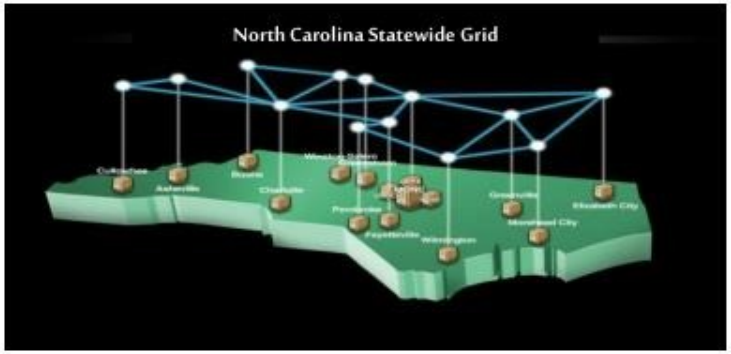
The Status of Grid Computing

NC State University, Jan 20, 2005



Wolfgang Gentzsch

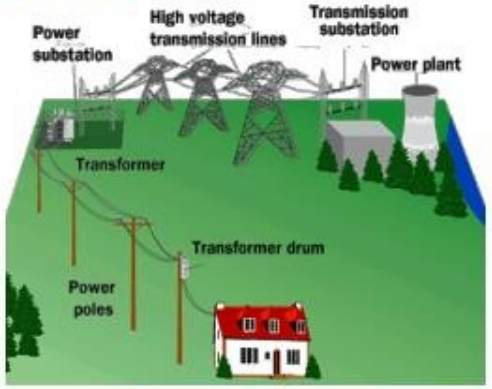
Managing Director MCNC, Grid Computing and Networking Services, and

Department of Computer Science at NC State University

This Lecture aims at complementing your knowledge in distributed computing technologies and developments, providing an actual overview of Grid Computing, and its applications and benefits for Education, Industry and Economy.

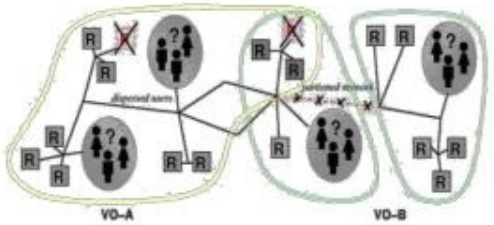
MCNC | Grid Computing & Networking Services

A simple View of Grid :



An IT Utility on a . . .

Grid Middleware (the glue) Managing . . .



Networked Distributed Resources

Integrated & "Integrate-able"

End-to-End, Javacard -> J2EE

The Grid

Partners

Innovation

Security

The underpinning technology for distributed computing

Flexibility

Scale

Enables coordinated sharing of distributed and dynamic resources, including computers, applications, data archives, visualization, sensors, and multiple remote instruments

Open public API's

Value

MCNC | Grid Computing & Networking Services

Why Should We Care about Grids ?

" It's the next big thing ! "

Grid technologies advance **Science and Education** in that we can do things which haven't been possible before.

Grid infrastructure attracts and enables **new businesses** and creates **new jobs**, especially in today's rural areas.

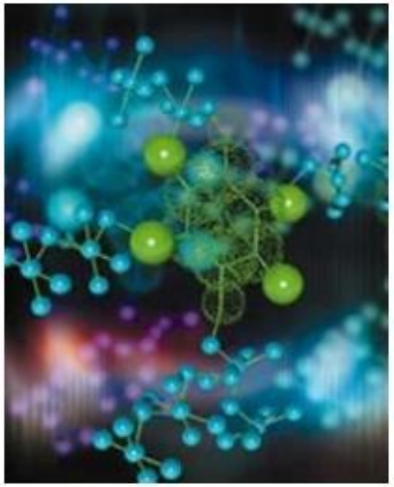
Grids make us **more competitive** by better utilizing resources, bringing results 'to market' faster, and delivering with higher quality.

An early pilot: NC BioGrid

One of the nation's first grid test beds for computing, data storage and networking resources for life sciences research

Established to research and implement new grid computing technologies that will enable researchers and educators throughout North Carolina to take full advantage of the genomic revolution

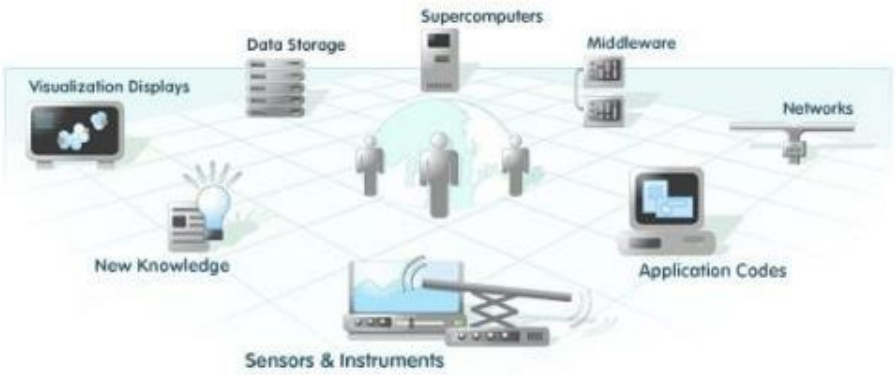
- Installed in Summer 2002, heterogeneous hardware and OS platforms
- More than 80 organizations
- Dedicated systems for testing grid middleware and developing grid applications for bioinformatics
- Spans multiple administrative domains with systems located at MCNC, NC State, UNC-CH & Duke
- Established a Certificate Authority



A Real Grid:

NEES, Network for Earthquake Engineering Simulation

NEESgrid will link earthquake researchers across the U.S. with leading-edge computing resources and research equipment, allowing collaborative teams to plan, perform, and publish their experiments.

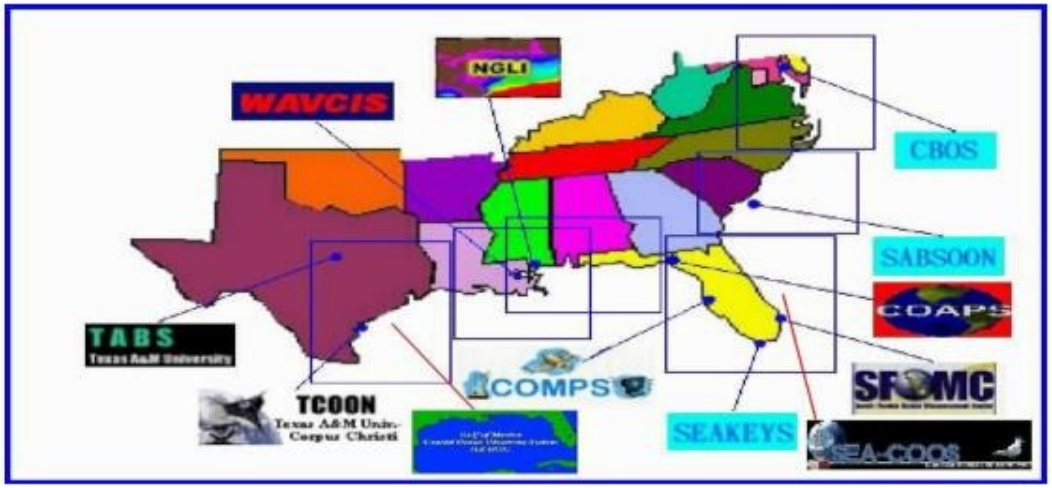


Through the NEESgrid, researchers will:

- perform tele-observation and tele-operation of experiments;
- publish to and make use of a curated data repository using standardized markup;
- access computational resources and open-source analytical tools;
- access collaborative tools for experiment planning, execution, analysis, and publication.

