SimX meets SCIRun: A Component-Based Implementation of a Computational Study System

Siu-Man Yau, Vijay Karamcheti, Denis Zorin
New York University

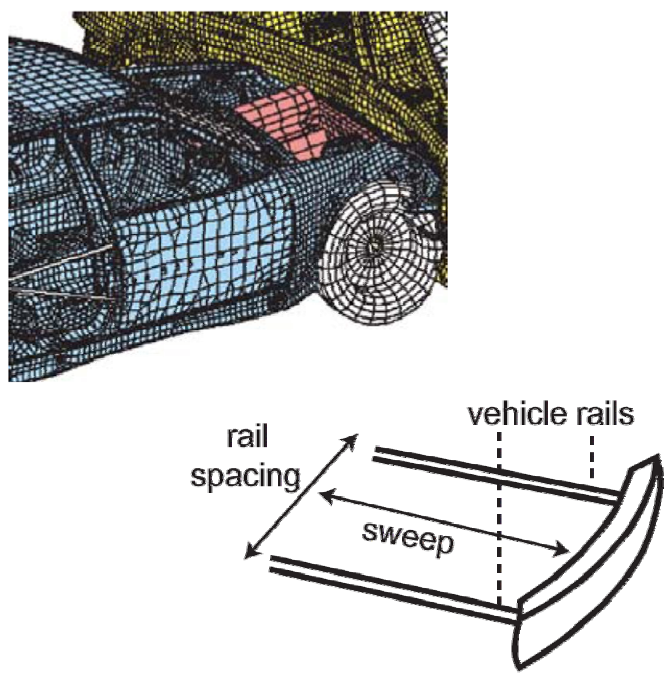
Eitan Grinspun
Columbia University

Also participating in the project:
Akash Garg (Columbia), Jeremy Archuleta (Utah), Steve Parker (Utah)

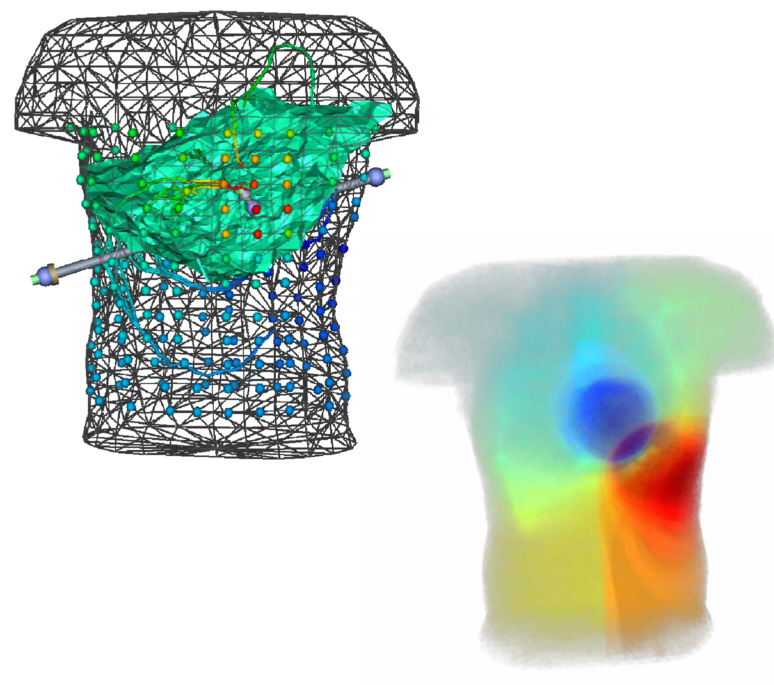
Multi-Experiment Computational Studies

- Computer Simulation is an integral part of the scientific method
  - Wide-spread use of computational studies
  - Involving 100s, 1000s, or 10,000s of simulations …
- Examples

