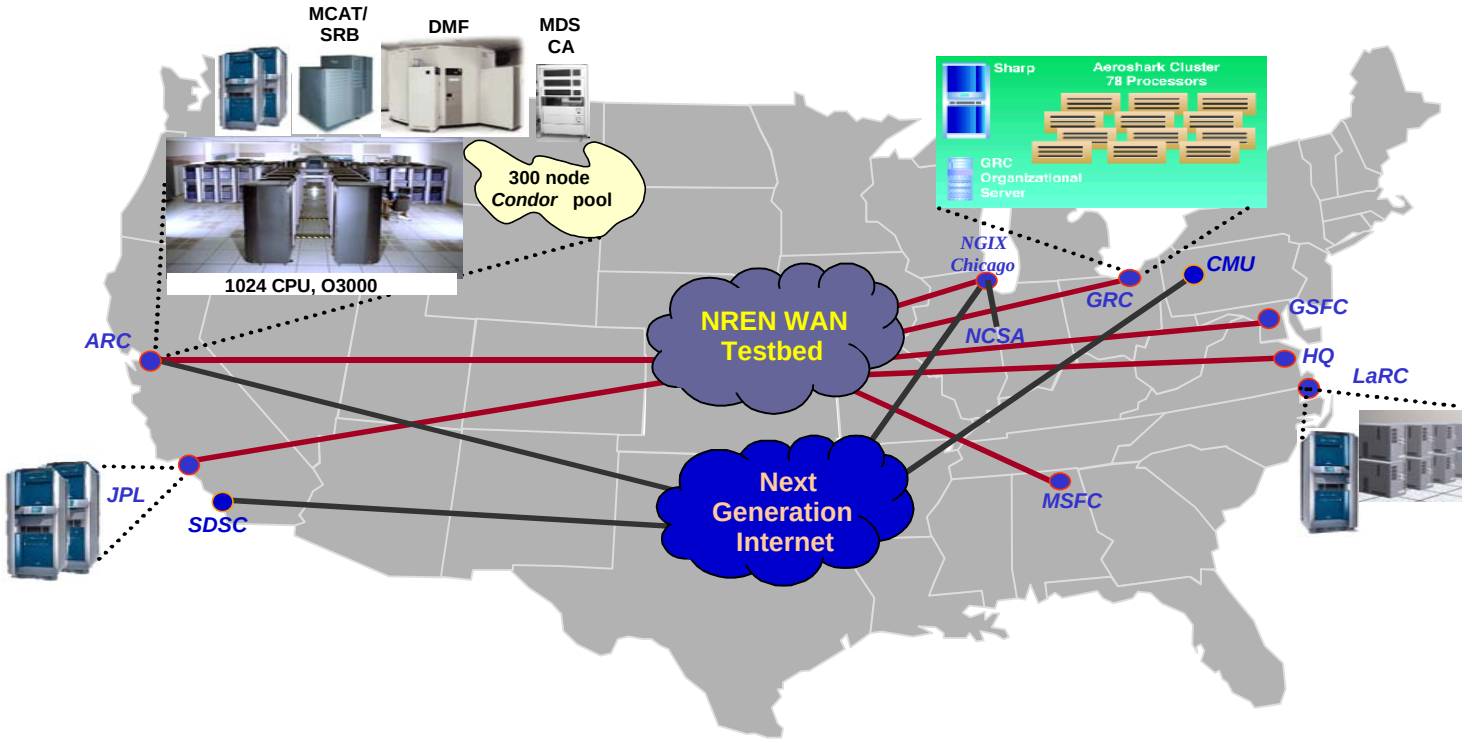
GSAW 2005 Breakout Session 9E :

Components, Frameworks, & Web/Grid Services for Ground Systems

Shirley Tseng
Infinite Global Infrastructures
www.igillc.com
Stseng@igillc.com
714-832-5373, 949-300-1899 cell

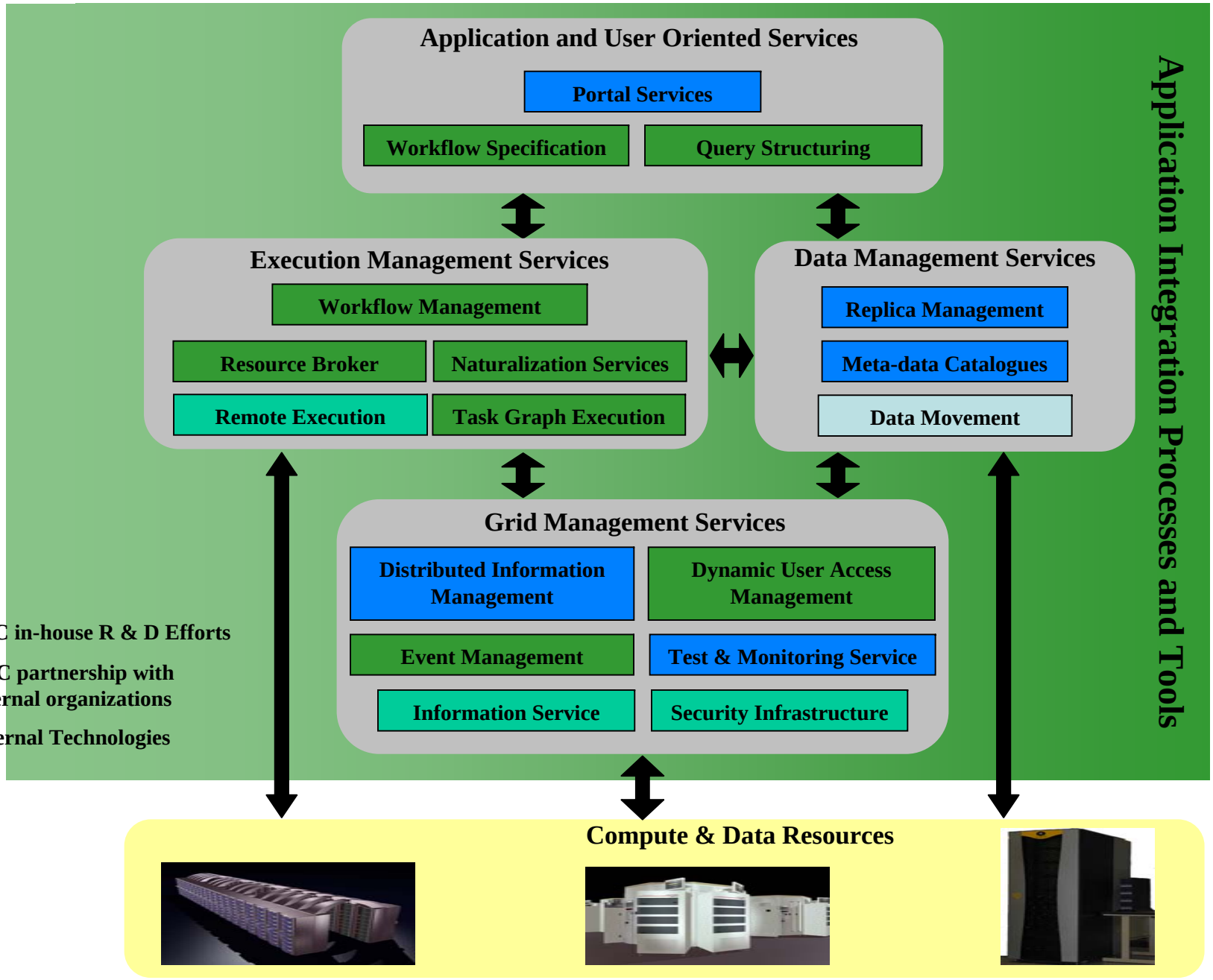
- Who is doing what? Where? Contact information
  - Project implementing Grid component layers
  - Projects implementing Web Services & SOA (WSDL, UDDI, SOAP)
  - XML in Space definition initiatives
  - Layer view to map projects to
- A few representative charts from the Projects
- Observations from Survey
  - Early US initiatives - NASA IPG, DOE
  - Concentrated European initiatives – SpaceGrid study (2001-2003), buyin by European management and companies, projects initializations
  - US Grid Rollout (driven by science data users/organizations)
    - driven by users close to the science data – Principle Investigators using satellite data Ex : NVO, CEOS, OGC
  - NASA ESTO SEEDS (Strategy for Evolution of ESE Data Systems study 1998)
    - ESDSWG (Earth Science Data Systems Working Groups)
  - Mixed pockets of implementations – OODT, MERS CIP

NASA'S Information Power Grid (IPG)



Source: Dec 2001 workshop presentation: http://www.ipg.nasa.gov/

NASA Grid Development



NASA IPG Applications

- The IPG, funded by the Computing, Information and Communications Technology (CICT) program at NASA Ames Research Center, is a collaborative effort between NASA Ames, NASA Glenn, and NASA Langley Research Centers, and the NSF PACI programs at SDSC and NCSA
- IPG Applications:

- Remote Access to High Data-Rate Instruments
  - The DARWIN system at Ames has users scattered across the country
  - UCSD TeleScience system has a NASA user at Wallops manipulating the instrument at UCSD.
- Management and access to massive data sets
- The IPG Data Mining application:
  - Persistent and uniform access to heterogeneous, multi-organizational archival storage systems
  - The SDSC Storage Resource Broker (SRB - an IPG Grid service) provides a standard data access interface for heterogeneous data archive systems
  - SRB's MCAT is a catalogue service that provides a standard way to define, manage, and search metadata for all files in a collection, where a collection may span many data archive systems
- Aviation Safety (Glenn on distributed simulation)

