Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

RISC-V accelerators? computing model

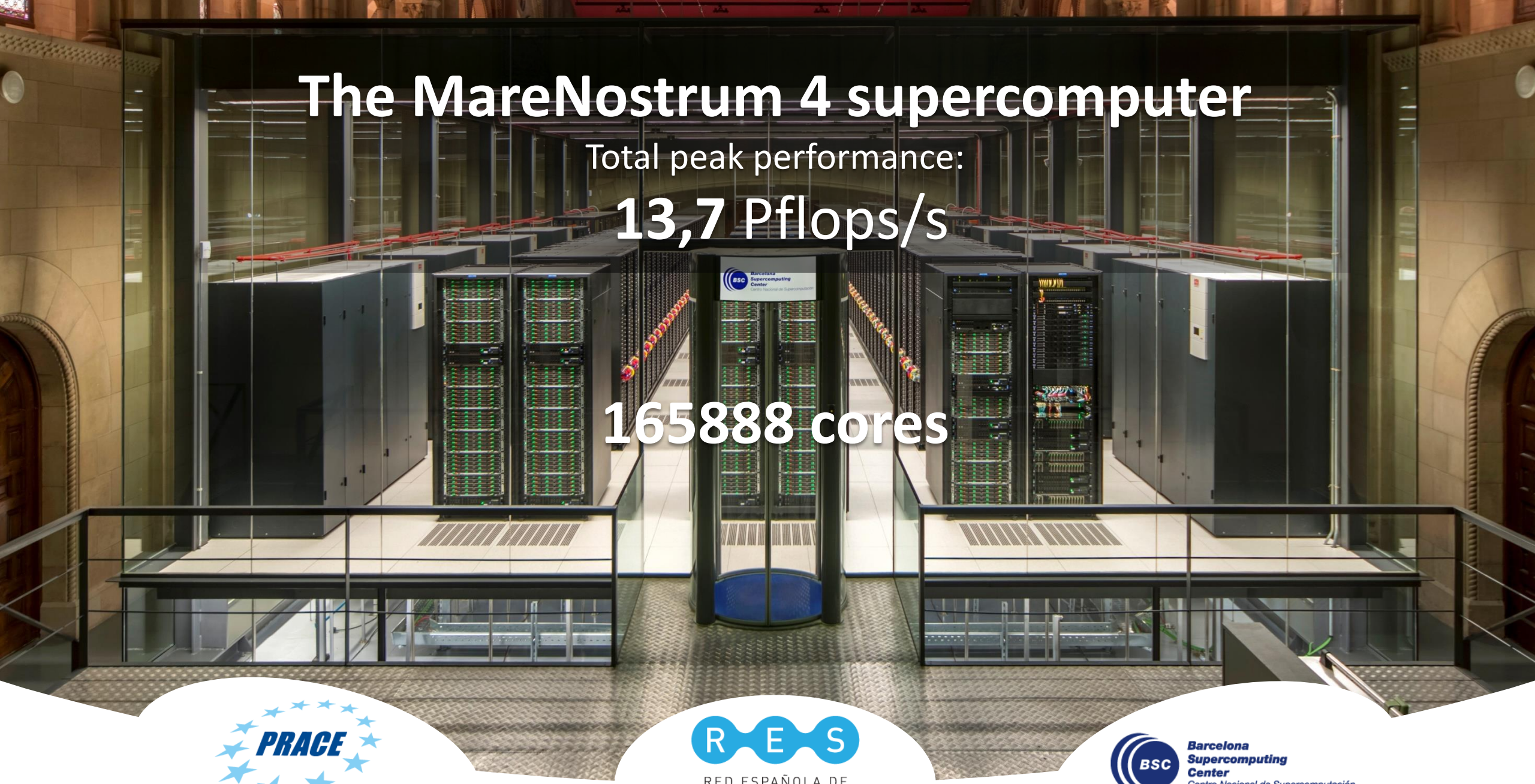
Carlos Álvarez, Jaume Bosch, Antonio Filgueras, JuanMiguel de Haro, Daniel Jiménez, Xavier Martorell, Miquel Vidal

CAPAP-H

06/02/2020

HPC Evolution

Barcelona Supercomputing Center
Centro Nacional de Supercomputación



The MareNostrum 4 supercomputer

Total peak performance:

13,7 Pflops/s

165888 cores

PRACE

Access: prace-ri.eu/hpc-access

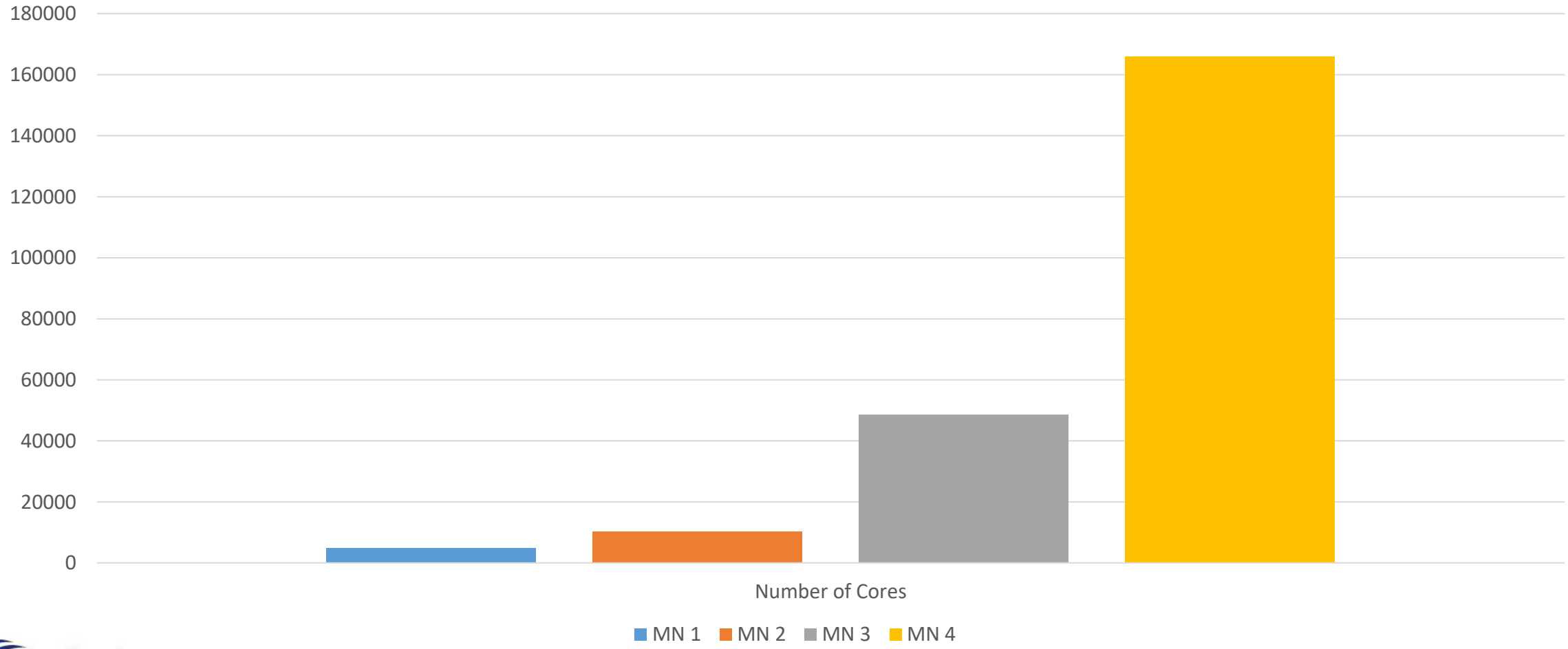
RES

RED ESPAÑOLA DE SUPERCOMPUTACIÓN

Access: bsc.es/res-intranet

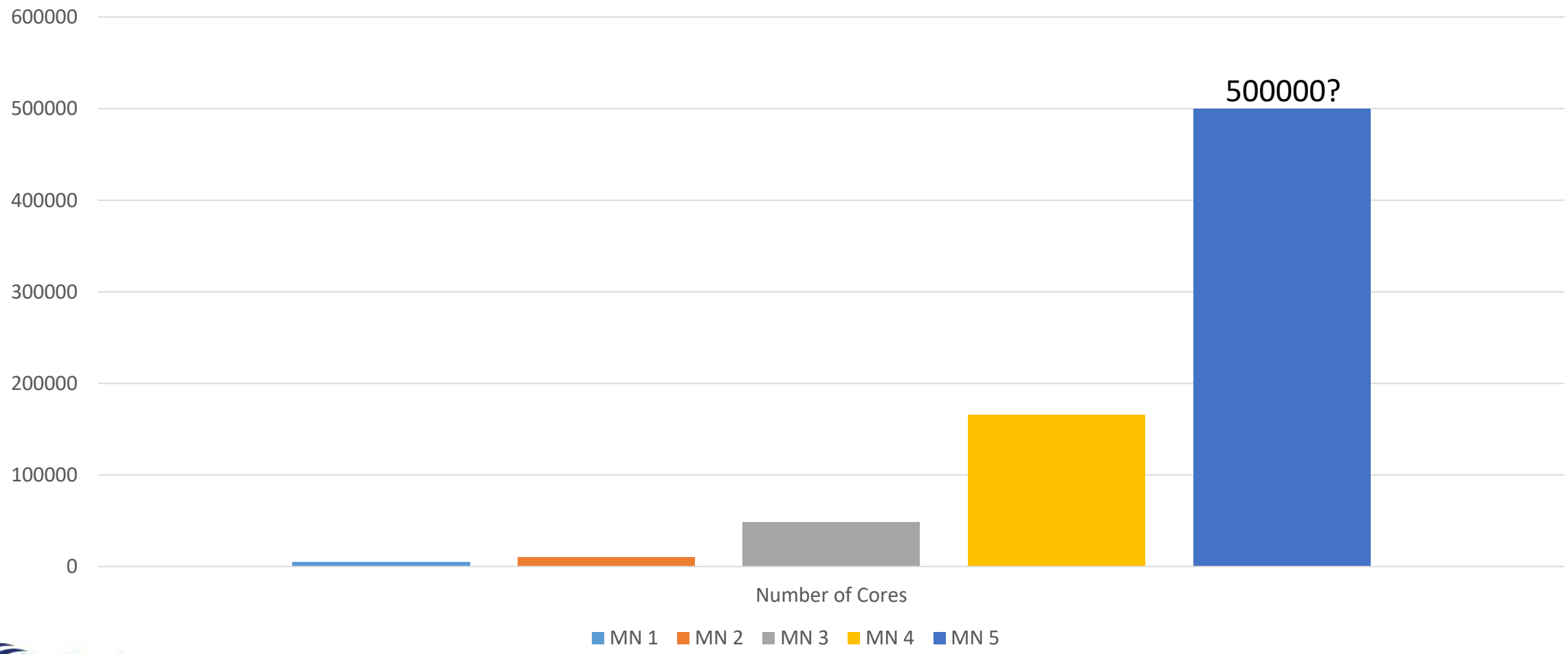
BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MareNostrum #cores evolution