Car Chassis Design



Defibrillator Design



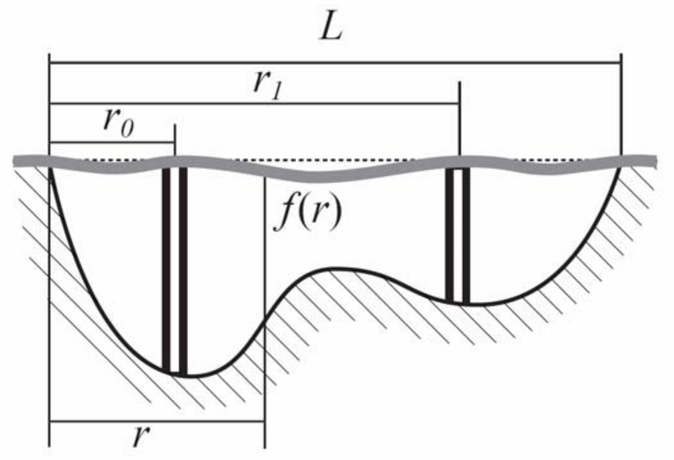
SimX: A Computational Study System

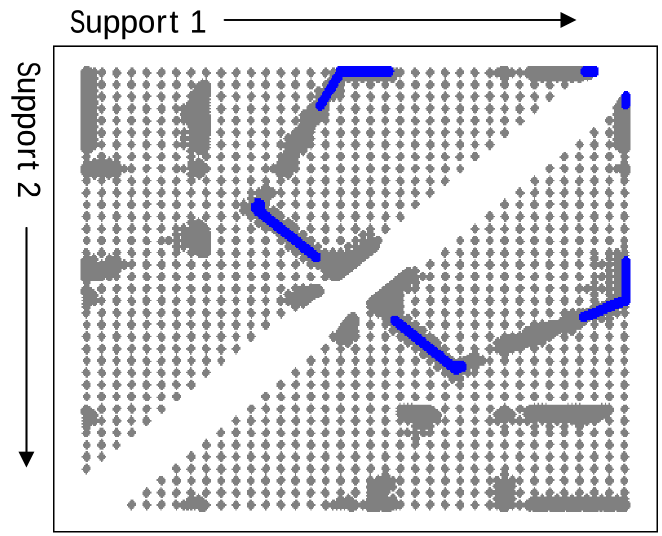
- Requirements
  - High-level control of the study, not individual experiments
  - Support for interactivity
    - Visualization and steering mechanisms
    - Dynamic reallocation of resources in response to changing objectives
- Existing systems fall short
  - Parameter Sweep schedulers
    - Examples: Condor, Globus, Virtual Instrument, Nimrod/O
    - Provide poor support for interactivity, high-level study objectives
  - Computational Steering Infrastructures
    - Examples: Falcon, CUMULVS, SCIRun, CSE
    - Poor support for macro-management of studies, resource allocation
- SimX: Interactive, high-level management of computational studies
  - Users interact at the level of aggregate studies
  - Aggressive management of system resources to reflect changing priorities

SimX: Technical Ideas

- Computational study: Exploration of a parameter space of computational experiments to identify a target set
  - Target set specified as constraints on experiment observations

