SAND2012-2185 C

SAND2012-2185C

Improvements to the Structural Simulation Toolkit

Arun Rodrigues, Keren Bergman, David Bunde, Elliot Cooper-Balis, Kurt Ferreira, K. Scott Hemmert, Brian Barrett, Cassandra Versaggi, Robert Hendry, Bruce Jacob, Hyesoon Kim, Vitus Leung, Michael Levenhagen, Mitchelle Rasquinha, Rolf Riesen, Paul Rosenfeld, Maria del Carmen Ruiz Varela, Sudhakar Yalamanchili

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.

View of the Simulation Problem

Scale.....

Many
Cores
+
Memory

X

Many
Many
Nodes

X

Many
Many
Many
Threads

Multiple Audiences.....

Network
Processor
System

X

Application writers
purchasers
designers

X

system procurement
algorithm co-design
architecture research
language research

X

present systems
future systems

Complexity.....

Multi-Physics Apps
Informatics Apps

X

Communication Libraries
Run-Times
OS Effects

X

Existing Languages
New Languages

Constraints.....

Performance
Cost

Power
Reliability

Cooling
Usability

Risk
Size

Worldwide Impact

"Total power used by servers [in 2005] represented ... an amount comparable to that for color televisions. "

-ESTIMATING TOTAL POWER CONSUMPTION BY SERVERS IN THE U.S. AND THE WORLD, Jonathan G. Koomey

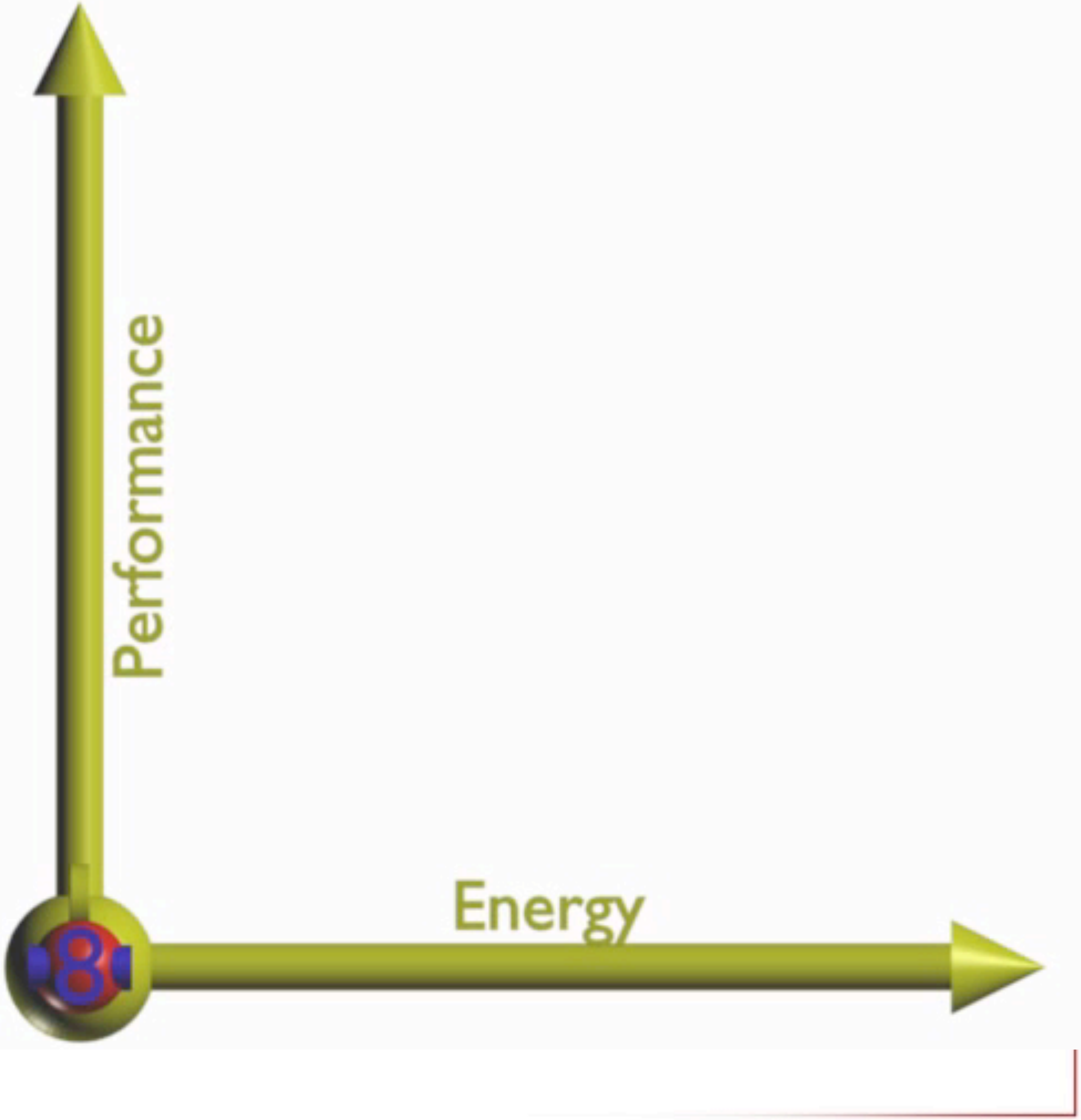
3741e9 KW-Hrs	Total US power consumption
* 3-4%	used by computers (>2% servers, >1% household computer use)
= 112 - 150e9 KW-Hrs	US Computer power consumption
* $0.1 $/KW-Hr	Retail cost, US Average 2009
= $11 - $15	Billion US$ in compute power
* 3-5	in 2005 US was roughly 1/3 of servers, by power.
= $33 (🇶🇦) - $75 (🇲🇲)	Billion US$ in worldwide computer power
* 15-35%	DRAM memory power
= $5 (🇲🇼) - $25 (🇱🇻)	BIllion in US$ in DRAM power

Major Simulation Challenges

- Multiple Objectives
  - Performance used to be only criteria
  - Now, Energy, cost, power, reliability, etc...
- Scale & Detail
  - Many system characteristics require detail to measure
  - Detailed simulation takes too long (10^4-10^5 slower than realtime)
- Accuracy
  - Systems more complex
  - Vendors don’t reveal necessary details

Major Simulation Challenges

- **Multiple Objectives**
  - Performance used to be only criteria
  - Now, Energy, cost, power, reliability, etc...
- **Scale & Detail**
  - Many system characteristics require detail to measure
  - Detailed simulation takes too long (10^4-10^5 slower than realtime)
- **Accuracy**
  - Systems more complex
  - Vendors don't reveal necessary details



SST Simulation Project Overview

Goals

- Become the standard architectural simulation framework for HPC
- Be able to evaluate future systems on DOE workloads
- Use supercomputers to design supercomputers

Status

- Current Release (2.1) at code.google.com/p/sst-simulator/
- Includes parallel simulation core, configuration, power models, basic network and processor models, and interface to detailed memory model

Technical Approach

- Parallel
  - Parallel Discrete Event core with conservative optimization over MPI
- Holistic
  - Integrated Tech. Models for power
  - McPAT, Sim-Panalyzer
- Multiscale
  - Detailed and simple models for processor, network, and memory
- Open
  - Open Core, non viral, modular

Consortium

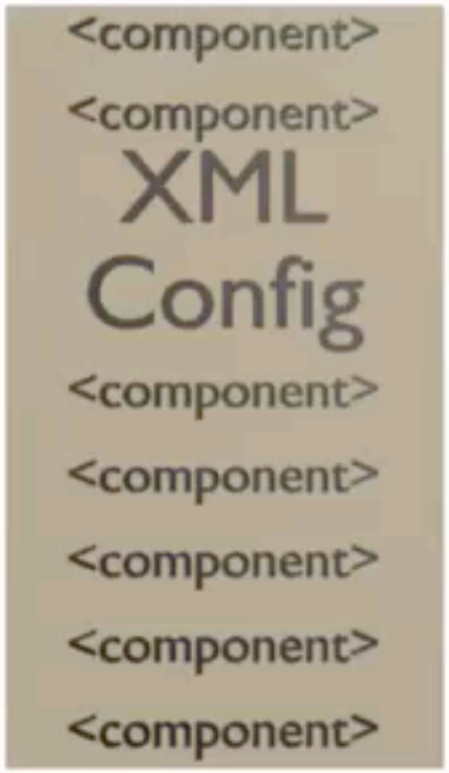
- "Best of Breed" simulation suite
- Combine Lab, academic, & industry

1. Parallel Implementation

- Implemented over MPI
- Configuration, partitioning, initialization handled by core
- Conservative, distance-based optimization