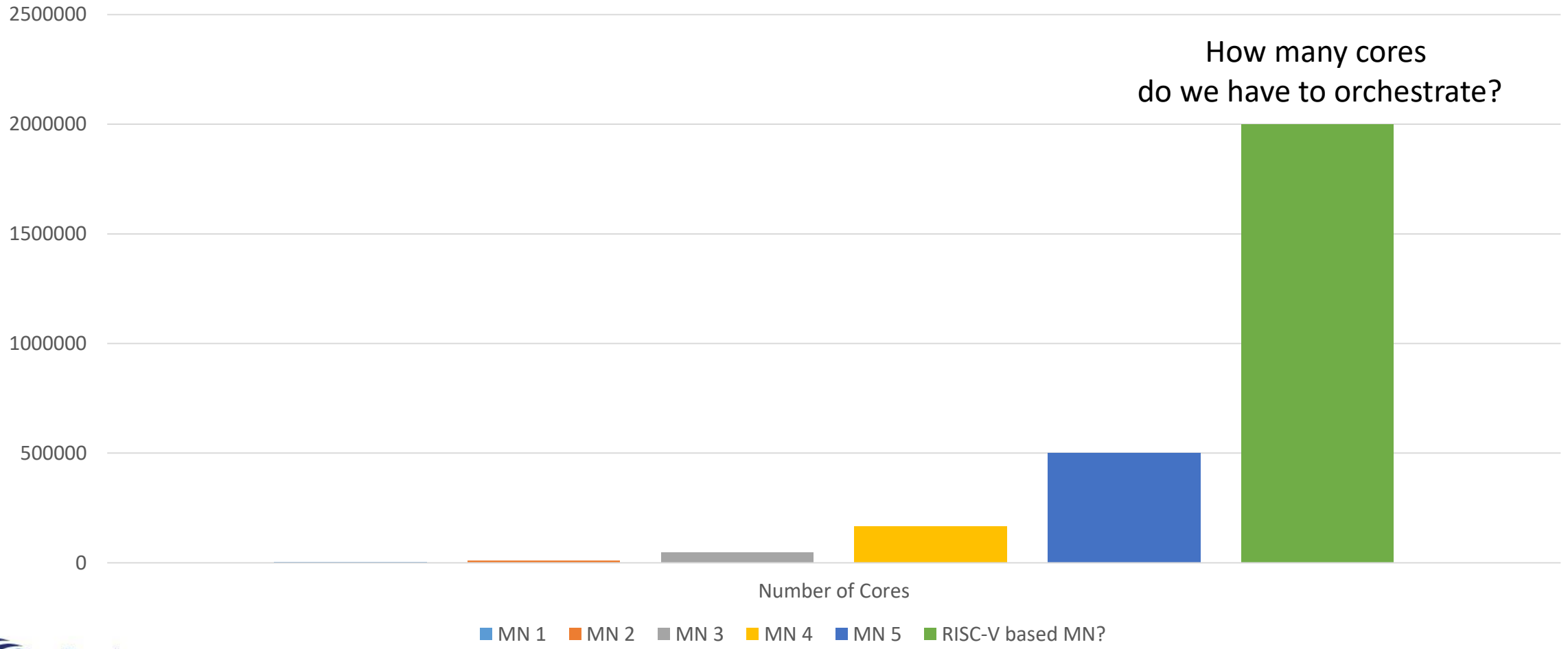
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MareNostrum #cores evolution



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

MareNostrum #cores evolution



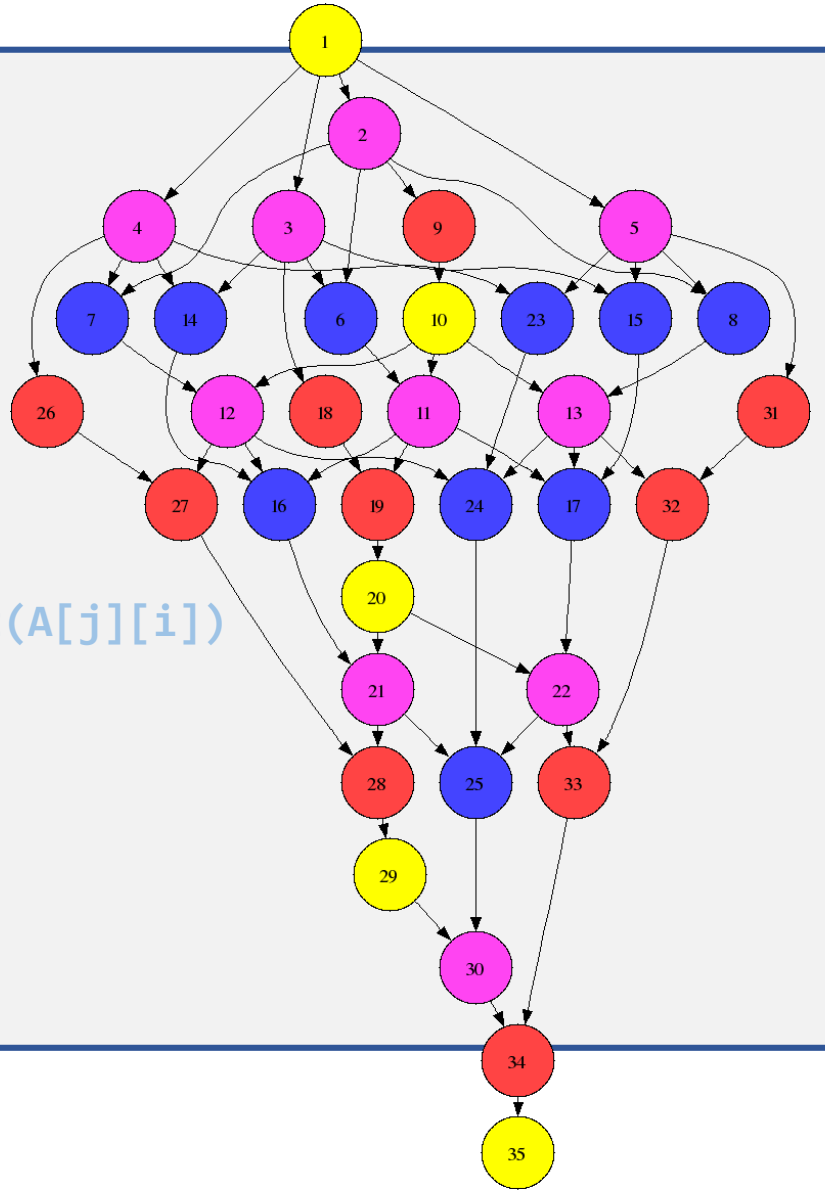
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Current Solution

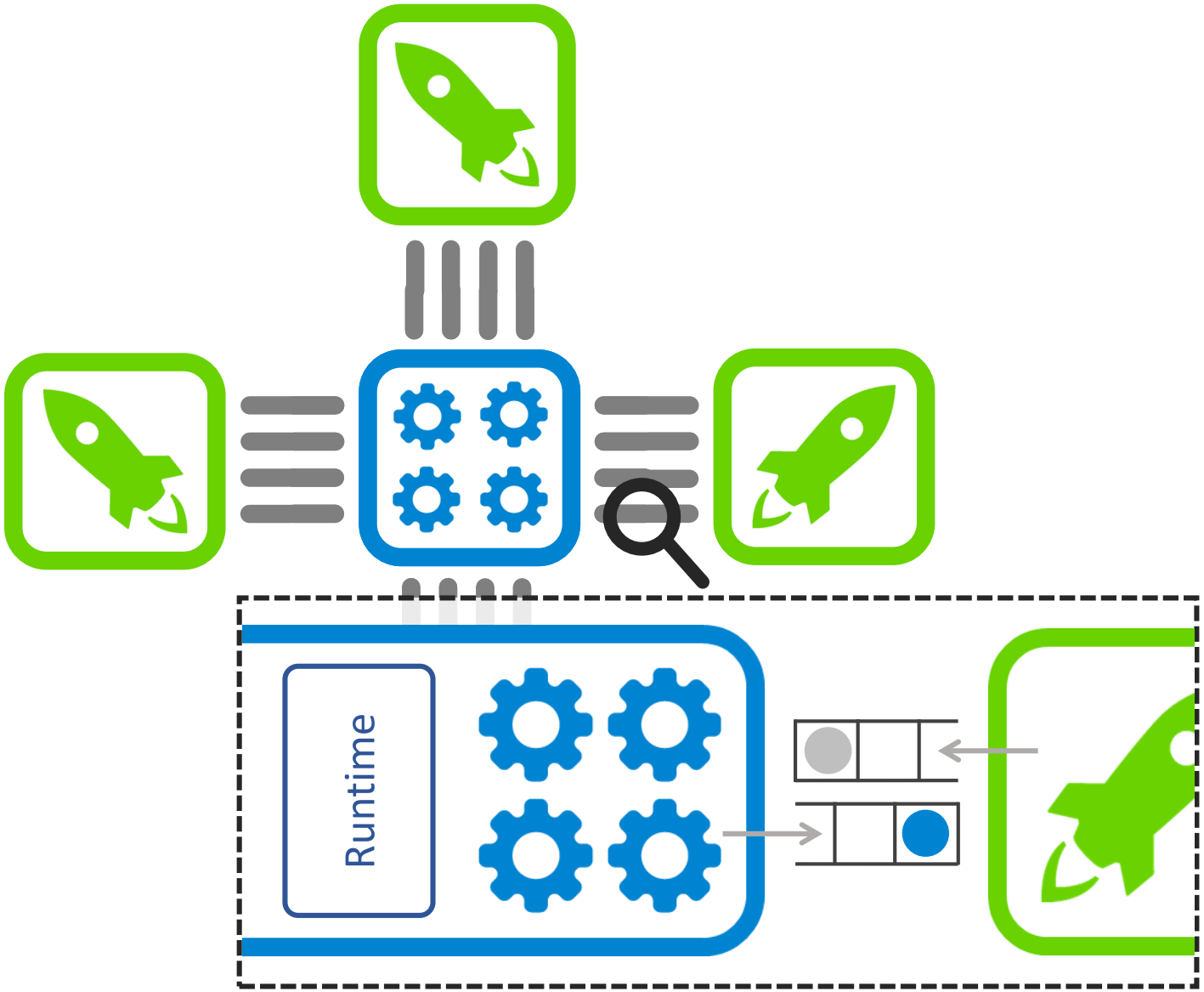
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

