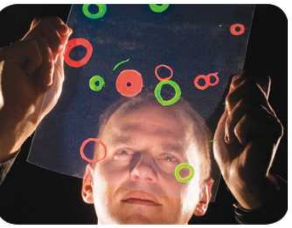
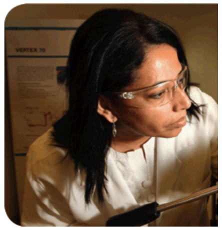
SAND2015-9699C





Analysis of Critical Infrastructures with the Pyomo Modeling Software

William Hart

Sandia National Laboratories

February 11, 2014

Exceptional service in the national interest

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Overview

- About Sandia National Laboratories

- Optimization modeling for critical infrastructures
  - Water security
  - Electrical energy planning and management

- Optimization modeling with Pyomo

DOE’s National Laboratories

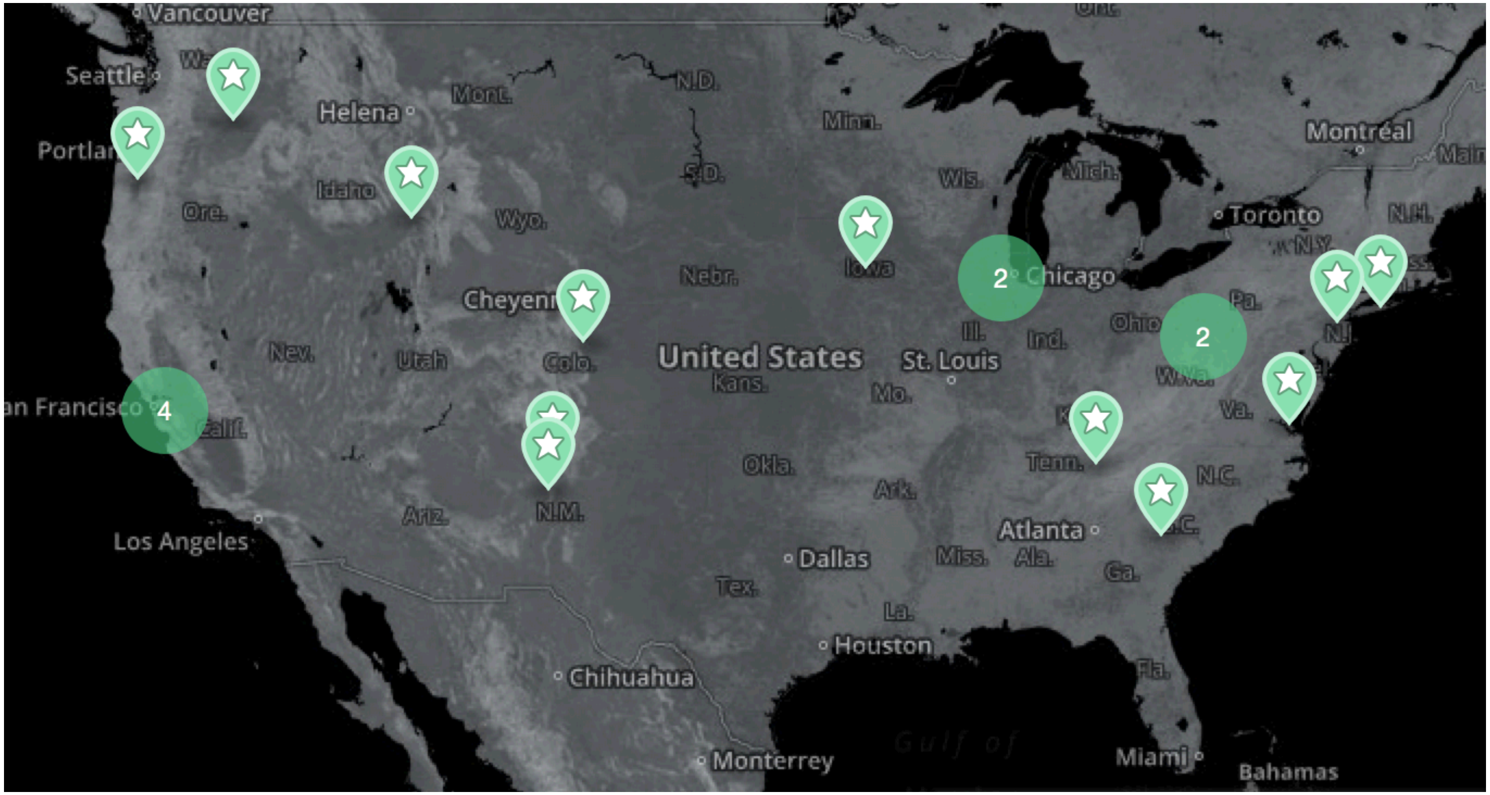
DOE National Labs address …

- large scale, complex research and development challenges …
- with a multidisciplinary approach …
- that places an emphasis on translating basic science to innovation.

The Energy Department's National Labs tackle critical scientific challenges:

- Conduct research of the highest caliber
- Advance U.S. energy independence and leadership
- Enhance global, national, and homeland security
- Design, build, and operate distinctive scientific instrumentation and facilities

