

Modeling and Simulation: Providing Answers

DOE MODSIM 2015

Dan Ernst
Advanced Technology Development
Cray, Inc.



Using Modeling & Simulation at Cray

- **Focus is on finding answers to specific questions**
- **Only answer the questions required**
 - One-off tools, some public tools, spreadsheet models
- **Constant stream of questions in an area → Invest in tools**
 - Application characterization / Node interaction
 - Network traffic modeling
 - Infrastructure models
- **Not limited to “Research”**



Generic Node Simulation

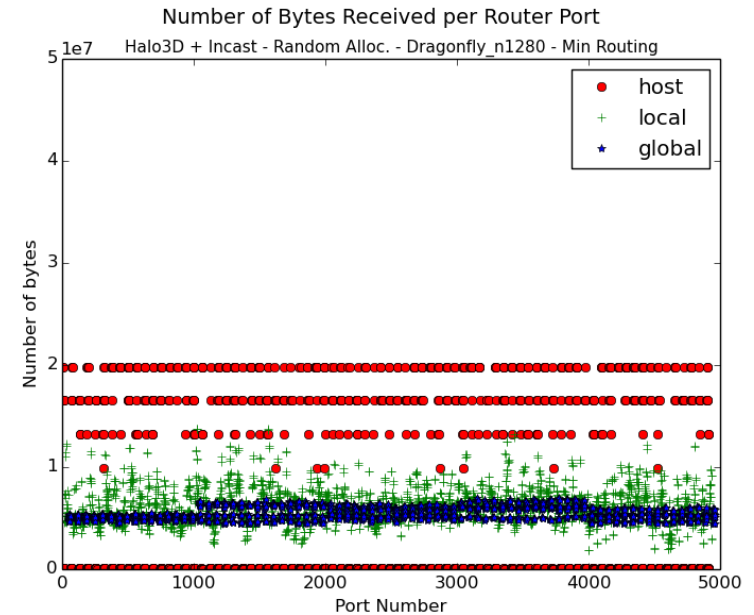
- **Cray has internal simulators that model a *generic* microarchitecture.**
 - Focus is on creating a performance profile for an application as quickly as possible (10-100 MIPS)
- **Gain behavioral understanding of application performance by capturing:**
 - Processor instruction stream behaviors
 - Memory system patterns and behaviors
- **Infrastructure heavily used by our performance team**
 - Also used for R&D purposes (FF2)

Network Modeling with SST

- Cray adopted SST as our simulation platform part way through our DesignForward program

