CCMT

Center Extreme Scale CS Research

Center for Compressible Multiphase Turbulence

University of Florida

Sanjay Ranka

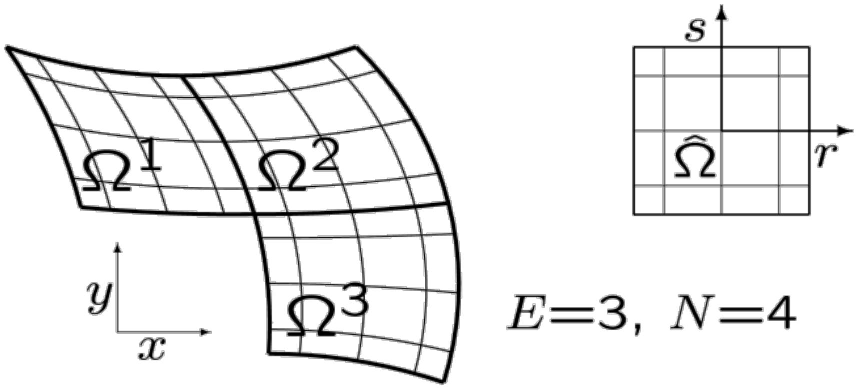
Herman Lam

Outline

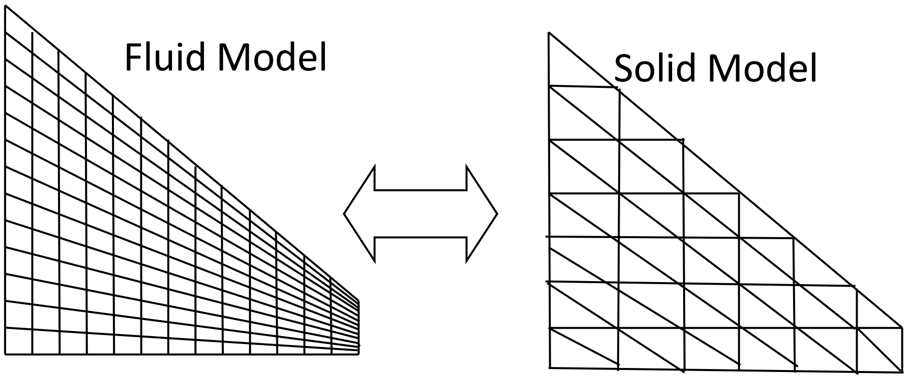
$10^6 \longrightarrow 10^7 \longrightarrow 10^8 \longrightarrow 10^9$ cores

- Parallelization and UQ of Rocfun and CMT-Nek beyond a million cores
  - Parallel performance and load balancing
  - Single processor (hybrid) performance
  - Energy management and thermal Issues
- Exascale Emulation with Novo-G
  - Exascale studies using fabrics of Behavioral Emulation Objects
  - Multiscale approach to study archs, apps, and systems
  - Multiobjective studies (performance, environment, dependability)

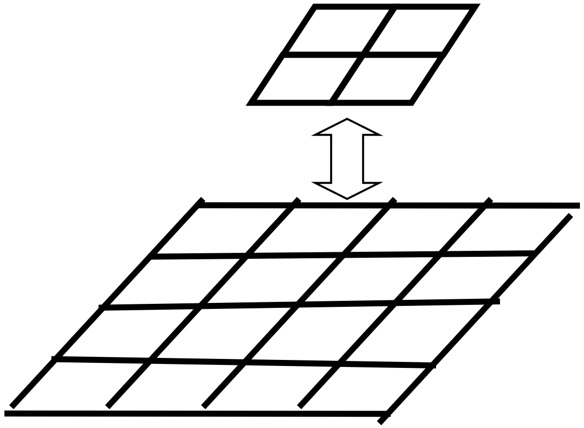
Communication Mapping: Types of Interactions



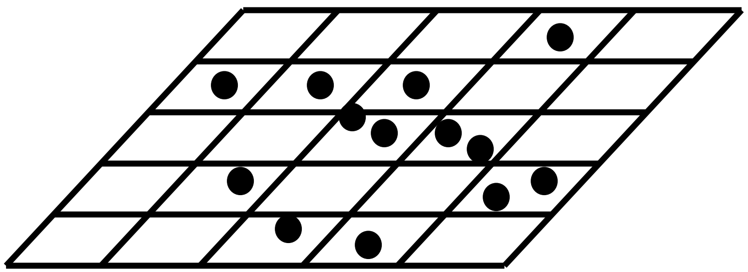
Across cells and elements



Across immersed interfaces

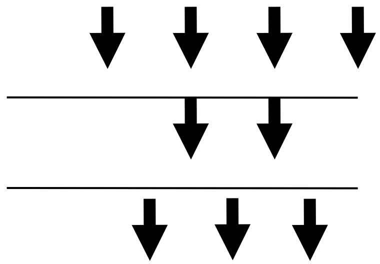


Levels of AMR

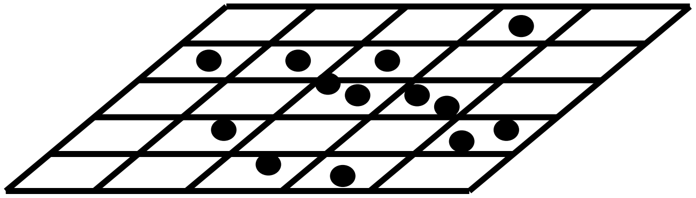


Lagrangian/Eulerian

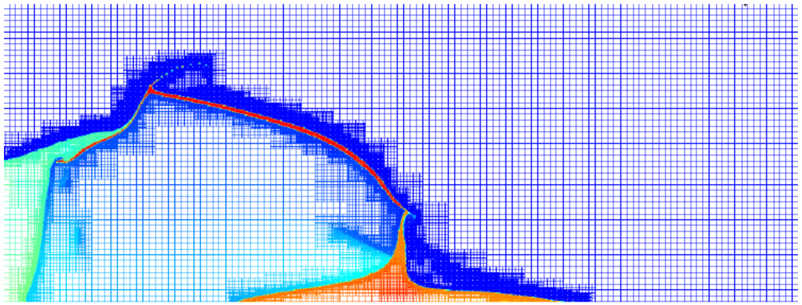
Load Balancing: Types of Adaptivity



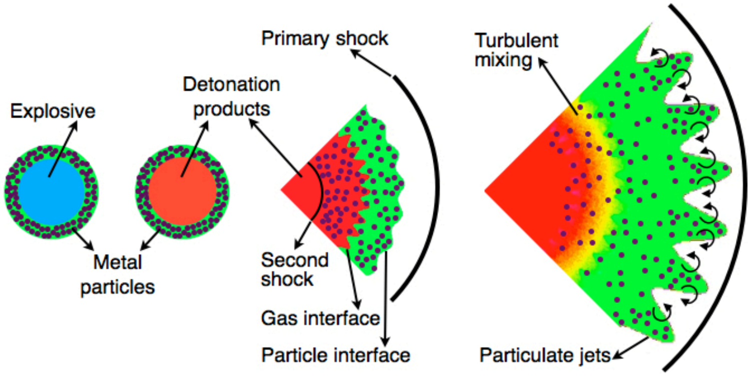
Extreme event UQ-driven
Computational steering



Preferential particle clustering
Lagrangian remap



Adaptive mesh refinement



Computational power focusing

Single Processor Performance (Hybrid Multicores)

Multiple Cores (e.g. Xeon Phi)

