SAND2016-8706C

Exceptional service in the national interest

Sandia National Laboratories

Structural Simulation Toolkit: Macroscale Components

Nov 18,

SC 2014, New Orleans, LA

U.S. DEPARTMENT OF ENERGY

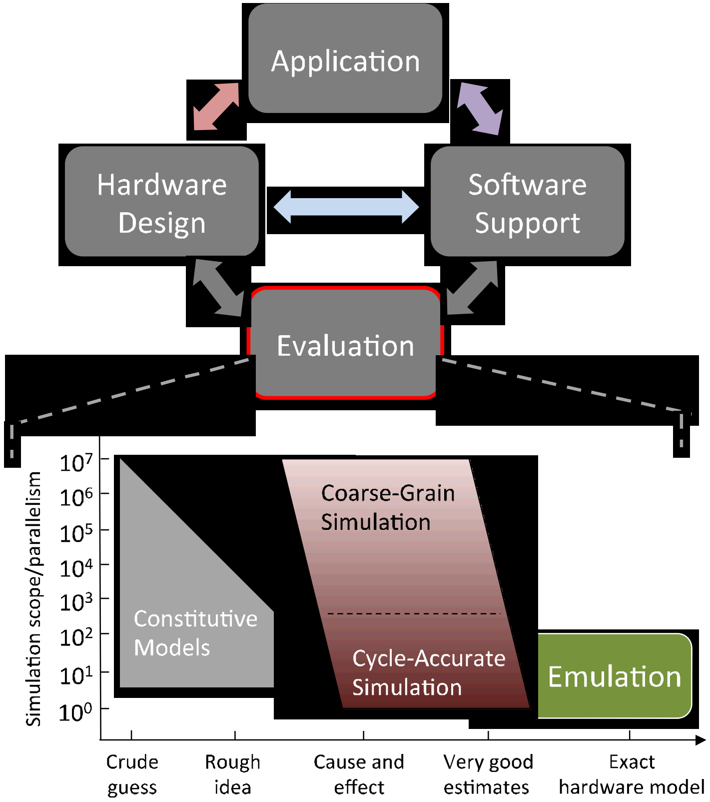
NNSA National Nuclear Security Administration

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. SAND NO. 2011-XXXXP

Aims of the Tool

Structural Simulation Toolkit is a *framework* for parallel discrete event simulation
The macroscale components explore a 2x2 simulation space of Hardware x Software
Hardware = Analytic or Structural Models   Software = Trace or On-line

- Aim #1: Macroscale co-design of skeleton applications/traces
  - Topology? Network? Memory? Where to spend money?
  - Deeper analysis of congestion, bandwidth tapering, comm pattern
- Aim #2: Runtime system development
  - How does computation evolve over time? How does runtime handle e.g. dynamic load balancing?
  - Deterministic debugging of parallel programs?



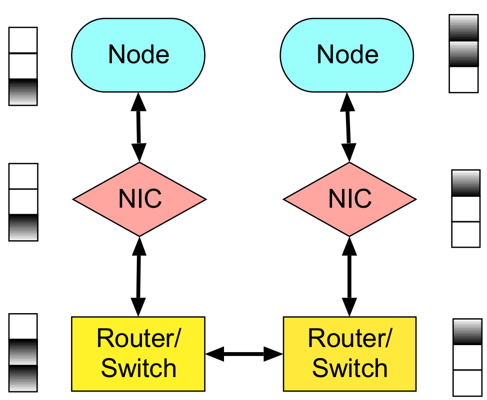
Typical Workload

Workload #1

MPI skeleton app or trace replay on coarse-grained network:

- 10K-100K MPI Ranks
- Basic question: what network and MPI layout best supports application?
- Not just comm/network congestion! Need compute and sync delays
- “Multi-scale”: Compute event = ~1ms Network = ~100ns
- Can be direct compilation of native C/C++ code