OmpSs programming model

```
void cholesky(float *A[NT][NT]) {
   for (int k=0; k<NT; k++) {
      #pragma omp task inout(A[k][k])
      spotrf( A[k][k] ) ;
      for (int i=k+1; i<NT; i++) {
         #pragma omp task in(A[k][k]) inout(A[k][i])
         strsm( A[k][k], A[k][i] );
      }
      for (i=k+1; i<NT; i++) {
         for (int j=k+1; j<i; j++) {
            #pragma omp task in(A[k][i], A[k][j]) inout(A[j][i])
            sgemm( A[k][i], A[k][j], A[j][i] );
         }
         #pragma omp task in(A[k][i]) inout(A[i][i])
         ssyrk( A[k][i], A[i][i] );
      }
   }
}
```

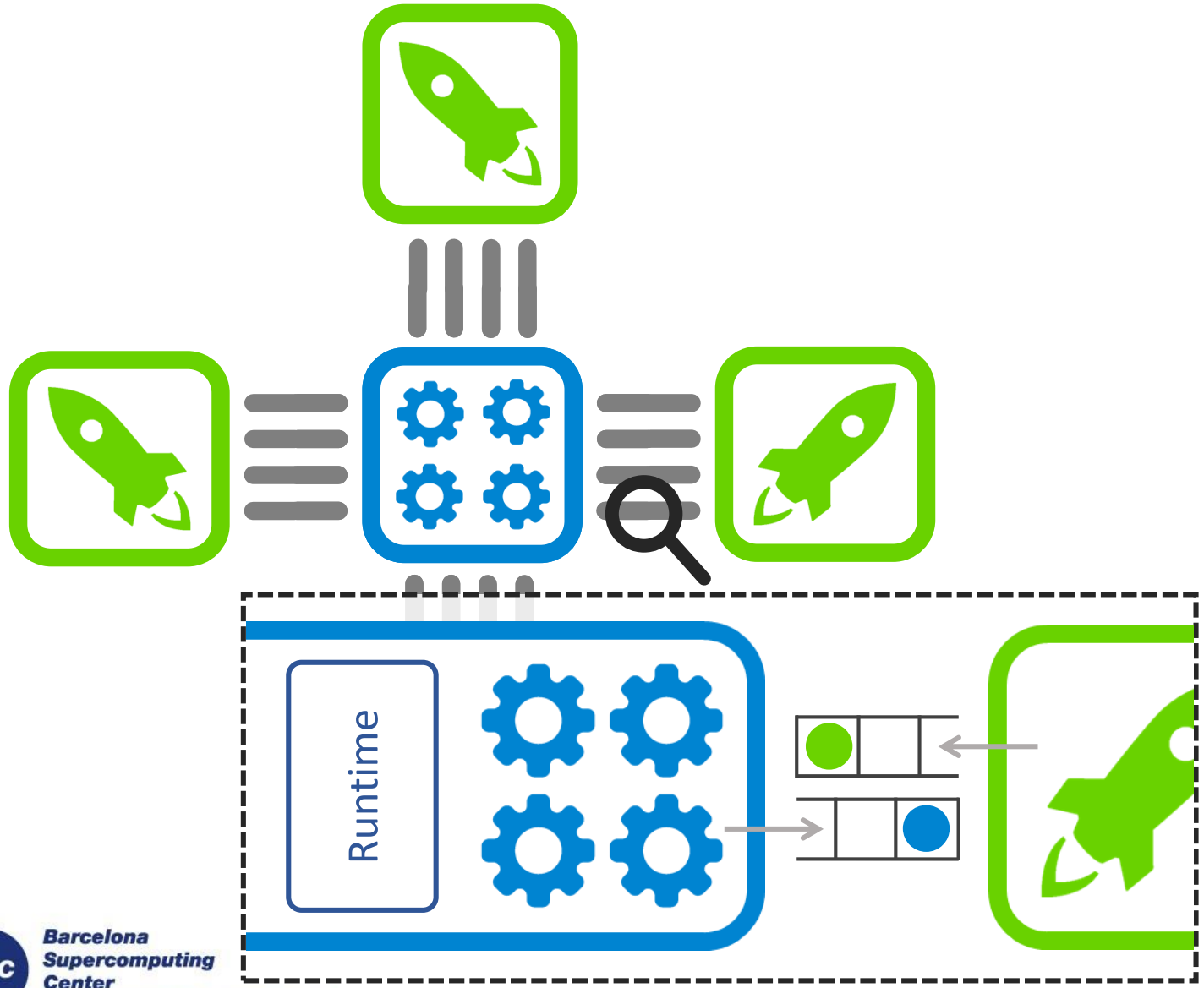


Current accelerator model



Application Acceleration on FPGAs with OmpSs@FPGA. FPT 2018: 70-77

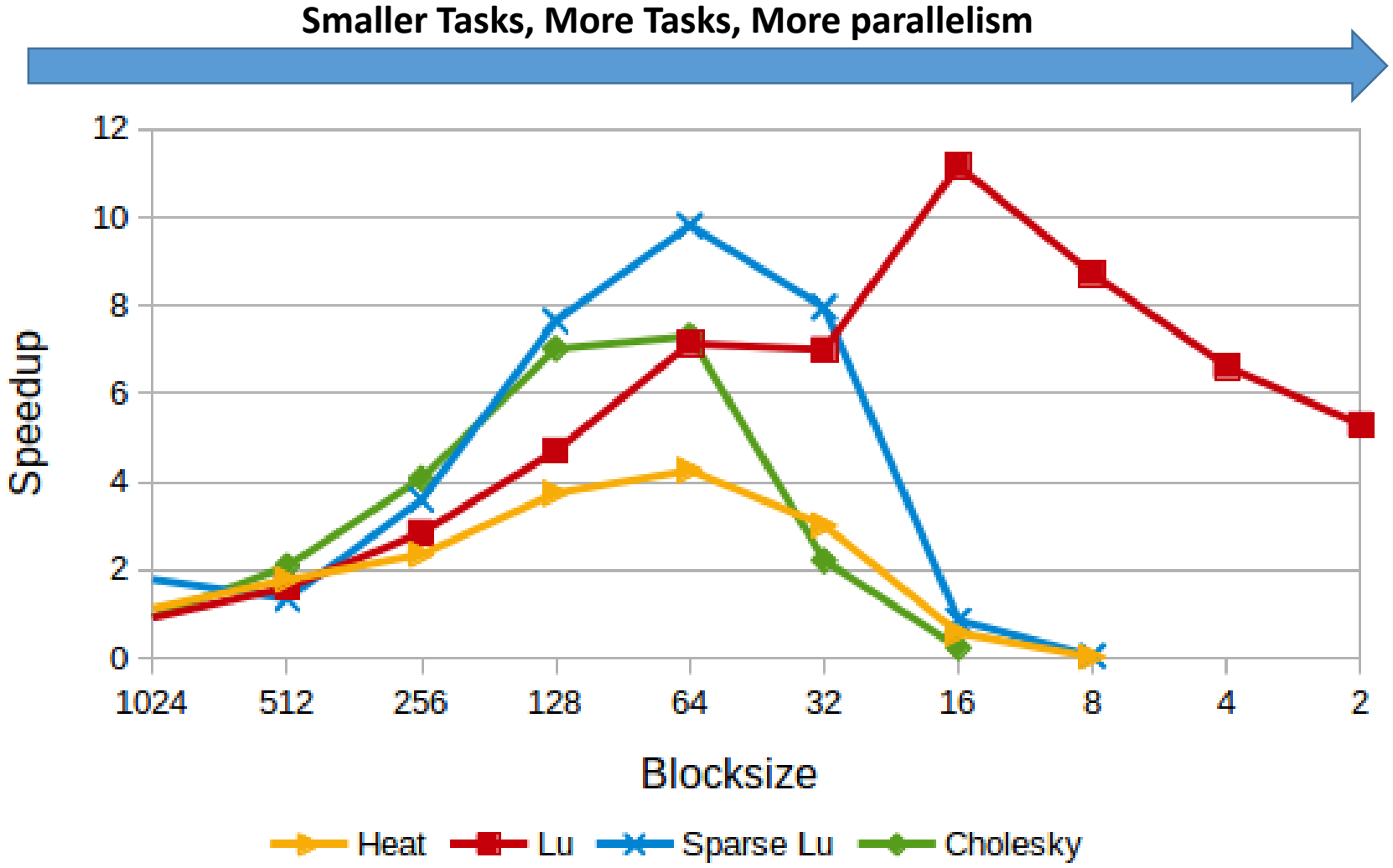
Some simple maths



- Each core feeds K tasks per time unit
- Each accelerator executes R tasks per time unit