Remote Access to High Data-Rate Instruments

2005

Users

Web user interface

Tele-Science Portal

UC San Diego

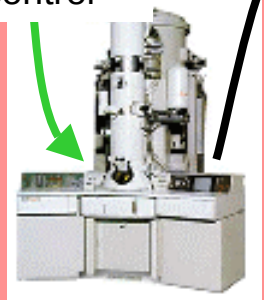
authentication

compute resource access

user data access

instrument data storage

remote instrument control



UC San Diego
National Center for Microscopy and Imaging Research

NASA Ames

authentication

user data access

Web user interface

DARWIN/DREAM data server / portal

metadata

instrument data storage

Ames Wind Tunnels:
- National Full-Scale Aerodynamics Complex
- 9x7 ft Supersonic and 11 ft Transonic
- 12 ft Pressure

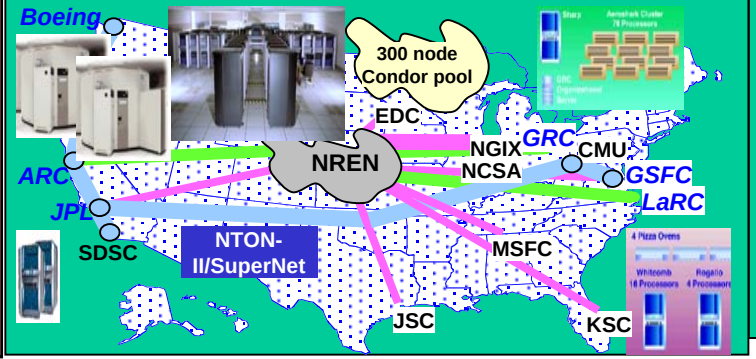
Grid Services: Uniform access to distributed

Grid Information Service | Uniform Resource Access | Brokering | Global Queuing | Co-Scheduling | Global Event Services | Data Cataloguing | Uniform Data Access | Collaboration and Remote Instrument Services | Network Cache | Communication Services | Authentication Authorization | Auditing | Monitoring | Fault Management

Security Services

SDSC compute and data storage resources

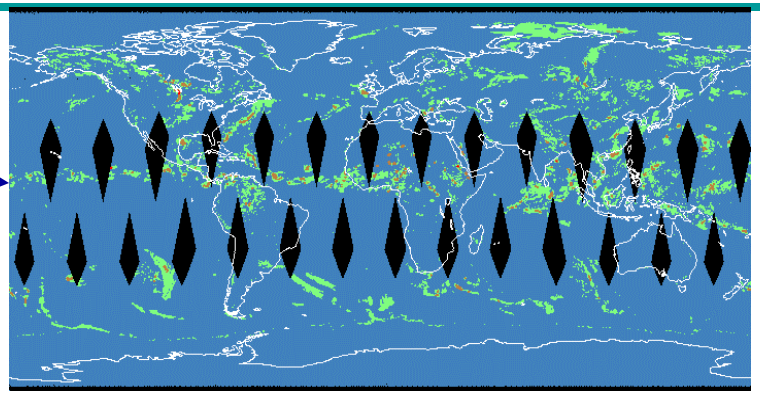
IPG compute and data resources

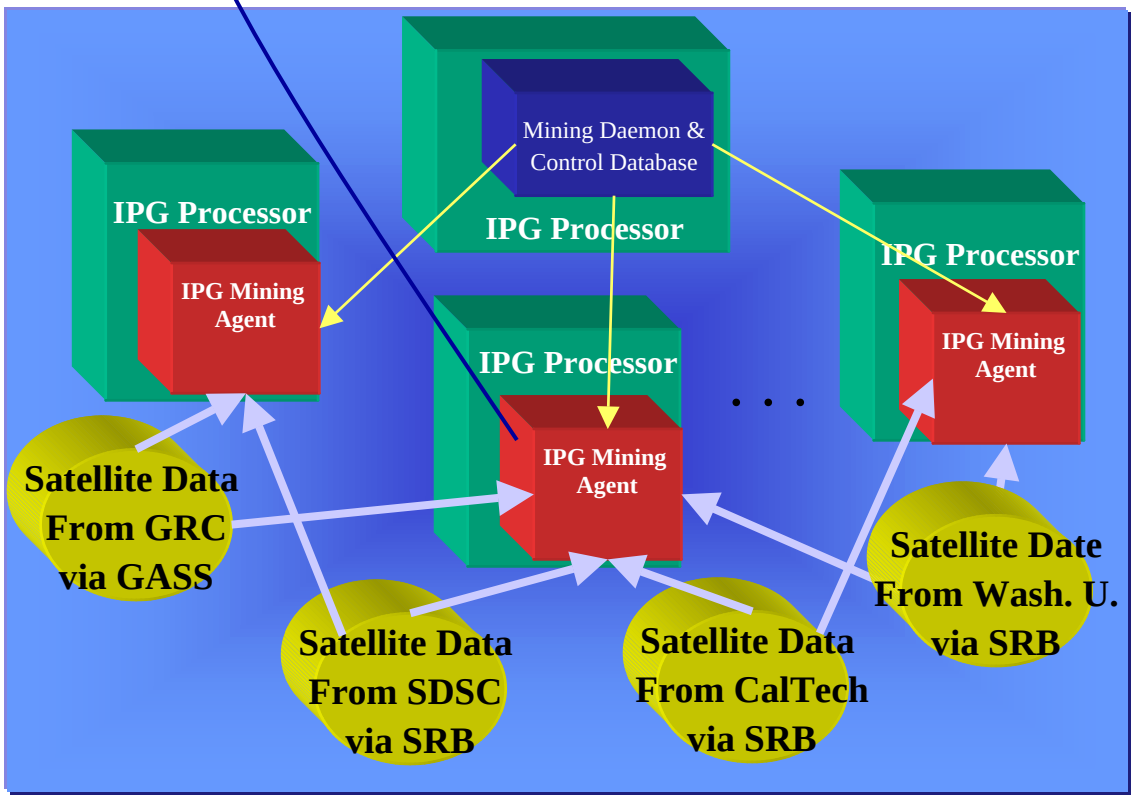


High Speed Distributed Data Access: IPG Milestone Completed 3/2000

- Data access capabilities of IPG are demonstrated by *parallel data mining*
- 512 node SGI Origin at Ames uses IPG uniform interface data access tools (SRB) to simultaneously mine hydrology data from four sites
  - SDSC
  - CalTech
  - GRC
  - Washington U.

Result from one agent





Tom Hinke, NASA Ames

Who is using Grids? ….CEOS Grid Prototypes

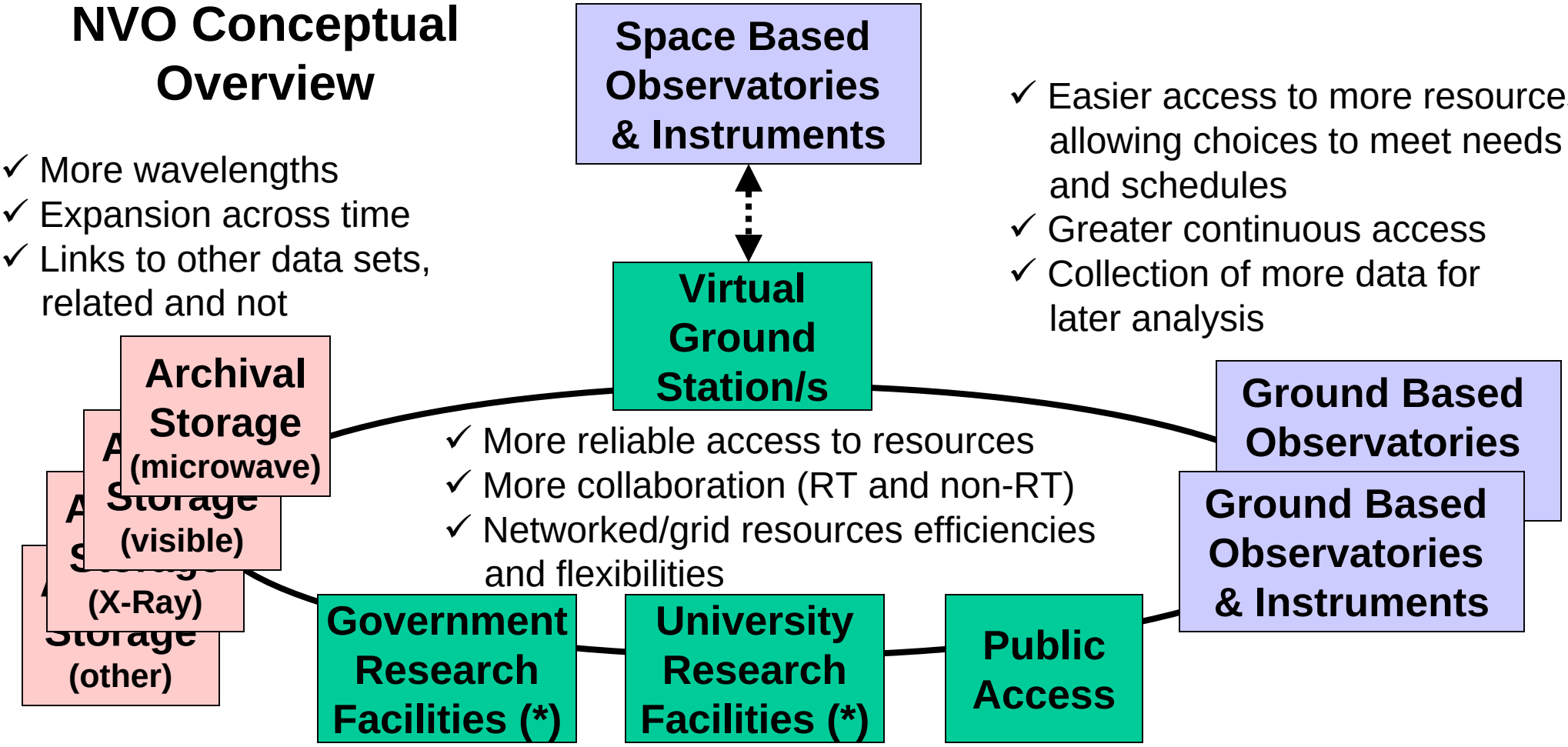
- Committee on Earth Observation Satellites (CEOS http://wgiss.ceos.org/ceos.htm)
- Next Generation Prototyping AI: Dick desJardins, NASA Ames
- CEOS Grid Team Objectives:
  - Establish CEOS Grid Testbed with multiple nodes
  - Demonstrate Grid enable applications
  - Infuse grid technologies into CEOS information systems and test facilities
- 4 CEOS Grid Prototypes
  - GSFC Advance Data Grid, Debbie Ladwig, Prototype advanced data management technologies like Storage Resource Broker with nodes at NASA GSFC, NASA Ames, Aerospace Corp., El Segundo, CA
  - USGS Data delivery utilizing GridFTP, Stuart Doescher, Prototype Globus Replica Catalog for keeping track of replicated files and the Globus Replica Management for managing files in a high performance, wide area data storage environment.
  - NOAA Operational Model Archive and Distribution System (NOMADS)". Glenn K. Rutledge, Prototype remote access to high volume numerical weather prediction and global climate models and data, the National Climatic Data Center (NCDC), along with the National Centers for Environmental Prediction (NCEP) and the Geophysical Fluid Dynamics Laboratory (GFDL)
  - ESA Data Integration via Grid Portal, ES data warehouse, Grid service management for Ozone data application

