Evaluating the Performance and Energy Efficiency of the COSMO-ART Model System

Exa2Green
energy-aware numerics

Joseph Charles & William Sawyer (CSCS),
Manuel F. Dolz (UHAM), Sandra Catalán (UJI)

EnA-HPC, Dresden
September 1-2, 2014

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

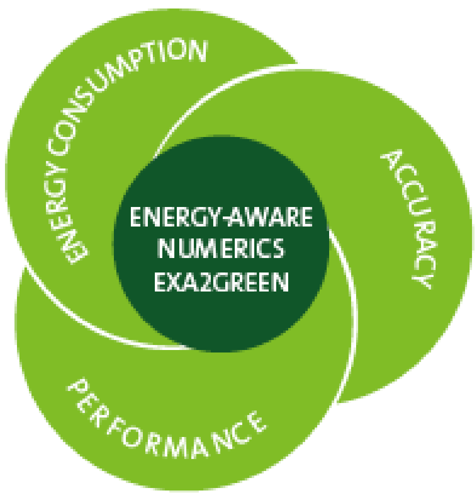
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

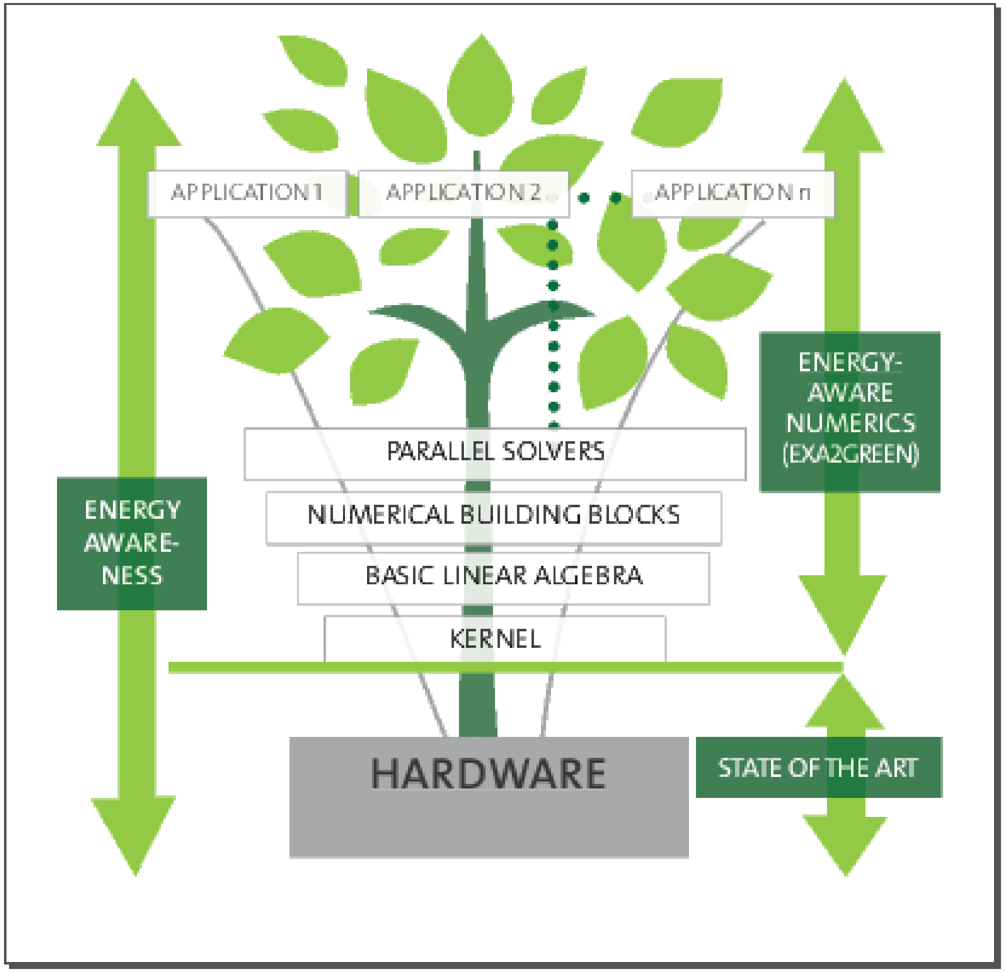
TECHNISCHE UNIVERSITÄT DRESDEN

ZIH
Zentrum für Informationsdienste und Hochleistungsrechnen

EU FP7-funded Exa2Green project

"Energy-aware sustainable computing on future technology"





http://exa2green-project.eu

Exa2Green energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

EU FP7-funded Exa2Green project

"Energy-aware sustainable computing on future technology"

- ***Seven European partners:***
  - University of Hamburg,
  - University of Jaume,
  - University of Heidelberg,
  - ETH Zurich / CSCS,
  - IBM Rueschlikon,
  - Karlsruhe Inst. of Tech.,
  - Steinbeis Innovation gGmbH
- ***Human resources:***