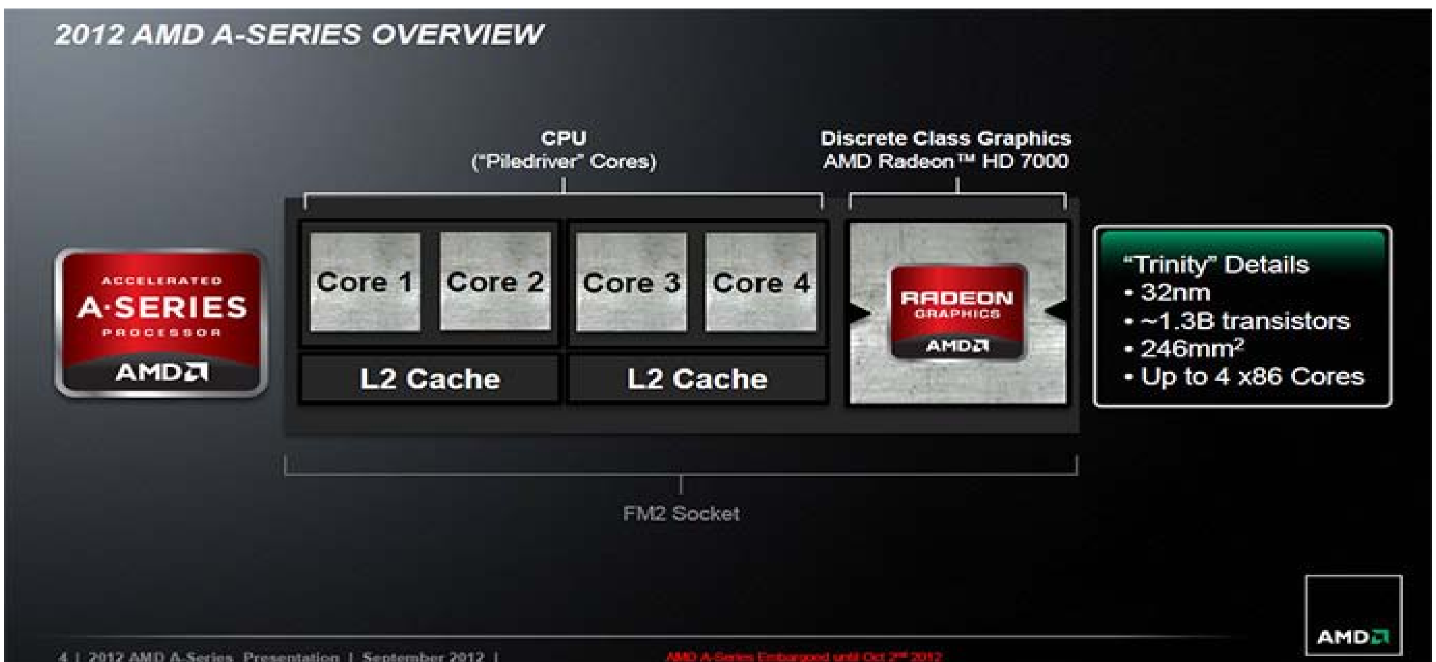
Common
- Multiple flows of Control
- Multiple Local Memories

Differences
- Synchronization
- Communication

GPU Cores

Common
- Single/Multiple flows of Control
- Multiple Local Memories

Differences
- Amount of Local memory
- Communication



Proposed Research: Hybrid Multicores

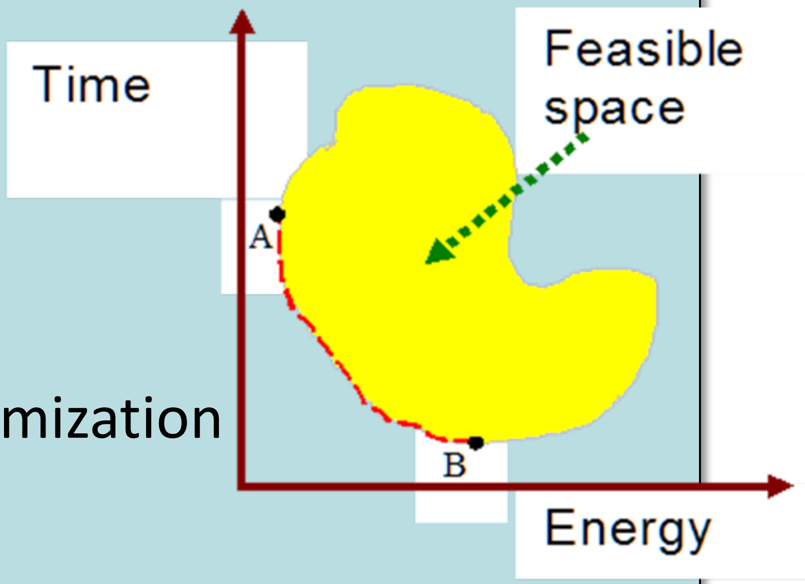
$10^1 \rightarrow 10^2 \rightarrow 10^3 \rightarrow 10^4$ cores

- Code Generation for hybrid cores
  - Support for multiple types of cores
  - Support for Vectorization
  - Auto-tuning
- Load Balancing
  - Non-uniform decomposition
- Local Data Movement
  - Movement between different cores and memories
- Multiple levels of Hierarchy
  - Small Caches, Large Caches, Level 2 Caches …

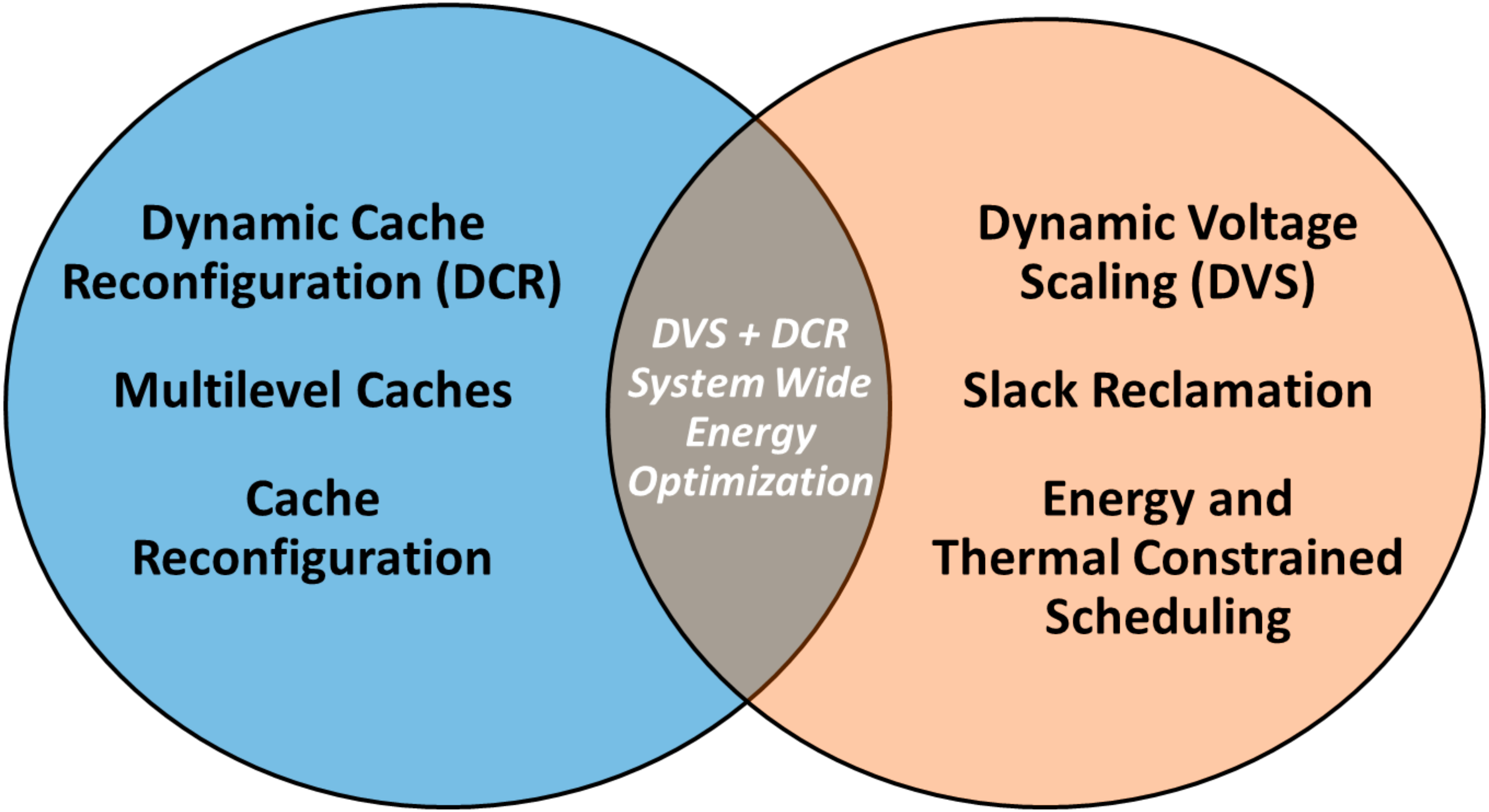
Proposed Research: Energy and Thermal Management