The Future just Began: SURA SCOOP

South-eastern Coastal Ocean Observing Program



The Challenge:

- More than half of the nation's tidal shores home to 80 million people
- coastal zone is undergoing environmental and ecological changes
- threaten the sustainability of the region's economies and marine resources

The Solution:

- develop a Grid of sensors and linked computers
- fully integrating several observing systems in the southern region
- provide data, in real-time and at high speed, for more reliable, accurate and timely information

To help guide effective coastal stewardship, plan for extreme events, facilitate safe maritime operations, and support coastal military security.

Customer Value Proposition (Benefits) Department, Enterprise, and Global Grids

- Access: transparent, remote, secure, wireless
- Sharing: enable collaboration over the network
- Failover: migrate/restart applications automatically
- On Demand: get resources, when you need them
- Productivity: more work done in shorter time
- Virtualization: access compute services, not servers
- Heterogeneity: platforms, OSs, devices, software
- Resource Utilization: increase from 20% to 80+%
- Virtual Organizations: build & dismantle on the fly

However, There is Still a Long Way to Go !

- Grids are over-hyped: currently, they promise much more than they can really offer.
- Grid technology is far from mature and complete.
- Grid standards are (mostly) still missing.
- Grids are very complex IT infrastructures.
- Grids bring new challenges: sharing resources, loosing direct control, security, intellectual property, legal, social, political issues . . .

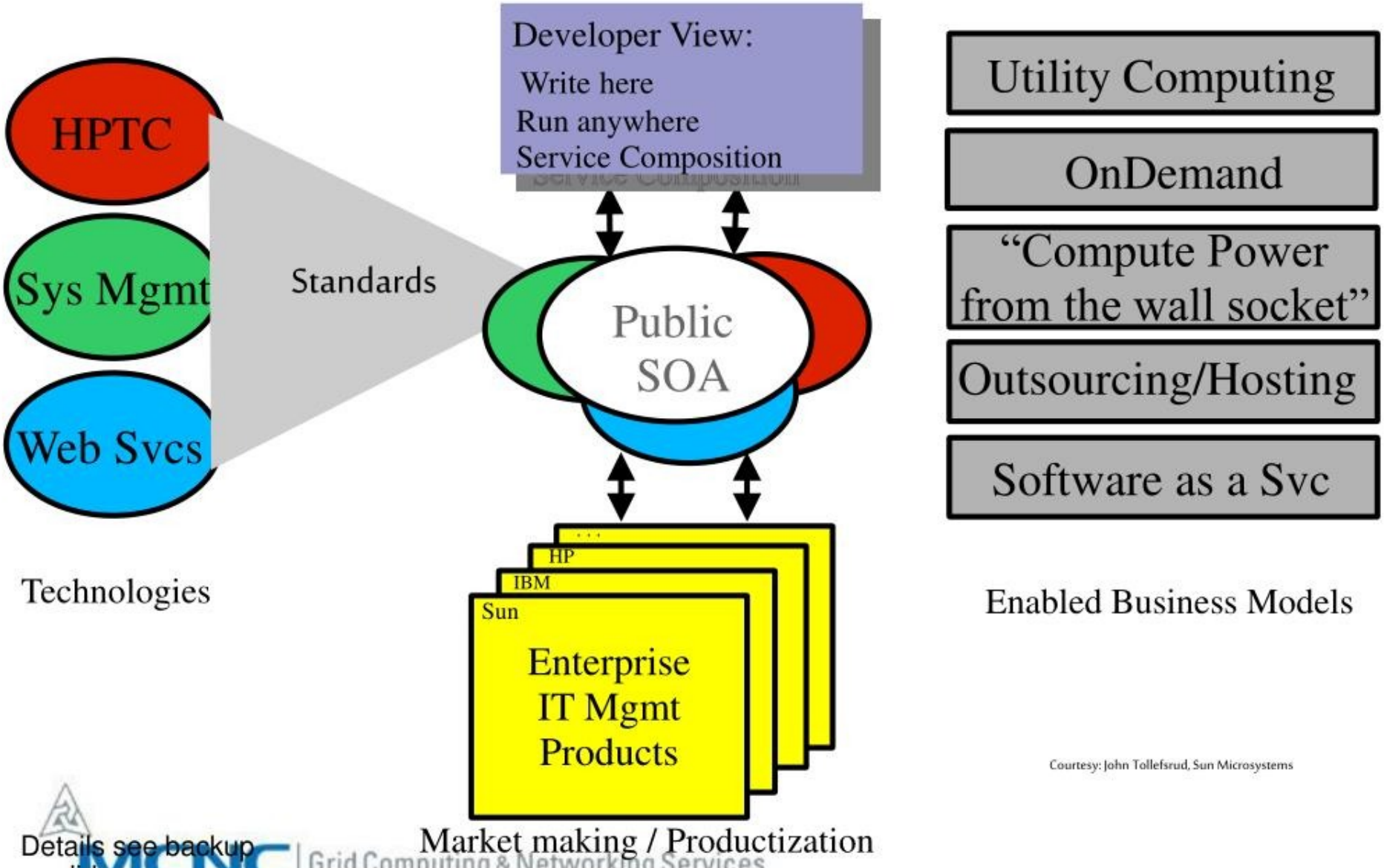
Grid Architecture, Technology, Standards

MCNC | Grid Computing & Networking Services

Architectural Requirements are huge

- ***Availability***
  - ***Downtime impact***
  - ***Individual jobs***
  - ***Maintenance windows***
- ***Scalability***
  - ***Growth 1-3-5 years***
  - ***Scaling strategy and***
  - ***Response to peak loads***
  - ***Technology refresh, evolution***
- ***Manageability***
  - ***Skill set / workload of SA***
  - ***Expected stability***
  - ***Code management, software distribution***
- ***Security***
  - ***User authentication***
  - ***Internet access***
  - ***Data Security requirements***
- ***Data Distribution***
  - ***Location, volume, refresh***
  - ***Security of data***
- ***Usability***
  - ***Administrative Skill set and***
  - ***Client environment***
  - ***Psychological factors important***
- ***Operations Management***
  - ***100s CPUs / SA***
  - ***Resources added in large blocks***
  - ***Change control is critical***