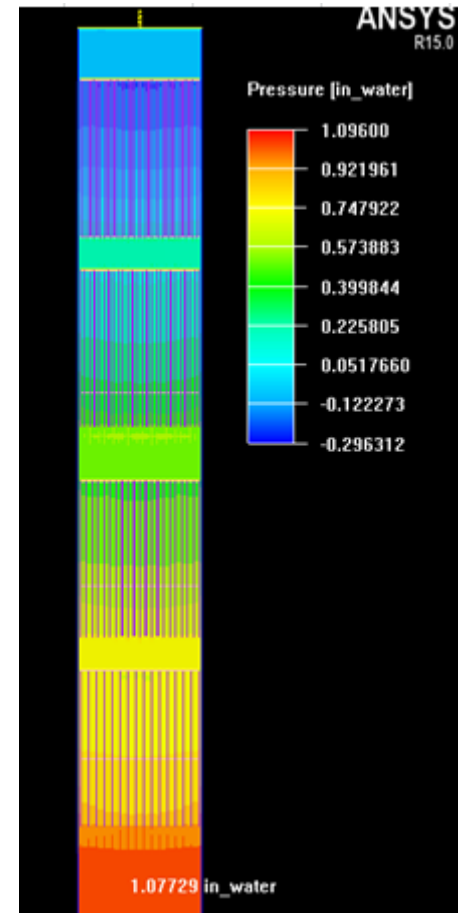
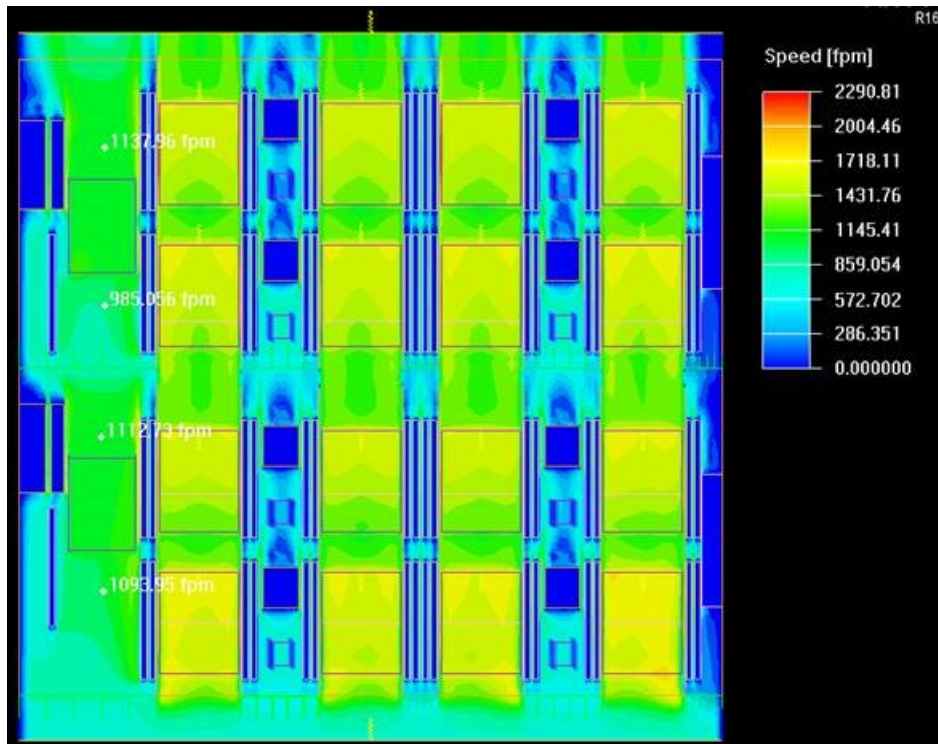
- We are using SST for internal system architecture investigation work
 - Using SST to study mixed workloads
 - The impact of one application on another
 - The impact of I/O traffic on applications
 - Dependencies on how jobs are distributed over nodes

- Working with Sandia
 - Enhancements and bug fixes are being returned to Sandia
 - Developing interfaces that allow Cray to use our own models



Infrastructure Modeling

- **CFD analysis of cooling solutions**
 - ANSYS Icepak
 - Examine airspeed, pressure, temperature, etc.