Develop Static and Dynamic Algorithms for Rockfun and CMT-Nek that can exploit

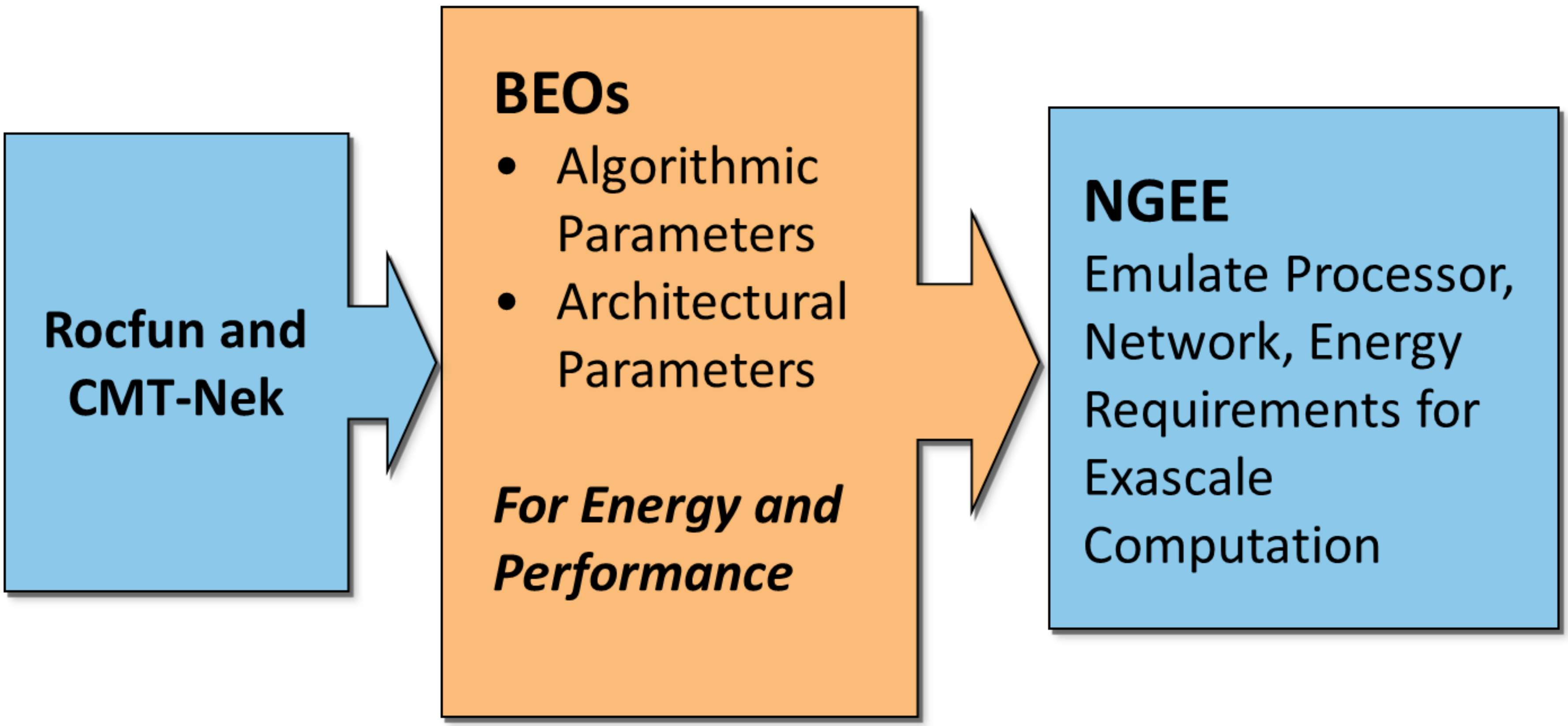
- Reconfiguration of
  - Processor (DVS)
  - Caches (DCR)
  - Buses
  - Memory
- Support Multi-objective optimization
  - Energy
  - Performance
- Support Multiple Constraints
  - Thermal Issues



