Tuning Alya with READEX for Energy-Efficiency

Venkatesh Kannan[1], Ricard Borrell[2], Myles Doyle[1], Guillaume Houzeaux[2]

[1]Irish Centre for High-End Computing (ICHEC)
[2]Barcelona Supercomputing Centre (BSC)

Context

- Application Dynamism
  - HPC applications exhibit dynamic behaviour at run time
    - Dynamic workload characteristics
    - Dynamic resource requirements
- READEX
  - Runtime Exploitation of Application Dynamism for Energy-efficient eXascale computing
  - EU Horizon 2020 FETHPC project to develop tuning techniques and tool suite
  - Tune HPC stack (hardware, system-software and application parameters) during an application run to improve energy consumption and/or performance
- Alya
  - High performance computational mechanics application
  - Developed by EoCoE at BSC with hybrid parallelization (MPI, OpenMP, CUDA, OpenACC)
  - Presence in multiple European application benchmark suites

READEX

Project and Tool Suite

Technische Universität Dresden/ZIH (Coordinator)

Norges Teknisk-Naturvitenskapelige Universitet

Technische Universität München

IT4Innovations, VSB-Technical University of Ostrava

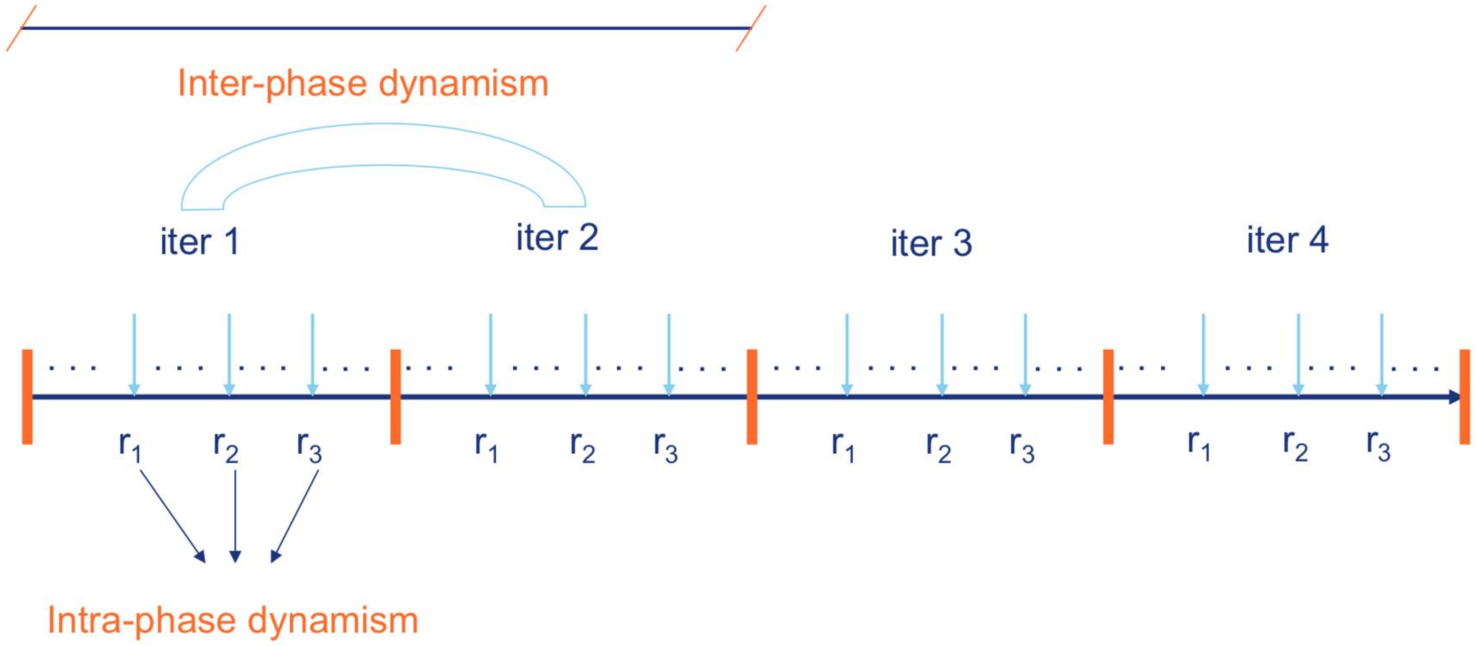
Irish Centre for High-End Computing

