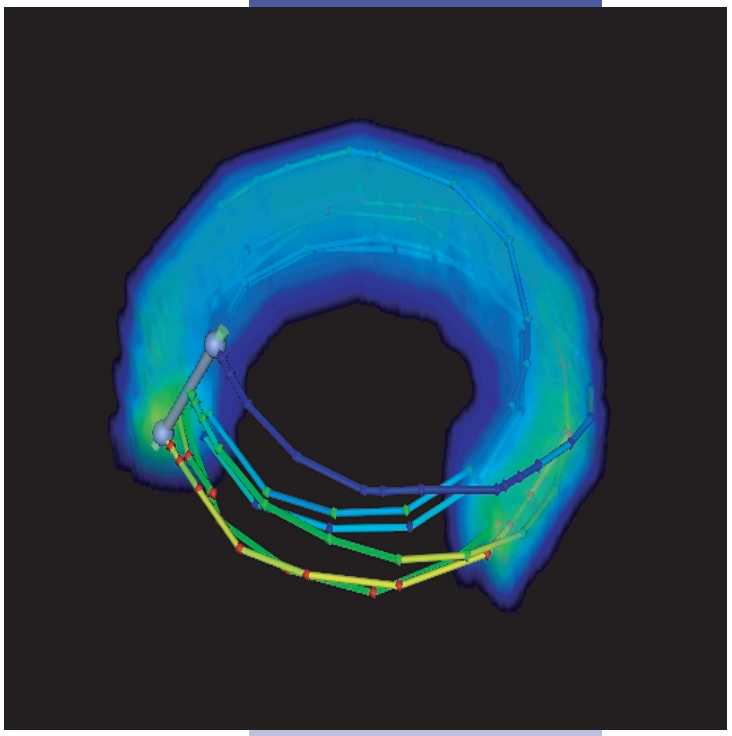
THE NATIONAL FUSION COLLABORATORY

D.P. Schissel, A. Finkelstein, I.T. Foster, T.W. Fredian, M.J. Greenwald, C.D. Hansen, K. Keahey, C.R. Johnson, S.A. Klasky, K. Li, D.C. McCune, M. Papka, Q. Peng, R. Stevens, M.R. Thompson



NIMROD simulated pressure stored in MDSplus and visualized with SciRUN

Presented at
US–Korea Bilateral Advanced Physics and Control Collaboration Meeting

April 11, 2002
San Diego, CA

A NEW THREE YEAR PROJECT HAS STARTED TO CREATE A NATIONAL FUSION COLLABORATORY

- The fusion Collaboratory represents a fundamental paradigm shift for the fusion community

- All data, analysis and simulation codes, and visualization tools will be thought of as network services

- The use of resources (data, codes, visualization tools) is separated from their implementation freeing the researcher from the need to know, in detail, how resources are implemented

FusionGRID
www.fusiongrid.org

THE GOAL OF THE FUSION COLLABORATORY IS TO ADVANCE SCIENTIFIC UNDERSTANDING & INNOVATION IN FUSION RESEARCH

- Enable more efficient use of existing experimental facilities through more powerful between pulse data analysis resulting in a greater number of experiments at less cost
- Allowing more transparent access to analysis and simulation codes, data, and visualization tools, resulting in more researchers having access to more resources
- Enable more effective integration of experiment, theory, and modeling
- Facilitate multi–institution collaborations
- Create a standard tool set for remote data access, security, and visualization allowing more researchers to build these services into their tools

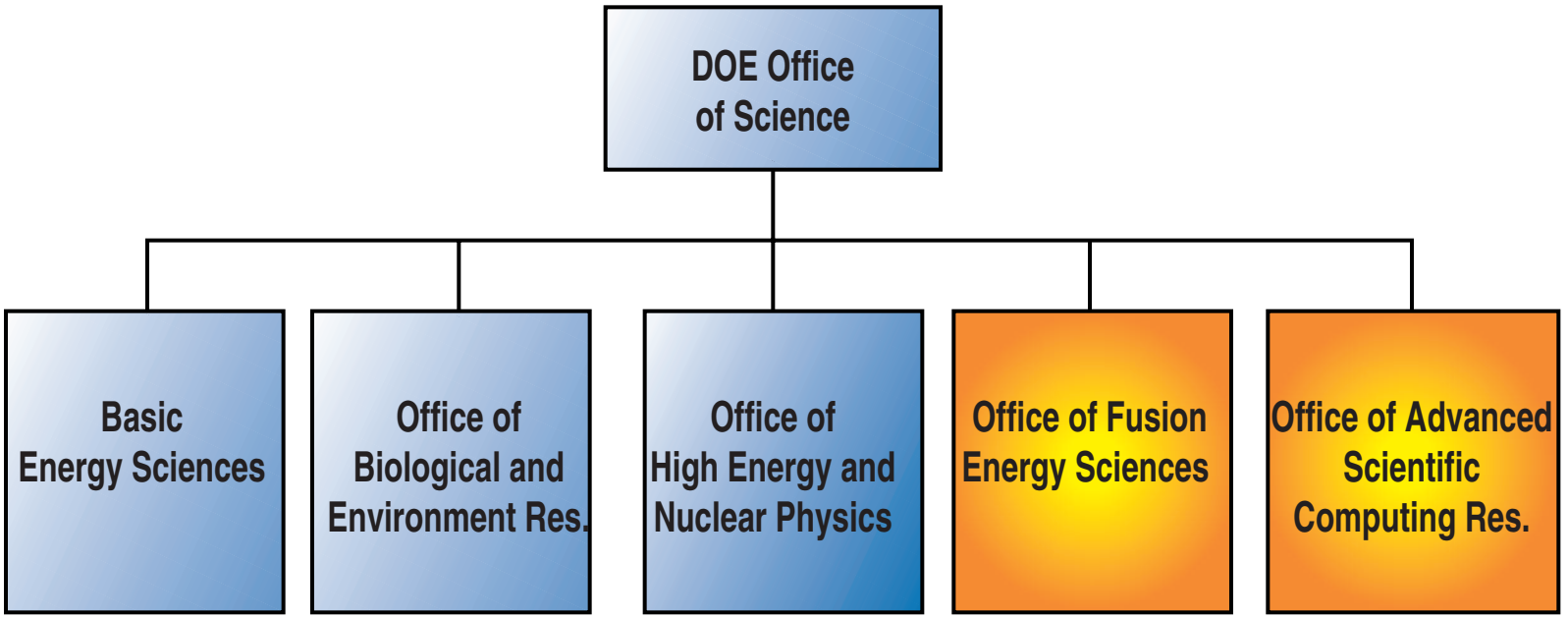
FusionGRID
www.fusiongrid.org

THE COLLABORATORY WILL CREATE & DEPLOY COLLABORATIVE SOFTWARE TOOLS FOR THE FUSION COMMUNITY

- Create transparent and secure access to local/remote computation, visualization, and data servers
- Develop collaborative visualization that allows interactive sharing of graphical images among control room display devices, meeting room displays, and with offices over a wide area network
- Enable real–time access to high–powered remote computational services allowing such capabilities as between pulse analysis of experimental data and advanced scientific simulations