The Grid Landscape is Changing



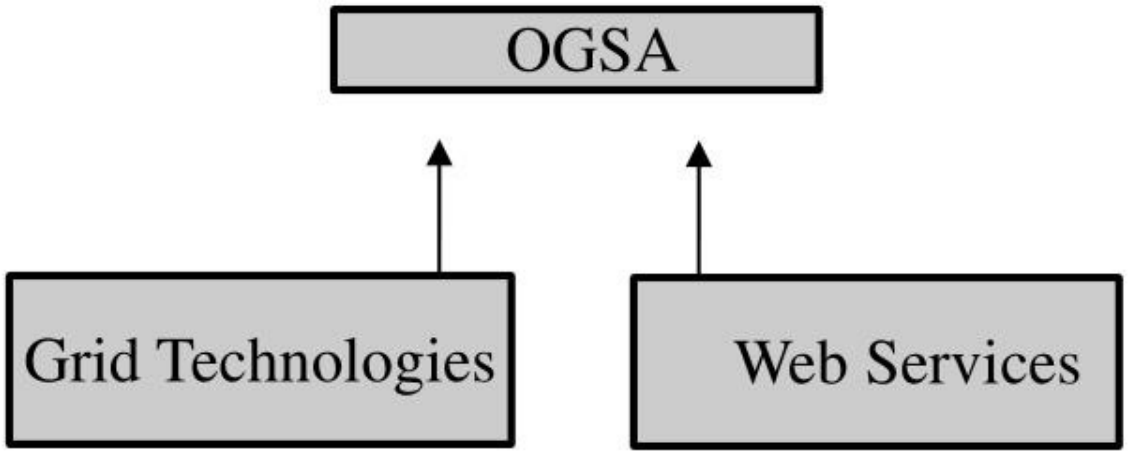
Courtesy: John Tollefsrud, Sun Microsystems

Details see backup slides

2001: Global Grid Forum (GGF)

- Community-driven set of working groups that are developing standards and best practices for distributed computing ("Grids") efforts
- 2001: Merger of US, APAC, Euro Grid Forums
- 2002: Open Grid Services Architecture (OGSA)
- 2004: WS-RF Web Services Resource Framework
- Standards: IETF, DMTF, OASIS, WS-I, W3C, EGA,…

2002: Open Grid Services Architecture

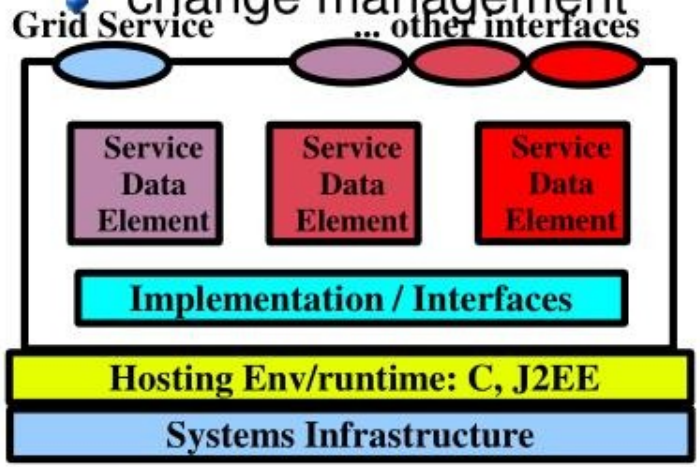


- OGSA Open Grid Service Architecture
  - Integrates grid technologies with Web Services (OGSA => WS-RF)
  - Defines the key components of the grid

2005: Globus Toolkit 4.x
Open Grid Services Architecture (OGSA)
Web Services Resource Framework (WSRF)

- ***Integration of Grid technologies and Web Services***
- ***OGSA defines a "Grid Service"***
- ***In terms of WSDL interfaces, defines mechanisms:***

- creating distributed systems
- lifetime management
- change management



- credential management
- notification

- Notification
- Authorization
- Service Creation
- Service Registry
- Manageability
- Concurrency

NCGrid Globus Toolkit 4 Workshops

Friday, January 28th, and Friday, February 4th, 2005

Presenter: Pawel Plaszczak, Gridwise Technologies; Contact: Chuck Kesler jckesler@mcnc.org

Day 1 - Friday, January 28th

Morning Session (9am - noon)

- Introduction to grid computing and the Globus Toolkit (brief)
- Review of GT 2.4, 3.2, 4.0 and counterpart standards (OGSA, WSRF)
- Overview of Globus components, including topics not covered (e.g. OGSA-DAI, CSF, CAS)
- Discussion of GSI security (X.509 certs, chain of trust, CA, signing certificates, proxies, secure WS components)

Afternoon Session (1pm - 4pm)

- Web services basics
- Intro to Java tools for Grid/Web Service/WSRF programming (Log4j, Ant, Apache Axis)
- GT4 Java WS core (WSRF)
- Overview of the WSRF standard
- Overview of GT4 Grid service programmer environment (schema, WSDL, compiling, stub generation, deployment, testing)

Day 2 - Friday, February 4th

Morning Session (9am - noon)

- GRAM and job submission / purpose and architecture overview
- Web Services GRAM vs. pre-Web Services GRAM
- Adapting to customer scheduler interfaces (e.g. SGE, LSF, PBS)
- Information Services (MDS4)
- GridFTP, RLS, XIO (theory)

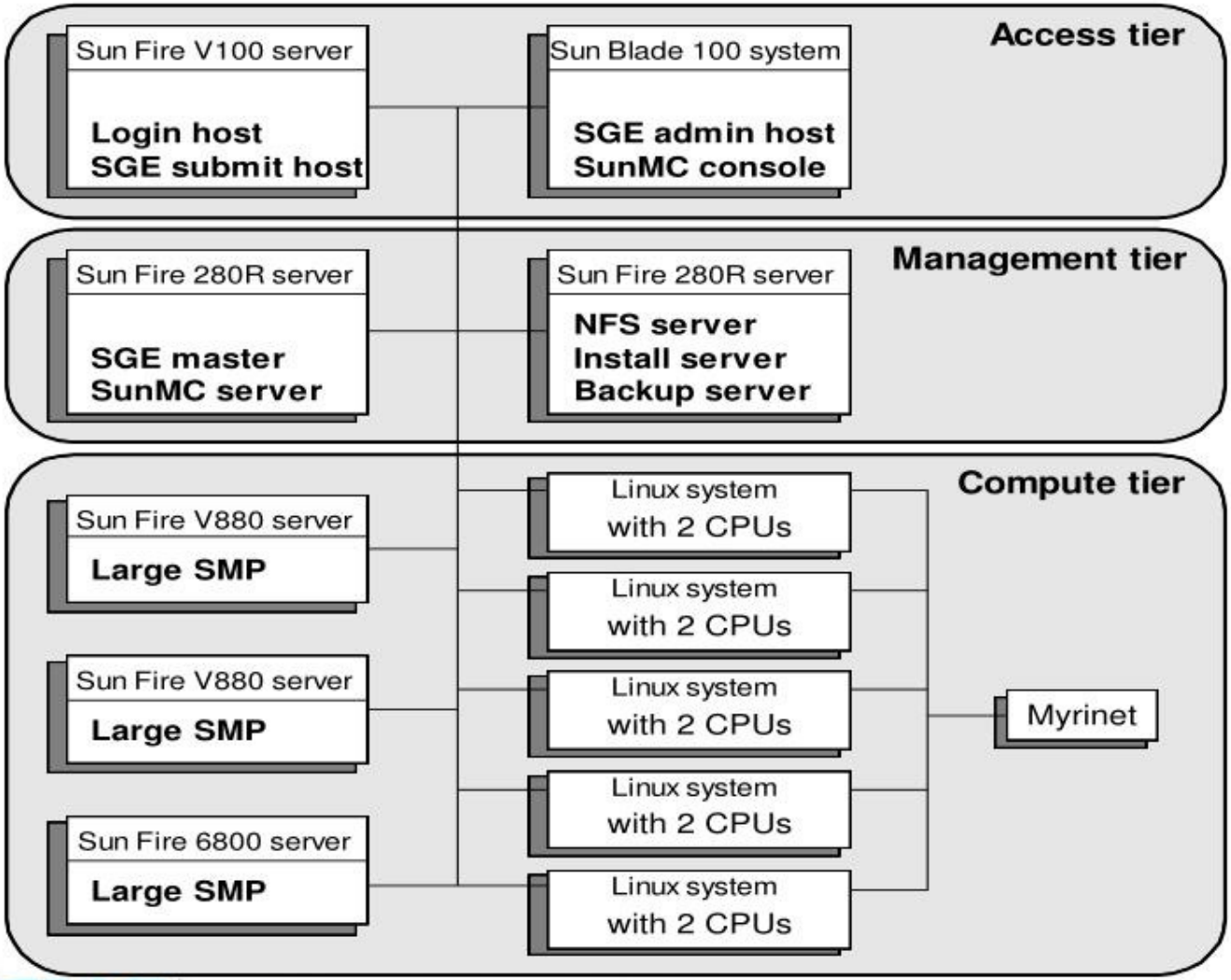
Afternoon Session (1pm - 4pm)

