

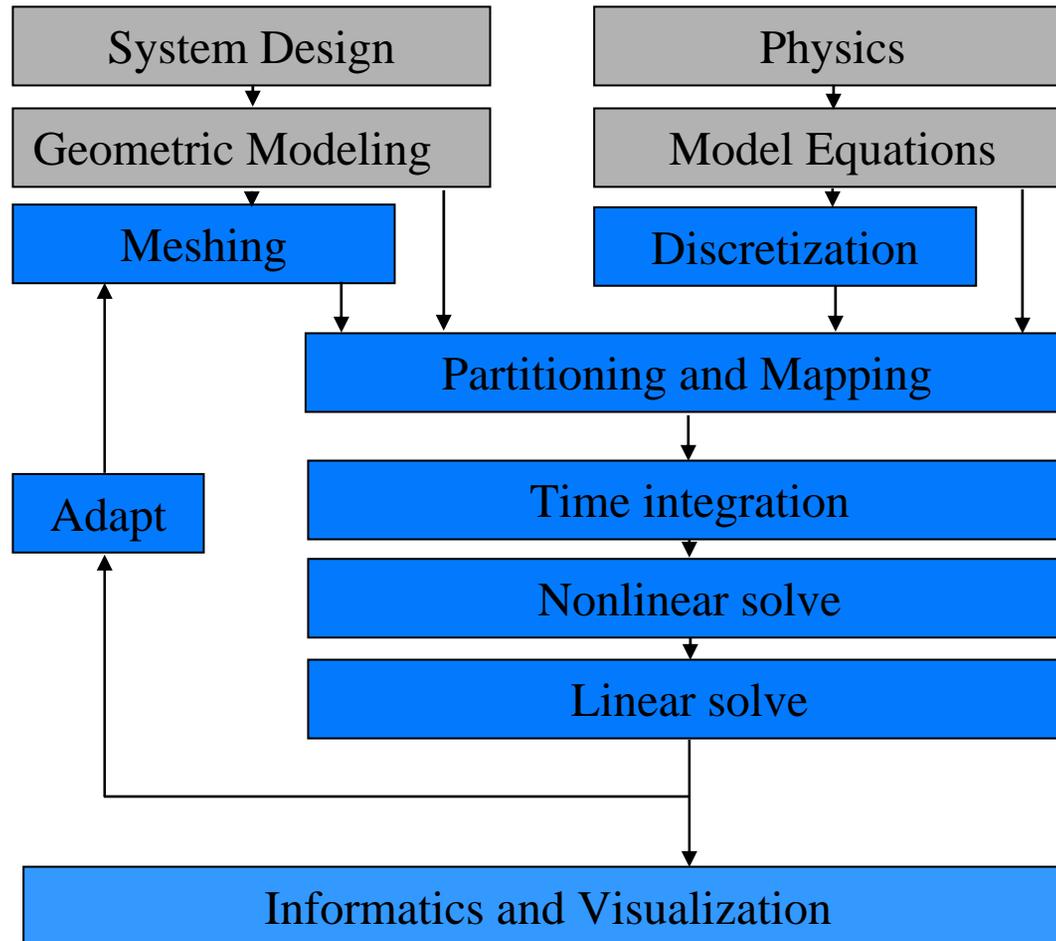
---

# Algorithms for Scientific Computing and Informatics

David E. Womble  
Sandia National Laboratories

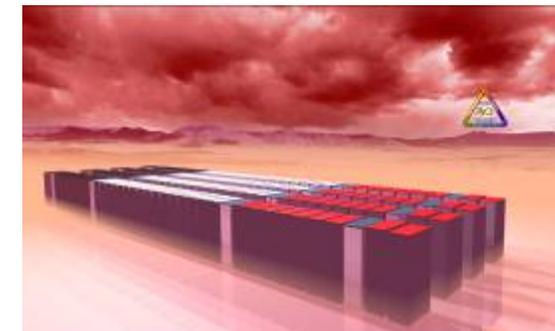
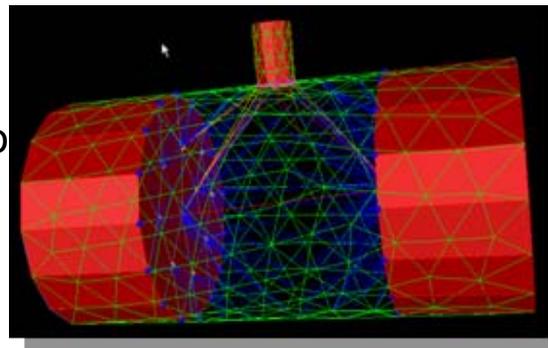
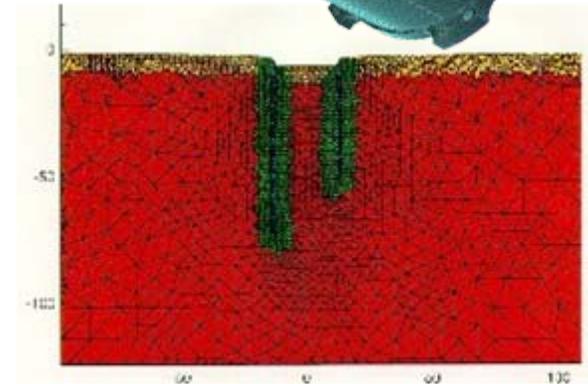
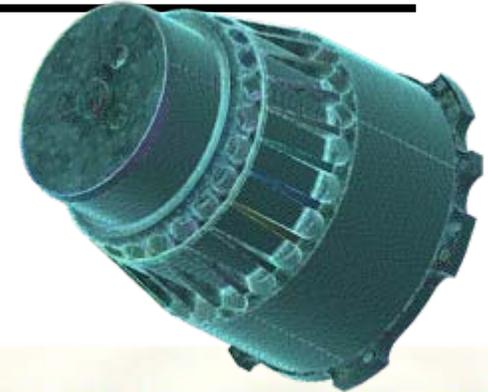
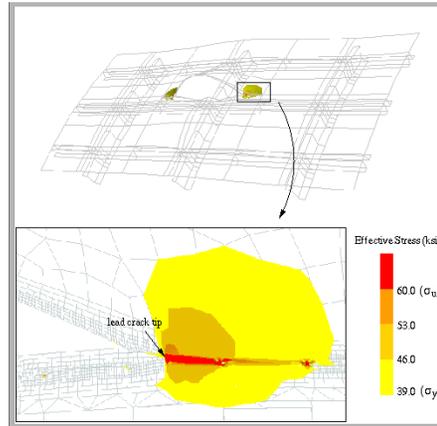
SOS10  