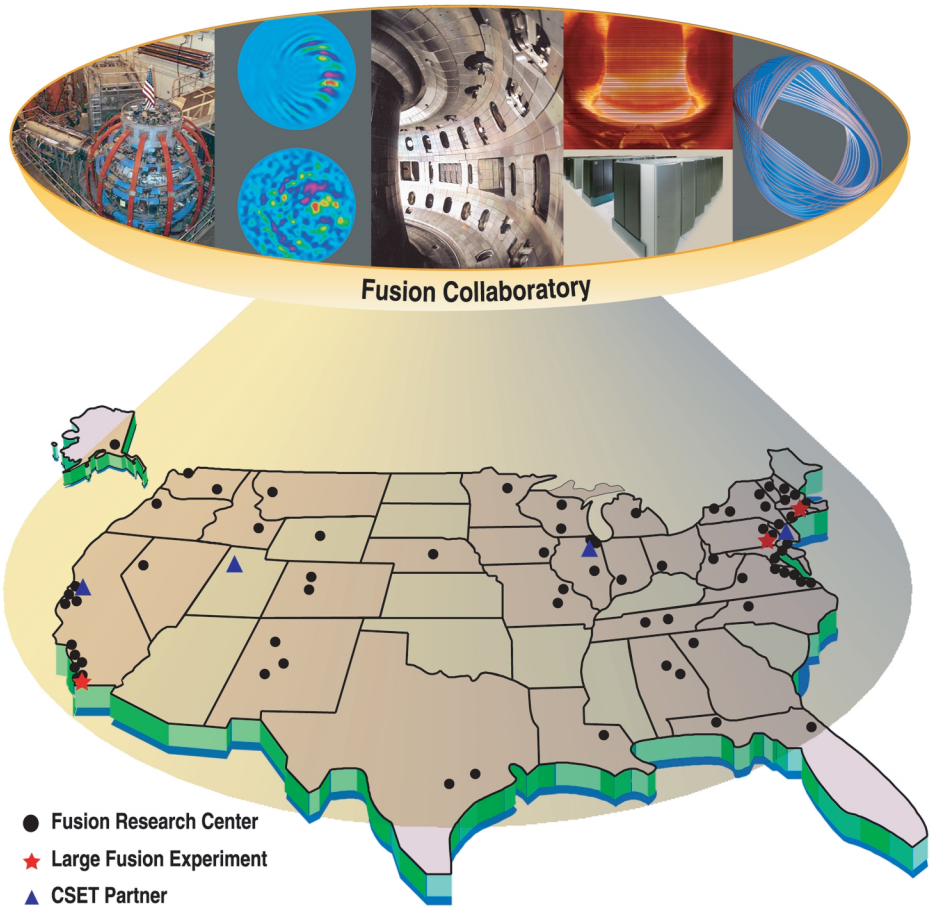
FusionGRID
www.fusiongrid.org

DEPARTMENT OF ENERGY/OFFICE OF SCIENCE: SCIENTIFIC DISCOVERY THROUGH ADVANCED COMPUTING (SciDAC)



FusionGRID
www.fusiongrid.org

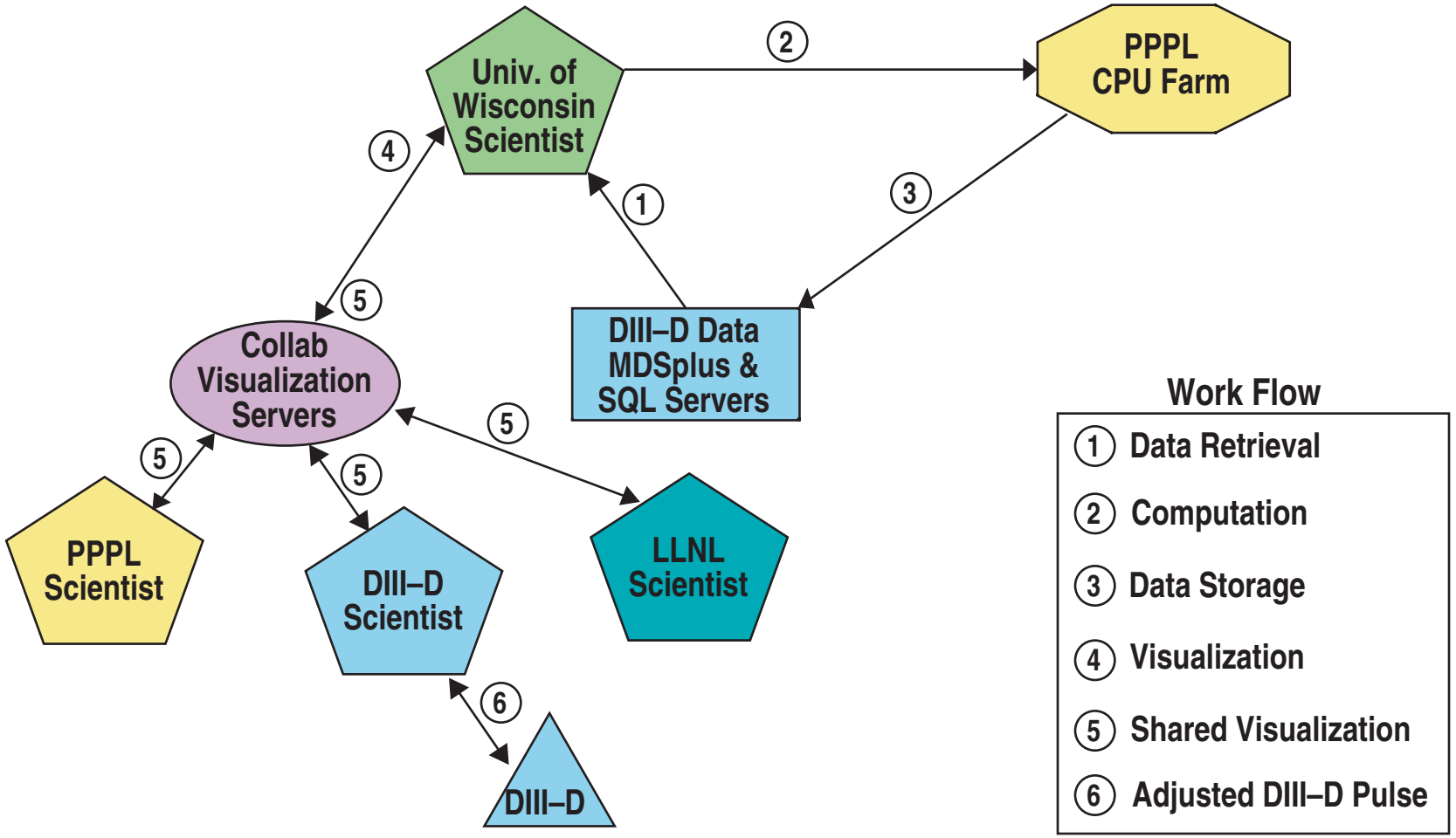
THE COLLABORATORY WILL EMBRACE 40 US SITES IN 37 STATES