- Hands-on workshop - details to be announced

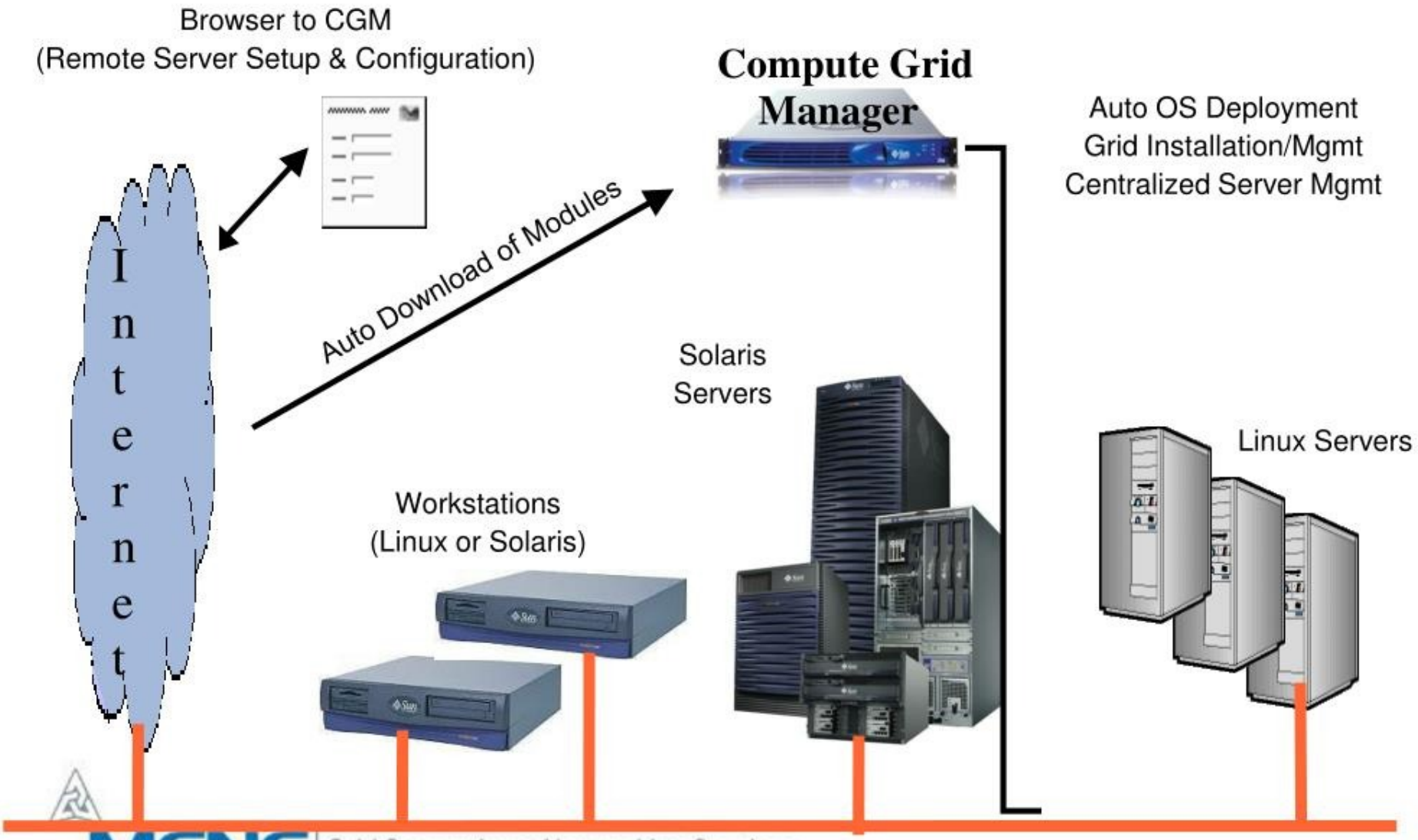
Grid Computing Today in Industry

MCNC | Grid Computing & Networking Services

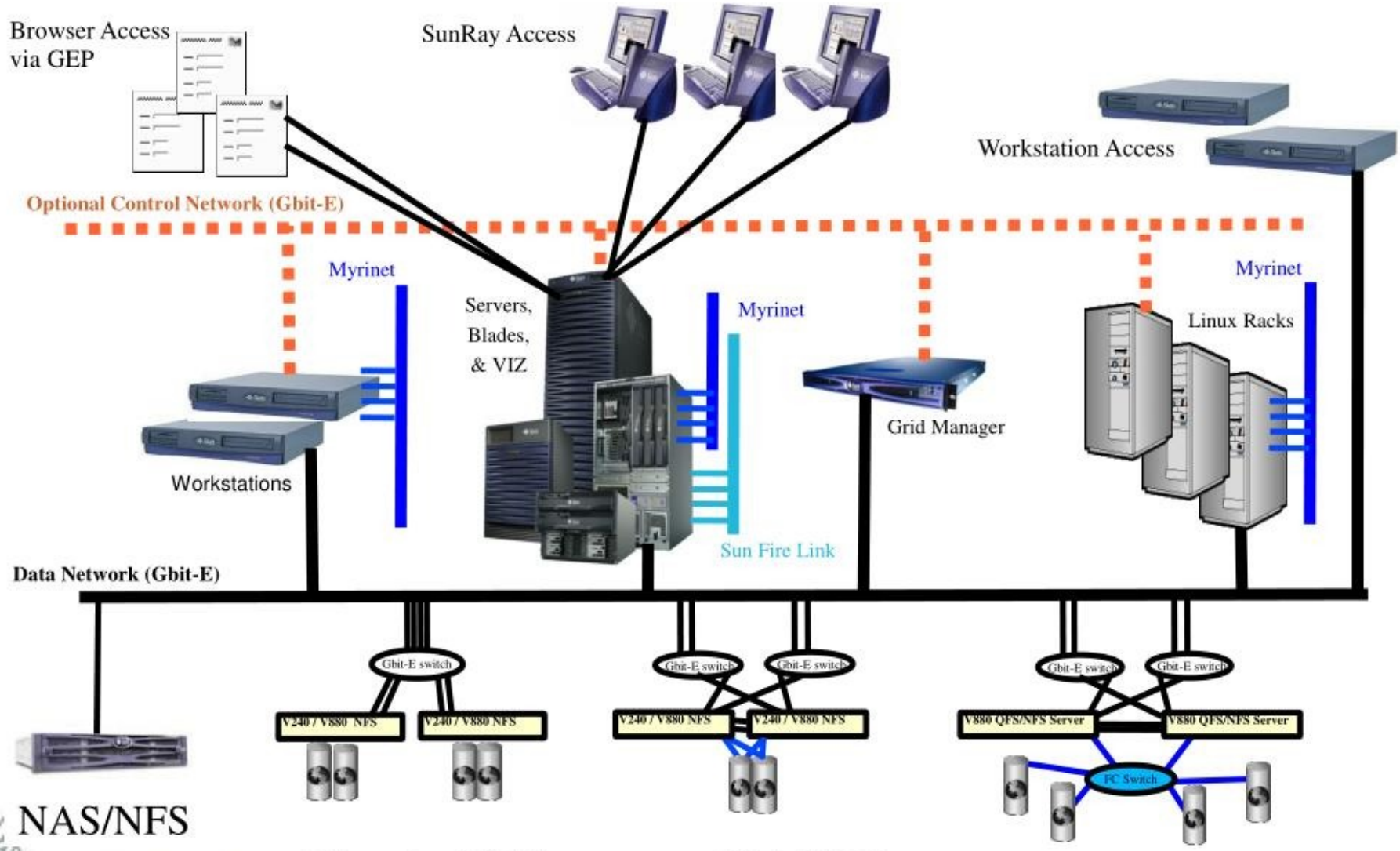
Mid 90s: Cluster Grids



Late 90s: Department Grid

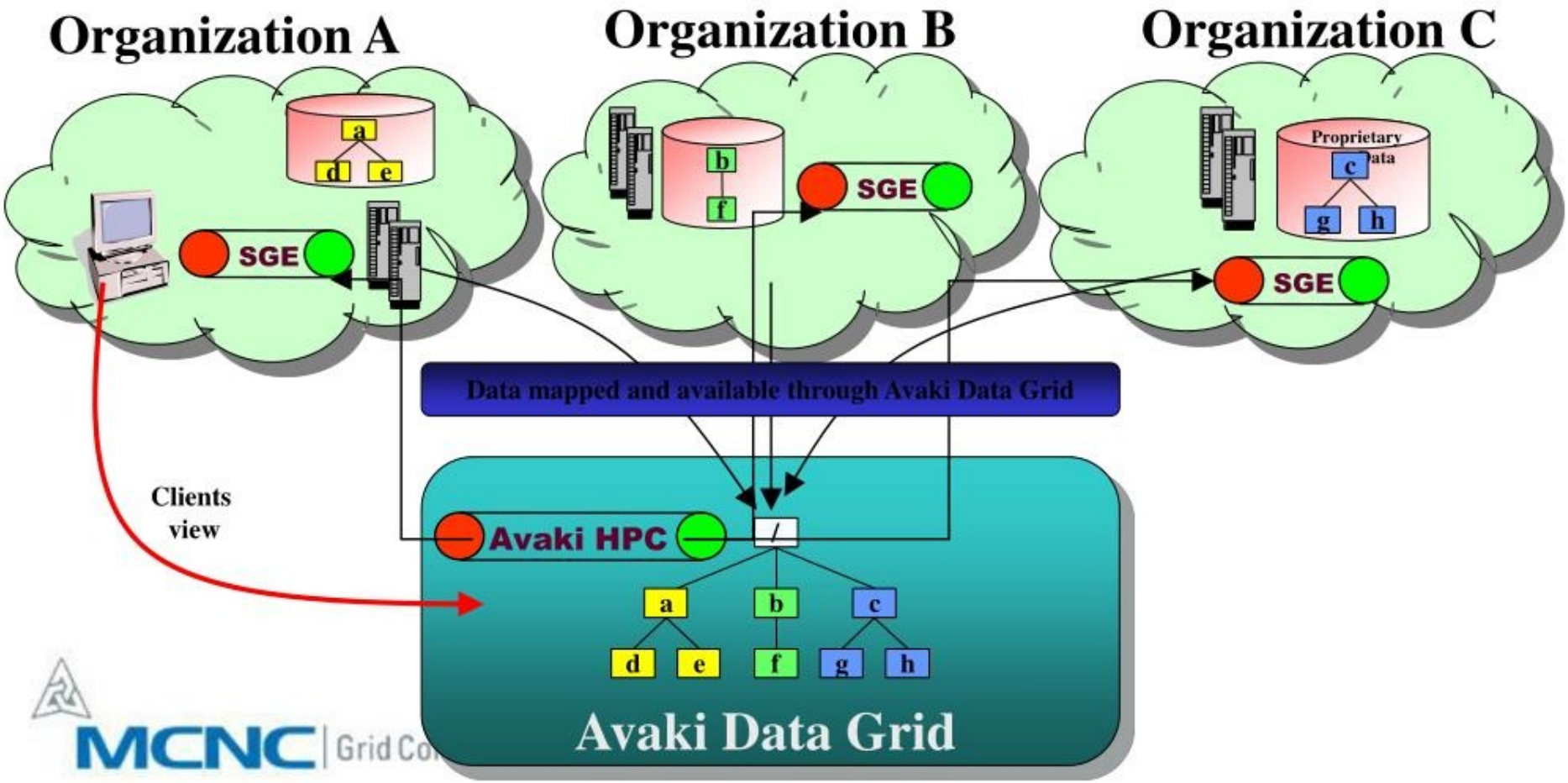


2001: Enterprise Grid Reference Architecture



2002: Globus & Avaki over mult Department Grids

1. SGE cluster mgmnt within an admin domain & file system area
2. Globus/Avaki connects resources, handles files, binary management, and high level resource selection



MCNC | Grid Co

GOAL: Computing as a Utility

=>> Enhancing The Grid with a Business Model

What's a Utility?