March 6-9, 2006

# The Forward Simulation Problem



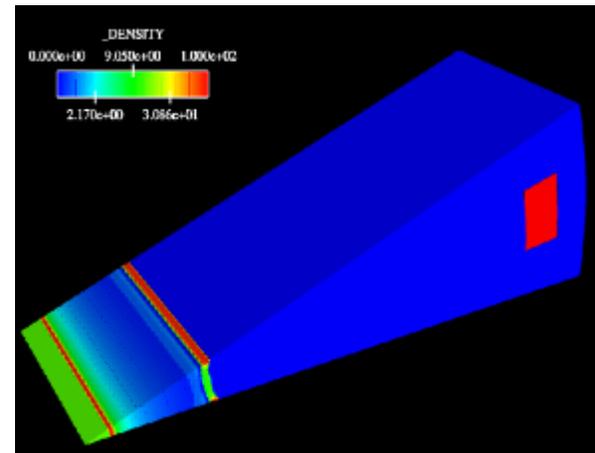
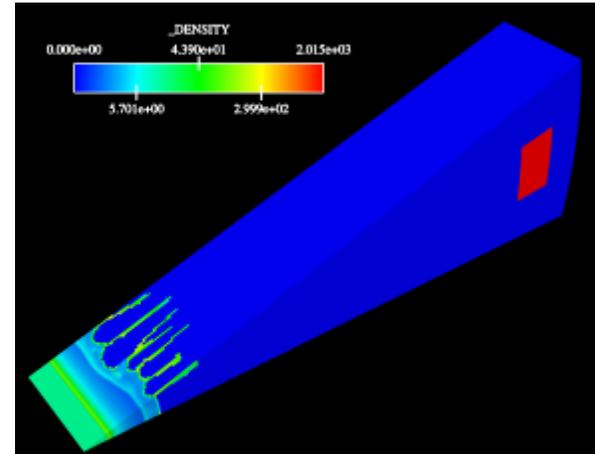
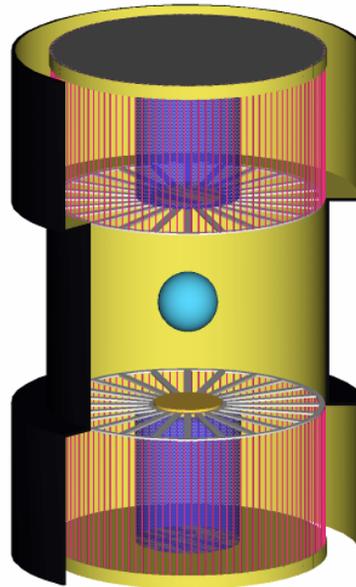
# Drivers

- Physics
  - Multiscale
- Scale
  - Billions of unknowns
  - Coupled physics
- Quality
  - Mesh
  - Discretization
- Implementation
  - Code and framework co
  - Machine architectures