Intel Corporation SAS

Gesellschaft für numerische Simulation mbH

READEX Tool Suite

Dynamism Detection



READEX Tool Suite

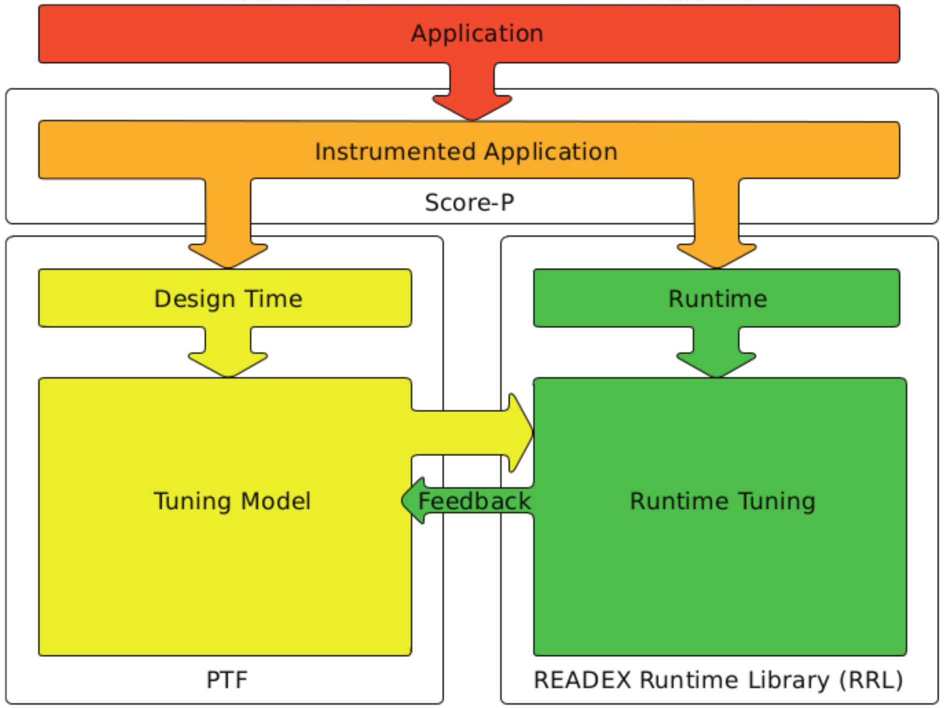
Tuning Parameters

Level	Tuning Aspect	Tuning Parameter	Scope
Hardware Parameters	CPU Frequency Controls	Core Frequency (DVFS)	Core
		Uncore Frequency	Socket
		Energy Performance Bias (EPB)	
System Software Parameters	OpenMP Parallelism	Dynamic Concurrency Throttling	Process
		Workload Scheduling Algorithm	
Application Parameters	User-specified Code-paths	Decomposition Routines	Application
		Types of Solvers	
		Preprocessing Stiffness & Course-problem matrices	
		…	

READEX Tool Suite

Steps to Apply on a HPC Application

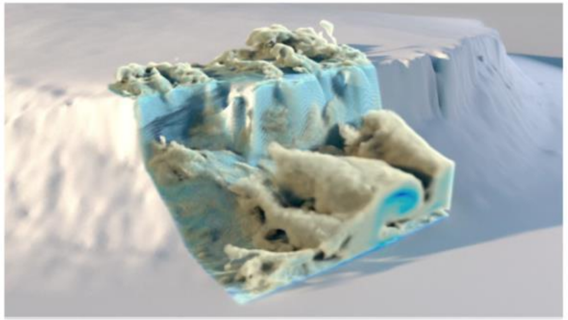
1. Application Instrumentation
   - Identify key application regions
2. Dynamism Detection
   - Check application for exploitable dynamism
3. Design-Time Analysis
   - Perform experimental runs
   - Identify optimal configurations
     - Summarised in *tuning model*
4. Runtime Application Tuning
   - Production run of application
   - Apply optimal configurations from *tuning model*
   - (Update *tuning model* using *calibration* for new scenarios)