Bridge design study, IPDPS'06





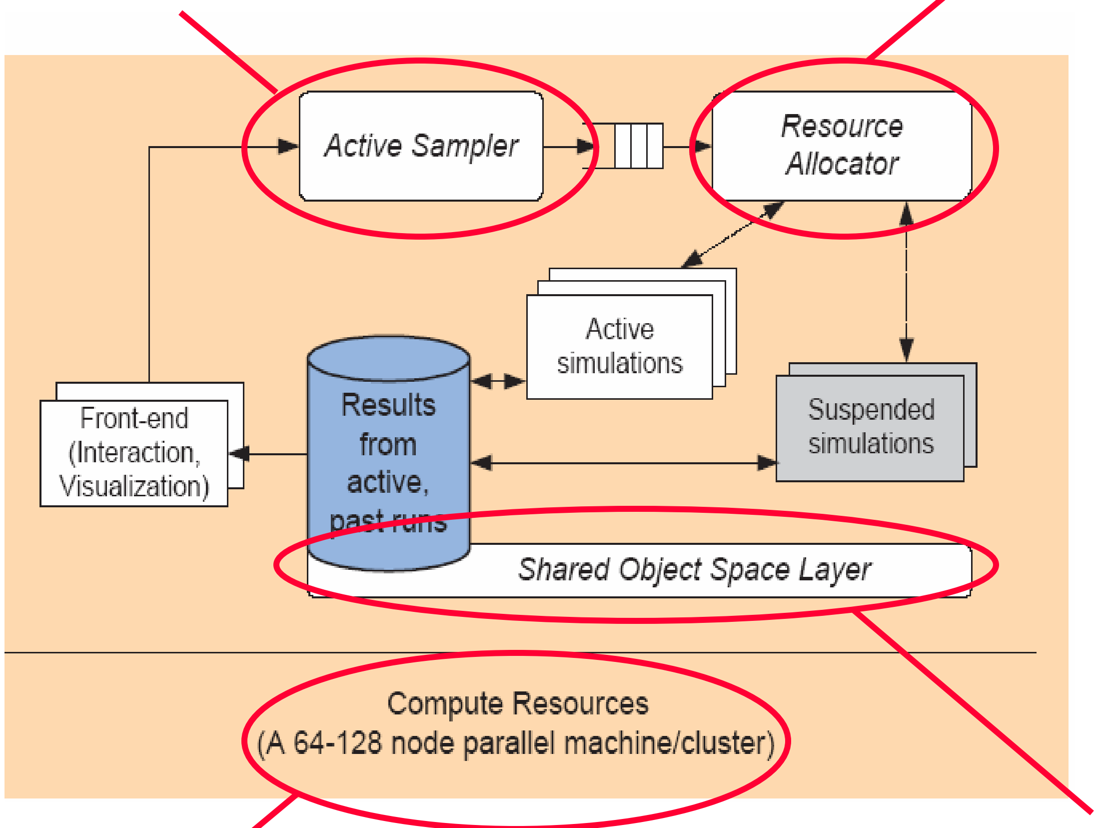
Study time reduced from 5678s to 13s

- SimX exploits a permeable interface between the system software and the application to deliver efficiency
  - Chooses which simulation to run and in which order
  - Maintains history, and aggressively reuses results

SimX: Architecture Components

"Brain" of the system: uses domain info to identify experiments to run, their order, and associated accuracy/efficiency tradeoffs

Value- and locality-aware assignment of experiments to machine resources



Targets growing availability of "personal" parallel computing (multicore CPUs, GPGPUs, …)

Machine-wide repository of study state (checkpoints, observations, profiles …)

Spatially-indexed, relaxed consistency, neigborhood queries

This Paper: Componentized SimX

Ongoing implementation of the SimX platform on top of Univ. of Utah's SCIRun/SCIRun2 problem-solving environment

- SCIRun/SCIRun2
  - Modular, CCA-compliant, component-based system for interactive scientific computing applications
  - Expresses applications as dataflow nets
    - Runtime system responsible for transfer of data between modules, and execution of module code
  - Rich library of solver, visualization, and steering modules
- Benefits from SimX/SCIRun integration
  - Natural component-level integration of SimX runtime modules with existing simulation functionality
    - Easier to express the permeable interface
  - Plug-and-play benefits for SimX policy modules
  - Access to existing scientific computing software
  - Transition path for SimX technology to broader scientific community