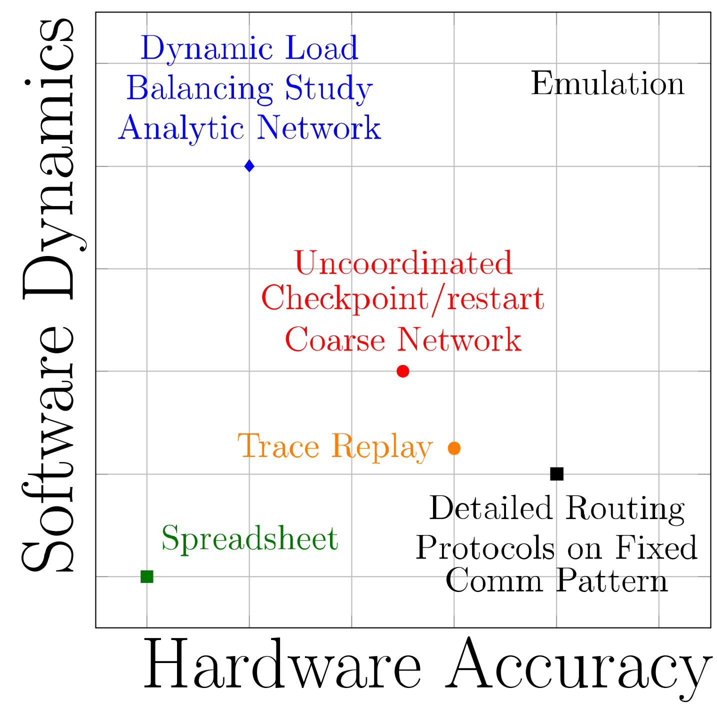
Buffers and queues for congestion. Coarse-grained, not flit level



Workload #2

Async many-task runtimes

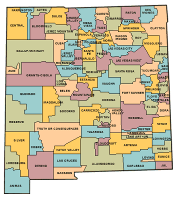
- ~1M threads = 100K ranks x 10 threads
- Congestion less important than dynamic load balancing
- Simulate things like Charm++, Legion, UPC – still working on adapting APIs



Ecosystem and Integration

Next-gen programming models for NNSA codes via Mantevo mini-apps
Janine Bennett (SNL-CA)

Uncertainty quantifcation
Development of discrete event UQ
Khachik Sargsyan, Habib Najm, Bert Debusschere (SNL-CA)



Keren Bergman
Optical Network Simulation

DOE co-design centers
Algebraic multi-grid and adaptive mesh refinement simulations
John Bell and John Shalf (LBL), Jacqueline Chen (SNL-CA)

Sudhakar Yalamanchili
Generation of compute/energy models with Eiger toolkit

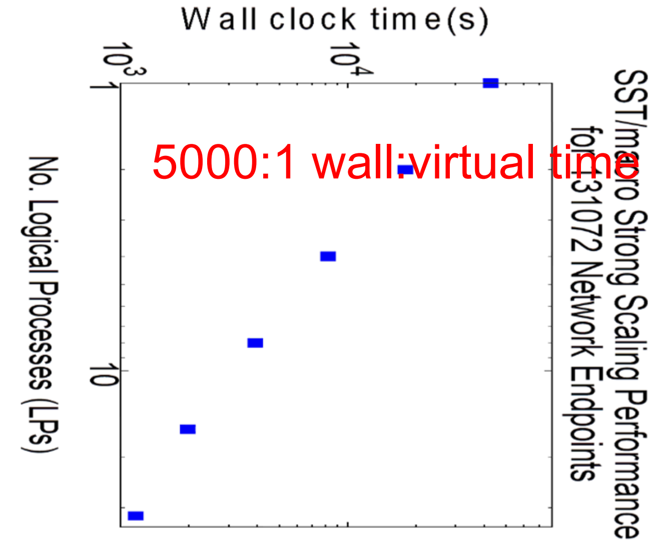
Self-assessment

Sandia National Laboratories

```
int
main(int argc, char** argv)
{
  ...
  for (int i=0; i < niter; ++i){
    MPI_Irecv(...);
    MPI_Irecv(...);
    MPI_Isend(...);
    MPI_Isend(...);
    DGEMM(...);
    MPI_Waitall(...);
  }
}
```