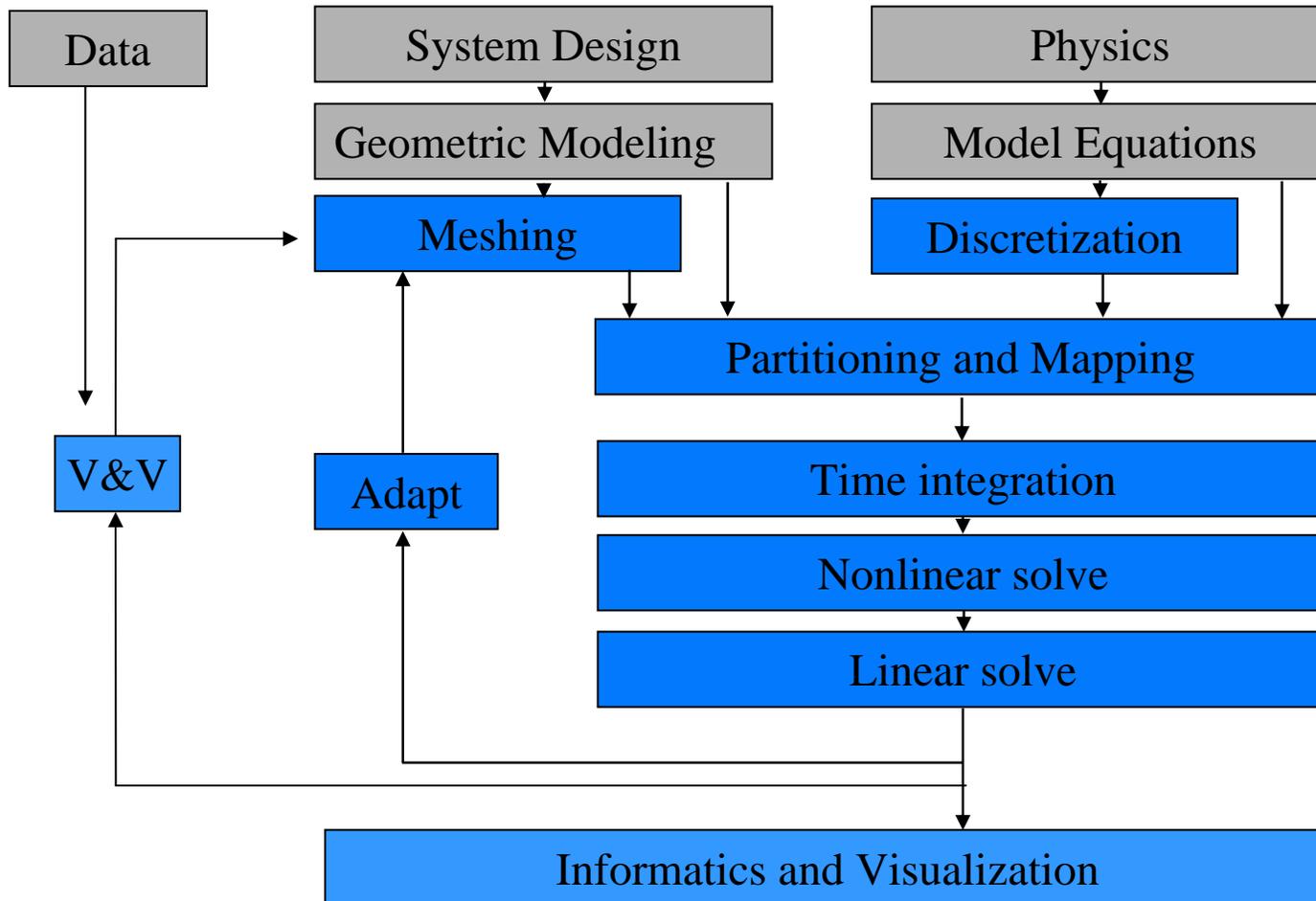
# Drivers

- Interactions between components
  - Solvers with
  - Discretization with
  - Meshing with
  - Load balancing with
  - Error estimation with
  - ...
- Example
  - Z-pinch simulations