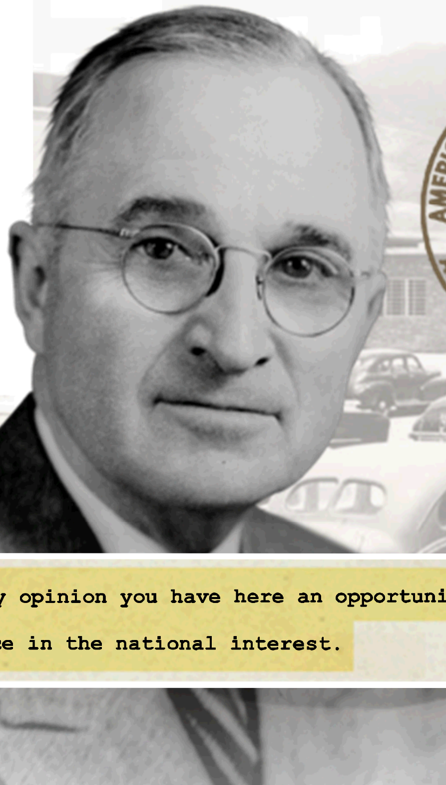
DOE National Laboratories

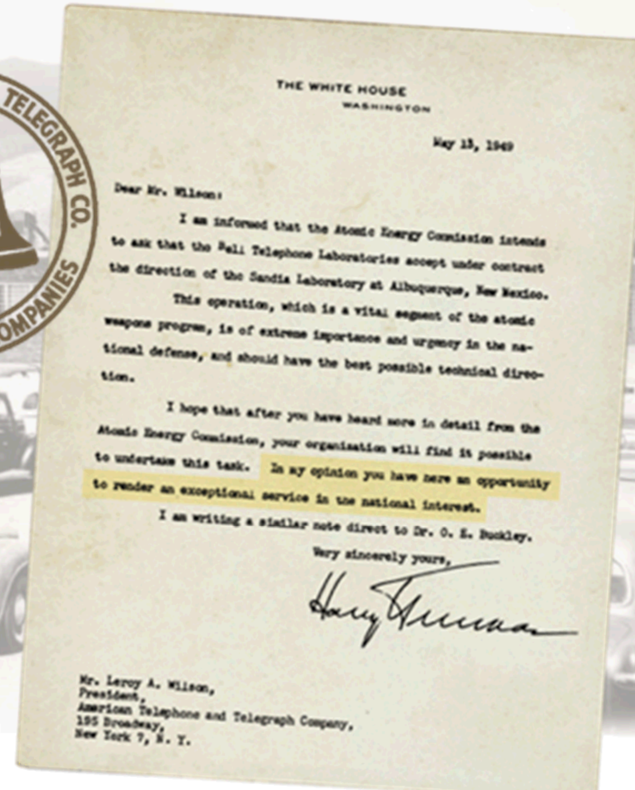


Sandia's History

Exceptional service in the national interest

- July 1945: Los Alamos creates Z Division
- Nonnuclear component engineering
- November 1, 1949: Sandia Laboratory established





THE WHITE HOUSE
WASHINGTON

May 13, 1949

Dear Mr. Wilson:

I am informed that the Atomic Energy Commission intends to ask that the Bell Telephone Laboratories accept under contract the direction of the Sandia Laboratory at Albuquerque, New Mexico.

This operation, which is a vital segment of the atomic weapons program, is of extreme importance and urgency in the national defense, and should have the best possible technical direction.

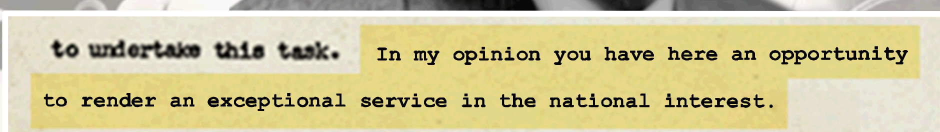
I hope that after you have heard more in detail from the Atomic Energy Commission, your organization will find it possible to undertake this task. In my opinion you have here an opportunity to render an exceptional service in the national interest.

I am writing a similar note direct to Dr. O. E. Buckley.

Very sincerely yours,

Harry Truman

Mr. Leroy A. Wilson,
President,
American Telephone and Telegraph Company,
195 Broadway,
New York 7, N. Y.



to undertake this task. In my opinion you have here an opportunity to render an exceptional service in the national interest.

Sandia Sites

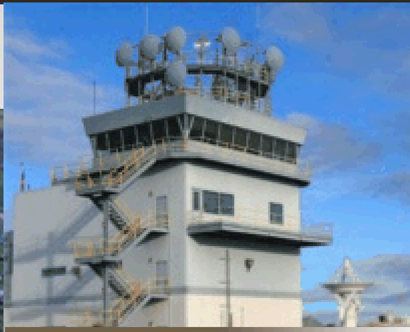
Albuquerque, New Mexico

Livermore, California

Kauai, Hawaii

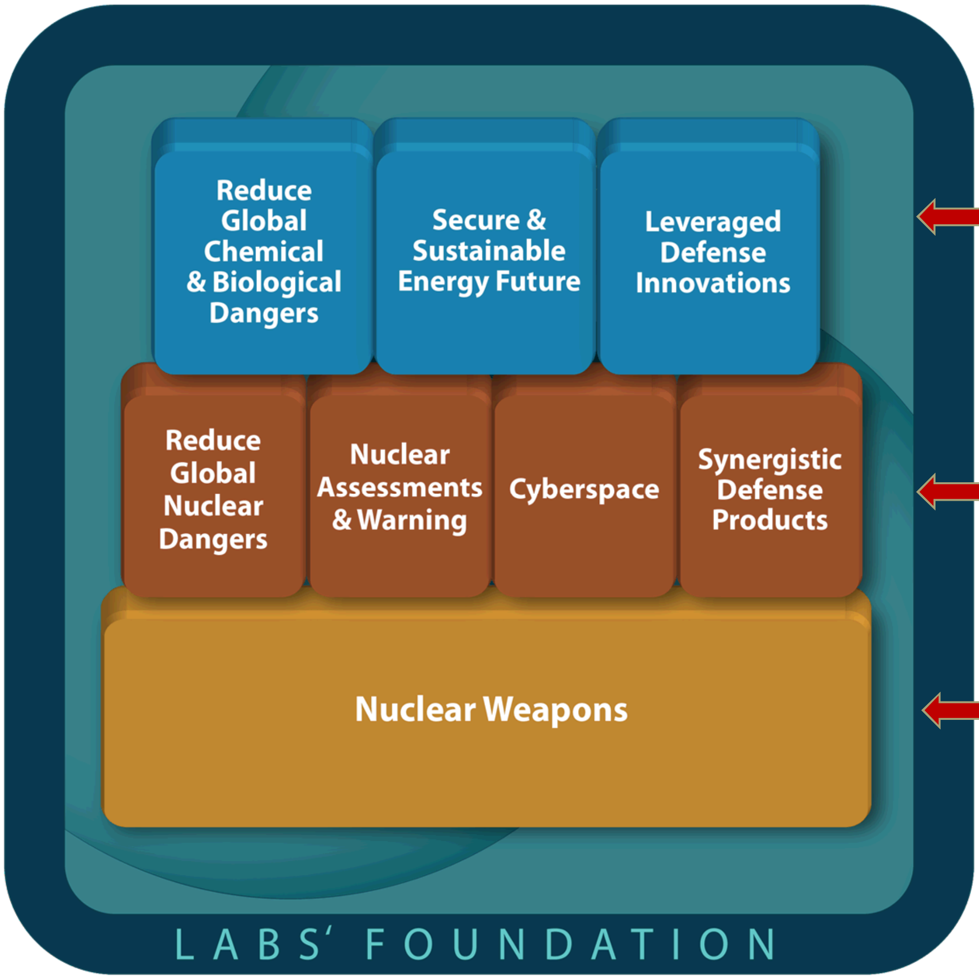
Waste Isolation Pilot Plant, Carlsbad, New Mexico