Exploiting Reconfiguration for a Single Processor



Performance and Energy BEOs



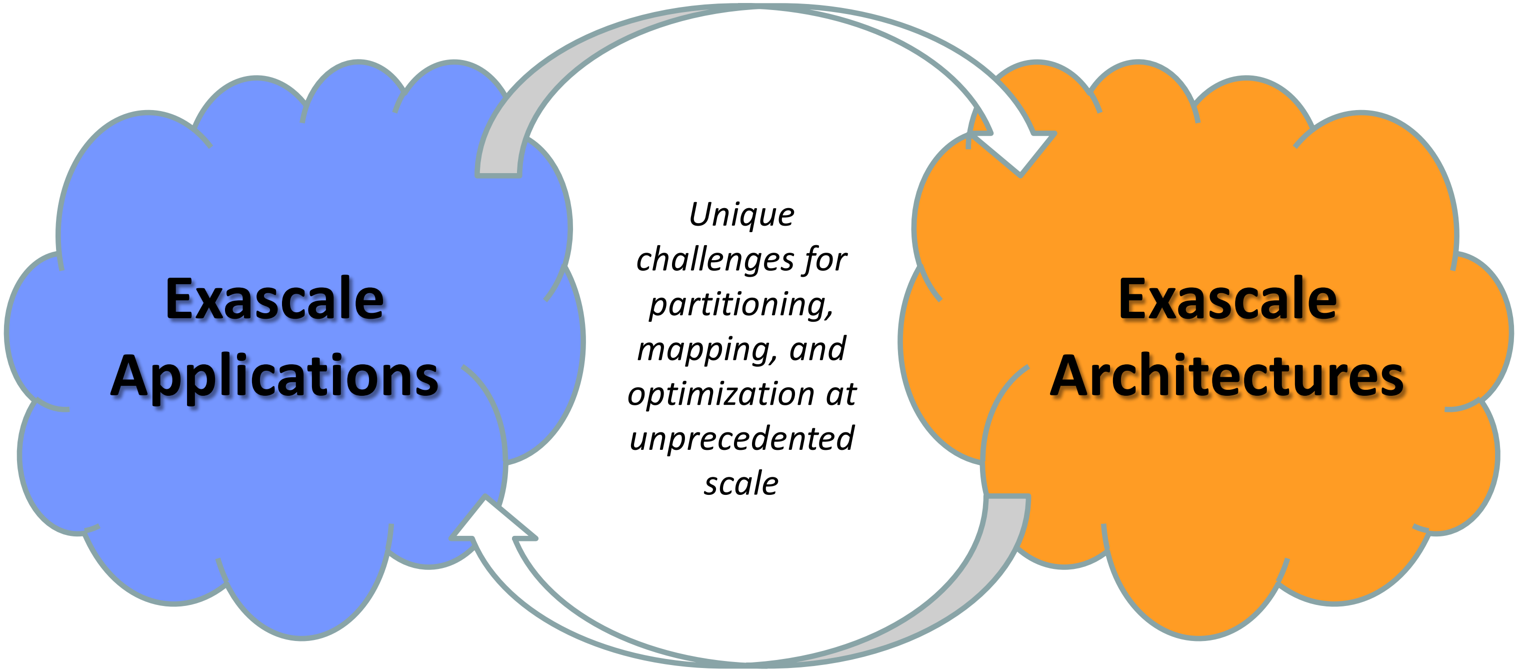
Outline

$10^6 \longrightarrow 10^7 \longrightarrow 10^8 \longrightarrow 10^9$ cores

- Parallelization and UQ of Rocfun and CMT-Nek beyond a million cores
  - Parallel performance and load balancing
  - Single processor (hybrid) performance
  - Energy management and thermal Issues
- Exascale Emulation with Novo-G
  - Exascale studies using fabrics of Behavioral Emulation Objects
  - Multiscale approach to study archs, apps, and systems
  - Multiobjective studies (performance, environment, dependability)

Why Study Exascale Systems?

- Wide variety of major research challenges
  - Design-space exploration and optimization of parallel applications and architectures at Exascale



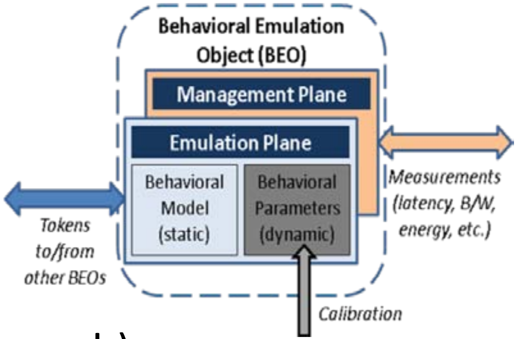
How to Study Exascale Systems?

- How may we study Exascale w/o Exascale?
  - Analytical studies – *systems too complicated*
  - Software simulation – *simulations too slow at scale*
  - Behavioral emulation – *to be defined herein*
  - Functional emulation – *systems too massive and complex*
  - Prototype device – *future technology, does not exist*
  - Prototype system – *future technology, does not exist*

- Many pros and cons with various methods
  - We believe **behavioral emulation** is most promising in terms of balance (accuracy, timeliness, scale, versatility)

Approach: Behavioral Emulation with BEOs

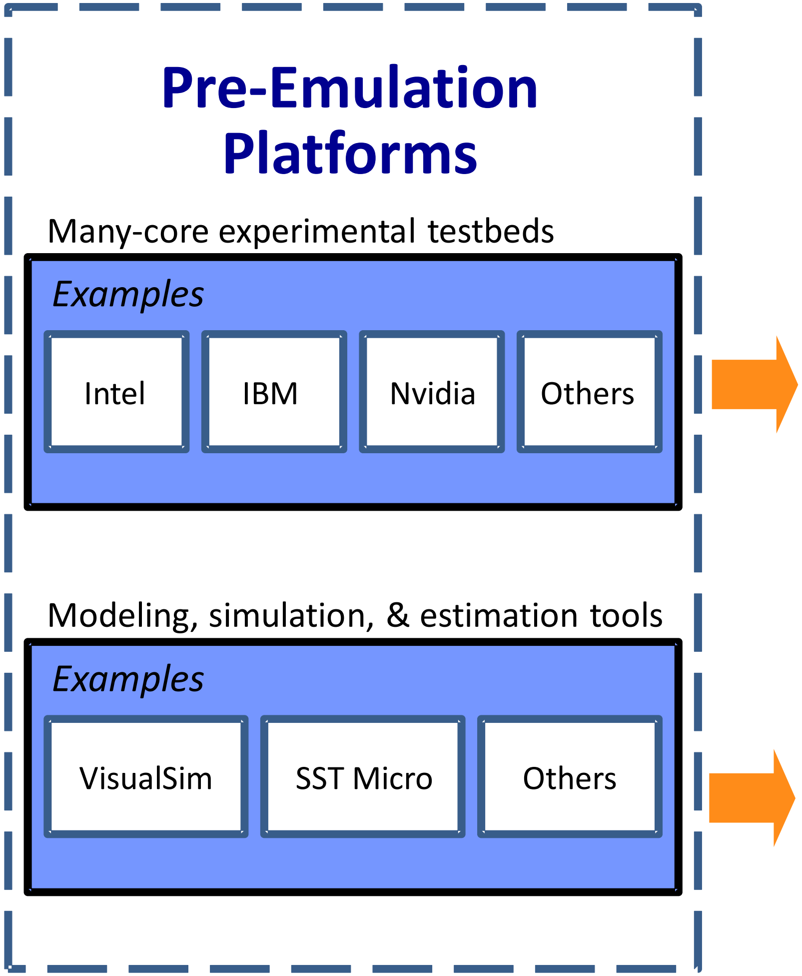
- Behavioral Emulation Objects (BEOs)
  - Characterize and explore Exascale devices, nodes, & systems, represented by fabrics of interconnected Architecture BEOs
  - Architecture BEOs stimulated by corresponding set of Application BEOs



- Multiscale, Multiobjective (different domain, but same approach)
  - Hierarchical method based upon experimentation and exploration

	CMT Apps	NGEE	NGEE BEO Models
Macro Level	CMT Skeleton-apps	***System*** BEO fabrics	▪ Models abstracted from NGEE-Meso ▪ Testbed experimentation in support ▪ Notional ***Exascale system*** exploration
Meso Level	CMT Mini-apps	***Node*** BEO fabrics	▪ Models abstracted from NGEE-Micro ▪ Testbed experimentation in support ▪ Notional ***Exascale node*** exploration
Micro Level	CMT Kernels	***Device*** BEO fabrics	▪ Architectural studies ▪ Testbed experimentation as foundation ▪ Notional ***Exascale device*** exploration