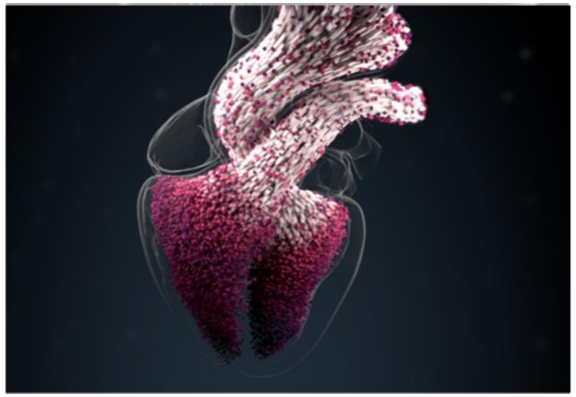


Alya

HPC Application

- Developed at BSC (Barcelona Supercomputing Centre)
- Simulation code system
  - High-performance computational mechanics
  - Hybrid parallelization (MPI, OpenMP, CUDA, OpenACC)
  - Present in industrial applications and European benchmark suites
- EoCoE (Energy oriented Centre of Excellence)
  - Renewable energy analysis (wind resource assessment)
  - Air quality simulation in cities
  - Full aircraft and turbine simulation for aeronautics
  - Cardiovascular and respiratory system simulation for biomechanics
  - And more





READEX on Alya

Dynamism Detection

```
Significant region information
===============================
Region name               Min(t)        Max(t)       Time       Time Dev.(%Reg)     Ops/L3miss     Weight(%Phase)
NSI_EIGEN_TIME_STEP_ALL   0.234         0.245        2.360          1.4             115143011            5
PAR_BARRIER               0.000         1.074       31.029          0.0                 0               65
DEFLCG                    0.233         0.403        8.355         16.9              4183031            18
NSI_DOMMAS                0.216         0.216        2.159          0.0                200               5
```

```
Phase information
=================
Min                 Max                 Mean                Time                Dev.(% Phase)       Dyn.(% Phase)
4.45103             7.02546             4.76904             47.6904                 0                  53.9821
```

```
SUMMARY:
========
Inter-phase dynamism due to variation of the execution time of phases

Intra-phase dynamism due to time variation(%) of the following important significant regions
    DEFLCG
Intra-phase dynamism due to variation in the compute intensity of the following important significant regions
    PAR_BARRIER
    DEFLCG
```

READEX on Alya

Tuning Parameters

Level	Tuning Parameter	Range	Default
Hardware Parameters	CPU Core Frequency	1.2 GHz to 2.5 GHz; step 300 MHz	2.5 GHz
	CPU Uncore Frequency	1.2 GHz to 3.0 GHz; step 300 MHz	3.0 GHz
System Software Parameters	Number of Active OpenMP Threads	1 to *fill node*; step 2 threads	*Fill nodes*
Application Parameters	RENUMBERING (Ordering used for local unknowns defined in vertices)	METIS, CUTHILL_MC_KEE	METIS
	GROUPS (# degrees of freedom of coarse system in Deflated CG)	100, 200, 400, 800, 1600, 3200	800
	COARSE_SOLVER (Direct solver used for coarse system in Deflated CG)	CHOLESKY, SPARSE	CHOLESKY
	ELEMENT_CHUNK (Elements chunk size used in OpenMP dynamic scheduling)	100, 200, 400, 600, 800, 1000	400
	SOLVER_CHUNK (Chunk size used in the OpenMP dynamic scheduling for solver)	100, 200, 400, 600, 800, 1000	400

