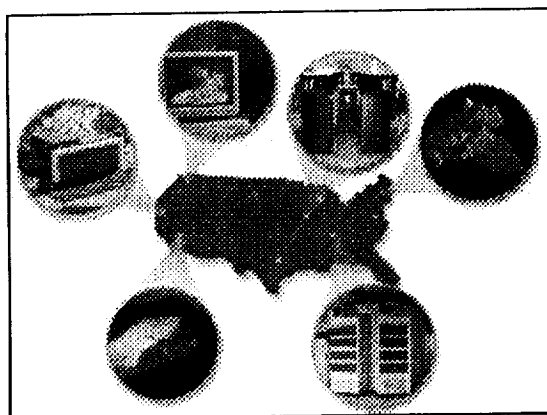Global Grid Forum - 1
The State of NASA's Information Power Grid

William E. Johnston, Project Manager
Arsi Vaziri, Deputy Project Manager
Leigh Ann Tanner, Implementation Task Manager

**The NAS Division
NASA Ames Research Center**
William J. Feiereisen, Division Chief
William Thigpen, Engineering Branch Chief,
http://www.ipg.nasa.gov

Application Classes Motivating Widely Distributed Computing Environments

Multi-disciplinary Simulations

Multi-disciplinary simulations provide a good *example of a class of applications that are very likely to require aggregation of widely distributed computing, data, and intellectual resources.*

Such simulations – e.g. whole system aircraft simulation and whole system living cell simulation – *require integrating applications and data that are developed by different teams of researchers* frequently in different locations.

The research teams are the only ones that have the expertise to maintain and improve the simulation code and/or the body of experimental data that drives the simulations. This results in an inherently distributed computing and data management environment.

Consider a vision for Aviation Safety:

How do we simulate the entire commercial airspace of the country?

(Yuri Gawdiak (VNAS) and Bill McDermott, NASA Ames, John Lytle and Gregory Follen, NASA Glenn (NPSS)).

This vision is being approached through a set of increasingly complex and computationally intensive integrations:

◆ **Component simulations are combined to get a system simulation.**
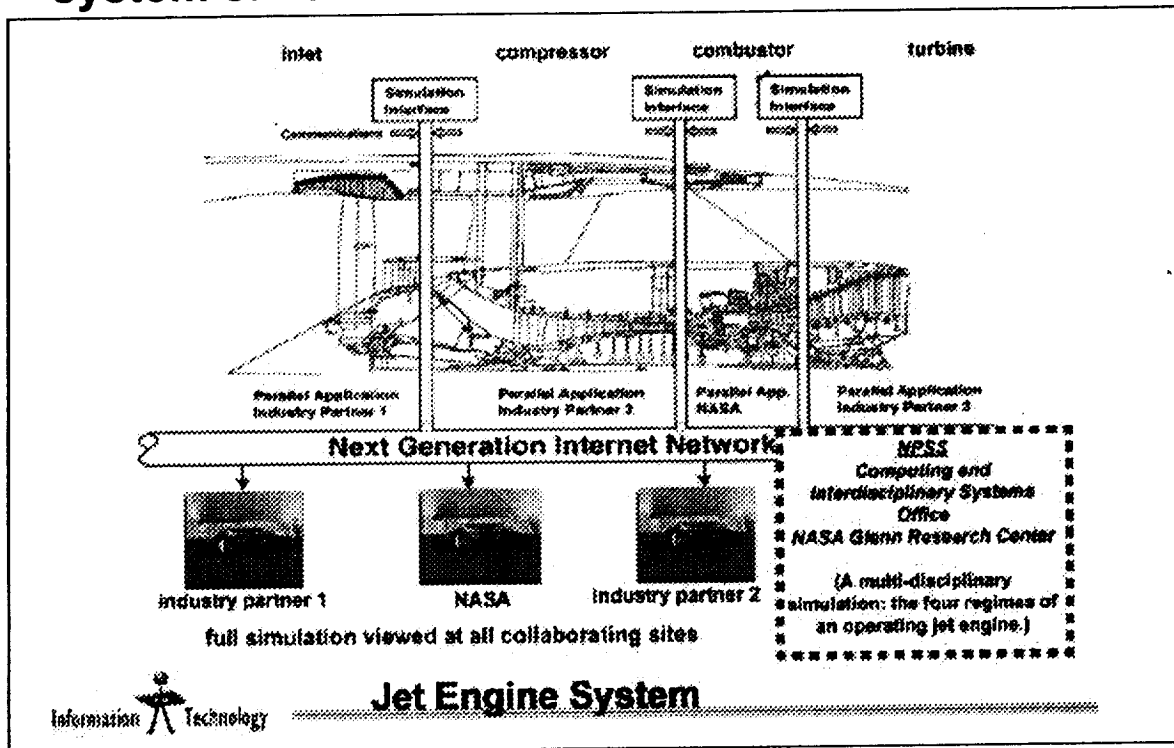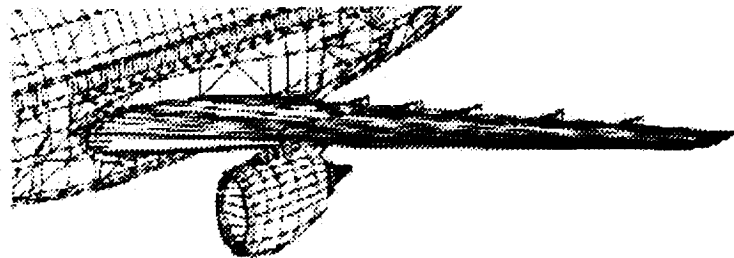