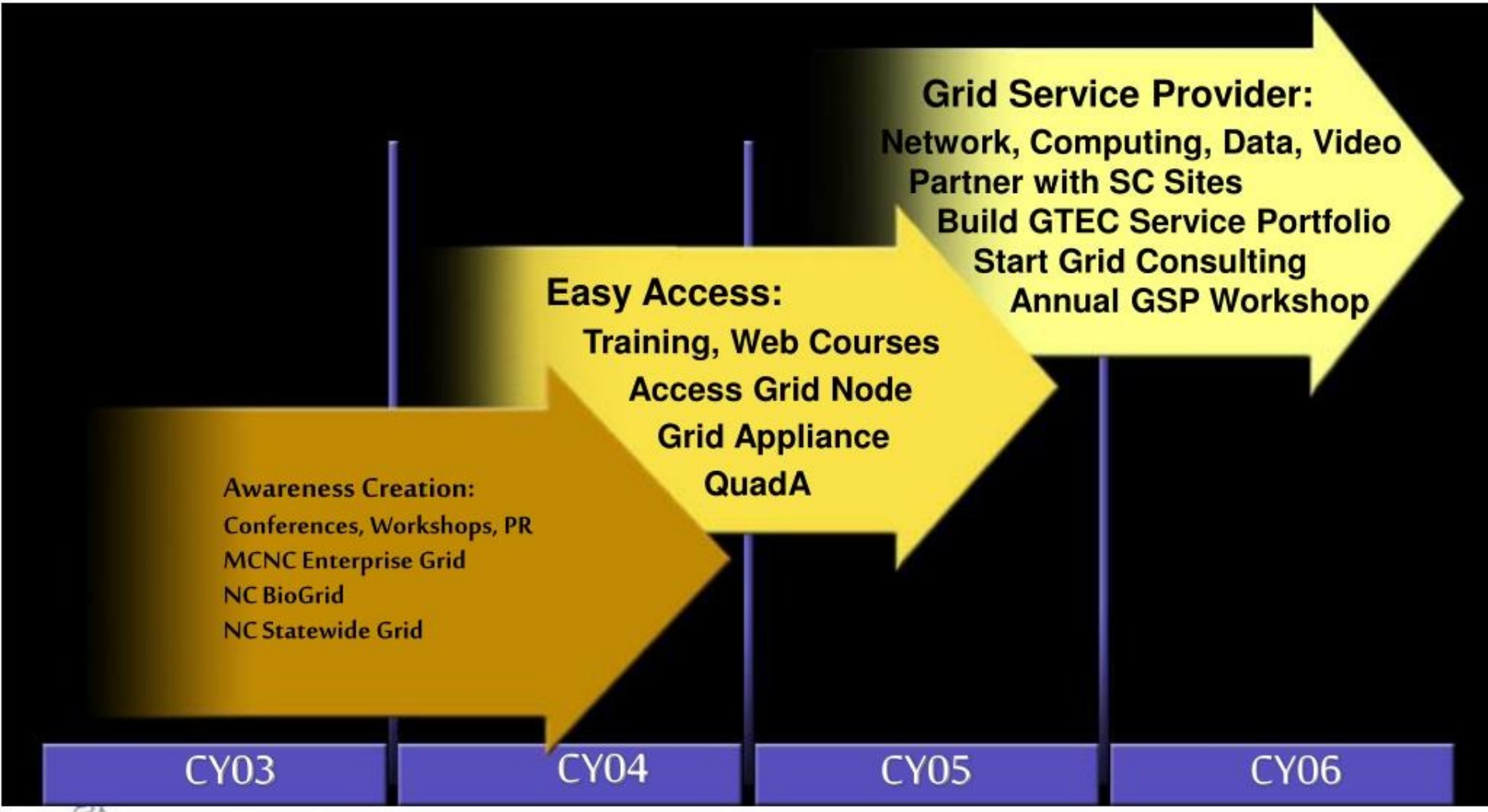
- **On Demand: Get a service at your finger tip**
- **From the Wall Socket: Don't care about the infrastructure**
- **Metering & Billing: Pay as you go, for what you use**

Like electricity, water, gas, heat, telephony

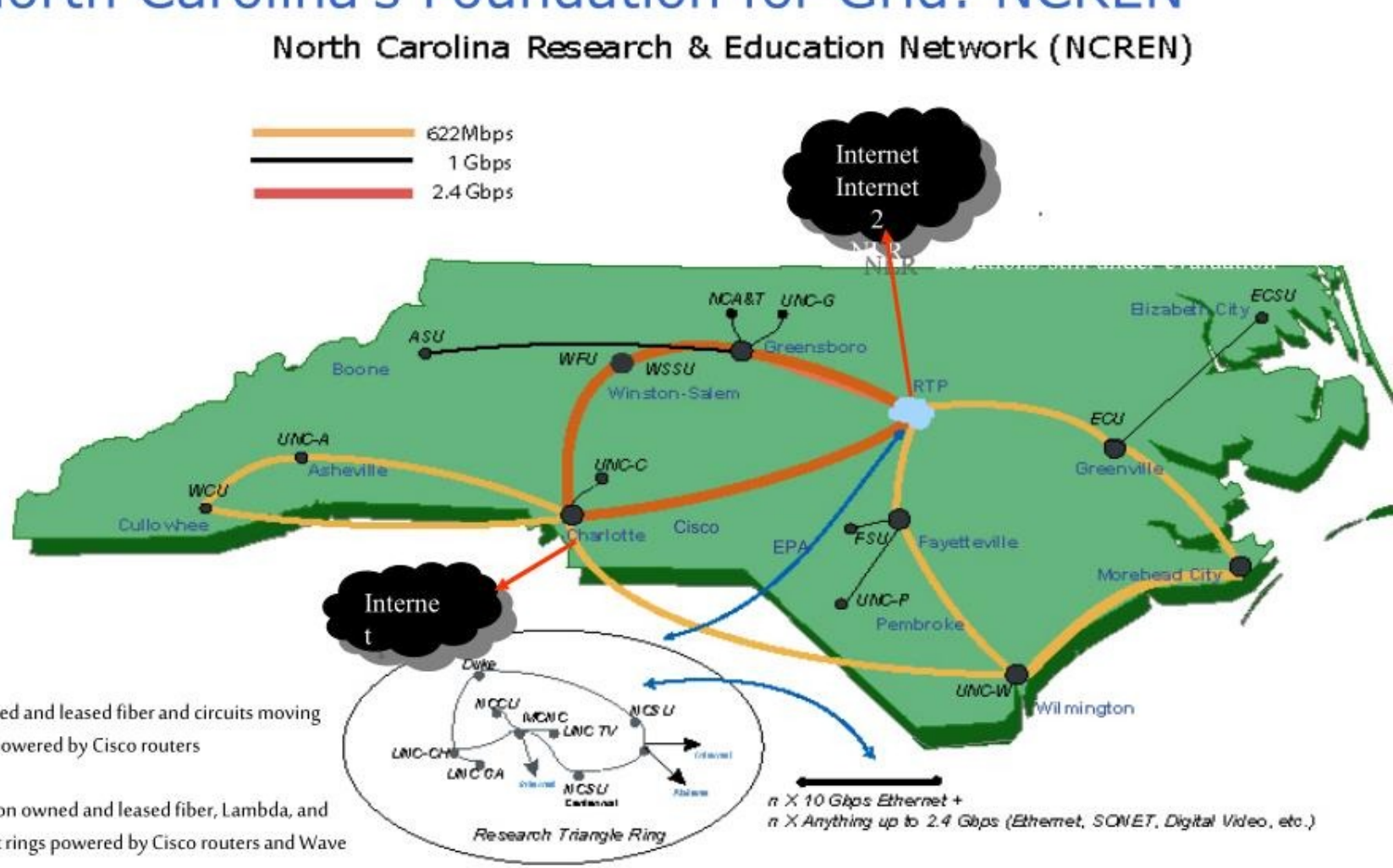
Who Uses Grids today?