Pantex Plant, Amarillo, Texas



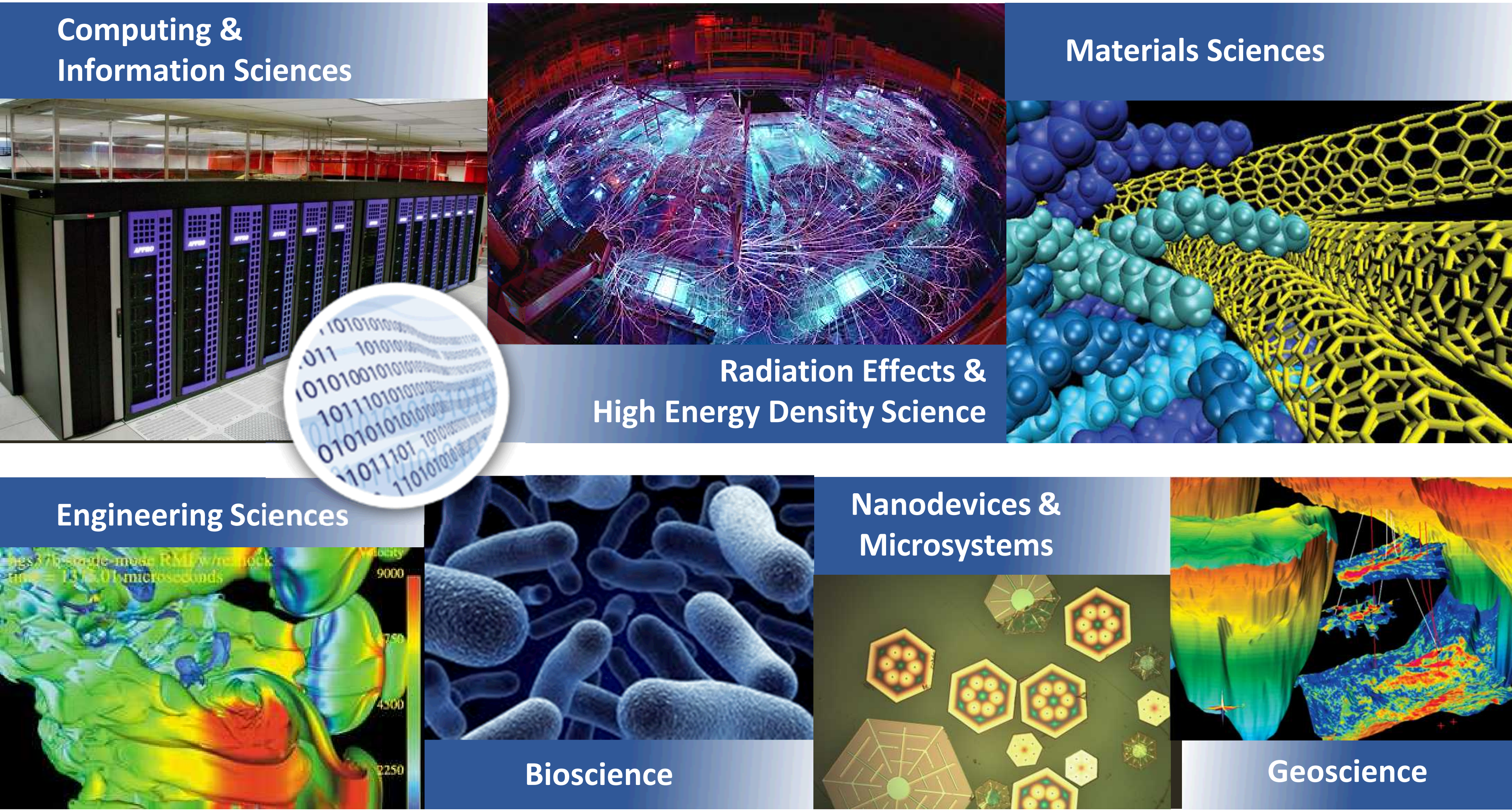
Tonopah, Nevada

National Security Mission Areas



- These mission areas are critical to national security, and they leverage, enhance, and advance our capabilities.
- These mission areas are essential to sustaining Sandia's ability to fulfill its NW core mission.
- *Our core mission*: nuclear weapons (NW).

Research Foundations



Computing and Information Sciences

Computing Sciences

" …the scientific approach to computation and application and specifically to the design of computing machines and processes"

Computational simulation is now considered a third fundamental tool of science, complementing theory and experiment.

Information Sciences

"" … the analysis, collection, classification, manipulation, storage, retrieval and dissemination of information … e.g. mathematics, cognition"

We live in the "information age" and increasingly substitute information for materials and energy to increase societal efficiency

Focus Areas

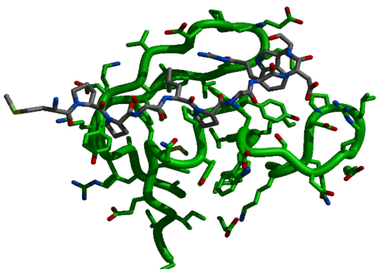
- Computational Engineering
- Scalable supercomputing
- Trusted information systems

Optimization Research at Sandia

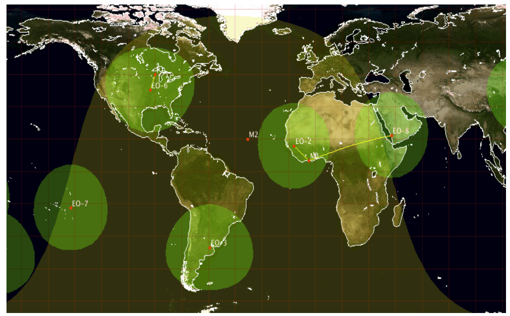
- Parallel optimization algorithms
  - Branch-and-bound, simulation-based optimization, etc
- PDE-constrained optimization
- Stochastic programming
- Parameter estimation
- Direct search
- Surrogate-based optimization
- Optimization under uncertainty
- *and more …*

- Optimization software frameworks
  - Engineering design and uncertainty quantification
  - Modeling frameworks
  - Solver libraries