Jet Engine System

◆ **Multiple system simulations are coupled to represent pieces of a device.**

Multi-System Simulation

Engine System + Wing System

◆ Whole device simulations are produced by coupling all of the subordinate system simulations.
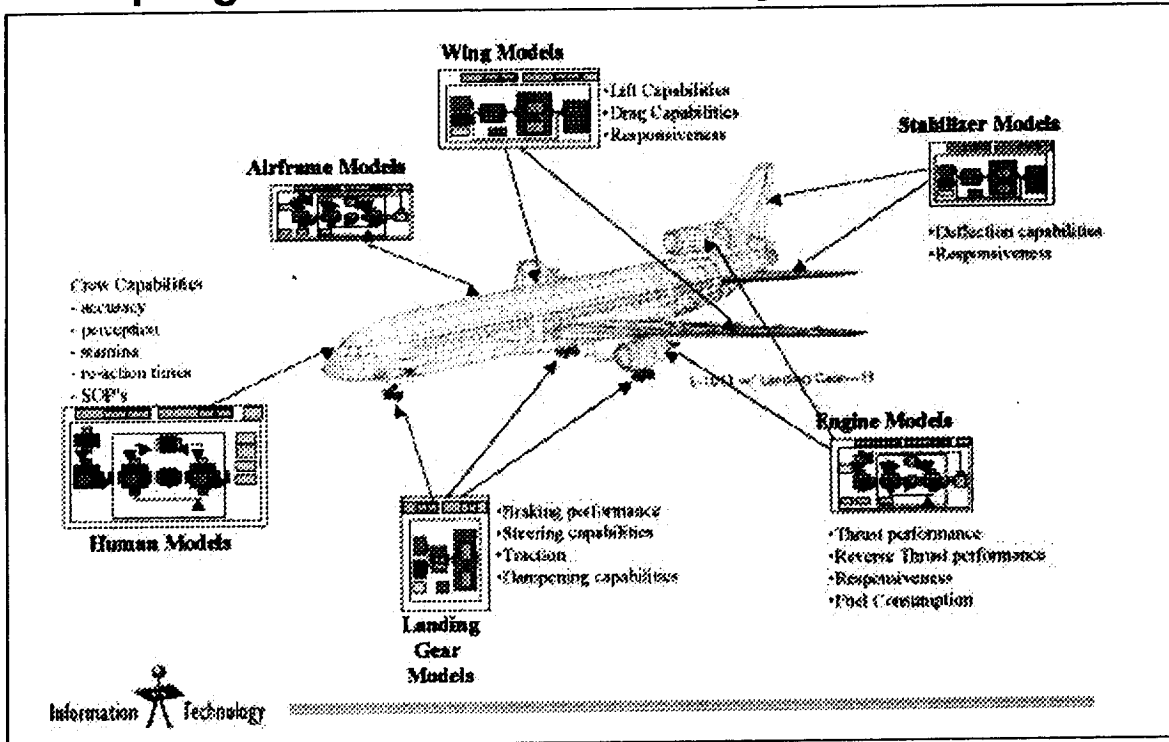
◆ Devices are inserted into a realistic environment.

