



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

MUSA Tutorial

Session 3: Advanced MUSA configuration

Francesc Martínez
RoMoL at BSC-CNS

Barcelona, 11th May 2018

Outline

- « Introduction + Session objectives
- « Experiment setup
- « Modify the simulated infrastructure parameters
- « Observe how it affects simulation results

Experiment configuration:

- Benchmark: HPCCG
 - Very memory dependant
 - MPI+OpenMP
- Simulated Systems:
 - MareNostrum 4
 - High performance ARM system (ThunderX2 + HBM)
 - Low power ARM

Experiment setup

« Best to use a new folder (hpccg-sim)

« To copy the initial scripts and config files:

```
$> musa_third_session.bash
```

« The HP-CG Trace is already generated:

```
$> musa_get_third_session_trace.bash
```

« Trace generation parameters:

- Data matrix size 104x104x104
- 16 MPI Ranks
- Traced ranks 7 and 8, 6 computation phases [7-12]

Experiment setup

Simulation configurations

	MareNostrum 4	High perf. ARM	Low power ARM
Cpu freq.	2100MHz	2000MHz	1500MHz
Commit width	4	6	2
L1 Cache	32KB	32KB	32KB
L2 Cache	1MB	1MB	256KB
L3 Cache	32MB (shared)	96MB (shared)	32MB (shared)
RAM	DDR4	HBM2	DDR3
- channels	4	8	2
- speed	DDR4_2400R	HBM_1Gbps	DDR3_1600K

Modify them at your leisure

Pre-simulation environment setup

« In order to generate the pre-simulation environment:

```
$> vim generate_sim_environment.bash
```

- Has the same parameters as job_scripts did
- Change the configuration name list

« To generate the simulation environment:

```
$> ./generate_sim_environment.bash
```

« This has generated the pre-simulation scripts

- At TRACE_HPCCG_000016_MEMO/SIMULATION/A1_PRESIM
- 3 Hardware configurations
- 7 number of threads per MPI Rank
- 2 simulation modes

Execute the pre-simulations

« We run the three architecture simulations together

```
$> cd TRACE_HPCCG_000016_MEMO/SIMULATION/A1_PRESIM
```

```
$> ./generate_musa_presim.bash
```

« We modify the job time limit

```
$> vim launch_all_presims.bash
```

« We submit the pre-simulations to the queue system

```
$> sbatch launch_all_presims.bash
```


