Department Review:

Invasive Gradient-Based Optimization

Improving Speed, Accuracy, and Capability

Roscoe A. Bartlett

Department 9211: Optimization and Uncertainty Estimation

Sandia National Laboratories, 2005-3650P

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

Sandia
National
Laboratories

Overview of Roscoe Bartlett (9211)

My Focus: Foster the development and use of fast, invasive, gradient-based optimization algorithms for large-scale and massively parallel Sandia applications

- Why optimization is important
  - Inversion / Parameter estimation
  - Design
- Why intrusive optimization is important
  - Speed: Faster times to solution
  - Accuracy: More accurate solutions
  - Reliability: More reliable solvers
  - Capability: Handing of more complex constraints
- Technologies for invasive optimization for which I am involved
  - Massively parallel, invasive, gradient-based optimization algorithms: MOOCHO (Trilinos?)
  - Transient sensitivities for optimization: Rythmos (Trilinos)
    - Collaborators: Todd Coffey (9214), Curt Ober (9233)
  - Accurate, inexpensive model derivatives: Automatic Differentiation
    - Collaborators: Eric Phipps (9233), David Gay (9211)
  - Interoperability of numerical software: Thyra (Trilinos), project leader
    - Collaborators: Mike Heroux (9214), Heidi Thornquist (9214), Roger Pawlowski (9233), Todd Coffey (9214), Rob Hoekstra (9237), Paul Boggs (8962), Kevin Long (8962), Steve Margolis (8962), Allan Williams (9243), Victoria Howle (8962)

Sandia National Laboratories

Overview of Roscoe Bartlett (9211)

Highlights for 2004-2005

- Became project leader for the new Thyra consortium and Trilinos package
  - Collaborators: Mike Heroux (9214), Heidi Thornquist (9214), Roger Pawlowski (9216), Todd Coffey (9214), Rob Hoekstra (9237), Paul Boggs (8962), Kevin Long (8962), Steve Margolis (8962), Allan Williams (9243), Victoria Howle (8962)
  - Operator/vector interfaces finished
- Real-time transient source inversion for internal spaces: MPSalsa/MOOCHO
  - Collaborators: Andy Salinger (9216), John Shadid (9237), Bart van Bloemen Waanders (9211)
- Started work on new Trilinos time integration and sensitivity package Rythmos (PI Todd Coffey (9214))
- Completion of sensitivity LDRD and final report (SAND2004-6574)
- Rewrote QPSchur Journal Paper and got it accepted
  - R. Bartlett and L. Biegler, “QPSchur: A Dual, Active-Set, Schur-Complement Method for Large-Scale and Structured Convex Quadratic Programming”, accepted to *Engineering and Optimization*

Transient Source Inversion for Airport Terminal Model

Protection of Internal Facilities

Goal: Develop numerical methods for the detection and remediation of chemical, biological and radiological (CBR) releases (e.g. terrorist attack).

Collaborators:

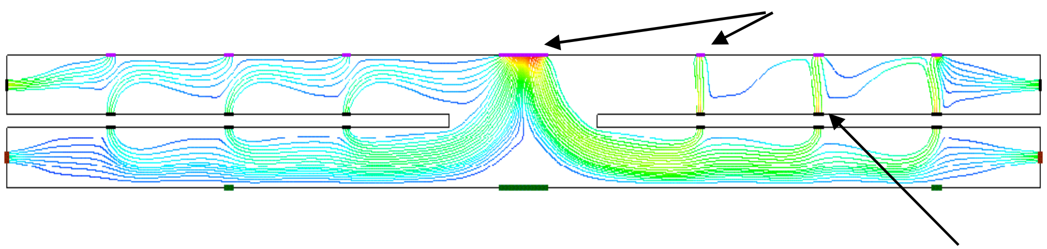
- Andy Salinger
- John Shadid
- Bart van Bloemen Waanders
- **Roscoe A. Bartlett**

