High Performance Computing Applications Dynamism Evaluation for Energy Tuning

O. Vysocký[1], M. Beseda[1], L. Říha[1], J. Zapletal[1], V. Nikl[2], M. Lysaght[3] and V. Kannan[3]

1 IT4Innovations, VŠB – Technical University of Ostrava, Ostrava, Czech Republic

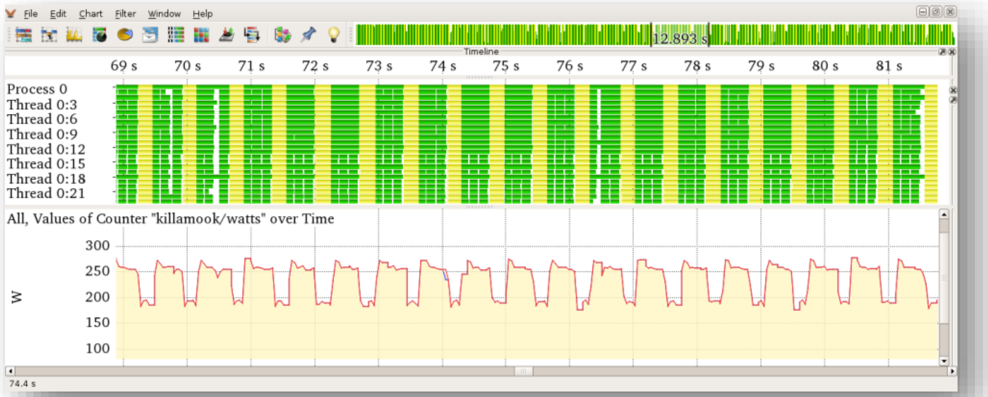
2 Faculty of Information Technology, Brno University of Technology, Czech Republic

3 Irish Centre for High End Computing, Ireland

25.6. - 30.6. 2017, IHPCSS

Applications exhibit dynamic behaviour

- Changing resource requirements
- Computational characteristics
- Changing load on processors over time



READEX creates a **tools-aided methodology for automatic tuning** of parallel applications

= dynamically adjust system parameters to actual resource requirements

READEX
Runtime Exploitation of Application Dynamism for Energy-efficient eXascale computing

European Commission | Horizon 2020 European Union funding for Research & Innovation

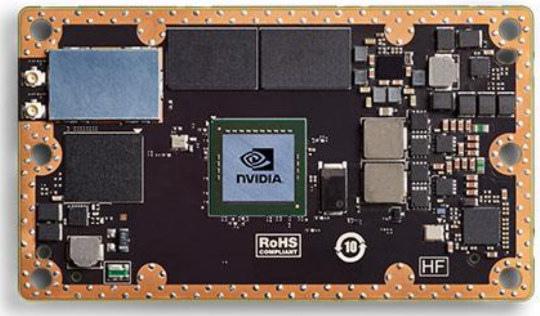
Parameters tuning

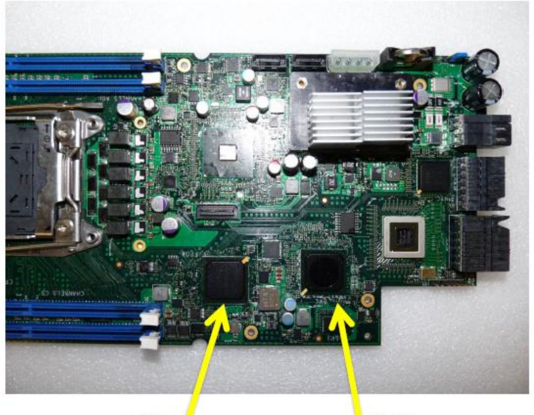
- Hardware parameters - CPU core frequency, uncore frequency
- System software parameters - number of OpenMP threads, thread placement
- Application-level parameters - depends on the specific application

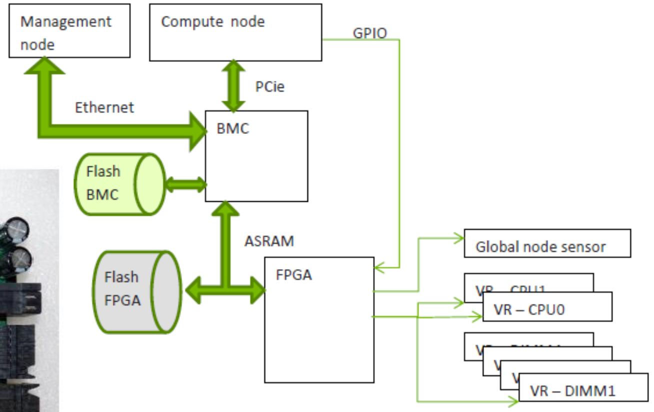
- Static tuning, Inter-phase dynamic tuning, Intra-phase dynamic tuning