$$Speedup \leq \frac{T_{exec}}{T_{overhead}} = \frac{K}{R}$$

Some simple example



Same problem size vs sequential execution

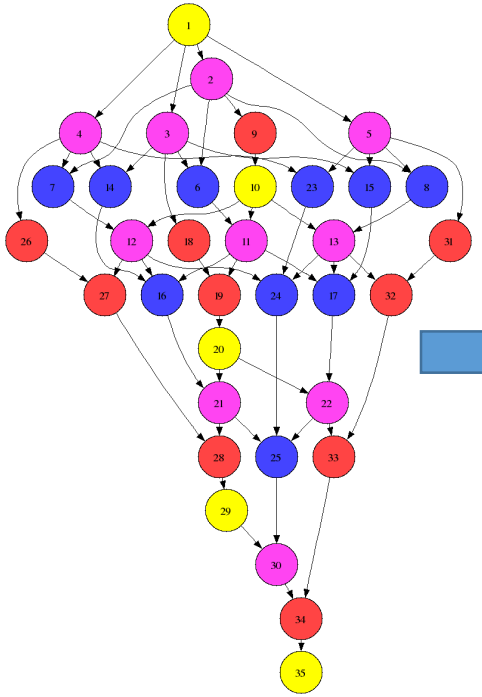
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

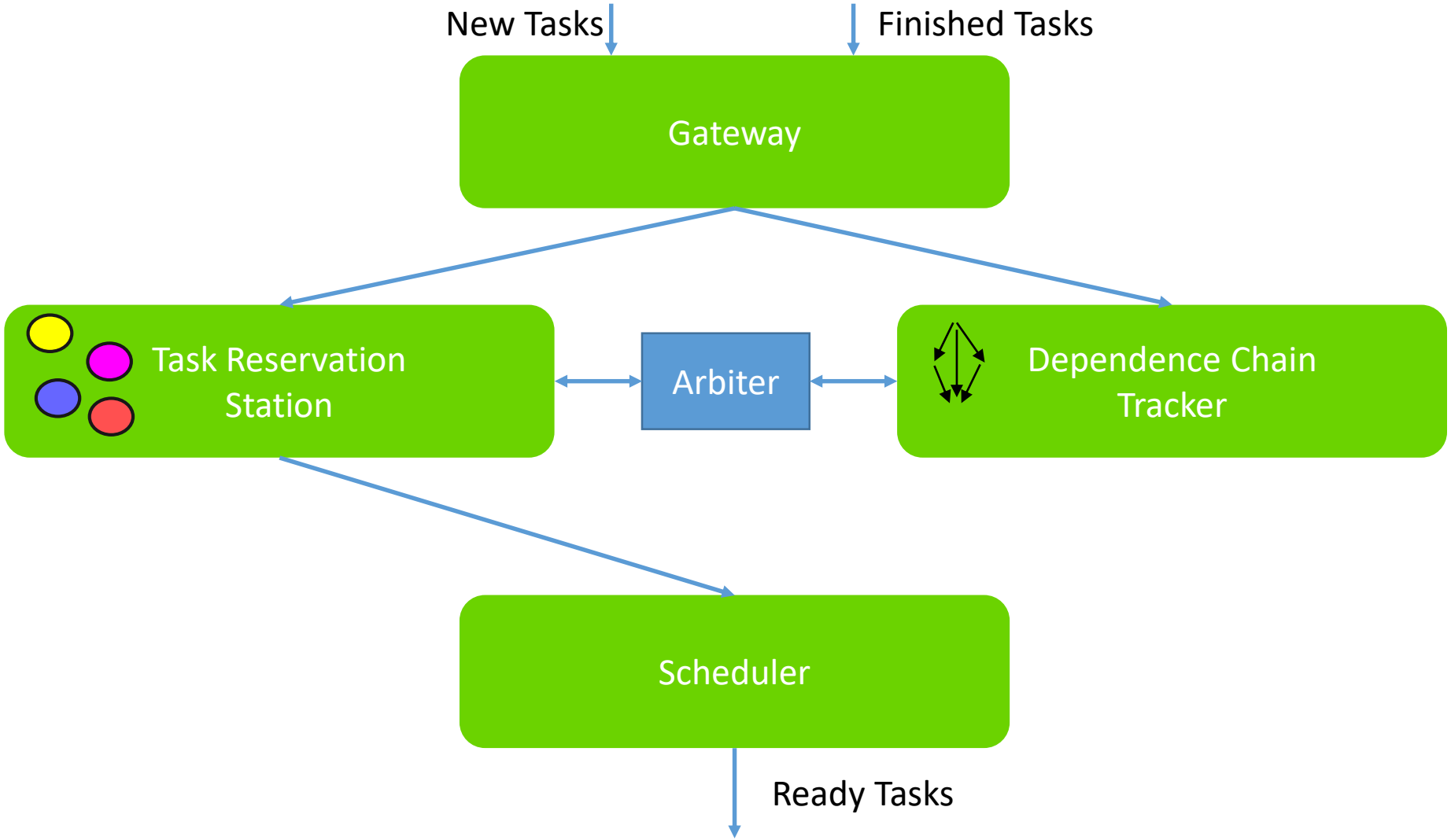
Picos: A hardware runtime architecture support for OmpSs. Future Generation Comp. Syst. 53: 130-139 (2015)

A Hardware Runtime

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

PICOS: A Hardware Task Dependence Manager

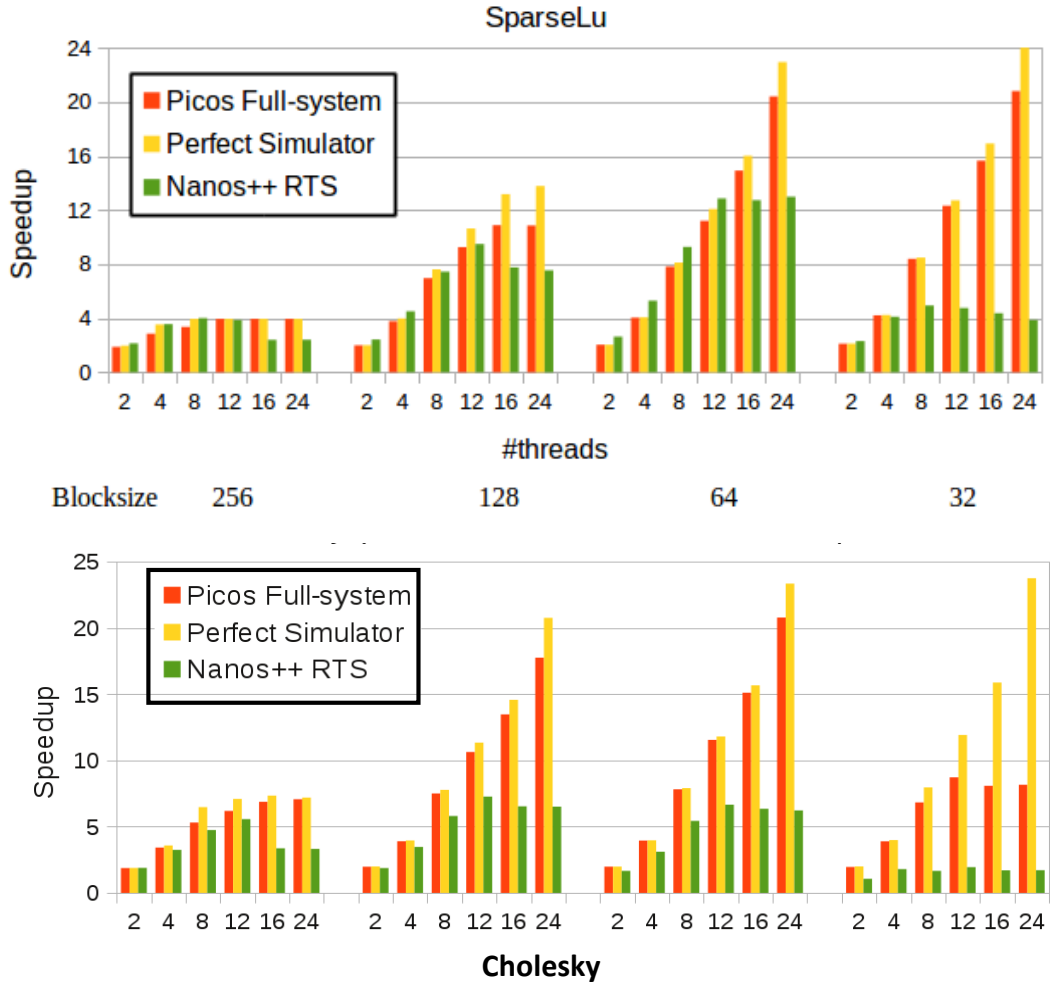




Barcelona Supercomputing Center
Centro Nacional de Supercomputación

General Purpose Task-Dependence Management Hardware for Task-Based Dataflow Programming Models. IPDPS 2017: 244-253