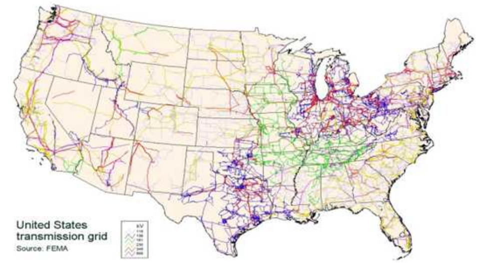
Optimization Application Examples



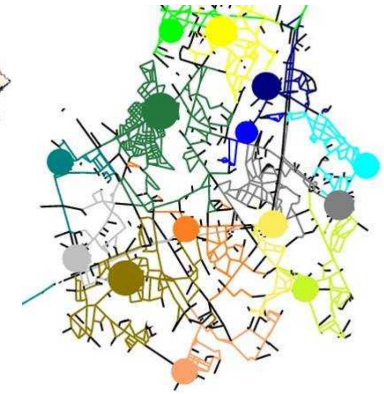
Molecular Docking



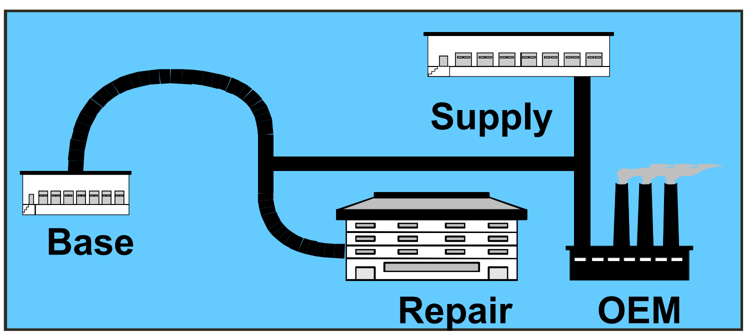
Satellite Scheduling



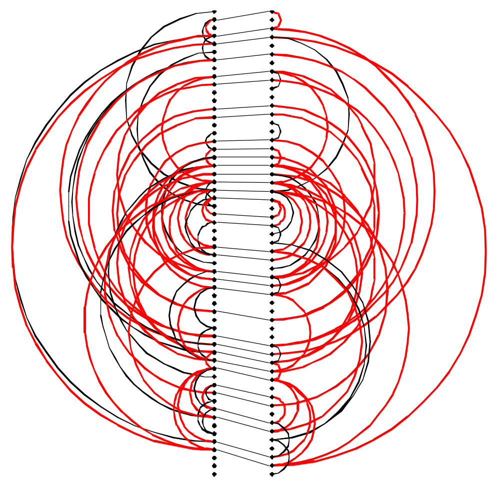
Power Grid Planning



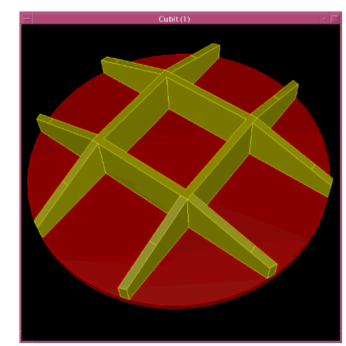
Water Security



Inventory Logistics



Protein Comparison



Topology Design

Overview

- About Sandia National Laboratories

- Modeling critical infrastructures
  - Water security
  - Electrical energy planning and management

- Optimization modeling with Pyomo

Resilient Critical Infrastructures

Presidential Policy Directive 21 (PPD-21): Critical Infrastructure Security and Resilience advances a national policy to strengthen and maintain secure, functioning, and resilient critical infrastructure.

- PPD-21 identifies 16 critical infrastructure sectors
- http://www.dhs.gov/critical-infrastructure-sectors

DHS provides broad guidance for supporting decisions and investments in infrastructure that will enhance the resilience of critical infrastructure systems, including:

- Leveraging science and predictive tools on future trends and risks
- Utilizing available risk assessment and scenario planning tools to make risk-informed decisions
- Mapping potential cascading effects from potential infrastructure disruptions

Some Practical Challenges

- The resiliency "problem" is poorly defined
  - *Need to explore alternative formulations*
  - *Need generic modeling/optimization strategies (e.g. MIP)*

- We may not have the data we need to solve "the problem"
  - *Some "good" formulations may not be practical*
  - *Inform stakeholders about the value of additional information*

- There are competing objectives for infrastructure resiliency
  - *Need multi-objective or goal-constrained techniques*

- Decision makers may need to make rapid decisions
  - *Need parallel techniques*
  - *Need good, fast heuristics*

Example: Electric Power Management

Unit commitment: a planning problem for scheduling electric power generation

Challenge: There is increased uncertainty in power generation

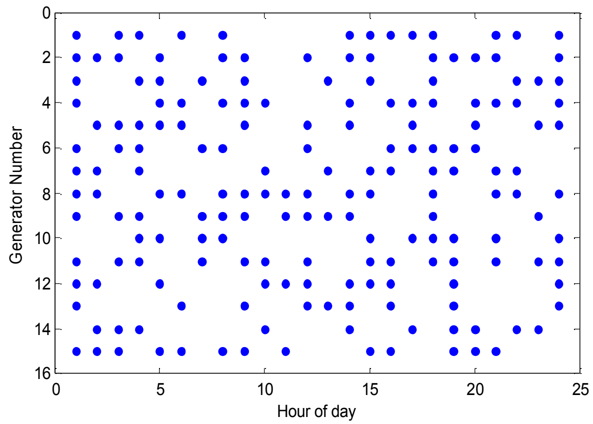
- Uncertainty and variability in alternative sources (e.g. wind, solar)
- Potential uncertainties in charging/discharging patterns for electric vehicles

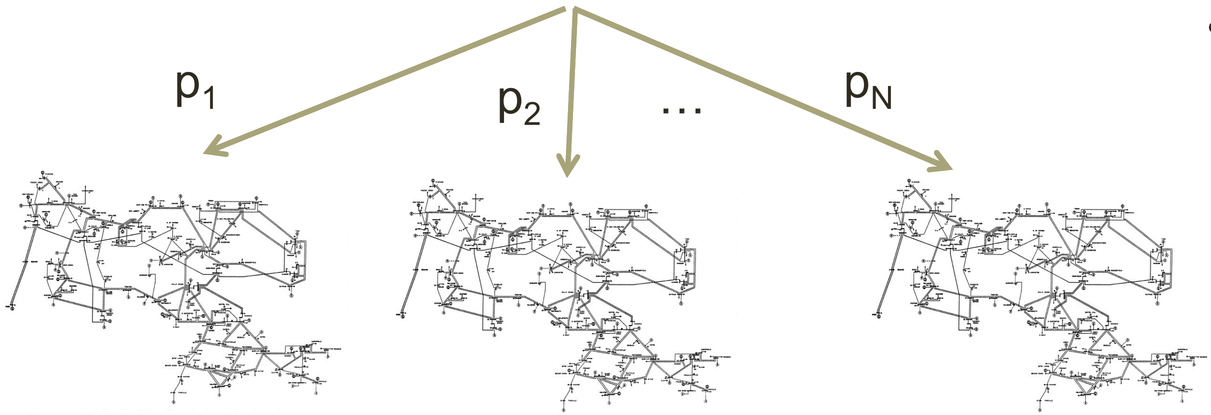
Idea: use stochastic unit commitment

- A scenario based uncertainty representation in the unit commitment formulation
- The objective is to minimize the expected cost

The General Structure of a Stochastic Unit Commitment Optimization Model

Objective: Minimize expected cost





Scenario 1 Scenario 2 … Scenario N

First stage variables:

- Unit On / Off

Nature resolves uncertainty

- Load
- Renewables output
- Forced outages

Second stage variables (*per time period*):

- Generation levels
- Power flows
- Voltage angles
- …

(Some) Historical Barriers to Adoption of Stochastic Unit Commitment

- We can't create sufficiently accurate sets of scenarios to capture load and renewables uncertainty