National Virtual Observatory

- NSF funded - $10M 5 year project starting Nov 01
- Objective: provide observers with 'transparent' access to astronomy data (across data centers, instruments, tools)
  - Enabling new science, new user classes, integrated tool, common data access
- WGs: System Architecture, Metadata, data model,
  - image services, registry, user interfaces,
- PI and project director: Alex Szalay (JHU); CoPI: Roy Williams (Caltech/CACR)
- Demo candidates
  - Brown Dwarf science search – distributed query across data archives, searching into raw data for new correlation
  - Galaxy morphology science demo – on the fly image analysis and recognition
  - Gamma Ray burst science demo – event follow-up service for interoperability demonstration
  - SkyQuery – dist query tool using set of reusable services
- International Virtual Observatory Alliance (IVOA) and roadmap
- **The National Academy of Sciences and Astronomical Survey Committee recommended in its decadal survey (NAS99) the establishment of a National Virtual Observatory (NVO) to utilize the latest computer and networking technologies to connect the archival and real-time resources of many earthbound and orbital astronomical observatories :**http://www.us-vo.org/

NVO Conceptual Overview

- ✓ More wavelengths
- ✓ Expansion across time
- ✓ Links to other data sets, related and not



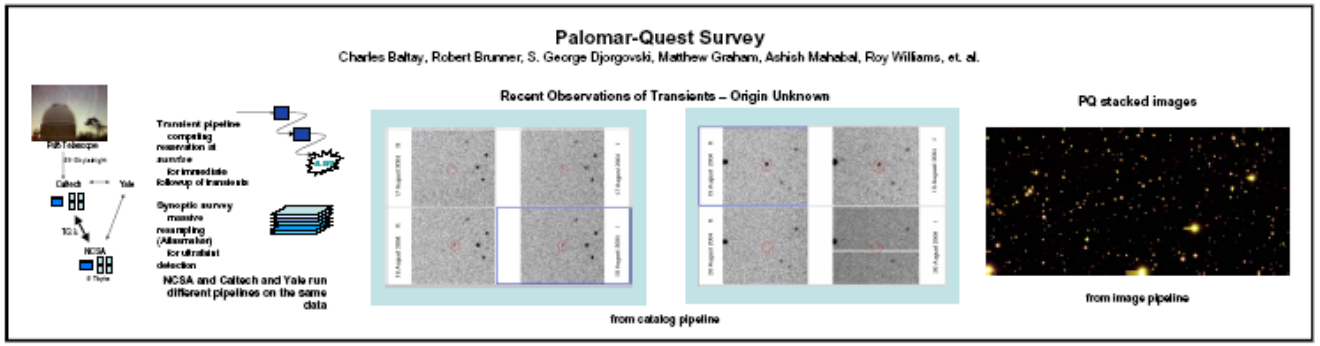
- ✓ Easier access to more resources, allowing choices to meet needs and schedules
- ✓ Greater continuous access
- ✓ Collection of more data for later analysis

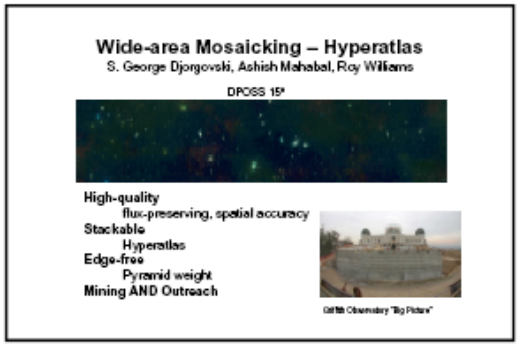
- ✓ More reliable access to resources
- ✓ More collaboration (RT and non-RT)
- ✓ Networked/grid resources efficiencies and flexibilities

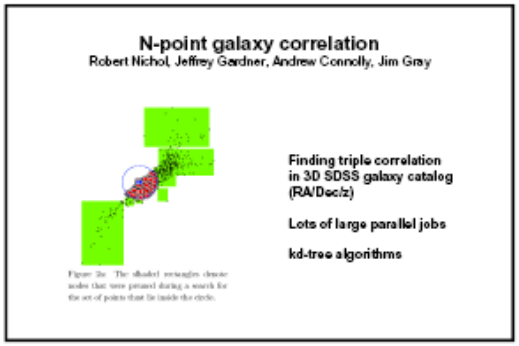
- ✓ More timely sharing of techniques and raw data
- ✓ Shared computing resources for greater efficiency and effectivity

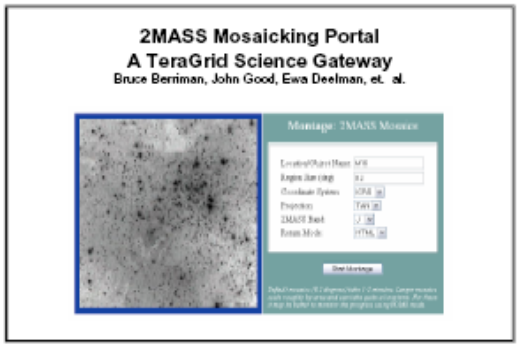
* *Includes computers, staff, local storage, etc.*