Liner implosion

# What's Missing?



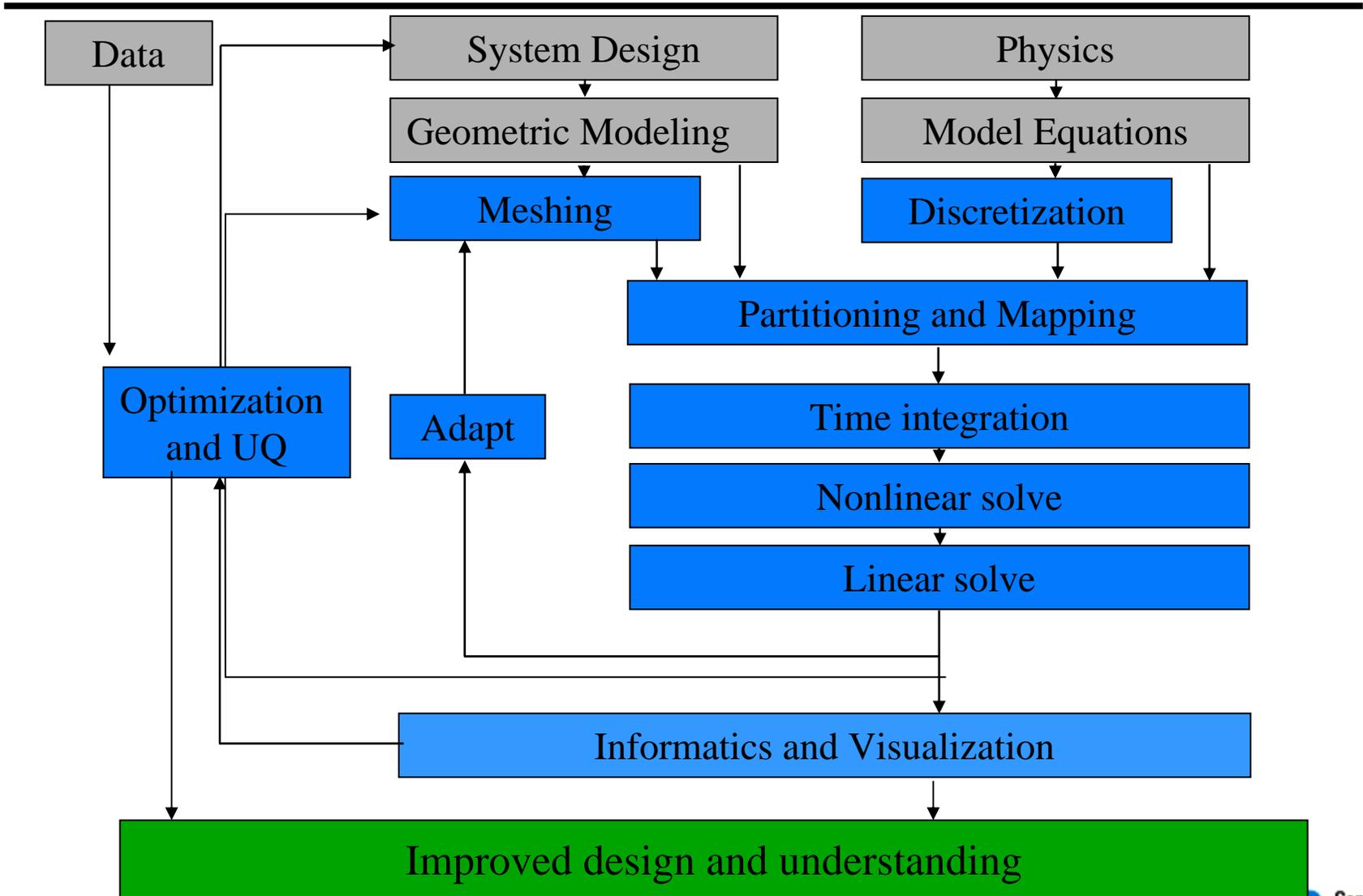


# Verification and Validation

---

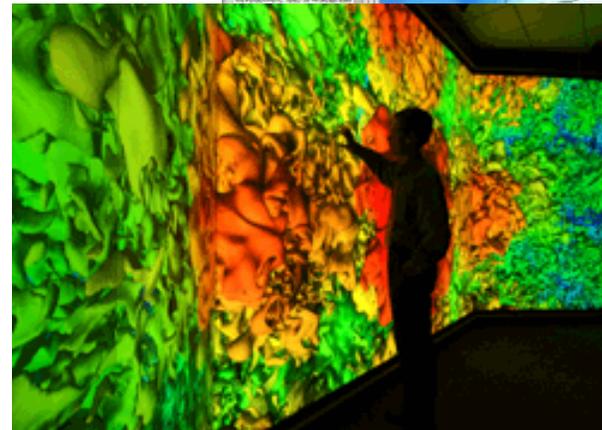
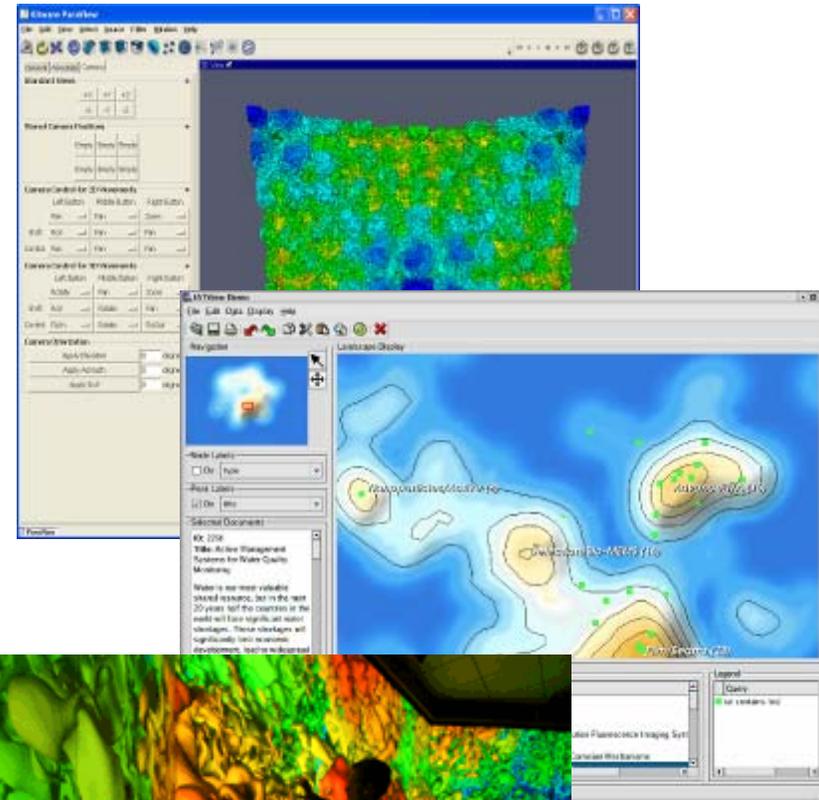
- Right now, we are using an ad-hoc, black box approach
  - Convergence studies and MMS
  - Point-wise comparisons to experimental data
- Need
  - Integration of V&V with code design, i.e., designed in.
    - » Element library with error estimators
    - » Symbolic manipulation of equations and automatic differentiation
    - » Designed to be part of a V&V framework (not a stand-alone code) and must allow some “intrusion” into the code (e.g., sensitivity calculations)
  - Algorithms
    - » Error estimators, statistical methods, input distributions and stochastic equations (including solvers)
  - Presentation of results
- Problem
  - Established infrastructure
  - Need knowledgeable users working with analysts and developers