Approaches

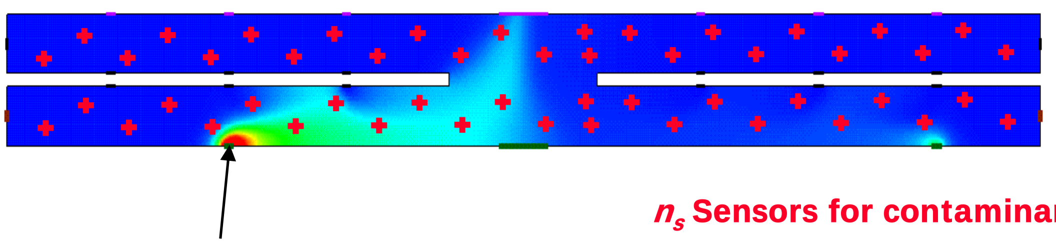
- Invert for source location of CBR release
- Remediation strategies

Transient Source Inversion for Airport Terminal Model

2D Cross Section of Airport Terminal Model



Contaminant Concentration (Forward Problem Solution)



Transient* Inversion Problem

- Given transient snapshots of sensor readings y_i^*, solve an optimization problem to invert for the transient source location and intensity of u along the floors!

Requirements

- Repeated online inversion of **transient** model (i.e. seconds!)

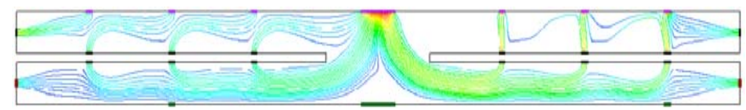
Real-time!

*** Steady-state inversion problem was completed in 2004**

Sandia National Laboratories

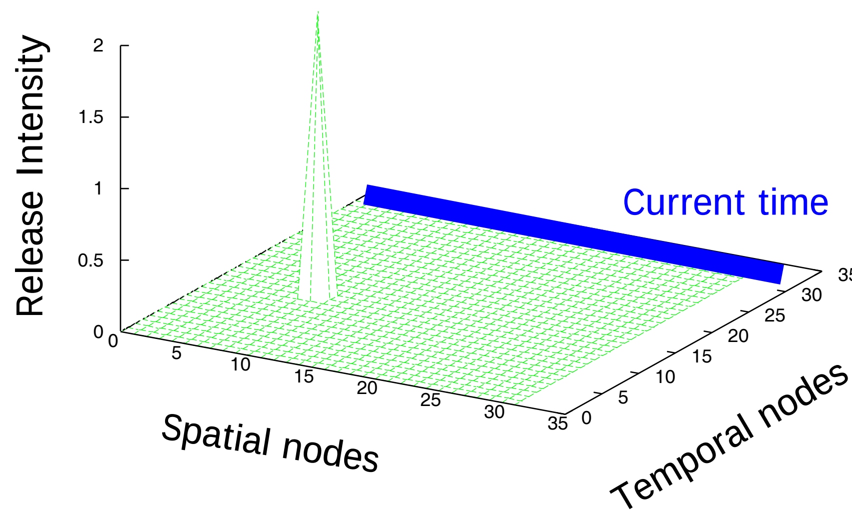
Overview of Problem Formulation

2D Transient Airport Model :
- Discretization of the advection-diffusion equations using MPSalsa
- Steady-state flow field precomputed
- 55,000 state ODE variables
- 128 time steps (0.5 sec/time step)
- 5 min simulation time on one processor



Discretization of the Source
- Piece-wise linear “hat” functions in space and time
- Limited to floor of facility
- 32 spatial x 32 temporal = 1024 total inversion parameters

Example of impulse source



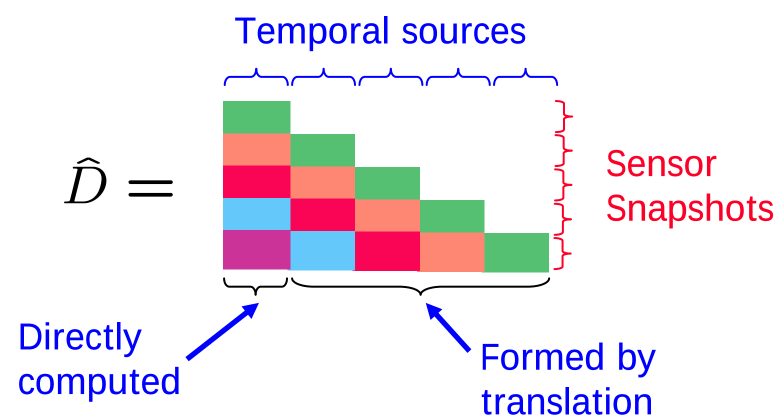
Final Optimization Problem
- Quadratic program
  - Linear PDE constraints
  - Quadratic least-squares objective function

Standard black-box optimization approach?
- (5 min)(1024 simulations/iteration)
  = 3.5 days / iteration!