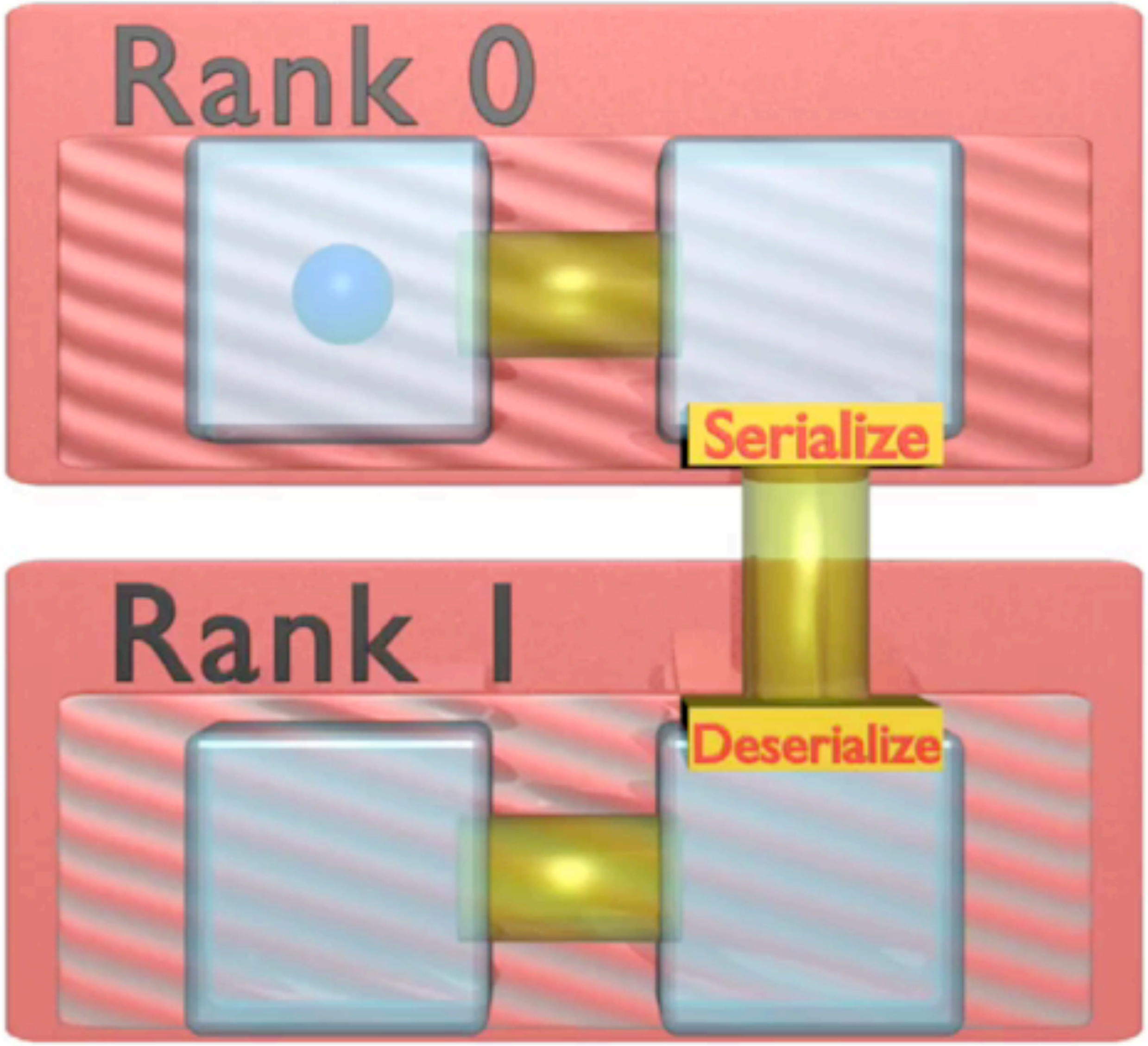
1. Parallel Implementation

- •Implemented over MPI
- •Configuration, partitioning, initialization handled by core
- •Conservative, distance-based optimization



Message Handling

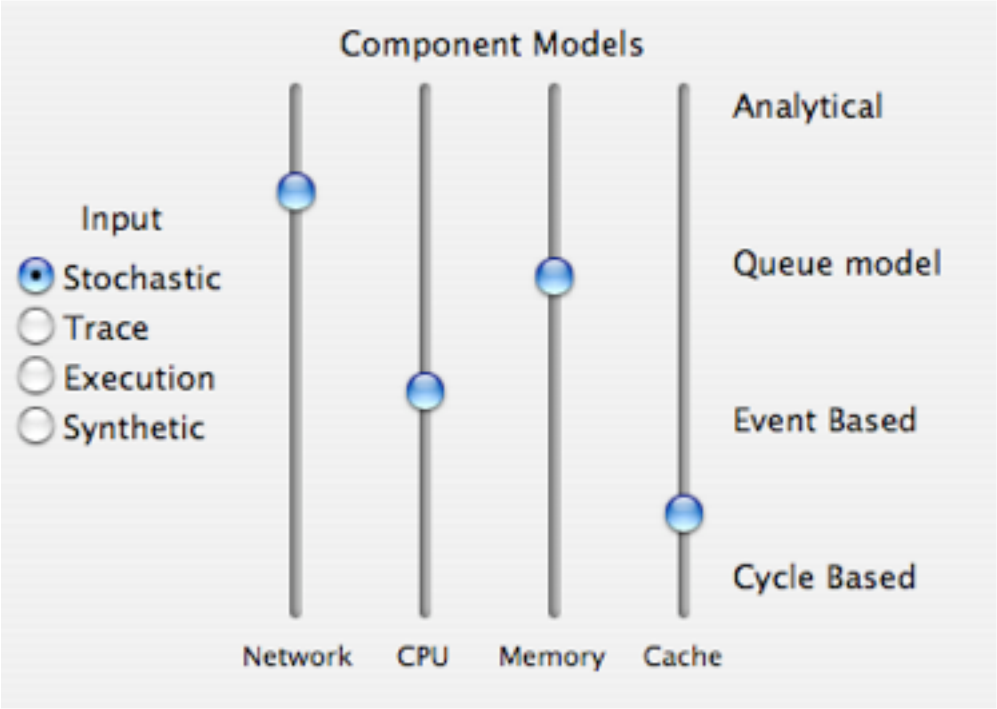
- SST core transparently handles message delivery
- Detects if destination is local or remote
- Local messages delivered to local queues
- Remote messages stored for later serialization and remote delivery
  - Boost Serialization Library used for message serialization
  - MPI used for transfer
- Ranks synchronize based on partitioning



2. Multi-Scale

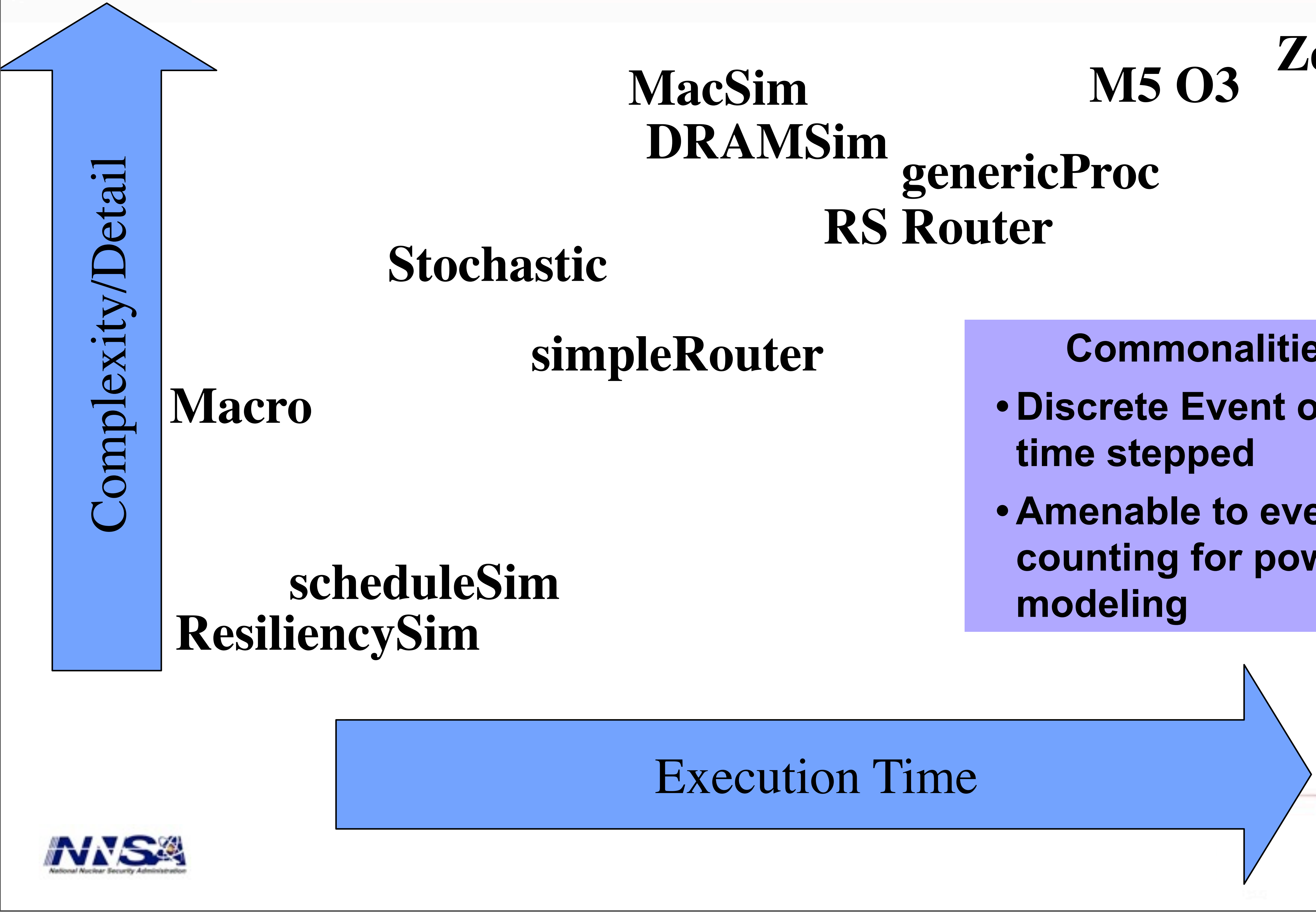
- Goal: Enable tradeoffs between accuracy, flexibility, and simulation speed
  - No single “right” way to simulate
  - Support multiple audiences
- High- & Low-level interfaces
  - Allows multiple input types
  - Allows multiple input sources
    - Traces, stochastic, state-machines, execution...