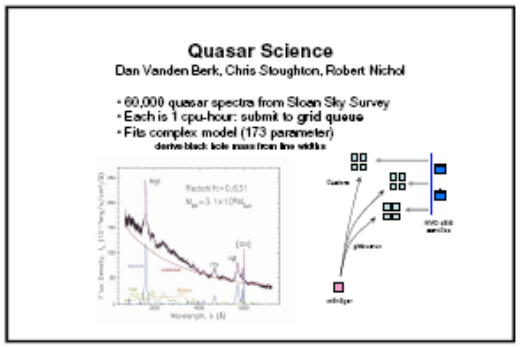
Astronomy Applications on the TeraGrid

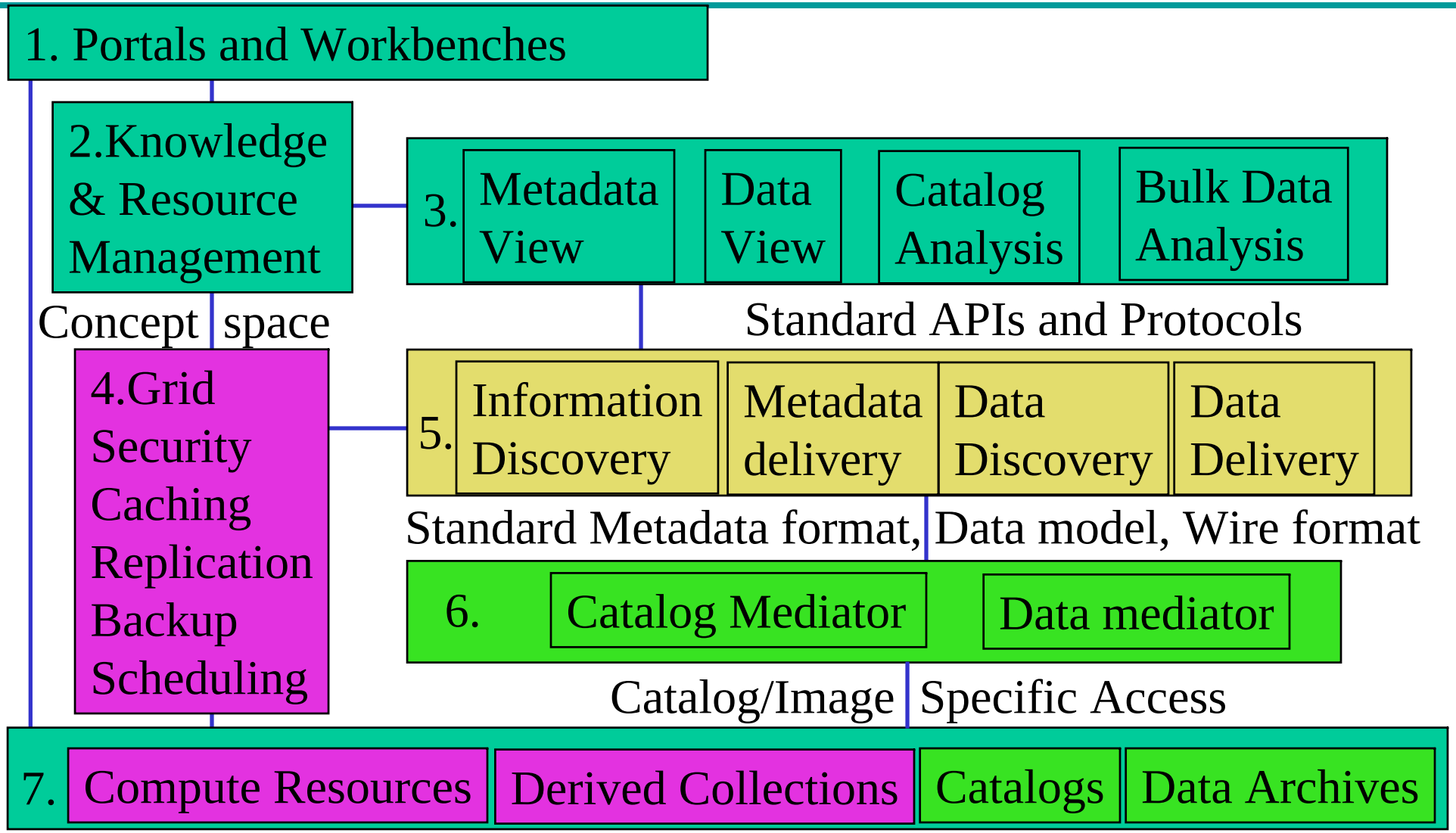






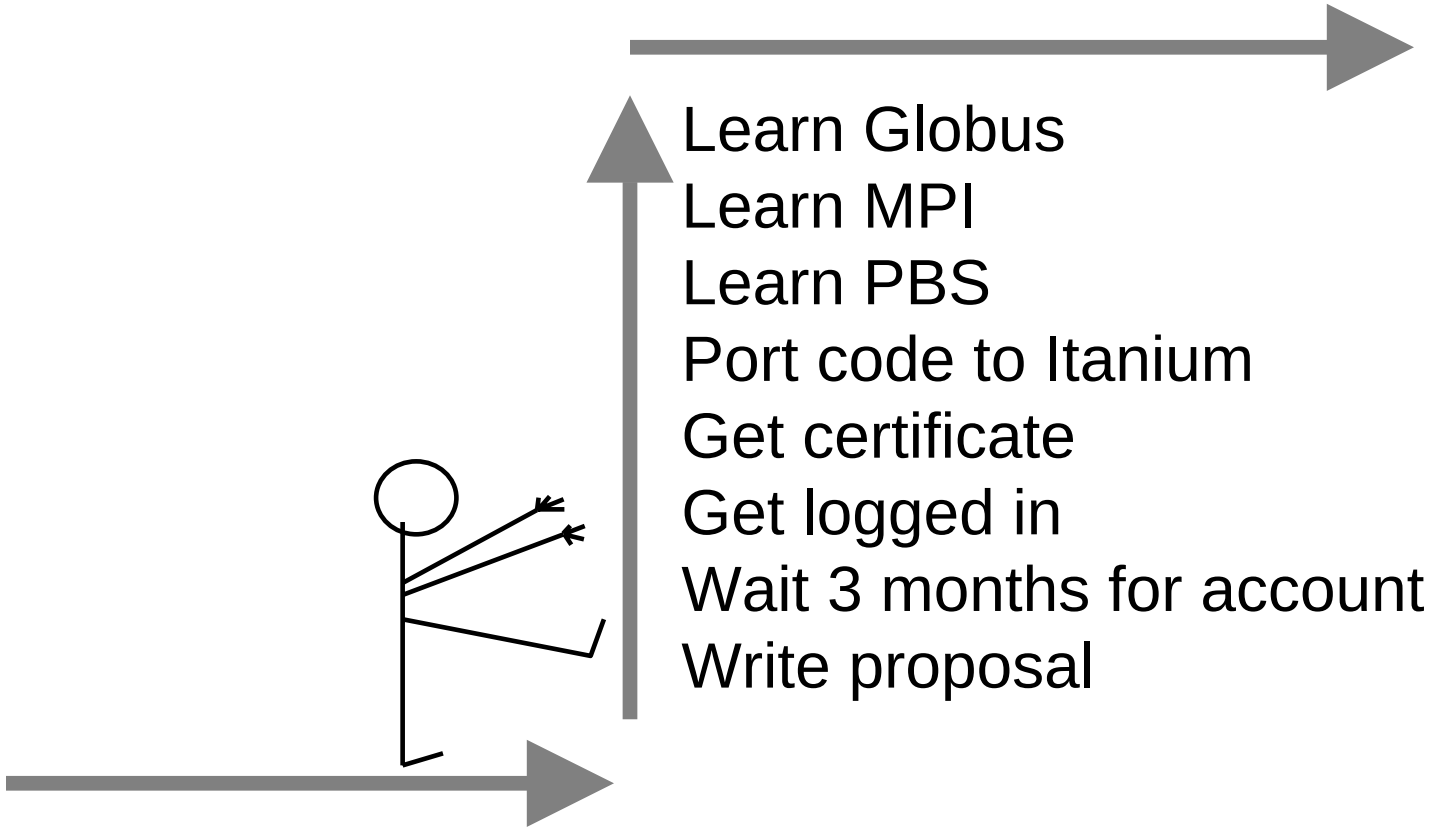





1. Portals and Workbenches
2.Knowledge & Resource Management
3.
Metadata View
Data View
Catalog Analysis
Bulk Data Analysis
Concept space
Standard APIs and Protocols
4.Grid Security Caching Replication Backup Scheduling
5.
Information Discovery
Metadata delivery
Data Discovery
Data Delivery
Standard Metadata format, Data model, Wire format
6.
Catalog Mediator
Data mediator
Catalog/Image
Specific Access
7.
Compute Resources
Derived Collections
Catalogs
Data Archives

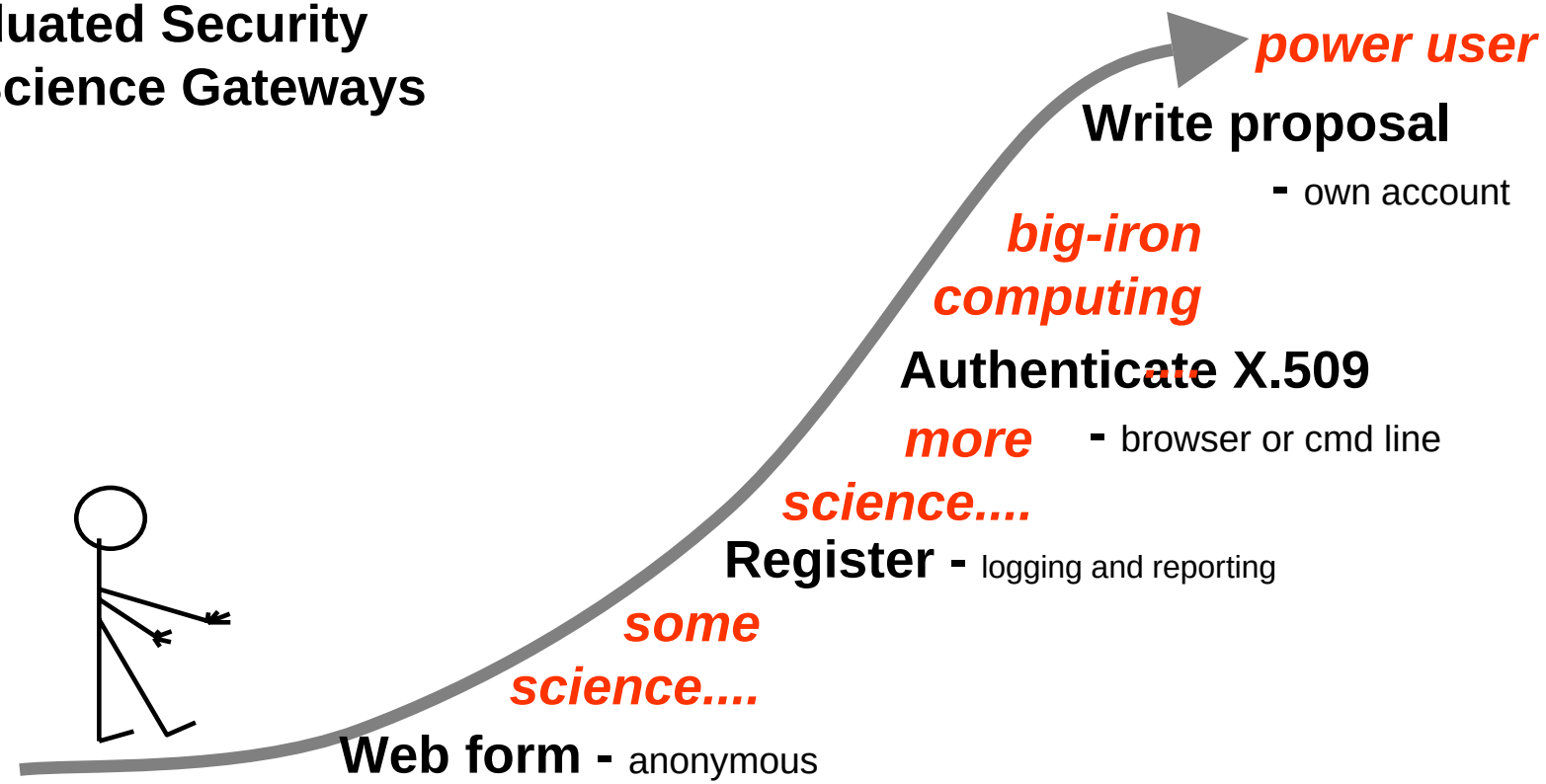
Teragrid Impediments

and now do some science....



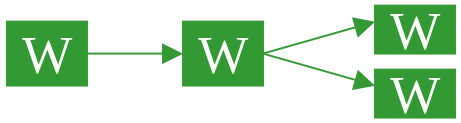
NVO Summer School 2004 Roy Williams presentation to user: Science Gateway - Future way

A better way:
Graduated Security for Science Gateways



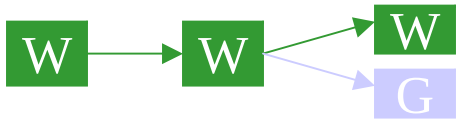
GGF tech in VObs? 4+1 options

1. No Grid, no way!



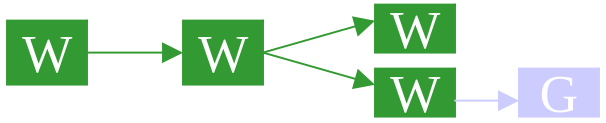
2. Grid throughout.

