Opportunities in Computational Applied Mathematics at Sandia National Laboratories

Brian M. Adams
Sandia National Laboratories
Optimization and Uncertainty Quantification

April 24, 2008

SIAM Student Chapter
North Carolina State University
Raleigh, NC

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Outline

A discussion of opportunities in computational science at Sandia National Laboratories

- Intro to Sandia; broad overview of CS and math
- How I got there
- Example projects
  - Social contact network model for disease spread
  - DAKOTA: from algorithm development to customer support
- Discussion

Questions are welcome at any time!

Sandia National Laboratories

SNL's Business: National Security

- **Core Purpose: to help our nation secure a peaceful and free world through technology.**
- **Highest Goal: to become the laboratory that the United States turns to first for technology solutions to the most challenging problems that threaten peace and freedom for our nation and the globe.**

Technologies for National Security

We develop technologies to:

- Sustain, modernize, and protect our nuclear arsenal
- Prevent the spread of weapons of mass destruction
- Provide new capabilities to our armed forces
- Protect our national infrastructures
- Ensure the stability of our nation's energy and water supplies.
- Defend our nation against terrorist threats

Distributed Facilities to Meet National Needs

Albuquerque,
New Mexico

Kauai Test Facility,
Hawaii

Tonopah Test Range,
Nevada

Yucca Mountain,
Nevada

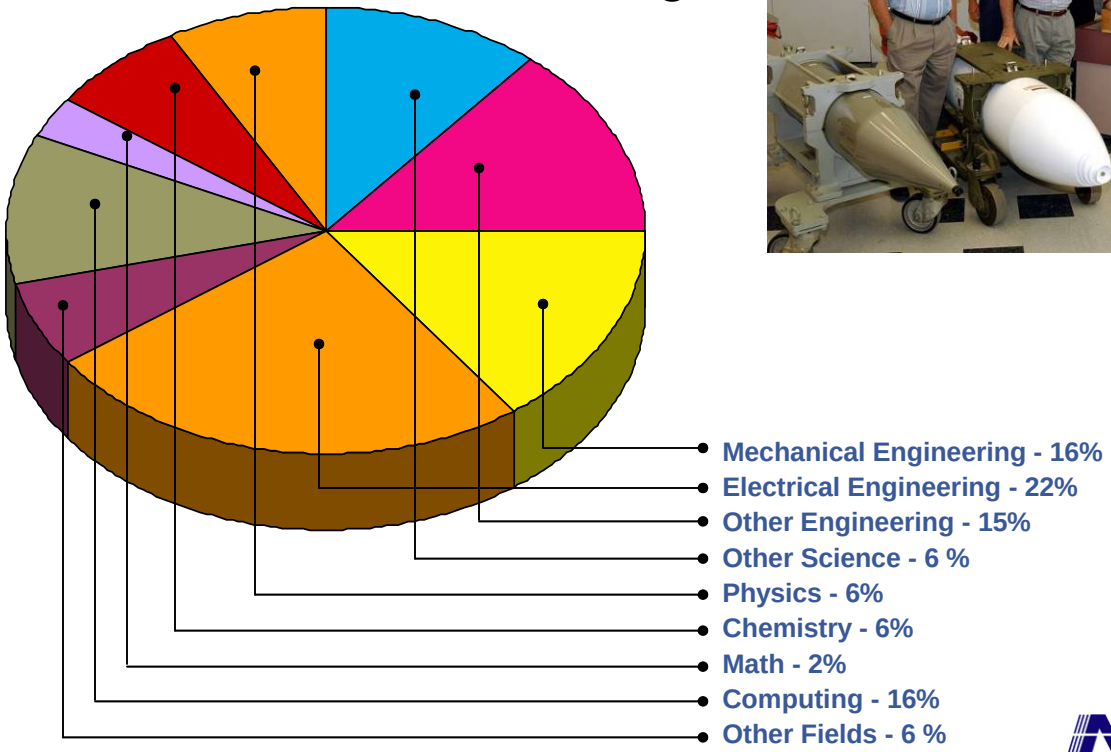
WIPP, New Mexico

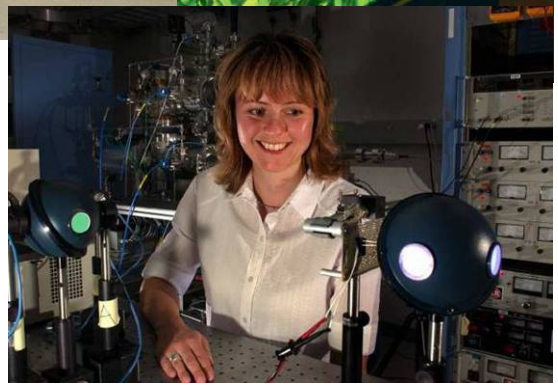
Pantex, Texas

Livermore, California

Highly Skilled Workforce

- More than 8,600 full-time employees
- More than 1,500 PhDs and 2,700 MS/MAs
- 2,200 on-site contractors
- $2.33 billion FY06 total budget





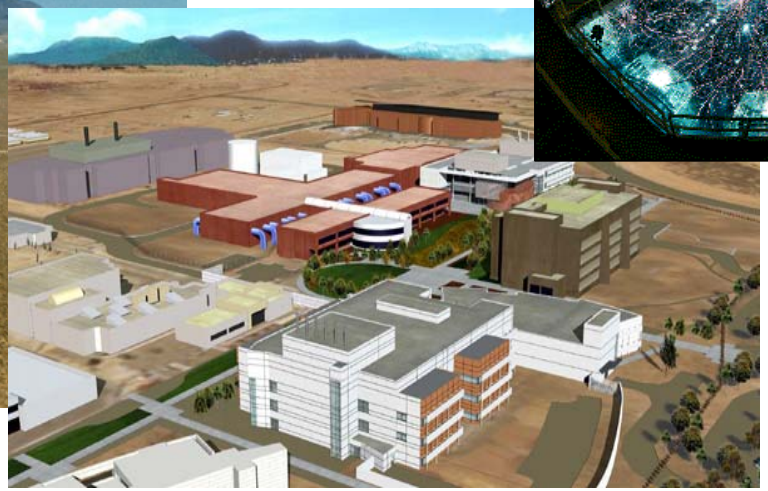
Nuclear Deterrence for National Security

Our Defense Programs Mission

Our Vision:

Credible deterrence built on both a safe, secure and reliable nuclear weapons stockpile that is capable of meeting all military requirements – now and in the future – and a science-based engineering infrastructure capable of responding to national security needs whenever they arise.

B1 bomber drops B83



Microsystems, Engineering Sciences and Applications (MESA) complex

Sandia's Z Pulsed Power Machine

Defense Systems & Assessments Programs

- Proliferation Assessment
- Remote Sensing and Verification
- Surveillance, and Reconnaissance
- Integrated Military Systems

Shuttle Heat Shield Inspection

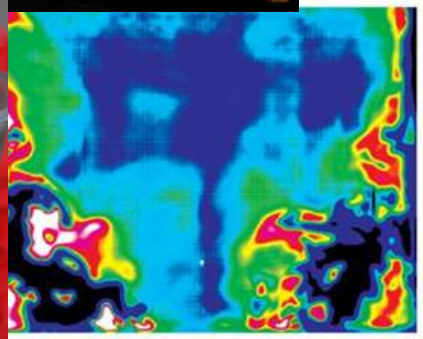
Future Combat System Integrated Support Team

- Information Operations
- Industrial Relations
- Intelligent Transformational Systems
- Science and Technology Products

Energy, Resources, and Nonproliferation

Energy, Water, and Security Enabled by Science & Technology

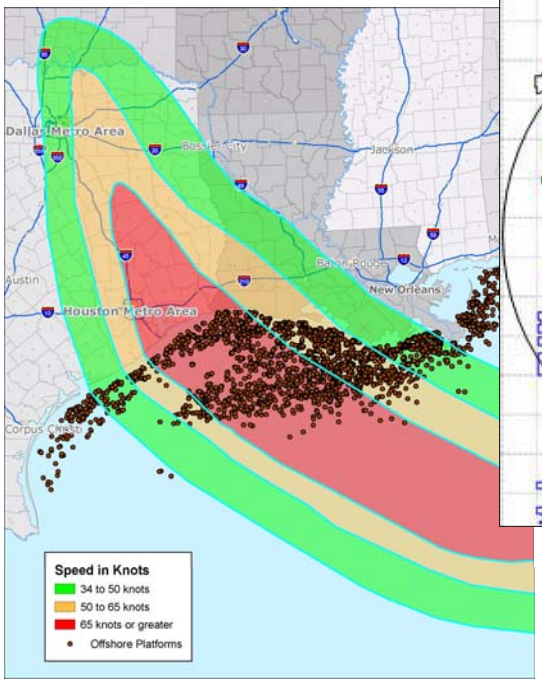
- **Secure energy supplies for national security**
- **Clean, abundant, affordable energy**
- **Water research**
- **Infrastructure protection**
- **Hydrogen & Nuclear Power**
- **Waste Management**

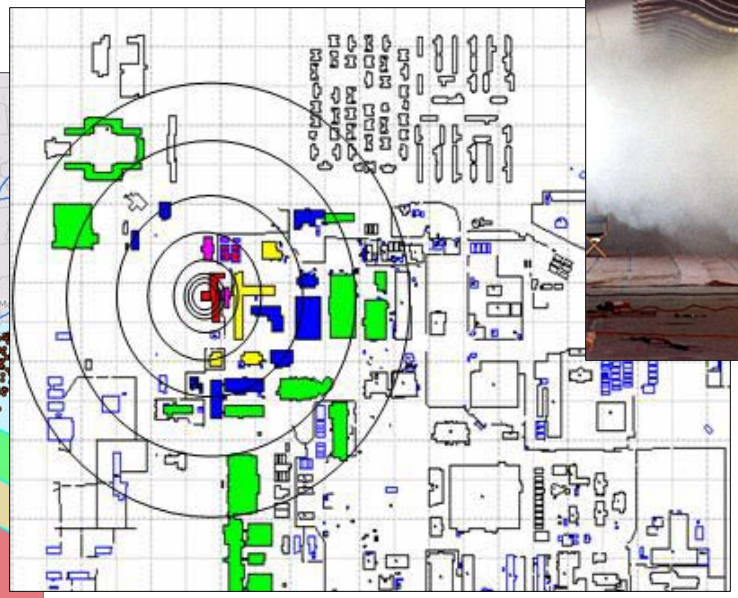


Homeland Security and Defense

Mitigating the risk of catastrophic events and enhancing the nation's ability to respond and recover

Risk Management and Infrastructure Protection

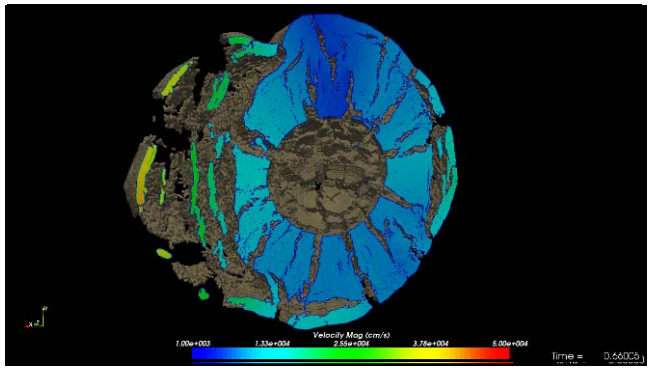




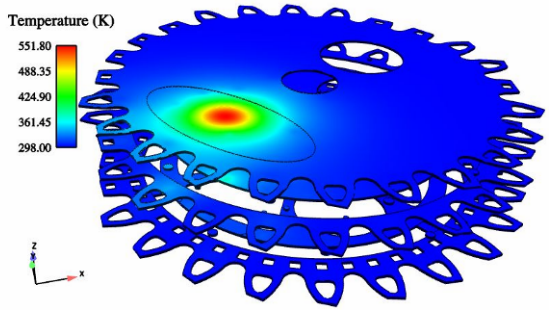
Homeland Defense and Force Protection

Catastrophic Event Mitigation

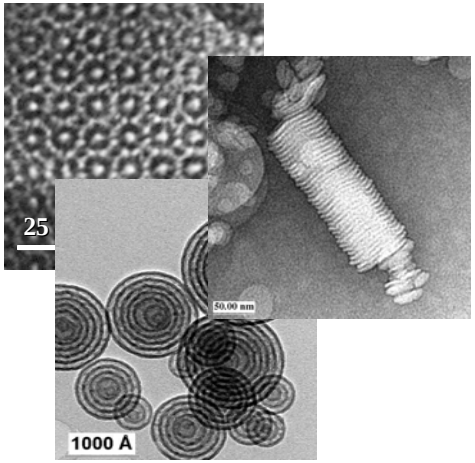
Our Mission Focus Relies on Strong Science and Engineering