	High-Level	Low-Level
Detail	Message	Instruction
Fundamental Objects	Message, Compute block, Process	Instruction, Thread
Static Generation	MPI Traces, MA Traces	Instruction Trace
Dynamic Generation	State Machine	Execution


Input
● Stochastic
○ Trace
○ Execution
○ Synthetic

Multiscale Parameters

Diversity!



Zesto

M5 O3

MacSim

DRAMSim

genericProc

RS Router

Stochastic

simpleRouter

Macro

scheduleSim

ResiliencySim

Complexity/Detail

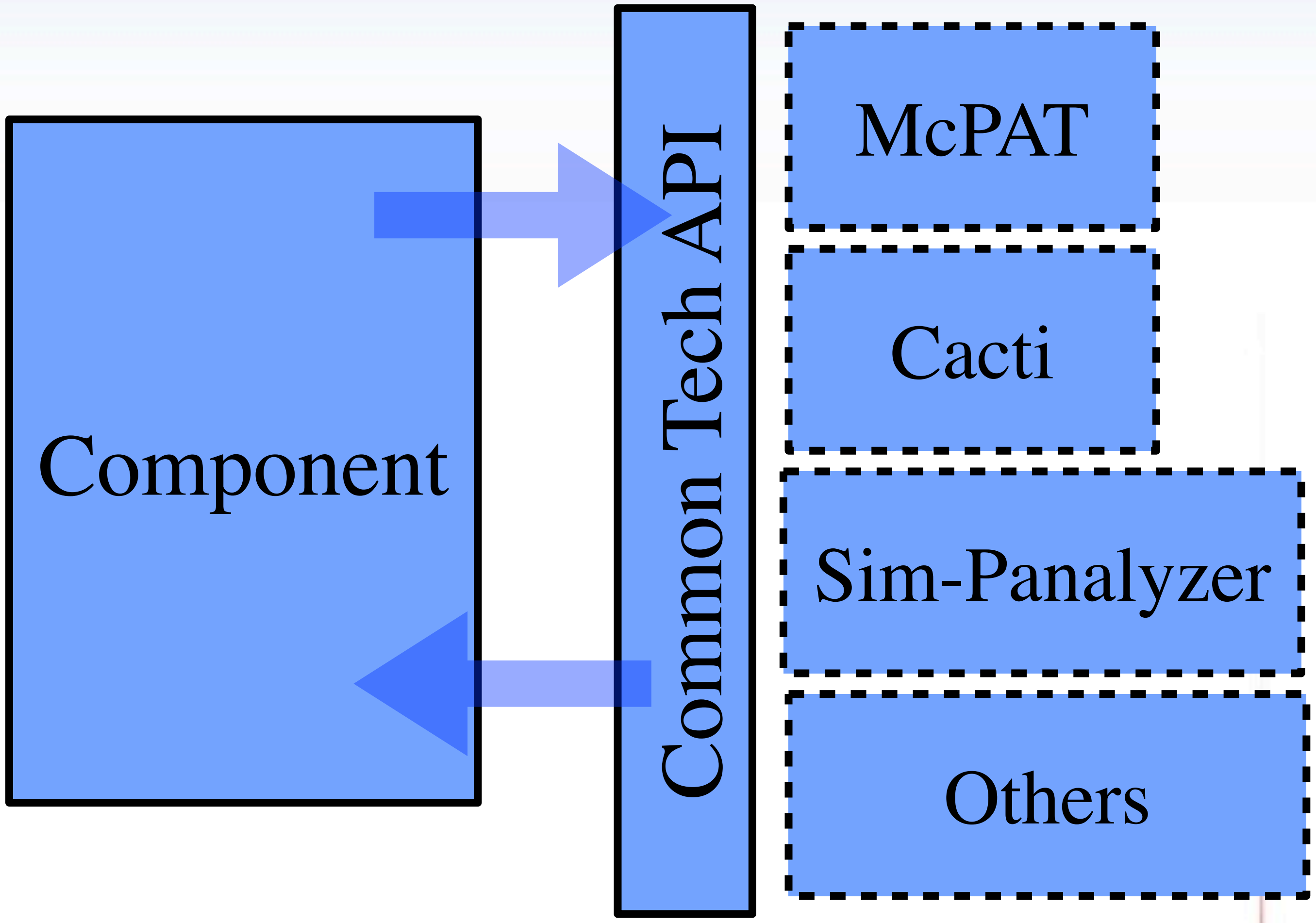
Execution Time

Commonalities

- Discrete Event or time stepped
- Amenable to event counting for power modeling

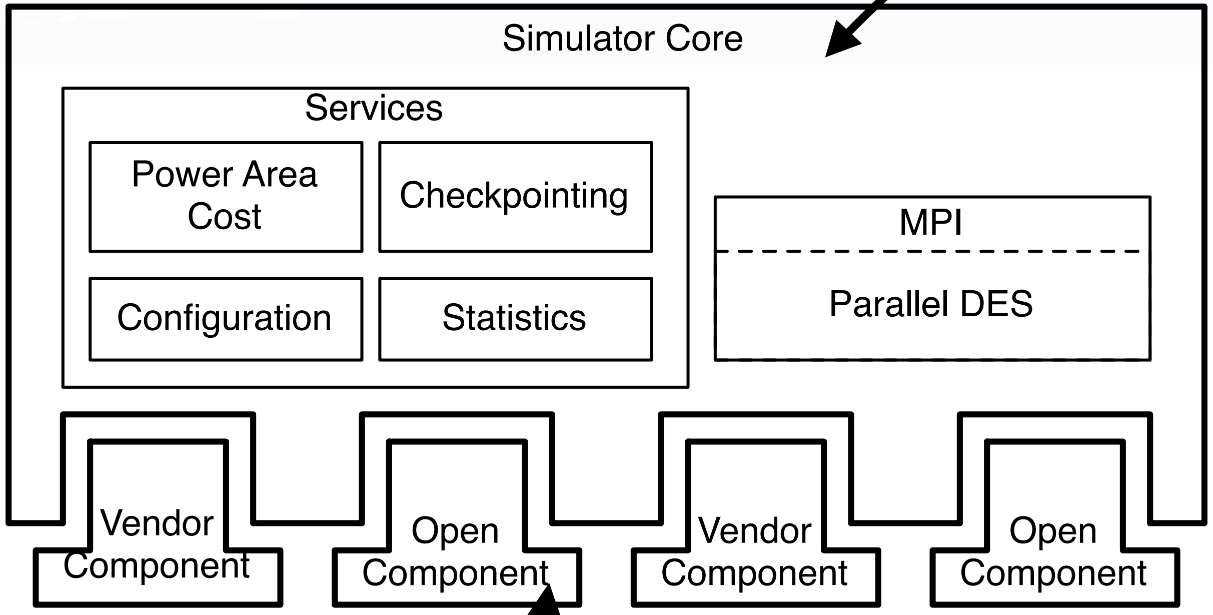
3. Holistic Simulation

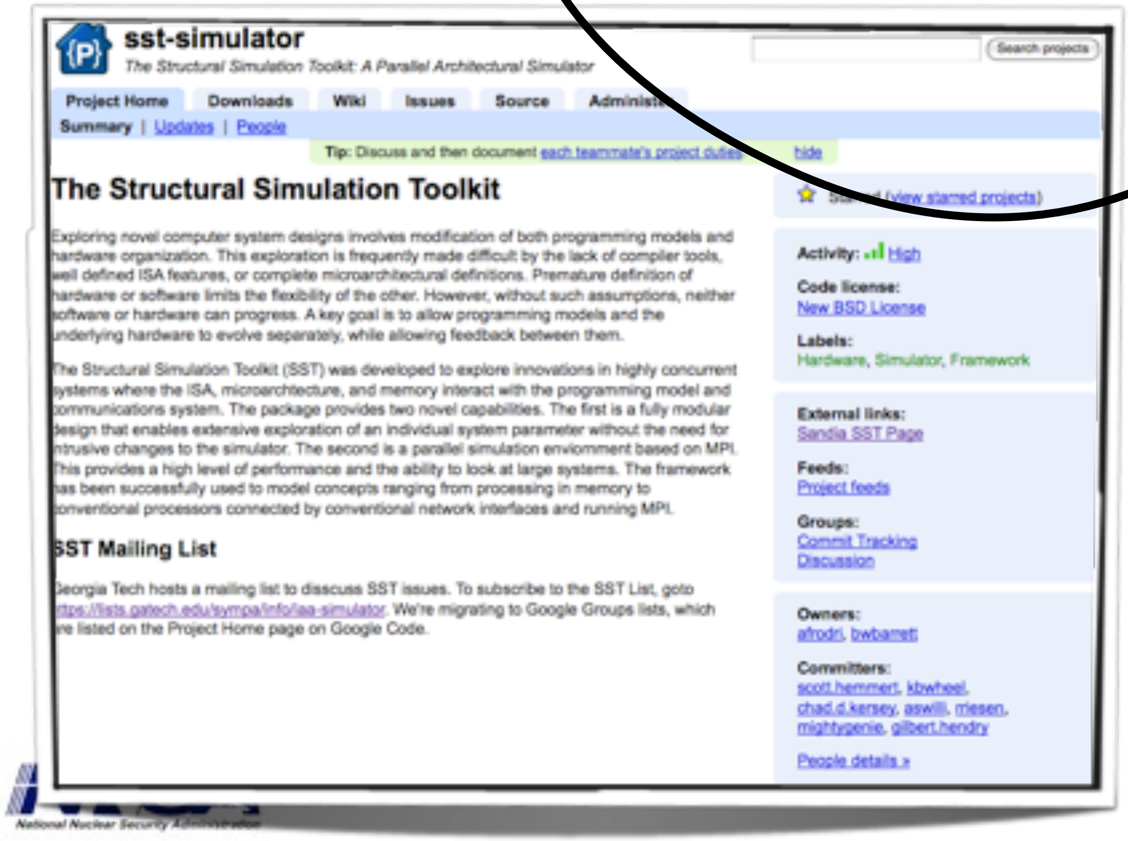
ApplicationScaling
NRE
NewApplicationSpaces
Reliability
Resiliency
MarketSize
ThermalEffects
Energy
Footprint
Power
Noise
Programmability
Mass
Cooling
Cost
OS



- Design space includes much more than simple performance
- Create common interface to multiple technology libraries
  - Power/Energy
  - Area/Timing estimation
- Make it easier for components to model technology parameters

4. Open Simulator Framework





The Structural Simulation Toolkit: A Parallel Architectural Simulator

Tip: Discuss and then document each teammate's project duties

The Structural Simulation Toolkit

Exploring novel computer system designs involves modification of both programming models and hardware organization. This exploration is frequently made difficult by the lack of compiler tools, well defined ISA features, or complete microarchitectural definitions. Premature definition of hardware or software limits the flexibility of the other. However, without such assumptions, neither software or hardware can progress. A key goal is to allow programming models and the underlying hardware to evolve separately, while allowing feedback between them.