- The Collaboratory will be created by a diverse team
  - 3 large fusion experiments
    - C–Mod, DIII–D, NSTX
  - 4 computer science centers
    - ANL, LBNL, Princeton U., U. of Utah
- Coordinated with the user community
  - Main experimental sites
  - Theory & simulation community
- A 3 year effort costing $5.4 million
  - For software, not hardware

FusionGRID
www.fusiongrid.org

EXAMPLE OF COLLABORATORY BENEFITS: ENHANCED EXPERIMENTAL OPERATIONS



FusionGRID
www.fusiongrid.org

MDSplus PROVIDES A COMMON, SHARED NETWORK ENABLED INTERFACE TO ALL DATA

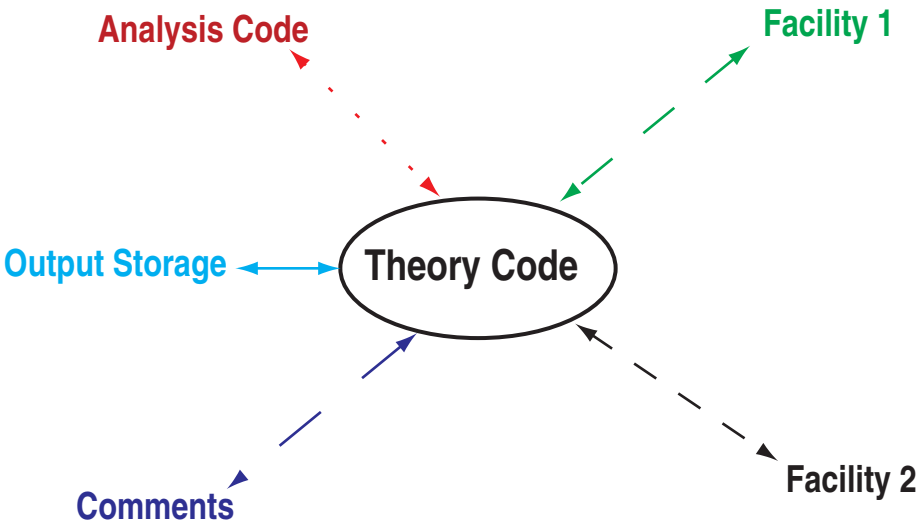
- MDSplus is a data system jointly developed by MIT, LANL, & Padova Italy
  - Provides for acquisition, storage, access, and organization of data
  - Client/server system utilizing TCP/IP
  - Can store experimental, simulation & theoretical data
- Presently used to serve data at 4 sites in US and 8 worldwide
  - Includes the 3 large US tokamaks
  - Clients at many sites
- Many physics analysis codes have been or are being integrated with MDSplus
  - Transport: MIST, ONETWO, TRANSP
  - MacroStability & Equilibrium: DCON, EFIT, M3D, NIMROD, PEST
  - MicroTurbulence: GS2

MDSplus CHARACTERISTICS

- Hierarchical and self descriptive
- Supports a variety of primitive data types
  - Byte, word, long, float, double, complex, string, expressions, actions
  - Built in expression evaluator (TDI)
- Remote access available currently from
  - Fortran, C/C++, Java, Python, IDL, Matlab, and labview
- Tools exist for quick display of data and structure
  - X–windows and Java scope & traverser
  - IDL tools (ReviewPlus, Pslice, JETDSP, etc.)
  - Matlab tools
- Supported platforms (so far)
  - AIX, Cray Unix, Digital Unix, HP/UX, Irix, Linux, Mac OS, Sun OS, VMS, Win32 (windows 9x, NT, 2000)

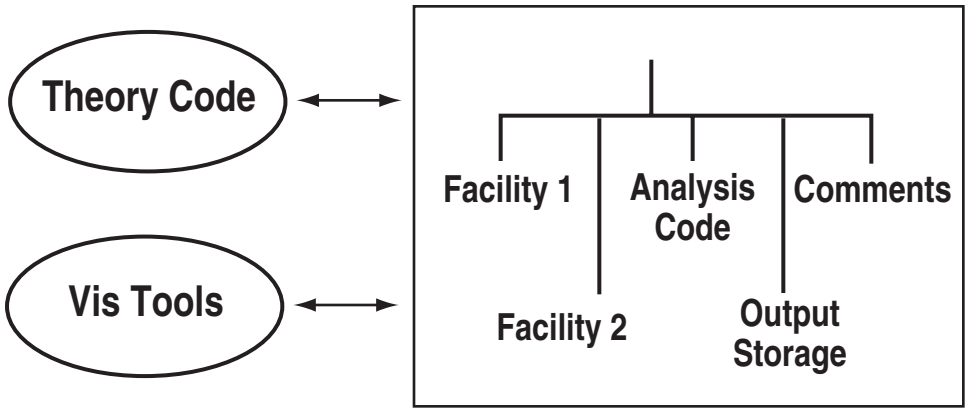
MDSplus WILL UNIFY DATA ACCESS FOR THE THEORETICAL COMMUNITY JUST AS IT HAS FOR THE EXPERIMENTAL COMMUNITY

Conventional Storage



- Each code needs it own interface
- Must know data format and file location
- Each code has its own graphics tool
- Hard to share results

