

Design-Time Analysis to improve Energy-efficiency of HPC Applications

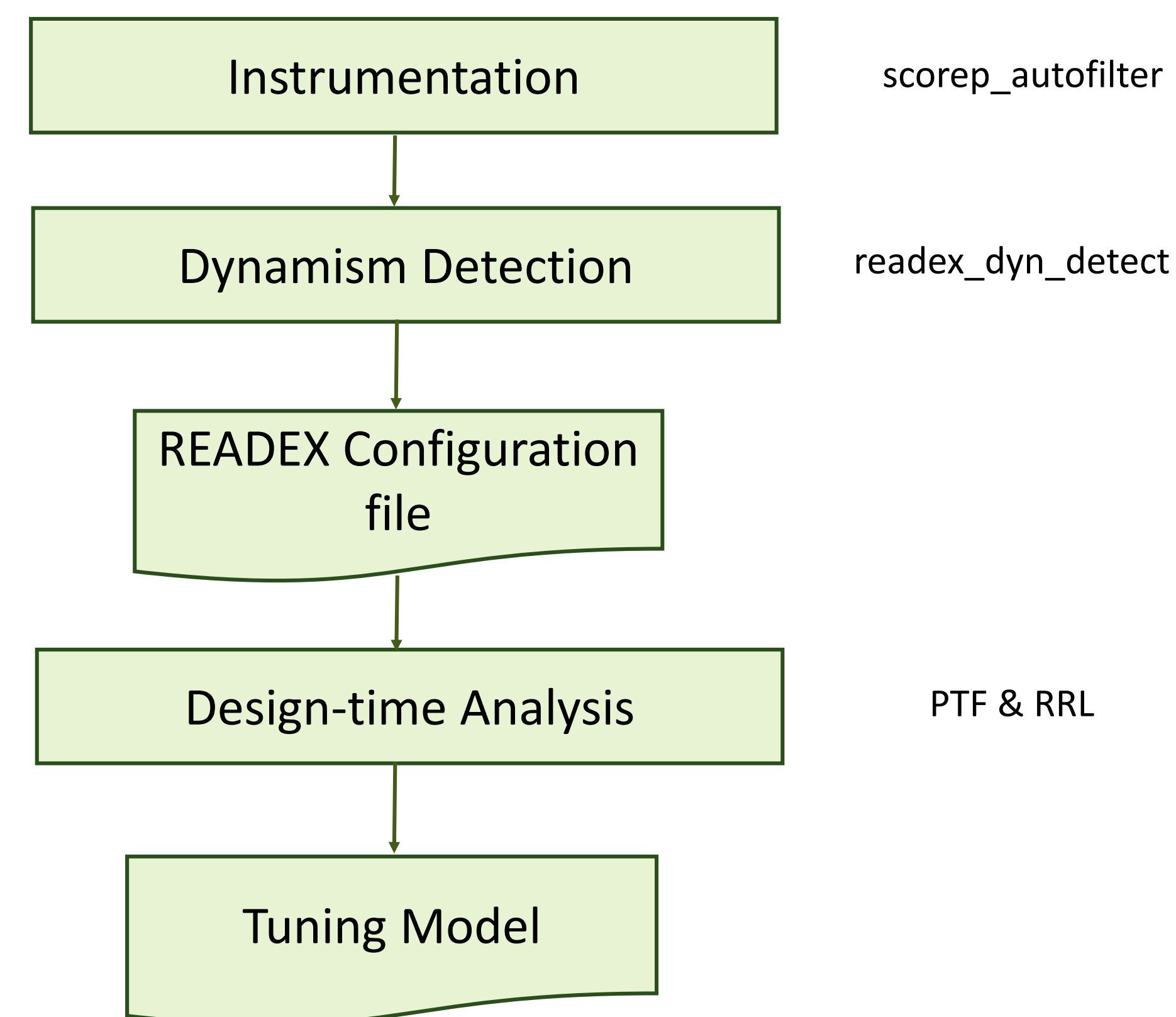
Anamika Chowdhury; Madhura Kumaraswamy; Michael Gerndt
Technical University of Munich (TUM)



Abstract

- Tune HPC applications dynamically for improved energy-efficiency and performance.
- Switching between configurations by exploiting dynamic characteristics of HPC applications.
- Develop tool aided auto-tuning methodology.
 - Design-time analysis
 - Runtime Application Tuning
- Detect during design-time, exploit during runtime.

