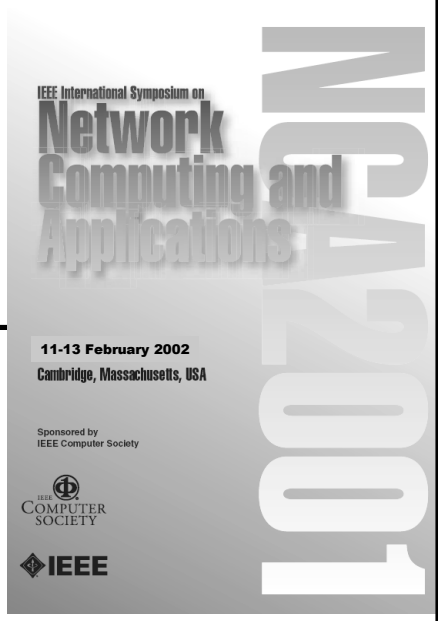
The Grid, NetSolve, and Its Applications

Jack Dongarra
Computer Science Department
University of Tennessee

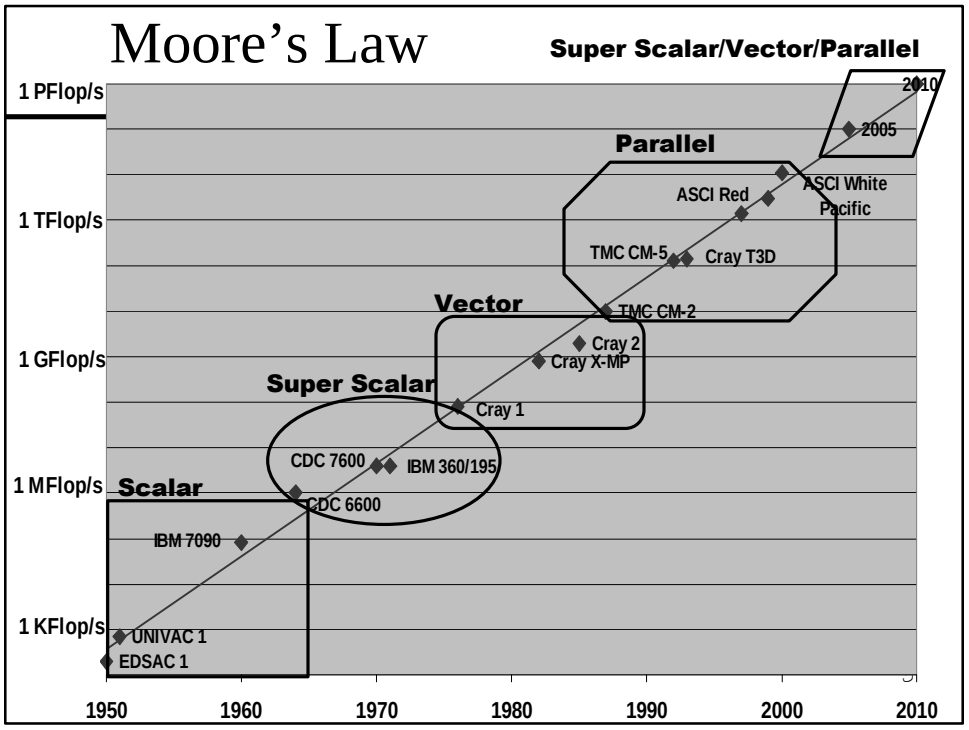


1

Outline

- **Overview of High Performance Computing**
- **The Grid**
- **NetSolve**

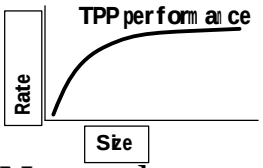
2



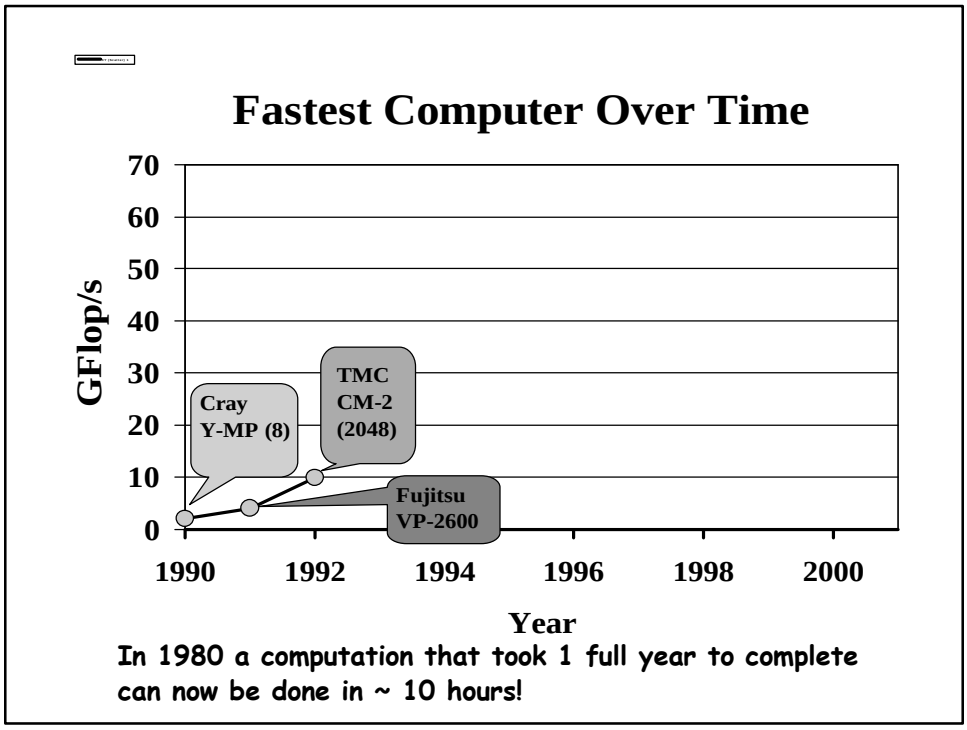
H. Mauer, H. Simon, E. Strohmaier, & JD

- Listing of the 500 most powerful Computers in the World
- Yardstick: Rmax from LINPACK MPP

