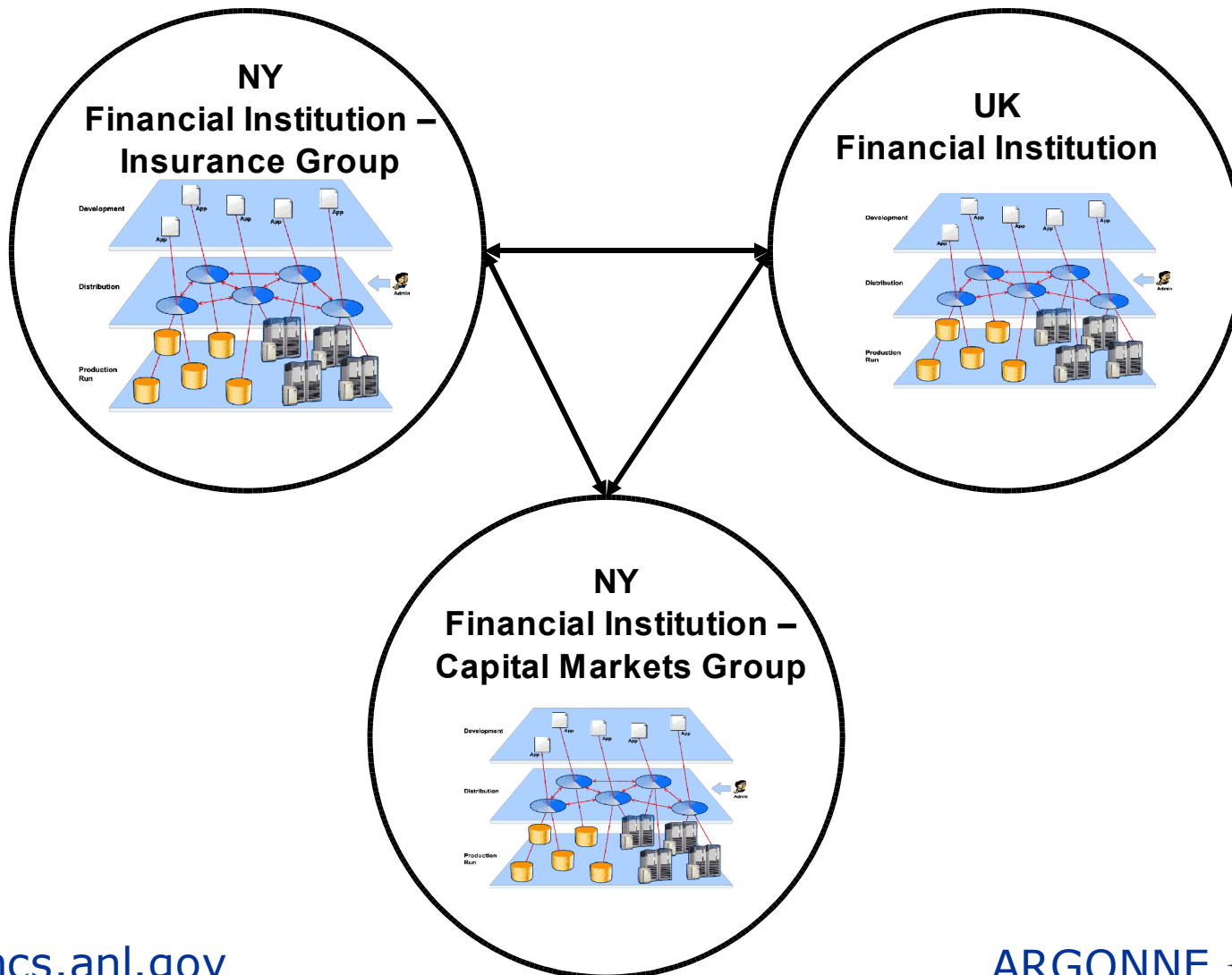
- Dynamic & intelligent provisioning
- Automatic failover