Energy measurement

- Running Average Power Limit (RAPL) interface
- High Definition Energy Efficiency Monitoring (HDEEM) system
- ARM Jetson TX1 energy measurement interface



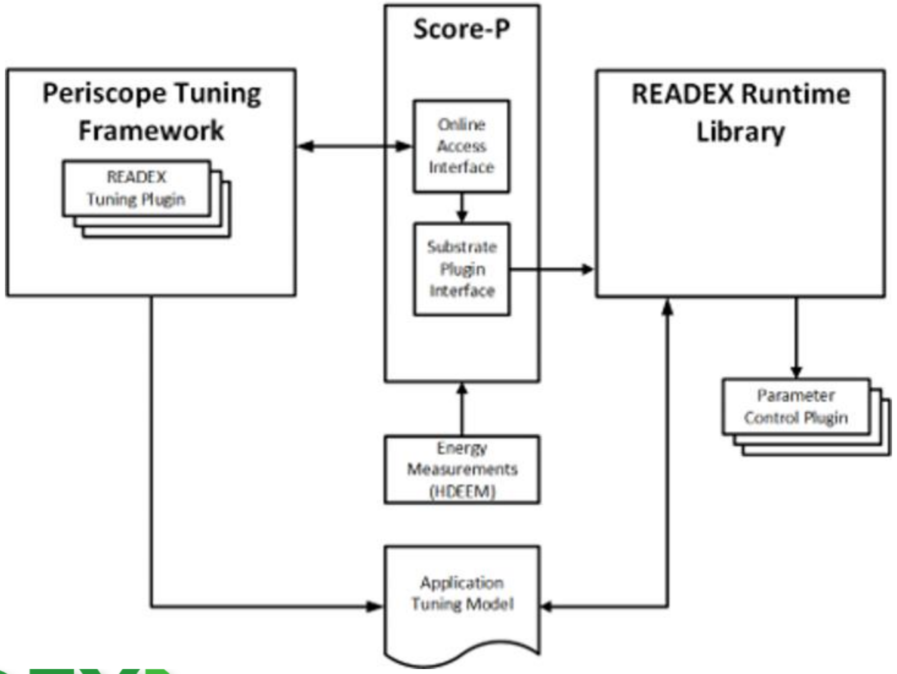




Tuning tools

- Score-P
- Periscope Tuning Framework
- READEX Runtime Library

- MERIC library
- RADAR generator



MERIC instrumentation:

```
MPI_Init(&argc, &argv);
MERIC_Init();
…
        MERIC_MeasureStart("RegionA");
        …
        MERIC_MeasureStop();
…
MERIC_Close();
MPI_Finalize();
```

1. Identification of significant regions
2. Insertion of MERIC probes
3. Compilation of MERIC and a user code
4. Setting MERIC parameters
5. Running complete energy measurement
6. Processing the results
7. Dynamic tuning