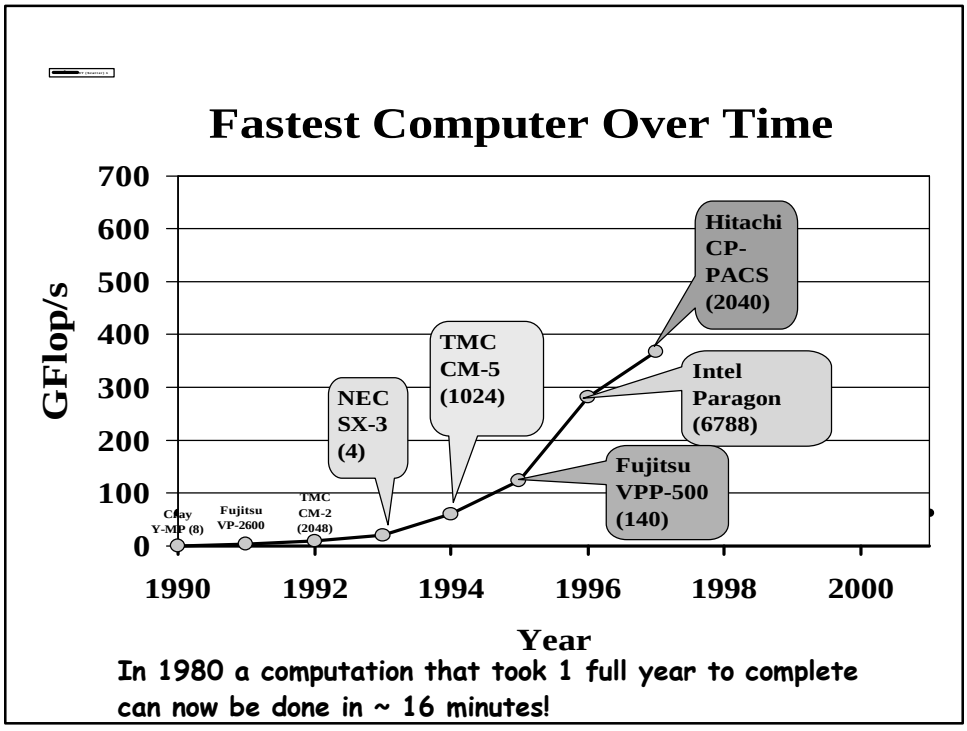
  $Ax=b$, *dense problem*



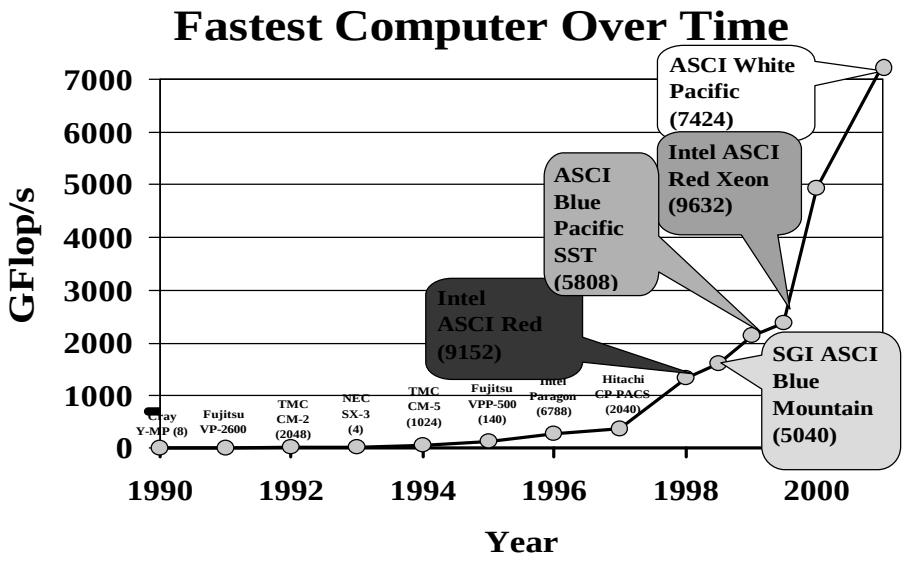
- Updated twice a year
  SC'xy in the States in November
  Meeting in Mannheim, Germany in June
- All data available from **www.top500.org**



In 1980 a computation that took 1 full year to complete can now be done in ~ 10 hours!

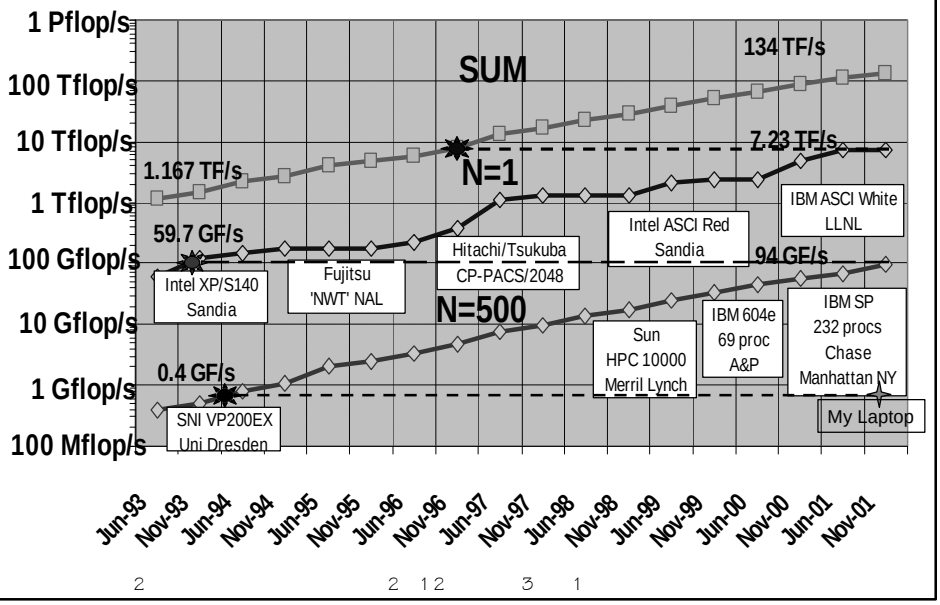


In 1980 a computation that took 1 full year to complete can now be done in ~ 16 minutes!