Grids: Multiple Independent Organizations



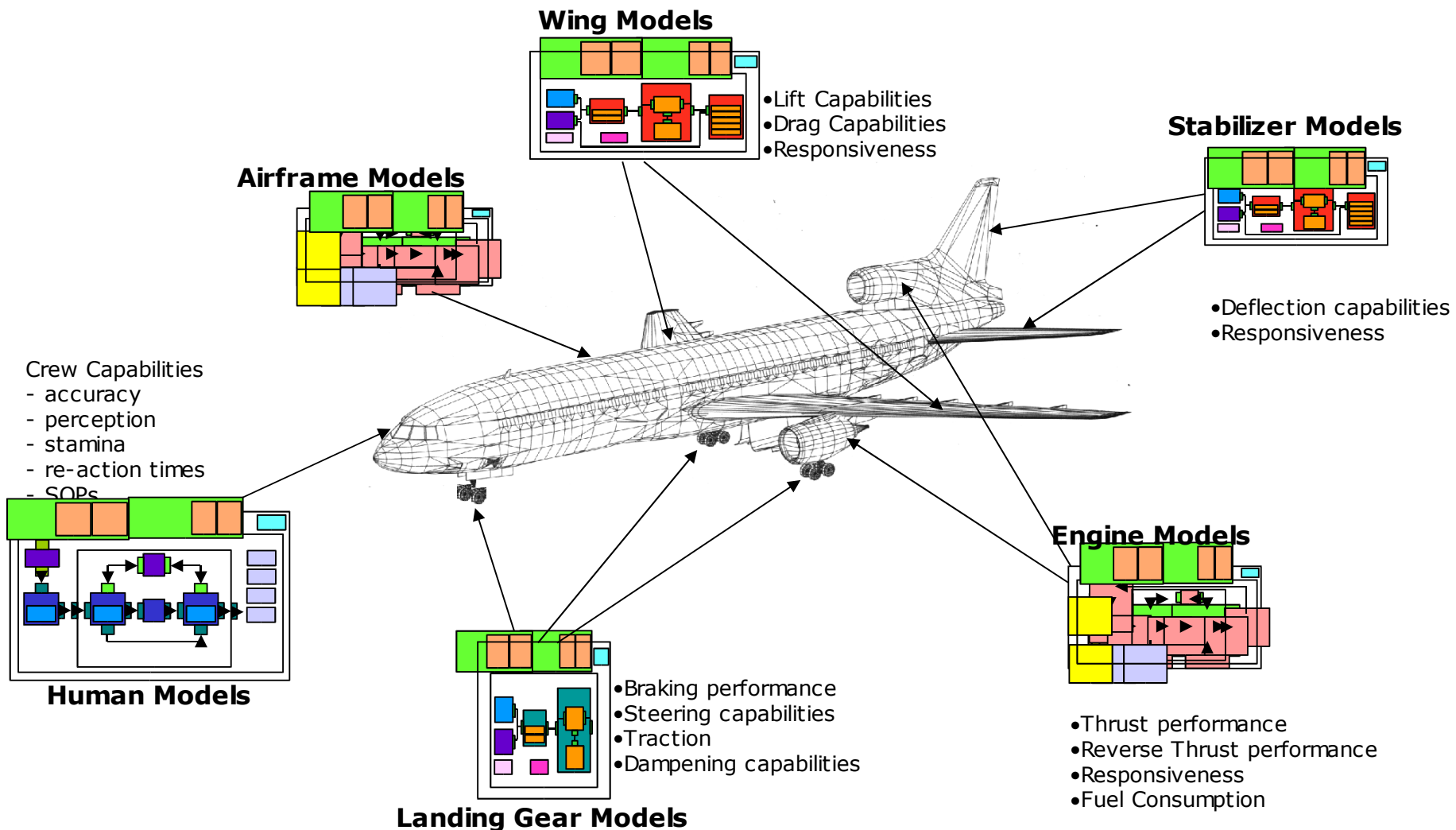
- Each organization enforces its own access policy
- Identity federation + authorization assertions
- Trusted third parties
- All part of QoS/QoP negotiation...