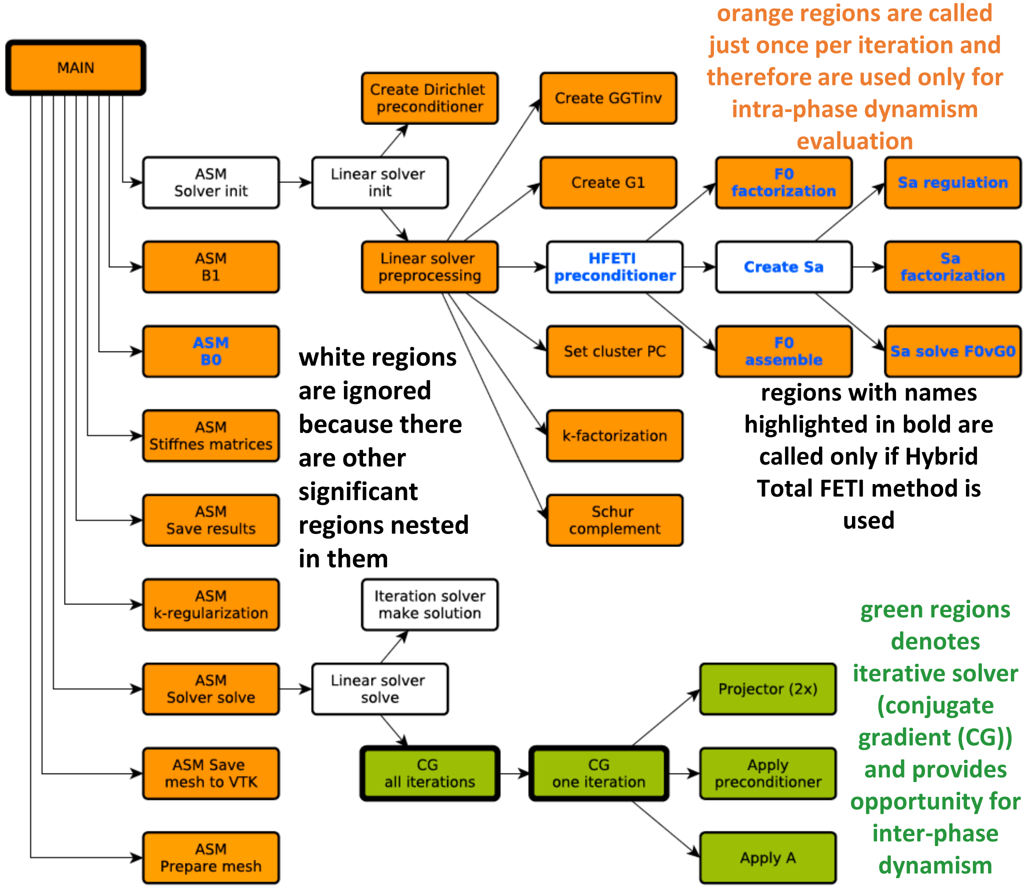
Example results - ESPRESO

static dynamic total

ESPRESO: 12.3% + 9.1% = 20.3%

- Structural mechanics code
- Finite element + sparse FETI solver

Region	% of 1 phase	Best static configuration	Value	Best dynamic configuration	Value	Dynamic savings
Assembler–AssembleStiffMat	14.32	18 threads, 1.8 GHz UCF, 2.5 GHz CF	733.73 J	20 threads, 2.0 GHz UCF, 2.5 GHz CF	731.22 J	2.51 J (0.34%)
Assembler–Assemble-B1	2.23	18 threads, 1.8 GHz UCF, 2.5 GHz CF	114.30 J	2 threads, 2.2 GHz UCF, 2.5 GHz CF	94.15 J	20.15 J (17.63%)
Cluster–CreateF0-FactF0	0.17	18 threads, 1.8 GHz UCF, 2.5 GHz CF	8.71 J	6 threads, 1.6 GHz UCF, 2.5 GHz CF	6.90 J	1.80 J (20.73%)
Assembler–SaveResults	3.10	18 threads, 1.8 GHz UCF, 2.5 GHz CF	158.81 J	2 threads, 1.2 GHz UCF, 2.5 GHz CF	147.66 J	11.16 J (7.03%)
			...			
Cluster–CreateSa-SaReg	0.17	18 threads, 1.8 GHz UCF, 2.5 GHz CF	8.59 J	8 threads, 2.0 GHz UCF, 2.5 GHz CF	7.03 J	1.56 J (18.15%)
Total value for static tuning for significant regions			733.73 + 114.30 + 8.71 + 158.81 + 278.39 + 113.87 + 14.23 + 658.07 + 325.69 + 99.93 + 74.70 + 641.88 + 1578.06 + 13.28 + 24.20 + 278.22 + 8.59 = 5124.66 J			
Total savings for dynamic tuning for significant regions			2.51 + 20.15 + 1.80 + 11.16 + 47.01 + 16.41 + 5.31 + 28.45 + 29.03 + 19.08 + 0.16 + 2.49 + 288.21 + 0.77 + 1.88 + 23.24 + 1.56 = 499.22 J of 5124.66 J (9.74 %)			
Dynamic savings for application runtime			499.22 J of 5493.55 J (9.09 %)			
Total value after savings			4994.33 J (79.72 % of 6265.18 J)			