HW implementation early results

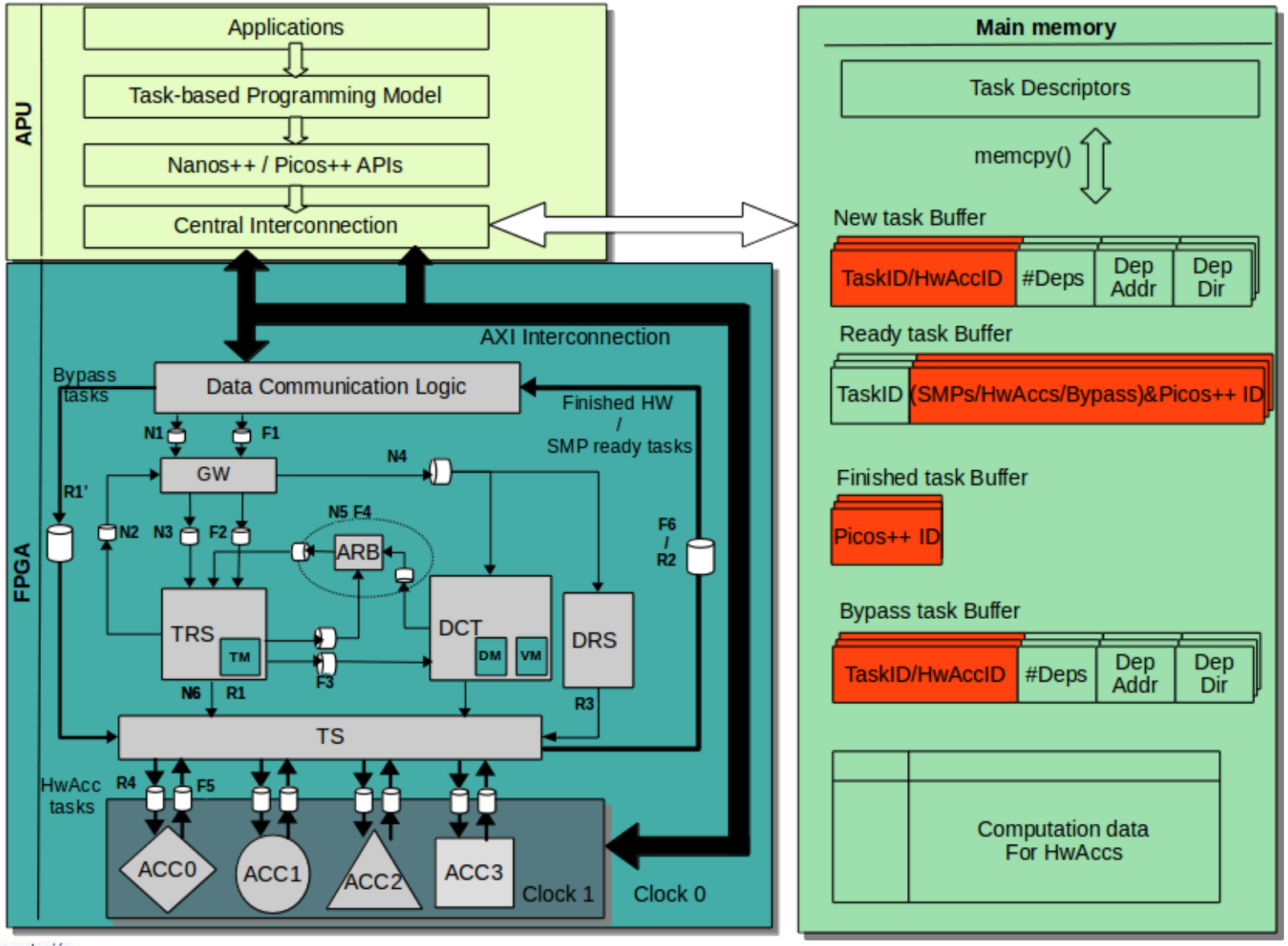


- Either smaller tasks or larger number of cores lead to diminishing performance returns
- The HW showed promising results
- Should test the real thing

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Performance analysis of a hardware accelerator of dependence management for task-based dataflow programming models. ISPASS 2016: 225-234

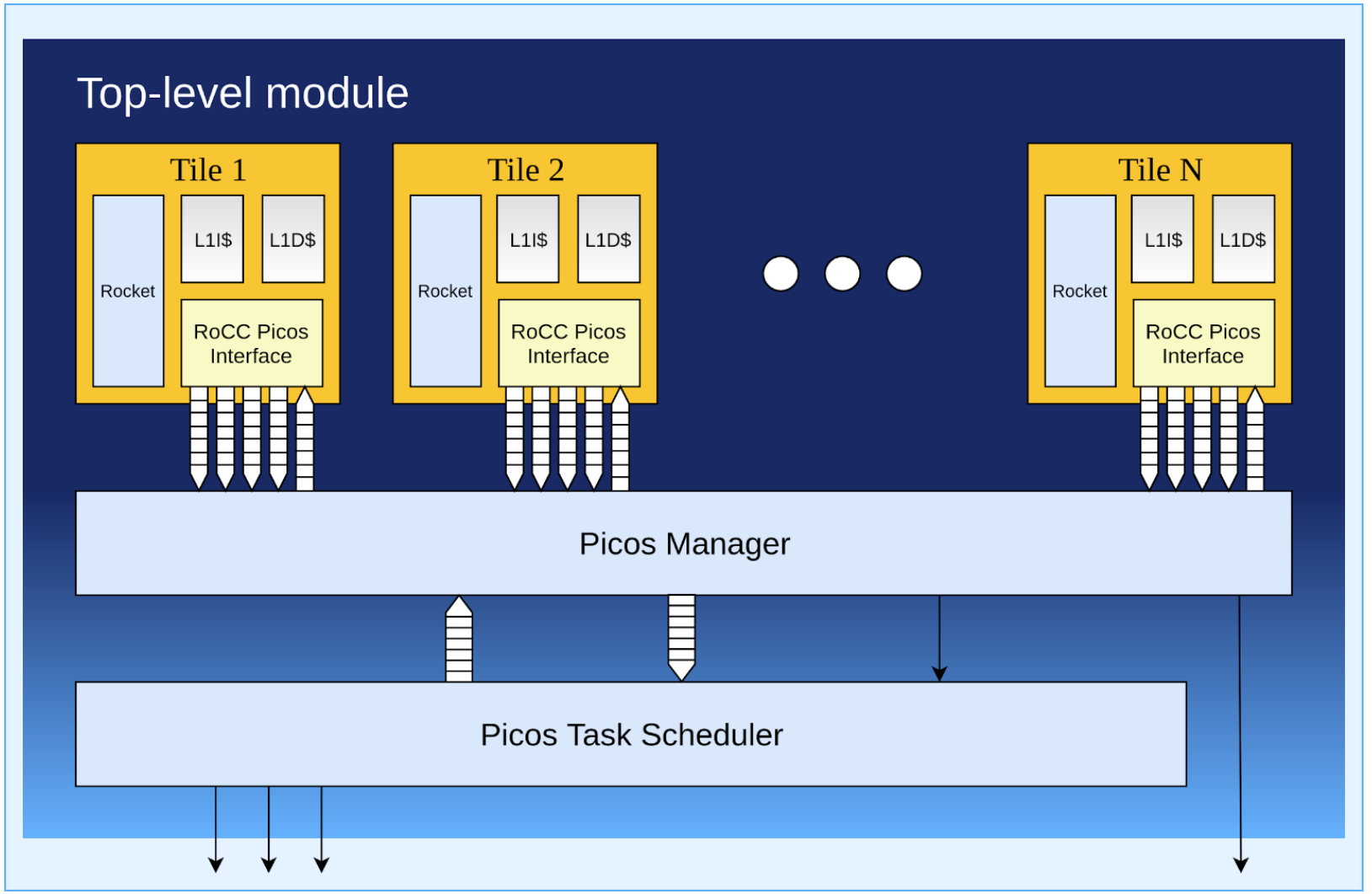
Full System Implementation



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

A Hardware Runtime for Task-Based Programming Models. IEEE Trans. Parallel Distrib. Syst. 30(9): 1932-1946 (2019)