Extending Symphony to a Cross-Institutional Grid

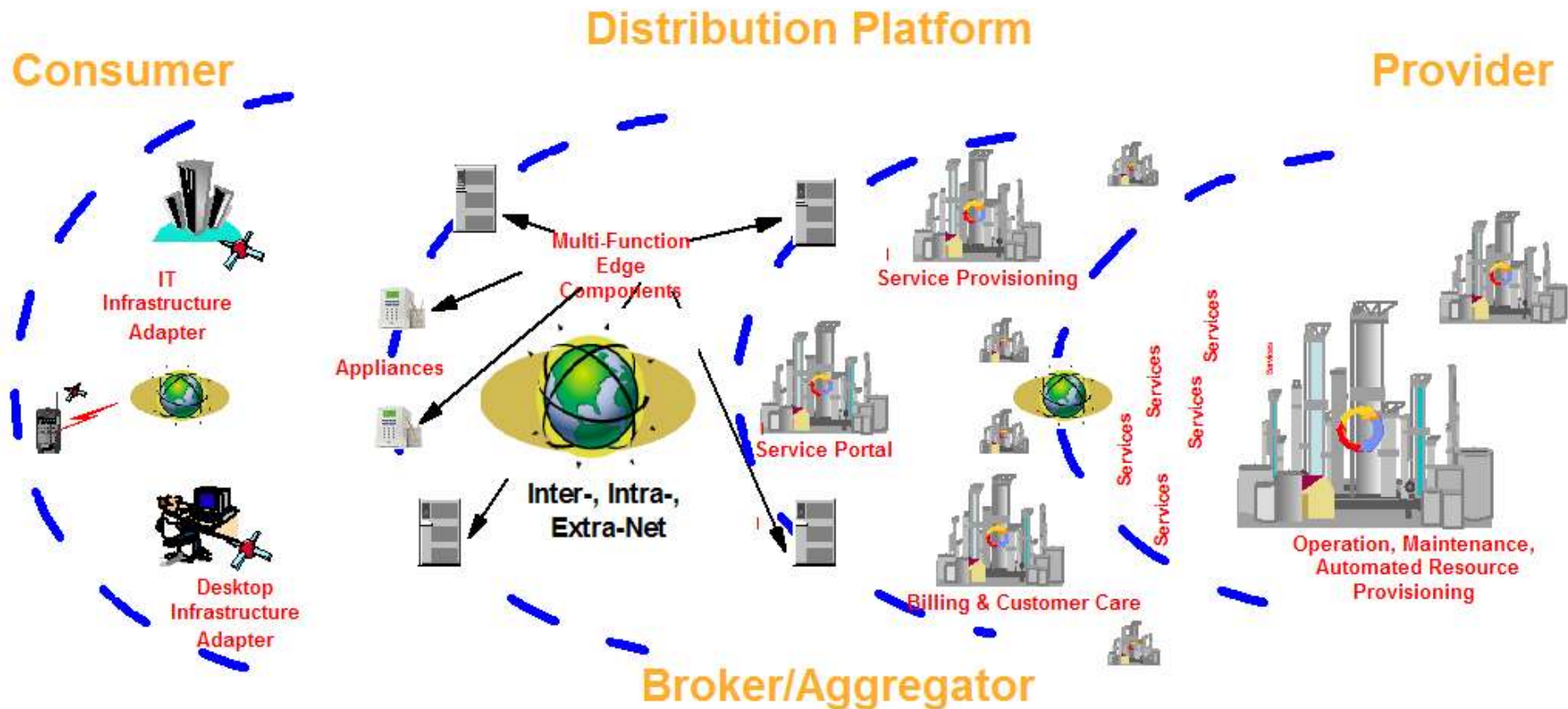




NASA: Aviation Safety



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

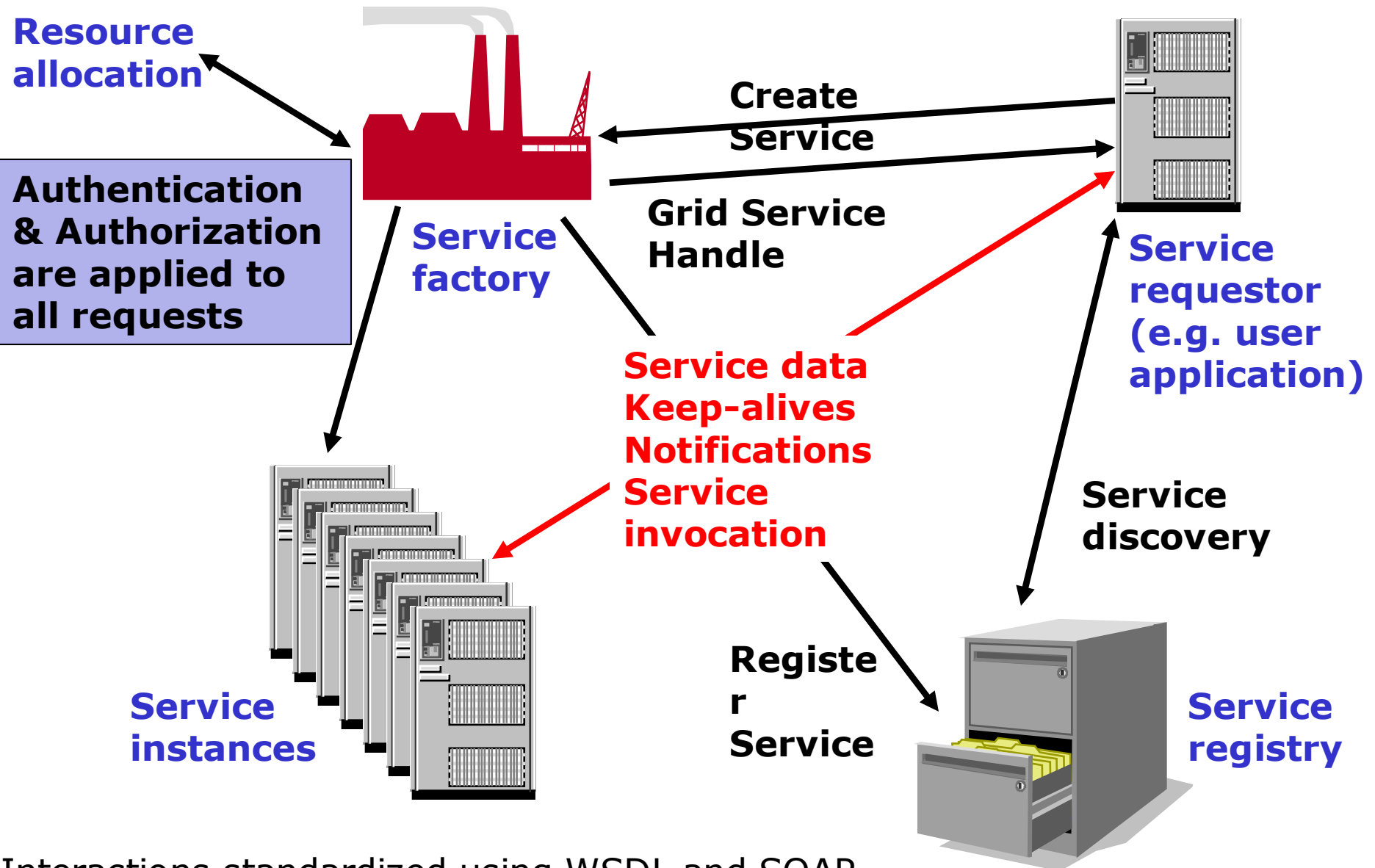
- Large number of Grid success stories
 - Many major projects in science
 - Growing number of commercial deployments
- Open source Globus Toolkit® a de facto standard for major protocols & services
 - Simple protocols & APIs for authentication, discovery, access, etc.: infrastructure
 - Information-centric 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)



Interactions standardized using WSDL and SOAP



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...