COMPUTE | STORE | ANALYZE

Evaluating Tools

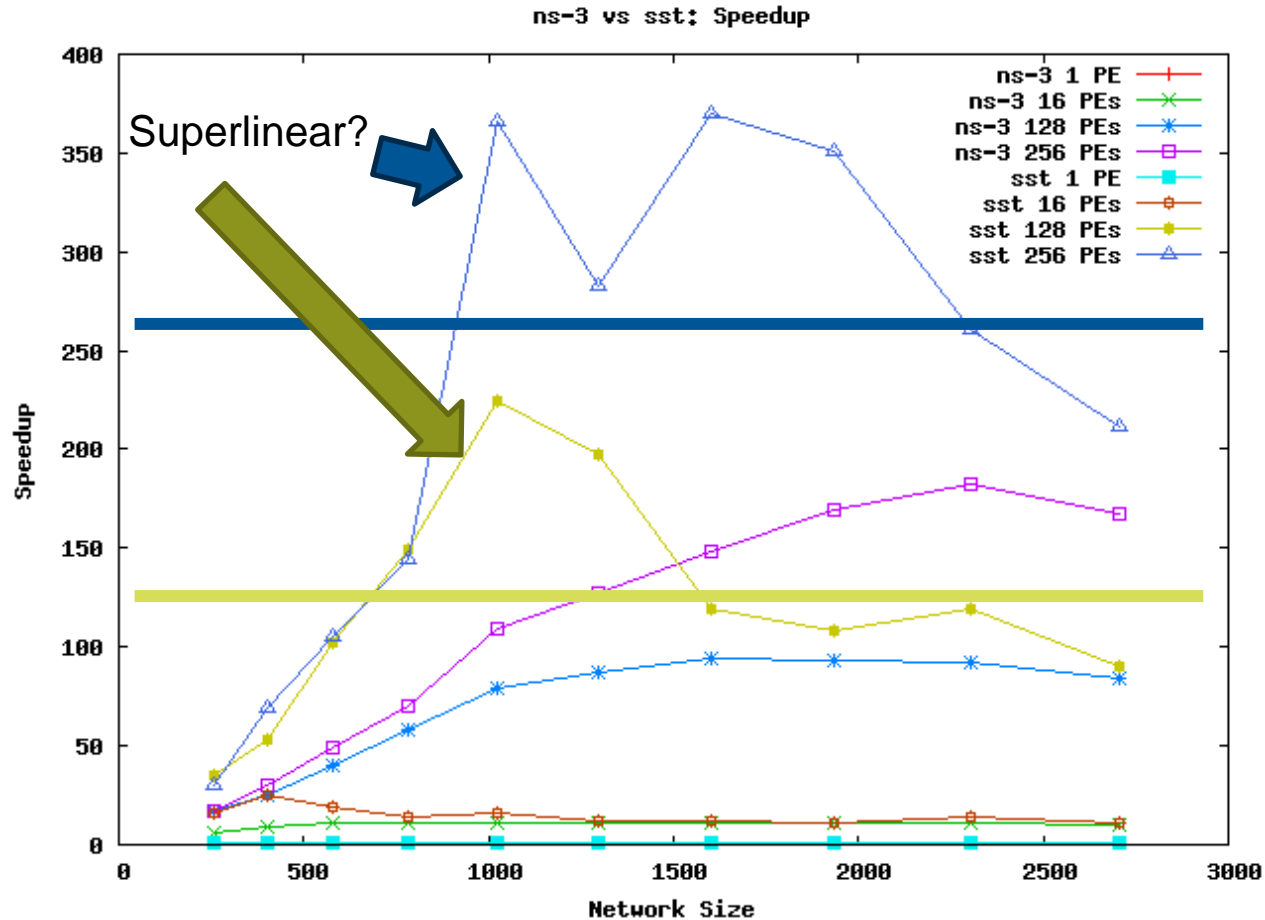
- **Do only what's needed**

- Not excessively complicated / flexible
- Positive implications on speed to completion