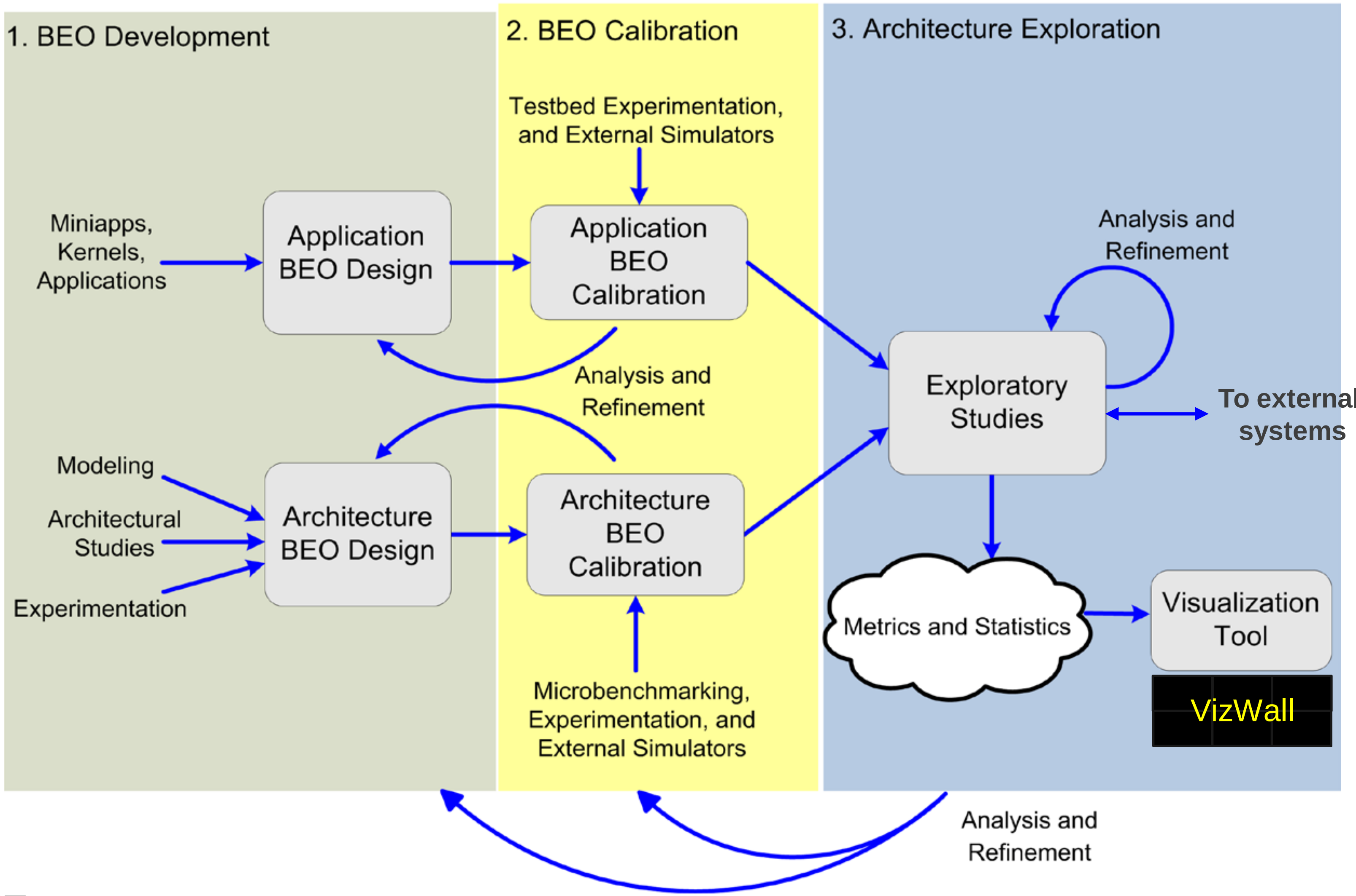
Proposed Exascale Emulator



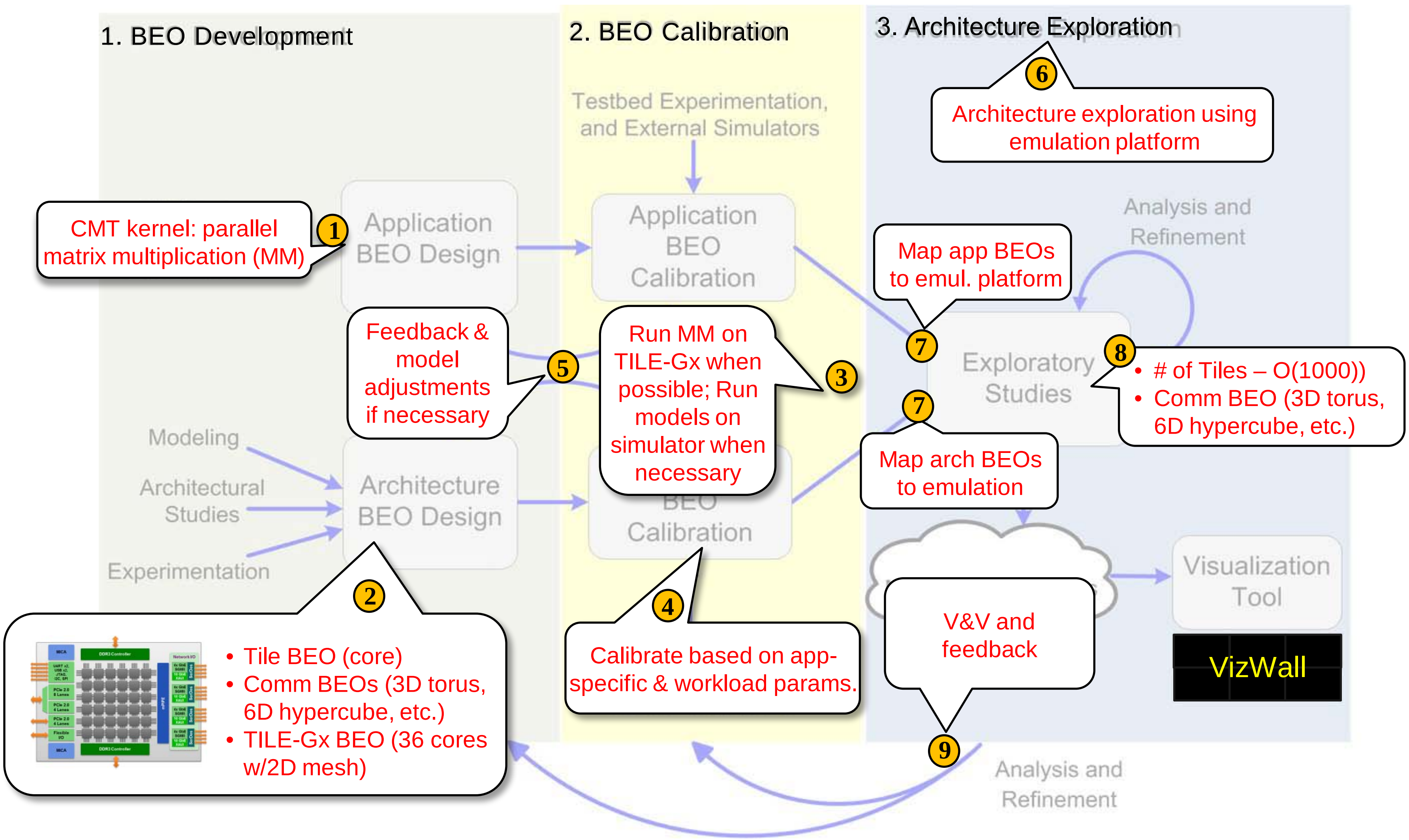
Emulation Platform

- BEOs representing Exascale devices, nodes, or systems mapped to emulation platform
- BEO method independent of emulation platform
  - Discrete-event simulation modeling tool (e.g., VisualSim)
  - Software on conventional (many-core) computer
  - Software-defined hardware on reconfigurable supercomputer (e.g., Novo-G)