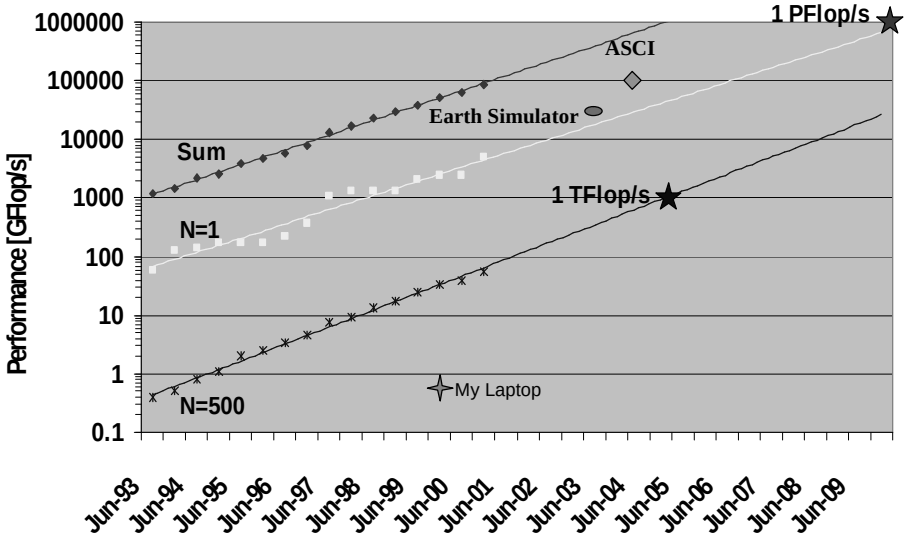


In 1980 a computation that took 1 full year to complete can today be done in ~ 27 seconds!

Performance Development

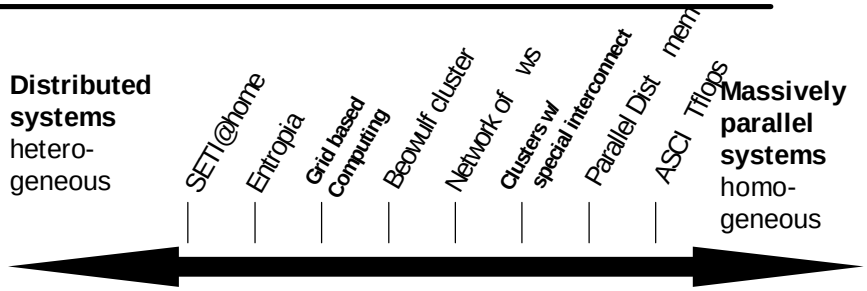


Performance Development



1 2 1 2 1

Distributed and Parallel Systems

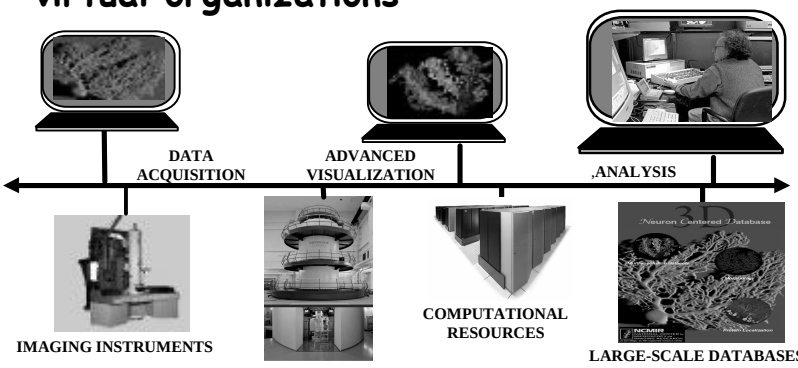


- Gather (unused) resources
- Steal cycles
- System SW manages resources
- System SW adds value
- 10% - 20% overhead is OK
- Resources drive applications
- Time to completion is not critical
- Time-shared
- SETI@home
  - ~ 400,000 machines
  - Averaging 27 Tflop/s

- Bounded set of resources
- Apps grow to consume all cycles
- Application manages resources
- System SW gets in the way
- 5% overhead is maximum
- Apps drive purchase of equipment
- Real-time constraints
- Space-shared
- ASCI White LLNL
  - 8000 processors
  - Averaging 7.2 Tflop/s

What is Grid Computing?

Resource sharing & coordinated problem solving in dynamic, multi-institutional virtual organizations



The Computational Grid is…

- **…a distributed control infrastructure that allows applications to treat compute cycles as commodities.**
- **Power Grid analogy**
  - **Power producers: machines, software, networks, storage systems**
  - **Power consumers: user applications**
- **Applications draw power from the Grid the way appliances draw electricity from the power utility.**
  - **Seamless**
  - **High-performance**
  - **Ubiquitous**
  - **Dependable**

Computational Grids and Electric Power Grids

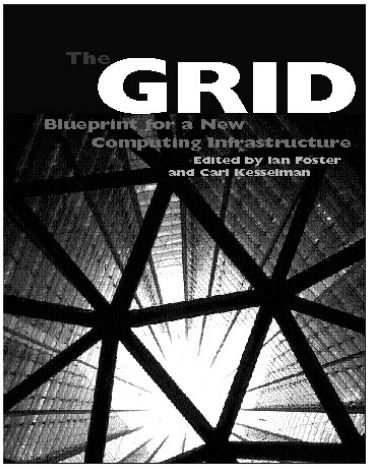
- **Why the Computational Grid is like the Electric Power Grid**
  - Electric power is ubiquitous
  - Don't need to know the source of the power (transformer, generator) or the power company that serves it

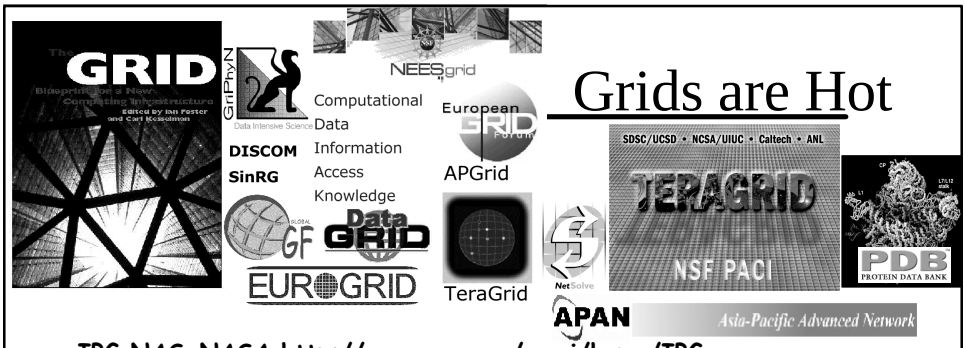
- **Why the Computational Grid is different from the Electric Power Grid**
  - Wider spectrum of performance
  - Wider spectrum of services
  - Access governed by more complicated issues
    - » Security
    - » Performance
    - » Socio-political factors