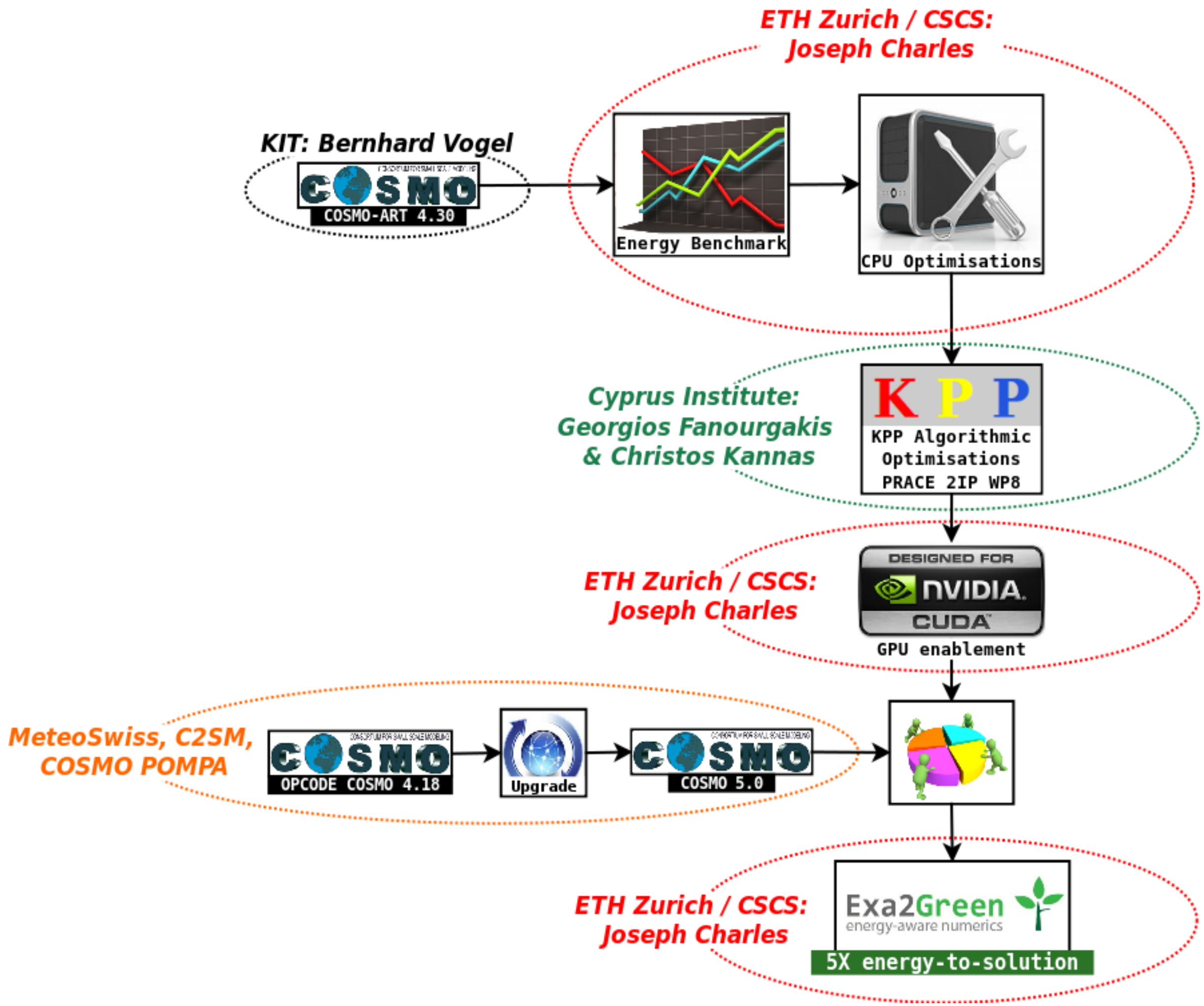
  36 PMs for CSCS, 261.6 PMs overall
- ***Framework:***

  Covers all essential fields of expertise in energy-efficient computing
- ***Showcase application:***

  COSMO-ART
- ***Ultimate goal:***

  5x improvement in energy-to-solution over baseline
- ***Boundary conditions:***

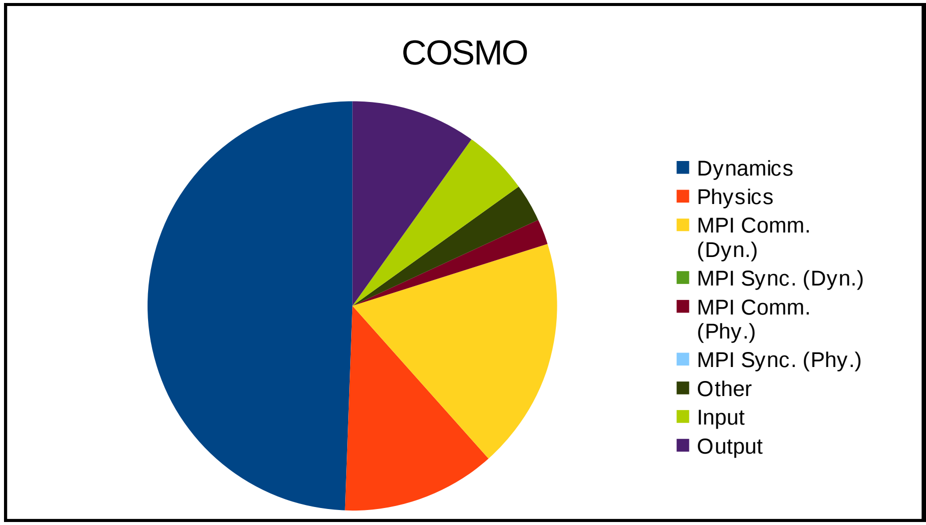
  leverage off of HP2C COSMO work, maximize benefit for Swiss climate community