- **Life Sciences**
  - *Startup and cost efficient*
  - *Custom research or limited use applications*
  - *Multi-day application runs (BLAST)*
  - *Exponential Combinations*
  - *Limited administrative staff*
  - *Complementary techniques*
- **Electronic Design**
  - Time to Market
  - Fastest platforms, largest Grids
  - License Management
  - Well established application suite
  - Large legacy investment
  - Platform Ownership issues
- **Financial Services**
  - Market simulations
  - Time IS Money
  - Proprietary applications
  - Multiple Platforms
  - Multiple scenario execution
  - Need instant results & analysis tools
- **High Performance Computing**
  - Parallel Reservoir Simulations
  - Geophysical Ray Tracing
  - Custom in-house codes
  - Large scale, multi-platform execution

MCNC: 3-Year GCNS Grid Roadmap



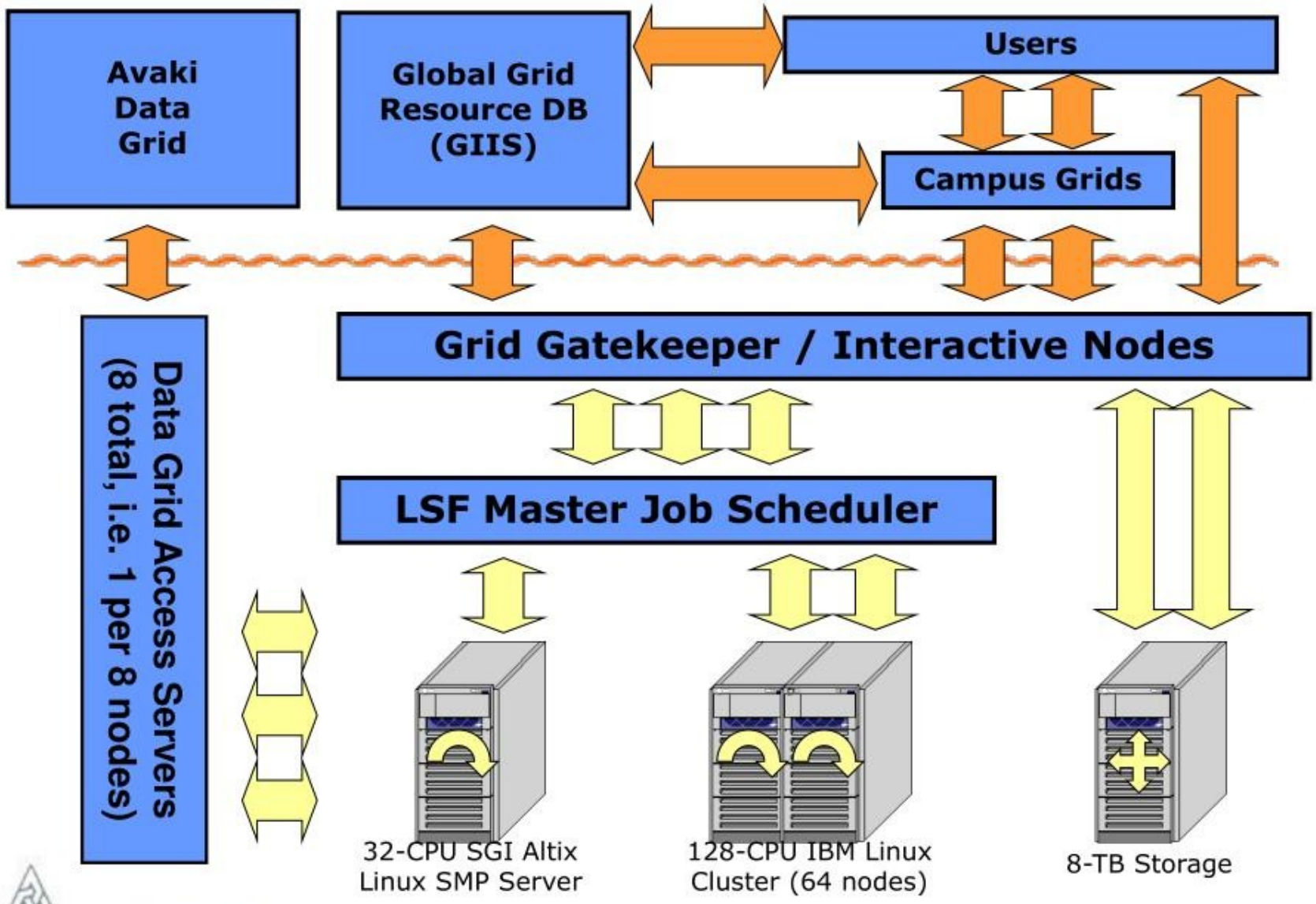
North Carolina's Foundation for Grid: NCREN



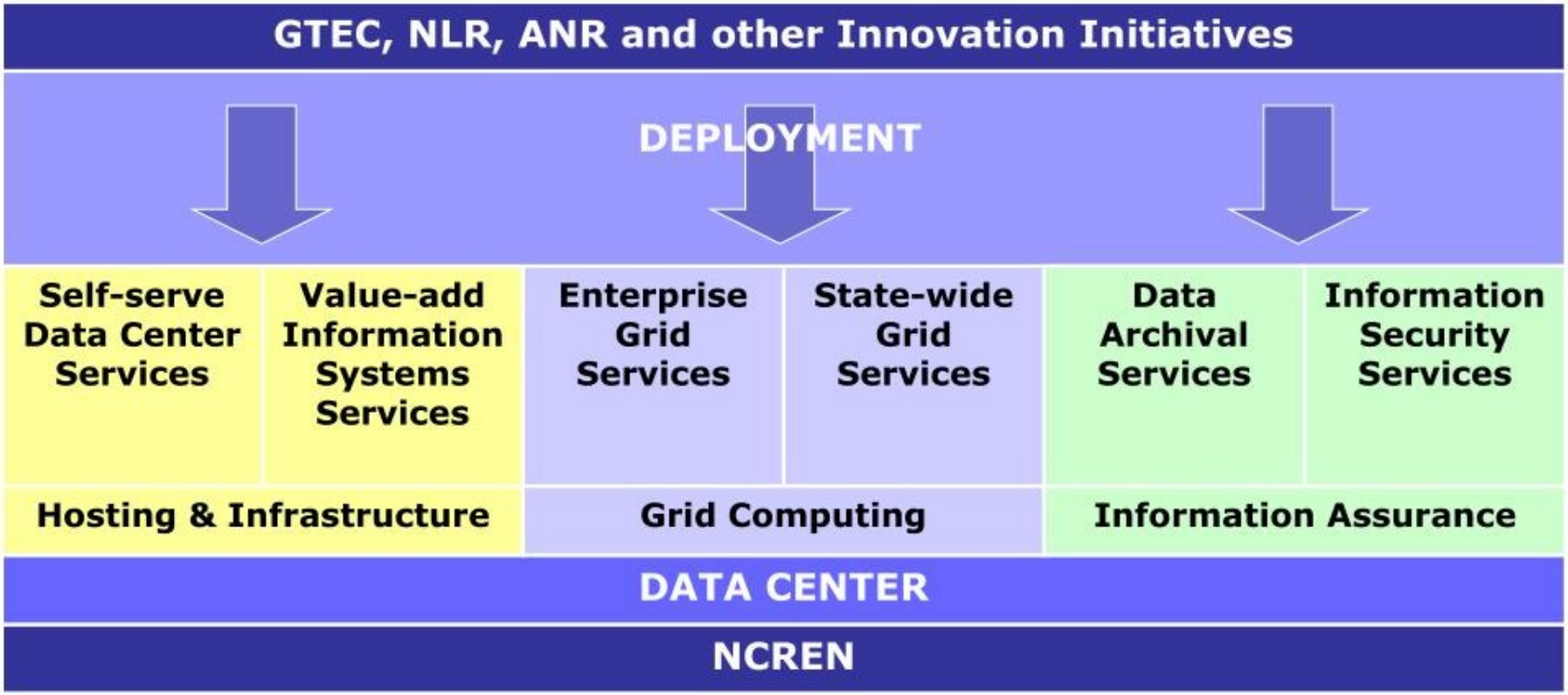
Existing: Blend of owned and leased fiber and circuits moving toward resilient rings powered by Cisco routers