# But What Do We Really Want To Do



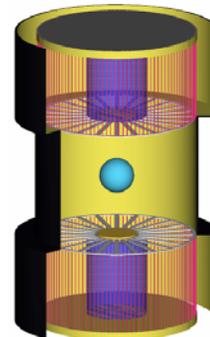
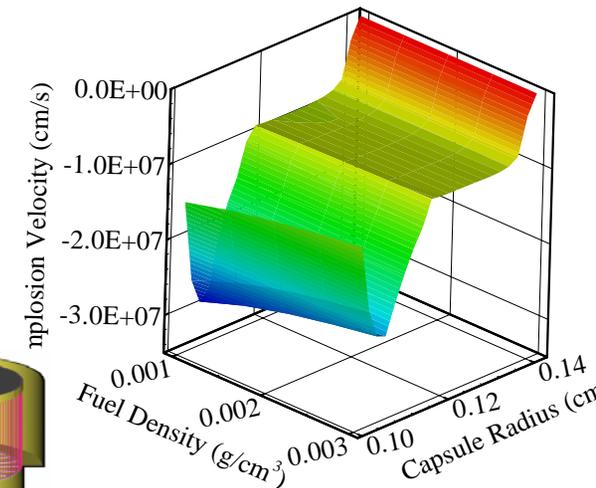
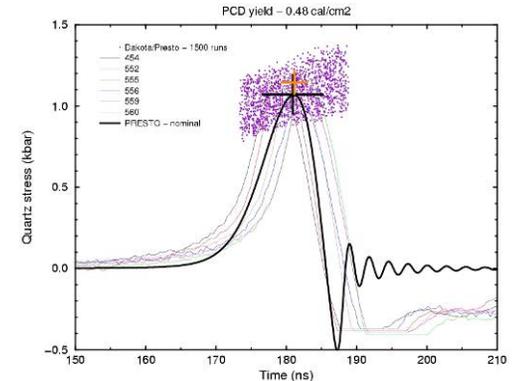
# Visualization

- Geometric information (current)
  - Large scale
  - Desktop
- Non-geometric information (future)
  - Aggregate information (many simulations)
  - Risk and uncertainties
- “Data-mining”
  - Need to think about how people really use information
  - Tools to help extract information (“features”)
  - Integration with simulation