National Air Space (NAS) Simulation Environment

Virtual
National Air
Space
VNAS
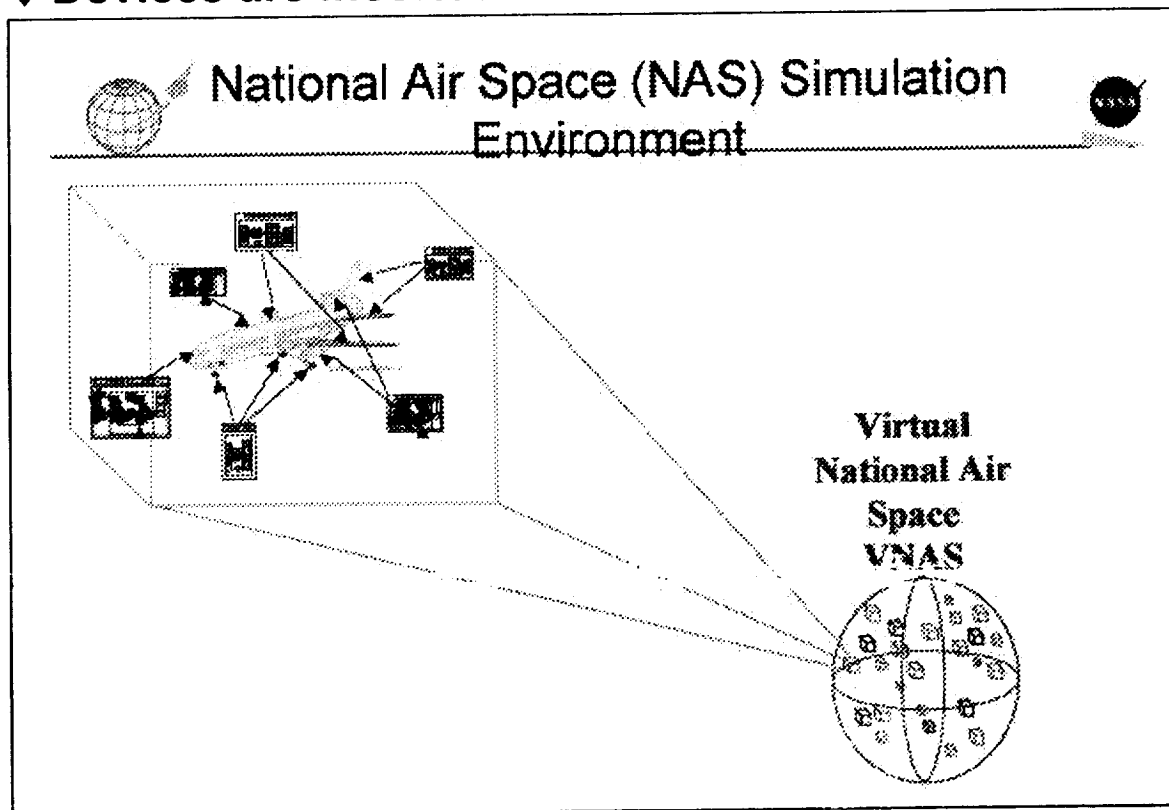
♦ Devices and environment are combined for operational systems simulation.

Daily NAS Simulation Baseline Generation
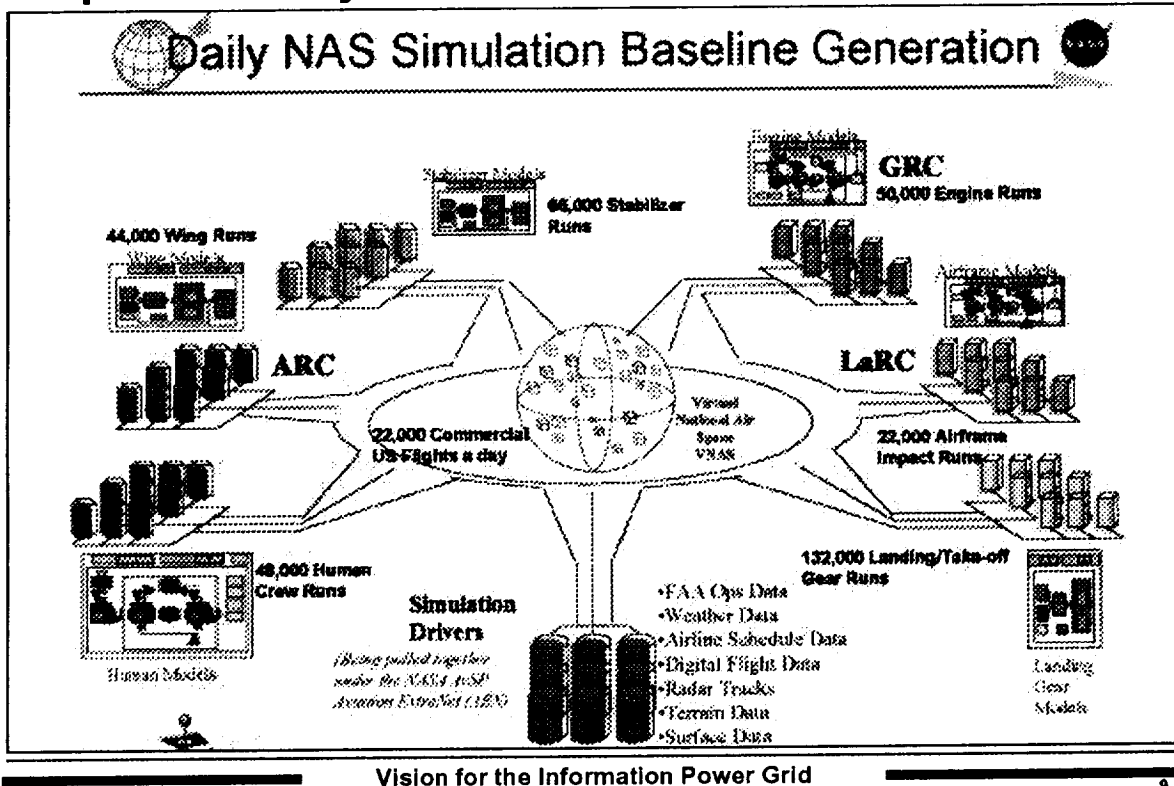
Multi-Disciplinary Simulations

Clearly such simulations will need to use aggregated computing, data, instrument, and intellectual resources across multiple NASA Centers.