Execute the pre-simulations

⌘ This takes ~50 minutes

- Mainly due to the 64 threads per rank configs

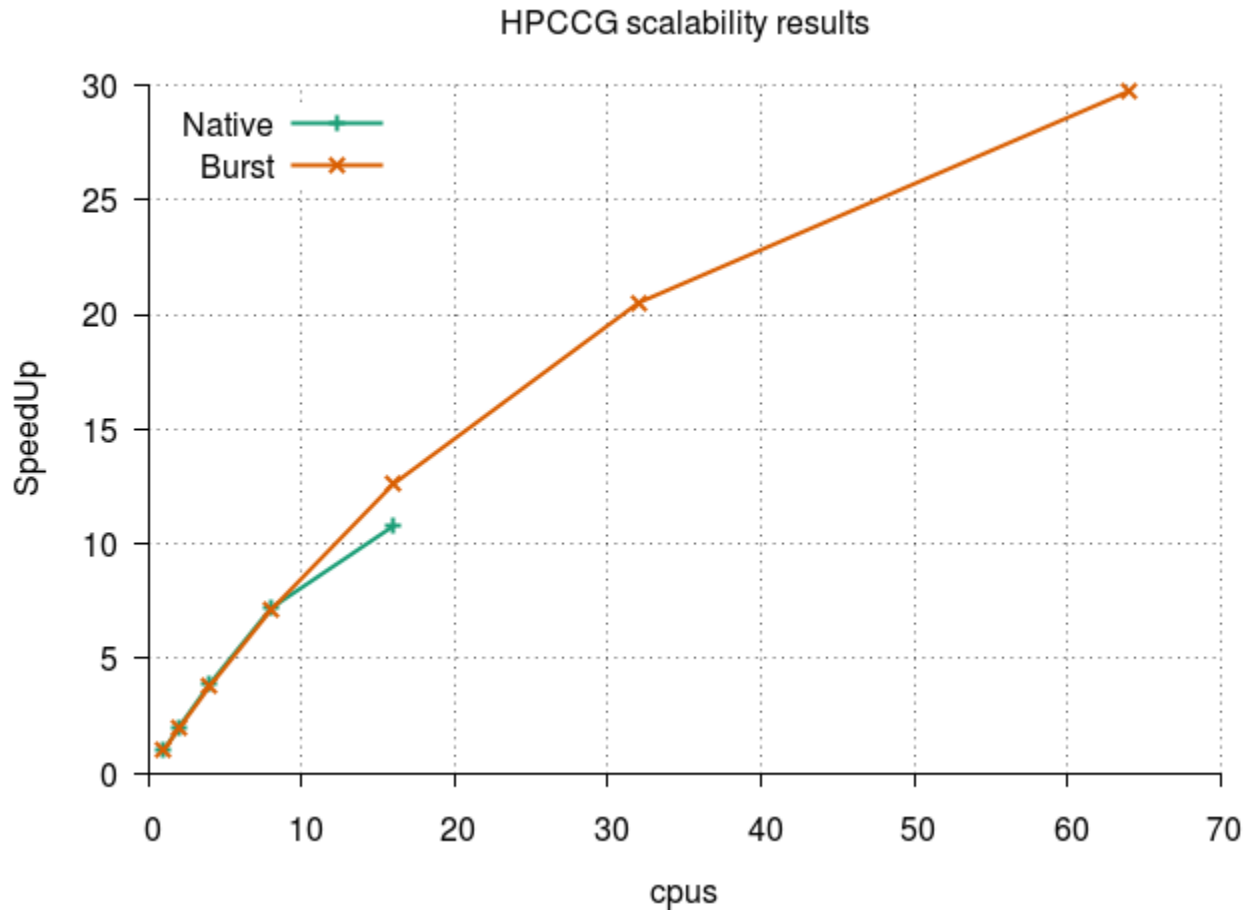
⌘ Over 10k simulations

- 6 memory phases x 2 ranks x 21 configs = 252 memory simulations

⌘ (We can extrapolate and integrate before all the results are in)

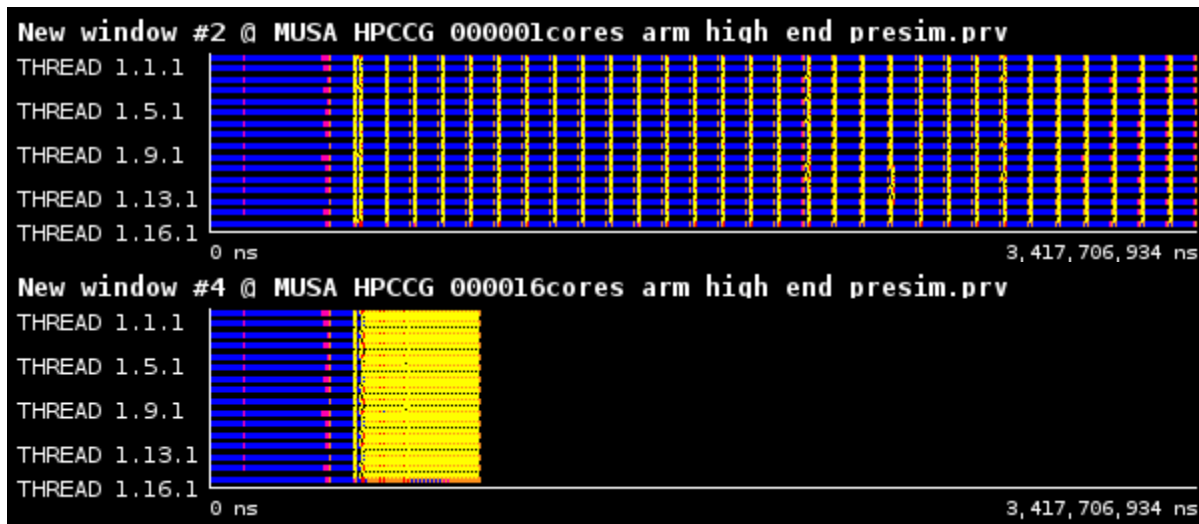
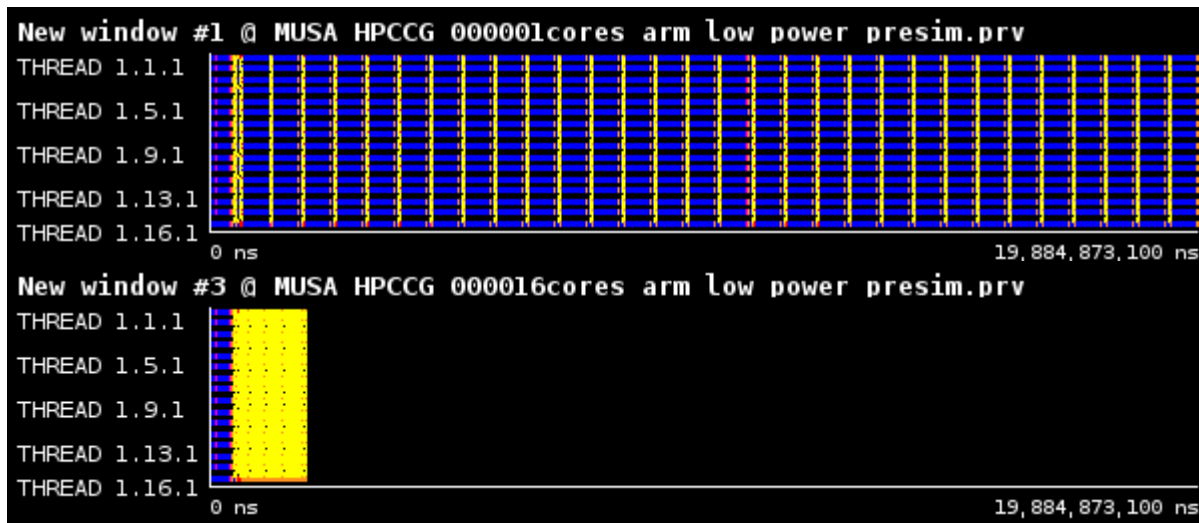
SpeedUp Native vs. Burst simulation