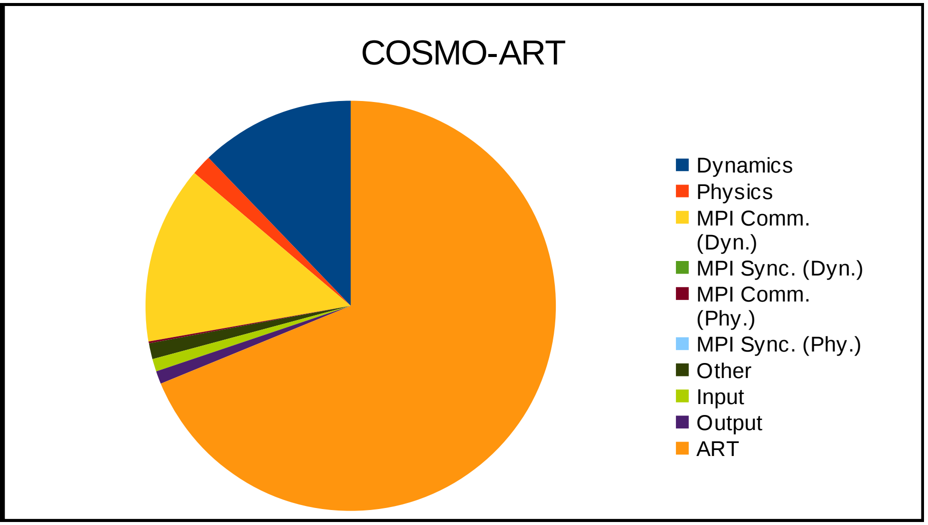


Exa2Green energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

Atmospheric chemistry as showcase





- ***COSMO:* *an ubiquitous weather forecast model in Europe***

Widespread use in federal weather forecast stations in Germany, Switzerland, Italy, Greece, Poland, Romania and Russia and large number of agencies including military and research institutions

- ***COSMO-ART:* *COSMO extended for Aerosols and Reactive Trace gases***

Massive increase in computational expense due to atmospheric chemistry and additional tracers to advect (only relatively short simulation times currently viable)

Exa2Green energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

COSMO-ART model setup

- 24-hour forecast simulations over Europe from April 13th 2010 (near the equinox).
- CORDEX-EU-44 calculation domain (grid of 222 x 216 x 40 points).
- ECMWF global spectral model IFS with an update frequency of 3h for meteorological initial and boundary conditions.
- IFS-MOZART output at 6h temporal resolution for boundary data for gas-phase species.
- 34 2D and 45 3D fields to be written out every hour.
- No data assimilation methods.
- Semi-Lagrangian horizontal advection scheme with tricubic interpolation and selective filling diffusion option in combination with the dynamical core using Runge-Kutta time stepping.
- Kinetic PreProcessor solver (KPP) for the solution of atmospheric chemistry ordinary differential equations.
- Includes indirect cloud feedbacks but does not take into account below cloud scavenging (washout) and does not include in-cloud scavenging (rainout).

Precipitation formation is performed by a two-moment cloud microphysics of Seifert and Beheng.

Environment setup

MONCH (CSCS – ETH Zurich) - 1040 cores using 20 MPI tasks per node

10-rack NEC-provided cluster composed of 312 standard compute nodes. A subset of 52 was used, each comprised of two Intel Xeon Ivy Bridge EP E5-2660v2 ten-core processors operating at 2.2GHz, equipped with 32GB of DDR3 1600MHz RAMand connected via InfiniBand Mellanox SX6036 and FDR switches.

