www.bsc.es

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MUSA: MUltilevel Simulation Approach

DEEP-SEA

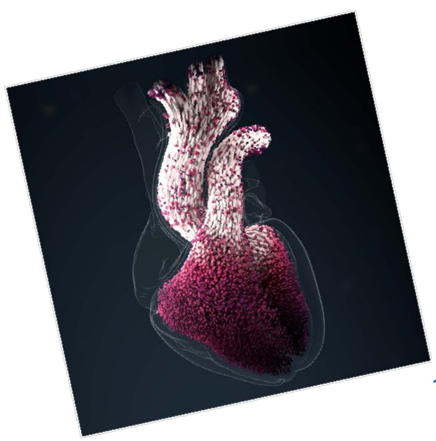
Dimitrios Chasapis (dimitrios.chasapis@bsc.es)
BSC-CNS

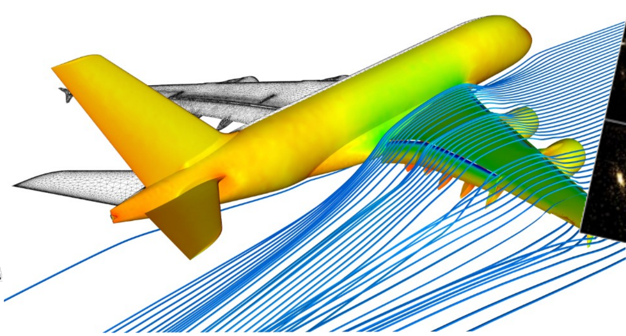
1st October 2021

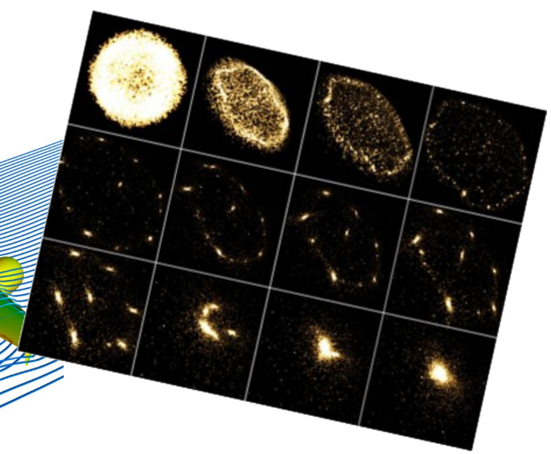
Motivation

Simulation in Science and Industry

- Abstract algorithmic representation of natural and human-made systems
- “Simulate” it with the use of computers
- Wide range of application:
  - Try industrial designs without the need to manufacture expensive prototypes (e.g. aerodynamics in aviation industry)
  - Observe and predict physical phenomenons at a fraction of both financial costs and time (e.g. nuclear fusion, astronomy)
  - Low risk (e.g. health care)

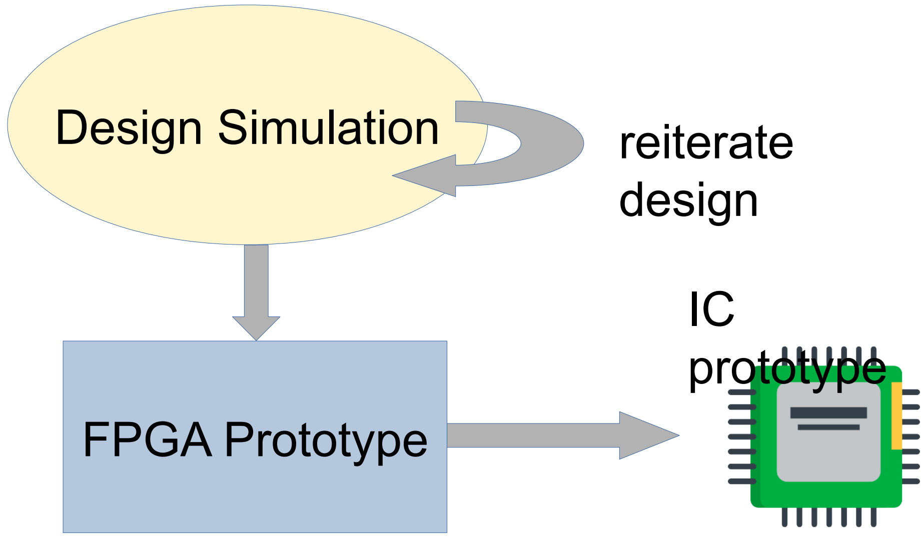






Motivation

- Guiding computer architecture design with simulations
  - Reduce development time and cost – writing software is cheaper than prototyping IC chips.
  - Intel, AMD and ARM all use in-house simulators in their R&D departments
  - Need to navigate through a constantly expanding design space
    - Try ideas fast and cheap