An Emerging Grid Community

1995-2000

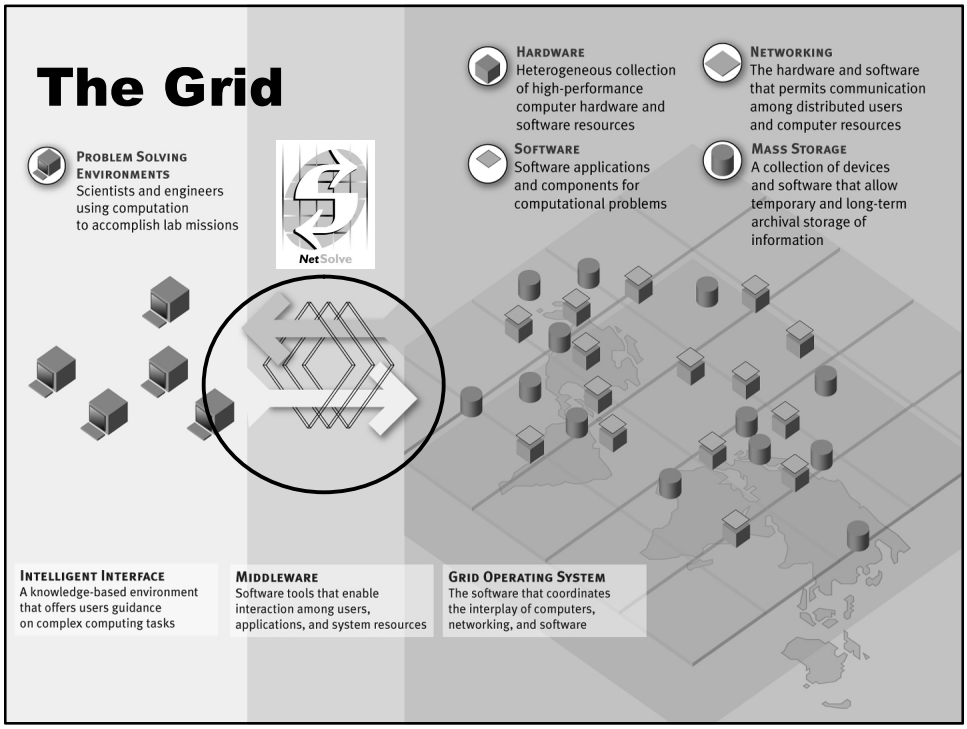
- **"Grid book" gave a comprehensive view of the state of the art**
- **Important infrastructure and middleware efforts initiated**
  - » Globus
  - » Legion
  - » Condor
  - » NetSolve, Ninf
  - » Storage Resource Broker
  - » Network Weather Service
  - » AppLeS, …



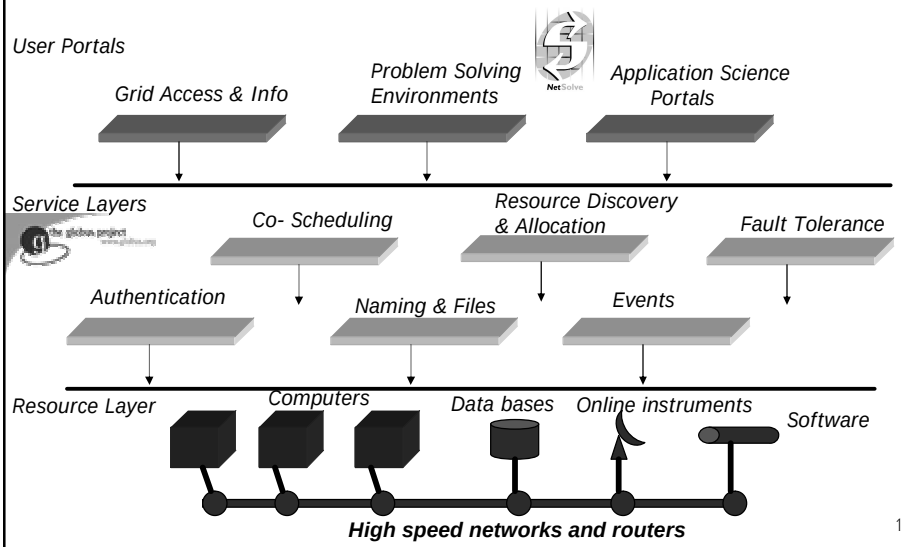


Grids are Hot

IPG NAS-NASA http://nas.nasa.gov/~wej/home/IPG
Globus http://www.globus.org/
Legion http://www.cs.virgina.edu/~grimshaw/
AppLeS http://www-cse.ucsd.edu/groups/hpcl/apples
NetSolve http://www.cs.utk.edu/netsolve/
NINF http://phase.etl.go.jp/ninf/
Condor http://www.cs.wisc.edu/condor/
CUMULVS http://www.epm.ornl.gov/cs/cumulvs.html
WebFlow http://www.npac.syr.edu/users/gcf/
LoCI http://loci.cs.utk.edu/



The Grid Architecture Picture

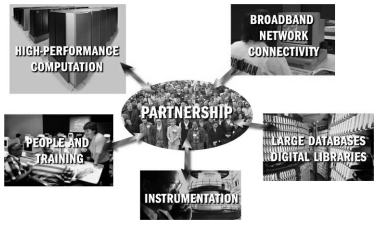


Globus Grid Services

- The Globus toolkit provides a range of basic Grid services
  - Security, information, fault detection, communication, resource management, ...
- These services are simple and orthogonal
  - Can be used independently, mix and match
  - Programming model independent
- For each there are well-defined APIs
- Standards are used extensively
  - E.g., LDAP, GSS-API, X.509, ...
- You don't program in Globus, it's a set of tools like Unix

Broad Acceptance of Grids as a Critical Platform for Computing

- **Widespread interest from government in developing computational Grid platforms**



NSF's Cyberinfrastructure

NASA's Information Power Grid

DOE's Science Grid

1

Broad Acceptance of Grids as a Critical Platform for Computing

- **Widespread interest from industry in developing computational Grid platforms**
- **IBM, Sun, Entropia, Avaki, Platform, …**

On August 2, 2001, IBM announced a new corporate initiative to support and exploit Grid computing.

AP reported that IBM was investing $4 billion into building 50 computer server farms around the world.