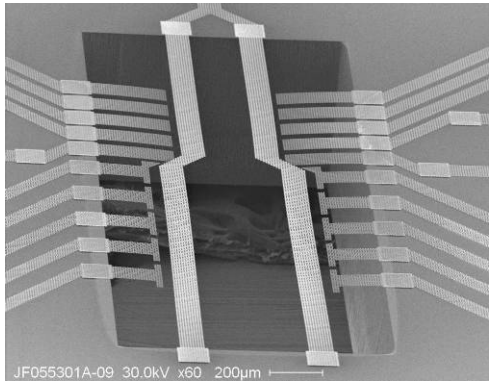
Computational and Information sciences



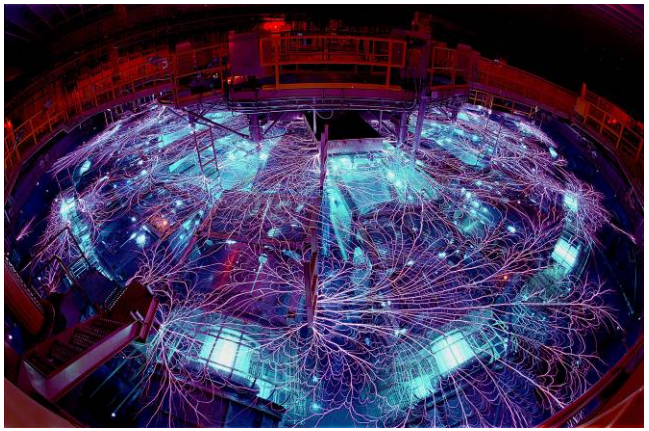
Engineering Sciences



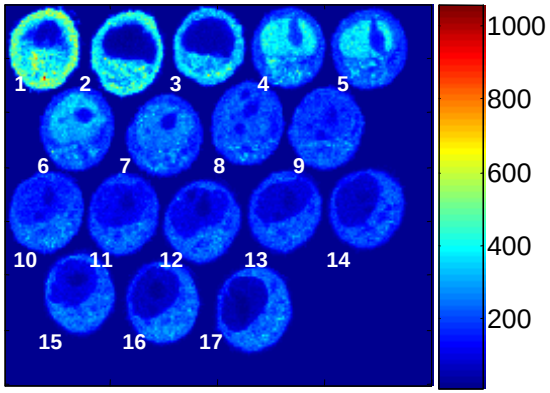
Materials Science and Technology



Microelectronics and Photonics



Pulsed Power

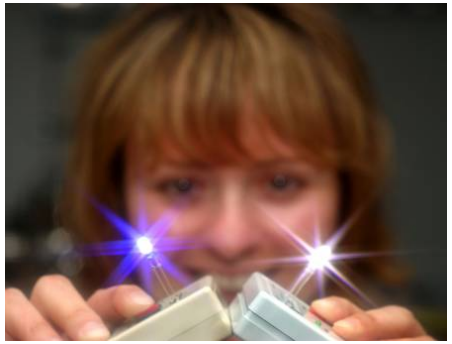


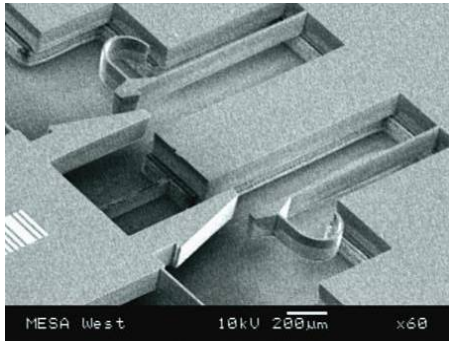
Bioscience

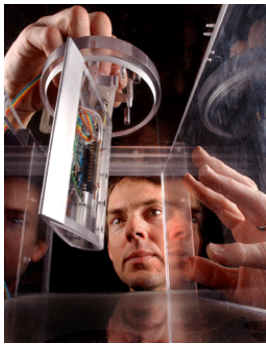
We Invest in Technologies for National Security

Sandia's Science and Engineering Advances

- **Our Science, Technology and Engineering has changed the world**
- **In 57 years, Sandia advances have revolutionized:**
  - **the electronics industry and medicine**
  - **the way computers work and problems are solved**
  - **the design-to-manufacture cycle of advanced products**
  - **our ability to use materials properties for enhanced national security**
- **ST&E supports all missions—advances for nuclear weapons safety can lead to a ten-fold reduction in lighting homes and businesses**

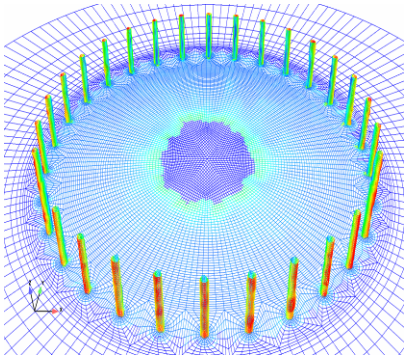






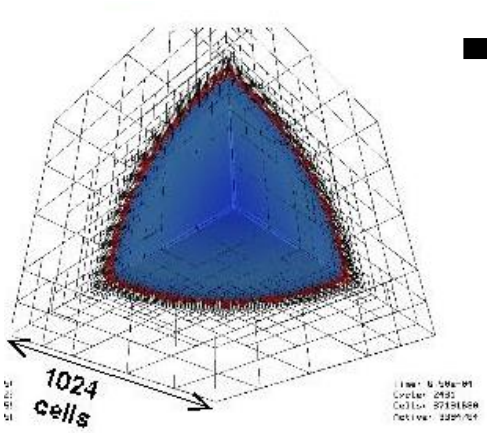
Components of Predictive Computational Capability

Computing Architectures And Systems



Computational Science Applications

Algorithms and Enabling Technologies



+

Workhorse Hardware

Sandia National Laboratories

Computer Science Research Institute

- Our collaboratory with the world (academia, industry)
- Mathematics and algorithms research
- A meeting place for new ideas to keep our research fresh
- Expands our focus on external partnerships for SNL computing R&D
- A resource for students and future staff

High Performance Computing: Sandia Systems and Software

CM-2
1988

nCUBE-2
1990

iPSC-860
1992

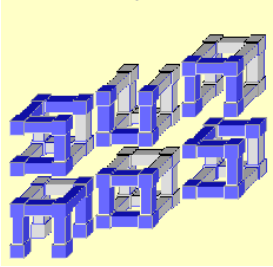
Paragon
1993

ASCI Red
1996

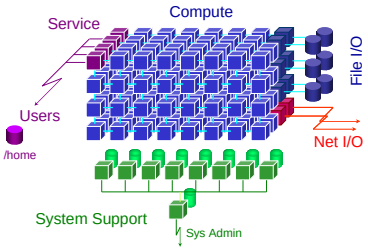
Cplant
1998

Red Storm
2005

1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005

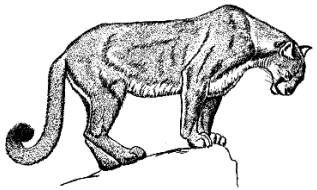


SUNMOS
1990 - 1997



Partition Model
1990 –

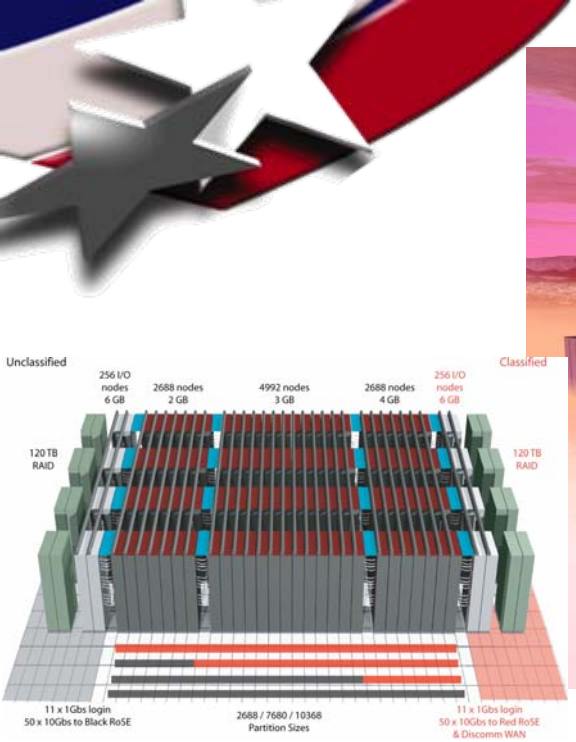
Portals
1992 –



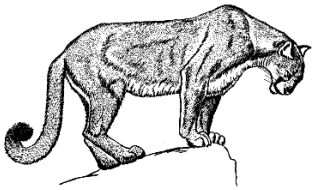
Puma, Cougar,
Catamount
1993 –

Cplant
1997 – 2005

FASTOS
2004 -

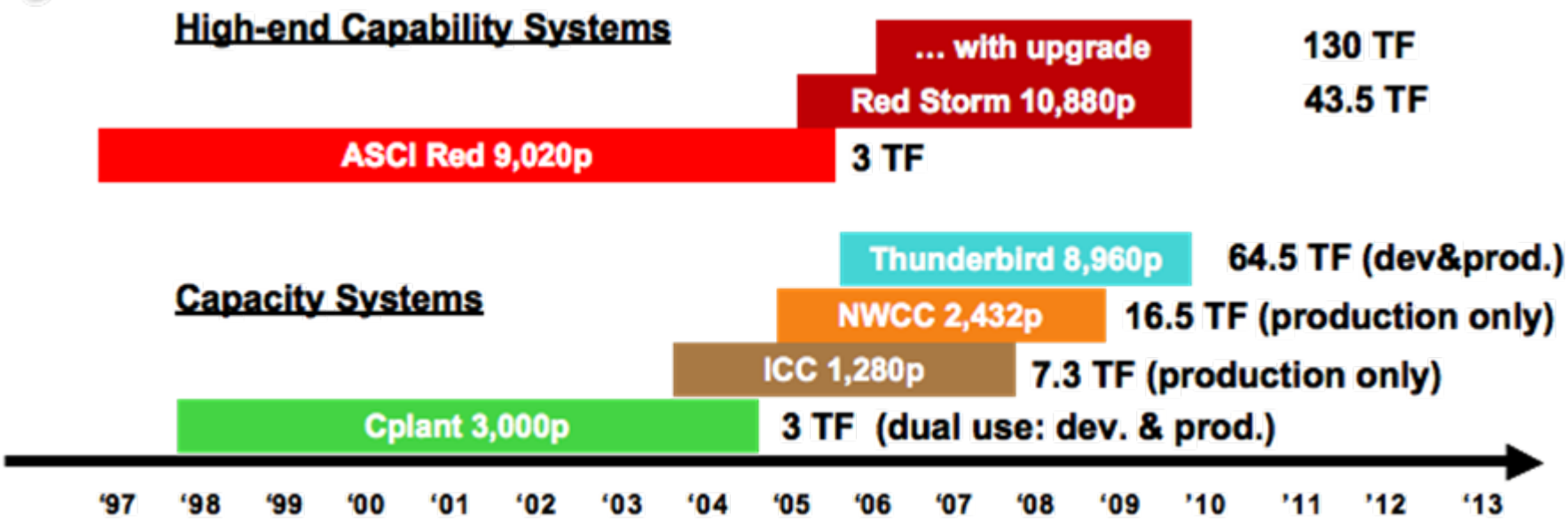


- **RedStorm: Sandia/Cray venture with "nearly linear" scaling John Daly (LANL)**
- **Currently #6 supercomputer in the world**
- **Sandia expertise:**
  - **Systems architectures**
  - **Middleware**
  - **Systems software**
  - **Supercomputer Installation & Maintenance**
  - **Performance analysis and tuning**