3. Grid services as leaf nodes.



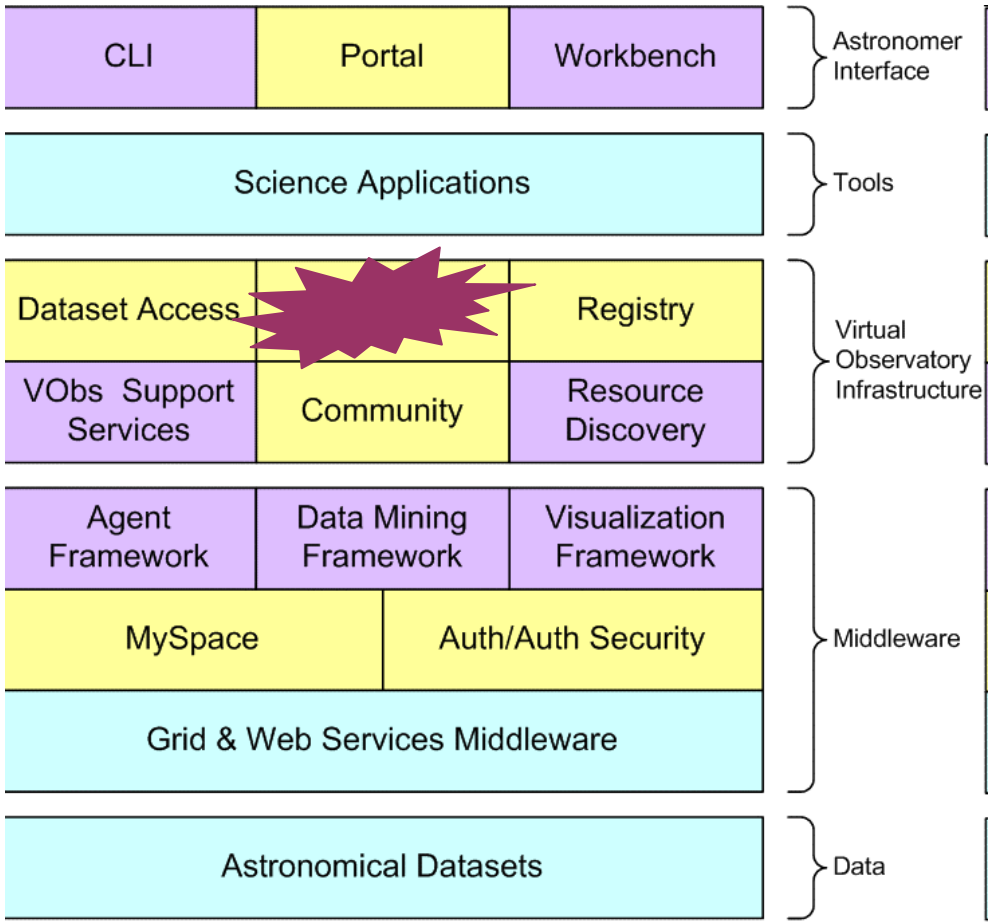
4. Leaf nodes + pervasive OGSA-SEC + pervasive GridFTP.

5. Grid only “within” web services



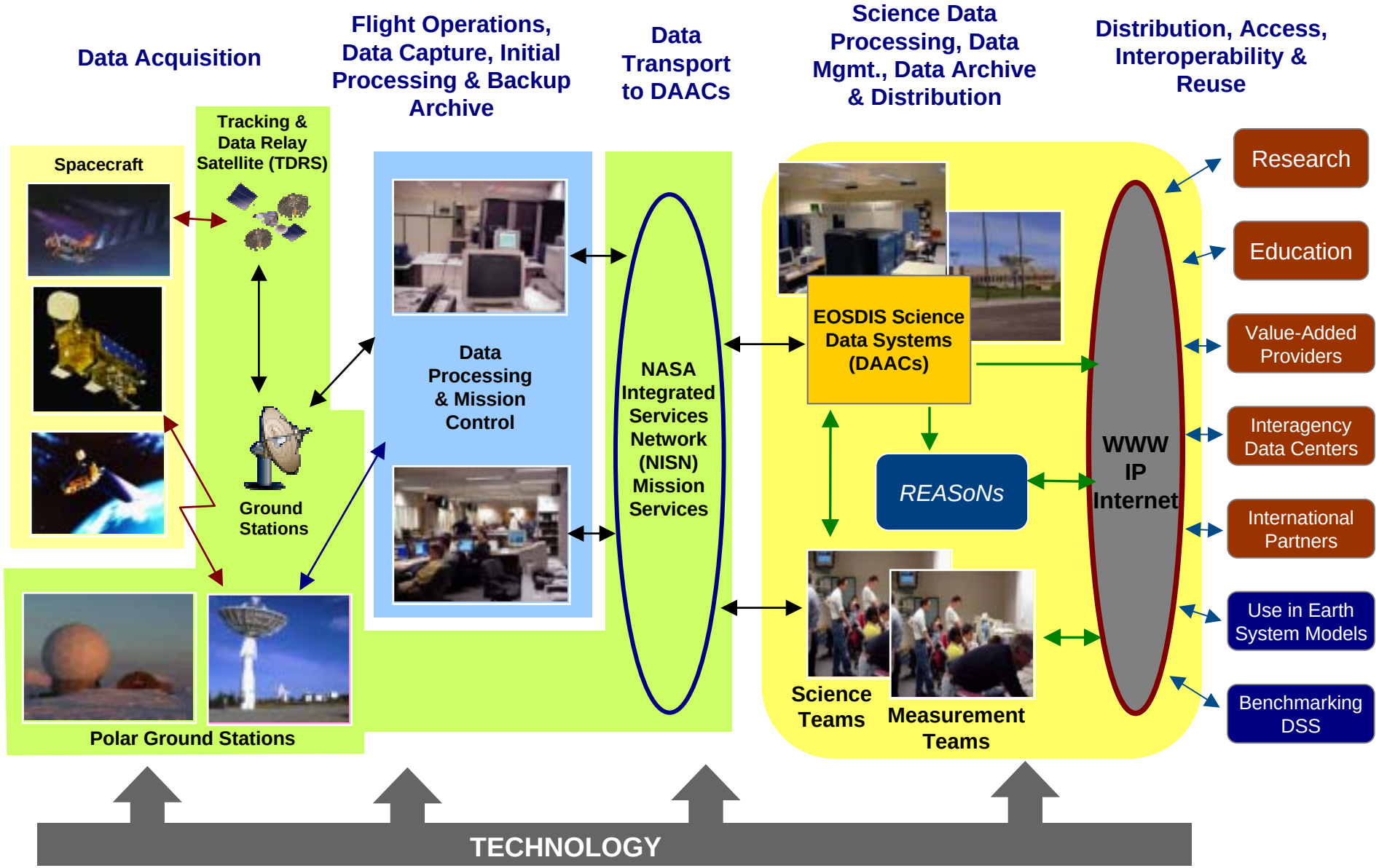
Preferred by GWS-WG of IVOA (options 4 & 5 are not mutually exclusive)

AstroGrid-2 Layer Diagram



ES Data System Architecture (current)