- Even if we could create accurate sets of scenarios, the resulting models are too difficult to solve

- Even if we could solve the resulting models, it would require significant HPC resources – which is a major impediment to industrial adoption

Solving with the Stochastic Extensive Form

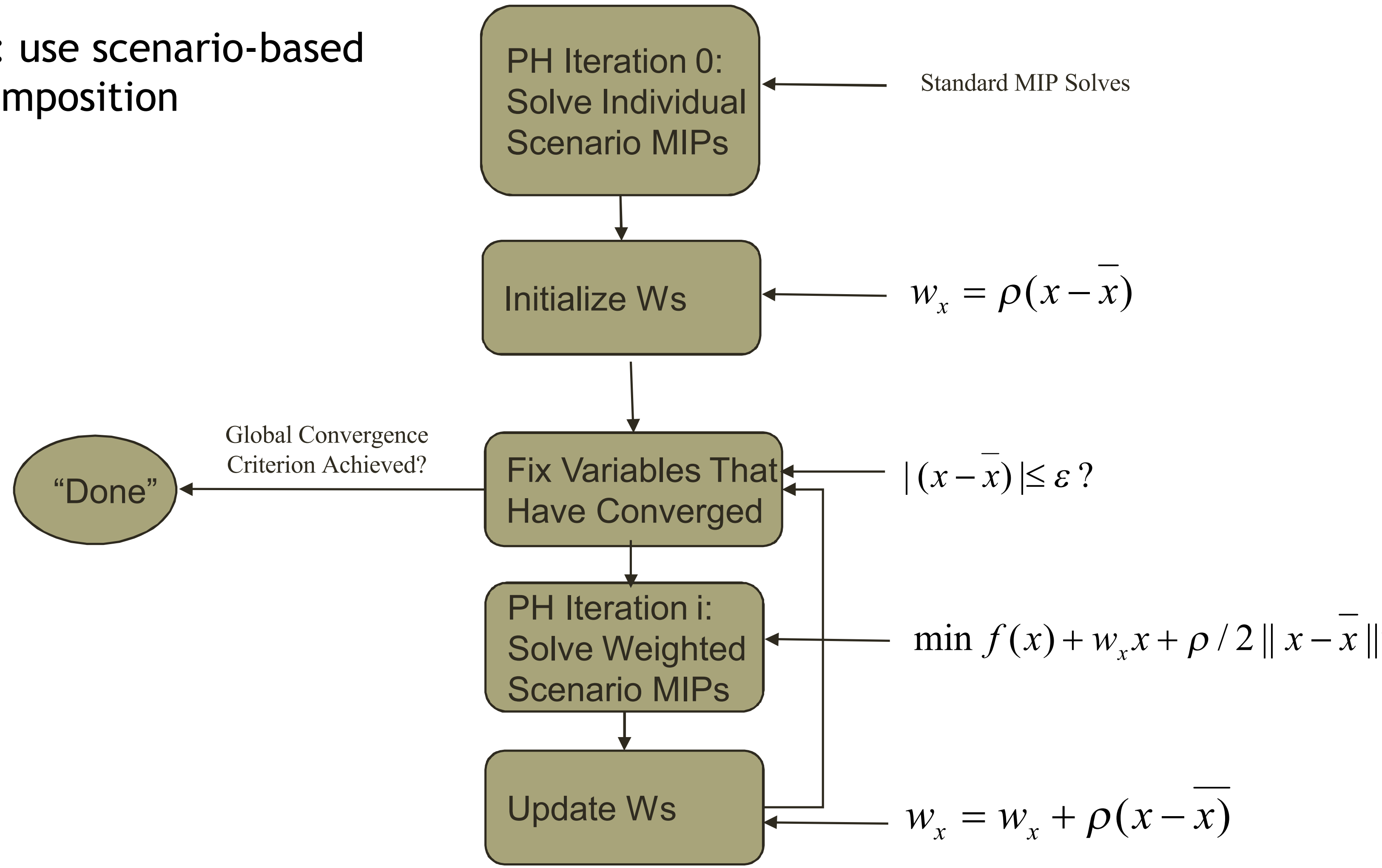
- Reliability Unit Commitment (RUC) Test Instance: WECC-240++
- J.E. Price, Reduced Network Modeling of WECC as a Market Design Prototype, 2011 IEEE PES General Meeting
- Changes necessary to create viable RUC test case
  - Addition of realistic ramping rates and min up/down time constraints
- Results

Table 3 Solution quality statistics for the extensive form of the *WECC-240-r1* instance, given 4 hours of run time.

# Scenarios	Objective Value	MIP Lower Bound	Gap %	Run Time (s)
3	64278.20	63797.72	0.75	14491
5	62740.67	62180.86	0.89	14723
10	61563.10	60835.45	1.18	14630
25	61455.55	59963.78	2.36	14960
50	61911.74	59540.87	3.83	15480
100	62388.85	59548.23	4.51	16562

Solving with Progressive Hedging

Idea: use scenario-based decomposition



Parallelization and Bundling

- Progressive Hedging is, at least conceptually, easily parallelized
  - Scenario sub-problem solves are clearly independent
  - Advantage over Benders, in that “bloat” is distributed
    - Critical in low-memory-per-node cluster environments
  - Parallel efficiency drops rapidly as the number of processors increases
    - But: Relaxing barrier synchronization does not impact PH convergence
  - Bundling scenarios might help with parallel scaling
    - May increase number of iterations required

- PH can provide bounds!
  - Now comes with (rather tight) lower bounds
  - See “Obtaining Lower Bounds from the Progressive Hedging Algorithm for Stochastic Mixed-Integer Programs” (Under review)

PH Results: Workstation and RedSky (HPC)

Table 4 Solve time (in seconds) and solution quality statistics for PH executing on the *WECC-240-r1* instance, with $\alpha = 1.0$, $\mu = 3$, and $\gamma = 0.03$

# Scenarios	Convergence Metric	Obj. Value	PH L.B.	# Vars Fx.	Time
		64-Core Workstation Results			
3	0.0 (in 23 iters)	64727.714	63188.709	4080	155
5	0.0 (in 26 iters)	62911.104	61609.576	4080	163
10	0.0 (in 26 iters)	61493.375	60347.220	4080	227
25	0.0 (in 27 iters)	60990.111	59875.661	4080	364
50	0.0 (in 17 iters)	60721.319	59527.252	4076	584
100	0.0 (in 23 iters)	61156.832	59880.559	4080	1218
		Red Sky Results			
50	0.0 (in 29 iters)	60676.383	59670.142	4062	514
100	0.0 (in 33 iters)	61122.781	60148.285	4073	672