AVAKI

2

Grids Form the Basis of a National Information Infrastructure

August 9, 2001: NSF Awarded $53,000,000 to SDSC/NPACI and NCSA/Alliance for TeraGrid

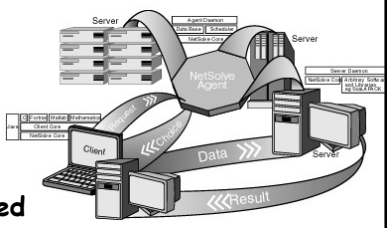
TeraGrid will provide in aggregate

- ***13.6 trillion calculations per second***
- ***Over 600 trillion bytes of immediately accessible data***
- ***40 gigabit per second network speed***
- ***Provide a new paradigm for data-oriented computing***
  - *Critical for disaster response, genomics, environmental modeling, etc.*

21

Motivation for NetSolve

Design an easy-to-use tool to provide efficient and uniform access to a variety of scientific packages on UNIX and Window's platforms



- Client-Server Design
- Non-hierarchical system
- Load Balancing and Fault Tolerance
- Heterogeneous Environment Supported
- Multiple and simple client interfaces
- Built on standard components

22

NetSolve Network Enabled Server

- NetSolve is an example of a Grid based hardware/software server.
- Based on a Remote Procedure Call model but with …
  - resource discovery, dynamic problem solving capabilities, load balancing, fault tolerance asynchronicity, security, …
- Easy-of-use paramount
- Other examples are NEOS from Argonne and NINF Japan.

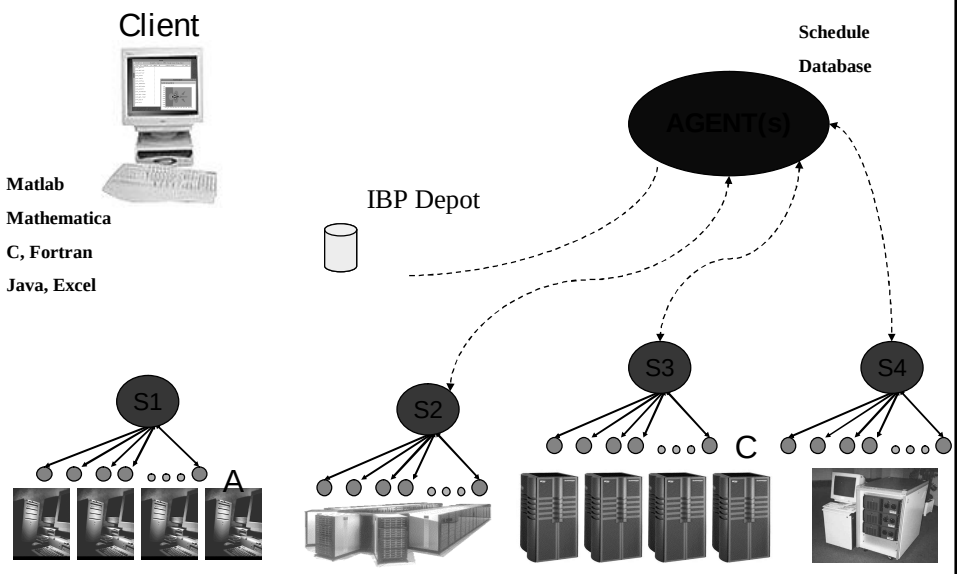
23

NetSolve

- Target not computer scientist, but domain scientist
- Hide logistical details
  - User shouldn't have to worry about how or where (issues about reproducibility)
- Present the set of available remote resources as a "multi-purpose" machine with a wealth of scientific software

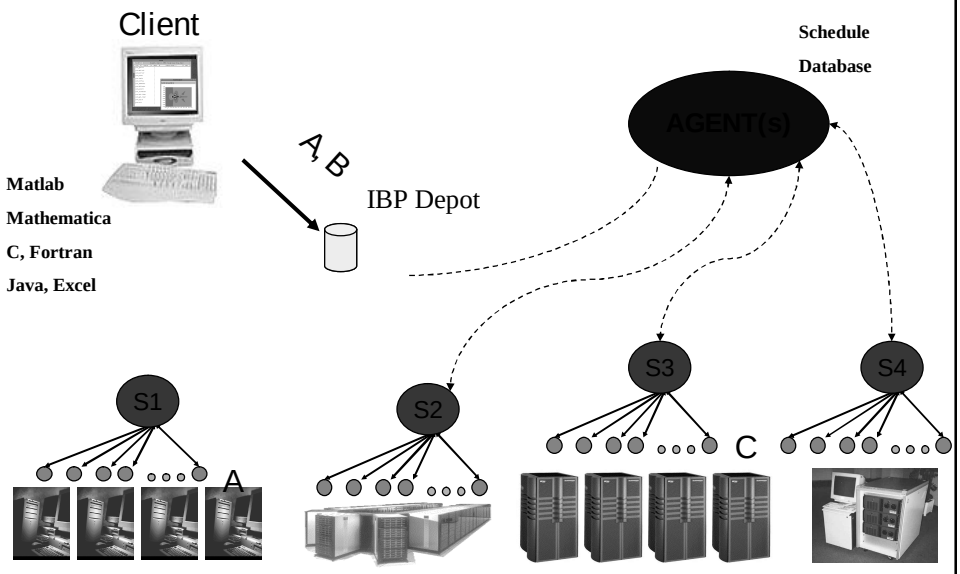
2

NetSolve: The Big Picture



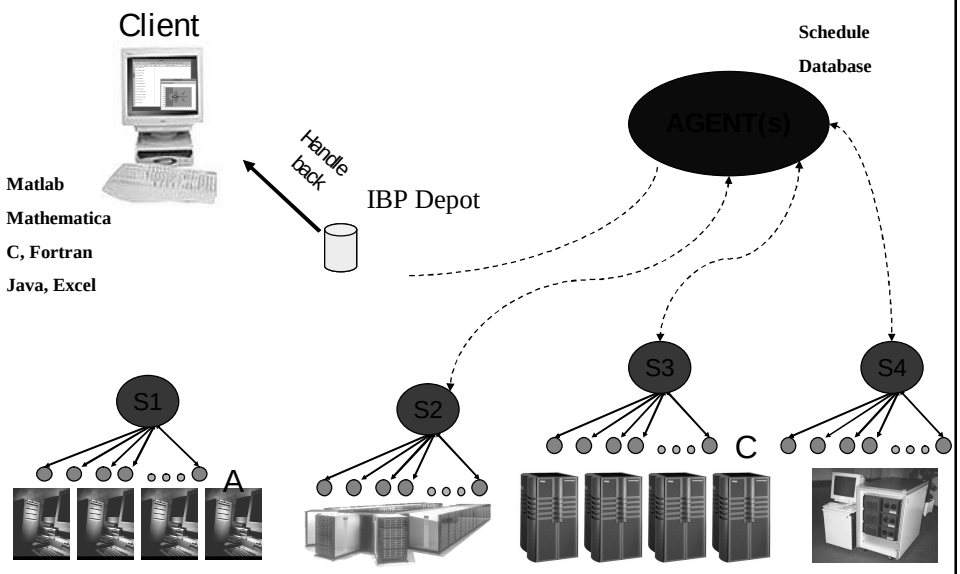
No knowledge of the grid required, RPC like.