Issues for combining component simulations

+ wrapping the simulation code

+ composing these codes

+ locating and coordinating resources for executing the multiple components and managing the resulting data (which is likely to be distributed)

Issues for multi system simulations

+ multi-Center interactions - component parameters maintained by discipline experts

+ multiple sources of data and data expertise

+ shared compute and data resources

CORBA and IDL are useful for addressing the first two points.

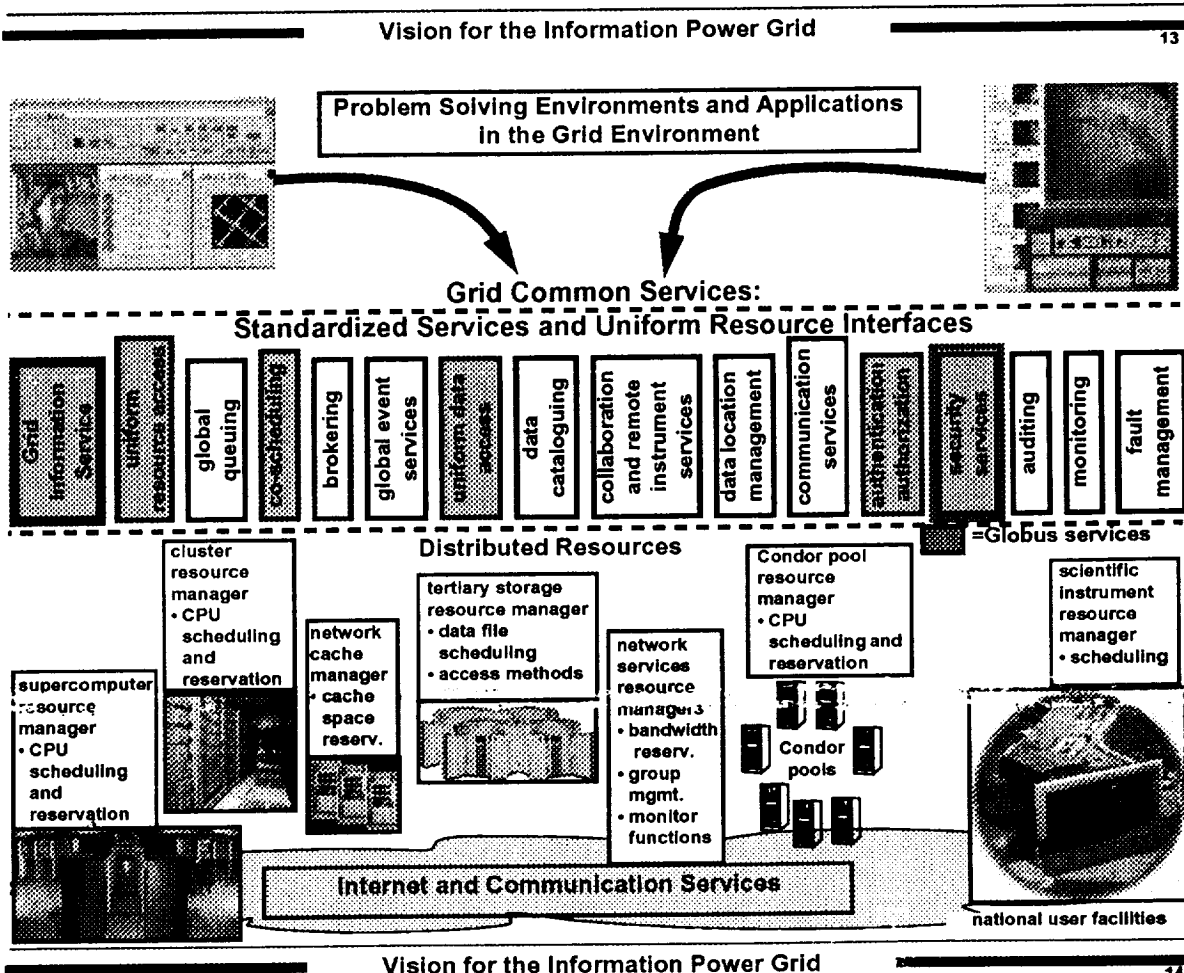IPG is intended to provide the basic framework for resource sharing and management across sites:

+ discovery

+ scheduling

+ access to, and

+ policy enforcement

with respect to computing systems, data management, and collaboration systems
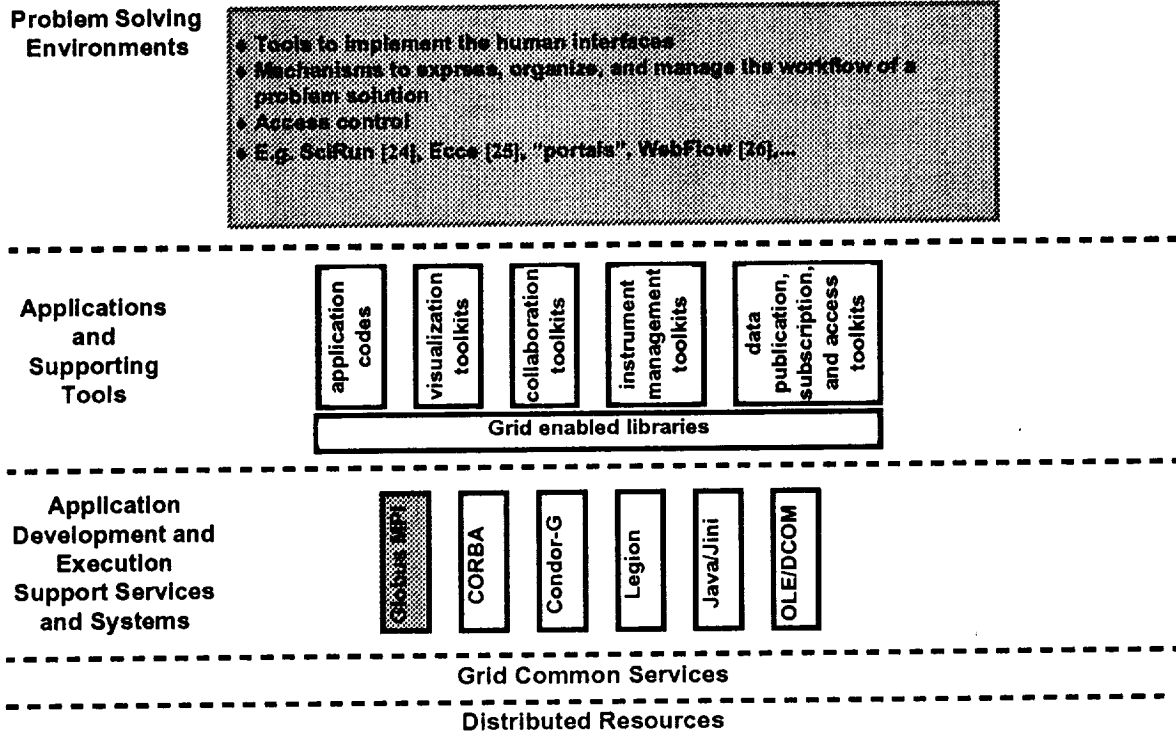
What are Grids?

Grids are tools, middleware, and services for

+ **providing a uniform look and feel to a wide variety of computing and data resources**

+ **supporting construction, management, and use of widely distributed application systems**

+ **facilitating human collaboration and remote access and operation of scientific and engineering instrumentation systems**

+ **managing and securing the computing and data infrastructure**

Problem Solving Environments and Applications in the Grid Environment

Grid Common Services:
Standardized Services and Uniform Resource Interfaces

Grid Information Service · uniform resource access · global queuing · co-scheduling · brokering · global event services · uniform data access · data cataloguing · collaboration and remote instrument services · data location management · communication services · authentication authorization · security services · auditing · monitoring · fault management

= Globus services

Distributed Resources

cluster resource manager
· CPU scheduling and reservation

network cache manager
· cache space reserv.

tertiary storage resource manager
· data file scheduling
· access methods

network services resource managers
· bandwidth reserv.
· group mgmt.
· monitor functions

Condor pool resource manager
· CPU scheduling and reservation

Condor pools

scientific instrument resource manager
· scheduling

supercomputer resource manager
· CPU scheduling and reservation

Internet and Communication Services

national user facilities

Problem Solving Environments
- Tools to implement the human interfaces
- Mechanisms to express, organize, and manage the workflow of a problem solution
- Access control
- E.g. SciRun [24], Ecce [25], "portals", WebFlow [36]...

Applications and Supporting Tools

application codes	visualization toolkits	collaboration toolkits	instrument management toolkits	data publication, subscription, and access toolkits

Grid enabled libraries

Application Development and Execution Support Services and Systems

Globus MPI	CORBA	Condor-G	Legion	Java/Jini	OLE/DCOM

Grid Common Services

Distributed Resources

What Grids Will and Will Not Do

Grids provide common resource access technology and operational services deployed across virtual organizations. This *allows the possibility of sharing resources, but does not automatically permit it*:

+ Local authorization models are not changed by the Grid.