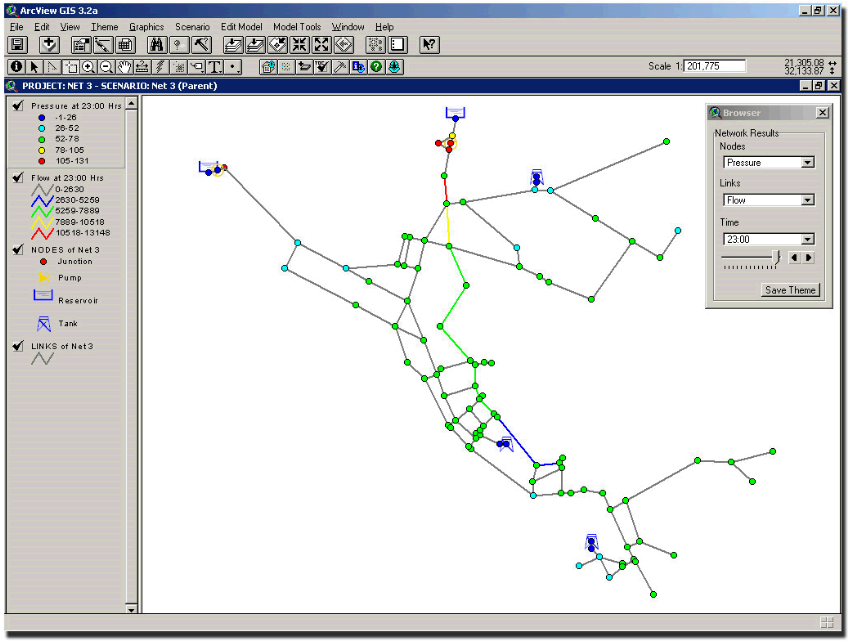
Example: Water Security

<u>Drinking Water</u>

- Water source
- Treatment facilities
- Transmission systems
- Distribution systems

<u>Wastewater</u>

- Wastewater source
- Collection system
- Treatment facility
- Receiving water body



Designing a Contamination Warning System

Technical Goal: placement of sensors in a water distribution system within a budget

Possible objectives:

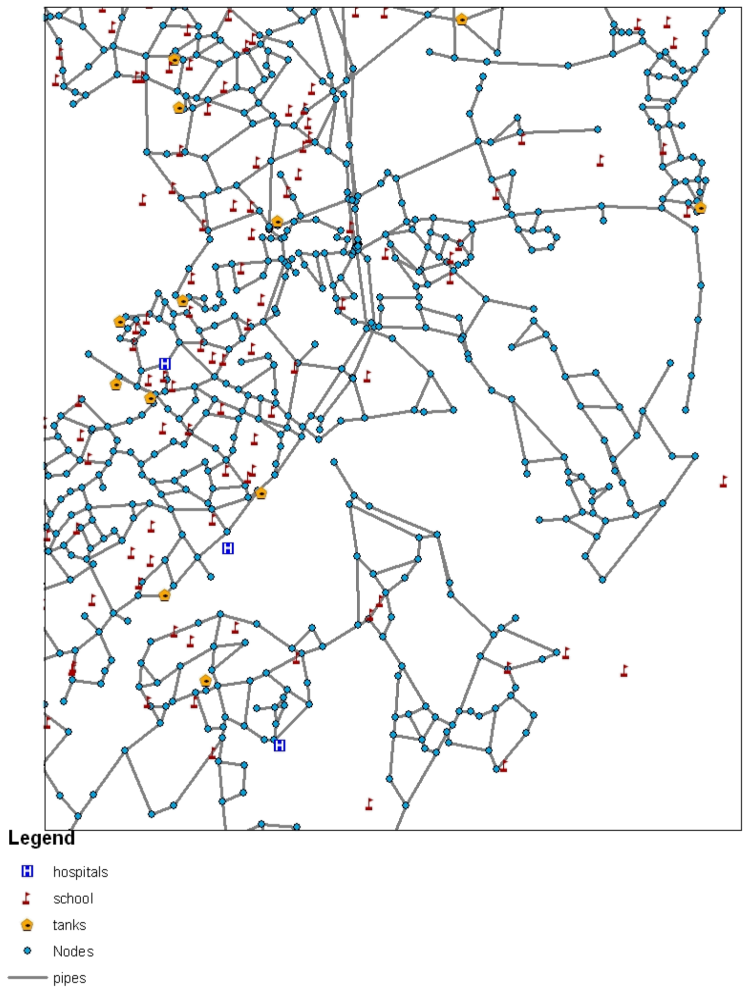
- Minimize response time
- Minimize health impacts
- Minimize contamination extent

Place sensors at networks junctions

- Public buildings, hospitals, etc

Sensors are expensive

- Cost of sensors
- Cost of installation



Contamination Scenarios

Water movement (direction, velocity in each pipe) determined by

- Demand (consumption), pumps, gravity, valves, sources/tanks, etc.

Current (most trusted) simulator

- EPANET code computes hydraulic equations to determine flows
- Discrete-event simulation for contaminant movement

Contamination impact

- Scenarios defined by start time, location and contaminant characteristics
- For each scenario, determine when/where the contaminant can be observed
- Compute cumulative impact statistics
  - Population exposed, # deaths, time of detection, mass consumed, etc

A Canonical Sensor Placement Problem

- Classic p-median facility location problem:
  - Open p facilities
  - Each of a set of customers served by closest facility
  - Minimize total distance

- Sensor placement is structurally equivalent
  - Assume: perfect sensors, a general alarm is raised that stops water consumption
  - Sensors = Facilities
  - Events = Customers to be "served" (witnessed)
  - "Distance" from an event *a* to a node *n* = impact if a sensor at node *n* witnesses event *a*.
  - If no sensors witness an event, the impact for that event is the maximum possible over the simulation time-horizon

Solving p-Median Sensor Placement

- Integer programming
  - IP model can capture different objectives/networks
  - Can be solved with COTS software on 64-bit computers

- GRASP heuristic
  - Runs quickly on real-world distribution networks (2000-20000 junctions)
  - In practice, often finds optimal sensor placements (!?!)