Sandia National Laboratories

Offline / Online Decomposition : Key to Real time performance!

- Offline: Compute spatial sensitivities and translate in time to eliminate transient state variables online
  - Sensitivities of state concentrations at sensor locations at w.r.t. each inversion parameter
  - Can be computed in parallel using MPP



- Online: Form reduced QP subproblem (compute reduced Hessian)
  - Computed on an SMP

$$G = \widehat{D}^T \widehat{Q}_y \widehat{D} + Q_u$$

Reduced Hessian

$$g = -\widehat{D}^T \widehat{Q}_y \widehat{y}$$

Reduced gradient

- Online: Solve bound-constrained reduced QP subproblem
  - QP solved using QPSchur* (Bartlett 2005)
  - Solved on an SMP

$$\begin{array}{ll} \text{min} & g^T u + \tfrac{1}{2} u^T G u \\ \text{s.t.} & u \geq 0 \end{array}$$

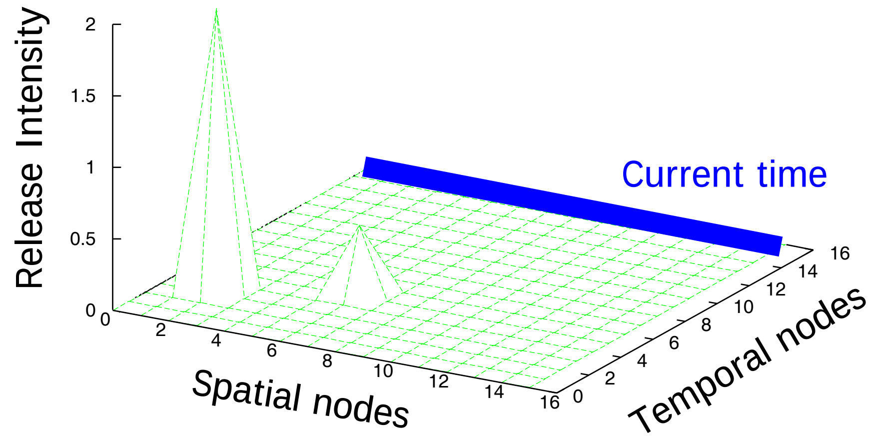
* R. Bartlett and L. Biegler, "QPSchur: A Dual, Active-Set, Schur-Complement Method for Large-Scale and Structured Convex Quadratic Programming", accepted to *Engineering and Optimization*

Sandia National Laboratories

Transient Source Inversion for Airport Terminal Model

Example of two impulse sources

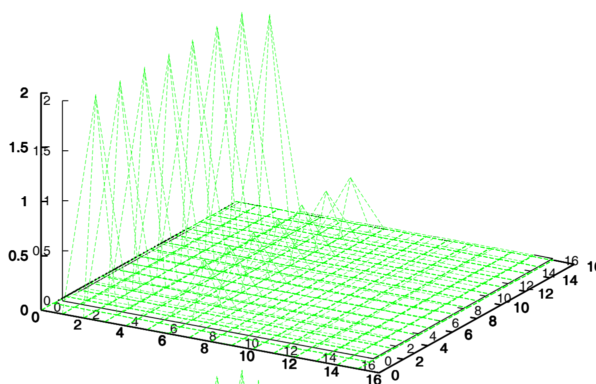
- 16 spatial x 16 temporal : $n_u = 256$ total inversion parameters
- Sensor readings taken every 8 seconds