NetSolve: The Big Picture



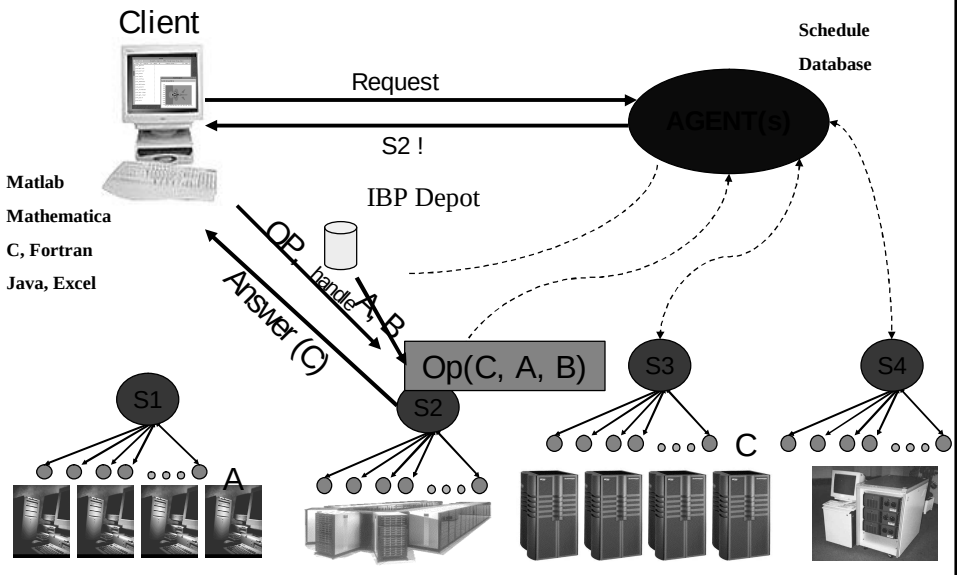
No knowledge of the grid required, RPC like.

NetSolve: The Big Picture



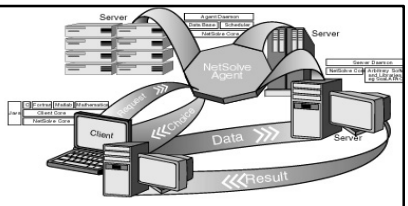
No knowledge of the grid required, RPC like.

NetSolve: The Big Picture



No knowledge of the grid required, RPC like.

Basic Usage Scenarios



- Grid based numerical library routines
  - User doesn't have to have software library on their machine, LAPACK, SuperLU, ScaLAPACK, PETSc, A TEC, ARPACK
- Task farming applications
  - "Pleasantly parallel" execution
  - eg Parameter studies
- Remote application execution
  - Complete applications with user specifying input parameters and receiving output
- "Blue Collar" Grid Based Computing
  - Does not require deep knowledge of network programming
  - Level of expressiveness right for many users
  - User can set things up, no "su" required
  - In use today, up to 200 servers in 9 countries
- Can plug into Globus, Condor, NINF, …

NetSolve Agent

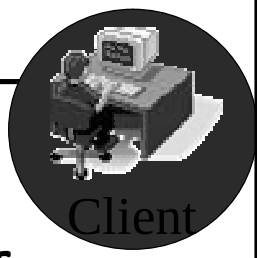
- Name server for the system.
- Information Service
  - client users and administrators can query the hardware and software services available.
- Resource scheduler
  - maintains both static and dynamic information regarding the NetSolve server components to use for the allocation of resources

NetSolve Agent

- Resource Scheduling (cont'd):
  - CPU Performance (LINPACK).
  - Network bandwidth, latency.
  - Server workload.
  - Problem size/algorithm complexity.
  - Calculates a "Time to Compute." for each appropriate server.
  - Notifies client of most appropriate server.

31

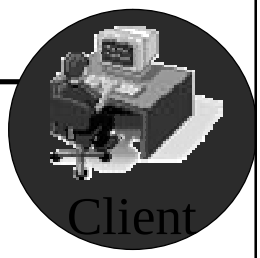
NetSolve Client



- Function Based Interface.
- Client program embeds call from NetSolve's API to access additional resources.
- Interface available to C, Fortran, Matlab, and Mathematica.
- Opaque networking interactions.
- NetSolve can be invoked using a variety of methods: blocking, non-blocking, task farms, …

32

NetSolve Client



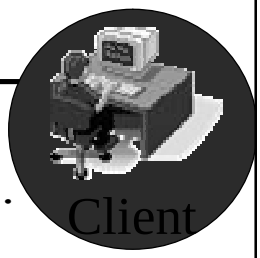
- **Intuitive and easy to use.**
- **Matlab Matrix multiply e.g.:**
  - **A matmul(B, C)**

A = netsolve('matmul', B, C);

- Possible parallelisms hidden.

33

NetSolve Client



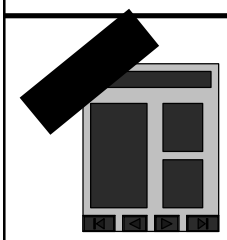
i. **Client makes request to agent.**

ii. **Agent returns list of servers.**

iii. **Client tries first one to solve problem.**

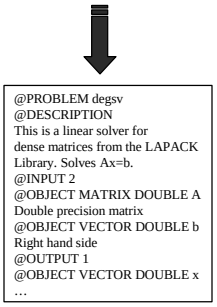
3

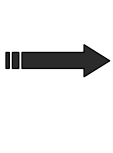
Generating New Services in NetSolve

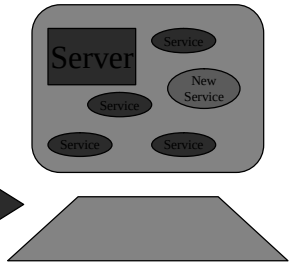


- **Add additional functionality**
  - **Describe the interface**
  - **Generate wrapper**
  - **Install into server**

New Service Added!