- Stochastic programming
  - The IP model can be viewed as the extensive form of a stochastic model
  - Can apply Progressive Hedging heuristic to solve in parallel and/or with limited memory
  - Can compute confidence interval for a sensor placement

Other Formulations

- Imperfect sensors
  - Allow sensors to fail with a fixed probability (perhaps dependent on location)
  - Is well-approximated by p-Median model
  - Need more data!
- Minimize number of sensors
  - Given a performance goal, minimize # of sensors
  - Requires a target goal …
- Multi-stage sensor placement
  - Place some sensors now and others later
  - Requires a model for the value of life today vs. tomorrow
- Multi-objective sensor placement
  - Goal-constrained sensor placement is harder for GRASP and IP solvers
  - Robust objectives (e.g. CVaR) are computationally challenging (*)

Overview

- About Sandia National Laboratories
- Modeling critical infrastructures
  - Water security
  - Electrical energy planning and management
- Optimization modeling with Pyomo

Optimization Modeling

Goal:

- Provide a natural syntax to describe mathematical models
- Formulate large models with a concise syntax
- Separate modeling and data declarations
- Enable data import and export in commonly used formats

Impact:

- Robustly model large constraint matrices (e.g. for MILPs)
- Integrated support of automatic differentiation for complex nonlinear models

Examples: AMPL, GAMS, OptimJ, AIMMS, FlopCPP, PuLP, …

Pyomo Overview

Idea: a Pythonic framework for formulating optimization models

- Provide a natural syntax to describe mathematical models
- Formulate large models with a concise syntax
- Separate modeling and data declarations
- Enable data import and export in commonly used formats

Highlights:

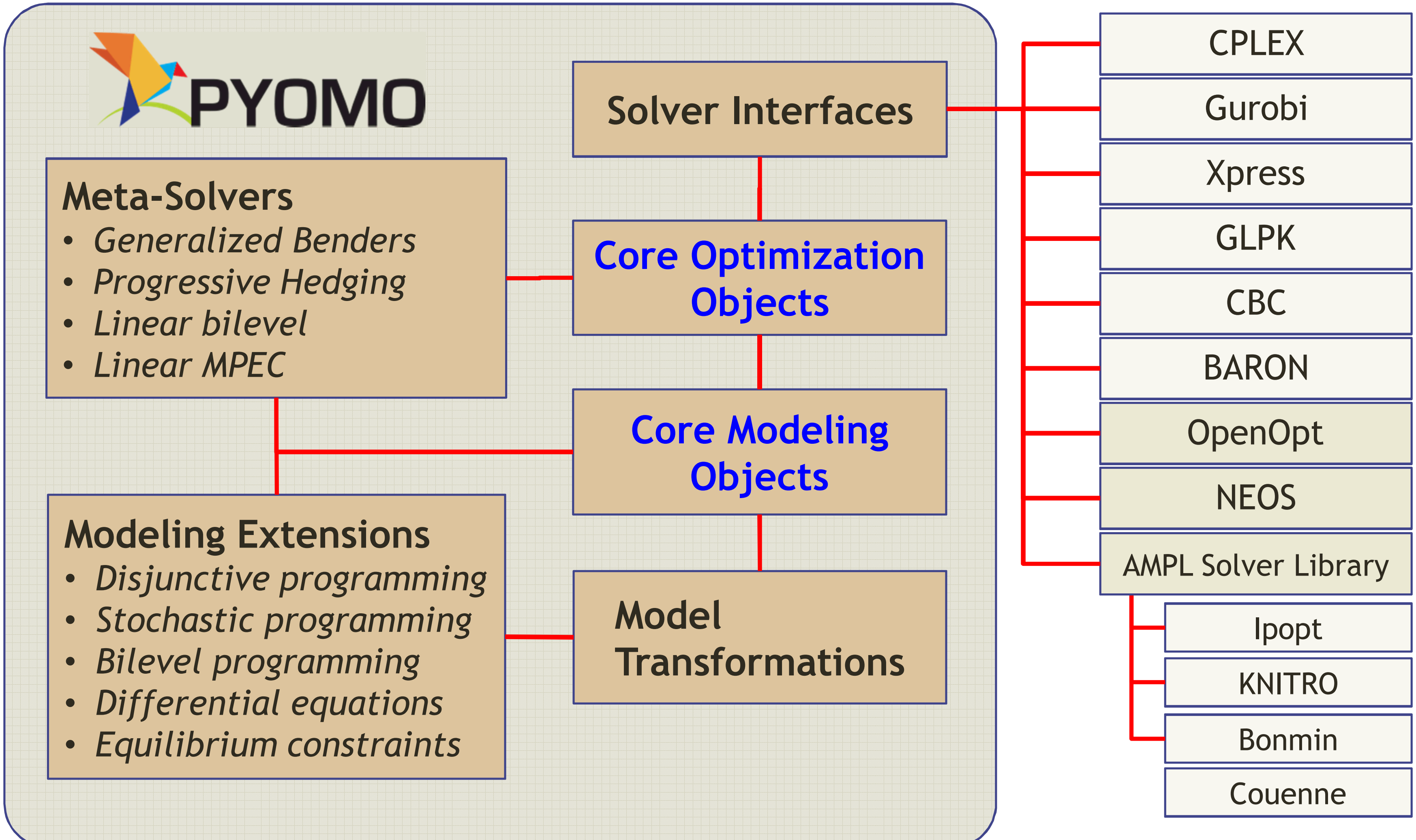
- Python provides a clean, intuitive syntax
- Python scripts provide a flexible context for exploring the structure of Pyomo models

```
# simple.py
from pyomo.environ import *

M = ConcreteModel()
M.x1 = Var()
M.x2 = Var(bounds=(-1,1))
M.x3 = Var(bounds=(1,2))
M.o  = Objective(
       expr=M.x1**2 + (M.x2*M.x3)**4 + \
           M.x1*M.x3 + \
           M.x2*sin(M.x1+M.x3) + M.x2)

model = M
```

Comparison with Other Python Modeling Tools

- Pyomo
  - Supports concrete/abstract modeling for LP/MILP/NLP models
  - Modeling extensions for stochastic programming, bilevel, MPEC, etc
  - Separate model objects
- PuLP
  - Supports concrete modeling for LP/MILP models
  - Separate model objects
  - Simple object model
- APLEpy
  - Supports concrete modeling for LP/MILP models
  - Single global model object
- PyMathProg, pyglpk, cplex, gurobi
  - Python interfaces for specific solver tools

Pyomo at a Glance



Why Model within a Programming Language?

Open Source License

- No licensing issues w.r.t. the language itself
- Can extend/refine the language in some cases

Extensibility and Robustness

- Highly stable and well-supported
- Simple model for integrating code developed by a user

Support and Documentation

- Extensive online documentation and several excellent books
- Long-term support for the language is not a factor

Standard Library

- Includes a large number of useful modules.

Scripting

- Language features includes functions, classes, looping, procedural constructs, etc.