READEX on Alya

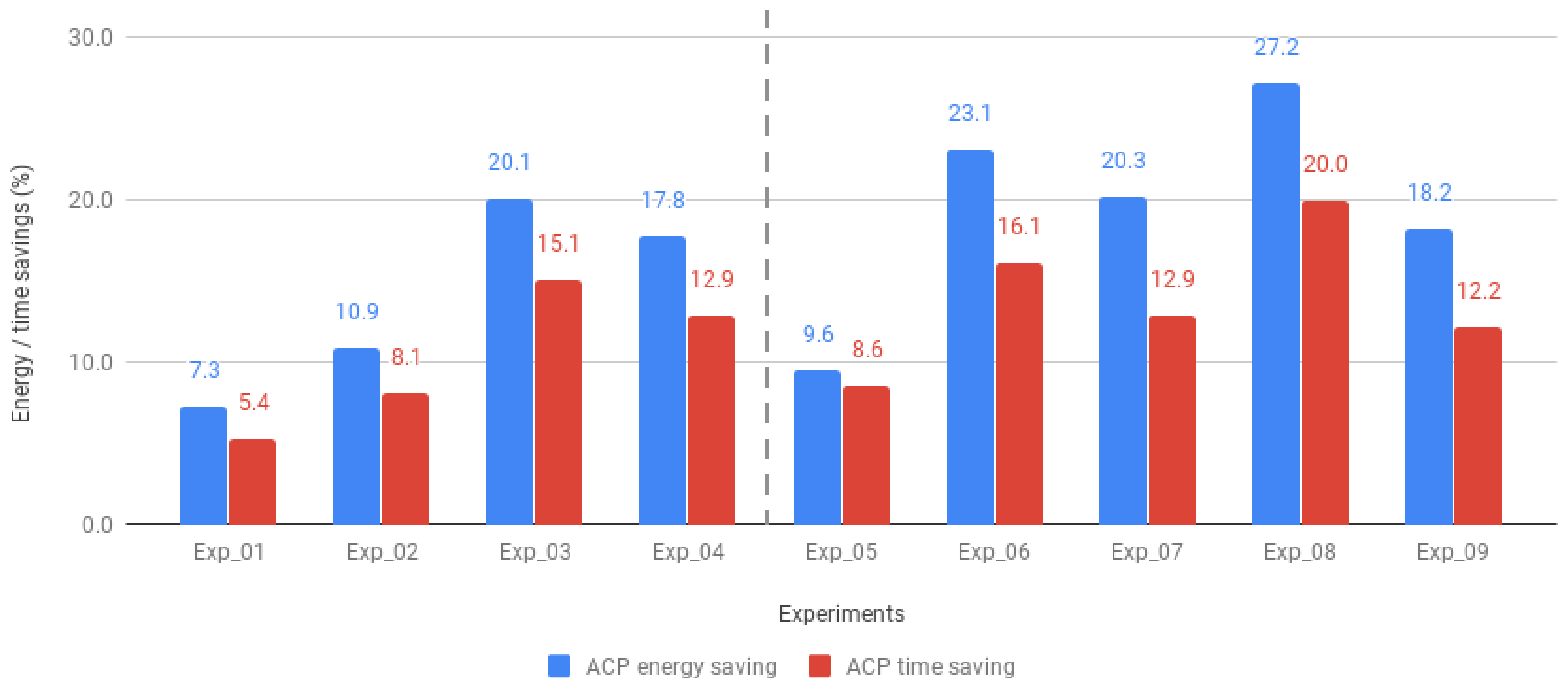
Experiments

Taurus cluster at TU Dresden | Intel Xeon CPU E5-2680 v3 processors | 24-core dual-socket nodes

Experiment label	# of nodes	# of MPI processes per node	# of OpenMP threads per process	Tuning model
Exp_01	10	2	12	Exp_01
Exp_02	10	4	6	Exp_02
Exp_03	20	2	12	Exp_03
Exp_04	20	4	6	Exp_04
Exp_05	10	2	12	Exp_02
Exp_06	20	2	12	Exp_02
Exp_07	20	4	6	Exp_02
Exp_08	40	2	12	Exp_02
Exp_09	40	4	6	Exp_02