MDSplus

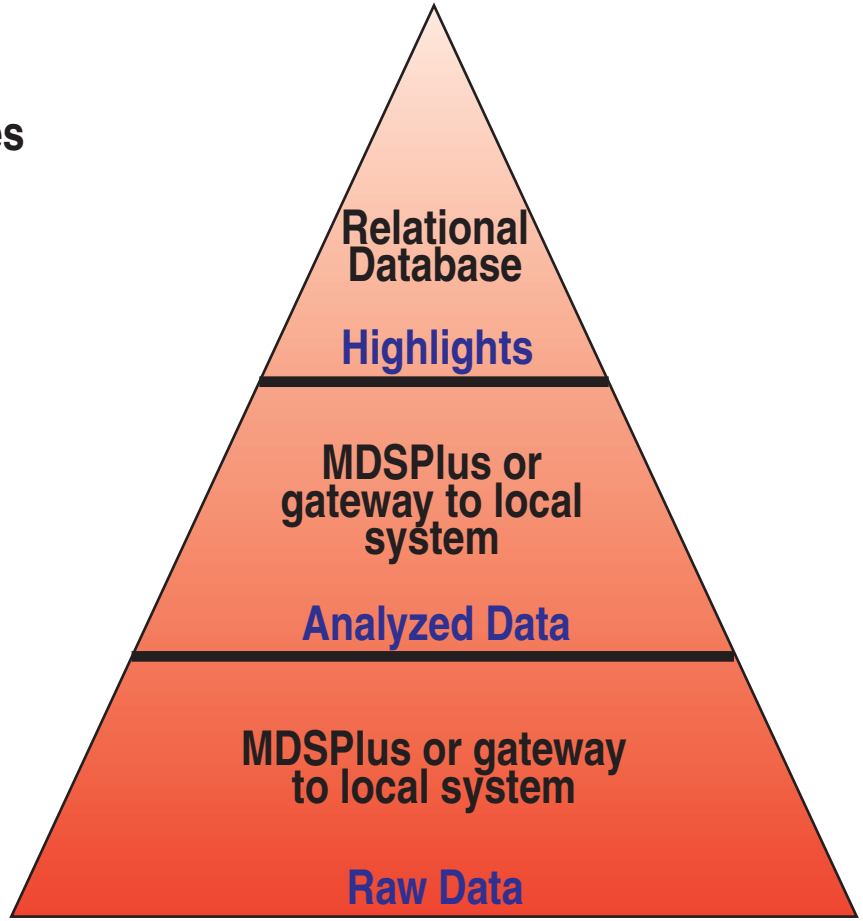


- One interface to many data types
- Only need location of data in tree
- Utilize existing visualization tools
- Easy to share results with both the theoretical and experimental community

FusionGRID
www.fusiongrid.org

A RELATIONAL DATABASE WORKS IN CONCERT WITH MDSplus

- MDSplus stores or is a gateway to all the data
  - Not optimized for queries across multiple pulses
- Relational DB stores highlights of the data
  - Optimized for queries
  - Drill down to smaller dataset
- Requirements of relational DB
  - Archival storage of data highlights
  - Track code runs
  - Interface to visualization software
  - Flexible schema evolution
  - Transparent access across WAN



FusionGRID
www.fusiongrid.org

THE COMPUTER SCIENCE RESEARCH NECESSARY TO CREATE THE COLLABORATORY IS CENTERED AROUND THREE AREAS

- **Security**
  - Valuable resources need to be protected: data, codes, & vis tools
  - Collaboratory will require authentication, authorization, and encryption
  - Fair use of shared resources
- **Remote and Distributed Computing**
  - Share the community's computational resources
  - Job scheduling, monitoring, exception handling, and accounting
- **Scientific Visualization**
  - Increased data quantities and ease of collaboration requires better visualization technology
  - Collaborative control rooms & meeting rooms, and enhanced vis tools

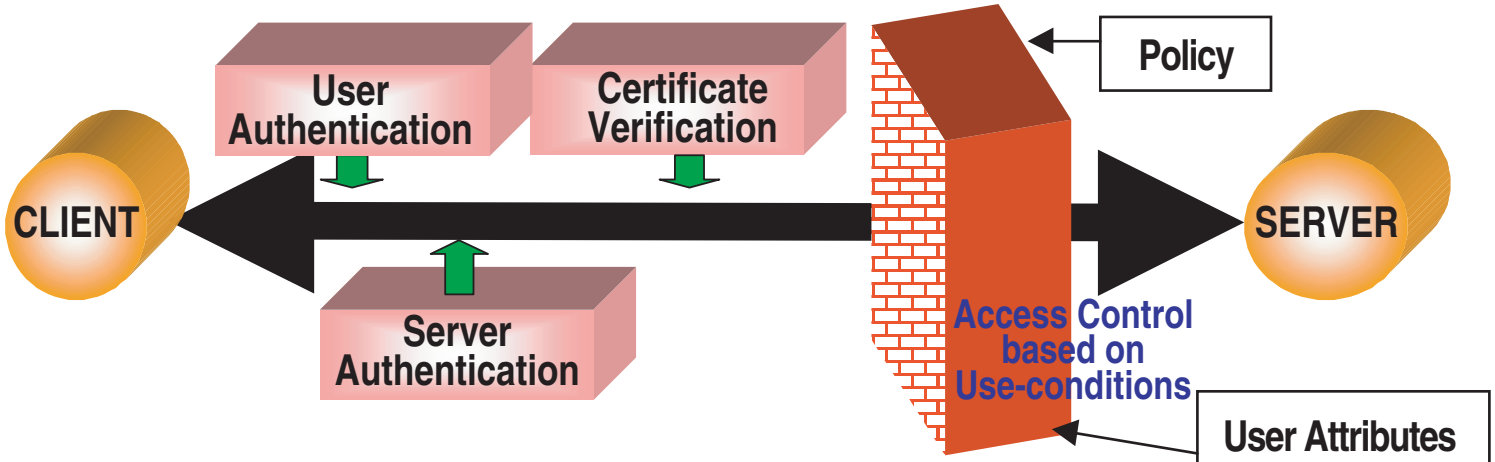
SECURITY: THE COLLABORATORY WILL UTILIZE THE GLOBUS SECURITY INFRASTRUCTURE & AKENTI AUTHORIZATION SERVICE

- Globus (www.globus.org) is a research and development project focused on enabling the application of Grid concepts to scientific computing
- The Grid refers to an infrastructure that enables the integrated, collaborative use of high–end computers, networks, databases, and scientific instruments owned and managed by multiple organizations
- Akenti (www-itg.lbl.gov/Akenti/) is an access control system designed to address the issues raised in allowing restricted access to distributed resources which are controlled by multiple organizations
- Existing fusion codes will be modified to use this infrastructure
- The middleware tools will be extended to meet Collaboratory needs

FusionGRID
www.fusiongrid.org

ACCESS CONTROL ENFORCES AN AUTHORIZATION POLICY

- Clients attempt to access resources controlled by servers
- Access control policy: prior authorization decisions for client access to resources
- Public-key infrastructure and secure message protocols provide confidentiality, message integrity, and user identity authentication