Sandia National Laboratories

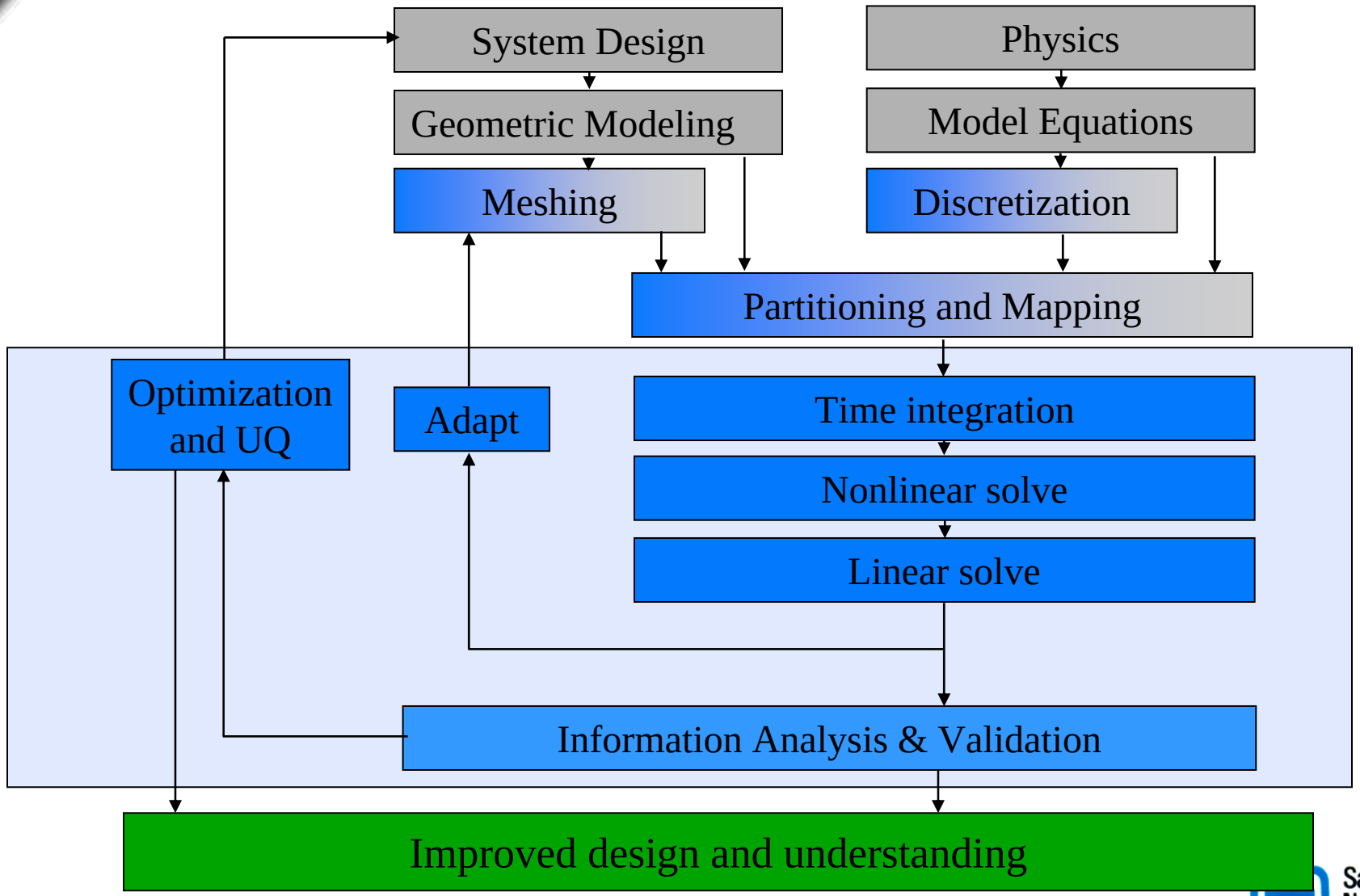
Near-term Capacity and Capability Timeline



- Red Storm Upgrade Underway
  - 54 TF to 130 TF (multicore)
  - Memory Upgrade (6GB/node)
- Red Storm Architecture Evolution
  - 1+ Petaflop in 2008

Sandia National Laboratories

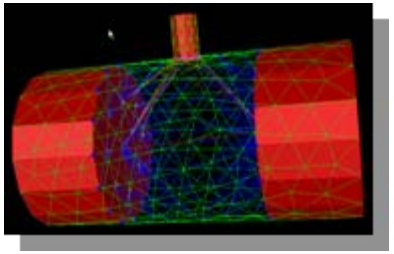
Algorithms for Computational Modeling & Simulation



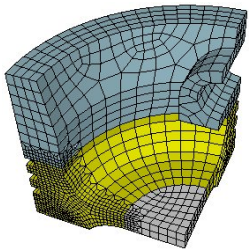
Sandia National Laboratories

Cubit Mesh Generation

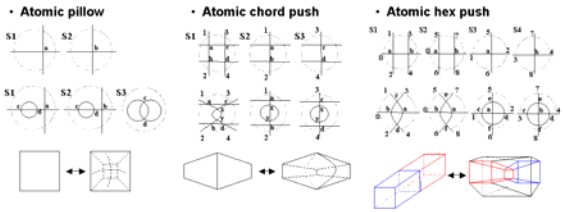
All-Hex Research



Conformal all-hex refinement

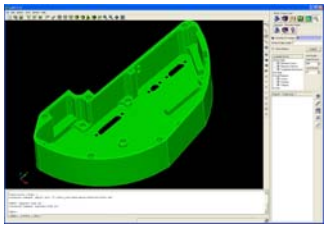


Cable weaving: all hex algorithm research



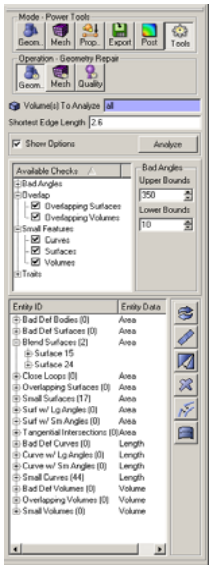
Hex topology research: improving connectivity of hex meshes

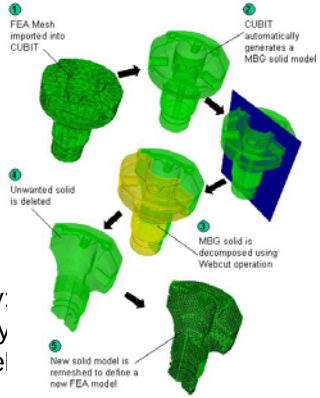
Geometric Editing



ACIS and Pro/E Native Geometry Kernels: Robust CAD model editing and management

Geometry Power Tool: Diagnose and fix problem geometry





Mesh-based Geometry: Build and edit geometry directly from FEA model

Component Architecture

Components used in CUBIT and in other NWC applications

CubitLib
Core CUBIT component

CLARO
Graphics and GUI environment

CGM
Geometry query and modification tools
CAD abstraction layer
Interfaces to ACIS, Pro/E, facet, virtual

CAMAL
Library for core meshing algorithms
Tet, Quad, Tri meshing components

Cholla
Facet-based geometry engine

Mesquite/Verde
Mesh improvement and mesh quality libraries

FacetBoolean
Facet-based boolean library

MOAB
Mesh Database
Large mesh out-of-core storage

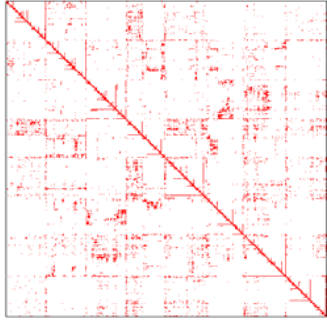
Large Mesh Capability

Routinely generates and handles 10s of Millions of Elements

Plans for FY05 to routinely handle 100s of Millions of Elements

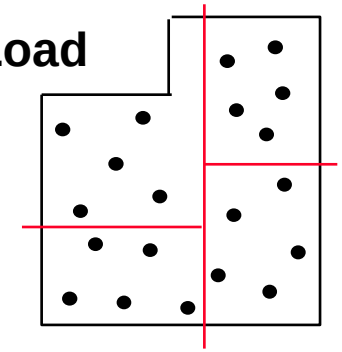
Sandia National Laboratories

Zoltan Parallel Data Services
Parallel Partitioning & Load Balancing

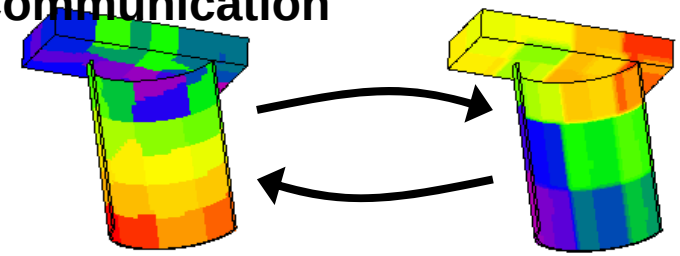


Xyce ASIC matrix (Stripped)
683K x 683K; 2.3M nonzeros

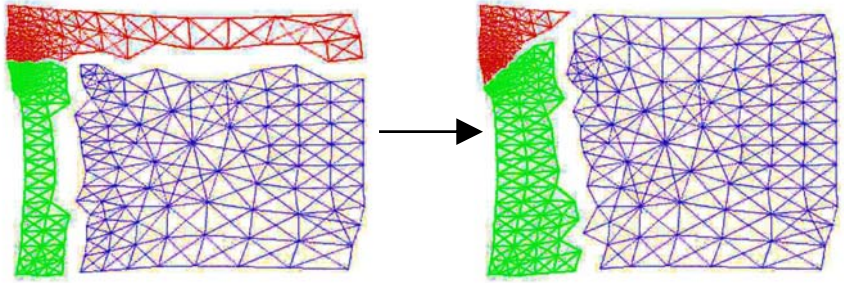
Dynamic Load Balancing



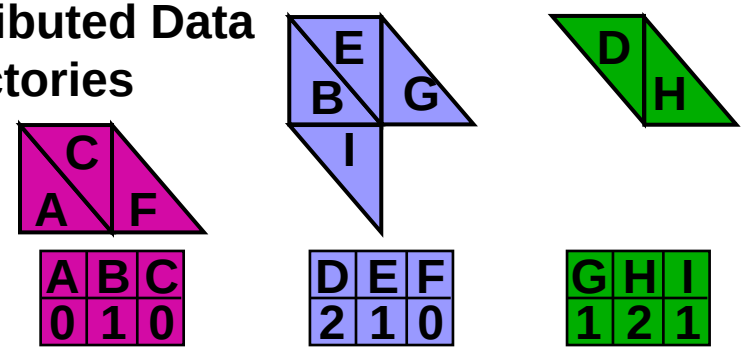
Efficient Unstructured Communication



Data Migration

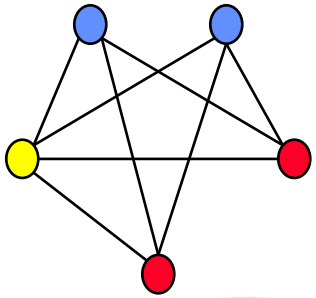


Distributed Data Directories



Dynamic Memory Debugging

Graph Coloring



- **Parallel hypergraph partitioning advantages: rectangular and nonsymmetric systems**

Open-source software at http://www.cs.sandia.gov/Zoltan.

Sandia National Laboratories

Trilinos HPC Solvers

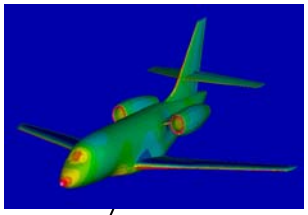
FY06

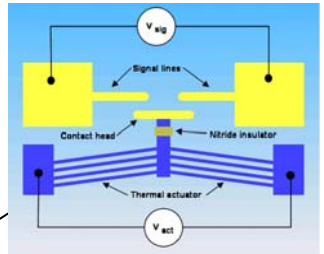
Trilinos 6.0 released; >1000 downloads

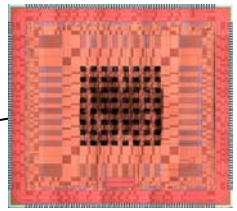
Focus on continued package development and software engineering

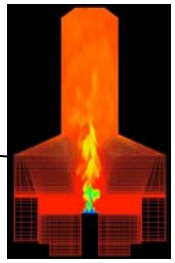
Awards

R&D 100 (in 2004)
IEEE HPC Software Challenge Award
Sandia ERA team award and NOVA nomination









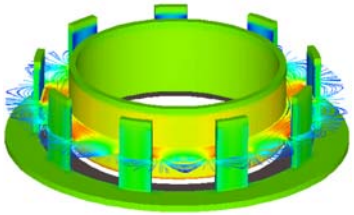
Trilinos

First software architecture to allow community development

Two-level design:

- Self-contained packages
- Leveraged common tools.
- Allows rapid algorithmic development and delivery

Leveraging investments in software infrastructure without compromising individual package autonomy



Sandia National Laboratories

(Some of)

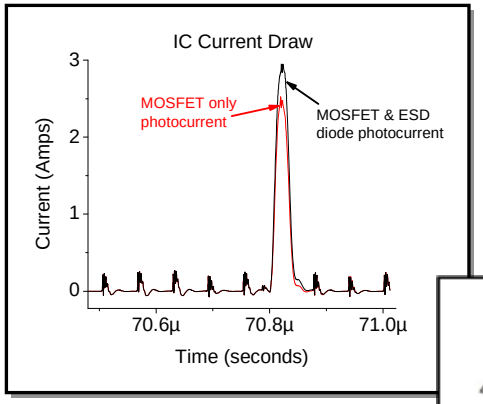
What can be done with Trilinos

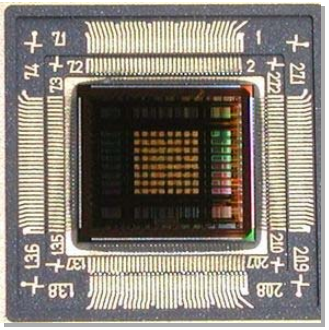
Objective	Package(s)
Distributed linear algebra objects	Epetra, Jpetra, Tpetra
Krylov solvers	AztecOO, Belos
ILU-type preconditioners	AztecOO, IFPACK
Smoothed aggregation, multilevel preconditioner	ML
Eigenvalue problems	Anasazi
Block preconditioners	Meros
Direct sparse linear solvers	Amesos
Direct dense solvers	Epetra, Teuchos, Pliris
Abstract interfaces	Thyra
Nonlinear system solvers	NOX, LOCA
Complex linear systems	Komplex
C++ utilities, (some) I/O	Teuchos, EpetraExt