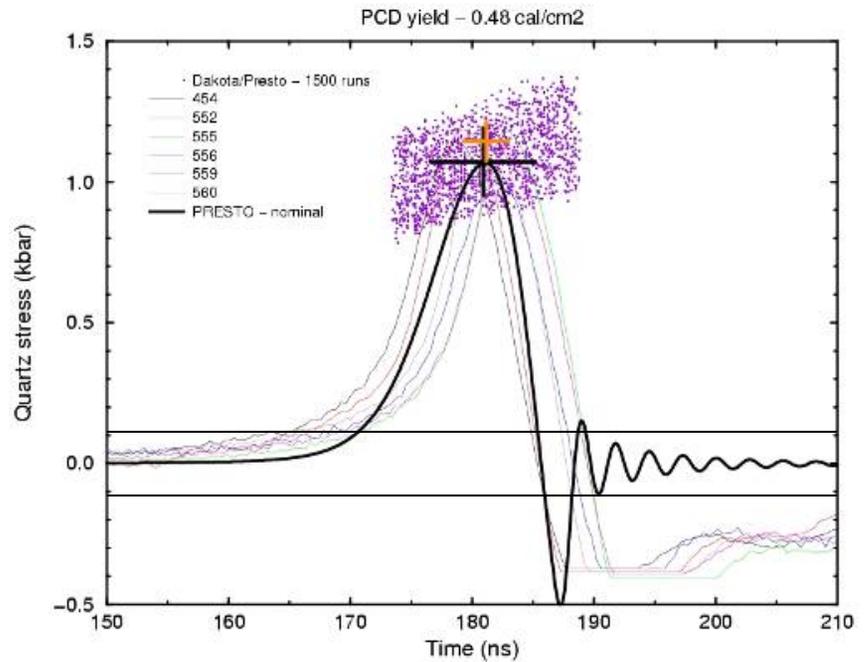
# Optimization and UQ

- **Can include V&V in this loop, but adds design and “understanding”**
  - Ensemble of “single-point runs insufficient
  - Need to design for system performance, not just assess system performance
- **Need integrated tools**
  - Many of the same tools as for V&V (e.g., error estimation, sensitivity analysis)
  - Identify critical phenomena (turning points, bifurcations points)
  - Rigorous exploration of design space
  - Integrated design and solvers
- **Need multifidelity**
  - Surrogate methods
  - Hierarchy of resolutions and scale/physics couplings
- **More integration of tools (designed in)**
  - Including V&V, meshing and visualization

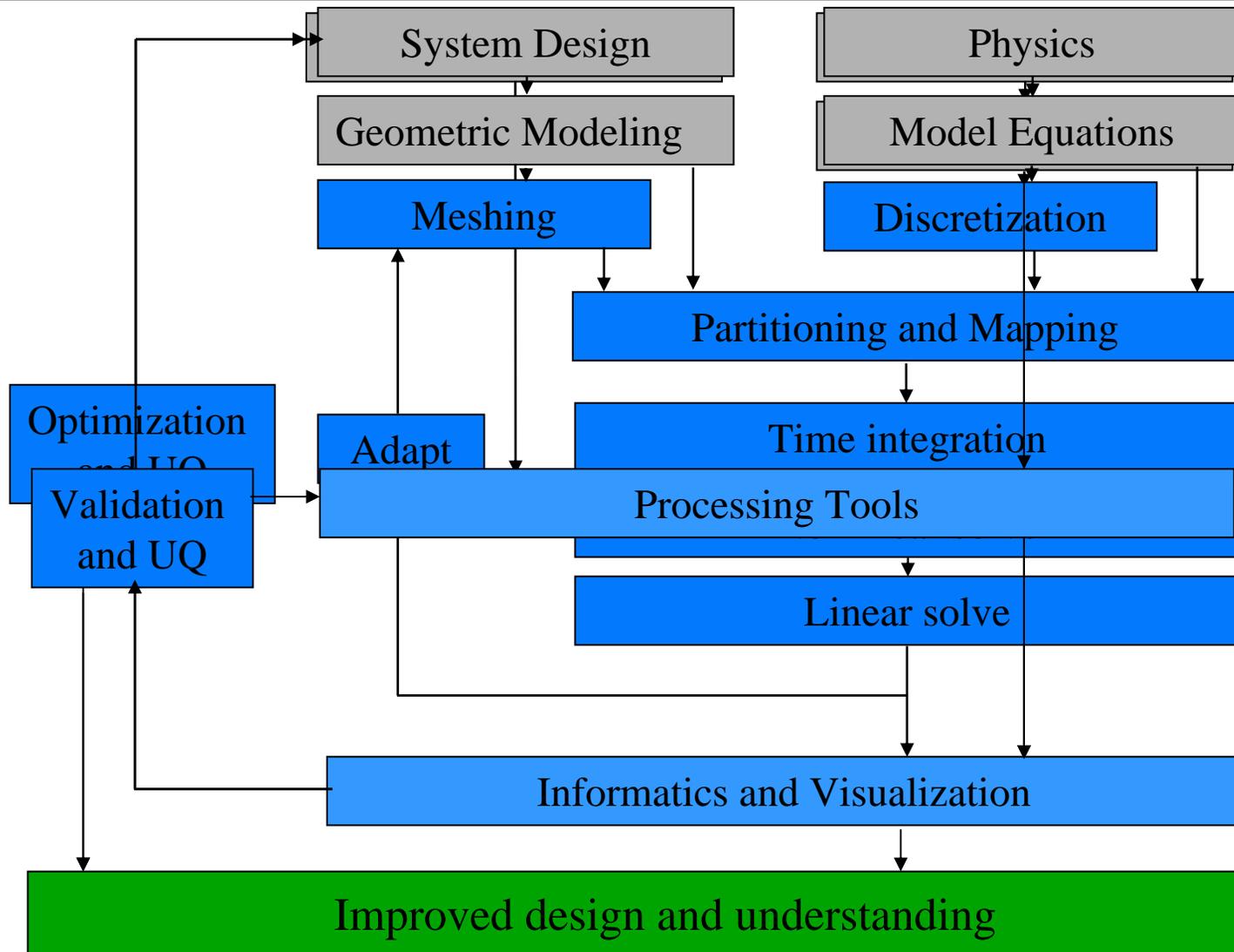


# Dealing With Uncertainty

- Simply propagating distributions and quantifying errors is not enough
- We need to be able to incorporate experimental data to further reduce uncertainties