S. S. Mudumbai, et al.

REMOTE & DISTRIBUTED COMPUTING: ACCESS TO POWERFUL DATA ANALYSIS & SIMULATION CODES AS NETWORK SERVICES

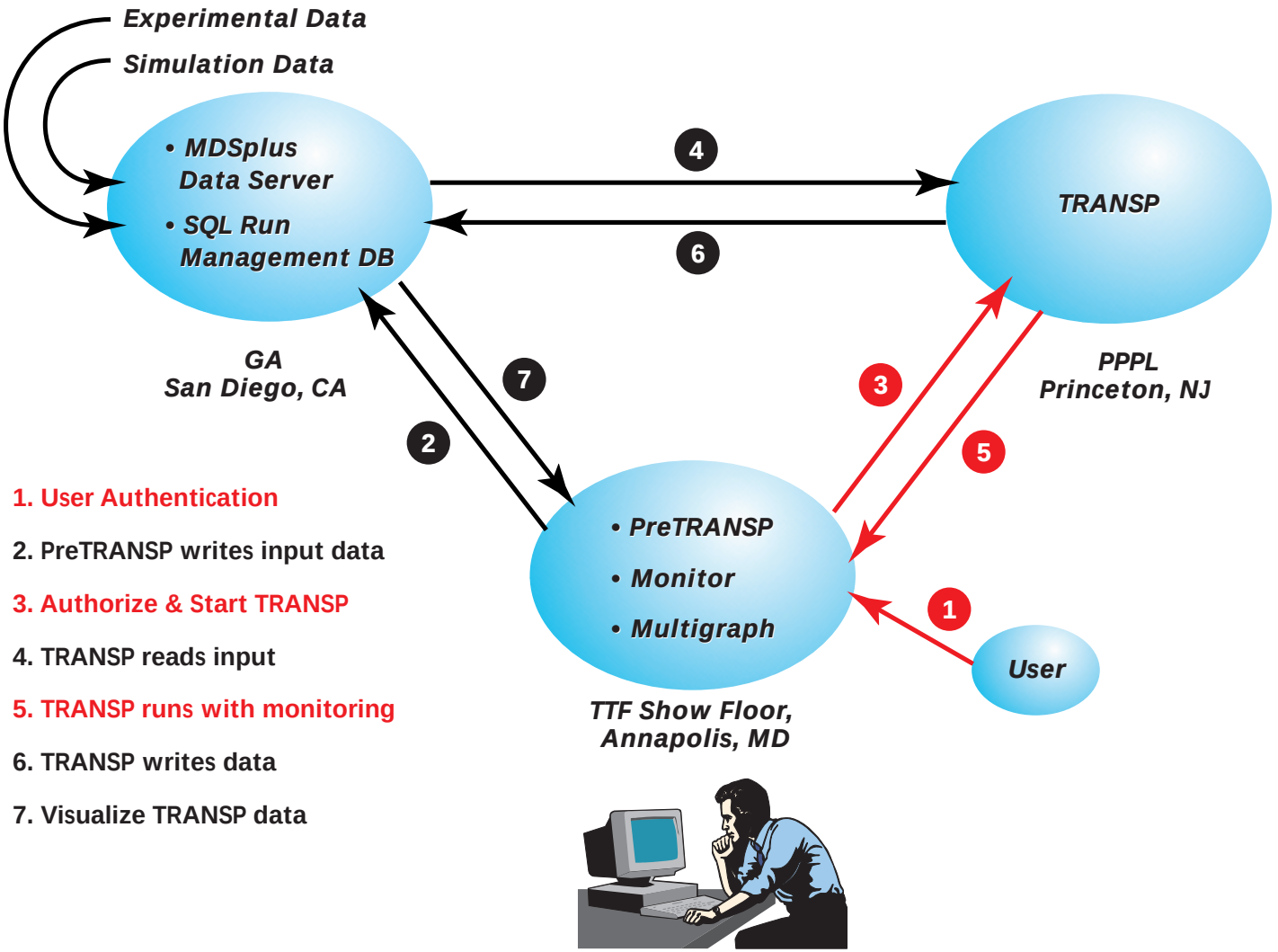
- Fusion codes will run on hardware appropriate for each code
  - Fast serial workstations – midrange parallel clusters – supercomputers
- Collaboratory will enable detailed time dependent transport and stability analysis between pulses
  - Equitable sharing and preemptive data analysis
- Globus can provide for the required capabilities
  - Create grids connecting computational resources with users
  - Track the capabilities of resources within a grid
  - Specify the resource needs of user's computing tasks
  - Mutually authenticate both users and resources