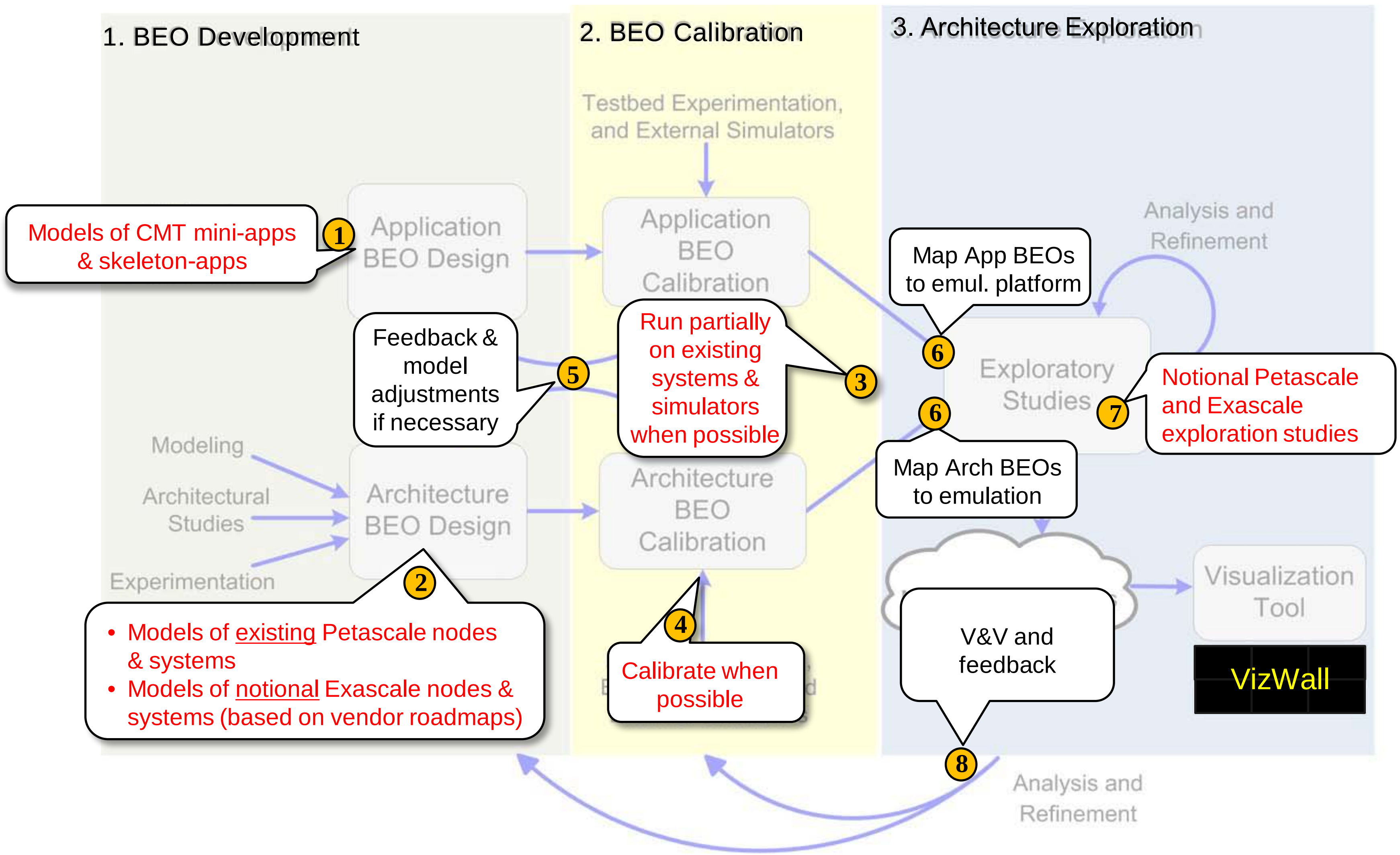
Three-Phase Workflow of NGEE



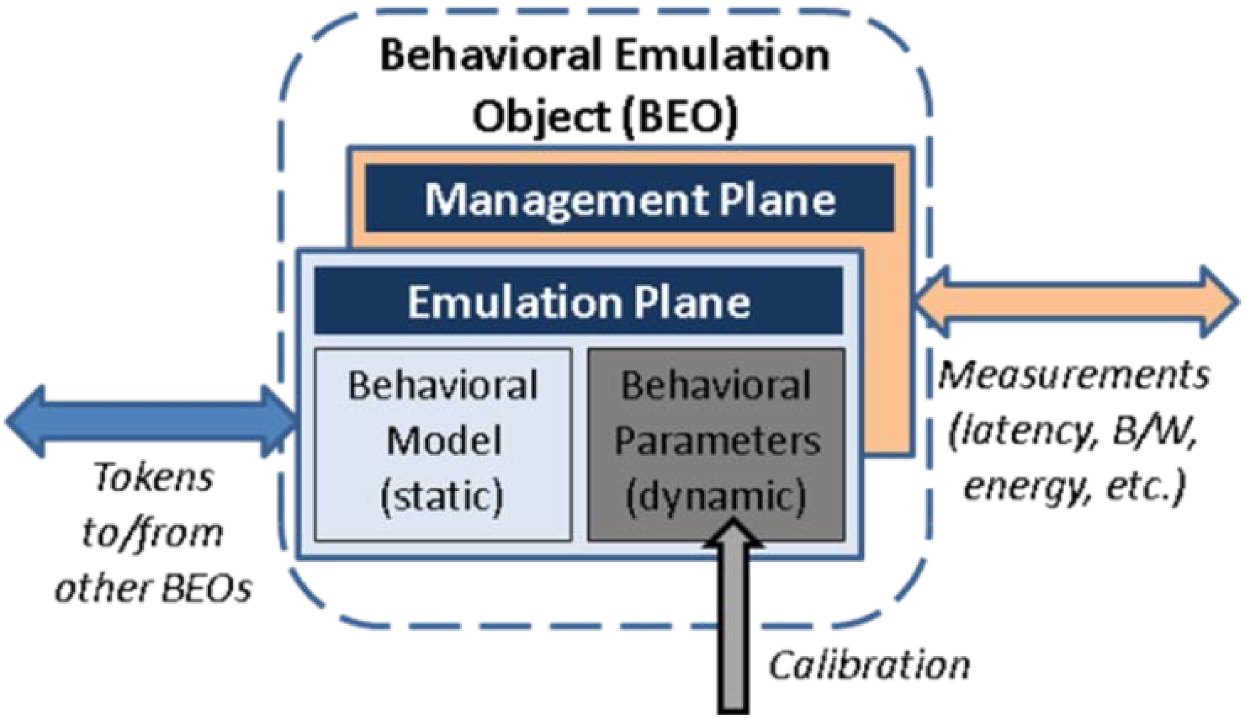
Example:
Micro Behavioral Emulation of Notional Device



Toward Emulation of Exascale Systems



Fundamental Design of a BEO

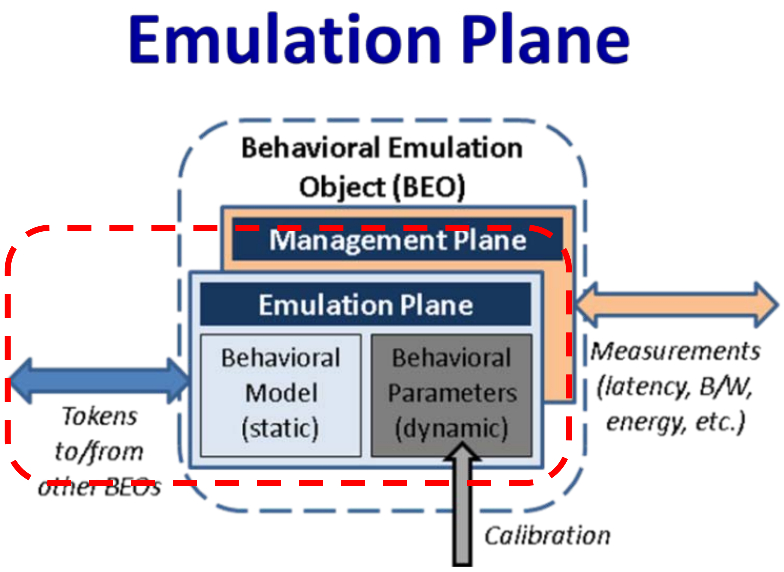


Multiobjective analysis and optimization

BEO is basic primitive in NGEE studies of Exascale systems

- **Performance factors** (execution time, speedup, latencies, throughputs, hotspots)
- **Environmental factors** (power, energy, cooling, temperature)
- **Dependability factors** (reliability, availability, redundancy, overhead)