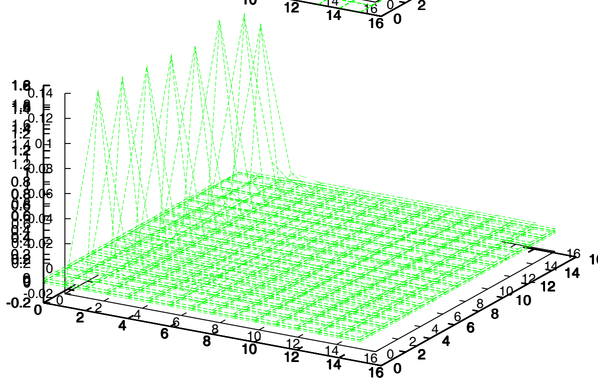
Inversion snapshots

- 16 randomly placed spatial sensors

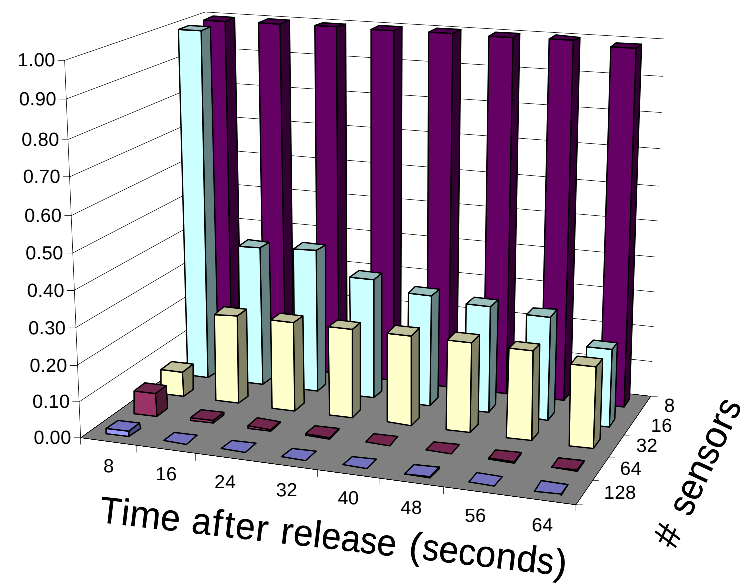
True source
*u**



Inverted source
u



Relative Error (between u* and u)