PILATUS (CSCS – ETH Zurich) - 672 cores using 16 MPI tasks per node

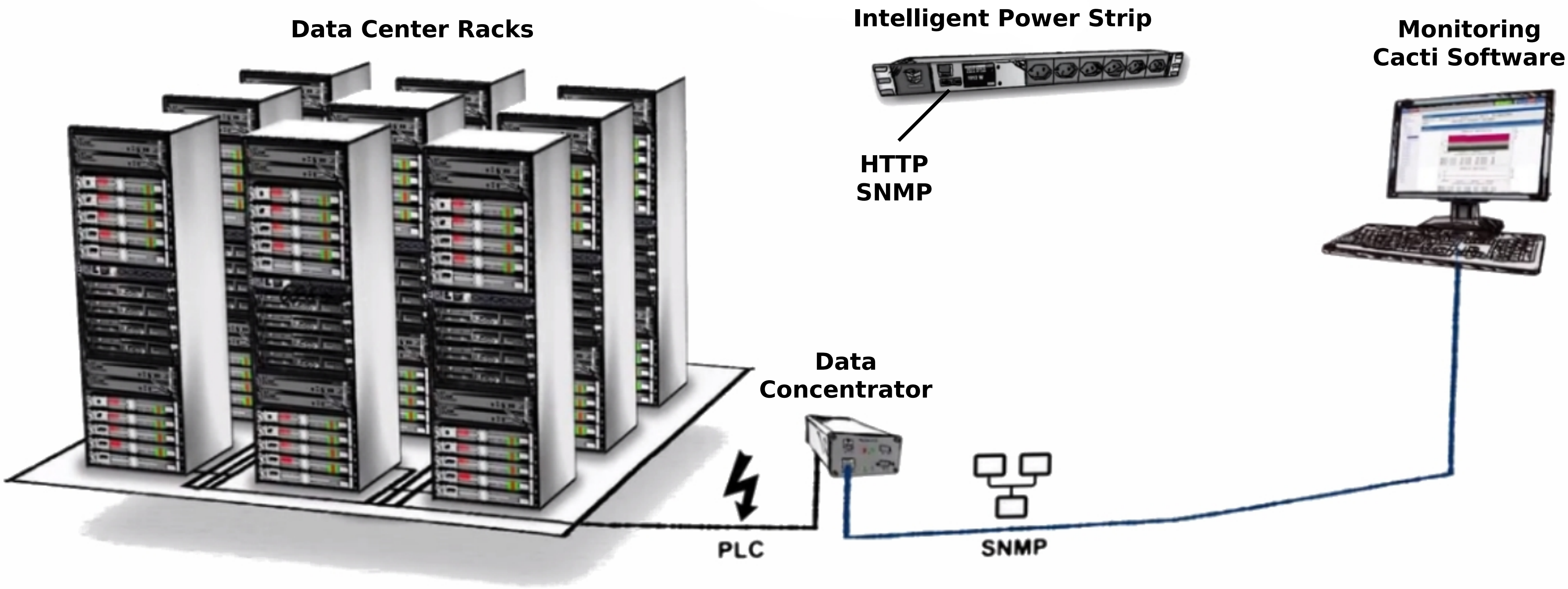
Cluster composed of 42 compute nodes. Each of them is comprised of two Intel Xeon Sandy Bridge EP E5-2670 eight-core processors operating at 2.6GHz equipped with 64GB of DDR3 1600MHz RAM and connected via InfiniBand Mellanox SX6036 and FDR switches.

TINTORRUM (UJI) - 192 cores using 12 MPI tasks per node

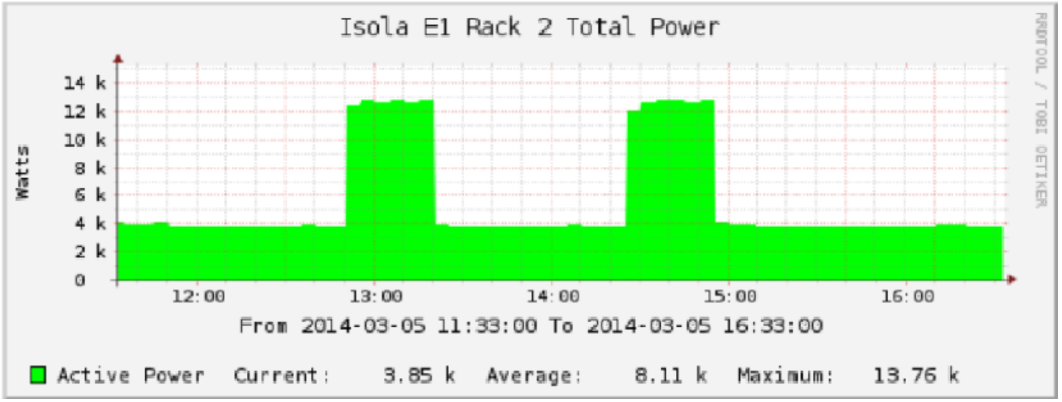
Heterogeneous cluster composed of 28 compute nodes. A subset of 16 homogeneous nodes was considered, each comprised of two Intel Xeon Westmere EP E5645 hexa-core processors running at 2.4GHz, equipped with 24GB of DDR3 1333MHz and connected via InfiniBand QDR with a Mellanox MTS-3600 switch.

Power measurement framework

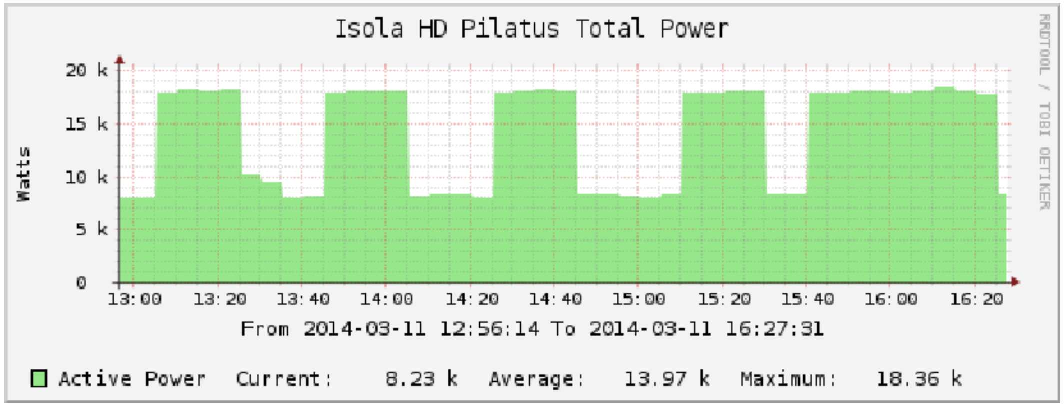
E3METER metering framework (Riedo Networks)