The Structural Simulation Toolkit (SST) was developed to explore innovations in highly concurrent systems where the ISA, microarchitecture, and memory interact with the programming model and communications system. The package provides two novel capabilities. The first is a fully modular design that enables extensive exploration of an individual system parameter without the need for intrusive changes to the simulator. The second is a parallel simulation enviornment based on MPI. This provides a high level of performance and the ability to look at large systems. The framework has been successfully used to model concepts ranging from processing in memory to conventional processors connected by conventional network interfaces and running MPI.

SST Mailing List

Georgia Tech hosts a mailing list to disscuss SST issues. To subscribe to the SST List, goto https://lists.gatech.edu/sympa/info/aa-simulator. We're migrating to Google Groups lists, which are listed on the Project Home page on Google Code.

Activity: High

Code license:
New BSD License

Labels:
Hardware, Simulator, Framework

External links:
Sandia SST Page

Feeds:
Project feeds

Groups:
Commit Tracking
Discussion

Owners:
afrodri, bwbarrett

Committers:
scott.hemmert, kbwheel, chad.d.kersey, aswill, rriesen, mightygenie, gilbert.hendry

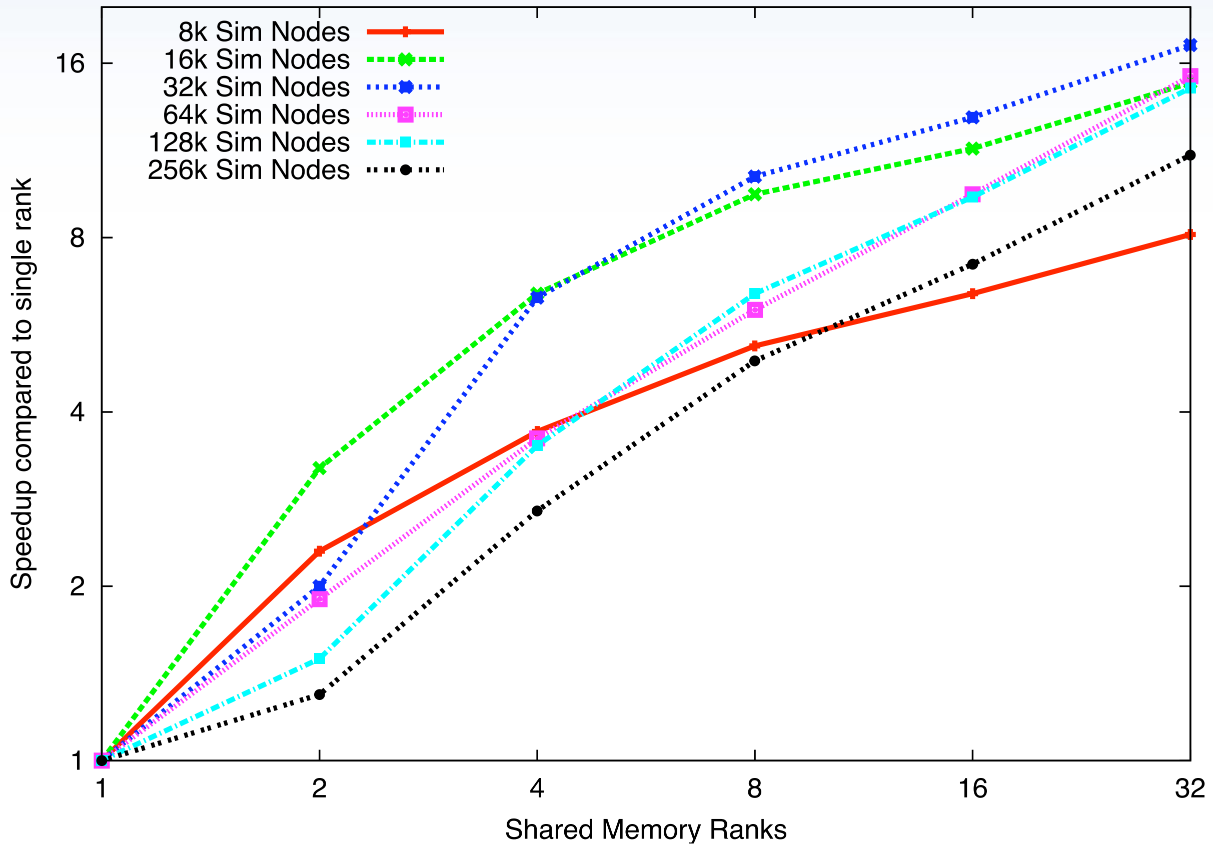
People details »