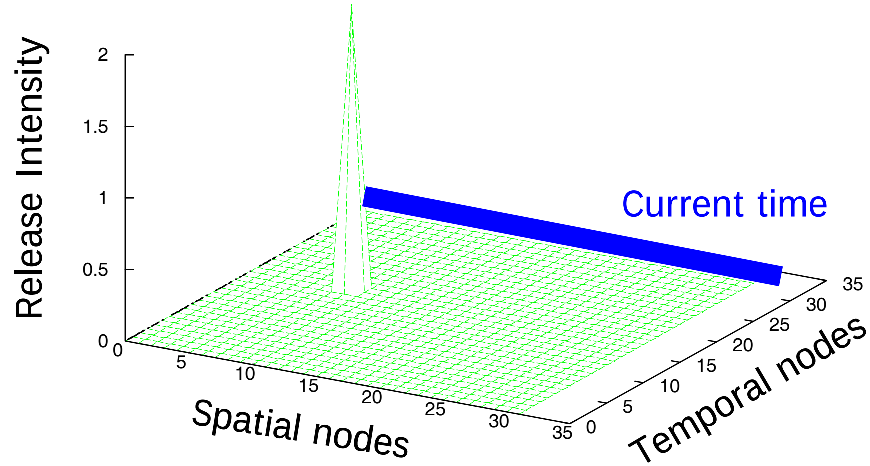
All CPU times < 1.0 sec

Sandia National Laboratories

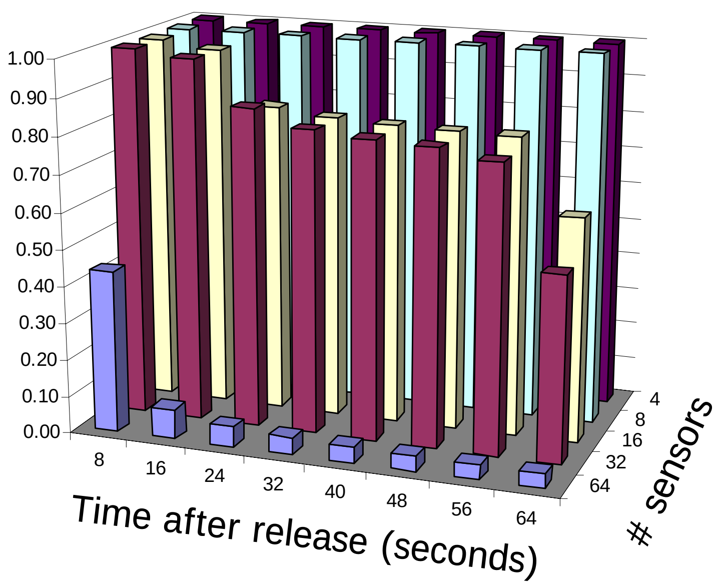
Transient Source Inversion for Airport Terminal Model

Example of single impulse source

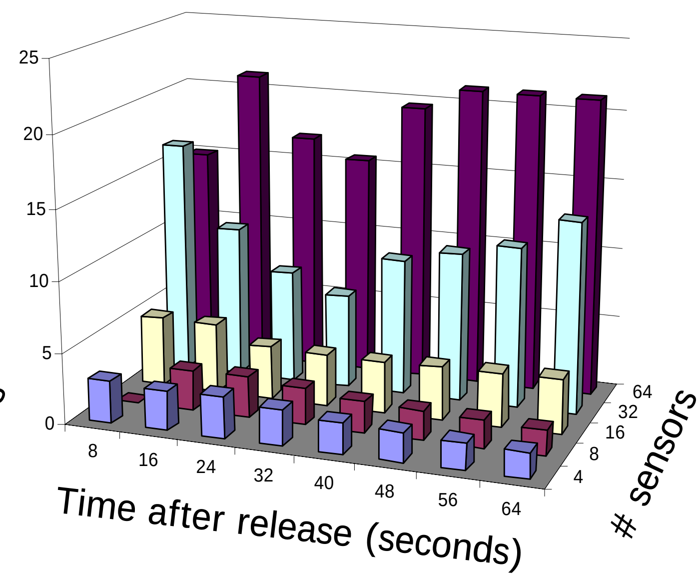
- 32 spatial x 32 temporal : $n_u = 1024$ total inversion parameters
- Sensor readings taken every 8 seconds



Relative Error



Total CPU time (in seconds)



20 Seconds may not be “real-time”

CPU Breakdown

Computation	Seconds
QP Preprocessing	1.10
QP Solve	20.00
Total	21.09

QP Solve is dominate time!

- Bound-constrained reduced QP solved to near optimality!
- CPU times on a 3.0 GHz Linux Box

Sandia National Laboratories

Speeding up the Online Solve : Inexact QP Solves

Bound-constrained reduced QP subproblem

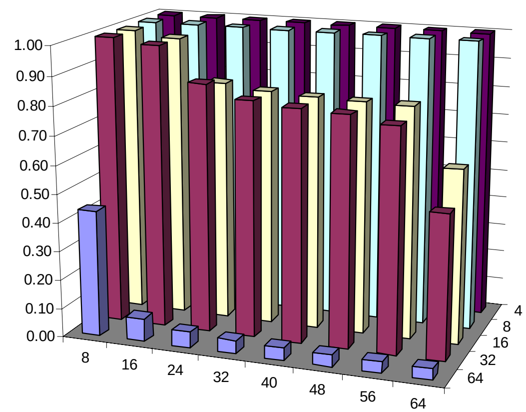
$$\min \quad g^T u + \tfrac{1}{2} u^T G u$$

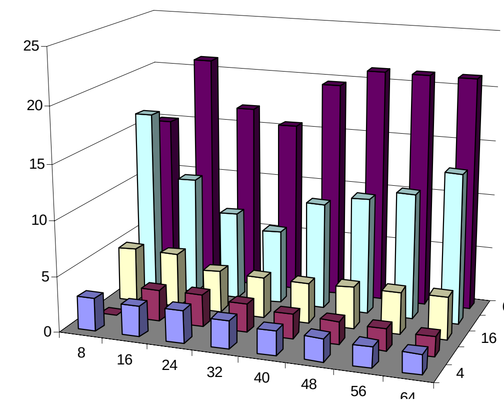
$$\text{s.t.} \quad u \geq 0$$

What if these bounds are not strictly satisfied?