Time-power-energy analysis

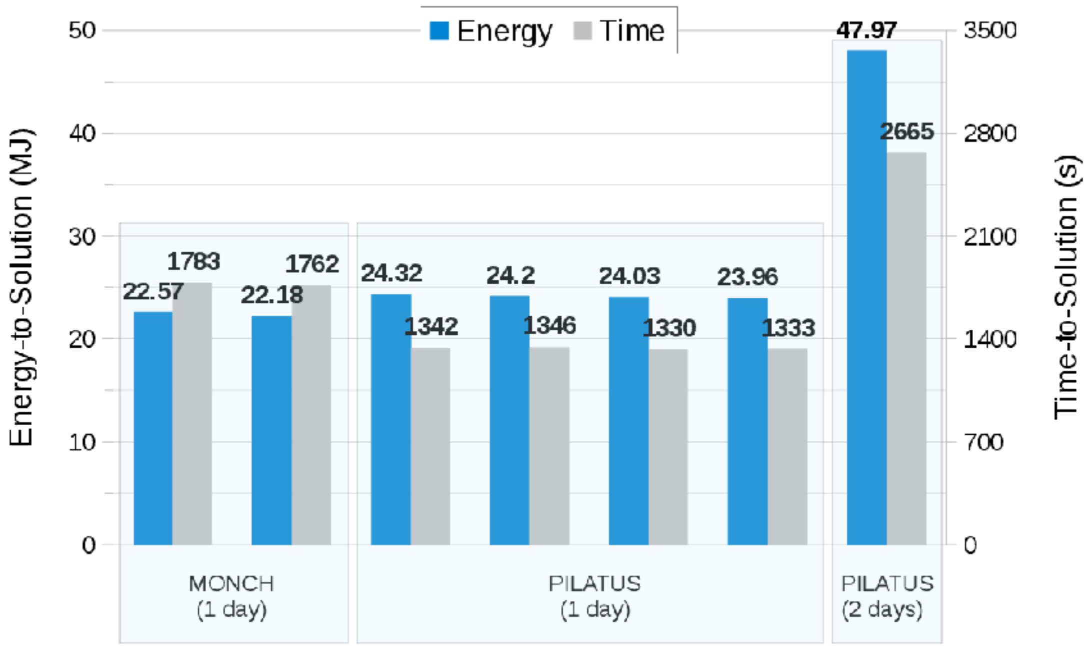


Isola E1 Rack 2 total power consumption of Monch



Isola HD total power consumption of Pilatus

Time-power-energy analysis



Time-to-solution and energy-to-solution comparison between Xeon E5 and Ivy Bridge-EP architectures for a 24h simulation

PILATUS	MONCH	TINTORRUM (*Aggressive*)	TINTORRUM (*Degraded*)
18035.0	12622.5	3713.6	3651.8

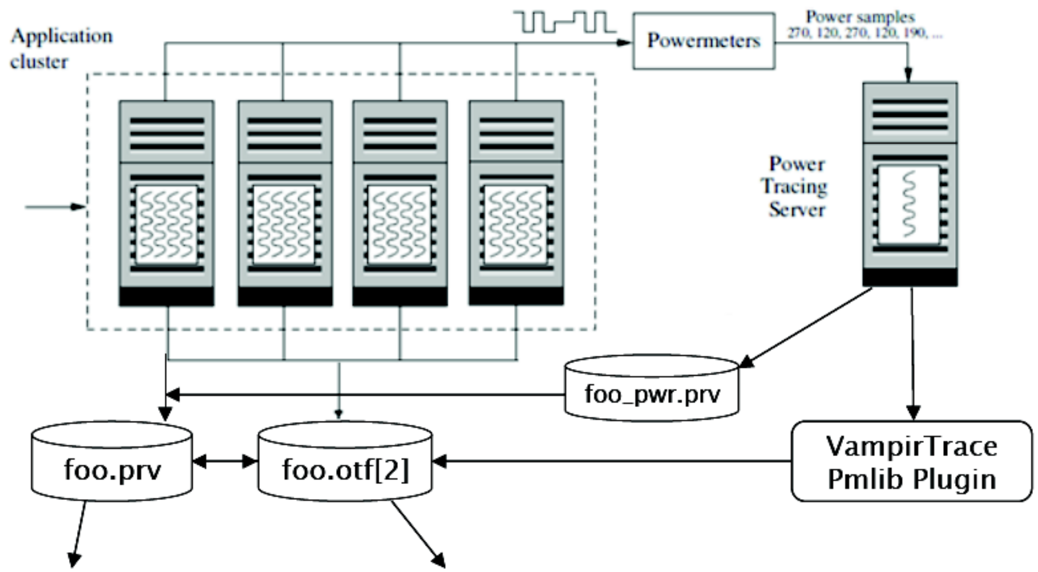
Averaged power consumption (W) of the platform