Design-Time Analysis Workflow



Design-Time Analysis (DTA)

- Performed by the **Periscope Tuning Framework (PTF)**
 - Tunes performance and energy.
 - Evaluates alternatives online.
 - Supports different tuning strategies.
- **The READEX Tuning Plugin**
 - Multiple objectives.
 - Configurable search space via READEX configuration file.
 - Multiple search strategies for searching.
 - Tuning Parameters:
 - core frequency, uncore frequency, # of threads
 - Experiments for selected configurations
 - Energy and time measured for all RTS's.
 - Identification of static best for phase and specific best configurations for RTS's.

Discussion

- The static configuration is applied for the entire phase.
- The worst configuration
 - worst energy consumption for the entire phase.
- The best configuration
 - best energy consumption for the entire phase.
- The best setting for the phase
 - not necessarily for all the significant regions.
- Static energy savings w. r. t. worst energy consumption
 - 43.60% for all the significant regions.
 - 24.60% for the phase
- Dynamic energy saving due to switching configuration dynamically w. r. t. best energy consumption.
 - 4.40% for all the significant regions.

THE READEX Tool-Suite

Design-Time Analysis

- Detect program regions having variations in characteristics.
- Determine different Runtime Situations (RTS) of the detected regions.
- Determine best configurations for RTSs.
- Classify RTS's based on similar configurations into scenarios.
- Encapsulate the scenario information into a tuning model.

Runtime Application Tuning

- Propagate the generated tuning model for the production run.
- Performed by the READEX Runtime Library
 - Lightweight.
 - Switch to the best configuration for a detected RTS retrieved from the tuning model.
- Calibration mechanism
 - Calibrates regions which were not seen during design-time.

Initial DTA Results

Table 1. Results of DTA for the BT-MZ benchmark using the READEX Tuning Plugin

- Dynamism Detection:
 - performed by **readex-dyn-detect**.
 - detected significant regions: `exch_qbc`, `x_solve`, `y_solve`, and `z_solve`.
- Performed by the READEX Tuning Plugin
 - 16 experiments
 - The exhaustive search strategy
 - Returns best configurations for the no. of threads and core frequency

Significant regions	Energy for worst configuration (1, 1.6)	Energy for best configuration (4, 1.6)	The READEX Energy	# of thread	Core frequency (GHz)
exch_qbc	3245	6649	2760	1	2.4
x_solve	74219	41341	39962	4	2.0
y_solve	73536	39497	39497	4	1.6
z_solve	76393	40699	40386	4	2.0
SUM	227393	128186	122605		
Energy for phase	376722	284223			

Conclusions

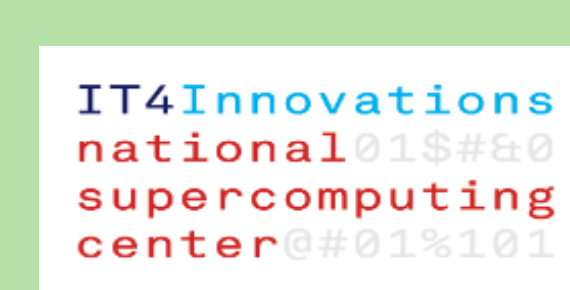
- Presents the Design-Time Analysis step of the READEX methodology for tuning to improve energy efficiency.
- READEX focuses on runtime tuning guided by a tuning model pre-computed during design-time.

Future Goals

- Inter-phase dynamism.
- Handling multiple input files.
- Domain knowledge specification.
 - Allows the user to provide domain knowledge as *identifiers*.
 - Application Tuning Parameters.
 - Input identifiers.

Contact Information

Robert Schoene
Technische Universitaet Dresden (TUD)
Email: Robert.schoene@tu-dresden.de
Phone: +49 (351) 463-42483



Collaborations



Acknowledgements

- Funded by the European Union's Horizon 2020
- Grant agreement no. 671657

More Info

- www.readex.eu
- www.researchgate.net/project/READEX