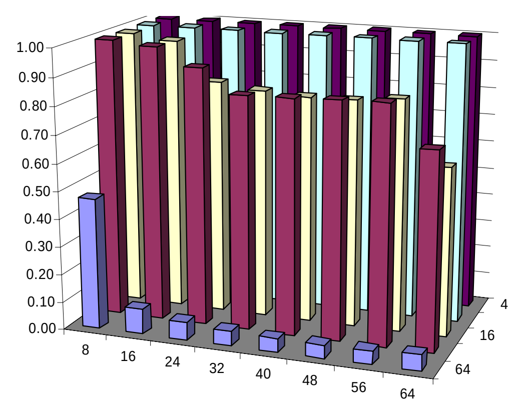
Relative Error | **Total CPU time (in seconds)**

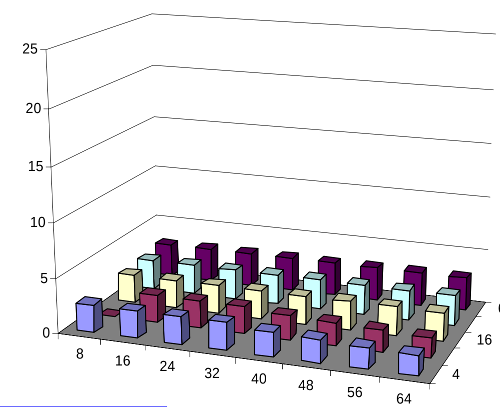
20 second QP Solve Solve Limit





2 second QP Solve Solve Limit





Key Point: Inexact QP solves greatly improve performance without much damage to inversion quality

Sandia National Laboratories

Transient Source Inversion : Summary

- Solved very large transient optimization problem in real time
  - $O(10^6)$ total state variables
  - $O(10^3)$ optimization variables
  - Offline/Online decomposition
    - Offline computations can be performed on MPP => Sandia Computers?
    - Online solution time in 3 seconds on serial computer => Cheap for facilities!
- Approach is also applicable to other inversion or least-squares problems which have many linear constraints!

Question: How do we apply optimization technology to Sandia applications?

Answer: Trilinos!

Trilinos Strategic Goals

- **Scalable Solvers**: As problem size and processor counts increase, the cost of the solver will remain a nearly fixed percentage of the total solution time.
- **Hardened Solvers**: Never fail unless problem essentially unsolvable, in which case we diagnose and inform the user why the problem fails and provide a reliable measure of error.
- **Full Vertical Coverage**: Provide leading edge capabilities from basic linear algebra to transient and optimization solvers.
- **Universal Interoperability**: All Trilinos packages will be interoperable, so that any combination of solver packages that makes sense algorithmically will be possible within Trilinos.
- **Universal Solver RAS**: Trilinos will be:
  - Integrated into every major application at Sandia (Availability).
  - The leading edge hardened, scalable solutions for each of these applications (Reliability).
  - Easy to maintain and upgrade within the application environment (Serviceability).

Thyra is being developed to address this issue (PI Roscoe Bartlett)

Question: Why am I the project leader for Thyra?