Exa2Green energy-aware numerics
SEVENTH FRAMEWORK PROGRAMME

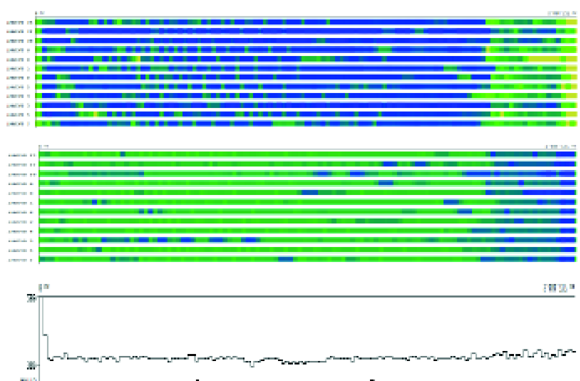
Power-performance tracing framework

Instrumented application:

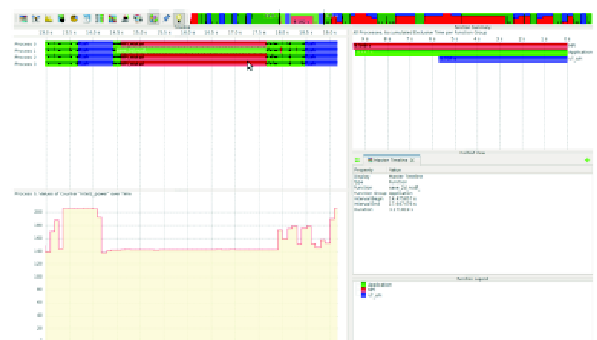
- **VampirTrace:**
  - Automatic instrumentation for serial, OpenMP, MPI, hybrid apps
  - Also Manual, Source and Binary instrumentations
  - Format: .otf
- **Score-P:**
  - Same instrumentation modes as VampirTrace
  - Format: .otf2 by default
- **Extrae:**
  - Automatic / Manual instrumentation
  - Format: .prv



2 Visualization tools:



Visualization with Paraver



Visualization with Vampir

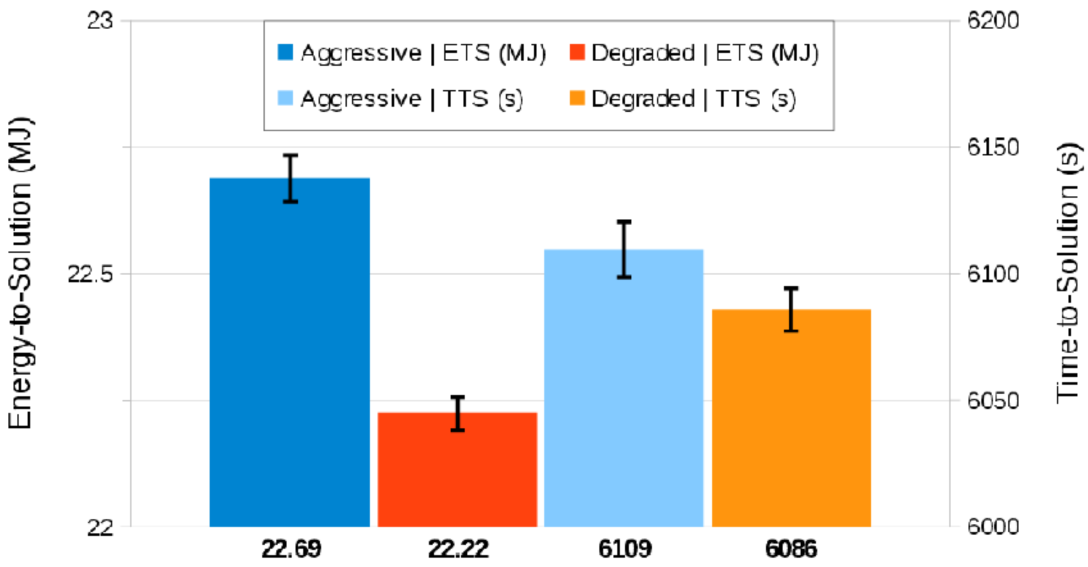
Exa2Green energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

Performance / Energy-efficiency of COSMO-ART

Use of a MPI energy-saving technique:

- MPI engine policies of OpenMPI 1.6.5
  - Aggressive: for exactly-/under-subscribed modes
    - Busy-waiting when waiting for a an incoming MPI message
  - Degraded: for over-subscribed modes
    - Repeatedly calls to sched_yield() to be picked again by the OS scheduler