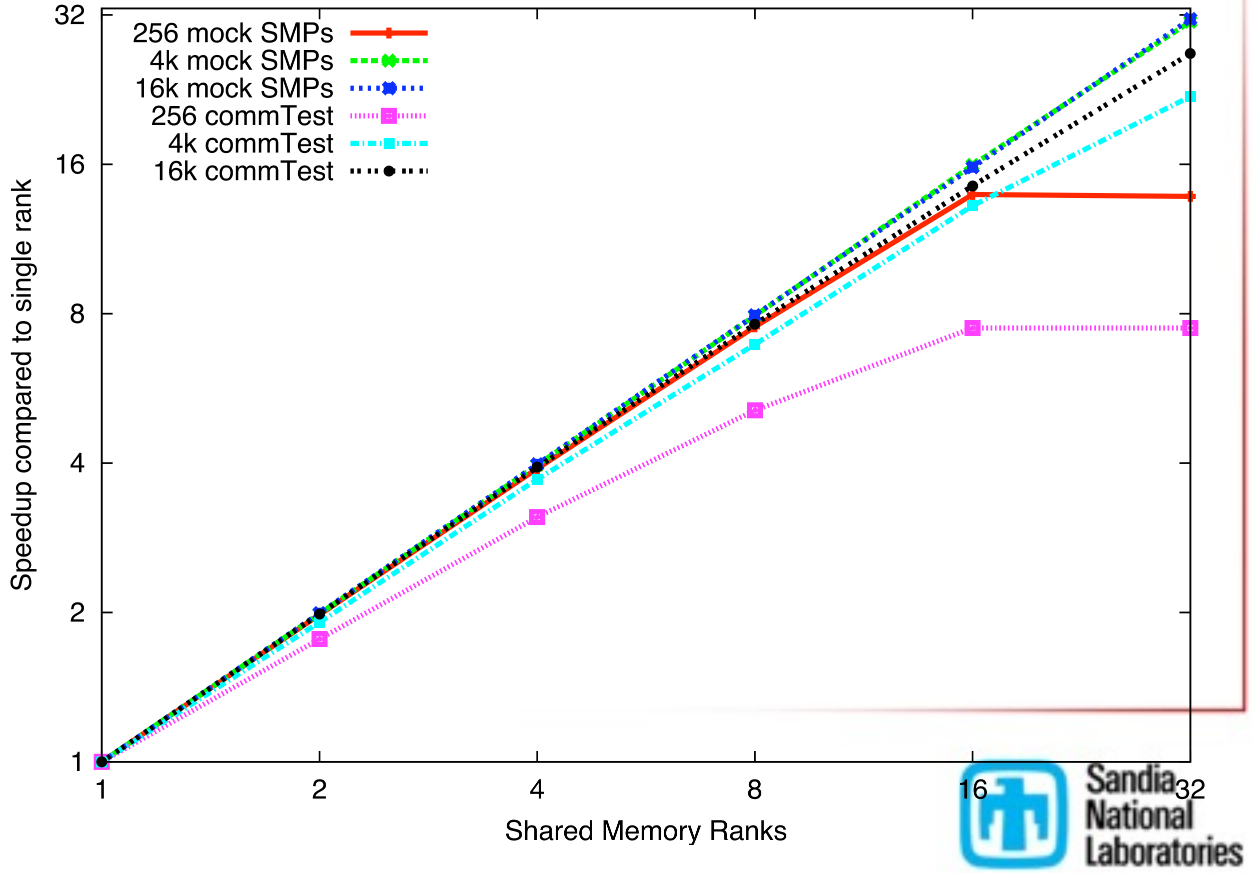
- Simulator Core will provide...
  - Power, Area, Cost modeling
  - Checkpointing
  - Configuration
  - Parallel Component-Based Discrete Event Simulation
- Components
  - Ships with basic set of open components
  - Industry can plug in their own models
    - Under no obligation to share
- Open Source (BSD-like) license
- SVN hosted on Google Code

Improvements

Scalability

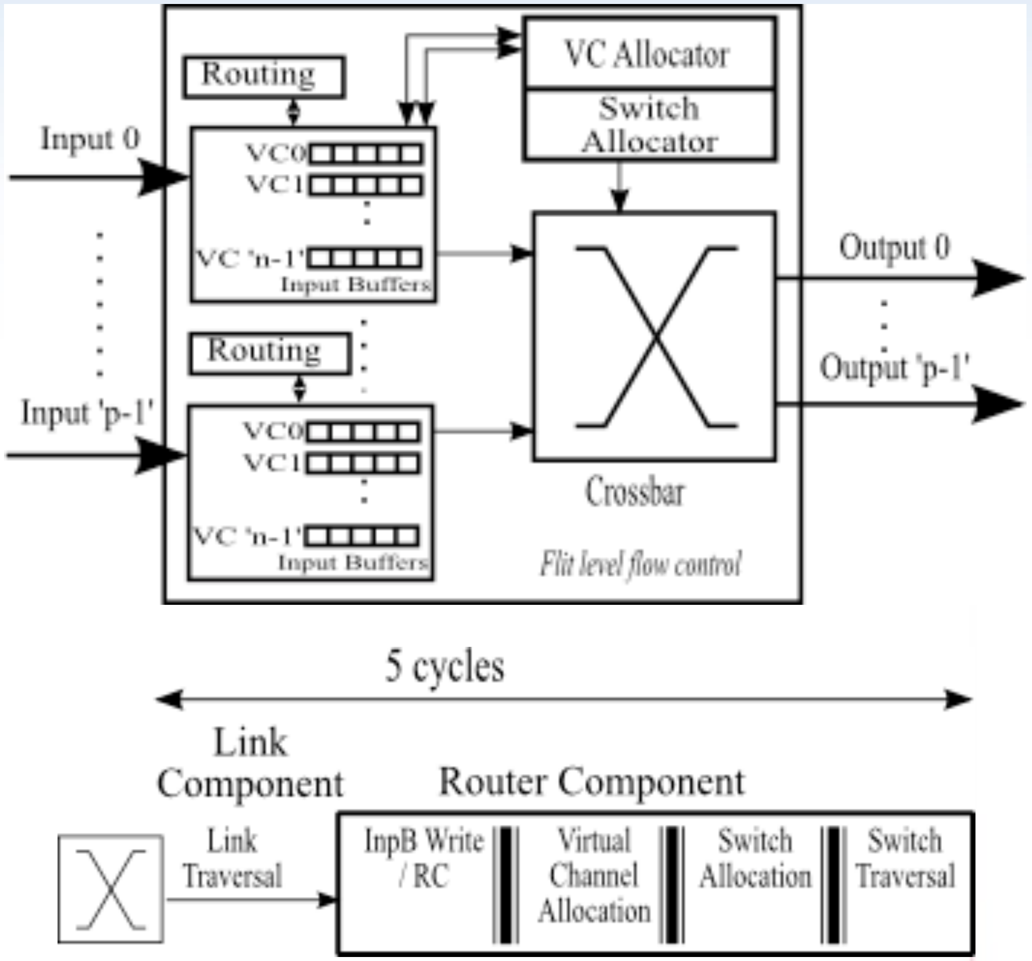
- Test setup: OpenMPI 1.4.3 Shared memory
- Portals NIC Simulation
  - State machine application, detailed NIC & Router
  - Best scaling w/ 32K simulated nodes (17.2x speed up on 32 host nodes)
  - May be limited by memory bandwidth
- Synthetic components
  - Poor scaling with 256 simulated SMP nodes
  - 85%+ scaling efficiency with 4K nodes

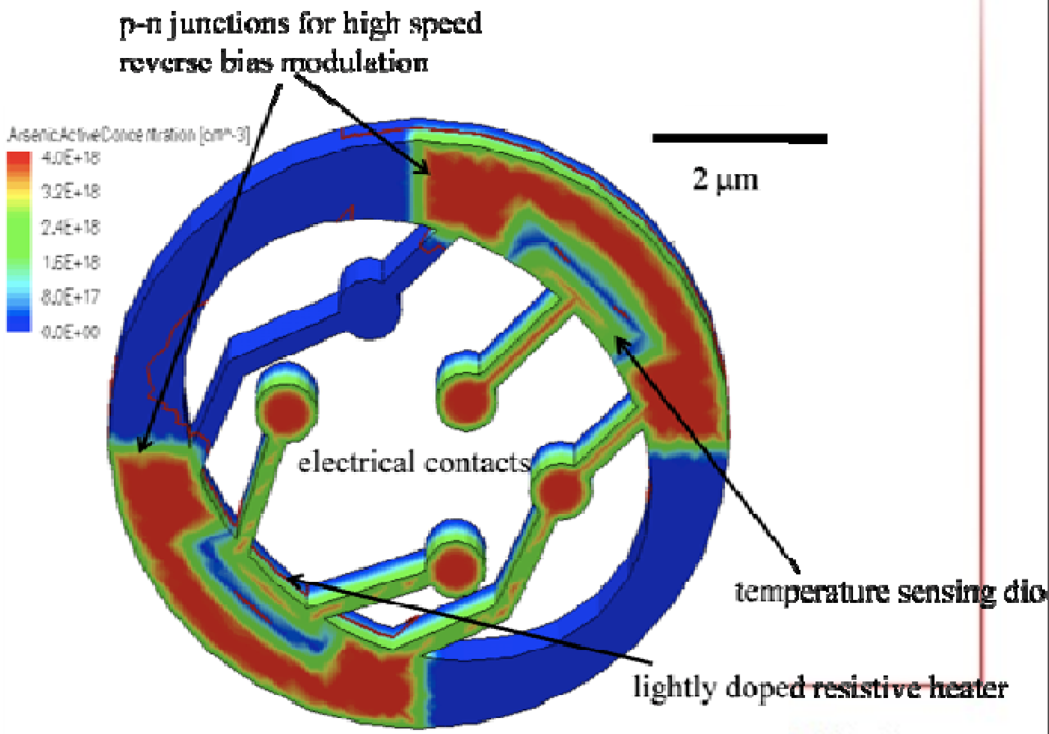




IRIS & PhoenixSim

- IRIS
  - Simulates pipelined, cycle-accurate router
  - Models variety of Network-on-Chip (NoC) and inter-node interconnects
  - Parameterized buffer sizes, virtual channels, routing & arbitration
  - Modular architecture
- Phoenix Sim
  - Physically detailed photonic network simulator
  - Component: waveguides, modulators, detectors, filters, switches, lateral couplers, and lasers