+ Common Grid technology will allow common views of resources and uniform access to resources, thereby permitting very large application systems to be built, and <u>if policy permits</u>, to share resources across sites and organizations.

Grids *will enable large scale applications based on*:

+ *Loosely coupled computations*: Simulation parameter sweeps and certain types of experiment data analysis involve initiating and managing 100s, 1000s, and 10000s of processes. Grids provide the access and mechanisms for using large numbers of computing and data resources for this type of calculation.

Grids *will enable large scale applications based on*:

+ *Large scale pipelined applications*: Multi-component simulations involve executing multiple, coupled, medium to large scale simulations on multiple computing resources. Grids provide co-scheduling and data stream management to support this.

Grids will not, in the near term, enable very large, single problems such as CFD calculations to be spread across distributed systems.

+ To accomplish this we will need new approaches and algorithms that are tolerant of high and variable latency. There is R&D going on to address this issue in the long term.

Grids will not provide a lot of "free" resources.

+ *To produce a highly capable science Grid organizations must place major resources on the Grid.*

Approach and Goals for NASA's IPG

- Grids are built through collaborative efforts, and at the same time facilitate collaboration: *IPG is a collaboration among several NASA Centers and the NSF Supercomputer Center consortia (PACIs), with the Grid Forum providing "coordination" of many institutions world wide*

- *Deployment of existing technology (Globus [1], Condor [23], Grid portals [13], etc.) is providing for relatively rapid impact – IT/ANCS computing and storage resources are providing the prototype production IPG environment*

- ***Strong security is being provided from the start*** in order to address authentication, authorization, and infrastructure assurance in open science networks for both applications and Grid services

How is IPG Being Accomplished?

Persistent operational environment that encompasses significant resources across the three Centers - there are groups at NAS whose responsibilities are:

- Grid Information Services
    - + MDS operations
    - + configuration databases

- IPG systems
    - + Globus configuration and deployment
    - + large system integration

- **Security**
  - **+ X.509 CA operation**
  - **+ security model**
  - **+ GSI services**

- **Remote Data Access**
  - **+ IPG Archival Storage, GSIftp, SRB/MCAT**
 - **+ security model**

- **Portals and monitoring**

- **User Services**
  - **+ consulting**
  - **+ support model**
  - **+ documentation**
  - **+ CORBA support**

- **Condor**

- **Testing**
  - **+. regression testing, verification suites, benchmarks, and reliability/sensitivity analysis**

- **PBS**
  - **+ Global queuing and user-level queue management capability on top of Globus**

- **Networking**
  - **+ IPG testbed connection Ames, GRC, LaRC**
  - **+ high-speed testbed**

- **Accounting**
  - **+ account management (automated generation and maintenance mechanisms)**
  - **+ standardized accounting records**

The State of IPG

- *Computing resources*: ≈600 CPU nodes in half a dozen SGI Origin 2000s and several workstation clusters at Ames, Glenn, and Langley, with plans for incorporating Goddard and JPL, and ≈200 nodes in a Condor pool

- Wide area *network interconnects* of at least 100 mbit/s

- *Storage resources*: 50-100 Terabytes of archival information/data storage *uniformly and securely* accessible from all IPG systems via MCAT/SRB and GSIftp / Gridftp