FusionGRID
www.fusiongrid.org

SUBSTANTIAL SETUP FOR APRIL FUSION SCIENCE MEETINGS BUT THE EXERCISE PROVED VALUABLE

FusionGRID
www.fusiongrid.org

SUCCESSFULL DEMONSTRATION OF GRID COMPUTING AT APRIL FUSION SCIENCE MEETINGS

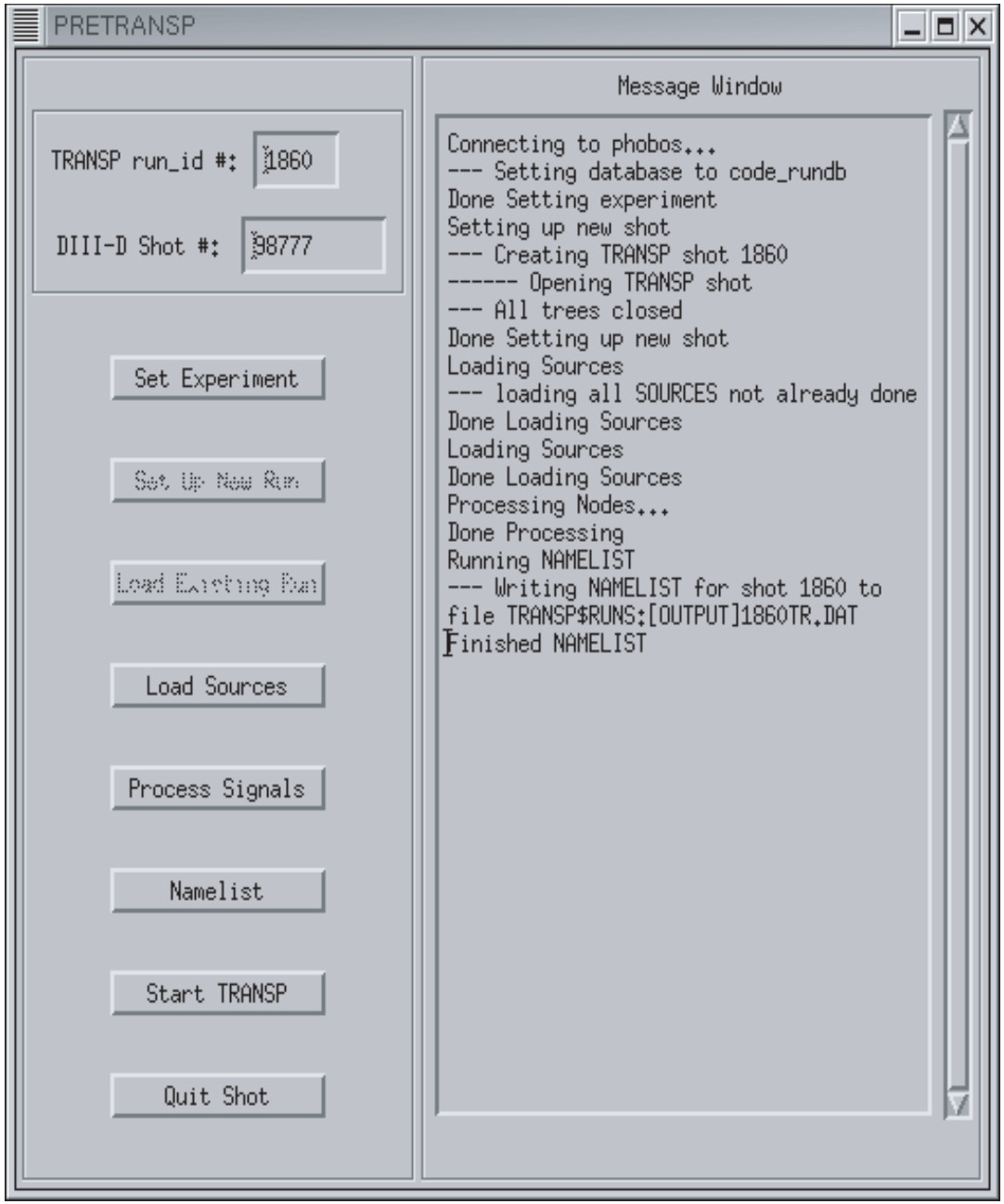


1. User Authentication
2. PreTRANSP writes input data
3. Authorize & Start TRANSP
4. TRANSP reads input
5. TRANSP runs with monitoring
6. TRANSP writes data
7. Visualize TRANSP data

FusionGRID
www.fusiongrid.org

A GUI WAS CREATED TO SETUP AND LAUNCH A TRANSP RUN

- First log onto the Grid
- Prepare data for TRANSP run and store inputs in MDSplus
- Submit TRANSP run
- Monitor state of run
- TRANSP writes data to MDSplus
- Visualize TRANSP data



FusionGRID
www.fusiongrid.org

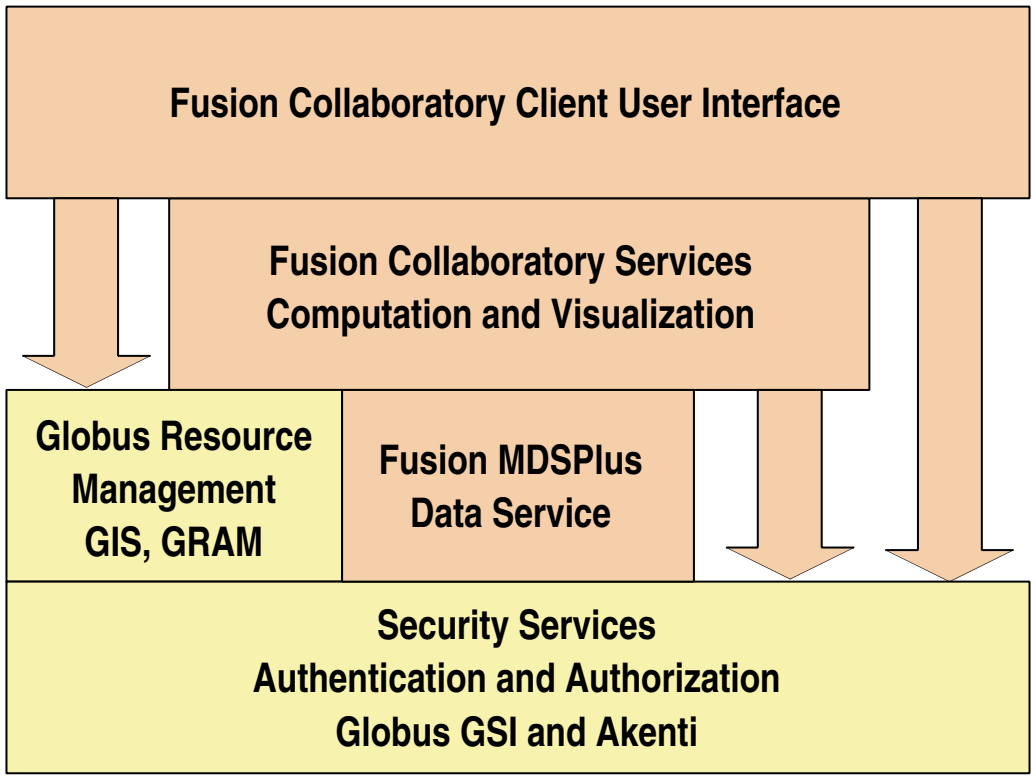
A VARIETY OF VISUALIZATION TOOLS BROUGHT THE DEMO ALIVE

- US map to visualize demo & make it real
- Monitor run via a web browser
- Visualize results using IDL based tool

FusionGRID
www.fusiongrid.org

SERVICES ARE LAYERED TO HIDE COMPLEXITY WHILE PROVIDING POWERFUL TOOLKITS FOR DEVELOPERS

- MDSplus will provide the common data access for the Collaboratory
- Globus and Akenti will provide the necessary middleware



FusionGRID
www.fusiongrid.org

VISUALIZATION: COLLABORATIVE NATURE OF FUSION RESEARCH NECESSITATES A SHARED VISUALIZATION ENVIRONMENT

- Strive to dramatically reduce the hurdles that presently exist for collaborative visualization
- Leverage existing technology where possible
  - Workspace docking using the Access Grid (AG)
  - Integrate existing AG collaborative tools with tiled display walls
- Collaborative Control Room
  - Large on–site group to interactively work with small to large off–site group
- New visualization software
  - Simultaneous sharing of complex visualizations
  - Error representation in complex experimental and simulation data

FusionGRID
www.fusiongrid.org

TILED DISPLAYS WALLS ALLOW A LARGE GROUP OF SCIENTISTS TO EXPLORE INFORMATION IN COLLABORATION MORE EFFECTIVELY

- Access Grid (www.accessgrid.org) compliments and extends the data grid
  - Ensemble of network, computing and interaction resources that supports group to group collaboration and communication
- Display wall research has focused on low–cost commodity components