Native Speedup vs. Burst simulation



Potential results

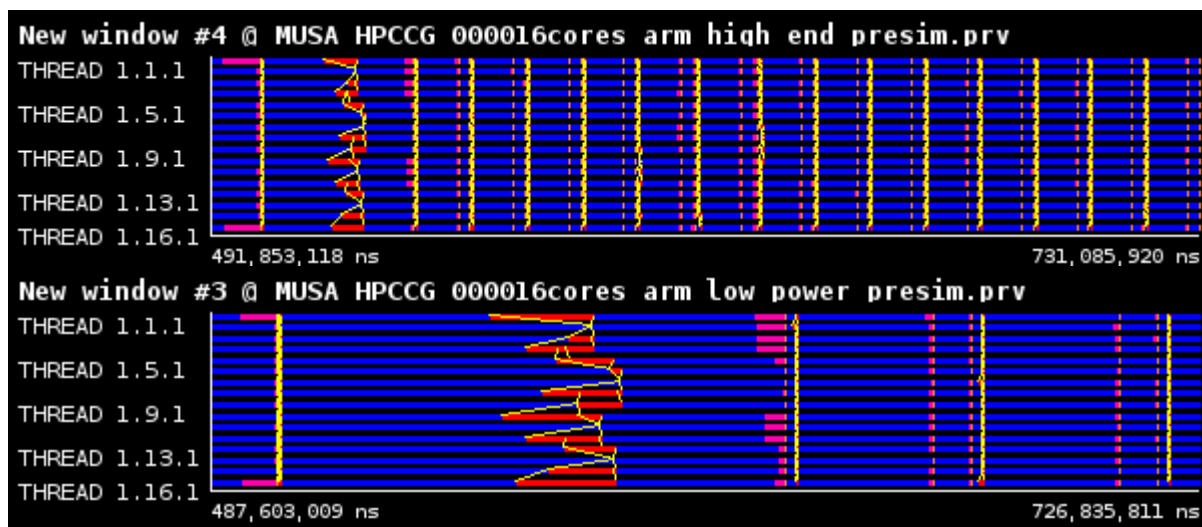
Original Paraver trace:



Potential results

« Original Paraver trace:

(Same total duration)