3

Task Farming - Multiple Requests To Single Problem

- **A Solution:**
  - any calls to netslnb( ) non-blocking
- **Farming Solution:**
  - Single call to netsl farm( )
- **Request iterates over an "array of input parameters."**
- **Adaptive scheduling algorithm.**
- **Useful for parameter sweeping, and independently parallel applications.**

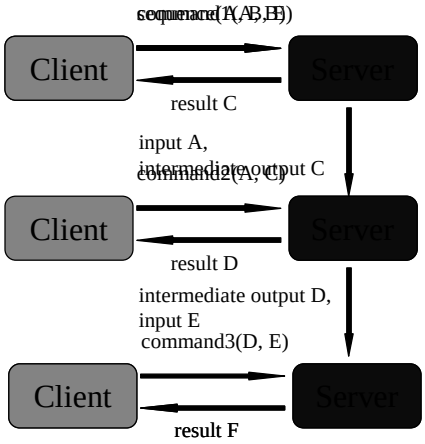
3

Data Persistence

- **Chain together a sequence of NetSolve requests.**
- **Analyze parameters to determine data dependencies. Essentially a DAG is created where nodes represent computational modules and arcs represent data flow.**
- **Transmit superset of all input/output parameters and make persistent near server(s) for duration of sequence execution.**
- **Schedule individual request modules for execution.**

3

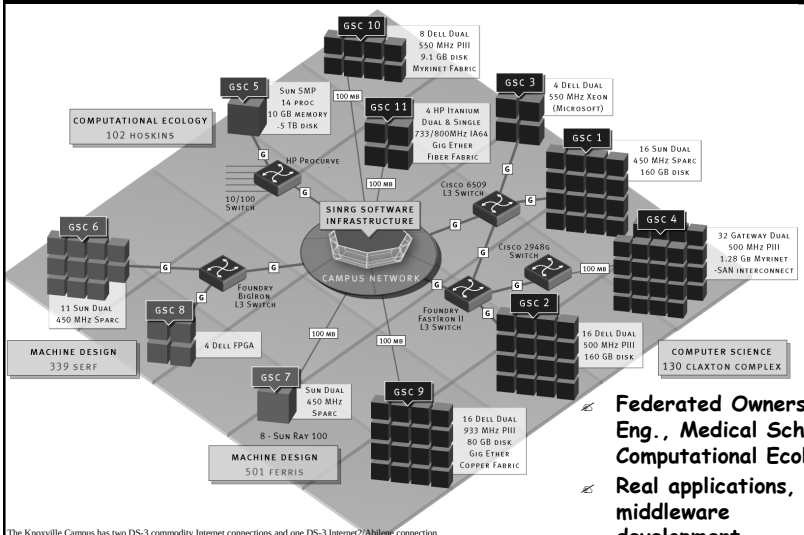
Data Persistence (cont'd)




```
netsl_begin_sequence( );
netsl("command1", A, B, C);
netsl("command2", A, C, D);
netsl("command3", D, E, F);
netsl_end_sequence(C, D);
```

3

University of Tennessee Deployment: Scalable Intracampus Research Grid SInRG



The Knoxville Campus has two DS-3 commodity Internet connections and one DS-3 Internet2/Abilene connection. An OC-3 ATM link routes IP traffic between the Knoxville campus, National Transportation Research Center, and Oak Ridge National Laboratory. UT participates in several national networking initiatives including Internet2 (I2), Abilene, the federal Next Generation Internet (NGI) initiative, Southern Universities Research Association (SURA) Regional Information Infrastructure (RII), and Southern Crossroads (SoX).

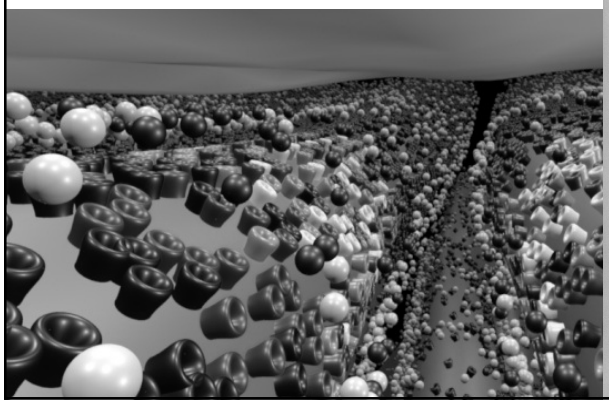
The UT campus consists of a meshed ATM OC-12 being migrated over to switched Gigabit by early 2002.

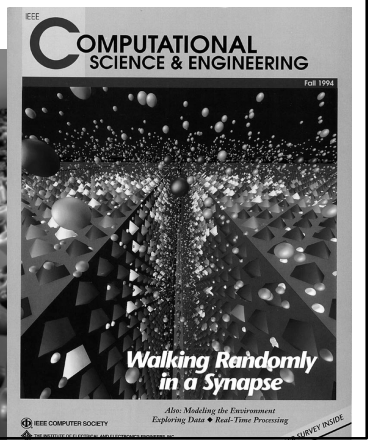
- Federated Ownership: CS, Chem Eng., Medical School, Computational Ecology, El. Eng.
- Real applications, middleware development, logistical networking

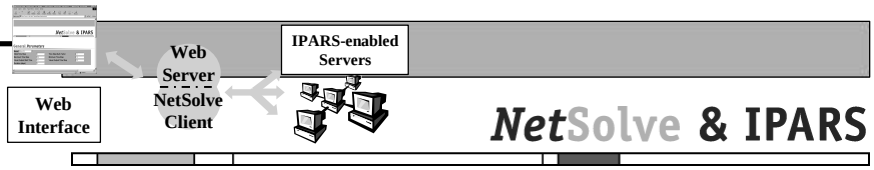
3

NPACI Alpha Project - MCell: 3-D Monte-Carlo Simulation of Neuro-Transmitter Release in Between Cells

- **UCSD (F. Berman, H. Casanova, M. Ellisman), Salk Institute (T. Bartol), CMU (J. Stiles), UTK (Dongarra, M. Miller, R. Wolski)**
- **Study how neurotransmitters diffuse and activate receptors in synapses**
- **blue unbounded, red singly bounded, green doubly bounded closed, yellow doubly bounded open**