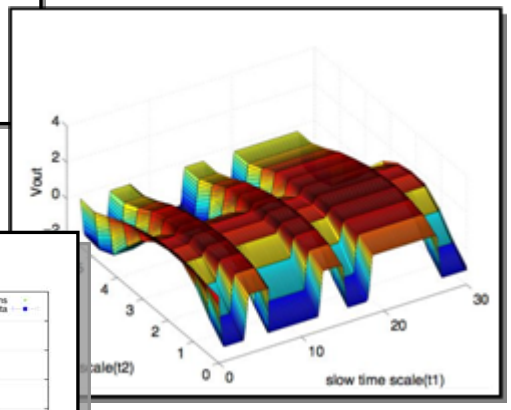
Xyce Circuit Simulation Code

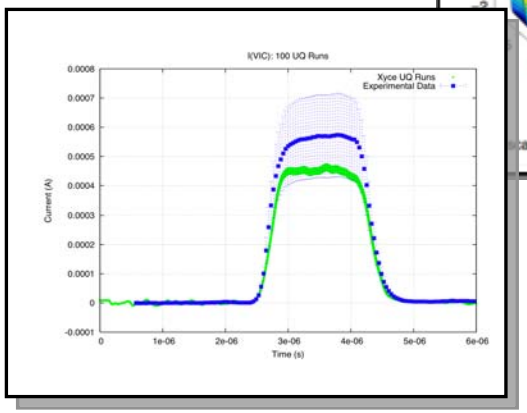
Sandia's SPICE-like circuit simulation code

- **Simulate extremely large circuits (e.g., ASICS+analog) at transistor level *to capture:***
  - ***Predictive radiation effects (x/γ–ray effects, transient neutron)***
  - ***Parasitics***
  - ***Power***
- **Massively parallel code allows:**
  - ***Simulation of very large-scale complex system-model***
  - ***Investigate circuit interaction***
  - ***Simulate large digital components***
- **Advanced solution methods & V&V:**
  - ***Multi-time PDE methods for oscillators***
  - ***Coupled advanced sampling & optimization with DAKOTA***
- **Impact:**
  - ***W76-1 AFS Qualification***
  - ***QASPR***
  - ***Integrated Stockpile Evaluation***





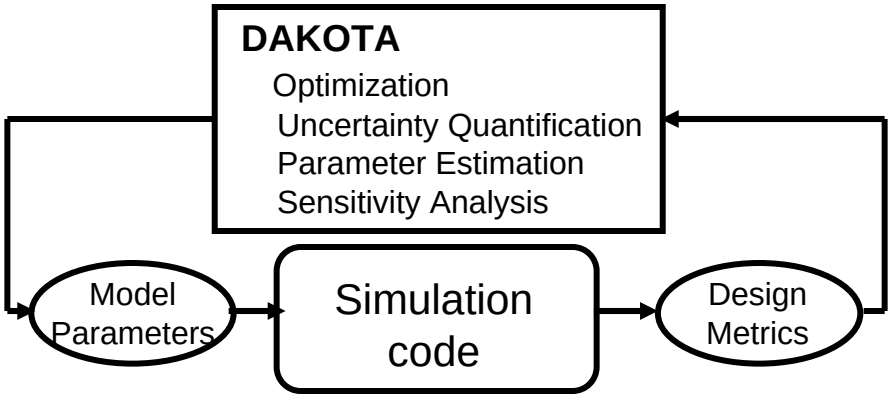




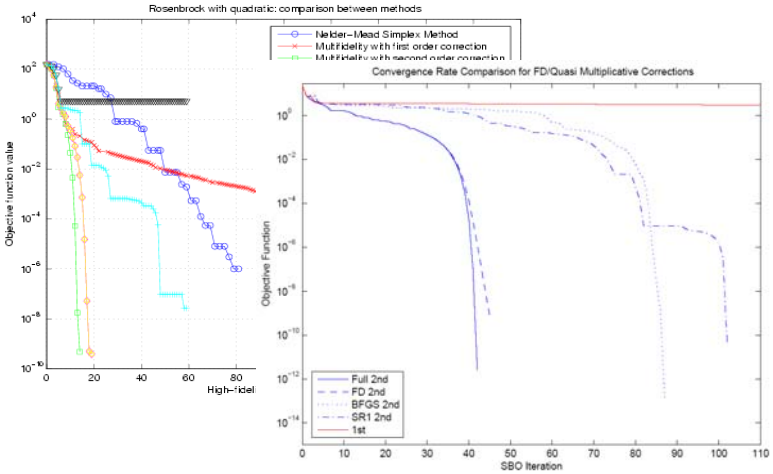
QASPR
QUALIFICATION ALTERNATIVES TO SPR

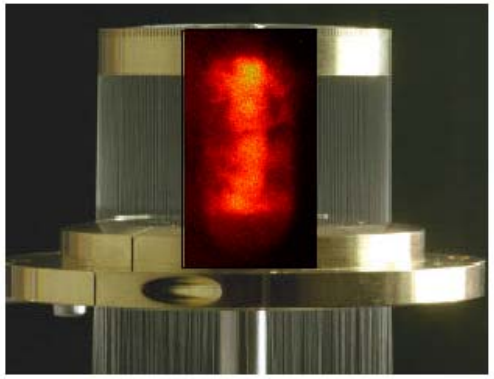
Sandia National Laboratories

Optimization and Design



Surrogate-based Optimization

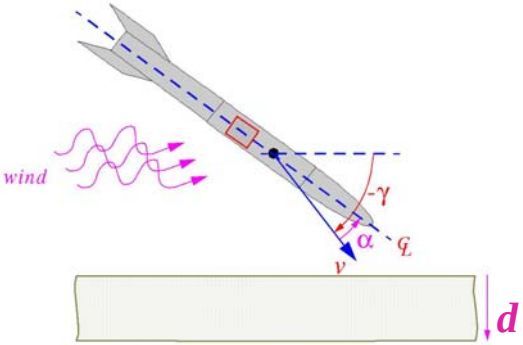


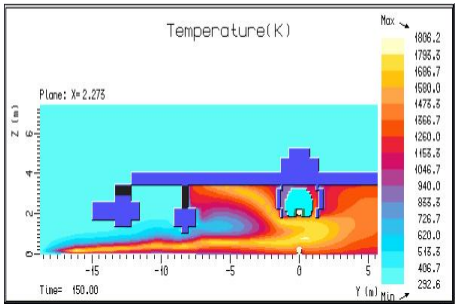


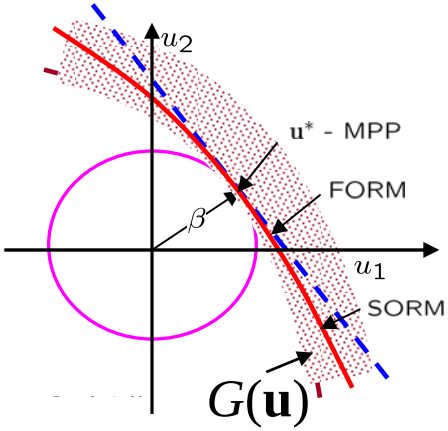
Sandia National Laboratories

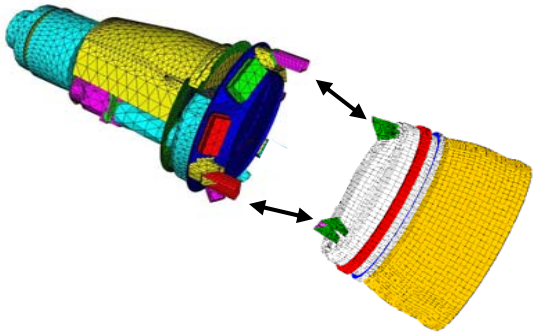
Uncertainty Quantification

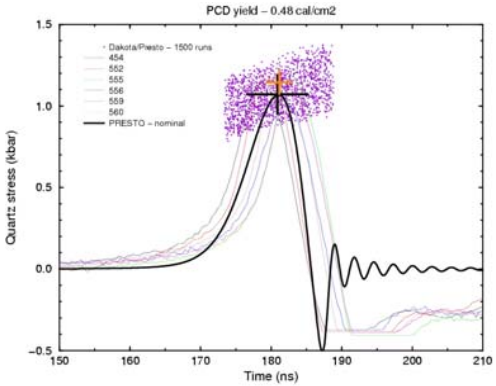
- Required for robust, “real” designs
- Required for verification and validation
- Algorithms delivered through the Dakota Framework







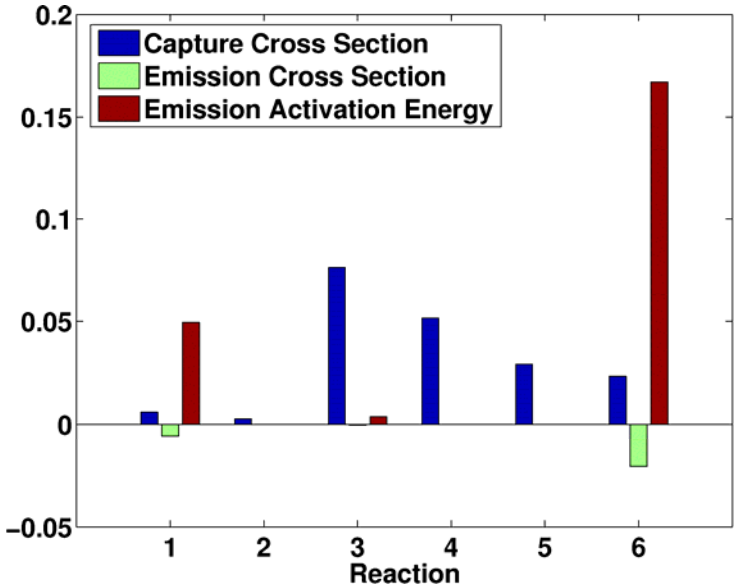




Automatic Differentiation

- Use C++ templates to automatically compute derivatives
- Has begun to change the way developers use derivatives in codes
  - Much reduced complexity
  - Reduced implementation time
  - Reduced errors
- Is critical to the increased effort in more predictive codes

Scaled Current Sensitivities



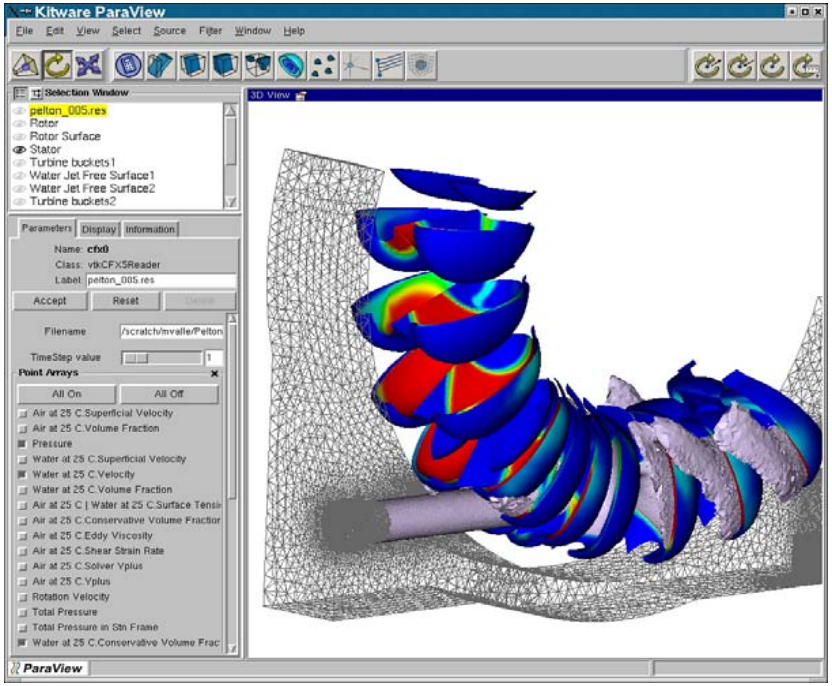
$$\left(\frac{dI}{dp}\right)^T = -\left(\frac{\partial f}{\partial p}\right)^T \left[\left(\frac{\partial f}{\partial x}\right)^{-T} \left(\frac{\partial I}{\partial x}\right)^T\right] + \left(\frac{\partial I}{\partial p}\right)^T$$

Sandia National Laboratories

ParaView

an open-source, scalable visualization toolkit

- Sandia develops large pieces of the world's most scalable visualization system…
  - Parallel Load Balancing/Data Distribution (D3)
  - Parallel Rendering (ICE-T)
  - Desktop Delivery (SQUIRT)
- Global Influence
  - Used at over 60 laboratories and universities worldwide
- Press release with NVIDIA
  - Fastest rendering system in the world (Sandia's Ice-T 1.5 Billion tris/sec)
  - Streaming Image delivery World's largest polygonal dataset



return

Sandia National Laboratories

Route to Sandia

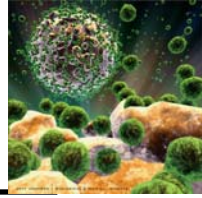
Ph.D., Computational and Applied Mathematics, NC State

- mathematics, statistics, computer science, immunology
- nondeterministic model calibration (HIV)
- internship at Fred Hutchinson CRC (insufficient!)