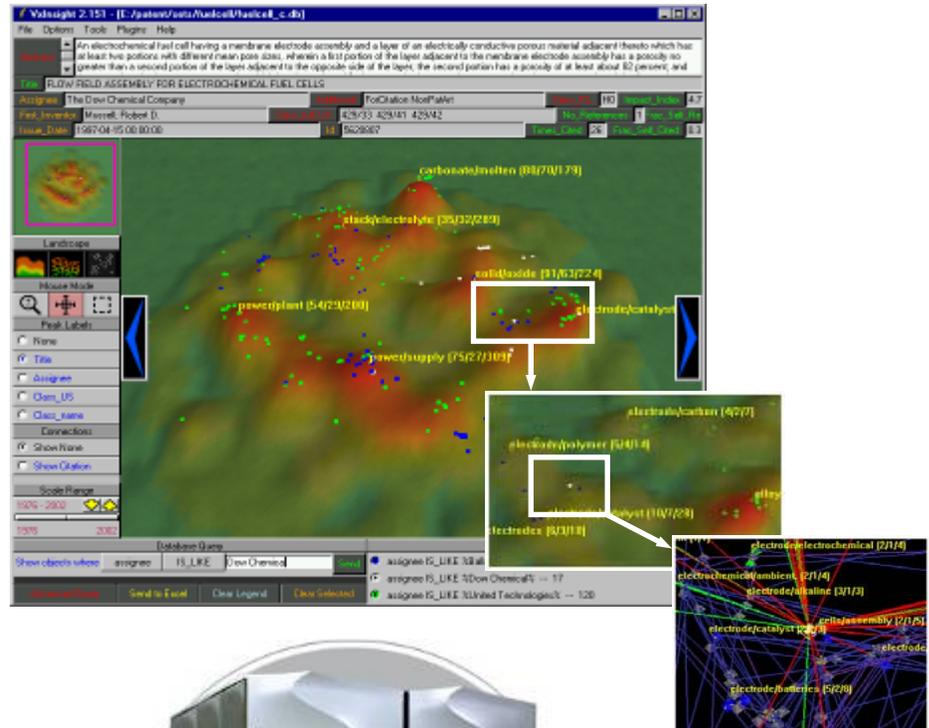


# What About The Informatics Problem?



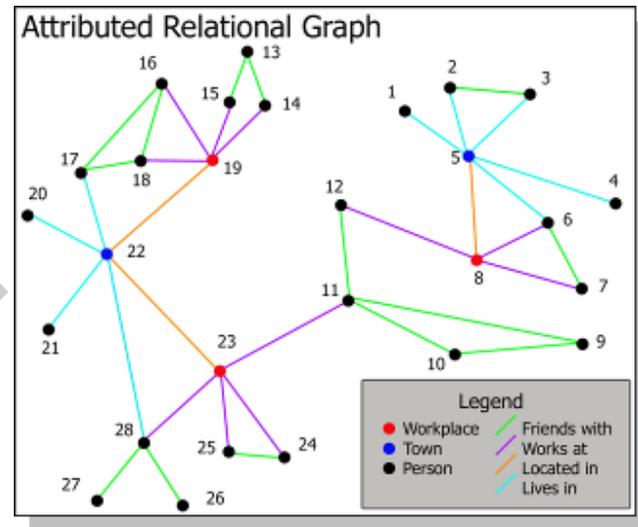
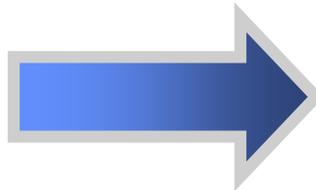
# Informatics Drivers

- Scale
  - Huge amounts of data
  - Visualization on multiple scales
  - Streaming data
  - “online” processing
- Presentation of Information
  - Making it meaningful
- Implementation
  - Code complexity
  - Architectures
- Quality
  - Representation of data
  - Dealing with model uncertainty
  - Dealing with missing data



# Representing Data

- Graphs represent relationships and are the link between the model and processing



- *Challenge*: Piecing the data together and extracting critical, relevant information in a timely manner
- Semantic Graphs or Attributed Relational Graphs (ARGs) are one way to integrate data from disparate sources
  - Vertices represent people, places, locations, events, etc.
  - Edges represent the relationships between the vertices
  - ARG encodes web of relationships