TINTORRUM (*Aggressive*)	TINTORRUM (*Degraded*)
3713.6 W	3651.8 W

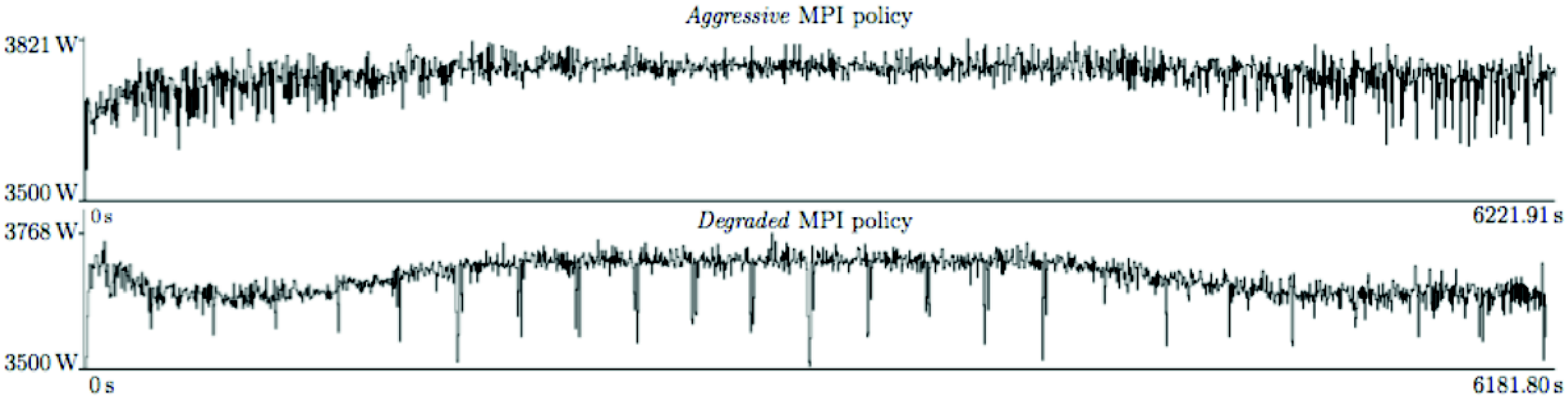
Aggressive utilizes 100% of CPU when busy-waiting!

Degraded policy only employs 50% of CPU!

Performance / Energy-efficiency of COSMO-ART

Use of a MPI energy-saving technique:

Power profile of a 24-hour simulation of COSMO-ART with Aggressive and Degraded policies



Systematic power drops each simulated hour with degraded mode

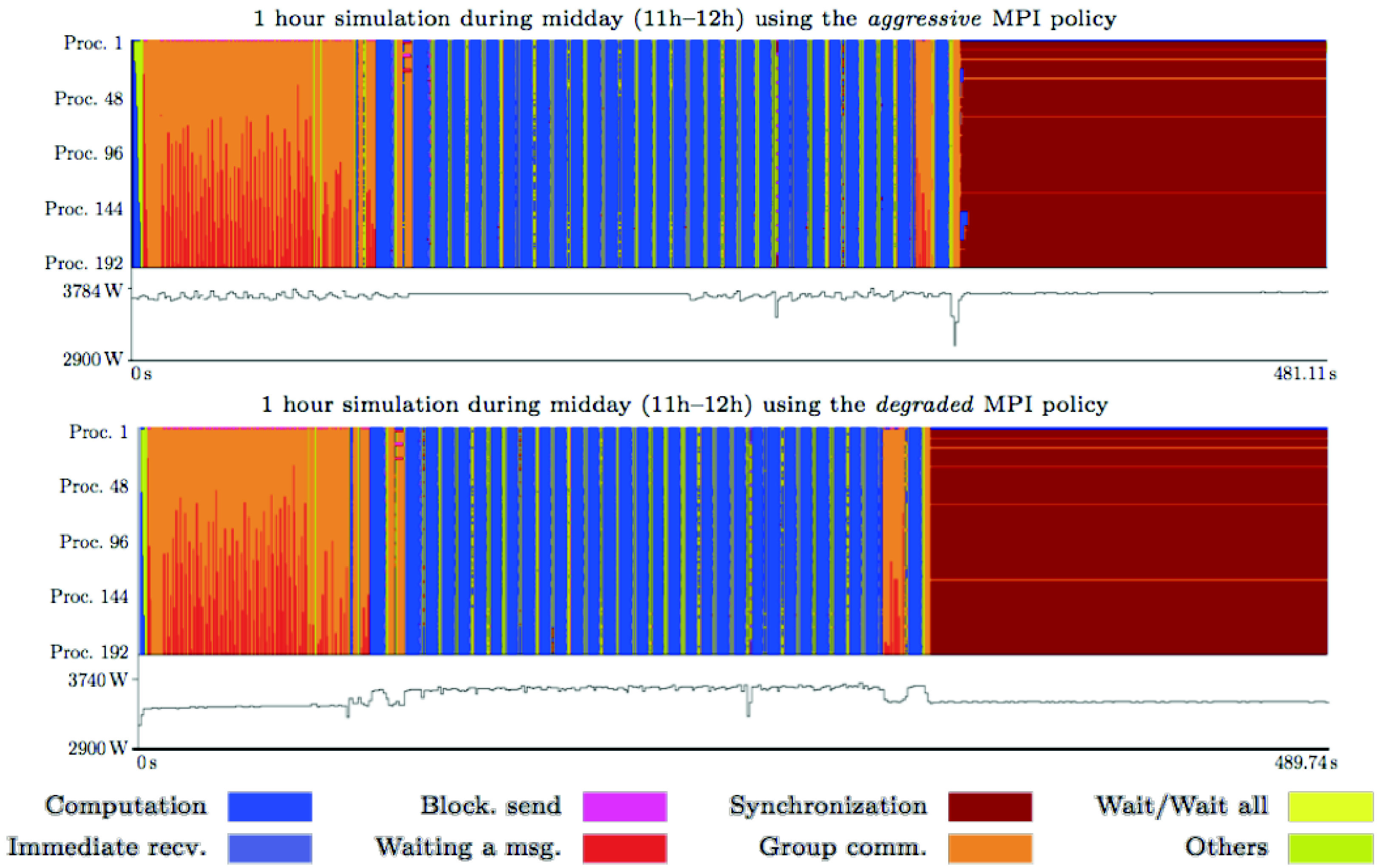
The cores were utilized (load) only ≈ 50%!