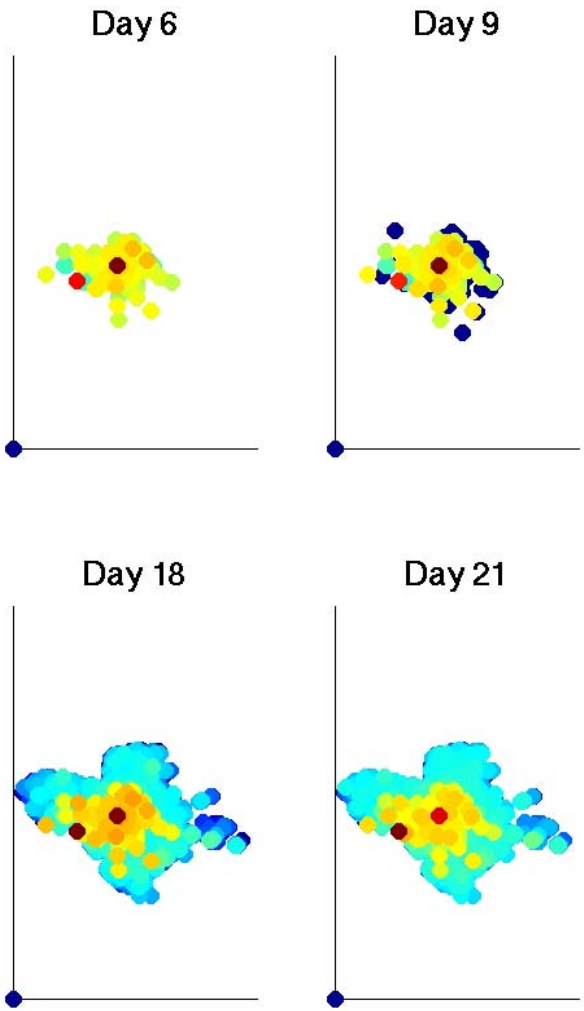
SNL since 2005 to fulfill goals:

- optimization focus (surprise: uncertainty quantification)
- develop algorithms; production software implementation in DAKOTA
- work with science/engineering application customers; *let their unmet needs drive research and software*

Disease Outbreak Characterization



❖ ***GOAL: Determine source and magnitude of natural or terrorist disease outbreak, given patients presenting for treatment***

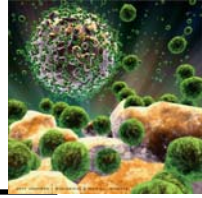


- **3-year effort with Jaideep Ray (PI), Karen Devine, Youssef Marzouk, Michael Wolf (UIUC), others**
- **Bayesian inverse problem to determine initial conditions**
- **My primary contribution: agent-based social contact network disease propagation simulator**
  - **model geographic spread**
  - **model chain of contacts**

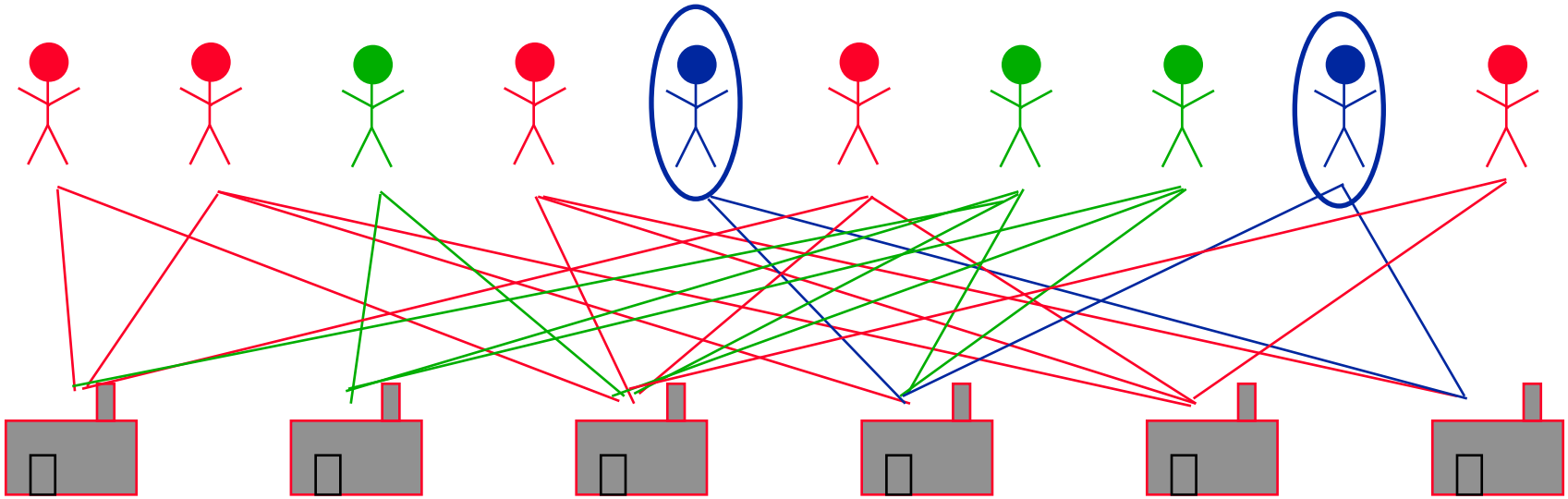
More depth on this at ACS08 Saturday, 9:15 am

Sandia National Laboratories

Network-based Disease Model



Each person's health modeled with differential equations; dynamic social network connecting people:



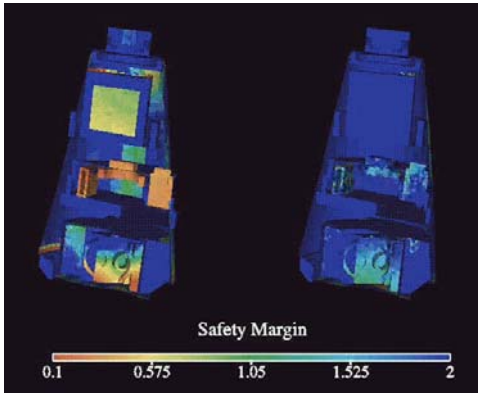
- **Graph theory for parallel partitioning and model reduction (cluster analysis on bipartite graph)**
- **Math biology for in-host disease models (pathogenesis) and transmission (epidemiology)**
- **Scalable parallelism via MPI for efficient simulations**

Sandia National Laboratories

DAKOTA Motivation

Goal: perform iterative analysis on (potentially massively parallel) simulations to answer fundamental engineering questions:

- **What is the best performing design?**
- **How safe/reliable/robust is it?**
- **How much confidence do I have in my answer?**



Nominal **Optimized**

DAKOTA
optimization, sensitivity analysis, parameter estimation, uncertainty quantification

parameters
(design, UC, state)

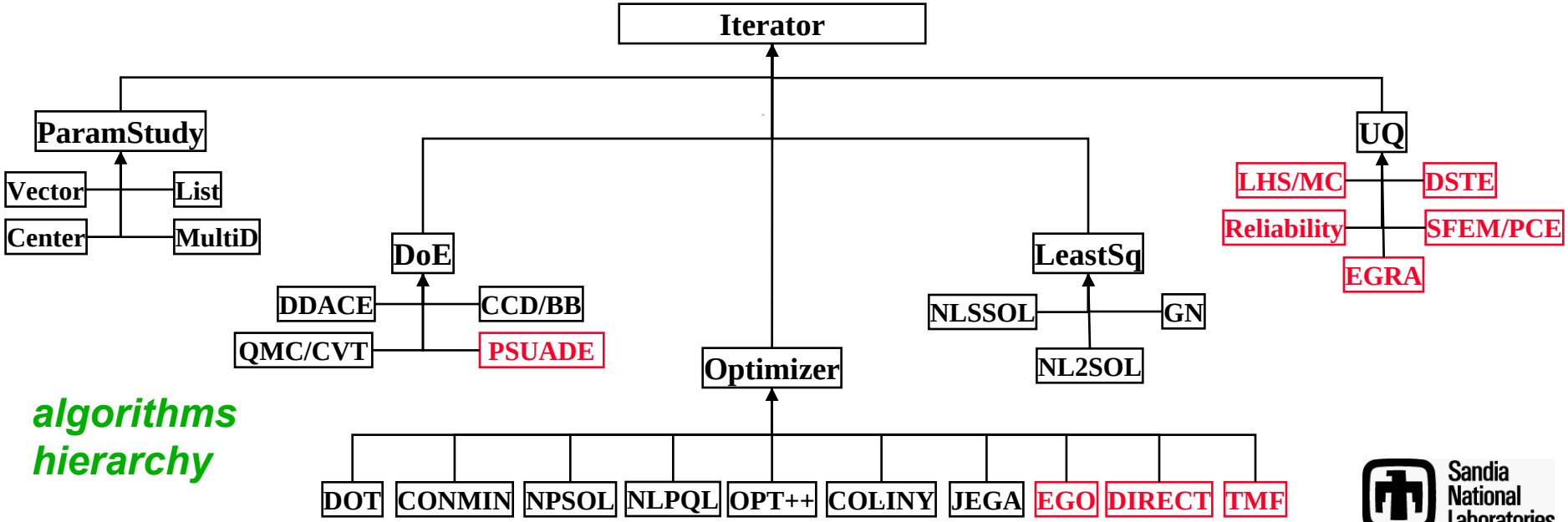
response metrics

Computational Model (simulation)
- **Black box:** any code: mechanics, circuits, high energy physics, biology, chemistry
- **Semi-intrusive:** Matlab, ModelCenter, Python SIERRA multi-physics, SALINAS, Xyce

Sandia National Laboratories

DAKOTA C++/OO Framework Goals

- **Unified software infrastructure:** reuse tools and common interfaces; *integrate commercial, open-source, and research algorithms*
- **Enable algorithm R&D**, e.g., for non-smooth/discontinuous/multimodal responses, probabilistic analysis and design, mixed variables, unreliable gradients, costly simulation failures
- **Facilitate scalable parallelism:** ASCI-scale applications and architectures; *4 nested levels of parallelism possible*
- **Impact:** tool for DOE labs and external partners; broad application deployment; *free via GNU GPL* (>3000 download registrations)



algorithms hierarchy

Sandia National Laboratories

Flexibility with Models & Strategies

DAKOTA models map inputs to response metrics of interest:

variables/parameters

- **design: continuous, discrete**
- **uncertain: (log)normal, (log)uniform, interval, triangular, histogram, beta/gamma, EV I, II, III**
- **state: continuous, discrete**

→

user application
(simulation)
system, fork, direct, grid

optional approximation (surrogate)

- **global (polynomial 1/2/3, neural net, kriging, MARS, RBF)**
- **local (Taylor); multipoint (TANA/3)**
- **hierarchical, multi-fidelity**

→

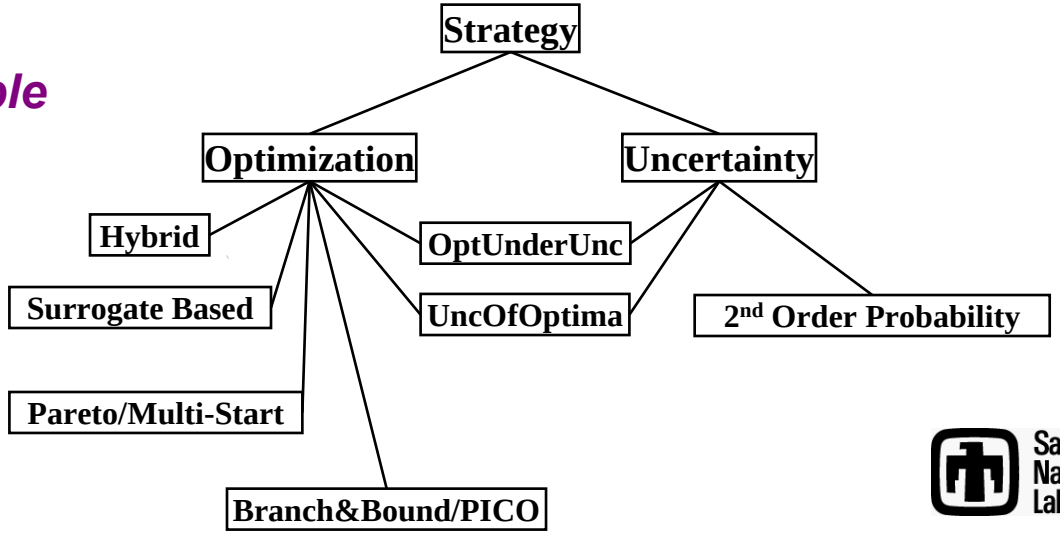
responses

- **functions: objectives, constraints, LSQ residuals, generic**
- **gradients: numerical, analytic**
- **Hessians: numerical, analytic, quasi**

app-specific: users often need initial help to ID key parameters, script, extract metrics

DAKOTA strategies enable flexible combination of multiple models and algorithms.