# Motif Finding

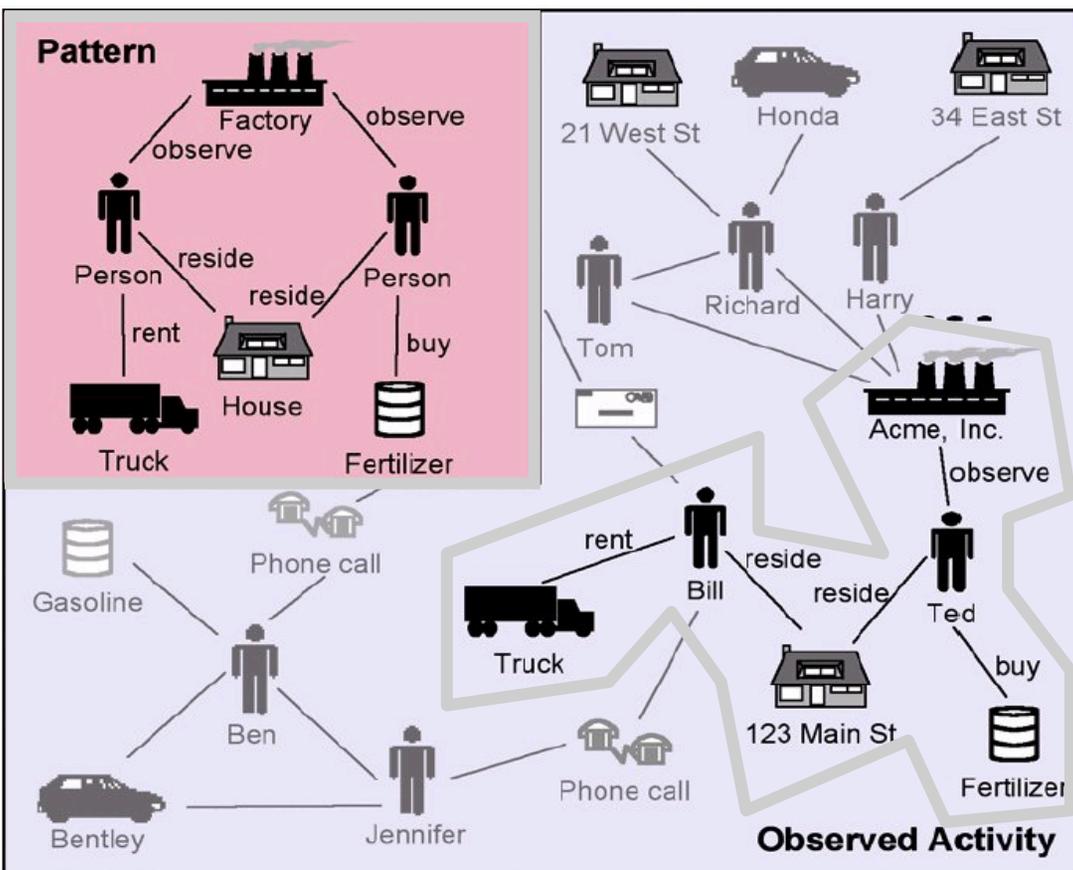
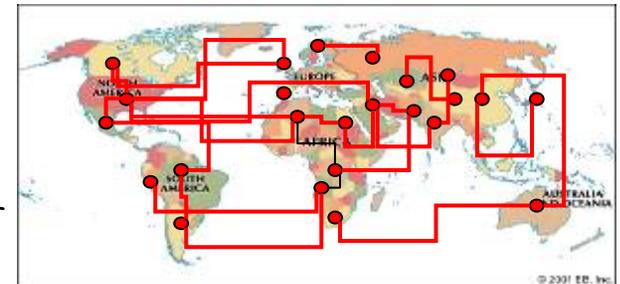
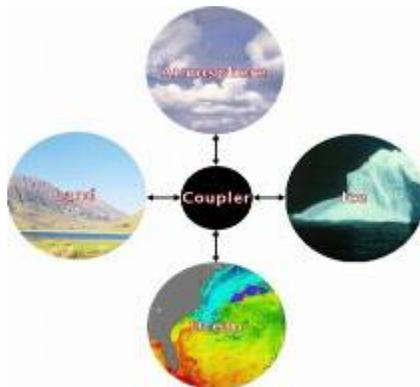


Image Source:  
T. Coffman,  
S. Greenblatt,  
S. Marcus,  
*Graph-based technologies for intelligence analysis*, CACM, 47 (3, March 2004): pp 45-47

# One More Challenge

- Merging data with large-scale modeling and simulation
  - Informatics and modeling and simulation are not separate problems
- Example - Climate simulation
  - Climate predictions will drive economic decisions
  - Climate and economics will both drive behavior





# Summary

---

- Scale will drive computing
  - Size of problem (in the traditional scientific computing model)
  - Amount of data
- Need to focus on how machines/simulation/information will be used
  - Presentation (visualization) of information
  - Support for design and decision making
- Software environments and complexity are limiting
  - Need a lightweight, simple environment that allows flexibility and linking of components. (No problem)
  - Data representation is a key question. This will drive both software and hardware design