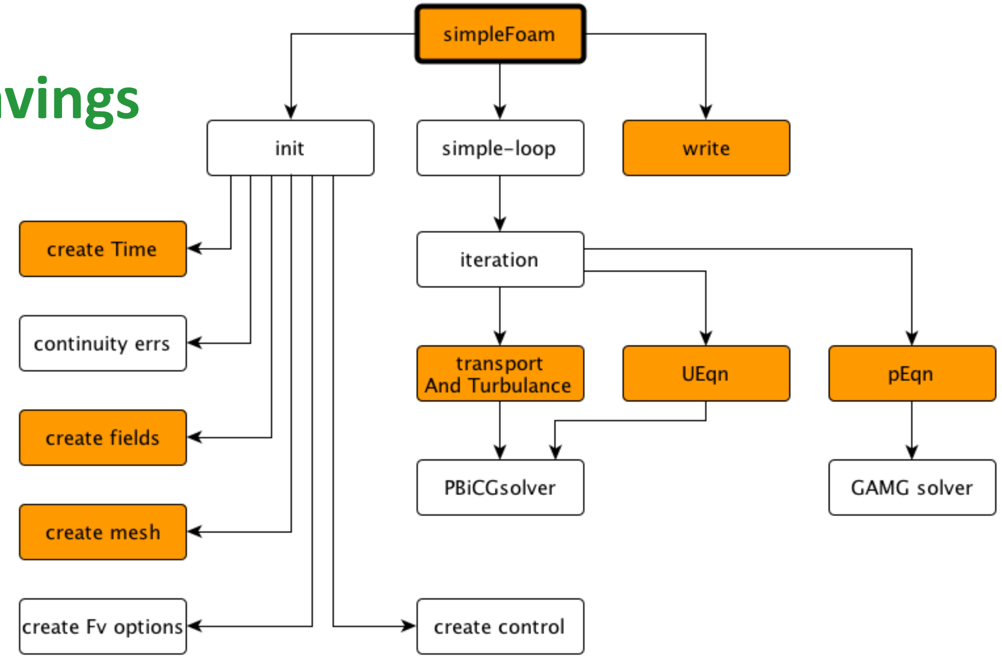
	Default settings	Default values	Best static configuration	Static Savings	Dynamic Savings
Energy consumption [J], Blade summary	24 threads, 3.0 GHz UCF, 2.5 GHz CF	6265.18 J	18 threads, 1.8 GHz UCF, 2.5 GHz CF	771.63 J (12.32%)	499.2 J of 5493.6 J (9.09 %)
Runtime of function [s], Job info - hdeem	24 threads, 3.0 GHz UCF, 2.5 GHz CF	29.55 s	22 threads, 3.0 GHz UCF, 2.5 GHz CF	0.01 s (0.04%)	0.82 s of 29.54 s (2.76 %)

Example results - OpenFOAM

static dynamic total

OpenFOAM: 15.9% + 1.8% = 17.4% energy savings

- Computational fluid dynamics
- Finite volume + multigrid solver



Region	% of 1 phase	Best static configuration	Value	Best dynamic configuration	Value	Dynamic savings
init-createTime	0.03	2.0 GHz UCF, 1.6 GHz CF	3.35 J	1.4 GHz UCF, 1.4 GHz CF	2.64 J	0.71 J (21.06%)
init-createFields	4.28	2.0 GHz UCF, 1.6 GHz CF	506.91 J	2.4 GHz UCF, 2.0 GHz CF	474.80 J	32.11 J (6.33%)
init-createMesh	2.26	2.0 GHz UCF, 1.6 GHz CF	267.33 J	1.4 GHz UCF, 1.4 GHz CF	194.38 J	72.96 J (27.29%)
UEqn	40.71	2.0 GHz UCF, 1.6 GHz CF	4820.82 J	2.2 GHz UCF, 1.6 GHz CF	4810.03 J	10.80 J (0.22%)
pEqn	19.15	2.0 GHz UCF, 1.6 GHz CF	2268.19 J	2.0 GHz UCF, 1.6 GHz CF	2268.19 J	0.00 J (0.00%)
trans-portAnd-Turbulence	25.70	2.0 GHz UCF, 1.6 GHz CF	3042.91 J	2.0 GHz UCF, 1.6 GHz CF	3042.91 J	0.00 J (0.00%)
write	7.88	2.0 GHz UCF, 1.6 GHz CF	932.59 J	1.2 GHz UCF, 1.4 GHz CF	841.62 J	90.97 J (9.75%)
Total value for static tuning for significant regions			3.35 + 506.91 + 267.33 + 4820.82 + 2268.19 + 3042.91 + 932.59 = 11842.12 J			
Total savings for dynamic tuning for significant regions			0.71 + 32.11 + 72.96 + 10.80 + 0.00 + 0.00 + 90.97 = 207.54 J of 11842.12 J (1.75 %)			
Dynamic savings for application runtime			207.54 J of 11966.36 J (1.73 %)			
Total value after savings			11758.82 J (82.63 % of 14231.30 J)			