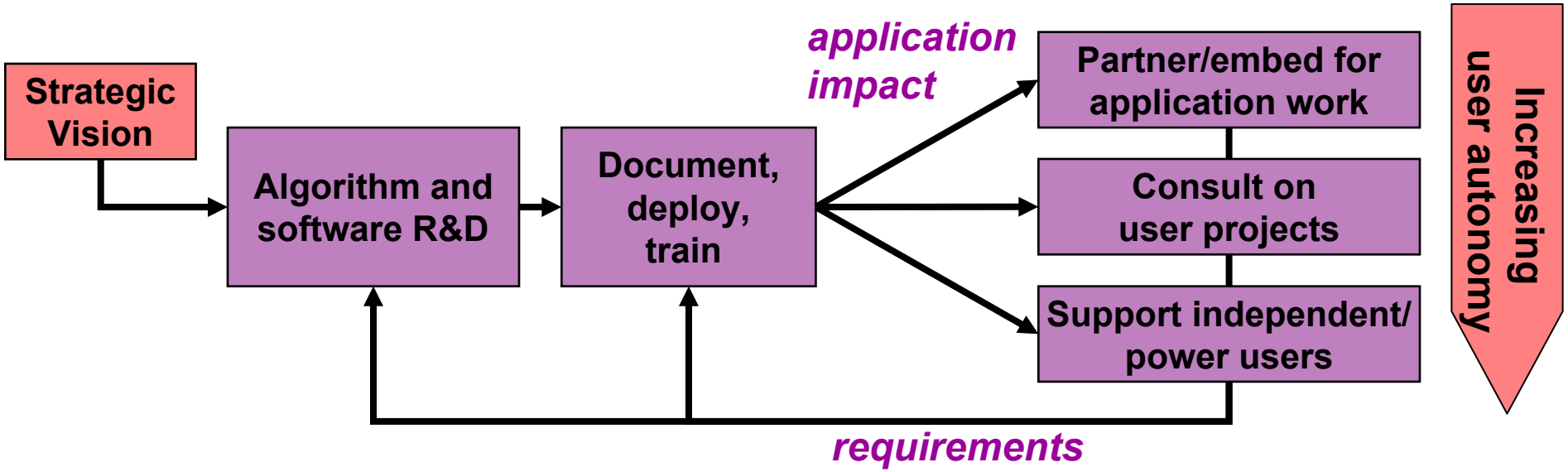
- ***nested***
- ***layered***
- ***cascaded***
- ***concurrent***
- ***adaptive / interactive***



Sandia National Laboratories

DAKOTA: Algorithm R&D to Impact

- *The DAKOTA team partners with system designers and analysts so their projects directly benefit from advanced algorithms and software.*

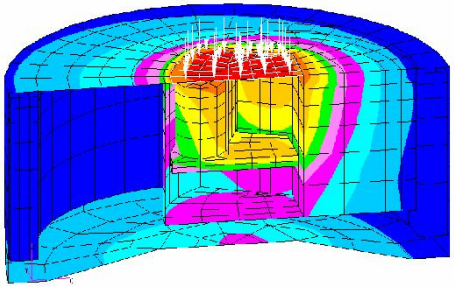


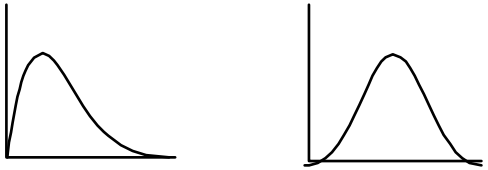
- **I develop algorithms, but also serve as DAKOTA's Product Manager: liaison to application customers.**
- **All modes yield vital feedback on requirements and usability, ultimately impacting algorithm & software research.**

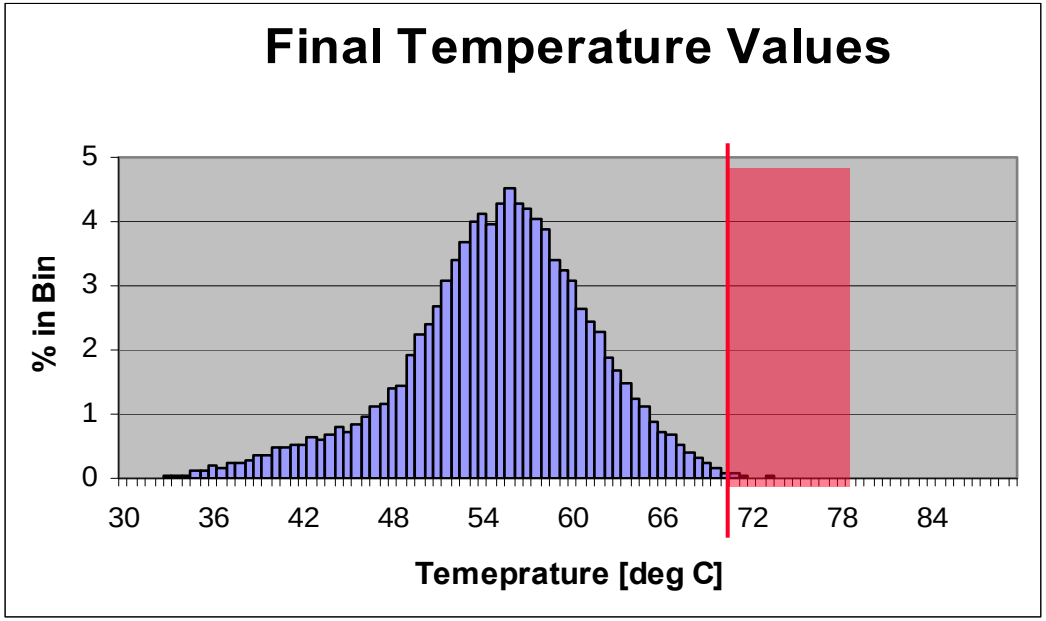
Sandia National Laboratories

Uncertainty Quantification Example

- **Device subject to heating** (experiment or computational simulation)
- Uncertainty in composition/ environment (thermal conductivity, density, boundary), parameterized by $u_1, \ldots, u_N$
- Response temperature $T(u_1, \ldots, u_N)$ calculated by heat transfer code







Given distributions of $u_1,\ldots,u_N$, UQ methods calculate statistical info on outputs:

- Probability distribution of temperatures
- Correlations (trends) and sensitivity of temperature
- Mean(T), StdDev(T), Probability($T \geq T_{critical}$)

Sandia National Laboratories

Uncertainty Quantification

Forward propagation: quantify the effect that uncertain (nondeterministic) input variables have on model output

Input Variables *u*
(physics parameters, geometry, initial and boundary conditions)

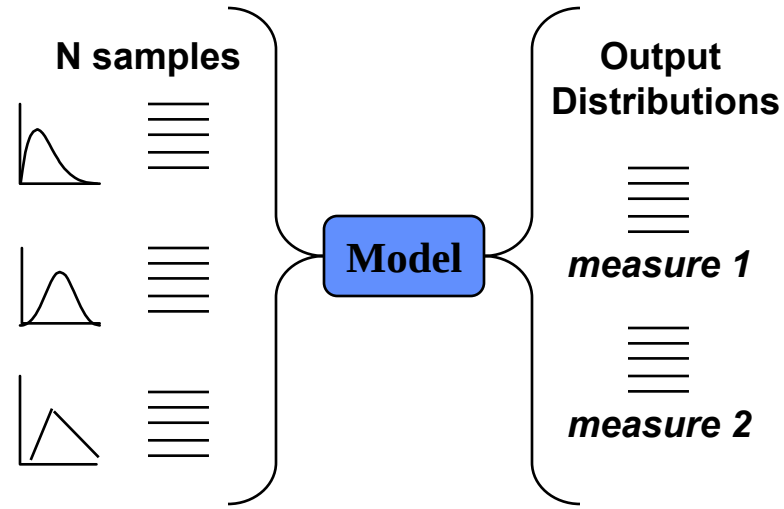
→ **Computational Model** →

Variable Performance Measures *G(u)*

(possibly given distributions)

Potential Goals:

- **based on uncertain inputs, determine variance of outputs and probabilities of failure (reliability metrics)**
- **identify parameter correlations/local sensitivities, robust optima**
- **identify inputs whose variances contribute most to output variance (global sensitivity analysis)**
- **quantify uncertainty when using calibrated model to *predict***

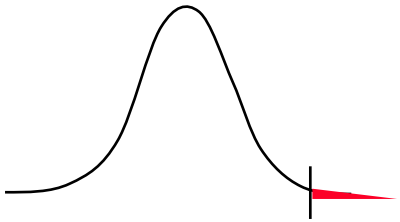


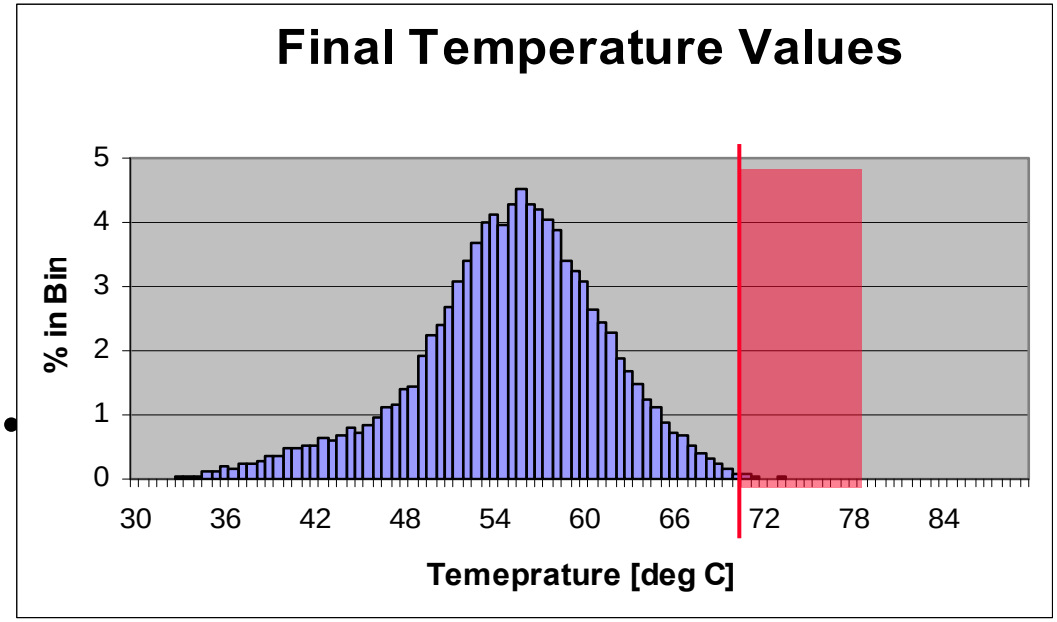
Typical method: Monte Carlo Sampling

Sandia National Laboratories

Calculating Probability of Failure

- ***Given uncertainty in materials, geometry, and environment, determine likelihood of failure Probability($T \geq T_{critical}$)***





- **Could perform 10,000 Monte Carlo samples and count how many exceed the threshold…**

- **Or directly determine input variables which give rise to failure behaviors by solving an optimization problem.**