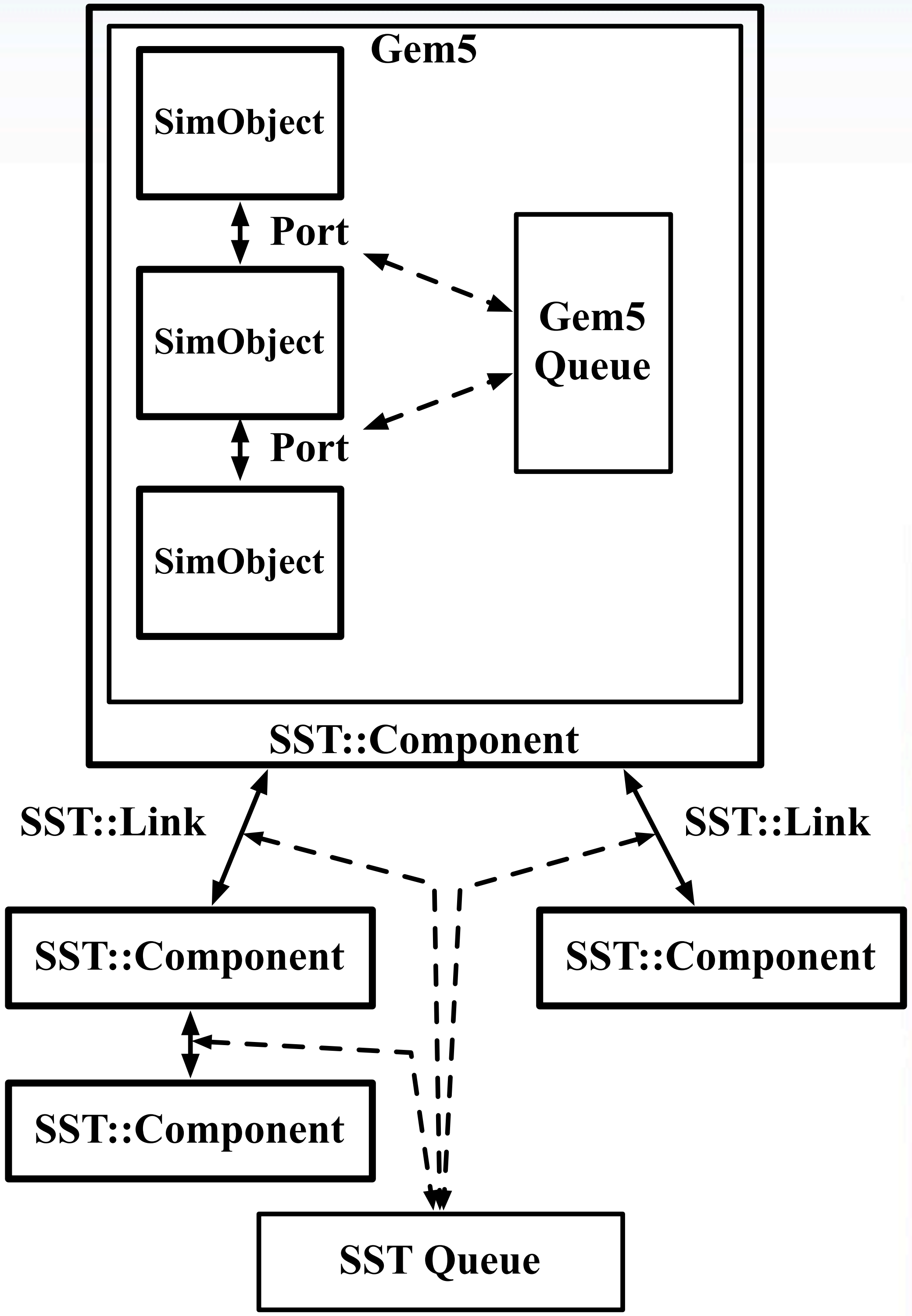




p-n junctions for high speed reverse bias modulation
ArsenicActiveConcentration [cm^-3]
4.0E+18
3.2E+18
2.4E+18
1.6E+18
8.0E+17
0.0E+00
2 µm
electrical contacts
temperature sensing dio
lightly doped resistive heater