Portability

- Widely available on many platforms

More than just mathematical modeling …

Scripting

- Construct models using native Python data
- Iterative analysis of models leveraging Python functionality
- Data analysis and visualization of optimization results

Model transformations (a.k.a. reformulations)

- Automate generation of one model from another
- Leverage Pyomo’s object model to apply transformations sequentially
- E.g.: relax integrality, GDP -> Big M

Meta-solvers

- Integrate scripting and/or transformations into optimization solver
- Leverage Python’s introspective nature to build “generic” capabilities
- E.g.: progressive hedging, SP extensive form -> MIP

Impact on Critical Infrastructure Analysis

- Agility
  - Optimization models were able to express a wide range of sensor placement formulations
    - Used both AMPL and Pyomo models
  - Pyomo facilitated the transition between IP and SP models
    - The SP representation supported by Pyomo requires a single-scenario model
    - Can easily transition between IP solvers optimizing the extensive form and PH
  - Transformations can be used to tailor analysis to available solvers
    - Commercial IP solvers are not available in some environments
    - Meta-solvers like PH can leverage weaker open-source IP solvers when analyzing subproblems

Impact on Critical Infrastructure Analysis

- Leveraging generic capabilities
  - Pyomo is developing generic SP analysis capabilities
    - The SP analysis is separate from the model specification
  - Consequently, extensions to these capabilities can be immediately leveraged by any/all SP applications

- Fast analysis
  - Leverage Python's capabilities to support scalable parallelism
  - We leverage portable parallel libraries in Pyomo, so no customization is required between workstation, cluster or HPC environments

Acknowledgements

- Sensor placement
  - Cindy Phillips, Jon Berry, Jean-Paul Watson, John Siirola, Regan Murray, Jim Uber, and more …

- Stochastic unit commitment
  - Jean-Paul Watson, Dave Woodruff, Roger Wets

- Pyomo
  - Jean-Paul Watson, John Siirola, Dave Woodruff, Carl Laird

The End

Questions?

For More Information

See the new Pyomo homepage

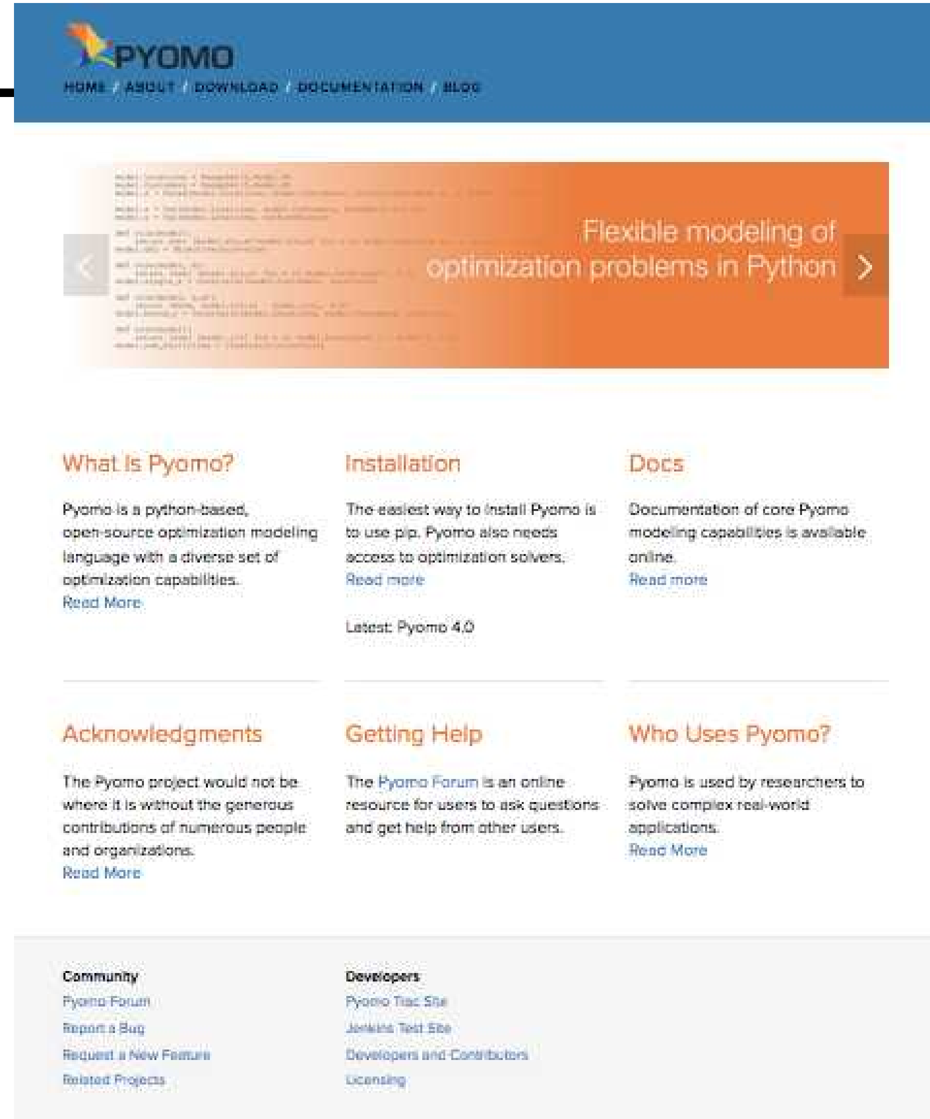
- www.pyomo.org

The Pyomo homepage provides a portal for:

- Online documentation
- Installation instructions
- Help information
- Developer links

Coming soon:

- Blogging about Pyomo capabilities and features
- A gallery of simple examples



Integer Programming Formulation

IPs can be used to model sensor placement for water security

- Berry et al (2003, 2006); Watson et al (2004)

Objective: $\boxed{\sum_{a \in A} \sum_{i \in L} \alpha_a w_{ai} x_{ai}}$

- α - contamination likelihood
- *w* - contamination impact
- *x* - witness variable
- *s* - sensor placement variable

(*)

$$
\boxed{
\begin{array}{ll}
\text{minimize } \sum_{a \in A} \sum_{i \in L} \alpha_a w_{ai} x_{ai} & \\
s.t. & \\
\sum_{i \in L} x_{ai} = 1 & \forall a \in A \\
x_{ai} \le s_i & \forall a \in A, i \in L \\
\sum_{i \in L} s_i \le S_{\max} & \\
s_i \in \{0,1\} & \\
0 \le x_{ai} \le 1 & \forall a \in A, i \in L
\end{array}
}
$$

What Happened to Coopr?

- Users were installing Coopr but using Pyomo
  - Pyomo modeling extensions were not distinct enough
  - Researchers cited “Coopr/Pyomo”
- Users/Developers were confused by the coopr and pyomo commands
- Developers were coding in Coopr but talking about Pyomo

We need to provide clear branding this project!

Who Uses Pyomo?

- Students
  - Rose-Hulman, UC Davis, U Texas, Iowa State, NPS

- Researchers
  - Sandia National Labs, Lawrence Livermore National Lab, Los Alamos National Lab, UC Davis, TAMU, Rose-Hulman, UT, USC, GMU, Iowa State, NCSU, U Washington, NPS, U de Santiago de Chile, U Pisa, Federal Energy Regulatory Agency, …

- Software Projects
  - TEMOA – Energy economy optimization models
  - Minpower – Power systems toolkit
  - Water Security Toolkit – Planning/Response for water contamination
  - SolverStudio – Excel plugin for optimization modeling