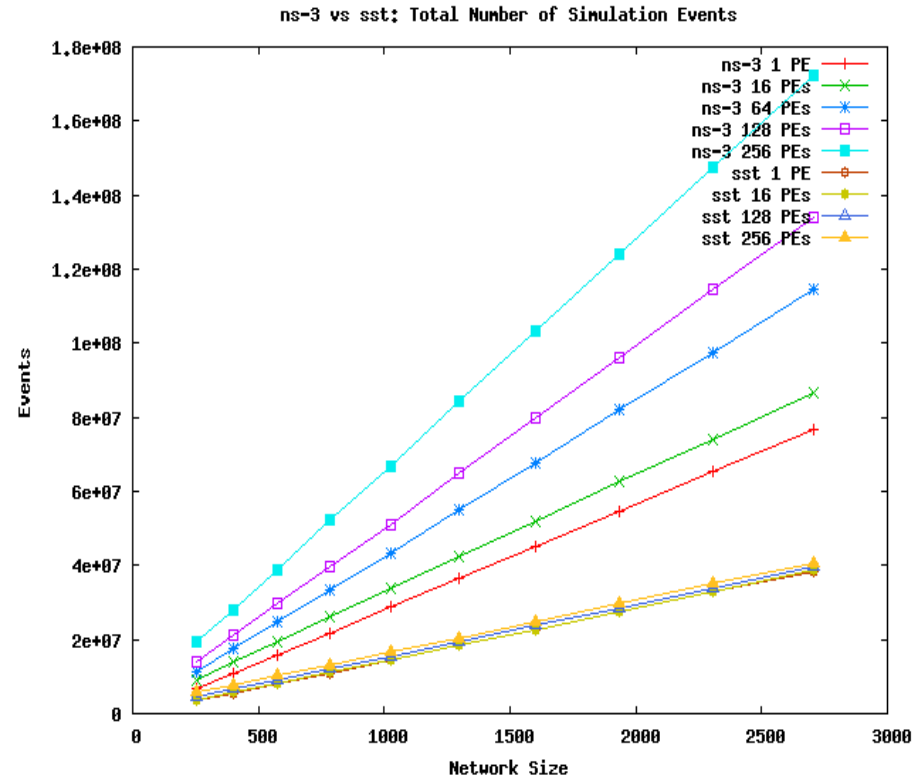
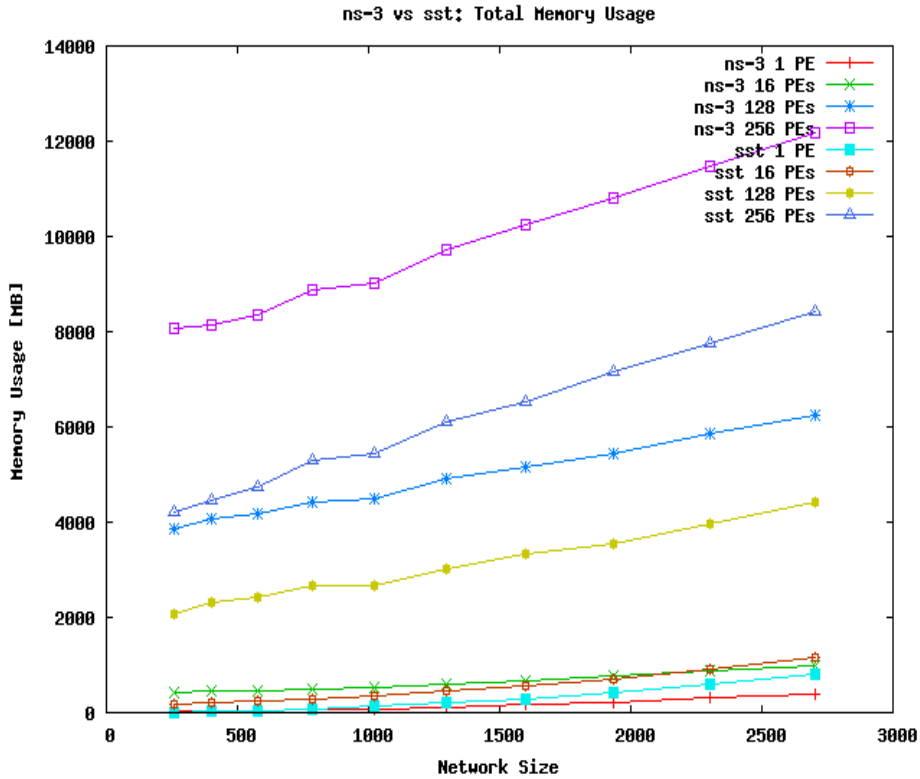
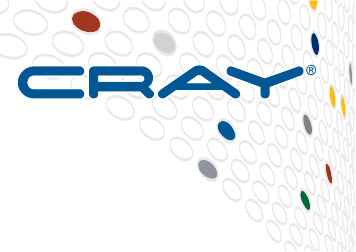
- **Scalability**

- Many of our questions are at macro scale
 - Full node (many cores) + memory hierarchy
 - Large network configurations
- Lots of detail limits the exploration space due to long runtimes

SST/ns-3 – Scalability (Speedup)



SST/ns-3 – Memory Usage & Simulation Events



COMPUTE | STORE | ANALYZE

The Importance of Accuracy (and Validation)



Estimation Accuracy:
Estimates compared to Initial Results on Delivered Hardware

- **Being wrong costs us money**
 - Cliff to the left
 - Slope to the right

Gaps?

- **There is no grand unified simulator that simulates full node details and network together**
 - (at a large scale and reasonable runtime)

What can the community do to address this?

- ***Please don't waste your time writing one!***
 - See previous notes about excessive complexity...