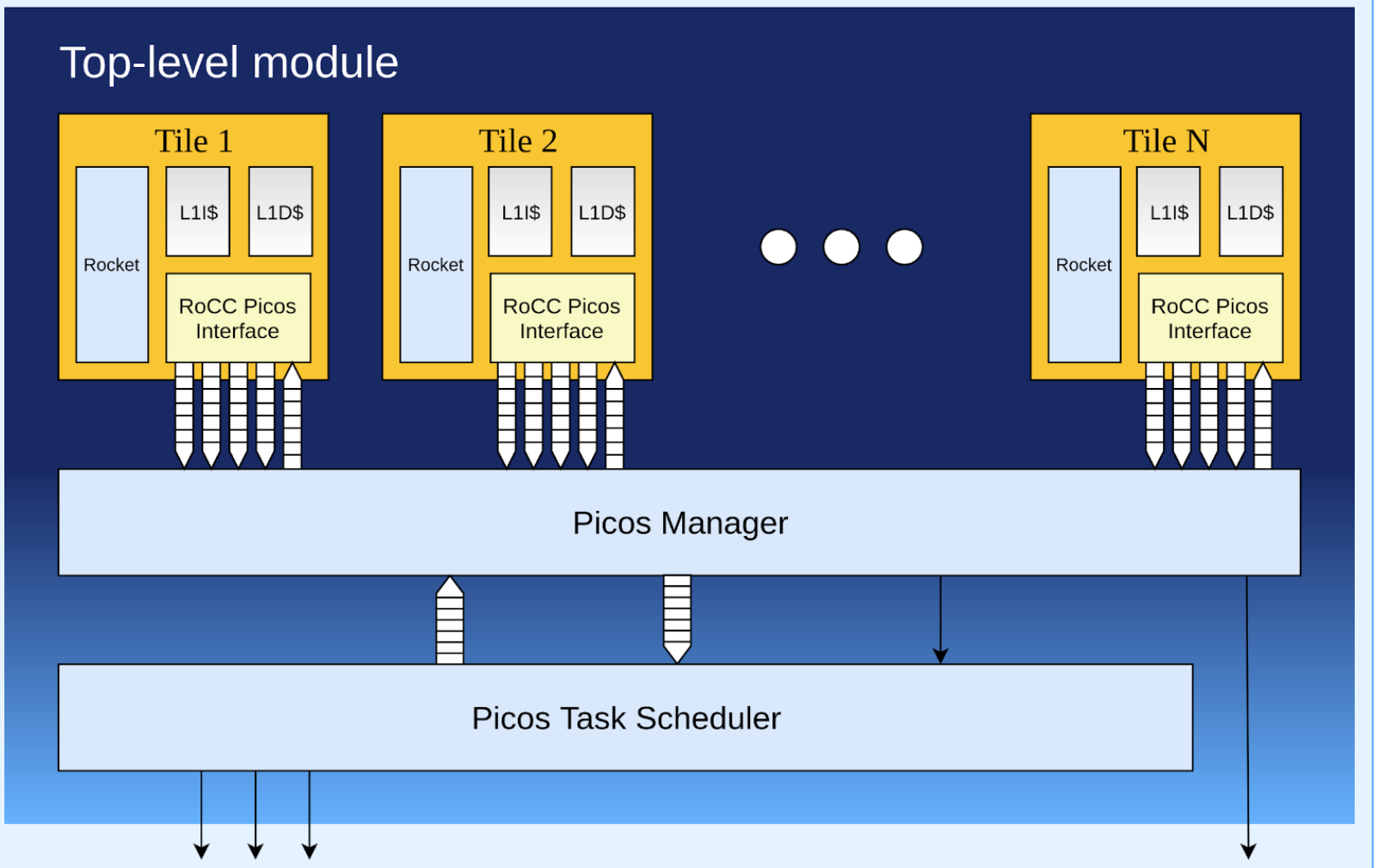
The “Real” Thing: RISC-V implementation



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

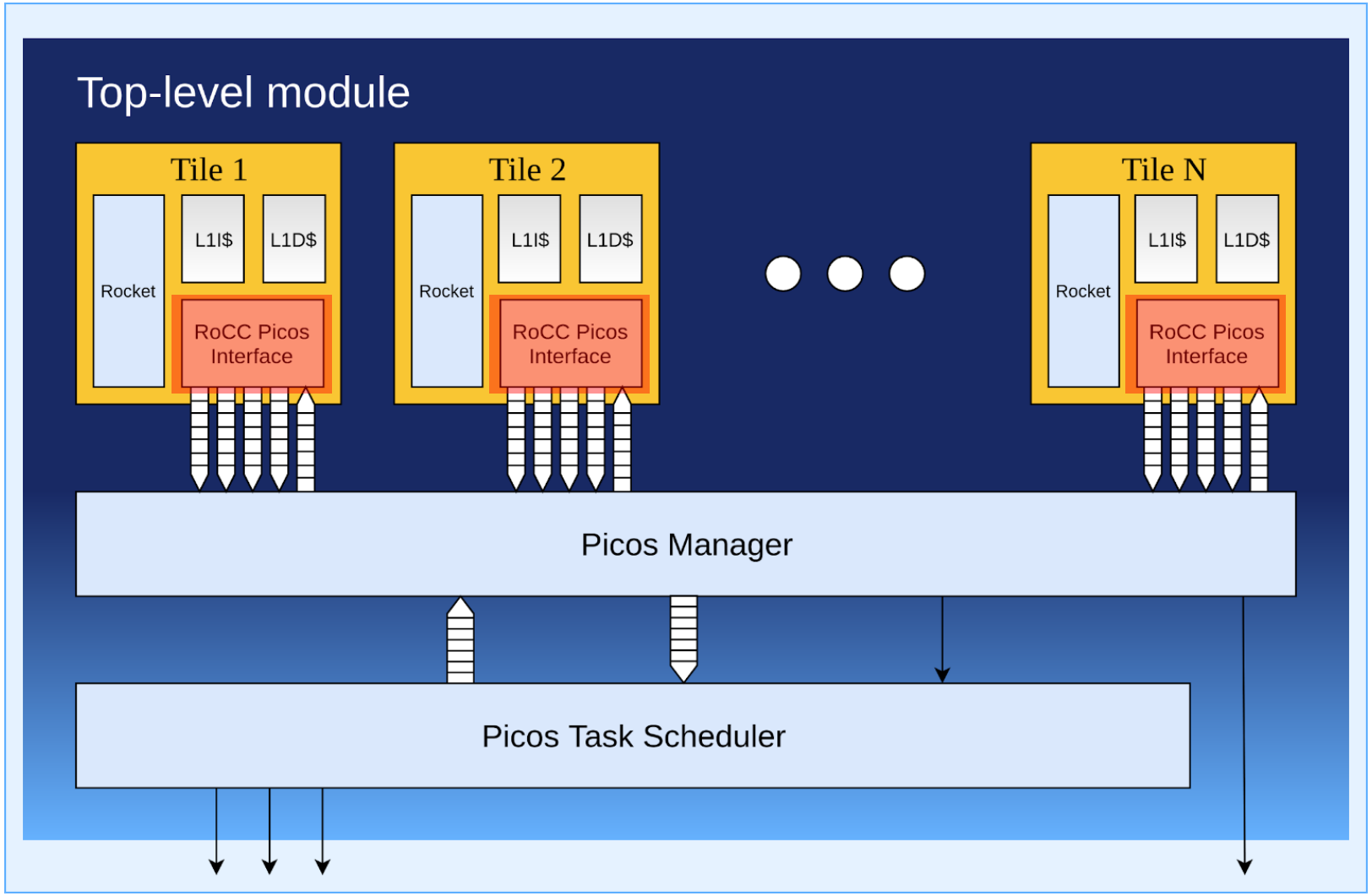
Adding Tightly-Integrated Task Scheduling Acceleration to a RISC-V Multi-core Processor. MICRO 2019: 861-872

The “Real” Thing: RISC-V implementation



Parameter	Value
Processor clock	80 MHz
# Cores	8
Processor characteristics	In-order, RISC-V priv. 1.10 compliant
RAM Size	1.25 GB
L1 I$ size	32 KB
L1 D$ size	32 KB
Cache characteristics	MESI cache coherence, L1-only, 8-way, Random Replacement
Compilation options	-O3