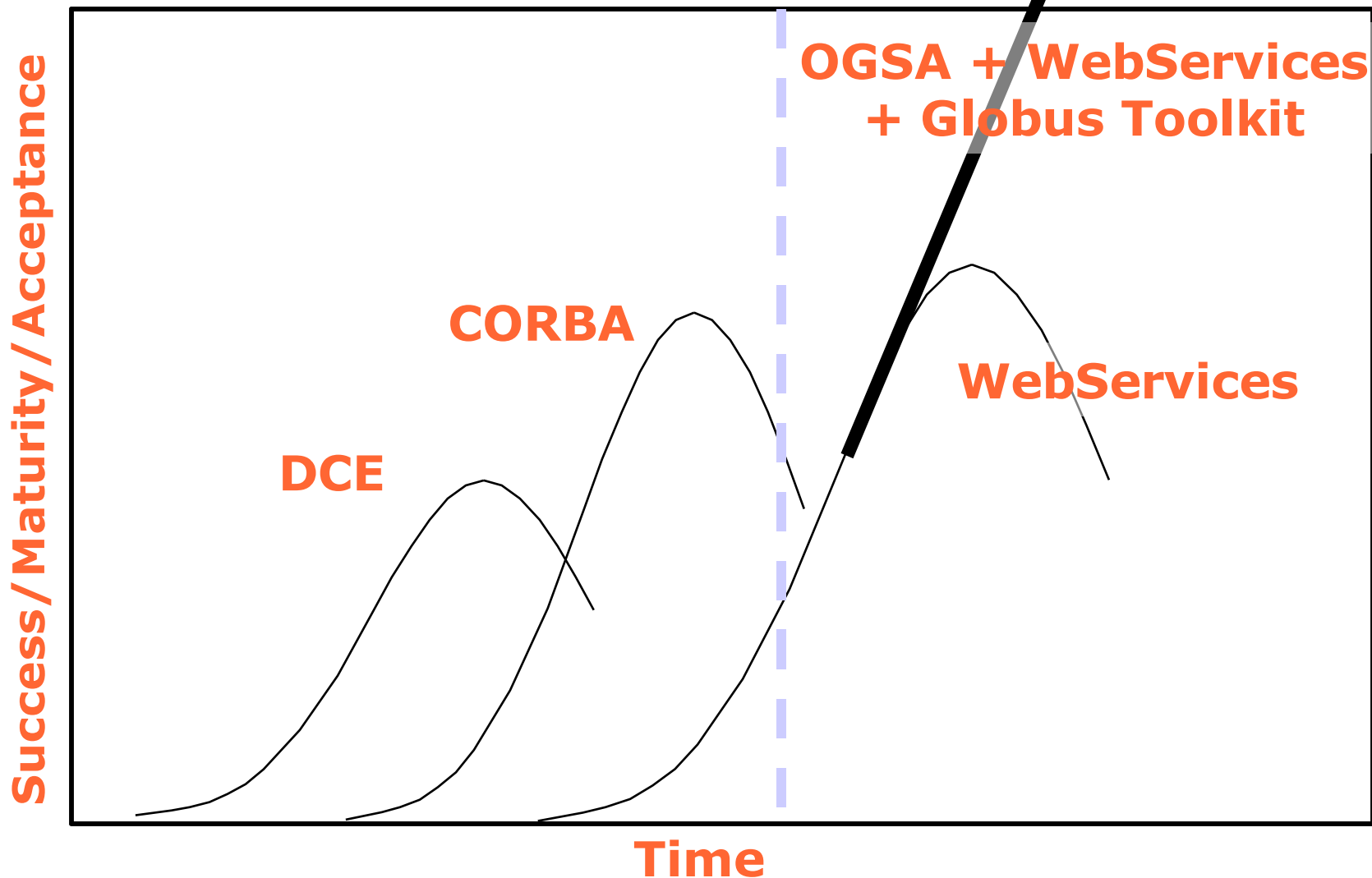
Why Grids will Succeed

- 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...



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**