3/18/2005



See SpaceGrid presentation 2001
… From Final SpaceGrid report

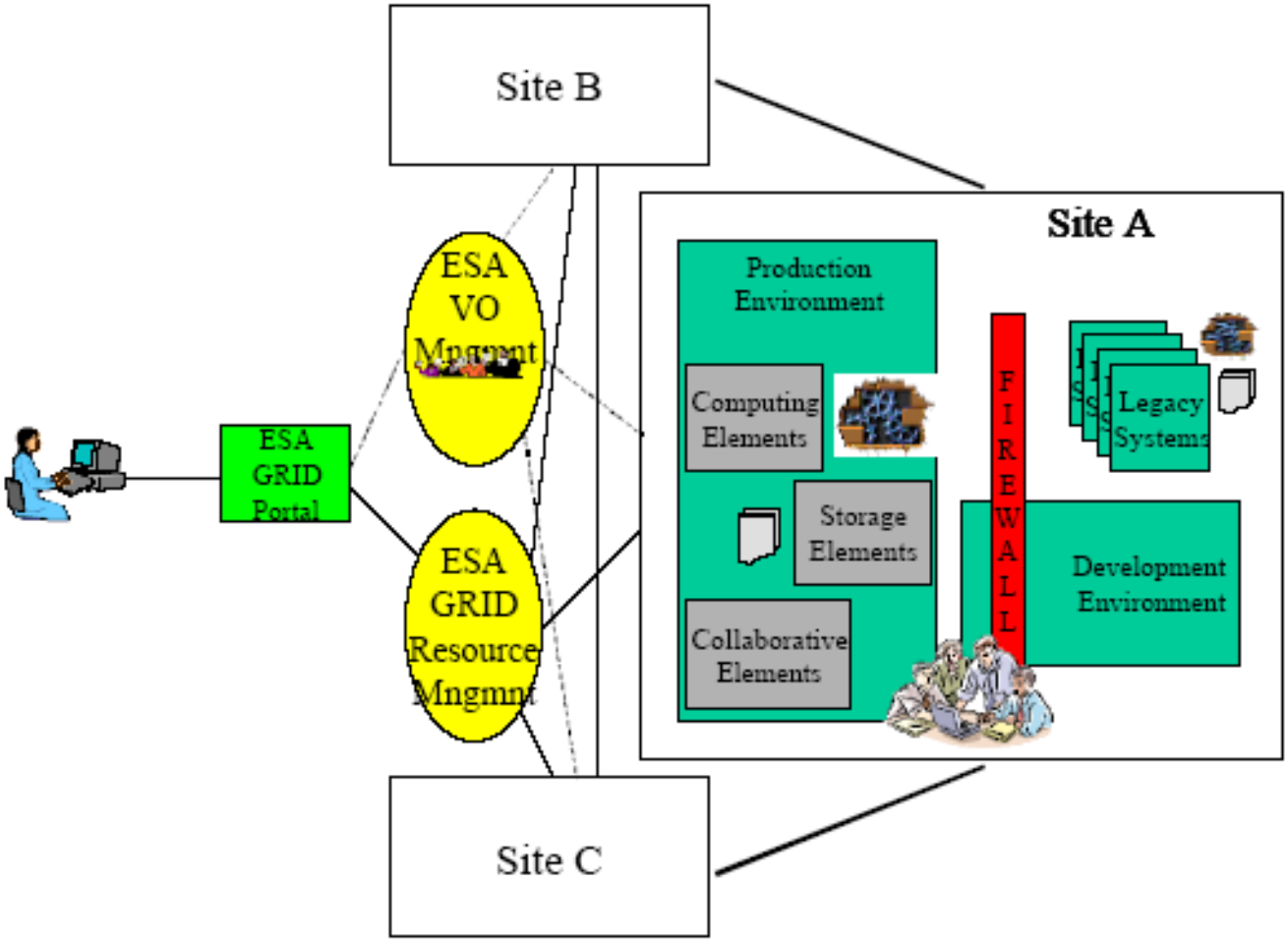


Figure 7: High-Level SpaceGRID Infrastructure

More from Final SpaceGRID report

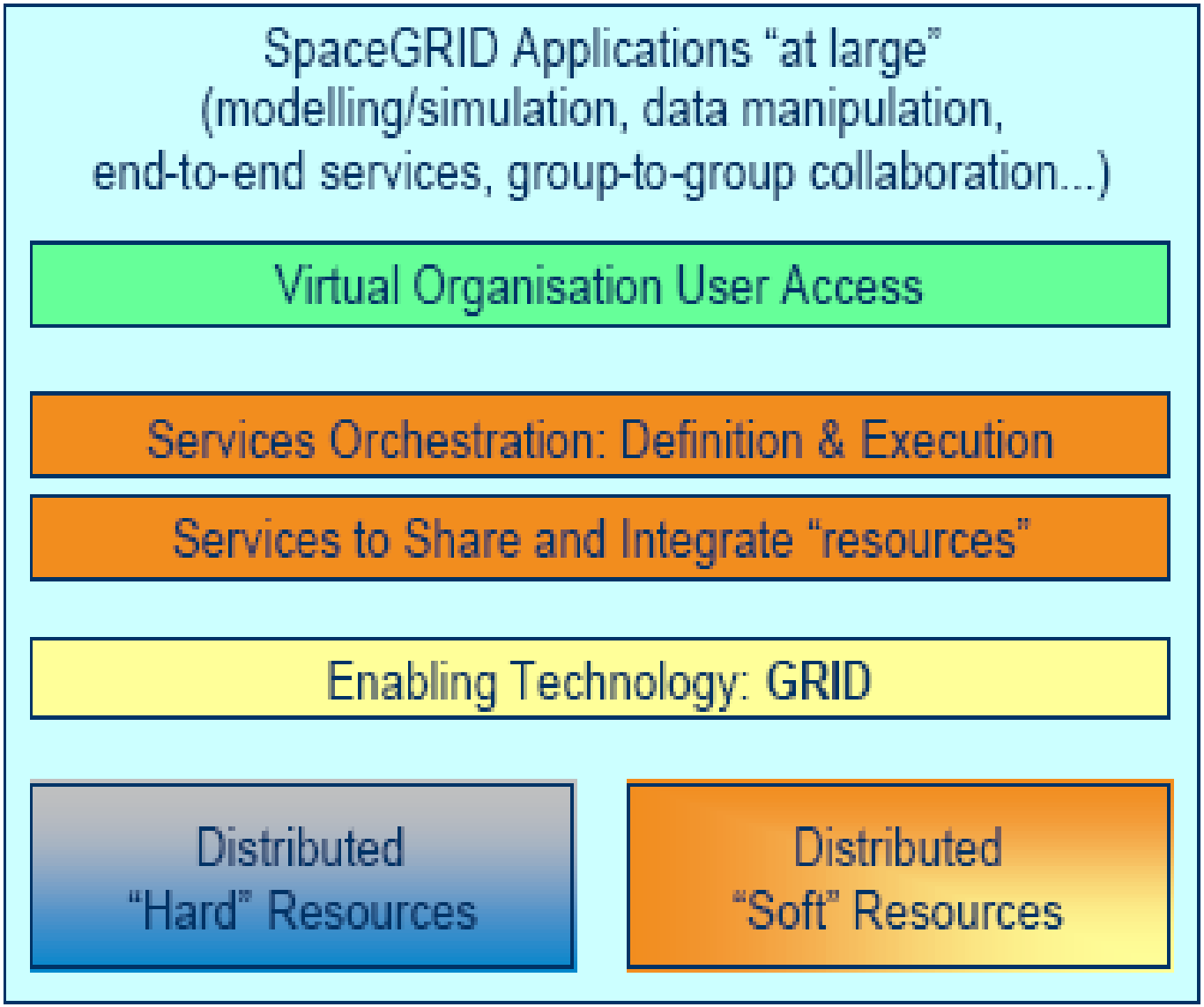


Figure 6: ESA wide SpaceGRID infrastructure with different flavours

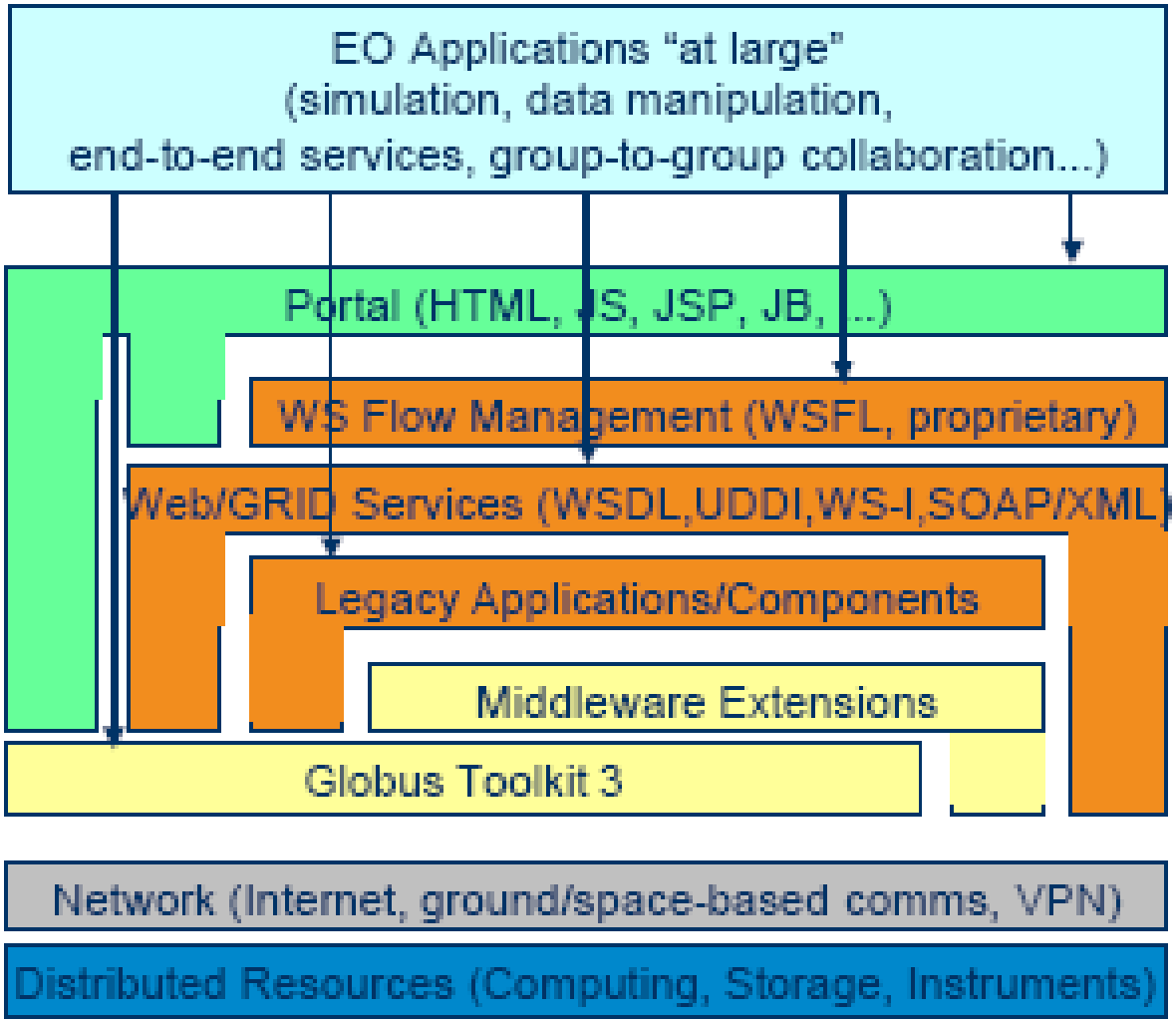
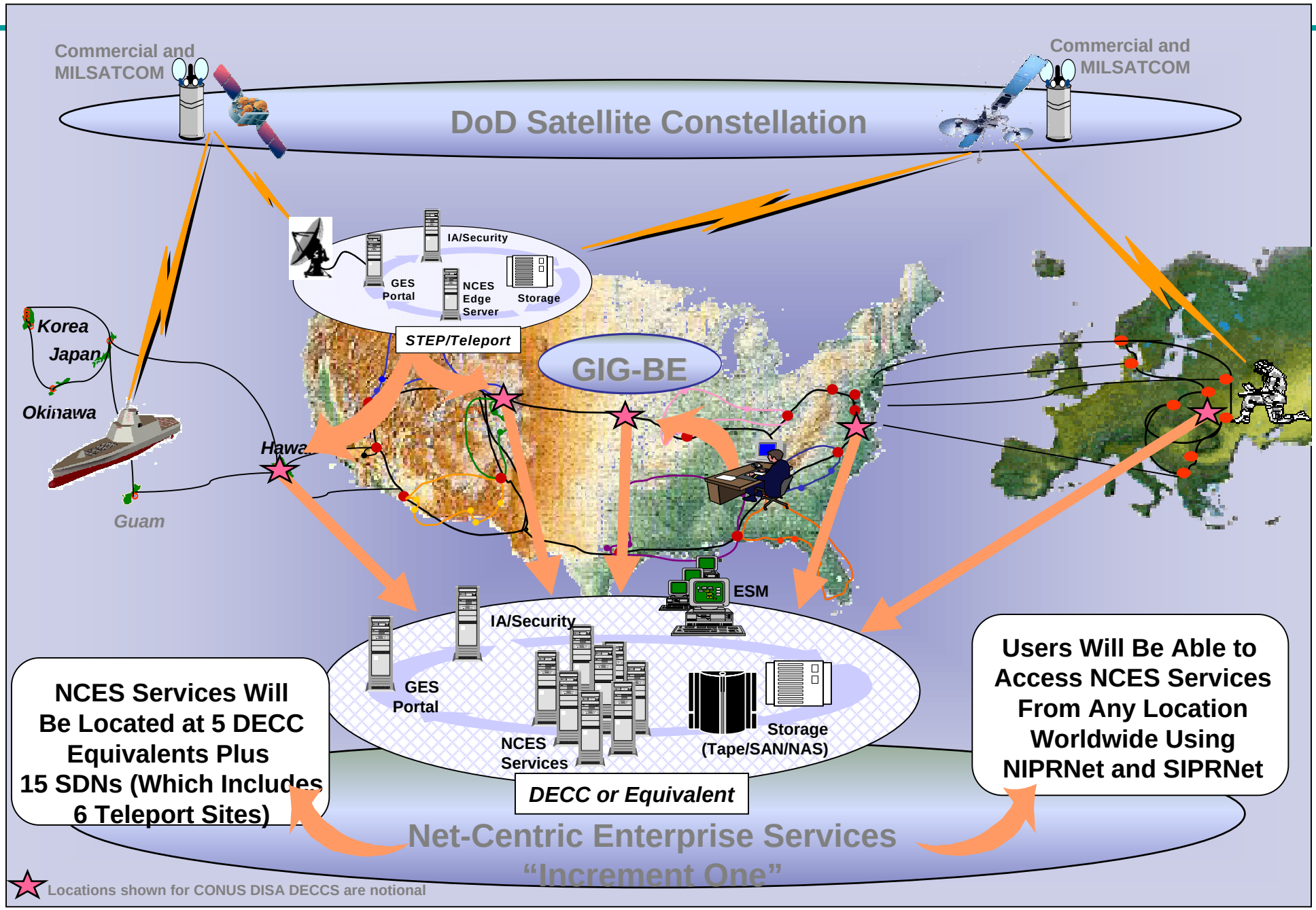
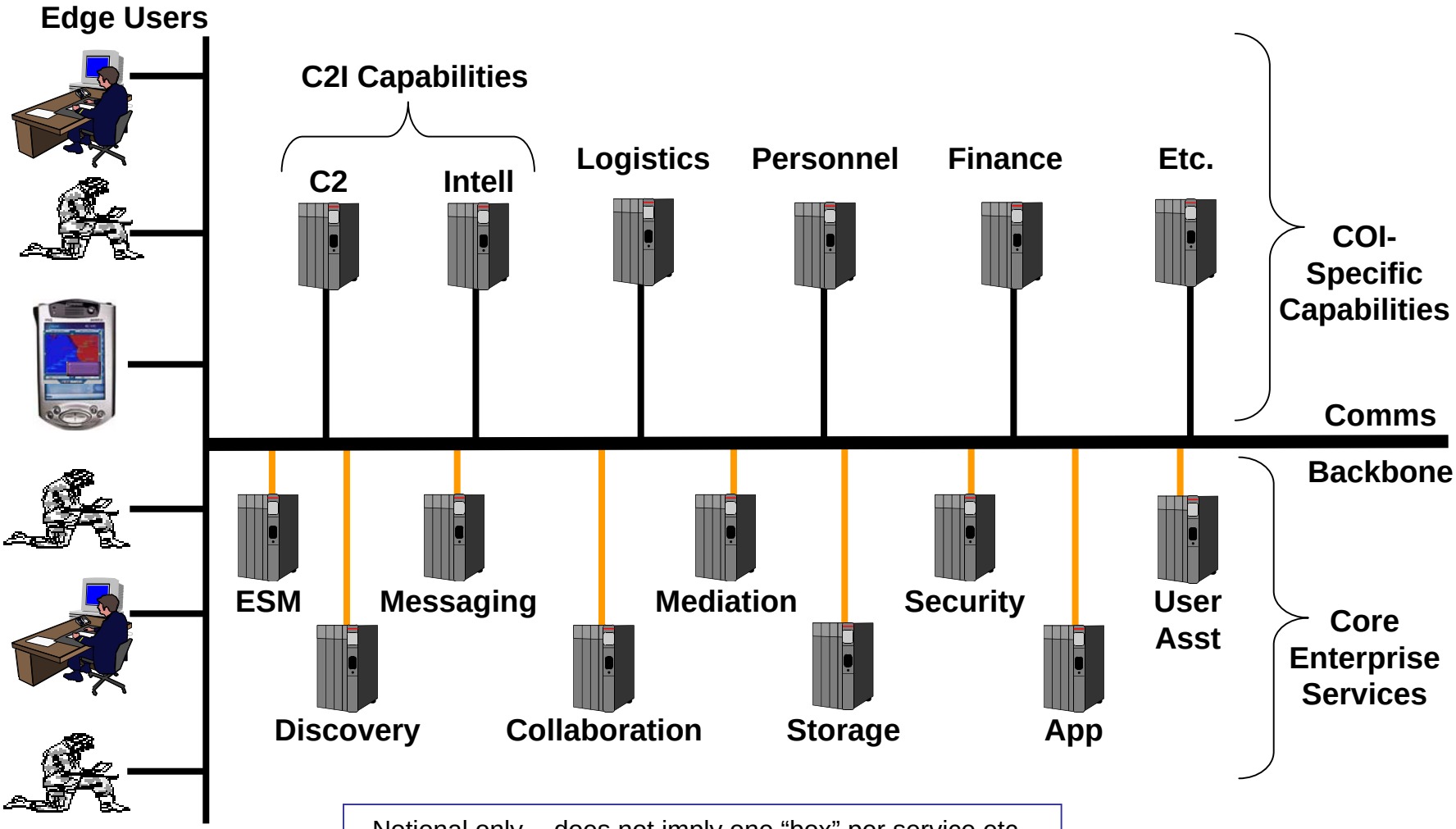


Figure 15: Technological view of Grid layered architecture for EO