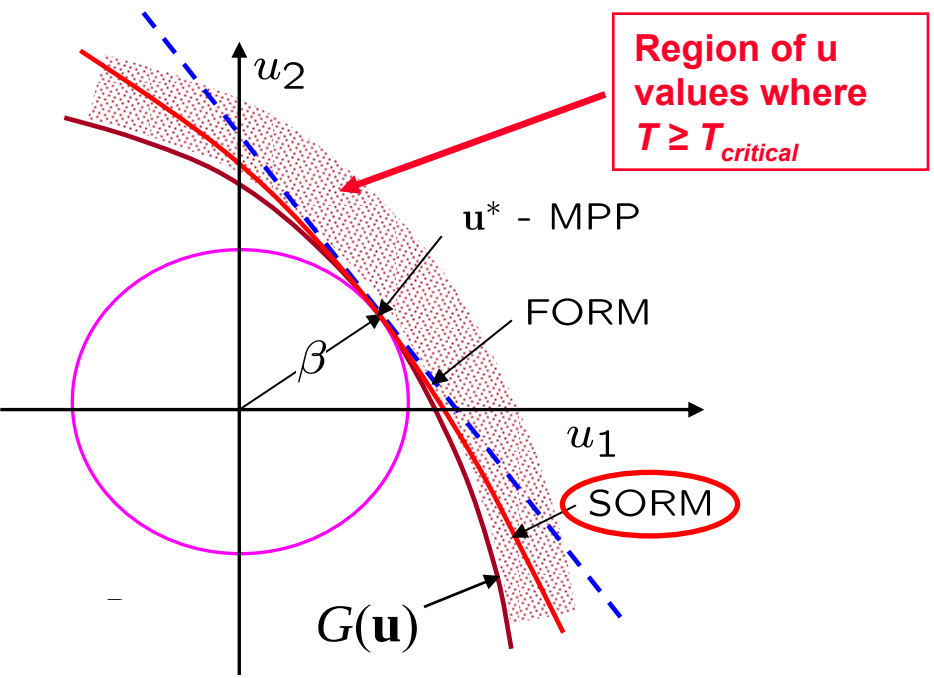
Sandia National Laboratories

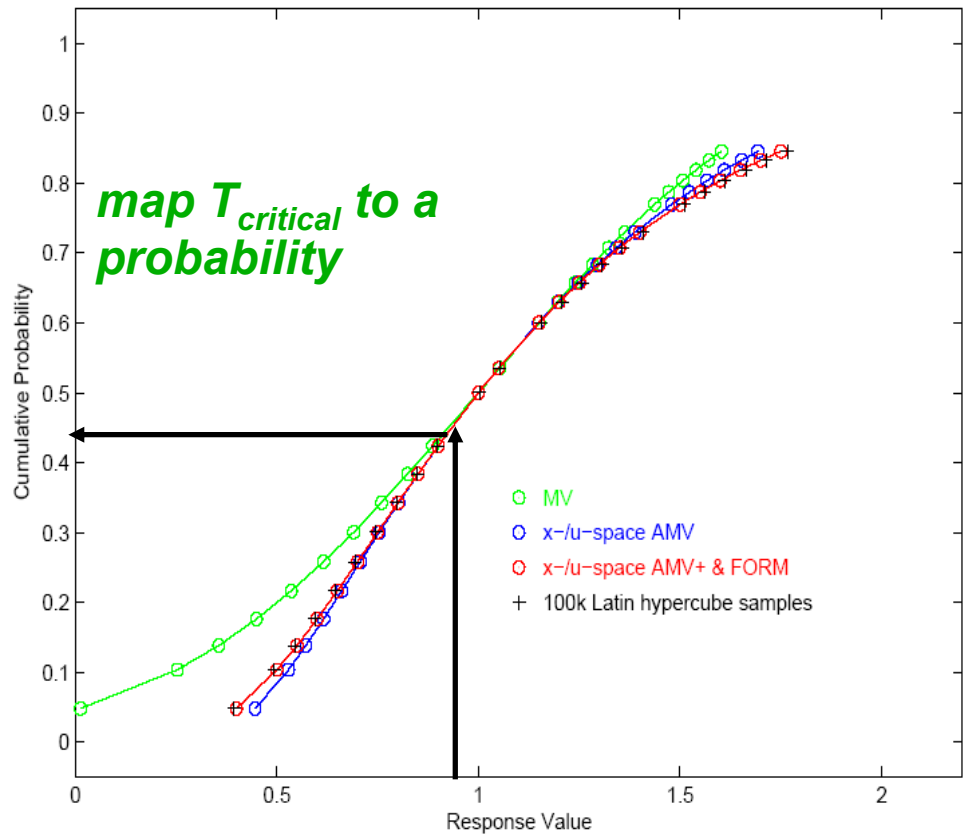
Analytic Reliability: MPP Search

Perform optimization **in uncertain variable space to determine Most Probable Point (of response or failure occurring) for *G(u) = T(u)*.**

Reliability Index Approach (RIA)

$$\text{minimize} \quad \mathbf{u}^T\mathbf{u}$$
$$\text{subject to} \quad G(\mathbf{u}) = \bar{z}$$





Reliability: Algorithmic Variations

Many variations possible to improve efficiency, including in DAKOTA…

- **<u>Limit state linearizations</u>: use a local surrogate for the limit state *G(u)* during optimization in u-space (or x-space):**

$$\text{u-space AMV:} \quad G(\mathbf{u}) = G(\mu_\mathbf{u}) + \nabla_u G(\mu_\mathbf{u})^T(\mathbf{u} - \mu_\mathbf{u})$$

$$\text{u-space AMV+:} \quad G(\mathbf{u}) = G(\mathbf{u}^*) + \nabla_u G(\mathbf{u}^*)^T(\mathbf{u} - \mathbf{u}^*)$$

$$\boxed{\text{u-space AMV}^2\text{+:} \quad G(\mathbf{u}) = G(\mathbf{u}^*) + \nabla_u G(\mathbf{u}^*)^T(\mathbf{u} - \mathbf{u}^*) + \tfrac{1}{2}(\mathbf{u} - \mathbf{u}^*)^T \nabla_u^2 G(\mathbf{u}^*)(\mathbf{u} - \mathbf{u}^*)}$$

(could use analytic, finite difference, or quasi-Newton (BFGS, SR1) Hessians in approximation/optimization – results here mostly use SR1 quasi-Hessians.)

- **<u>Integrations (in u-space to determine probabilities)</u>: may need higher order for nonlinear limit states**

$$1^{\text{st}}\text{-order:} \left\{ \begin{array}{lcl} p(g \le z) & = & \Phi(-\beta_{cdf}) \\ p(g > z) & = & \Phi(-\beta_{ccdf}) \end{array} \right. \qquad 2^{\text{nd}}\text{-order:} \left\{ p = \Phi(-\beta) \underbrace{\prod_{i=1}^{n-1} \frac{1}{\sqrt{1 + \beta\kappa_i}}}_{\text{curvature correction}} \right.$$

- **MPP search algorithm:** Sequential Quadratic Prog. (SQP) vs. Nonlinear Interior Point (NIP)
- **Warm starting (for linearizations, initial iterate for MPP searches):** speeds convergence when increments made in: approximation, statistics requested, design variables

Sandia National Laboratories

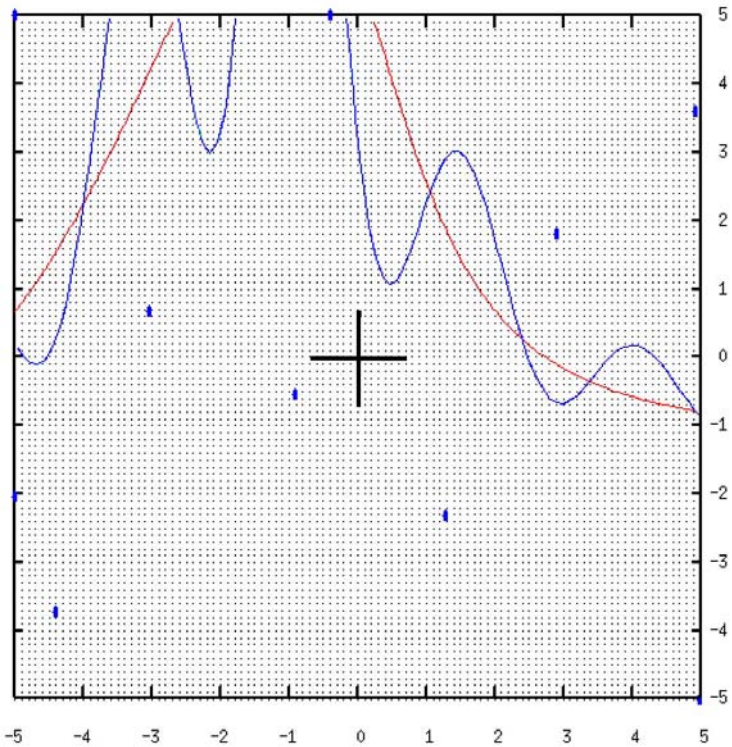
Efficient Global Reliability Analysis

- **EGRA (B.J. Bichon) performs reliability analysis with EGO (Gaussian Process surrogate and NCSU DIRECT optimizer) coupled with Multimodal adaptive importance sampling for probability calculation.**

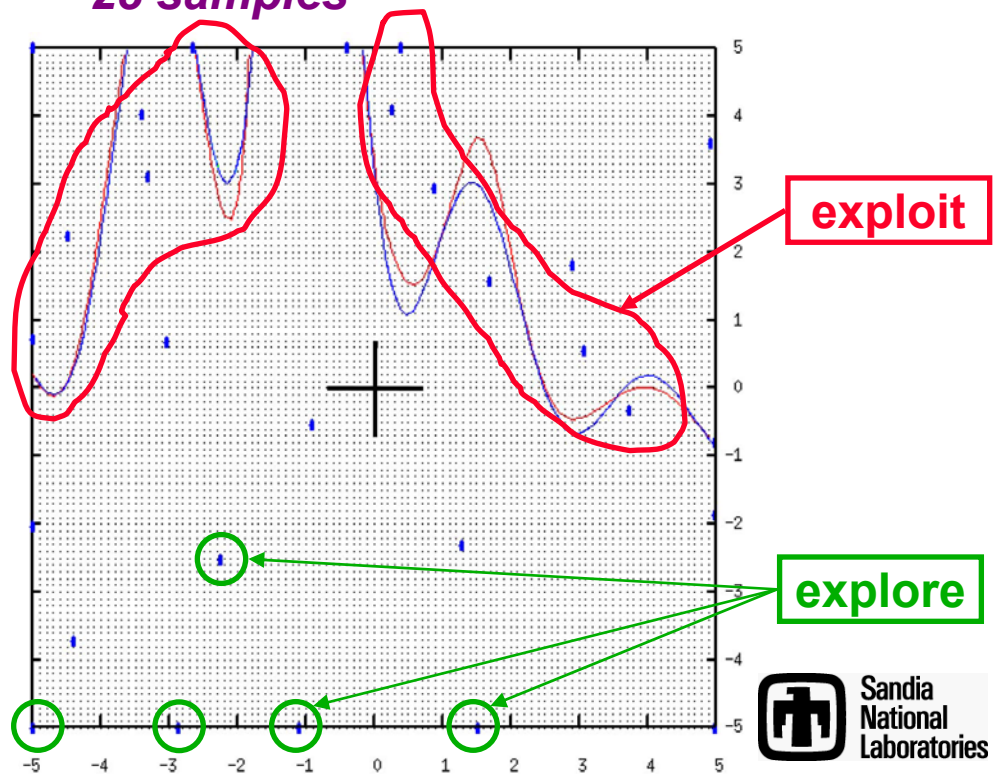
- **Created to address nonlinear and/or multi-modal limit states in MPP searches.**

Gaussian process model of reliability limit state with

10 samples



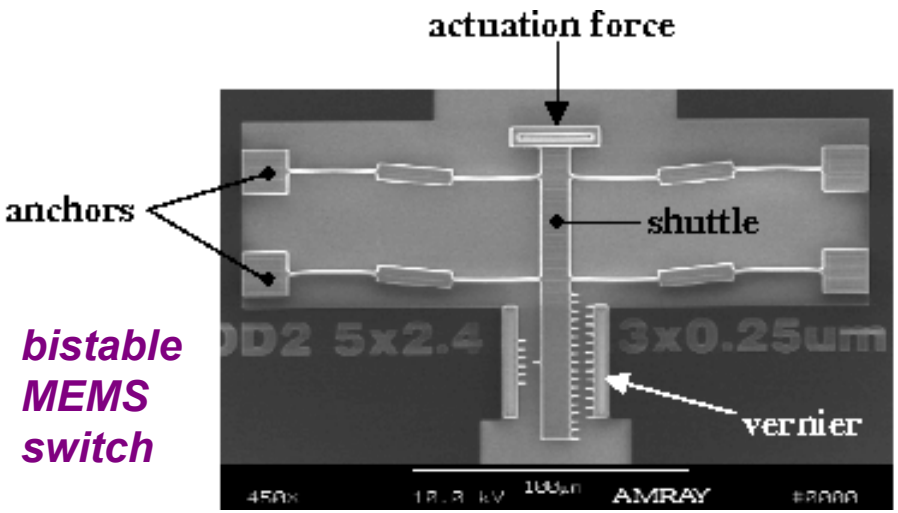
28 samples



Sandia National Laboratories

Shape Optimization of Compliant MEMS

- **Micro-electromechanical system (MEMS):** typically made from silicon, polymers, or metals; used as micro-scale sensors, actuators, switches, and machines
- **MEMS designs are subject to substantial variability** and lack historical knowledge base. Materials and micromachining, photo lithography, etching processes all yield uncertainty.
- Resulting part yields can be low or have poor cycle durability
- **Goal: shape optimize finite element model of bistable switch to…**
  - **Achieve prescribed reliability** in actuation force
  - Minimize sensitivity to uncertainties **(robustness)**