Integrate the pre-simulations

« To extrapolate the results:

(Still inside SIMULATION/A1_PRESIM)

```
$> ./extrapolate_burst_duration_arm_high_end.bash
```

```
$> ./extrapolate_burst_duration_arm_low_power.bash
```

```
$> ./extrapolate_burst_duration_mn4_musa.bash
```

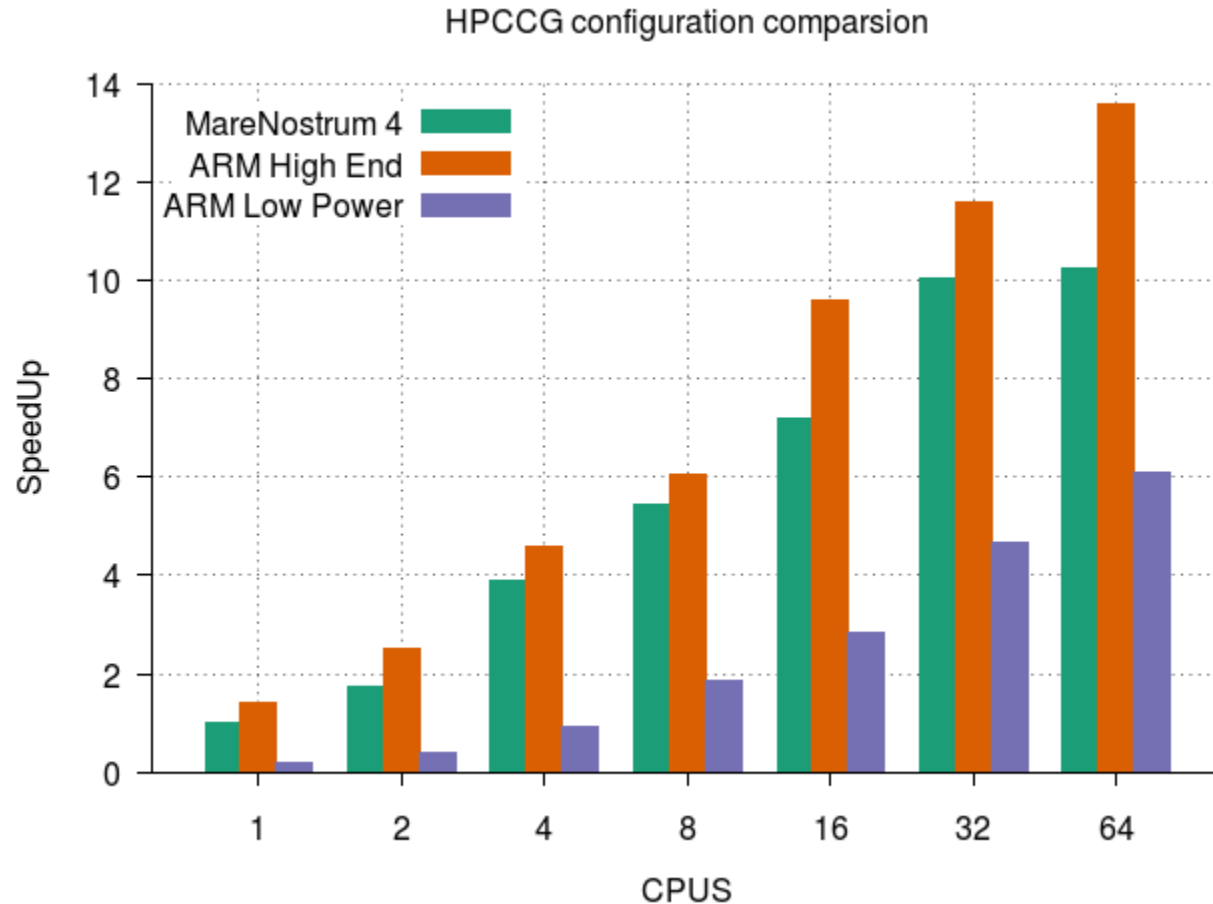
« To integrate the result traces:

```
$> cd ../A2_INTEGRATION_PRESIM
```

```
$> ./integrate_dimemas_simulations.bash
```

Simulation results (only Region of Interest)

⌘ MN4 vs ARM High End vs ARM low power



Questions?