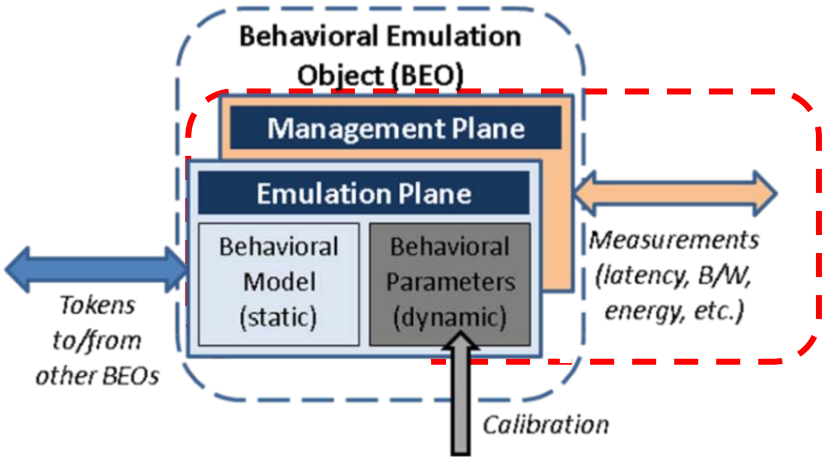
BEO Emulation Plane



- Contains
  - *Static behavioral model* (developed in development phase)
  - Set of *behavioral parameters* (specified dynamically during calibration phase)
- Performs functions to
  - *Mimic* appropriate behavior of BEO
  - *Interact* with other BEOs via tokens to support emulation studies

BEO Management Plane

Management Plane



- Responsible for
  - Measuring, collecting, and/or calculating *metrics and statistics* to support architectural exploration in terms of performance, energy, temperature, reliability, and scalability
  - Interacting with other BEOs
- Separation of management from emulation plane
  - *Minimize its interference* with emulation functions

NGEE: Novo-G Exascale Emulator

Emulation Platform

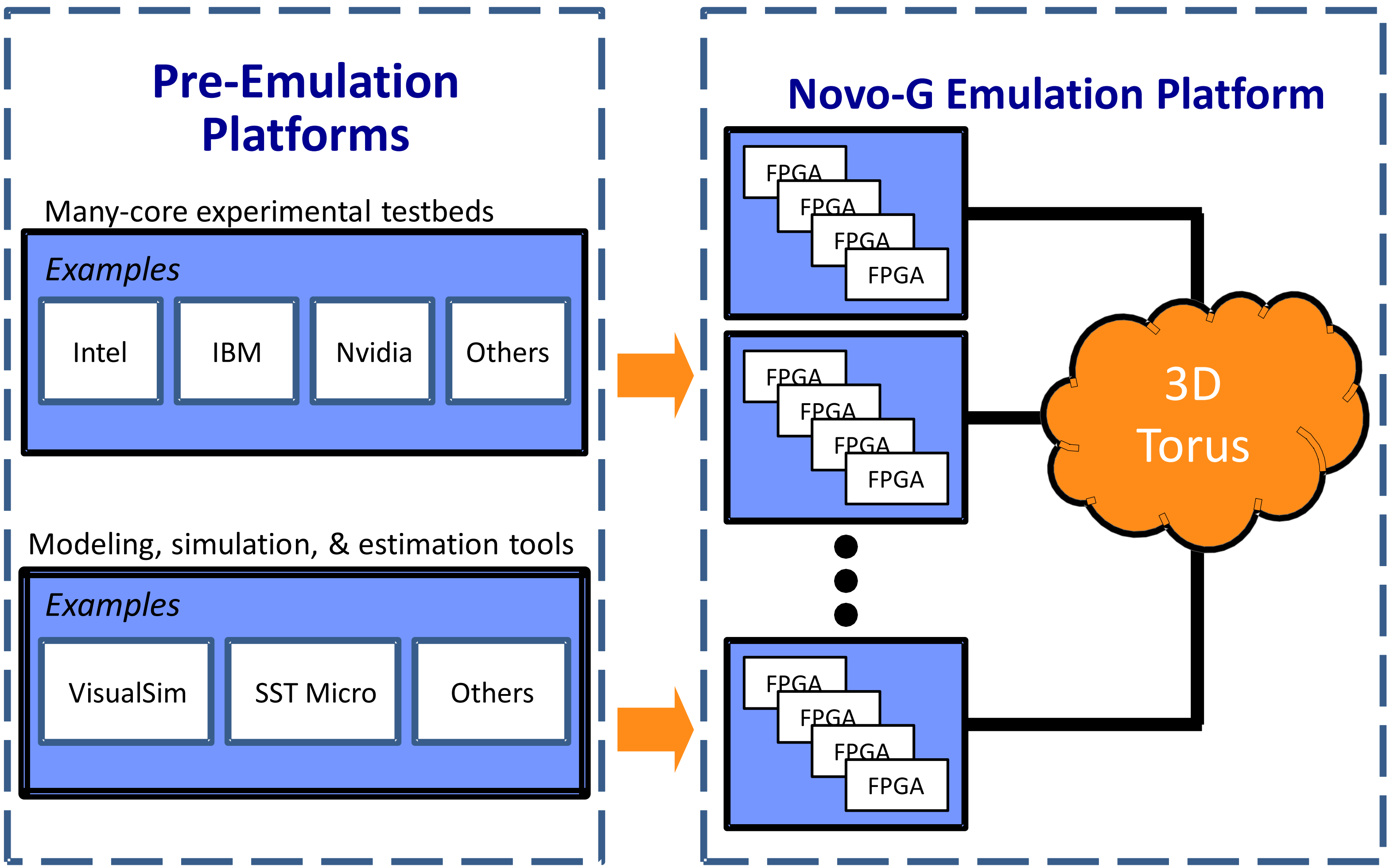
- BEOs representing Exascale devices, nodes, or systems mapped to emulation platform
- BEO method independent of emulation platform
  - Discrete-event simulation modeling tool (e.g., VisualSim)
  - Software on conventional (many-core) computer
  - Software-defined hardware on reconfigurable supercomputer (e.g., Novo-G)

- *Commercially available, flexible, ease of use*
- *For small-scale devices, nodes, and systems*

- *Emulation platform to be developed in software*
- *Higher performance than simulators, but not sufficient for Exascale*

- *Even the proposed BEO approach to emulation is challenging for studying Exascale systems*
  - ✓ Exascale, multiscale, multiobjective
- *High-risk, high-payoff, and only viable solution for full Exascale system emulation*

Novo-G Exascale Emulator (NGEE)



Novo-G Supercomputer

- Developed and deployed at CHREC (2009-present, PIs: George and Lam)
  - Supports a broad range of apps, tools, and systems research tasks in CHREC and globally
- Most powerful reconfigurable computer in known world
  - For some apps and uses, could be fastest computer of any kind in world!
  - Yet, 1000s of times less cost, size, power, cooling than high-end conventional supercomputers
- **2012 Schwarzkopf Prize**
  - CHREC and Novo-G recognized with 2012 Alexander Schwarzkopf Prize for Technology Innovation at NSF