Planned: Strong focus on owned and leased fiber, Lambda, and few circuits, in resilient rings powered by Cisco routers and Wave Division Multiplexers

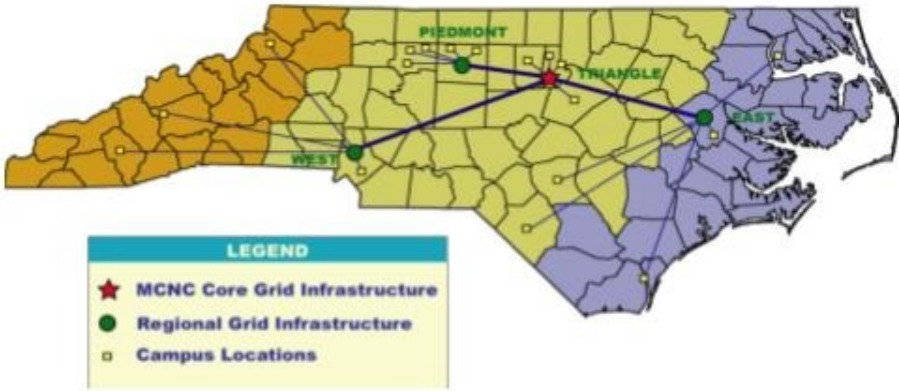
GCNS Enterprise Grid



Network, Grid and Data Center Services



The NC State-Wide Grid Roadmap

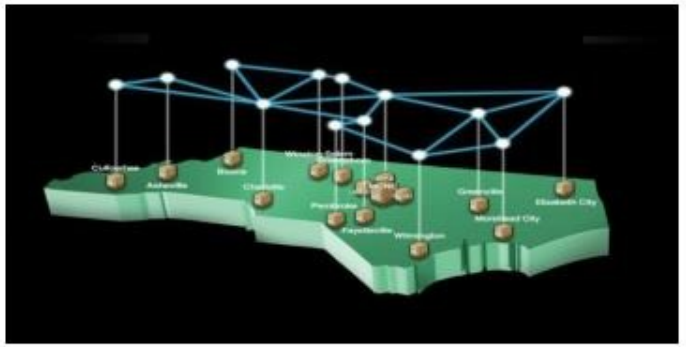


- 06/04: "Do-Grid-Yourself" Workshop
- 07/04: Phase 1, Awareness Creation
- 09/04: WCU's Barry Wilkinson Grid Lectures
- 10/04 – 06/05: Deliver/Connect Grid Appliance Clusters to University Partners in NC
- 10/04: Develop "Do-Grid-Yourself" Training Course
- 10/04: Start Quad**A** Project: **A**ccess, **A**ccounting, **A**uthentication, **A**uthorization
- 12/04: Start Deliver Grid Training to Partner Univs
- 02/05: Globus Toolkit Training
- 01/05 – 06/05: Work with Grid Users to Port Apps
- 03/05: Build Access Grid Node
- 06/05: 1st NC Statewide Grid Workshop

Example:

The North Carolina Start-up Grid

MCNC | Grid Computing & Networking Services



With NC State University &

Centennial Campus Incubator

- To help Start-ups overcome manifold obstacles
- To reduce Start-up risk and increase success
- To attract Start-ups to North Carolina
- To contribute to North Carolina's economic growth

Easy Start: Leverage what Exists

- All-Inclusive Office Space: Centennial Campus Incubator exists
- IT Resources: MCNC's Enterprise Grid and NC Statewide Grid exist
- Grid Access: Grid Appliance and Grid Portal prototype exist
- Web Presence: For company product information and marketing
- Business Platform: To offer any (grid-enabled) digital service to users, customers, and partners, over the Web/Grid
- Consulting: From NC State (e.g. College of Mgmnt) and the Triangle

Grid Vision

Our Vision :

The Three Waves of Grid Computing

The Research Wave

Technology, Prototypes
Virtual Organizations
Standards
GGF, IETF, OASIS

GCNS: "Awareness Creation"

The Industry Wave

Grid-Enabled Products
Enterprise Solutions
Interoperability
GGF, EGA, IETF, OASIS

GCNS: "Easy Access"

The Consumer Wave

Commodity
IT Utility
Integration
Legal, Ethical, Political Orgs

GCNS: "Grid Service Provider"

Grids Today - Grids in 3 - 5 Years

- Focus on Research
- Compute-oriented
- Computing Services
- Proprietary interfaces
- "Mental Firewall"
- Difficult to build
- Difficult to manage
- Difficult to implement & use
- Many technologies

- Focus on R&D and business
  - Petaflops linked w/ Petabytes
- Web Services
- Standards: GGF, OGSA, WS-RF, DRMAA
- Security, policies, identity
- Standards, services, solutions
- Sun N1, IBM Autonomic, HP DC,…
- Grid Portals: transparent, remote, secure
- Globus Toolkit 5.x, ???

Finally:

Anyone, anywhere, anytime, any device, any data, connected to The Grid

- Integration of new devices, data and information sources
- Cell phones, PDAs, smart sensors, sensor arrays, health monitors
- Devices embedded in cars, engines, roads, bridges, clothes,...
- Huge amount of data for real-time analysis
- Policies, grid economy, to maintain stability and efficiency
- Support organizational and societal structures, to bridge political and social boundaries

. . . very much like other vital infrastructures (roads, telecom,…)

Time Machines

The Innovation Engine

The Combustion Engine

The Steam Engine

Thank You !

wgentzsch@mcnc.org
http://www.mcnc.org