IPG Baseline System and
High Data-Rate (DX) Testbed
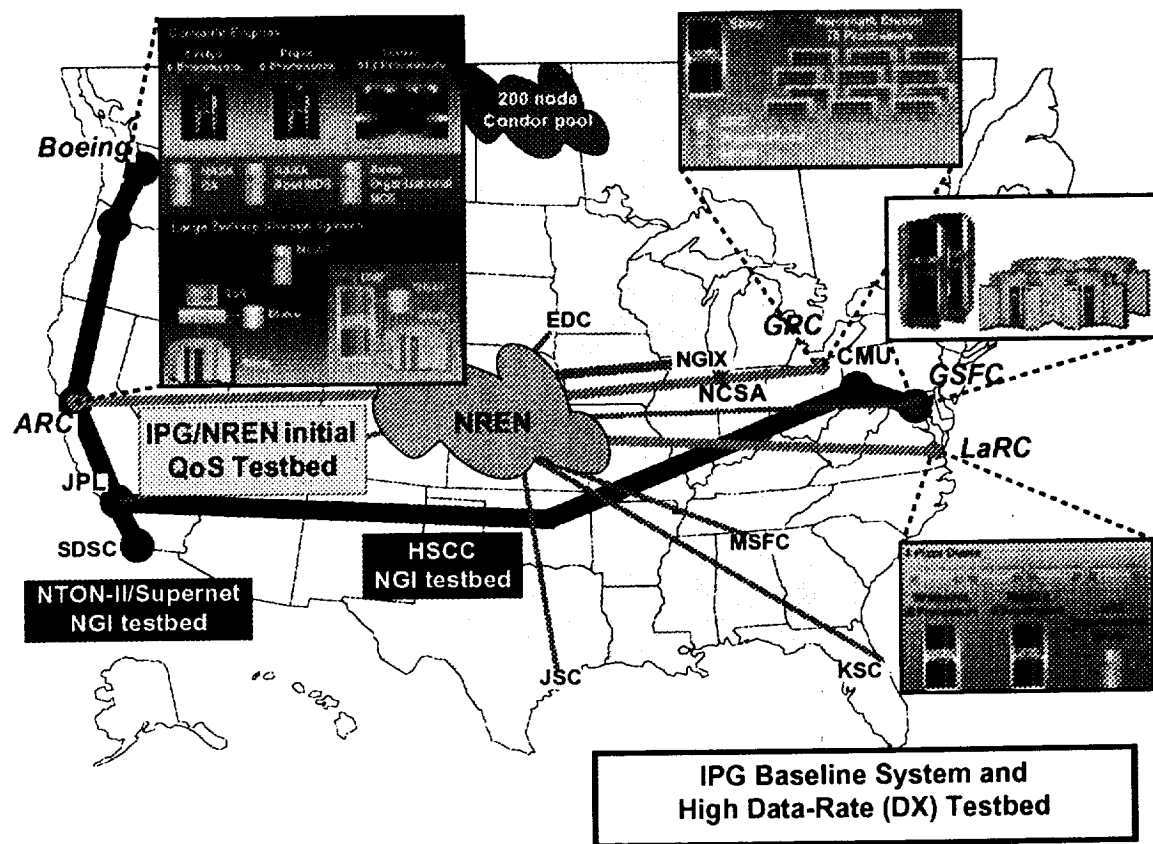
State of IPG: Baseline Operational System (cont.)

- *Globus* providing the Grid common services

- *Programming and program execution support*
  + Grid MPI (via the Globus communications library)
  + CORBA integrated with Globus
  + global job queue management
  + high throughput job manager
  + Condor [23] ("cycle stealing" computing)

- A *stable and supported operational environment*

- Several "benchmark" *applications operating across*
  IPG (multi-grid CFD code, parameter study)

- Multi-Grid operation (applications operating across
  IPG and NCSA)

- **IPG Functionality Tasks**
  - + **CORBA in the IPG environment**
  - + **Integration of Legion**
  - + **CPU resource reservation**
  - + **High Throughput Computing**
  - + **Programming Services**
  - + **Distributed debugging**
  - + **Grid enabled visualization**
  - + **Parameter study frameworks**
  - + **Network bandwidth reservation**

- **Characteristic Applications**
  - + **OVERFLOW**
  - + **NPSS**
  - + **INS3D**
  - + **molecular analysis**

IPG has met it's first three Level-1 milestones

References and Acronyms

[1] Globus is a middleware system that provides a suite of services designed to support high performance, distributed applications. Globus provides:
- Resource Management: Components that provide standardized interfaces to various local resource management systems (GRAM) manage allocation of collections of resources (DUROC). All Globus resource management tools are tied together by a uniform resource specification language (RSL).
- Remote Access: Components that enable remote access to files (GASS and RIO) and executables (GEM).
- Security: Support for single sign-on, authentication, and authorization within the Globus system (GSI) and (experimentally) authorization (GAA).
- Fault Detection: Basic support for building fault detection and recovery into Globus applications.
- Information Infrastructure: Global access to information about the state and configuration of system components of an application (MDS).
- Grid programming services: Support writing parallel-distributed programs (MPICH-G), monitoring (HBM), etc.

www.globus.org provides full information about the Globus system.

[2] *The Grid: Blueprint for a New Computing Infrastructure*, edited by Ian Foster and Carl Kesselman. Morgan Kaufmann, Pub. August 1998. ISBN 1-55860-475-8. http://www.mkp.com/books_catalog/1-55860-475-8.asp

[3] "Grids as Production Computing Environments: The Engineering Aspects of NASA's Information Power Grid," William E. Johnston, Dennis Gannon, and Bill Nitzberg. Eighth IEEE International Symposium on High Performance Distributed Computing, Aug. 3-6, 1999, Redondo Beach, California. (Available at http://www.nas.nasa.gov/~wej/IPG)

[4] "Vision and Strategy for a DOE Science Grid" - http://www.itg.lbl.gov/~wej/Grids

[5] See www.nas.nasa.gov/IPG for project information and pointers.

[6] See http://www-itg.lbl.gov/NGI/ for project information and pointers.