Uncore freq [GHz] / Core freq [GHz]	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
1.2	13,200.02	12,717.1	12,621.78	12,410.62	12,380.68	12,507.38	12,774.16	13,108.6	13,604.2	14,040.8
1.4	13,161.9	12,597.78	12,125.18	12,065.52	12,074.54	12,173.36	12,312.24	12,802.26	13,095.84	13,450.8
1.6	13,320.66	12,640.76	12,256.22	12,033.62	11,966.36	11,992.7	12,372.04	12,579.22	13,126.44	13,370.24
1.8	13,878.04	13,082.66	12,700.92	12,457.08	12,373.86	12,445.98	12,574.6	12,831.82	13,081.62	13,296.04
2	14,218.58	13,327.12	12,902.62	12,544.82	12,456.82	12,494.8	12,680.32	13,038.86	13,207.38	13,474.8
2.2	14,625.62	13,849.58	13,240.14	12,851	12,760.98	12,802.24	12,993.44	13,260.38	13,497.6	13,767.62
2.4	15,083.2	14,412.62	13,568.68	13,447.18	12,973.38	13,238.6	13,332.7	13,388.7	13,777.68	14,030.66
2.5	15,554.96	14,465.2	13,991	13,553.84	13,300.24	13,354.46	13,472.36	14,179.16	14,083.06	14,231.3

	Default settings	Default values	Best static configuration	Static Savings	Dynamic Savings
Energy consumption [J], Blade summary	3.0 GHz UCF, 2.5 GHz CF	14231.30 J	2.0 GHz UCF, 1.6 GHz CF	2264.94 J (15.92%)	207.54 J of 11966.36 J (1.73 %)
Runtime of function [s], Job info - hdeem	3.0 GHz UCF, 2.5 GHz CF	56.45 s	2.6 GHz UCF, 2.4 GHz CF	0.37 s (0.66%)	2.36 s of 56.08 s (4.20 %)

READEX
Runtime Exploitation of Application Dynamism for Energy-efficient eXascale computing

European Commission

Horizon 2020
European Union funding
for Research & Innovation

Other results

- Evaluation of HPC codes ranging from basic kernels to very complex applications
- **Key results**
    - Highly optimized applications tend to provide higher static and lower dynamic savings
    - Complex applications, such as ESPRESO, which contains variation on workload (not only compute) shows opportunity for dynamic tuning

Application	Static savings [%]	Dynam. savings [%]	Total Savings [%]
Parallel OpenMP I/O	56	—	56
Dense BLAS - DGEMV - without NUMA	5.6	—	5.6
Dense BLAS - DGEMM - without NUMA	10.4	—	10.4
Compute only kernel	12.8	—	12.8
Sparse BLAS Routines - without NUMA	3.1-12.3	—	3.1 – 12.3
Sparse BLAS Routines - with NUMA	4.2-66.2	—	4.2 – 66.2
ProxyApps 1 - AMG2013, configuration 1	6.53	2.89	9.23
ProxyApps 1 - AMG2013, configuration 2	25.66	2.80	27.74
ProxyApps 2 - Kripke, configuration 1	28.16	1.56	29.28
ProxyApps 2 - Kripke, configuration 2	12.63	7.04	18.78
ProxyApps 3 - LULESH, configuration 1	28.58	0.55	28.88
ProxyApps 3 - LULESH, configuration 2	25.81	1.23	26.72
ProxyApps 4 - MCB, configuration 1	4.13	1.42	5.51
ProxyApps 4 - MCB, configuration 2	3.40	4.18	7.44
ESPRESO - configuration 0	5.6	8.7	14.3
ESPRESO - configuration 1	12.3	9.1	21.4
ESPRESO - configuration 2	7.8	4.7	12.5
ESPRESO - configuration 3	7.8	5.4	13.1
OpenFOAM (Motorbike benchmark)	15.9	1.8	17.7
Indeed	17.6	to be evaluated	17.6
MiniMD	21.92	0.00	21.92

The research leading to these results has received funding from the European Union's Horizon 2020 Programme under grant agreement number 671657.

This work was supported by VSB-Technical University of Ostrava under the grants SP2017/165.