DOD NCES (Network Centric Enterprise Services) Operational Concept



NCES Vision



Notional only - does not imply one "box" per service etc.

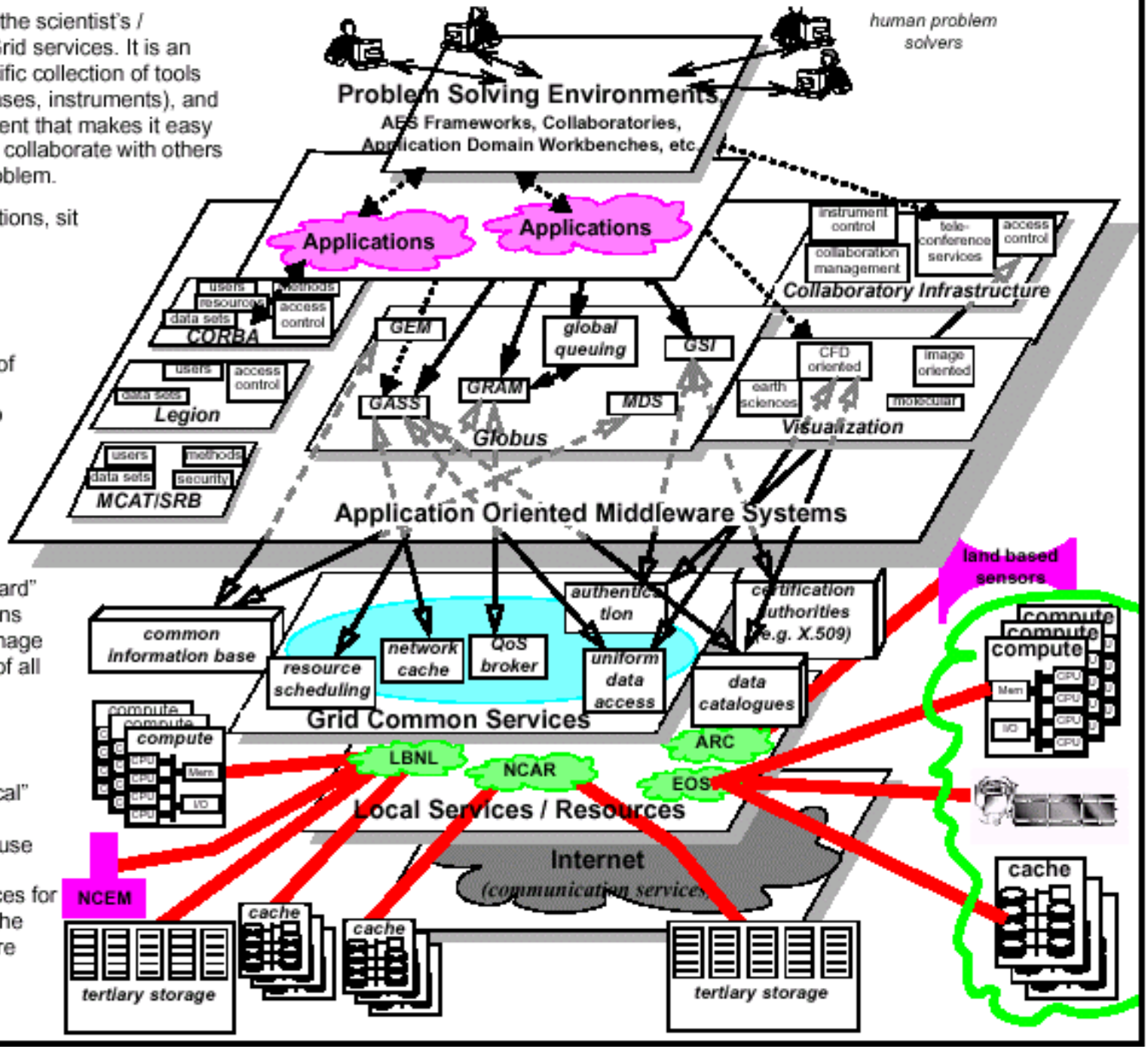
The PSE layer provides the scientist's / engineer's interface to Grid services. It is an application domain-specific collection of tools (e.g. simulations, databases, instruments), and a "workbench" environment that makes it easy to use those tools and to collaborate with others working on the same problem.