Motivation

- Simulating computers with… computers – what could go wrong?
  - Introduces some error
    - Abstraction may not describe a system with 100% accuracy
    - We cannot compare with native executions
  - Very slow compared to native execution
    - Every instruction introduces overhead
    - About 2 MIPS → e.g. 10 seconds of native execution can take many hours
- MUltilevel Simulation Approach – MUSA
  - Multiple levels of abstraction at varying levels of detail
  - Speedup simulation → can be faster than native execution
  - Can simulate thousands of cores

MUSA – Multilevel Simulation Approach

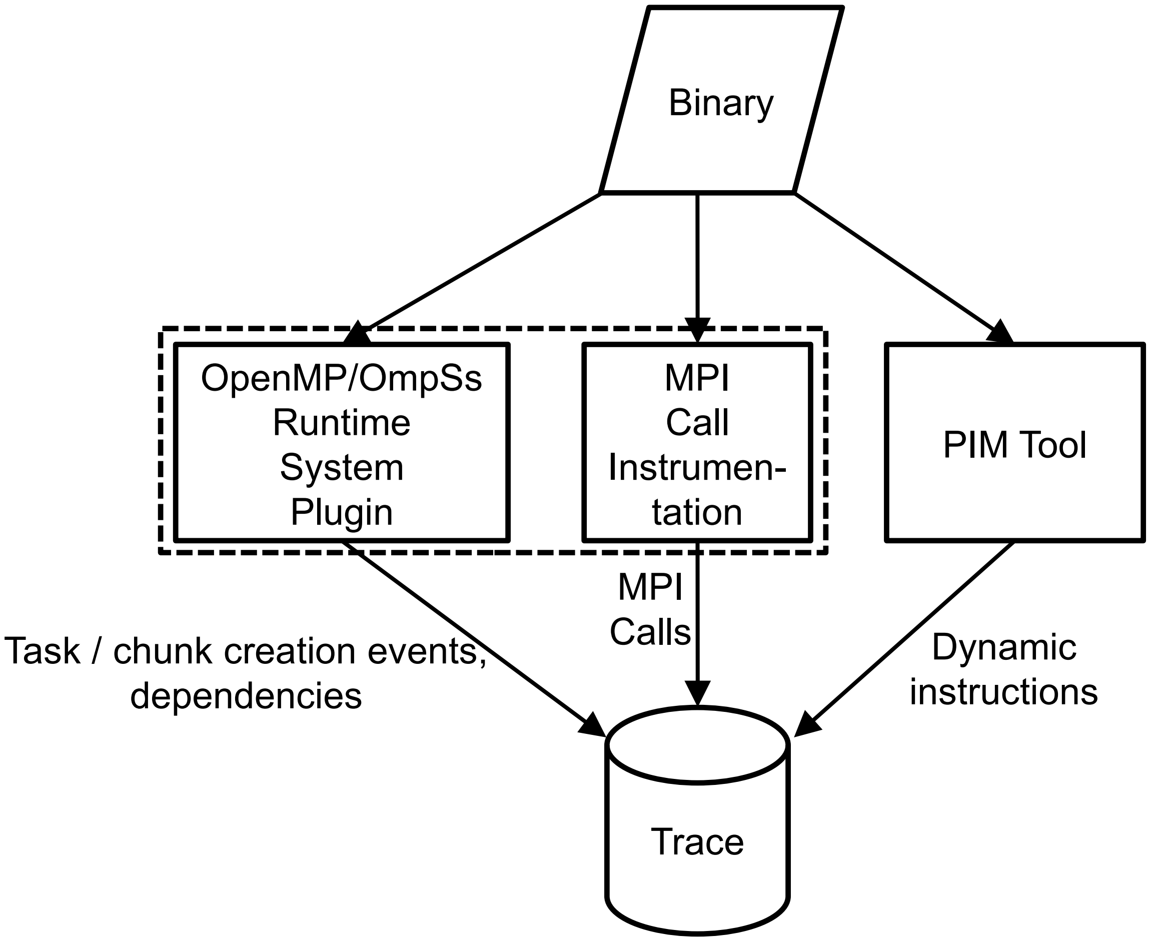
- HPC applications stress systems at multiple levels
  - Hardware: CPU, Memory, Network
  - Software: Scheduling, Synchronization, Communication
- Need a practical way to capture all the levels
- Combine multiple levels of simulation detail
  - Detailed mode: cycle accurate simulation
  - High level mode: analytical-model
- Allows us to simulate large scale machines with thousands of cores in reasonable time