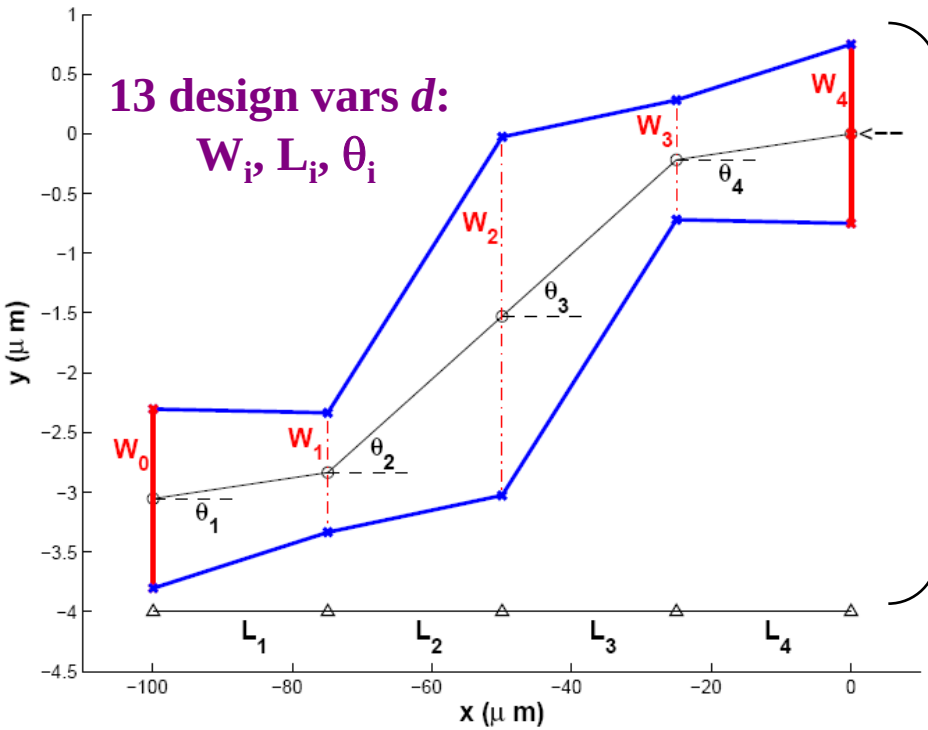
bistable MEMS switch

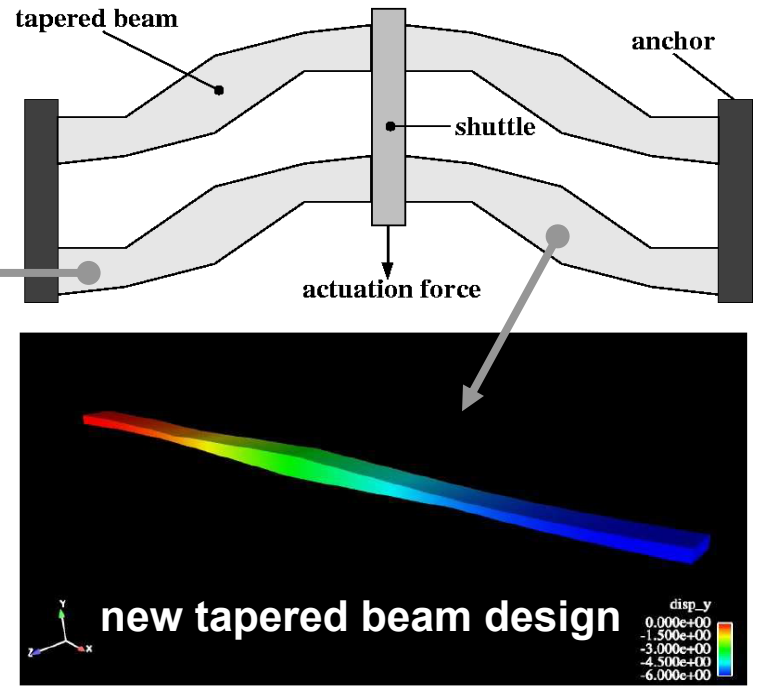
uncertainties to be considered (edge bias and residual stress)

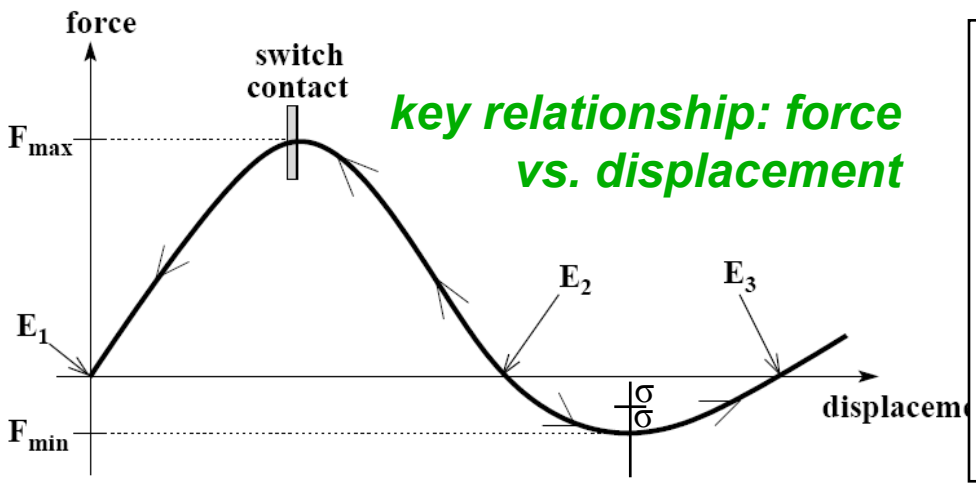
variable	mean	std. dev.	distribution
Δw	-0.2 μm	0.08	normal
S_r	-11 Mpa	4.13	normal

Sandia National Laboratories

Tapered Beam Bistable Switch: Performance Metrics







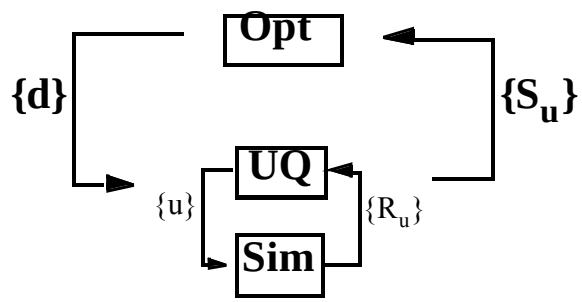
key relationship: force vs. displacement

Typical design specifications:

- **actuation force F_{min} reliably 5 μN**
- **bistable ($F_{max} > 0$, $F_{min} < 0$)**
- **maximum force: $50 < F_{max} < 150$**
- **equilibrium E2 < 8 μm**
- **maximum stress < 1200 MPa**

Optimization Under Uncertainty

Rather than design and then post-process to evaluate uncertainty…
actively design optimize while accounting for uncertainty/reliability metrics $s_u(d)$, e.g., mean, variance, reliability, probability:



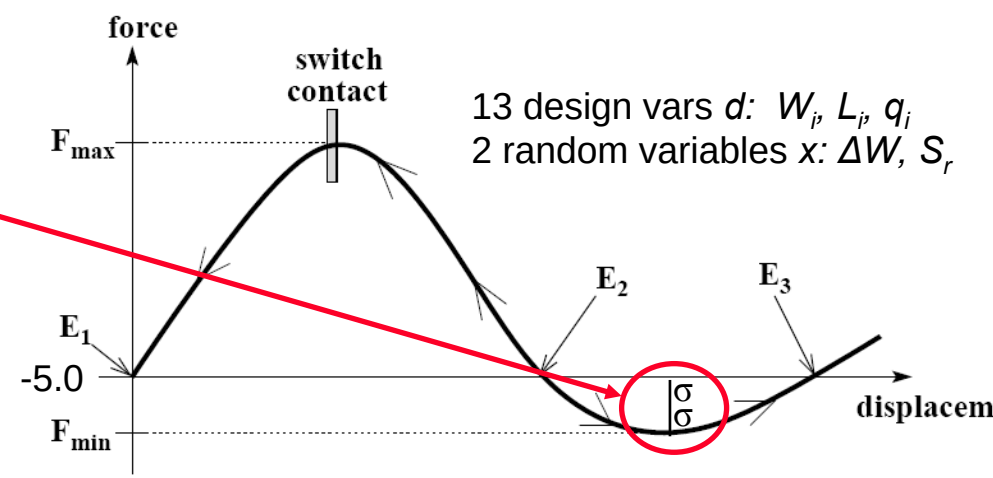
(nested paradigm)

$$
\begin{aligned}
\min \quad & f(d) + W s_u(d) \\
\text{s.t.} \quad & g_l \le g(d) \le g_u \\
& h(d) = h_t \\
& d_l \le d \le d_u \\
& a_l \le A_i\, s_u(d) \le a_u \\
& A_e\, s_u(d) = a_t
\end{aligned}
$$

Bistable switch problem formulation (Reliability-Based Design Optimization):

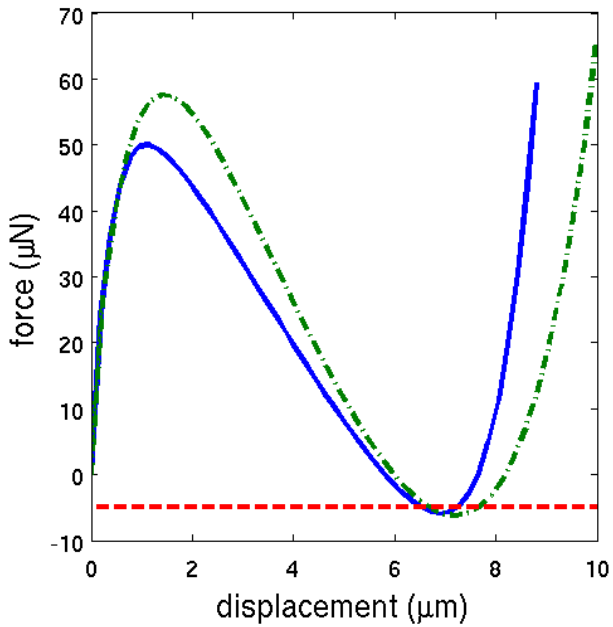
simultaneously reliable and robust designs

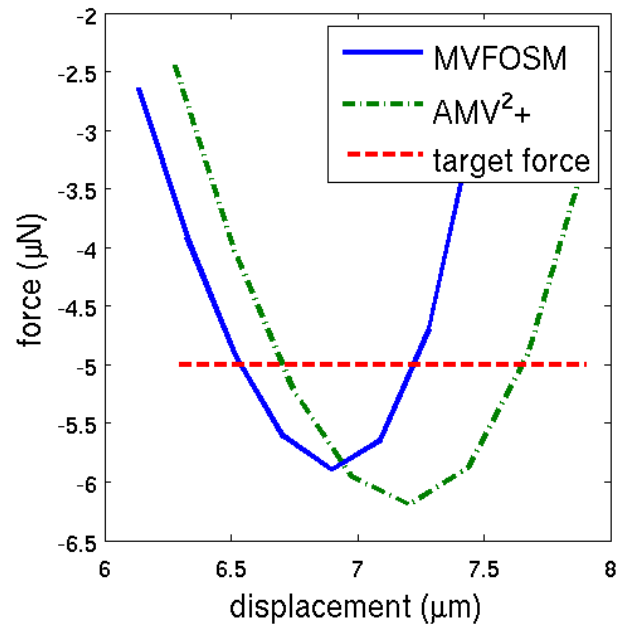
$$
\begin{array}{lrcl}
\max & & & \mathrm{E}\left[F_{min}(\mathbf{d},\mathbf{x})\right] \\
\text{s.t.} & 2 & \le & \beta_{ccdf}(\mathbf{d}) \\
& 50 & \le & \mathrm{E}\left[F_{max}(\mathbf{d},\mathbf{x})\right] \le 150 \\
& & & \mathrm{E}\left[E_2(\mathbf{d},\mathbf{x})\right] \le 8 \\
& & & \mathrm{E}\left[S_{max}(\mathbf{d},\mathbf{x})\right] \le 3000
\end{array}
$$



13 design vars *d:* W_i, L_i, q_i
2 random variables *x:* ΔW, S_r

RBDO Finds Optimal & Robust Design





Close-coupled results: DIRECT / CONMIN + reliability method yield optimal and reliable/robust design:

metric				MVFOSM	AMV^2+	FORM
l.b.	name	u.b.	initial $\mathbf{d}^0$	optimal $\mathbf{d}^*_M$	optimal $\mathbf{d}^*_A$	optimal $\mathbf{d}^*_F$
	E $[F_{min}]$ (μN)		-26.29	-5.896	-6.188	-6.292
2	β		5.376	2.000	1.998	1.999
50	E $[F_{max}]$ (μN)	150	68.69	50.01	57.67	57.33
	E $[E_2]$ (μm)	8	4.010	5.804	5.990	6.008
	E $[S_{max}]$ (MPa)	1200	470	1563	1333	1329
	AMV^2+ verified β		3.771	1.804	-	-
	FORM verified β		3.771	1.707	1.784	-

Sandia National Laboratories

Summary

- Sandia's broad mission areas span many disciplines
- Computational research spans
  - hardware, operating systems, theory, algorithms, software, modeling applications, uncertainty analysis…
  - ultimately to risk-informed decision making
- Pure/abstract and practical/application-driven research
- Collaborations with academia, industry, and government

Thank you for your attention!

briadam@sandia.gov
http://www.sandia.gov/~briadam

Sandia National Laboratories