Single-source for SST and real runs? Almost…

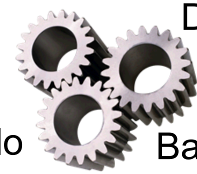
- Need better compiler tools
  - Auto-skeletonization of exisitng codes
  - Compute model generation
  - Global variables!
- PDES/extreme-scales
  - Is our PDES enough? Is any good enough?
  - Scalability of structural simulation with congestion modeling? (lookahead=100ns)
  - Scalability of analytic congestion models (lookahed=1µs)
- Data-dependent control flow
  - Model all the metadata, but not data
  - AMR/PIC needed to generate control flow
  - Some success with “box traces” for AMR
- Adaptive routing



Success stories

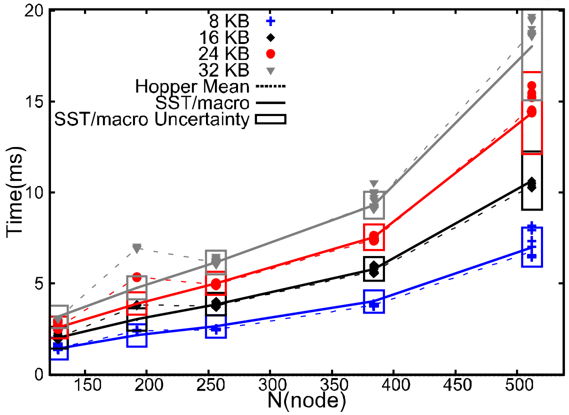
UQ Workflow:
Bracket errors!

Calibration Data

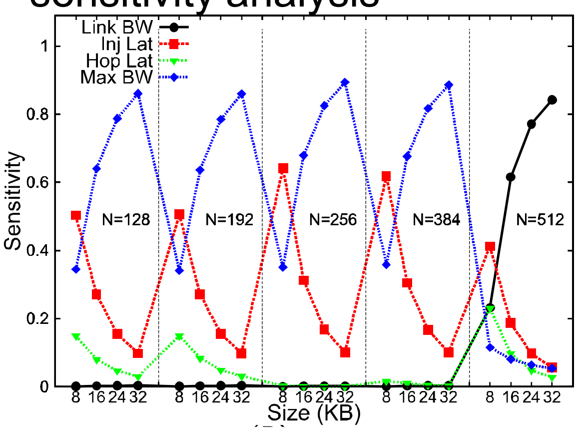


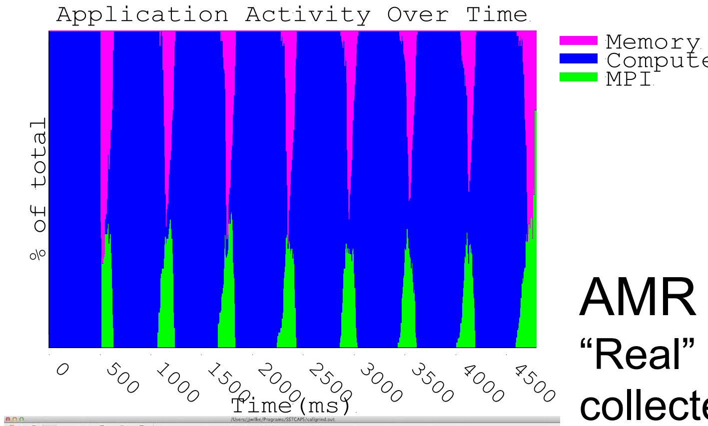
Adaptive Markov Chain Monte Carlo

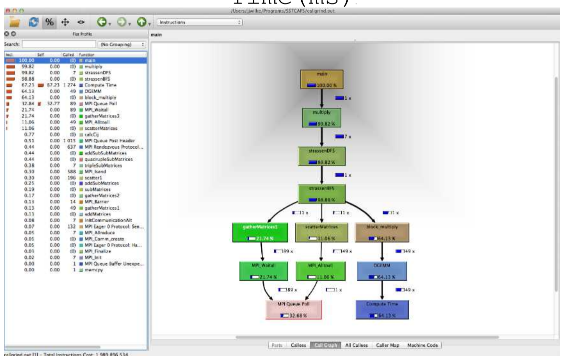
Bayesian Inference



Simulations with error bars and sensitivity analysis







AMR codes

"Real" performance metrics collected. Simulated data-dependent simulation via metadata trace

Development of many-task runtime

Over 2M threads simulated on 512 cores