

B-1 Modeling/Benchmarking/Simulation

DOE HPC Operations Review
San Francisco, November 5-6, 2013

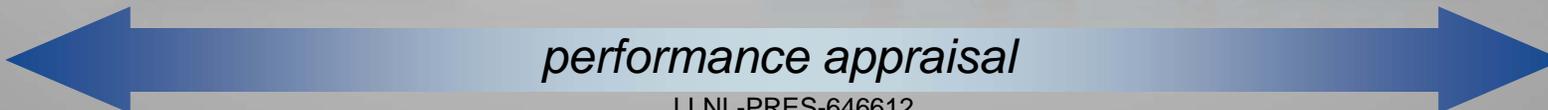


**Modeling &
Simulation**

Benchmarks

Proxy Apps

Real Apps



Breakout participants

- Brian Austin (LBL)
- Bob Ballance* (SNL)
- Scott Futral (LLNL)
- Wayne Joubert (ORNL)
- Kaki Kelly* (LANL)
- Jiayuan Meng (ANL)
- Dino Pavlakos (SNL)
- John Shalf (LBL)
- Harvey Wasserman (LBL)

* Denotes breakout session lead

Processes (scope of activity)

The usage of modeling, simulation, benchmarking is to:

- Provide outside entities with information about the workload
- Inform selection process
 - Set team expectations
 - Evaluate proposals
 - Manage risk that the system might be performant
 - Understand the effort required to port/tune their applications
- Help application developers know how/where to tune code
- Ensure that the delivered system operates at the level desired throughout the lifetime of the system
- Assist end users in understanding how to configure/parameterize their production runs
- Help vendors build better systems for DOE's needs

Processes (scope of activity), cont.

- What begins first: timeline for activities (before or after hardware)?
 - WRT procurement, most benchmark activities are pre-RFP
 - Develop/find representative benchmarks or models
 - Key difference: when are the acceptance criteria finalized?
 - When do benchmark requirements come into play?
 - Role of "market basket" approach

Processes, cont.

- Ongoing interaction with application teams
 - Gain understanding of customer need
 - Application characteristics
 - Structure of application
 - Scale
 - Time-varying nature of workload is a complication
- Ongoing benchmarking is a (system) lifetime activity

Processes, cont.

- What is the role of early hardware access (either locally or remotely) and prototype systems?

- **Vital**

Processes (scope of activity), cont.

- What is the role of vendor partnerships?
 - Access to HW simulators
 - Access to modeling information
 - Assistance in porting/optimizing benchmarks
 - Engineering samples

Processes (scope of activity), cont.

- What are the roles of research and design and engineering?
 - Researchers often lead the development of new benchmarks, modeling techniques, and models
 - There is a tight connection between system design features and the requirements for new benchmarks
 - E.g. Multicore, stacked memory, accelerators, etc.

Processes (scope of activity), cont.

- What resiliency activities are executed (for example, redundancy)
 - Benchmarks are often needed to wring out issues
 - Acceptance testing
 - Over Life
 - Refinement of benchmarks & models (over lives of systems) improves the resiliency of our processes

Organization and management

- What is the structure of the integration and preparation teams?
 - Ongoing teams vs. integration with application support
 - Differs by lab and team
 - Frequent interactions between research, operations, and support

Organization and management, cont.

- What are the necessary skills for the activity team?
 - Computer & software architecture
 - Application structure and motifs
 - Algorithms
 - Code and performance analysis
 - Consulting skills
 - Data analysis, presentation, & communication skills
 - Knowledge
 - Engineering, computer science, domain science
 - Modeling skill

Experiences and lessons learned

- What were the good and bad experiences and lessons learned?
 - Creating representative proxy apps is hard
 - Whatever mechanism (benchmark, proxy, full app) you choose today will prove problematic in the future.
 - Scaling up benchmarks is challenging
 - It would be a lot easier to pick a machine if we could test drive the real machine first

Experiences and lessons learned

- What were the good and bad experiences and lessons learned?
 - Benchmark suites contribute greatly to bringing up and wringing out the platform.
 - Would be nice to have models that confer similar confidence
 - Need more useful I/O models & benchmarks
 - Need improved access to hardware characteristics for modeling
 - Software layers introduce discontinuities into the measurement process

Experiences and lessons learned, cont.

- What were the most productive activities?
 - Strong synergy between the centers, the users, and the vendors in choosing, developing, and understanding the benchmarks and models
 - Deep knowledge of the benchmarks arises from the teams working on a wide variety of platforms
 - The structure of benchmarking can be productively automated --- (room for possible collaboration)
 - Common software harness for runs
 - Common API's for reporting results

Experiences and lessons learned, cont.

- What were the resiliency experiences?
 - Sharing common failure data
 - Cray (LANL, LBL, ORNL, SNL)
 - IBM (ANL, LLNL)
 - Using the benchmark and modeling tools to detect issues on operating platforms
 - Regression testing
 - Debugging

Experiences and lessons learned, cont.

- What was the highest risk? Was it a surprise or expected?
 - Selecting tomorrow's systems using today's codes
 - Selection bias
 - Scale issues
 - Model inadequacies
 - Getting the market basket wrong...

Most significant observation

- *Performance appraisal should be a continuous, on-going activity*
 - Not a limited, ephemeral activity, during procurement
 - Teams, code, and knowledge need to be developed over time
 - Systems benefit from ongoing evaluation

Novel systems bring great uncertainty

Effort estimate

- How big of an effort was this?
 - Currently hard to measure due to ramp up/ramp down
 - Probably $O(1-2 \text{ FTE})$
 - Benchmarking & Benchmark Development
 - Workload Characterization
- Collaborations can help
 - Aspen, SKOPE, ExaSAT
 - Models: SST
 - Proxy apps
 - Software
 - Vendors

Discussion