FusionGRID
www.fusiongrid.org

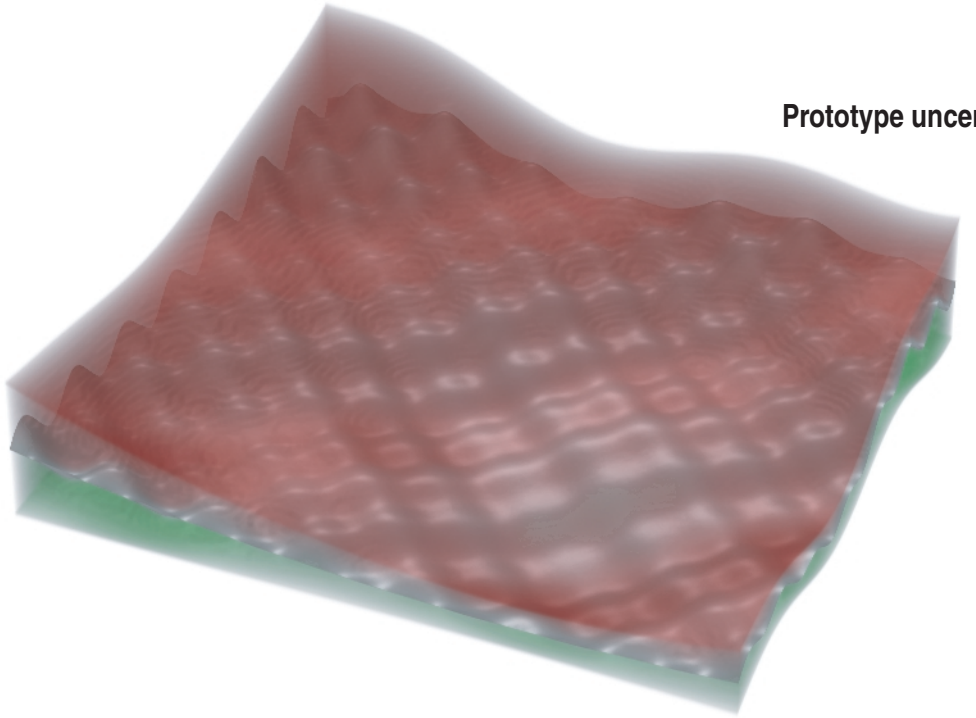
SMALL DESK TOP ACCESS GRID NODE HAS BEEN DEMONSTRATED

ANL

- Targed for the small research center
  - For one to one and one to many interactions
- Usage example: communication to a tokamak control room

FusionGRID
www.fusiongrid.org

VISUALIZATION TOOLS WILL BE CREATED THAT REPRESENT A SIGNIFICANT INCREASE IN EFFICIENCY FOR THE FUSION COMMUNITY

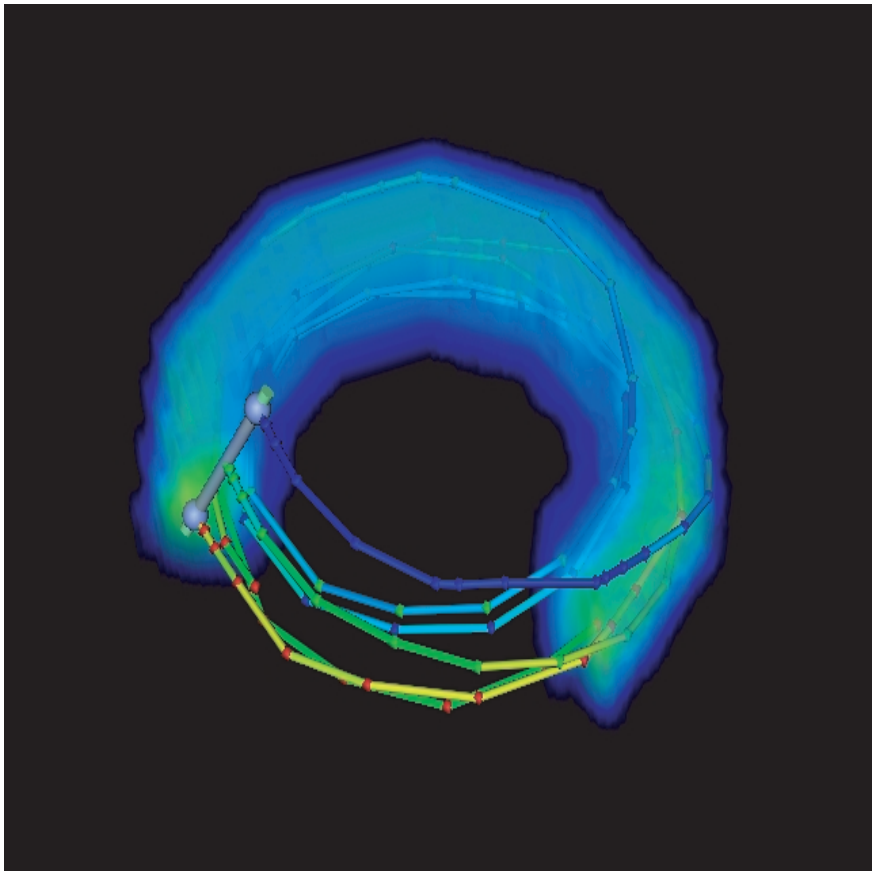


Prototype uncertainty visualization
courtesy U. of Utah

- Provide visualizations that incorporate and compare data from multiple experimental and simulation sources and to reflect uncertainty information to aid in data analysis
- Collaborative visualization tool will be created for experimental operations giving both local and remote teams an interactive, big picture view of the vast amount of data