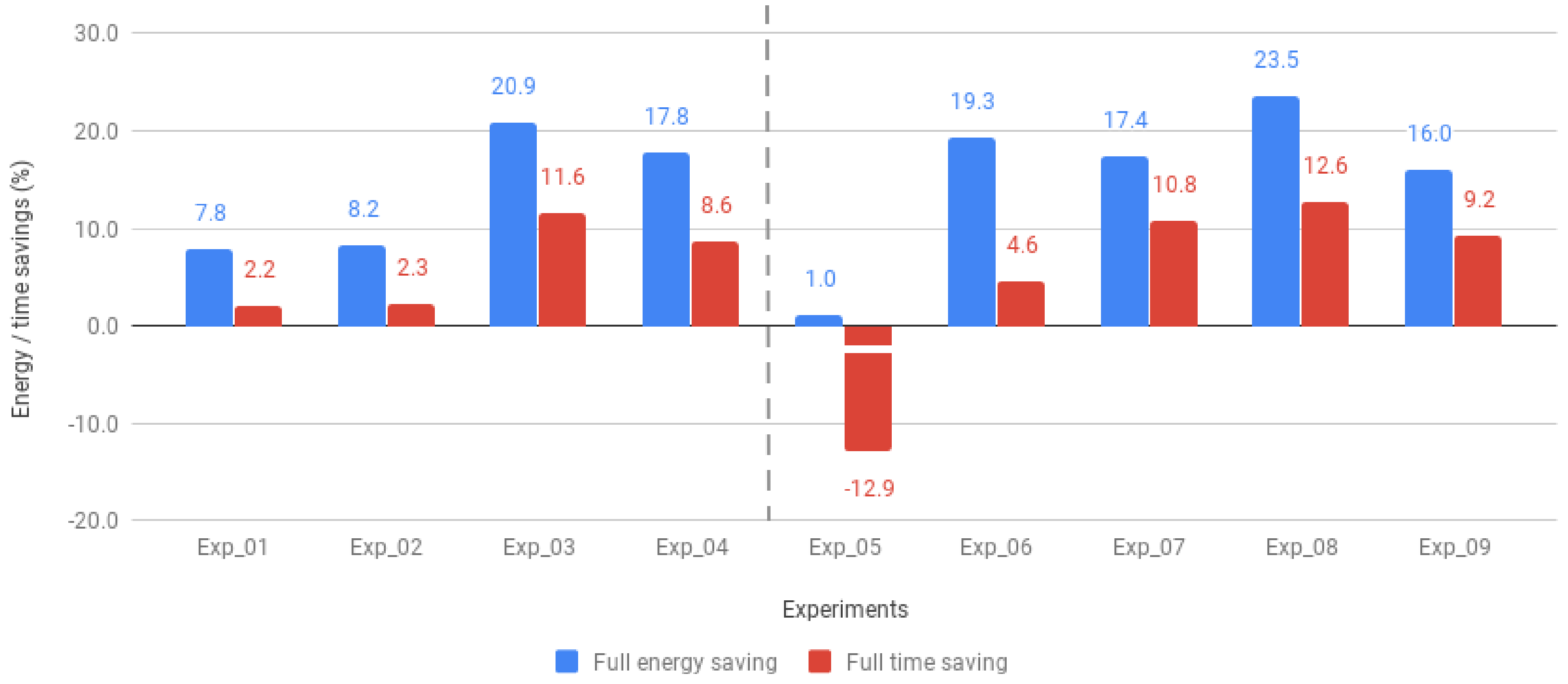
READEX on Alya

Results (tuning application parameters)



READEX on Alya

Results (tuning hardware, system-software & application parameters)



Summary & Credits

- Current evaluations show that energy consumption and execution time savings are achievable
  - For Alya, between 5-25% on up to 40 node (960 core) runs
  - Application dependent; exploitable dynamism
  - Indicative of a promising line of action to better understand the reasons for the dynamisms in Alya
- Further investigation into source of dynamism and potential optimisations on extreme scale clusters

PRACE

Myles Doyle
ICHEC

Ricard Borrell
BSC

Guillaume Houzeaux
BSC