MUSA – Multilevel Simulation Approach

- Targets hybrid HPC applications
  - MPI + OpenMP/OmpSs
- Dimemas analytical model used for MPI
- TaskSim simulator is used for OpenMP/OmpSs
- Trace-based simulation
  - Three levels of tracing
    - MPI level
    - OpenMP/OmpSs level
    - Instruction level

MUSA – Multilevel Simulation Approach

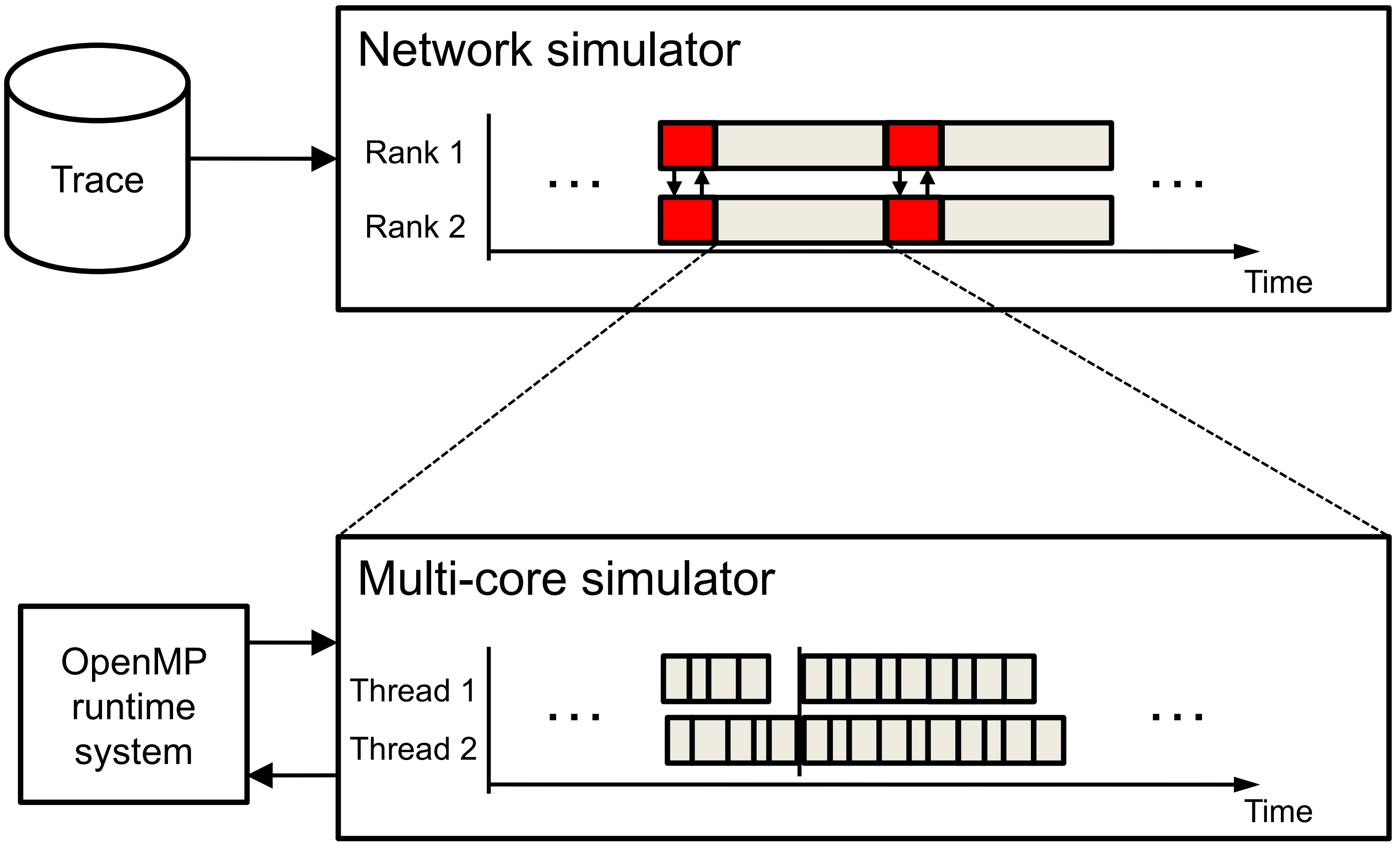
Trace-based Simulation



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MUSA – Multilevel Simulation Approach