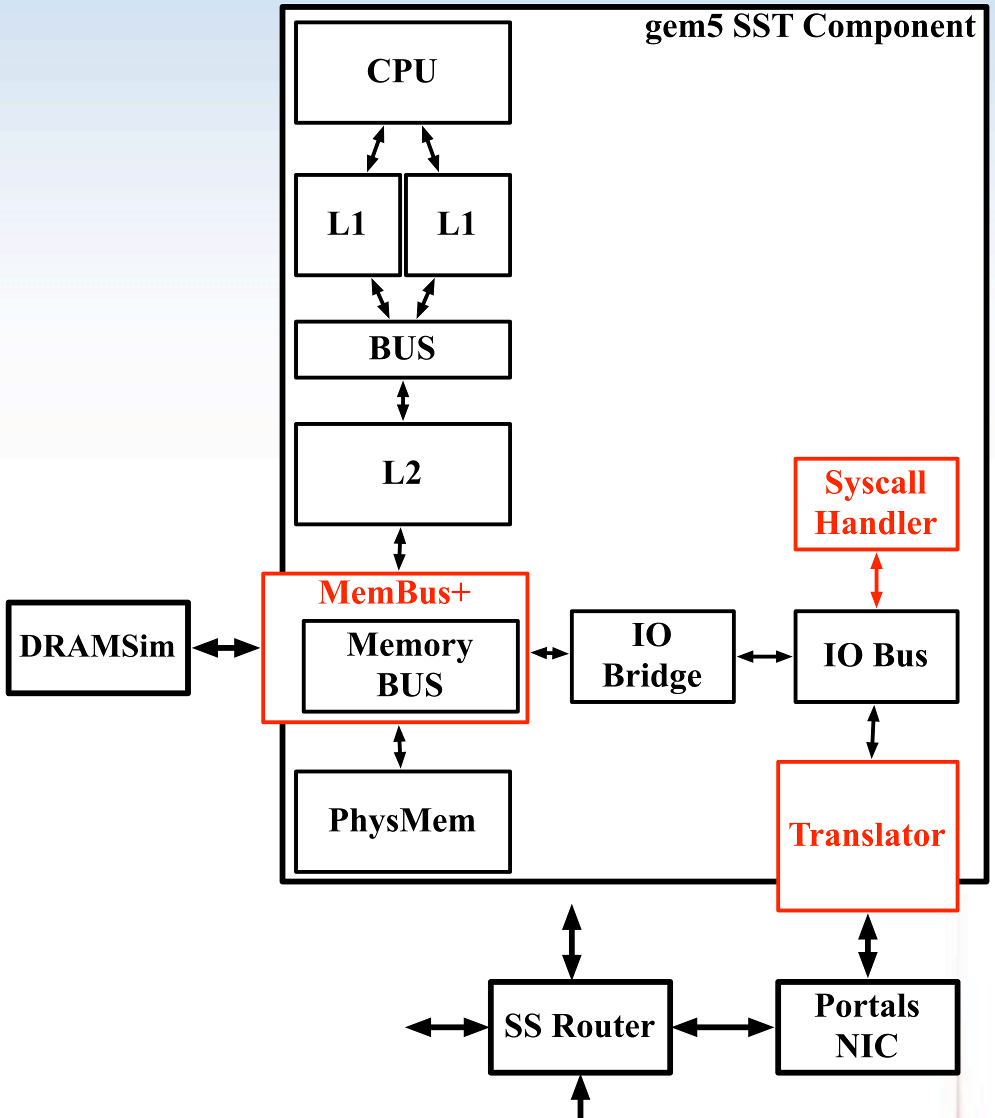
GeM5

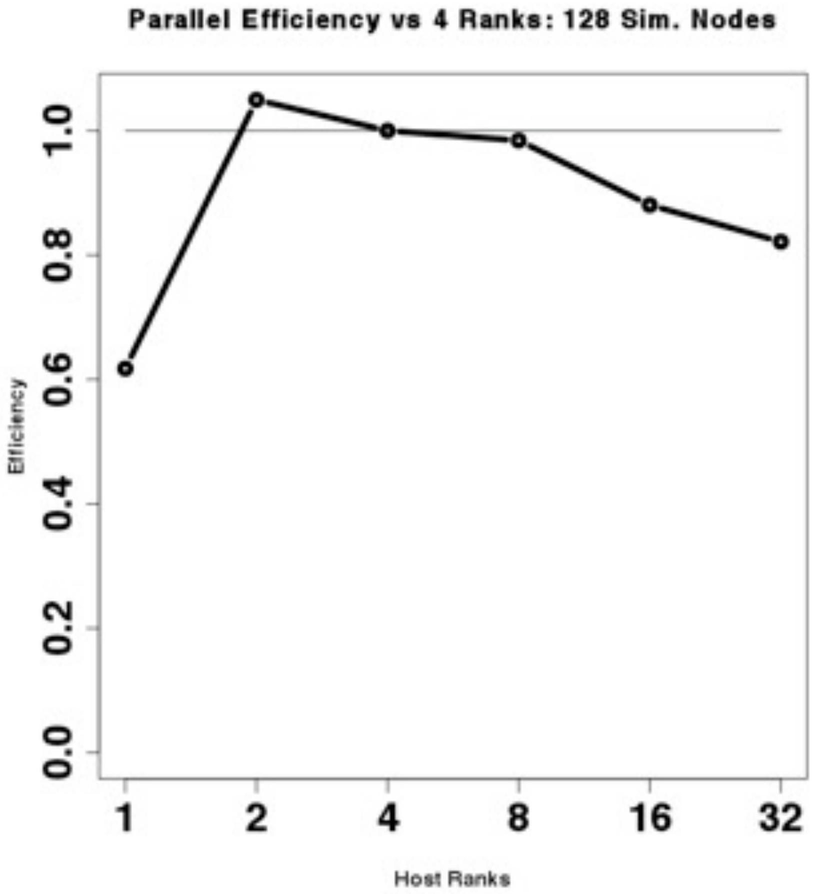
- M5: Modular platform for computer system architecture research, encompassing system-level architecture as well as processor microarchitecture.
- Provides detailed, full-system CPU models for x86, ARM, SPARC, Alpha
- Integrated at SST Component, allows interaction with SST models, and parallel execution
- Currently tested up to 256 nodes.



SST/GeM5 Integration

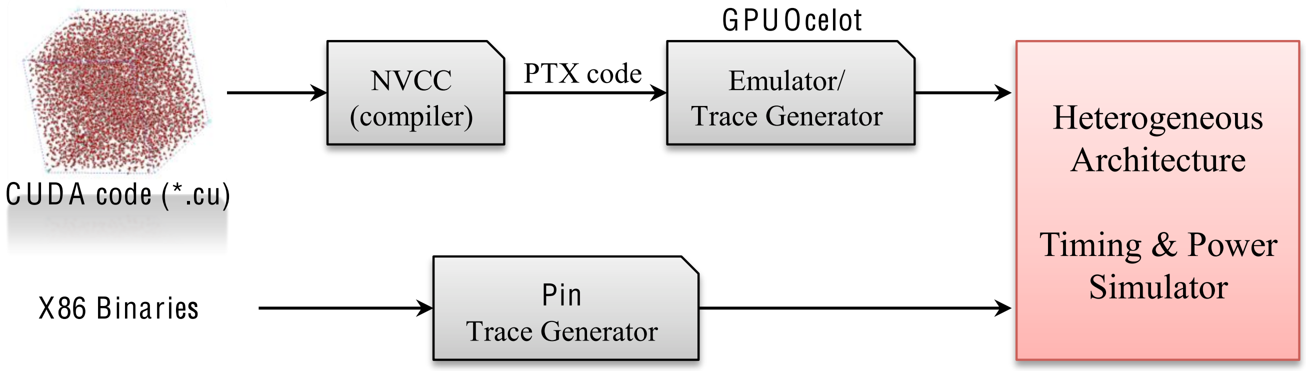
- Goals: Run GeM5 in parallel, connect with other SST components
- High parallel efficiency
- Changes
  - Replaced Python-based configuration w/ XML or C++-based system
  - Encapsulated GeM5 as an SST Component, GeM5 event Q driven by SST clock
  - Created translator SimObjects to connect to SST links
  - Changes made to GeM5 loader to avoid use of async (untimed out-of-band) messages





MacSim GPU Simulator

- Trace driven simulator
  - –Internal RISC-like uops, with threading information
  - –X86 Captured w/ PIN tool
  - –CUDA Captured through Ocelot emulation tool
- Models for
  - –OoO & GPU-like cores
  - –Cache memory (inc. texture & Constant caches)
  - –Timing & Power
- Allows ties in to MacPAT, thermal modeling, RAMP reliability model



CUDA code (*.cu)

NVCC (compiler)

PTX code

GPUOcelot

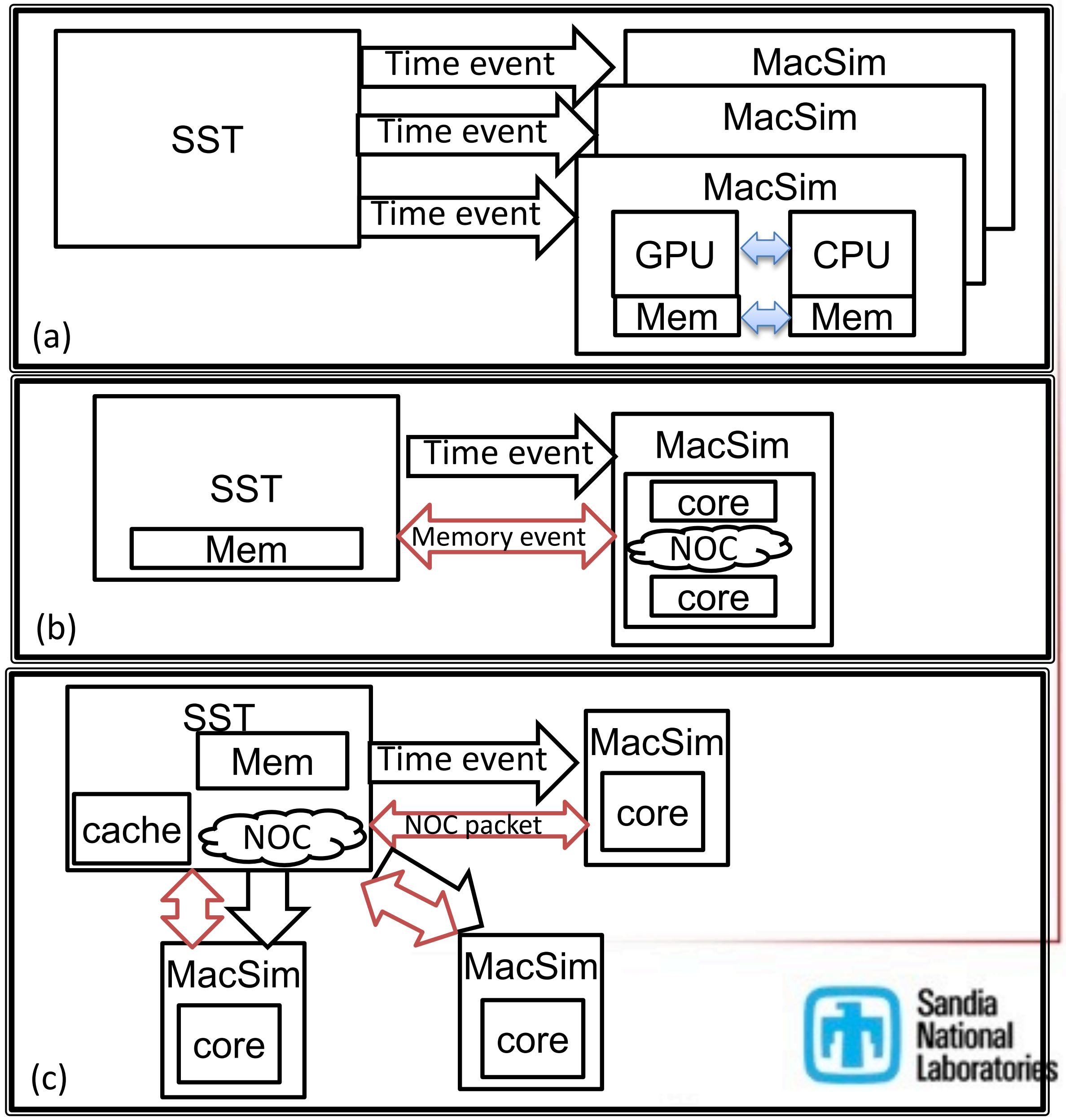
Emulator/ Trace Generator

X86 Binaries

Pin Trace Generator

Heterogeneous Architecture

Timing & Power Simulator



(a)

SST

Time event

Time event

Time event

MacSim

MacSim

MacSim

GPU

CPU

Mem

Mem

(b)

SST

Mem

Time event

Memory event

MacSim

core

NOC

core

(c)