Courtesy of Mike Heroux, Trilinos Project Leader

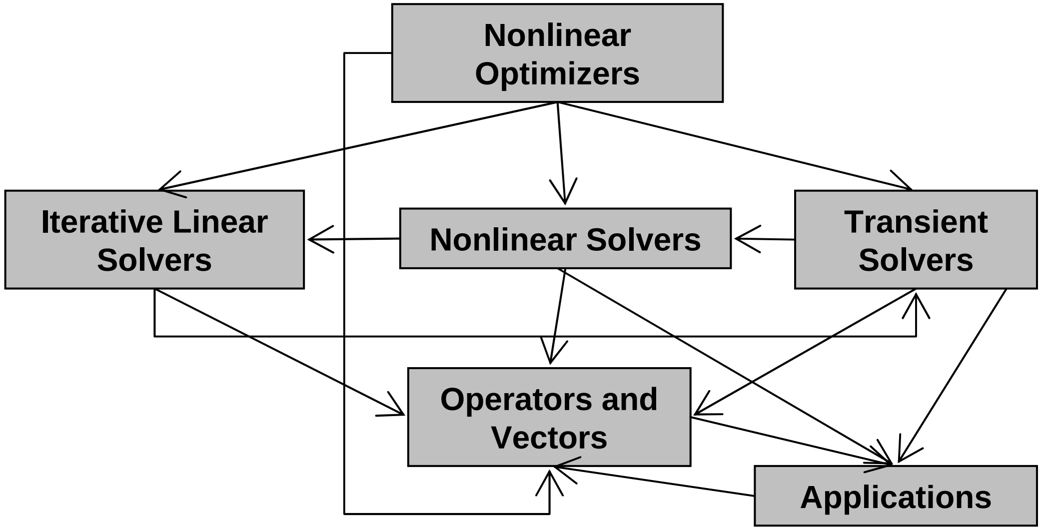
Key Point

- Universal Interoperability will not happen automatically and if not done carefully then can compromise the other strategic goals

Sandia National Laboratories

Interoperability is Especially Important to Optimization

Numerous interactions exist between abstract numerical algorithms (ANAs) in a transient optimization problem



What is needed to solve problem?

- Standard interfaces to break $O(N^2)$ 1-to-1 couplings
  - Operators/vectors
  - Linear Solvers
  - Nonlinear solvers
  - Transient solvers
  - etc.

This is hard! This level of interoperability for massively parallel algorithms has never been achieved before!

What background/skills are needed?

- Nonlinear constrained optimization
- Transient optimization
- Large-scale numerics
- Software skills
  - Object-oriented design expert
  - C++ expert

Key Points

- Higher level algorithms, like optimization, require a lot of interoperability
- Interoperability must be “easy” or these configurations will not be possible
- Many real problems even more complicated

Thyra : Organization, Scope and Impact

- Organization
  - Collaborators: Mike Heroux (9214), Heidi Thornquist (9214), Roger Pawlowski (9216), Todd Coffey (9214), Rob Hoekstra (9237), Paul Boggs (8962), Kevin Long (8962), Steve Margolis (8962), Allan Williams (9243), Victoria Howle (8962)
  - Project lead: Roscoe Bartlett (9211)
- Scope
  - Operators/vectors, linear solvers, nonlinear solvers, stability analysis, transient solvers and sensitivities, nonlinear optimization, transient optimization etc.
- Current Impact
  - Rythmos : Transient solves and sensitivities (Todd Coffey (9214))
  - CAPO : Picard-Newton methods for nonlinear equations (Andy Salinger (9216))
  - Siesmic inversion : Felix Hermann (UBC) [external]
- Future Impact
  - All of Trilinos
  - SIERRA
  - Nevada

Summary for Roscoe Bartlett

Highlights for 2004-2005

– Became project leader for the new Thyra consortium and Trilinos package

- Collaborators: Mike Heroux (9214), Heidi Thornquist (9214), Roger Pawlowski (9216), Todd Coffey (9214), Rob Hoekstra (9237), Paul Boggs (8962), Kevin Long (8962), Steve Margolis (8962), Allan Williams (9243), Victoria Howle (8962)

- Operator/vector interfaces finished

– Real-time transient source inversion for internal spaces: MPSalsa/MOOCHO

- Collaborators: Andy Salinger (9216), John Shadid (9237), Bart van Bloemen Waanders (9211)

– Started work on new Trilinos time integration and sensitivity package Rythmos (PI Todd Coffey (9214))

– Completion of sensitivity LDRD and final report (SAND2004-6574)

– Rewrote QPSchur Journal Paper and got it accepted

- R. Bartlett and L. Biegler, “QPSchur: A Dual, Active-Set, Schur-Complement Method for Large-Scale and Structured Convex Quadratic Programming”, accepted to *Engineering and Optimization*

Sandia National Laboratories