FusionGRID
www.fusiongrid.org

ADVANCED VISUALIZATION TOOLS BEING CREATED



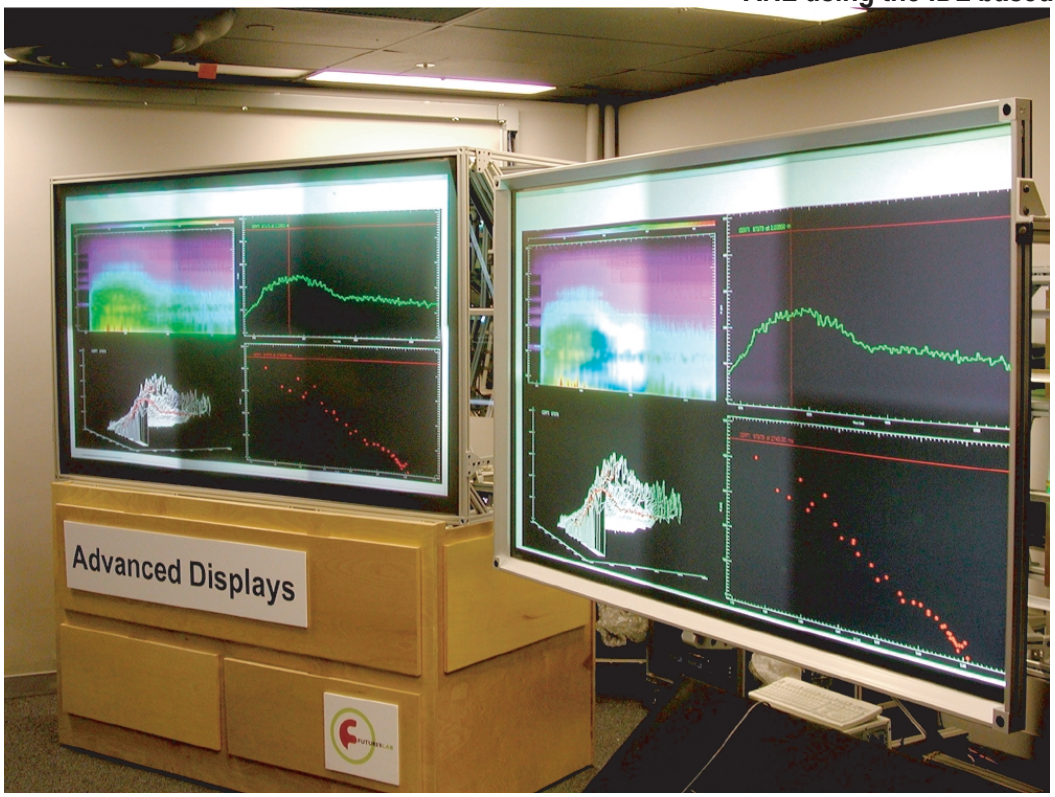
U. of Utah - NIMROD pressure with magnetic field stream lines

- Quantiative 3D visualization – SciRUN
- Designed for general use
- Can be remotely shared
- To be demonstrated at Sherwood

FusionGRID
www.fusiongrid.org

SHARED VISUALIZATION BETWEEN TILED WALLS HAS BEEN DEMONSTRATED

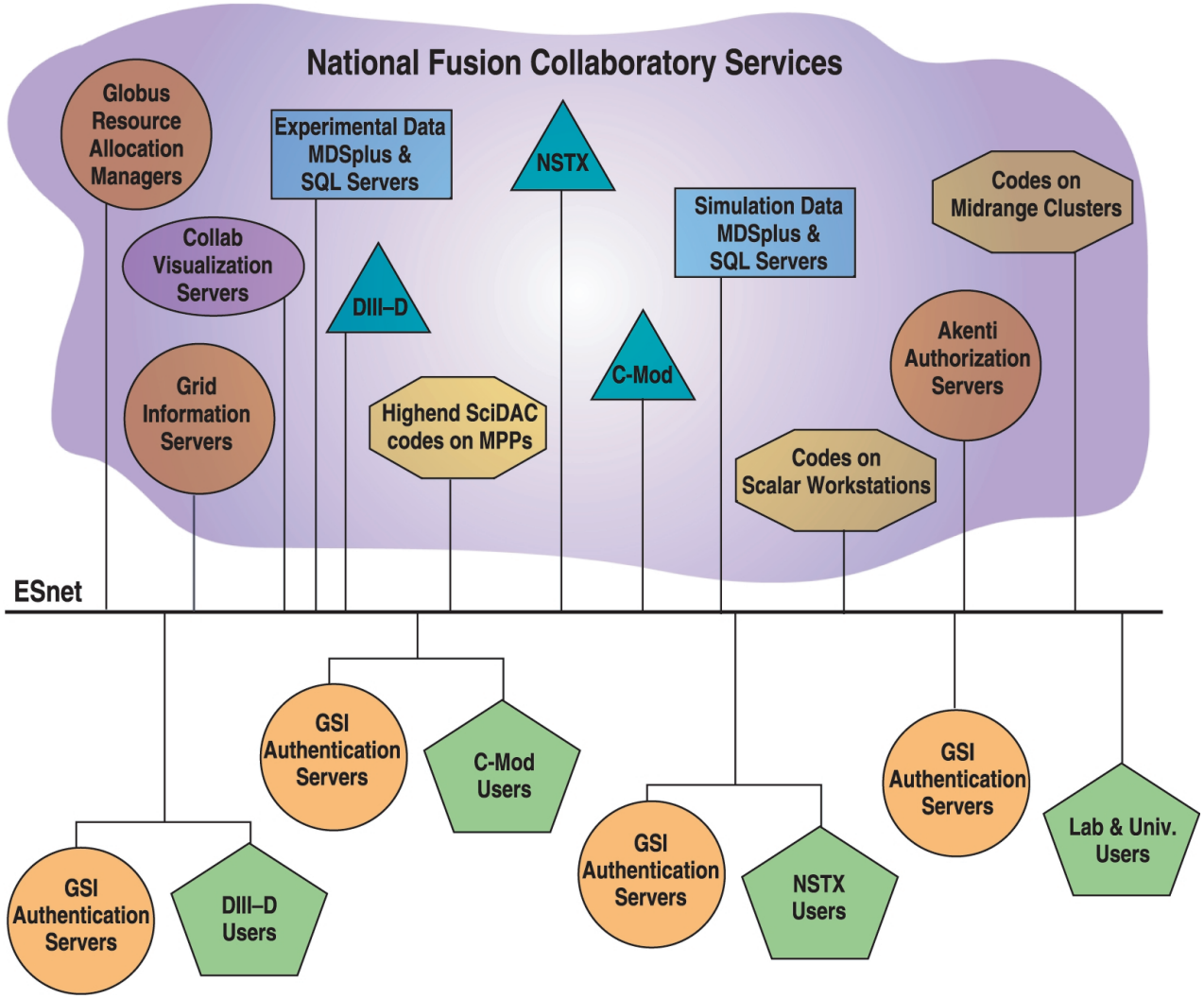
ANL using the IDL based tool ReviewPlus



- Workstation to Wall and Wall to Wall is possible - communicate to the control room
- To be demonstrated at APS/DPP 2002 in Orlando

FusionGRID
www.fusiongrid.org

CREATE A UNIFIED FRAMEWORK SO DATA, CODES, VIS TOOLS ARE AVAILABLE SECURELY & TRANSPARENTLY OVER THE INTERNET



FusionGRID
www.fusiongrid.org

SUMMARY

- A 3–year project to create a Fusion Collaboratory has begun
- Team effort comprised of fusion scientists and computer scientists
- The collaboratory will enable networked real–time data analysis and instantaneous communication amongst geographically dispersed teams of experimentalists and theoreticians
- The goals of the KSTAR and Collaboratory Projects are similar
  - Expect this can be a very fruitful area of collaboration
  - Would like to pursue areas of mutual interest for US–Korea bilaterial work
- More information at http://www.fusiongrid.org/

FusionGRID
www.fusiongrid.org