SST

Mem

cache

NOC

Time event

NOC packet

MacSim

core

MacSim

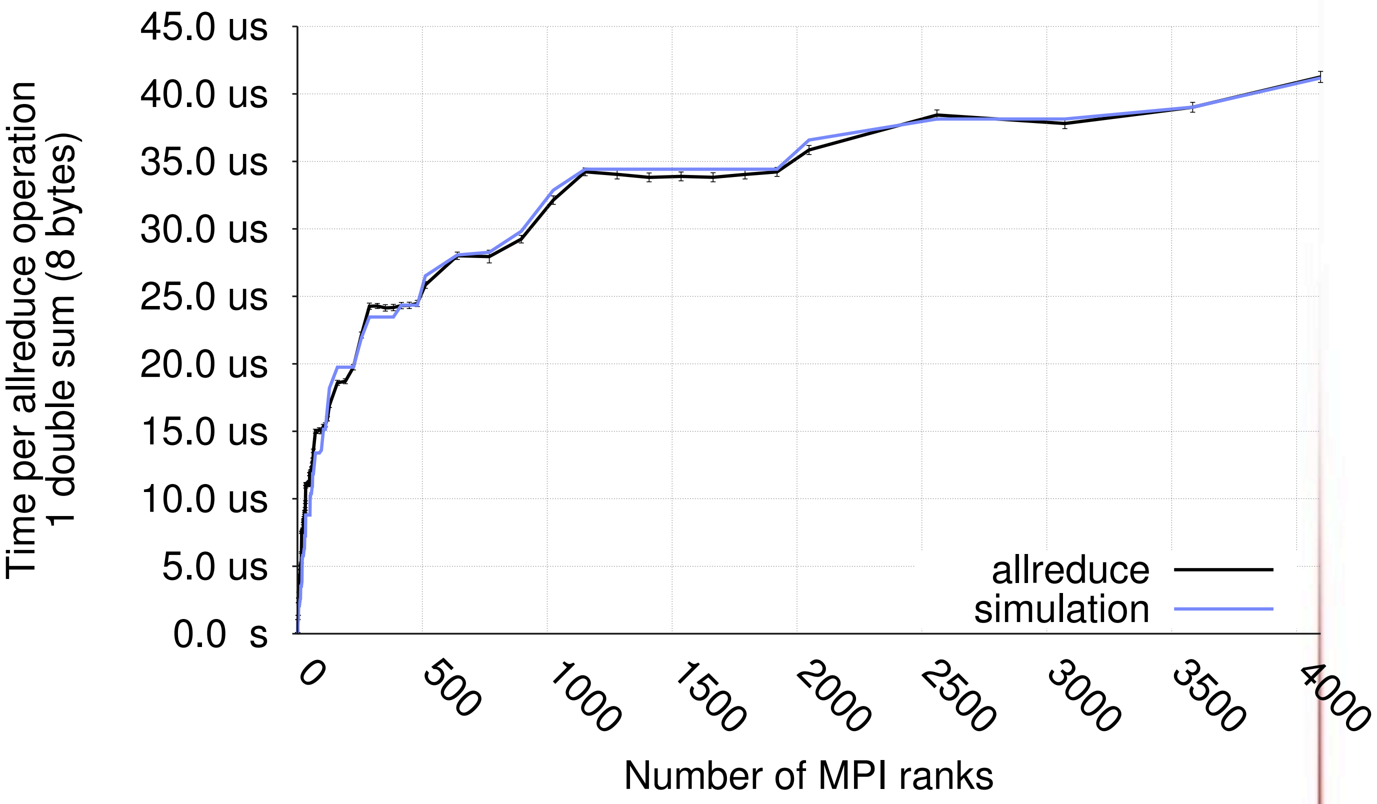
core

MacSim

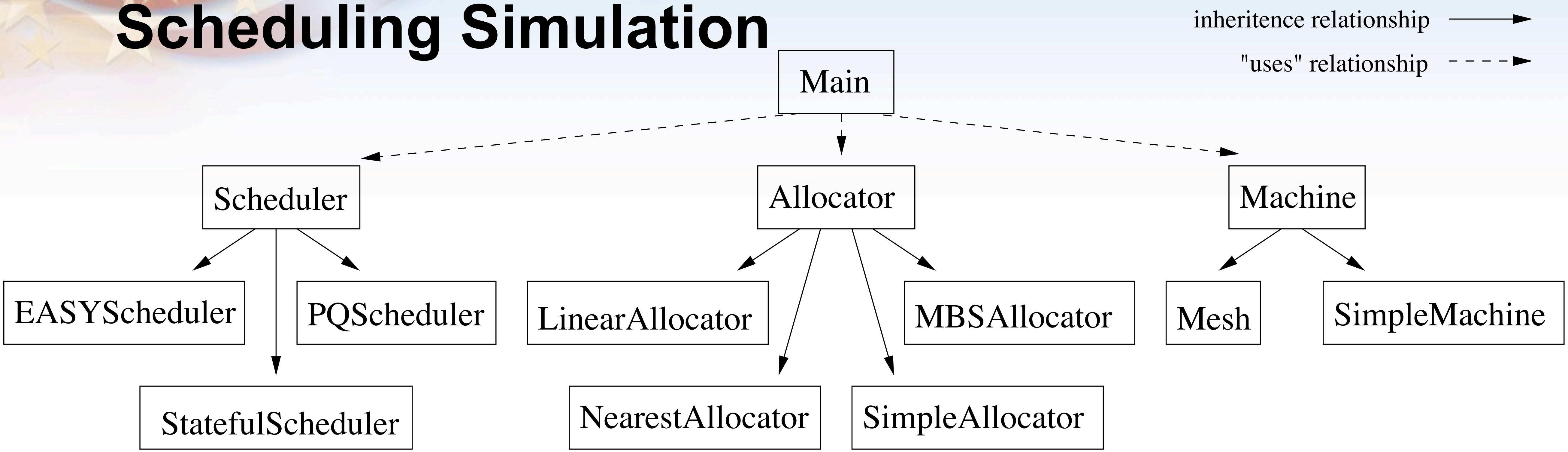
core

Additional Components

- BOBSim
  - –Models Buffer-on-Board Memory systems
  - –Hardware verified using RTL from Micron
  - –Includes detailed power model
- Reliability Simulation
  - –extend SSTfor simulating resilience mechanisms of a large-scale system (e.g. Checkpointing, use of NVRAM, burst buffers)
  - –Models: Router, Storage, Node
  - –Validated against large systems



Scheduling Simulation



- new high-level system simulation capabilities
- using job traces from large systems
- simulates the decisions being made
  - scheduling
  - allocation
- working on task mapping and tighter integration with the lower-level simulations that SST is already able to do

Summary

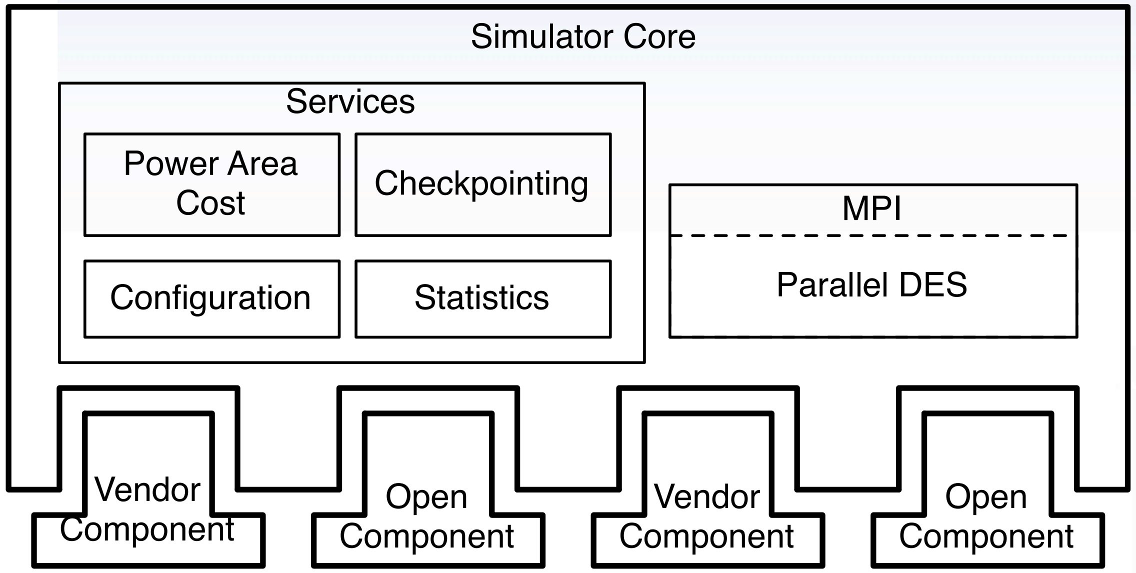
- The SST is parallel, scalable, modular, and open
- Recent Improvements
  - Scalability
  - Processor, network, memory, IO system models
- Future
  - Improving component interoperability
  - Validation
    - Emphasis on mixed-resolution simulation
  - Scaling
    - Emphasis on distributed memory systems
  - New models
    - Stacked memory
    - Exascale Supercomputers

Questions?

code.google.com/p/sst-simulator/

Component Library

- Parallel Core v2
  - Parallel DES layered on MPI
  - Partitioning
  - Configuration & Checkpointing
  - Power modeling
- Technology Models
  - McPAT, Sim-Panalyzer, IntSim, Orion and custom power/energy models
  - HotSpot Thermal model
  - Working on reliability models
- Components
  - Processor: Macro Applications, Macro Network, NMSU, genericProc, state-machine, Zesto, GeM5, GPGPU
  - Network: Red Storm, simpleRouter, GeM5
  - Memory: DRAMSim II, Adv. Memory, Flash, SSD, DiskSim



SST Simulator Core