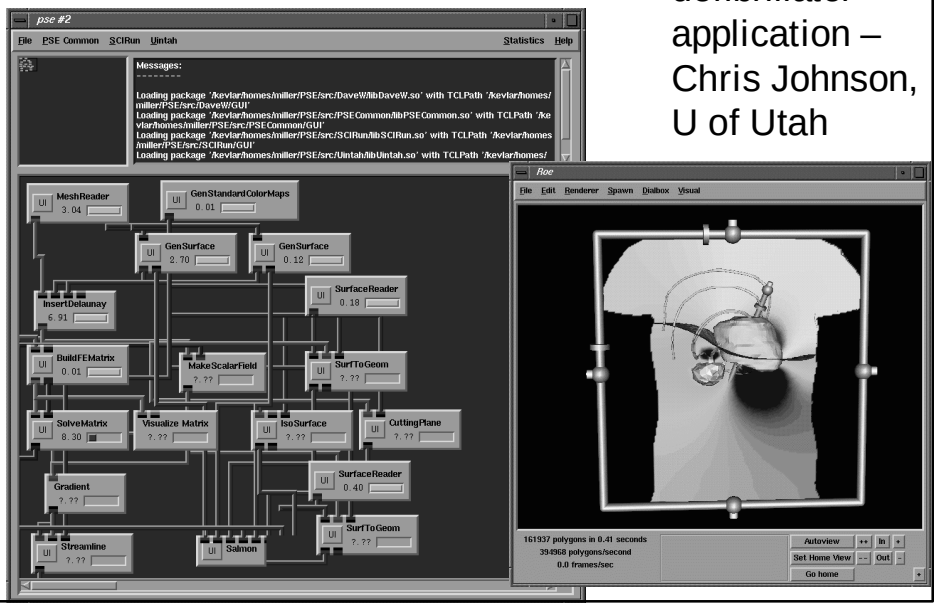


- Integrated Parallel Accurate Reservoir Simulator.
  - Mary Wheeler's group, UT-Austin
- Reservoir and Environmental Simulation.
  - models black oil, waterflood, compositions
  - D transient flow of multiple phase
- Integrates Existing Simulators.
- Framework simplified development
  - Provides solvers, handling for wells, table lookup.
  - Provides pre/postprocessor, visualization.
- Full IPARS access without Installation.
- IPARS Interfaces:
  - C, FORTRAN, Matlab, Mathematica, and Web.

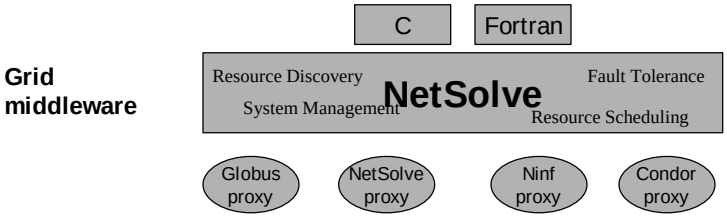
1

Netsolve and SCIRun

SCIRun torso defibrillator application – Chris Johnson, U of Utah

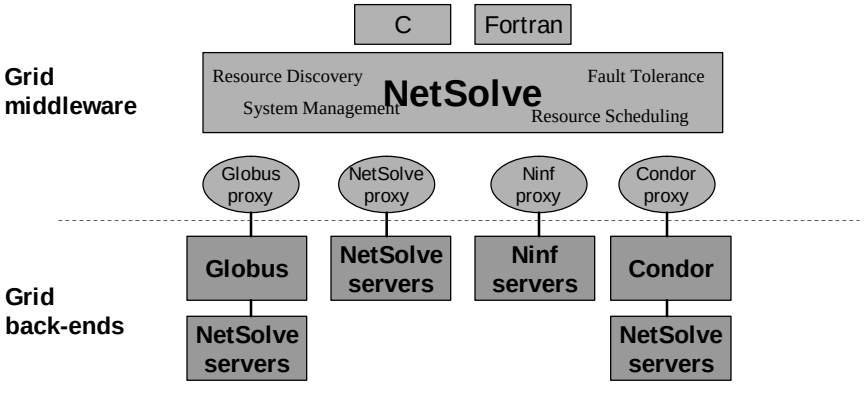


NetSolve: A Plug into the Grid

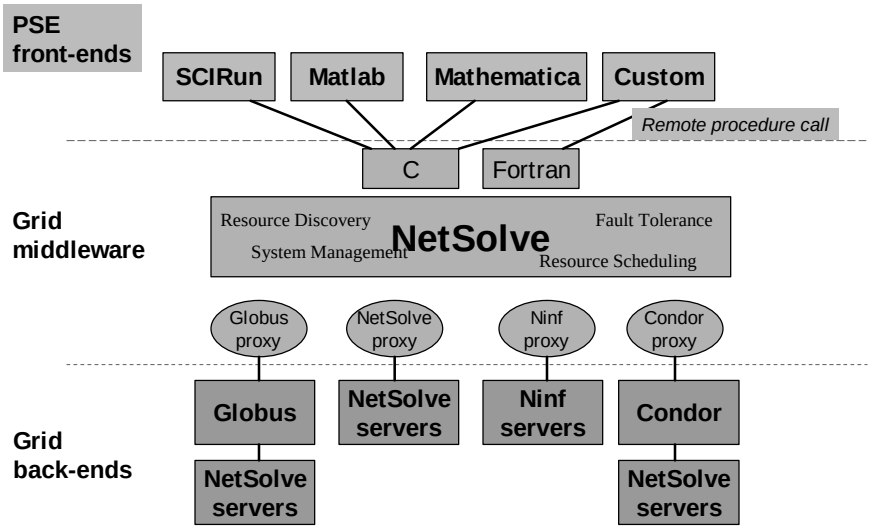


3

NetSolve: A Plug into the Grid

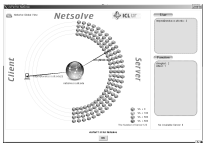


NetSolve: A Plug into the Grid



Things Not Touched On

- Security
  - Using Kerberos V5 for authentication.
- Separate Server Characteristics
  - Implementing Hardware and Software servers
- Hierarchy of Argents
  - More scalable configuration
- Monitor NetSolve Network
  - Track and monitor usage
- Network status
  - Network Weather Service



- Internet Backplane Protocol
  - middleware for managing and using remote storage.
- Fault Tolerance
- Local / Global Configurations
- Dynamic Nature of Servers
- Automated Adaptive Algorithm Selection
  - Dynamic determine the nest algorithm based on system status and nature of user problem

Conclusion

- Exciting time to be in scientific computing
- Network computing is here
- The Grid offers tremendous opportunities for collaboration
- Important to develop algorithms and software that will work effectively in this environment

Contributors to These Ideas

- Top500
  - Erich Strohmaier, NERSC
  - Horst Simon, NERSC
  - Hans Meuer, Mannheim U
- NetSolve
  - Henri Casanova, UCSD
  - Michelle Miller, UTK
  - Sathish Vadhiyar, UTK
- Fran Berman, UCSD/SDSC

For additional information see…

www.netlib.org/top500/

icl.cs.utk.edu/netsolve/

www.cs.utk.edu/~dongarra/

any opportunities within the group at ennessee