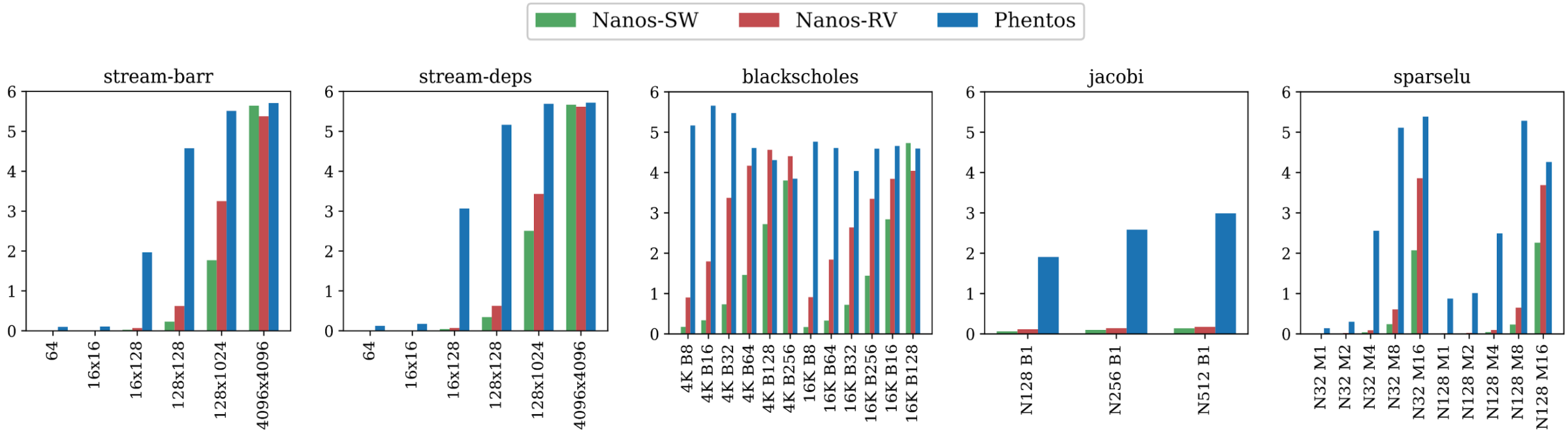
Scheduling through instructions



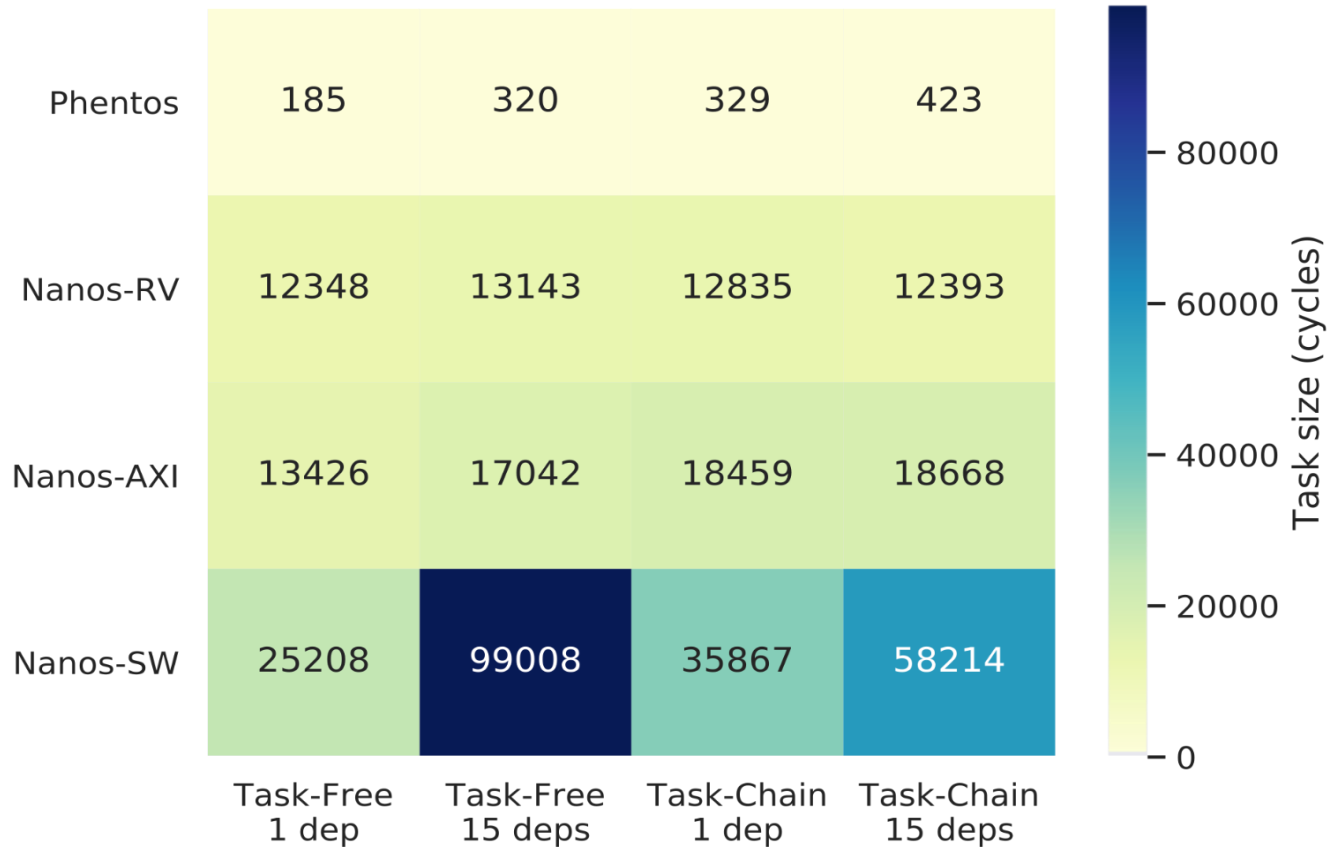
- Rocket RISC-V Chisel implementation (64 bits)
- Real system executing Linux
- New instructions use RoCC interface to Schedule, Retrieve and Retire Tasks
- Aim to reduce communication overheads between Hardware Task Manager and Risc-V cores

Speedup Results

Over serial execution

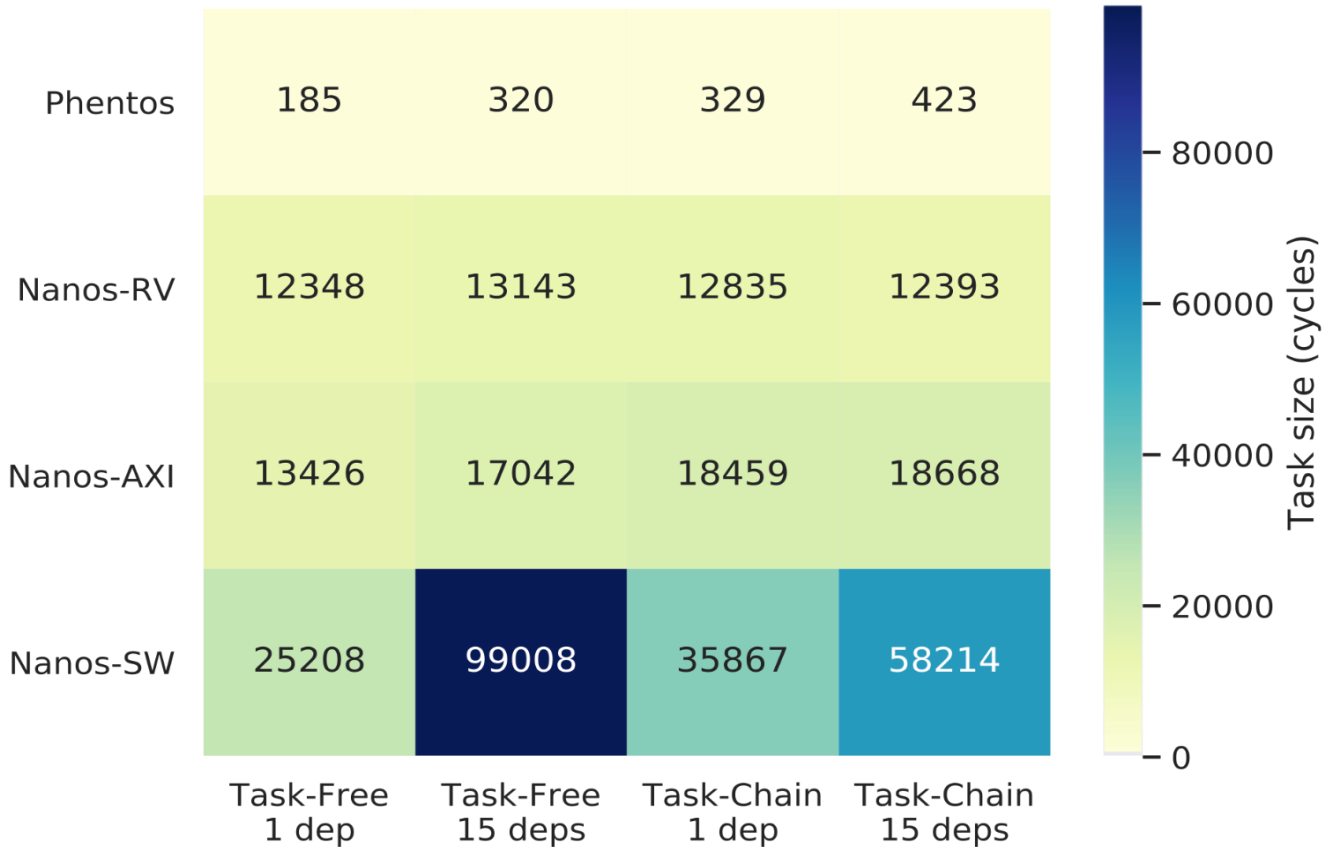


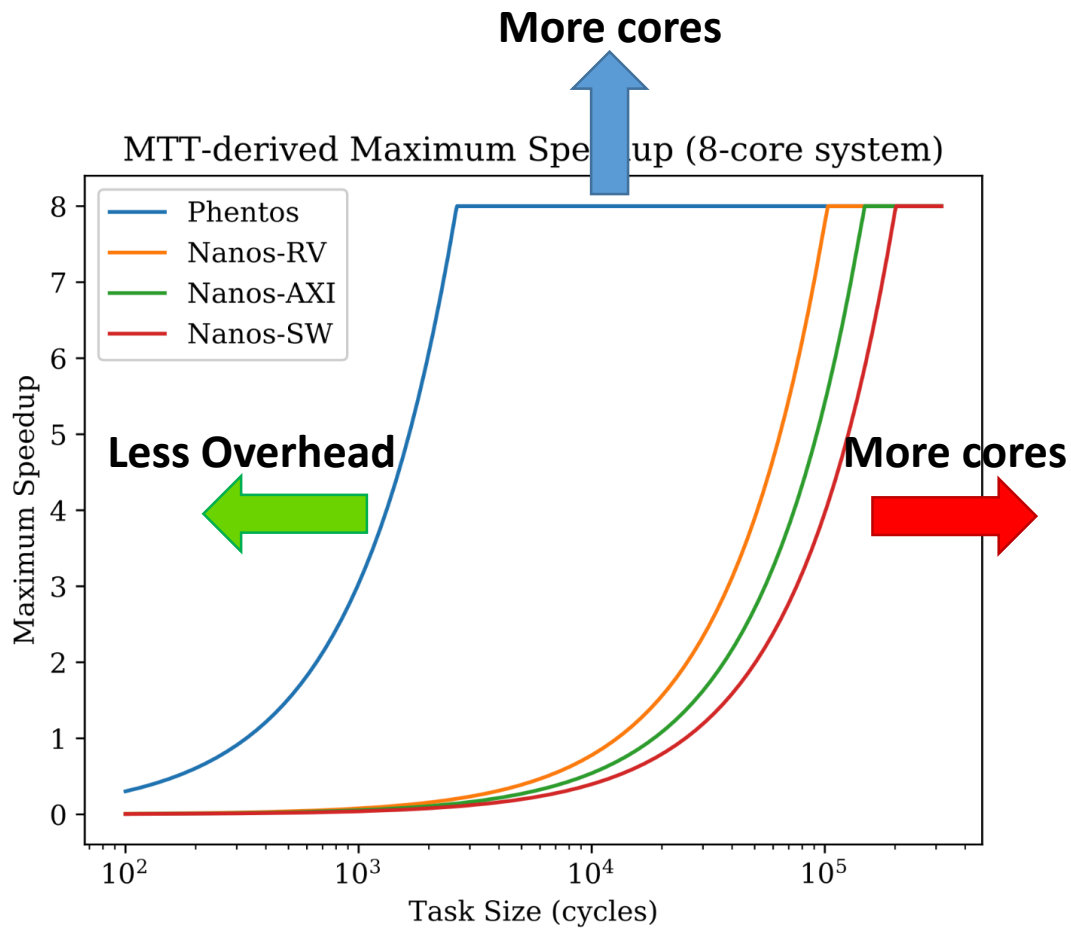
Life-time Overheads



- New simple runtime (Phentos) to use only HW runtime
- Nanos-AXI uses through memory communication
- Nanos-RV has still big software overheads (in the SW part of the runtime)
- Up to 308x lower overheads over Pure Software Runtime (Nanos-SW)