Novo-G Annual Growth

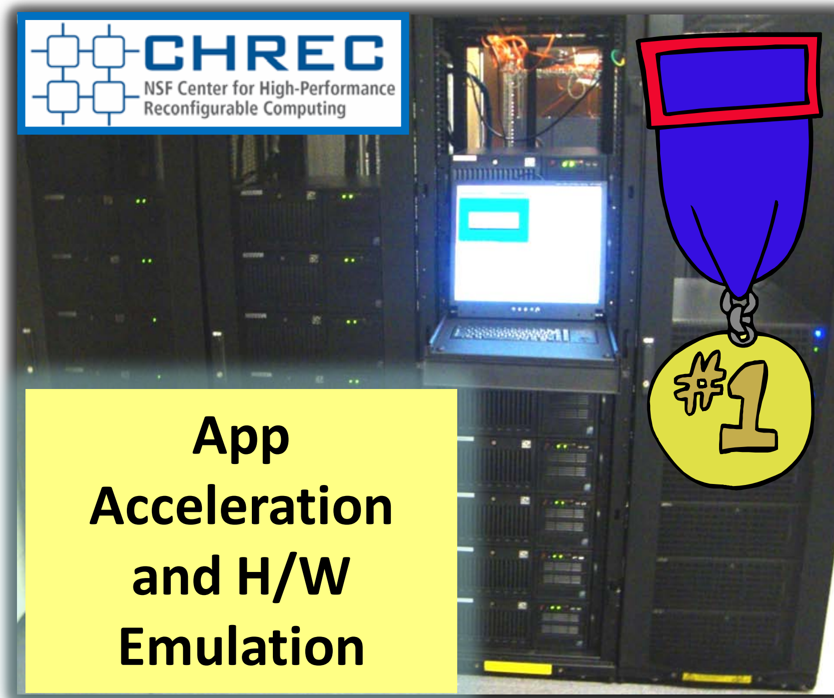
2009: 96 top-end Stratix-III FPGAs, each with 4.25GB SDRAM

2010: 96 more Stratix-III FPGAs, each with 4.25GB SDRAM

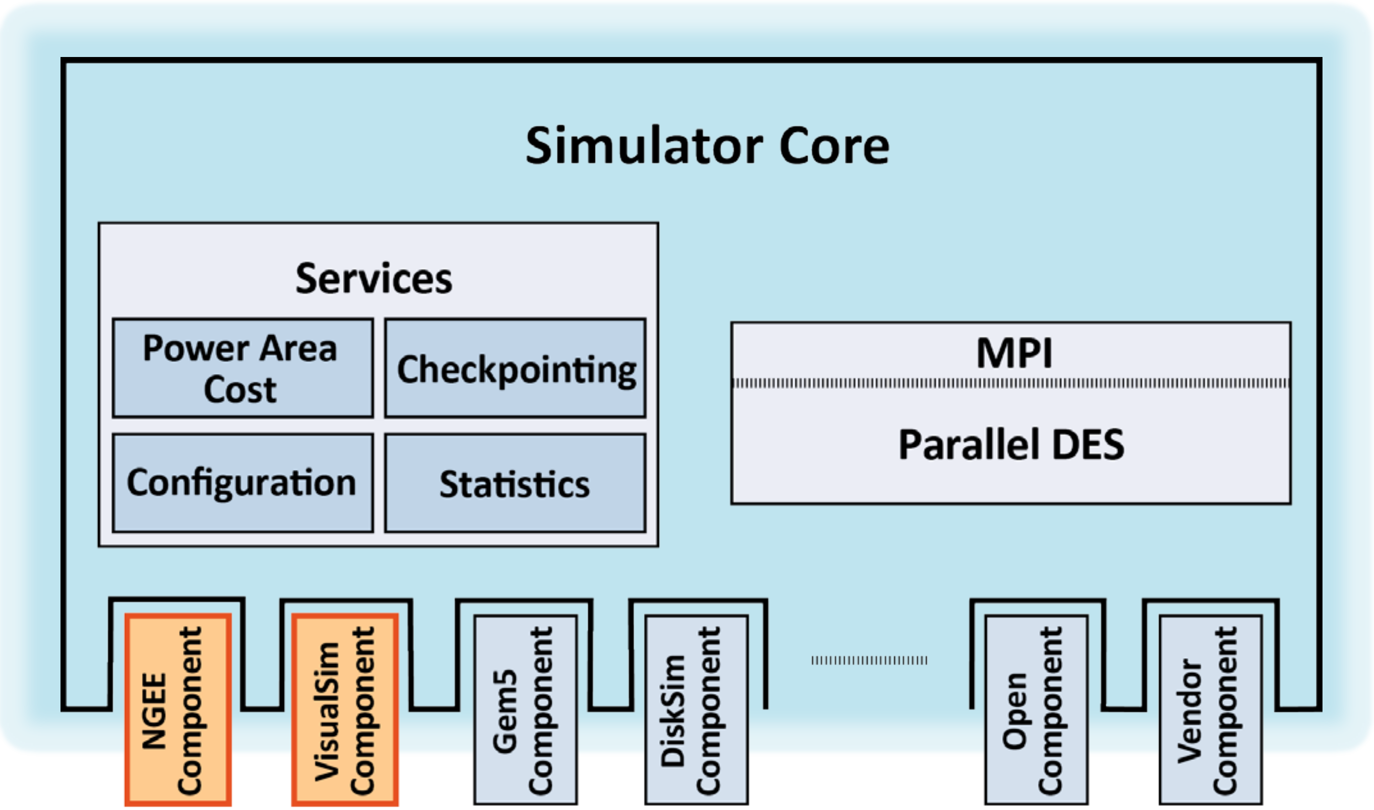
2011: 96 top-end Stratix-IV FPGAs, each with 8.50GB SDRAM

2012: 96 more Stratix-IV FPGAs, each with 8.50GB SDRAM

2013: 32 top-end Stratix-V FPGAs (4x4x2 torus or 5D h-cube)



SST Simulation Framework (c/o Sandia)

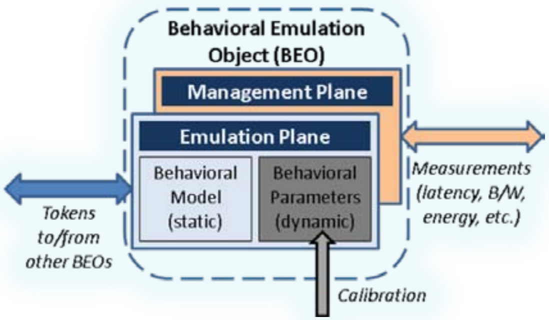


Integration with SST

- SST as pre-emulation tool
  - Support BEO development
  - BEO calibration
- SST as run-time tool
  - Dynamic support for emulation studies

Conclusions

- **Novel approach to study systems at unprecedented scale**
  - Behavioral emulation for Exascale studies with BEO fabrics
  - Ideal balance of accuracy, timeliness, scale, versatility
- **Multiscale approach to study archs, apps, and systems**
  - **Micro** = devices for Exascale
  - **Meso** = nodes for Exascale
  - **Macro** = systems for Exascale



- **Multiobjective studies**
  - **Performance** (runtime, speedup, latency, throughput, hotspots)
  - **Environment** (power, energy, cooling, temperature)
  - **Dependability** (reliability, availability, redundancy, o/h)

Do you have any questions?