Reduction of the total power/energy consumption of 2%

Exa2Green
energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

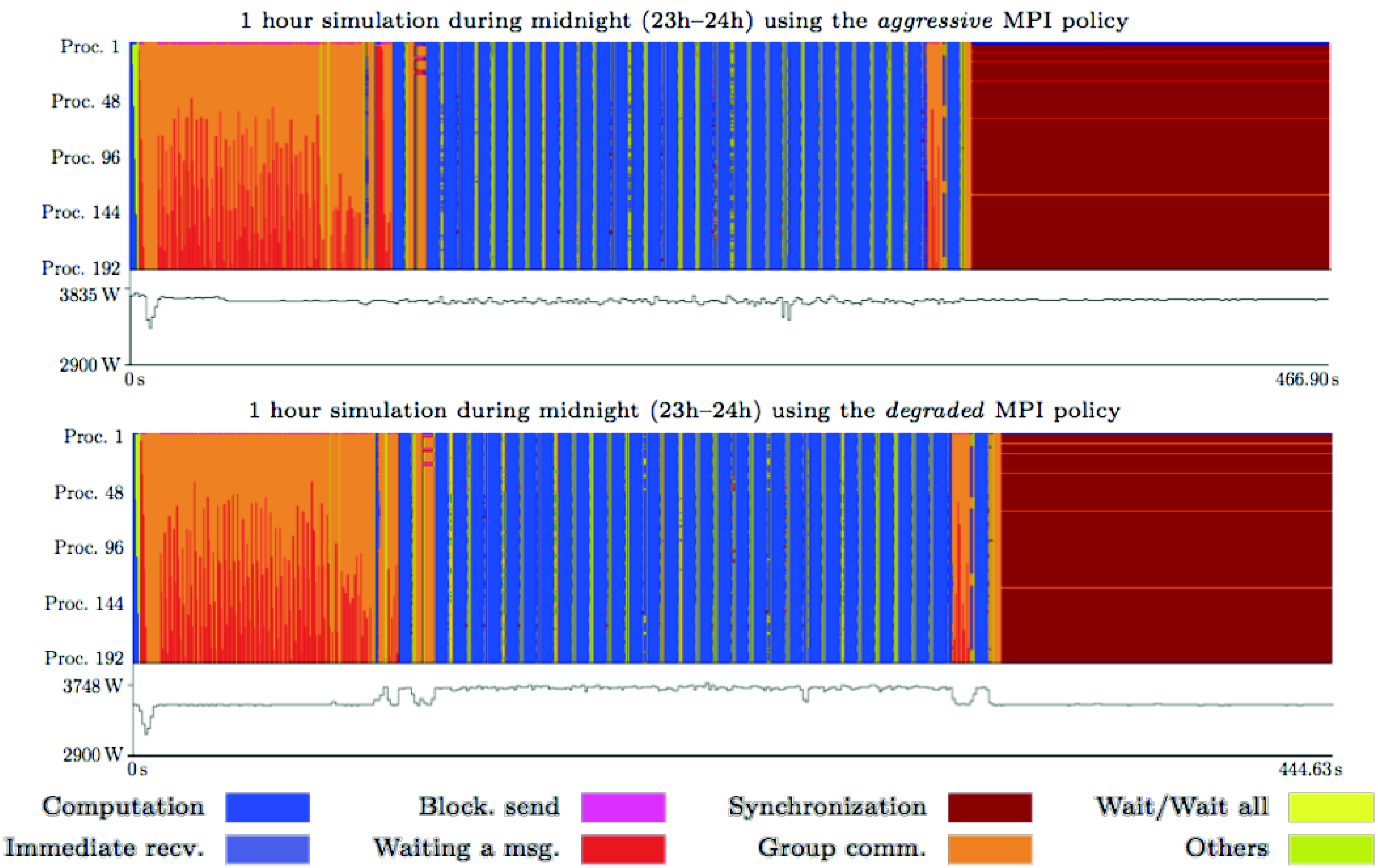
Performance / Energy-efficiency of COSMO-ART



Exa2Green energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

Performance / Energy-efficiency of COSMO-ART



Exa2Green
energy-aware numerics

SEVENTH FRAMEWORK PROGRAMME

Thank you for your attention!

Questions?