[7] The Particle Physics Data Grid has two long-term objectives. Firstly: the delivery of an infrastructure for very widely distributed analysis of particle physics data at multi-petabyte scales by hundreds to thousands of physicists. Secondly: the acceleration of the development of network and middleware infrastructure aimed broadly at data-intensive collaborative science. http://www.cacr.caltech.edu/ppdg/

[8] Tierney, B. Lee, J., Crowley, B., Holding, M., Hylton, J., Drake, F., "A Network-Aware Distributed Storage Cache for Data Intensive Environments", Proceeding of IEEE High Performance Distributed Computing conference (HPDC-8), August 1999.

[9] "Real-Time Generation and Cataloguing of Large Data-Objects in Widely Distributed Environments," W. Johnston, Jin G., C. Larsen, J. Lee, G. Hoo, M. Thompson, and B. Tierney (LBNL) and J. Terdiman (Kaiser Permanente Division of Research). Invited paper, International Journal of Digital Libraries - Special

Issue on "Digital Libraries in Medicine". May, 1998. http://www-itg.lbl.gov/WALDO/

[10] MAGIC: "The MAGIC Gigabit Network." See: http://www.magic.net

[11] TerraVision-2: VRML based data fusion and browsing -
www.ai.sri.com/TerraVision

[12] "A Monitoring Sensor Management System for Grid Environments," Brian Tierney,
Brian Crowley, Dan Gunter, Mason Holding, Jason Lee, Mary Thompson. To
appear, HPDC-9, July, 2000. Available at http://www-didc.lbl.gov/JAMM/

[13] A collaborative effort to enable desktop access to remote resources including,
supercomputers, network of workstations, smart instruments, data resources, and
more - computingportals.org

[14] "The Data Grid: Towards an Architecture for the Distributed Management and
Analysis of Large Scientific Datasets." A. Chervenak, I. Foster, C. Kesselman, C.
Salisbury, S. Tuecke, (to be published in the Journal of Network and Computer
Applications).

[15] "Storage Access Coordination Using CORBA," A. Sim, H. Nordberg, L.M.
Bernardo, A. Shoshani and D. Rotem. Proceedings of the International Symposium
on Distributed Objects and Applications. See http://gizmo.lbl.gov/sm/

[16] The Clipper Project: Computational Grids providing middleware that supports
applications requiring configurable, distributed, high-performance computing and
data resources. See http://www-itg.lbl.gov/~johnston/Clipper

[17] The Grid Forum (www.gridforum.org) is an informal consortium of institutions and
individuals working on wide area computing and computational Grids. Current
working groups include Security (authentication, authorization), Scheduling and
Resource Management, Grid Information Services, Application and Tool
Requirements, Advanced Programming Models, Grid User Services and
Operations, Account Management, Remote Data Access, Grid Performance

[18] "New Capabilities in the HENP Grand Challenge Storage Access System and its
Application at RHIC" http://rncus1.lbl.gov/GC/docs/chep292lp1.doc
"STACS is ... responsible for determining, for each query request, which events and
files need to be accessed, to determine the order of files to be cached dynamically
so as to maximize their sharing by queries, to request the caching of files from
HPSS in tape optimized order, and to determine dynamically which files to keep in
the disk cache to maximize file usage."

[19] "DeepView: A Collaborative Framework for Distributed Microscopy." IEEE Conf. on
High Performance Computing and Networking, Nov. 1998. See http://vision.lbl.gov/
(projects -> collaborative computing)

[20] Akenti: "Certificate-based Access Control for Widely Distributed Resources,"
Mary Thompson, William Johnston, Srilekha Mudumbai, Gary Hoo, Keith Jackson,
Usenix Security Symposium '99. Mar. 16, 1999. (See http://www-itg.lbl.gov/Akenti)

[21] GAA: "Generic Authorization and Access control API" (GAA API). IETF Draft.
http://ghost.isi.edu/info/gss_api.html)

[22] Storage Resource Broker (SRB) provides uniform access mechanism to diverse and distributed data sources. http://www.sdsc.edu/MDAS/

[23] Condor is a High Throughput Computing environment that can manage very large collections of distributively owned workstations. http://www.cs.wisc.edu/condor/

[24] SCIRun is a scientific programming environment that allows the interactive construction, debugging and steering of large-scale scientific computations. http://www.cs.utah.edu/~sci/software/

[25] Ecce - www.emsl.pnl.gov

[26] WebFlow - A prototype visual graph based dataflow environment, WebFlow, uses the mesh of Java Web Servers as a control and coordination middleware, WebVM. See http://iwt.npac.syr.edu/projects/webflow/index.htm

[27] "QoS as Middleware: Bandwidth Reservation System Design." Gary Hoo and William Johnston, Lawrence Berkeley National Laboratory, Ian Foster and Alain Roy, Argonne National Laboratory and University of Chicago. To appear, Eighth IEEE International Symposium on High Performance Distributed Computing, Aug. 3-6, 1999, Redondo Beach, California. (See http://www-itg.lbl.gov/Clipper/QoS)