Simulation



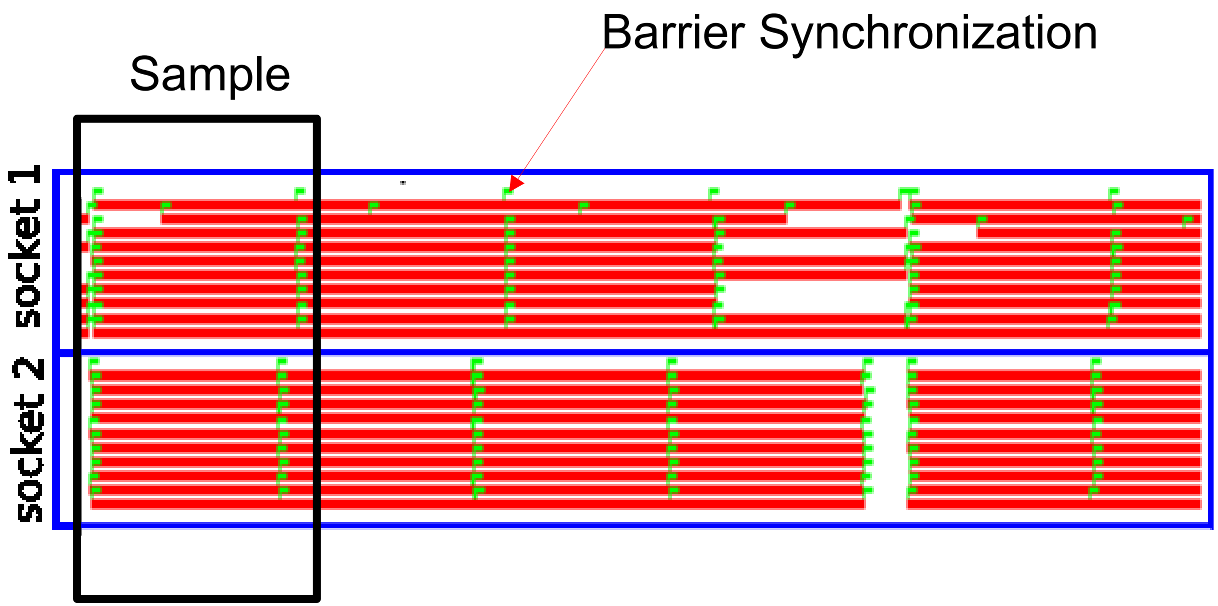
TaskSim – Architectural Simulator

- Two modes of simulation
- Burst mode
  - Execute higher level events (e.g. task creation, synchronization, scheduling)
  - Faster than native execution
  - Does not consider memory contention within node
  - Useful for quick exploration of large design spaces
- Detailed mode
  - Cycle-accurate simulation
  - Very slow
  - Higher accuracy than burst mode

TaskSim – Speeding up detailed simulations

- Use sampling to balance speed and accuracy
- Sampling can be applied orthogonally on all three levels
  - MPI:
    - Periodic Sampling of ranks (simulate 1 rank for every N ranks)
  - Burst and/or Detailed mode:
    - Manually or automatically (e.g. TaskPoint) identify and mark iterative phases
    - Tasks, loop constructs and synchronization primitives can be used by tools such as TaskPoint to identify repetitive patterns
    - User can use Paraver tool to visualize application traces and sample manually

TaskSim – Speeding up detailed simulations



- Use Paraver to visualize trace file
- Sample can be run once, and then fast-forward the result

MUSA – DEEP-SEA Extensions

- Vector engine extension
  - Important component in all HPC architectures
  - Currently supports simple vector instructions
    - 128 to 2048 vector lengths
  - Support for RISC-V, ARM, x86
- Heterogeneous Memory
  - More representative of modern systems
    - Accelerators, heterogeneous nodes
  - Important to study NUMA effects
    - Both hardware and software

www.bsc.es

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

MUSA: MUltilevel Simulation Approach

DEEP-SEA

Dimitrios Chasapis (dimitrios.chasapis@bsc.es)
BSC-CNS