Life-time Overheads



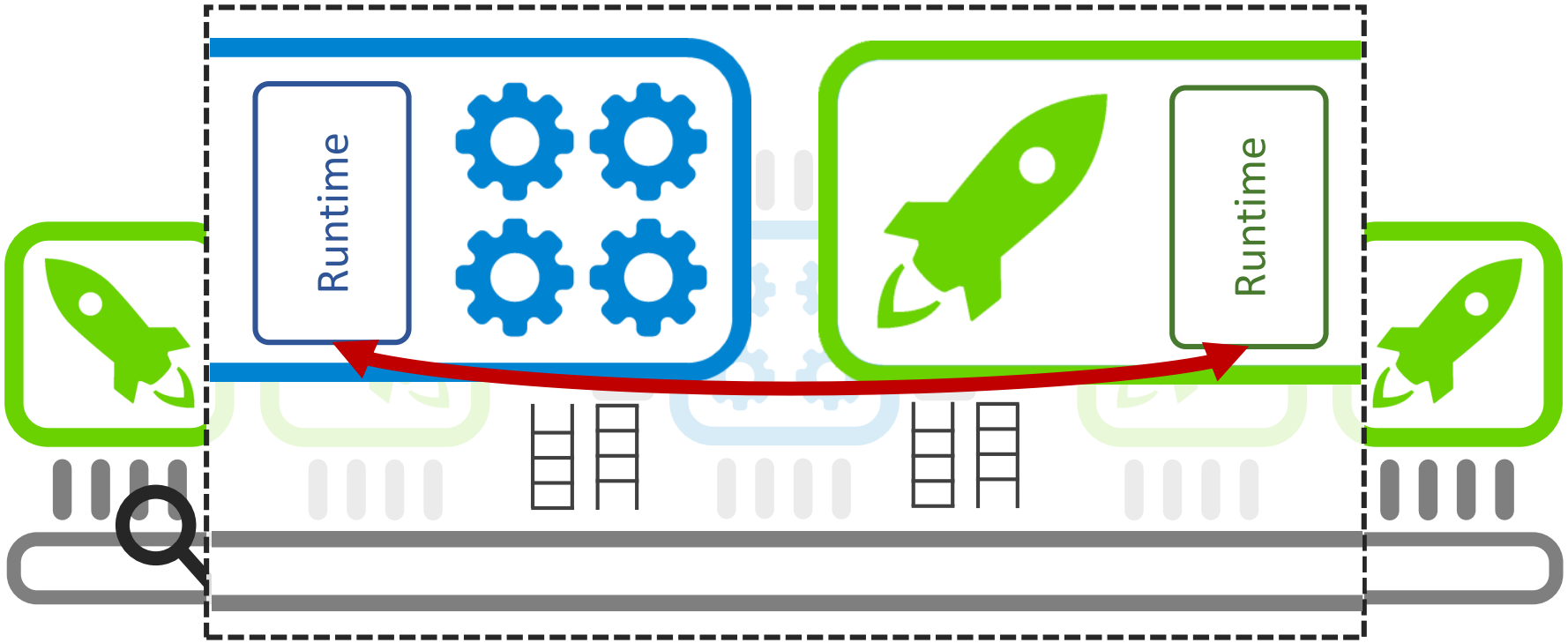


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Near Future

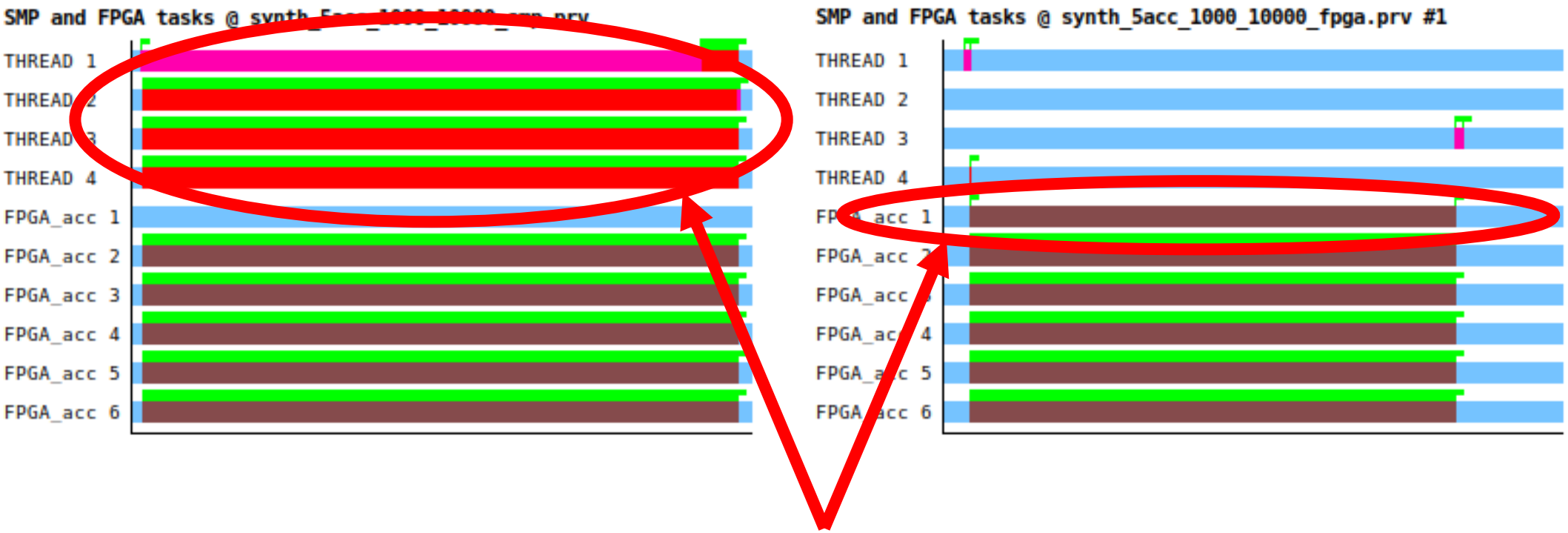
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Next accelerator model



Breaking Master-Slave Model Between Host and FPGAs. To appear at PPoPP 2020

Next accelerator model early results



One accelerator with task creation capabilities (instructions) + a task manager is **better** than 4 general purpose cores

Conclusions

- Task management is key to performance in future many-many-core systems
- RISC-V is the ideal target to introduce new HPC oriented instructions
- A Runtime Architecture can leverage several more cores than alternative software-only counterparts
- Designing a new hardware is a long and hazardous path… but worth it in the long run

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

Thank you

carlos.alvarez@bsc.es

References

- Xubin Tan, Jaume Bosch, Carlos Álvarez, Daniel Jiménez-González, Eduard Ayguadé, Mateo Valero: A Hardware Runtime for Task-Based Programming Models. IEEE Trans. Parallel Distrib. Syst. 30(9): 1932- 1946 (2019)
- Lucas Morais, Vitor Silva, Alfredo Goldman, Carlos Álvarez, Jaume Bosch, Michael Frank, Guido Araujo: Adding Tightly-Integrated Task Scheduling Acceleration to a RISC-V Multi-core Processor. MICRO 2019: 861-872
- Jaume Bosch, Xubin Tan, Antonio Filgueras, Miquel Vidal, Marc Mateu, Daniel Jiménez-González, Carlos Álvarez, Xavier Martorell, Eduard Ayguadé, Jesús Labarta: Application Acceleration on FPGAs with OmpSs@FPGA. FPT 2018: 70-77
- Xubin Tan, Jaume Bosch, Miquel Vidal, Carlos Álvarez, Daniel Jiménez-González, Eduard Ayguadé, Mateo Valero: Picos, A Hardware Task-Dependence Manager for Task-Based Dataflow Programming Models. HPCS 2017: 878-880
- Xubin Tan, Jaume Bosch, Miquel Vidal, Carlos Álvarez, Daniel Jiménez-González, Eduard Ayguadé, Mateo Valero: General Purpose Task-Dependence Management Hardware for Task-Based Dataflow Programming Models. IPDPS 2017: 244-253
- Jaume Bosch, Xubin Tan, Carlos Álvarez, Daniel Jiménez-González, Xavier Martorell, Eduard Ayguadé: Characterizing and Improving the Performance of Many-Core Task-Based Parallel Programming Runtimes. IPDPS Workshops 2017: 1285-1292

References II

- Xubin Tan, Jaume Bosch, Daniel Jiménez-González, Carlos Álvarez-Martínez, Eduard Ayguadé, Mateo Valero: Performance analysis of a hardware accelerator of dependence management for task-based dataflow programming models. ISPASS 2016: 225-234
- Fahimeh Yazdanpanah, Carlos Álvarez, Daniel Jiménez-González, Rosa M. Badia, Mateo Valero: Picos: A hardware runtime architecture support for OmpSs. Future Generation Comp. Syst. 53: 130-139 (2015)
- Fahimeh Yazdanpanah, Carlos Álvarez-Martínez, Daniel Jiménez-González, Yoav Etsion: Hybrid Dataflow/von-Neumann Architectures. IEEE Trans. Parallel Distrib. Syst. 25(6): 1489-1509 (2014)
- Fahimeh Yazdanpanah, Daniel Jiménez-González, Carlos Álvarez-Martínez, Yoav Etsion, Rosa M. Badia: Analysis of the Task Superscalar Architecture Hardware Design. ICCS 2013: 339-348
- Yoav Etsion, Felipe Cabarcas, Alejandro Rico, Alex Ramírez, Rosa M. Badia, Eduard Ayguadé, Jesús Labarta, Mateo Valero: Task Superscalar: An Out-of-Order Task Pipeline. MICRO 2010: 89-100