Applications, e.g. simulations, sit below the PSE and use middleware services.

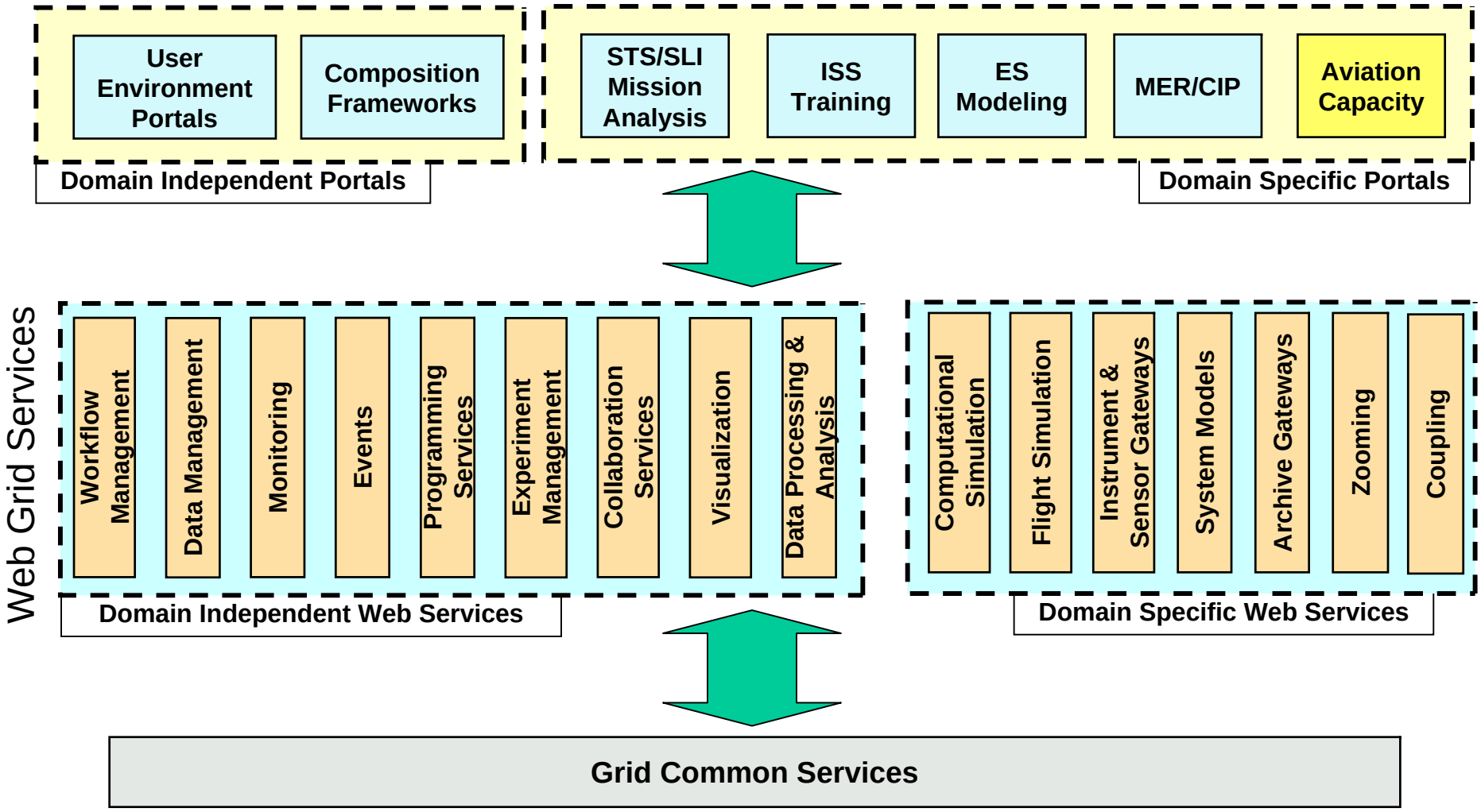
The middleware layer provides different styles of service interfaces for application developers to access the basic Grid services.

Grid services are "standard" interfaces for the functions needed to build and manage distributed applications of all sorts.

Most "resources" are "local" and will have their own resource managers and use policies. It is the use mechanisms and interfaces for the local resources that the Grid common services are intended to homogenize.



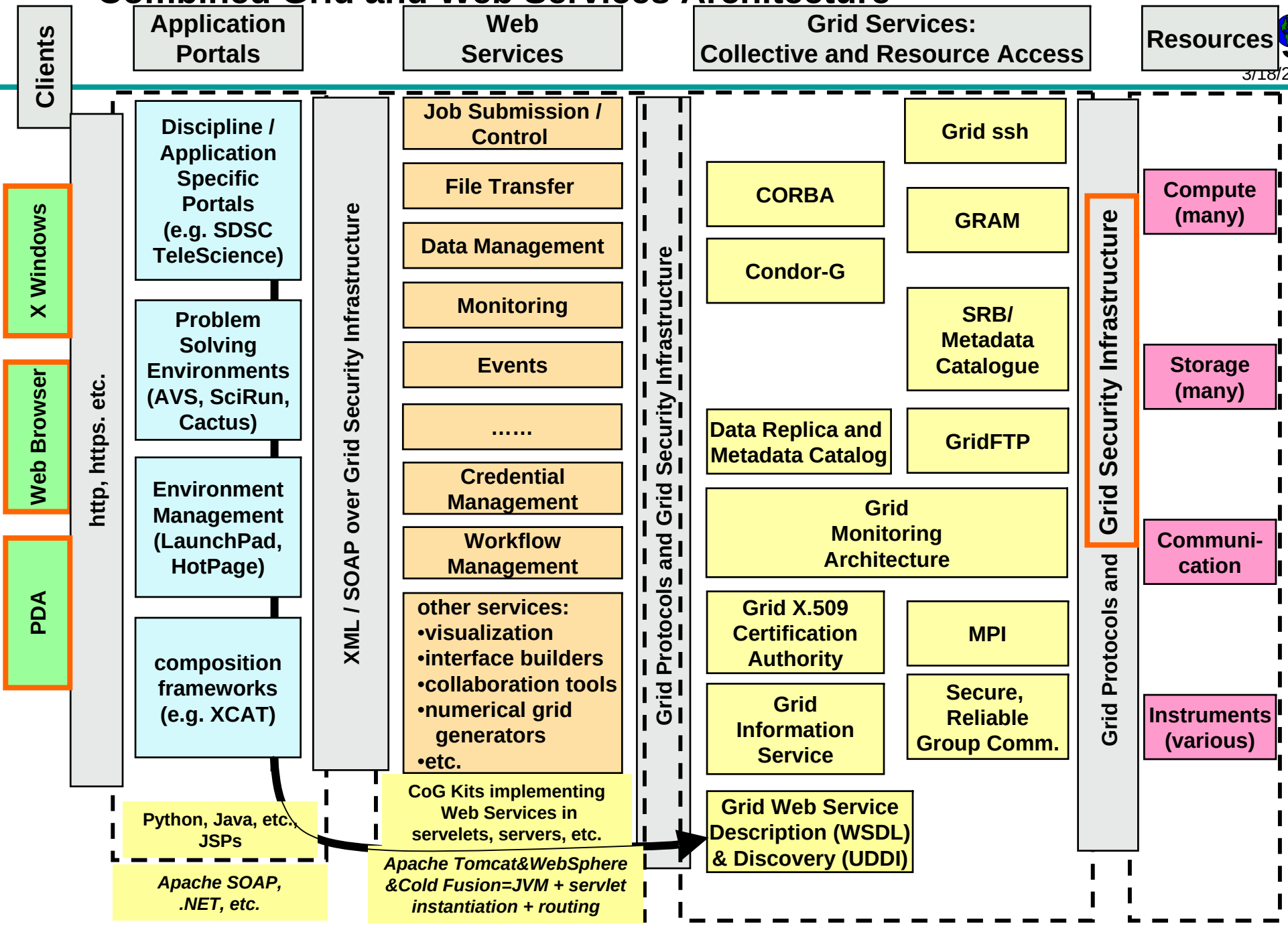
Grid Architecture – Information Environments



Piyush Mehrotra, NASA Ames

So: GGF4 Arch WG presentation http://grid.lbl.gov/GPA/GPA.GGF-4.1.ppt

Combined Grid and Web Services Architecture



So: GGF4 Arch WG presentation http://grid.lbl.gov/GPA/GPA.GGF-4.1.ppt

Grid Applications in a Generic Space Ops Architecture