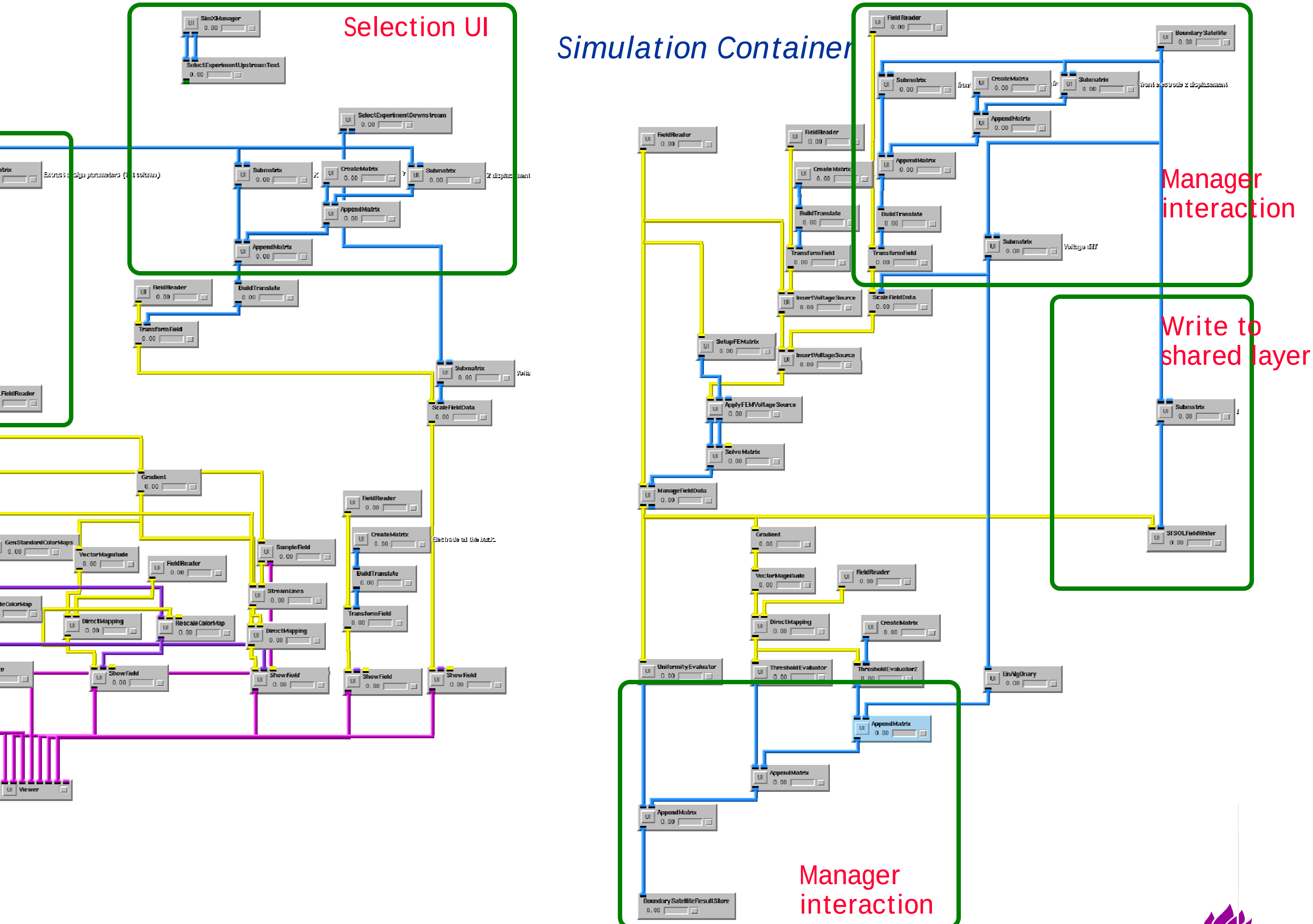
Realizing SimX on SCIRun

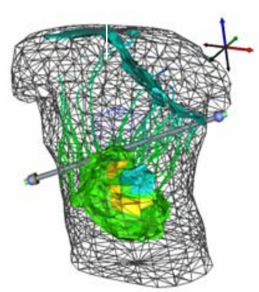
- One manager and multiple simulation container nets
  - Launched as separate SCIRun processes
    - Interact via sockets using communication components
  - SCIRun2 supports parallel/distributed nets, so more natural

SimX components

- Manager
  - SimXManager: Encapsulates the Active Sampler functionality
  - SelectExperiment: UI interface to permit visualization
    - Currently per-experiment, being extended to study level
- Simulation container
  - Arbiter: Manages interaction with the Manager,
    collection of performance measures
- Common
  - SharedObjectLayer{Reader,Writer}

Example: Defibrillator Design Study





DefibSim Study: Performance Evaluation

- 256 node cluster (dual 2GHz PowerPC, 2GB, GbE), 4 shared object servers
- Design space for study: Electrode positions, potential difference

2D case: Back electrode fixed,
Front electrode has one free dimension

	Grid Sampler	*Active Sampler*	
# Sim procs	Time (secs)	Time (secs)	# Exp Issued
1	9100	1790	760
2	5673	1074	769
4	2842	555.2	788
8	1427	315.6	771
16	818.6	146.4	787
32	359.3	83.80	795
64	181.4	42.66	881
128	134.0	39.31	1035

3D case: Back electrode fixed,
Front electrode has two free dimensions

	Grid Sampler	*Active Sampler*	
# Sim procs	Time (secs)	Time (secs)	# Exp Issued
1	9160	5217	2329
2	5702	3313	2365
4	2856	1371	2416
8	1278	771.4	2418
16	717.0	421.9	2330
32	362.5	218.3	2348
64	182.2	123.7	2390
128	95.62	95.35	2511

Next Steps

Short term

- Steering/visualization of entire studies
  - Abstractions to allow interactive navigation through design space
- More aggressive reuse
  - Study level
  - Experiment level
- Active sampling schemes
  - Sensitivity analysis along different design space dimensions
- Shared object layer improvements
  - Value-based management of resources

Longer term

- SCIRun2-based support for studies, parallel experiments
- Analytical resource, error estimates for core numerical solvers
- More sophisticated use of history information
  - Prediction, resource usage models