



SCALABLE PARALLEL ASTROPHYSICAL CODES FOR EXASCALE

Second report on Evaluation of Energy Efficiency

Deliverable number: D2.4

Version 1.0



**Co-funded by
the European Union**



EuroHPC
Joint Undertaking

Funded by the European Union. This work has received funding from the European High Performance Computing Joint Undertaking (JU) and Belgium, Czech Republic, France, Germany, Greece, Italy, Norway, and Spain under grant agreement No 101093441

Project Information

Project Acronym: SPACE
Project Full Title: Scalable Parallel Astrophysical Codes for Exascale
Call: Horizon-EuroHPC-JU-2021-COE-01
Grant Number: 101093441
Project URL: <https://space-coe.eu>

Document Information

Editor:	Lubomir Riha, Ondrej Vysocky – IT4I@VSB
Deliverable nature:	Report (R)
Dissemination level:	Public (PU)
Contractual Delivery Date:	31.12.2024
Actual Delivery Date	27.02.2025
Number of pages:	48
Keywords:	Energy efficiency, DVFS, MERIC, RAPL
Authors:	Pluto: Tomas Panoc – IT4I@VSB, Marco Rossazza – UNITO, OpenGADGET: Milan Jaros – IT4I@VSB, Luca Tornatore – INAF, RAMSES: Matej Spetko – IT4I@VSB, Benoît Commerçon – CNRS, iPIC3D: Tomas Panoc – IT4I@VSB, Pranab J Deka – KU Leuven, FIL: Joao Barbosa – IT4I@VSB, Khalil Pierre – GUF, BHAC: Joao Barbosa – IT4I@VSB, Georgios Doulis – GUF, ChaNGa: Ondrej Vysocky – IT4I@VSB, Robert Wissing – UiO
Peer review:	Fabrizio Magugliani – E4 Manolis Marazakis – FORTH

History of Changes

Release	Date	Author, Organization	Description of changes
0.1	18.10.2025	Gino Perna – ENGIN-SOFT, Ondrej Vysocky – IT4I@VSB	Started the document
0.7.0	7.02.2025	Lubomir Riha, Ondrej Vysocky, Tomas Panoc, Matej Spetko, Joao Barbosa, Milan Jaros – IT4I@VSB	Measurements results added and the text completed
0.8	15.2.2025	Manolis Marazakis – FORTH, Fabrizio Magugliani – E4 Computer Engineering	Internal review
0.9	21.2.2025	Lubomir Riha, Ondrej Vysocky – IT4I@VSB	new measurements for BHAC and ChaNGa
1.0	26.2.2025	Lubomir Riha – IT4I@VSB	Rewiever's comments incorporated into the document.

Scalable Parallel Astrophysical Codes for Exascale

DISCLAIMER

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European High Performance Computing Joint Undertaking (JU) and Belgium, Czech Republic, France, Germany, Greece, Italy, Norway, and Spain. Neither the European Union nor the granting authority can be held responsible for them.



The space above and below the message intentionally is left blank.

Executive Summary

This deliverable reports and evaluates the activities implemented to enhance the energy efficiency of all SPACE applications. Energy efficiency is quantified as the ratio of performance (measured in floating-point operations per second, or FLOPs) to power consumption (measured in Watts) for each application. FLOPs are determined using platform-specific performance counters and are measured throughout the entire application execution.

We examined measured energy consumption and efficiency by modifying specific hardware power settings. Our method involved static frequency tuning, where a single hardware configuration (CPU/GPU frequency) was set at the beginning of each application run and remained unchanged. We monitored and assessed execution time, energy consumption, and FLOPs.

This deliverable builds on the previous report (D2.2) on the energy efficiency of SPACE codes. It evaluates the energy efficiency of three new hardware platforms: (i) NVidia A100 GPU installed in the IT4Innovations Karolina supercomputer, (ii) Intel Sapphire Rapids processor with DDR and HBM memory, and (iii) NVidia’s ARM-based GRACE CPU. These platforms were chosen based on their relevance to SPACE CoE’s work, with a focus on GPU accelerators and co-design platforms. The results compare the performance per watt for CPU platforms, showing that the NVidia GRACE CPU is the most efficient CPU for all codes, see Table 1¹

Code energy efficiency	NVidia GRACE CPU [MFLOPS/W]	Intel Sapphire Rapids with DDR [MFLOPS/W]	Intel Sapphire Rapids with HBM [MFLOPS/W]
Pluto	805.4	264.0	309.0
OpenGADGET	716.2	138.2	149.2
iPIC3D	791.3	238.2	321.4
RAMSES	854.9	399.7	417.8
BHAC	292.9	121.9	125.5
FIL	522.2	223.9	248.9
ChaNGa	1478.9	779.8	1018.3

Table 1: The energy efficiency of the CPU platforms. Results are shown for hardware configurations generating very small runtime extensions (less than extra 3% of default runtime).

Table 2 shows the potential energy savings coming from adjusting hardware parameters. For the NVidia A100 GPU, savings reach up to 9% for Pluto with a 13% runtime extension, and 6% with minimal performance impact. OpenGADGET saves 7%, and iPIC3D between 3% and 5%. In general, the Nvidia GPUs do not achieve as large of a savings as we can see for CPUs if the GPU streaming multiprocessor frequency is tuned. This means that their internal power management systems perform more efficiently. We also observe that in all cases, the GPUs consume significantly lower power than its 400W TDP: Pluto up to 185W, OpenGADGET up to 74W, iPIC3D up to 159W, and ChaNGa up to 95W.

On the Sapphire Rapids platform, savings range from 6% to 14% without runtime extension, and 7% to 20% with runtime extension using DDR memory. HBM memory results in faster execution and higher performance per watt but requires higher CPU frequencies, leading to lower savings for some codes (Pluto, iPIC3D, BHAC,

¹Our tools do not have implemented measurements of FLOPS for NVidia GPUs, therefore these are not presented.

	Pluto	Open GADGET	iPIC3D	RAMSES	BHAC	FIL	ChaNGa
Nvidia A100	-6% / 103% -9% / 113%	-7% / 102% -7% / 102%	-3% / 104% -5% / 111%	- - - -	- - - -	- - - -	- - - -
SPR w. DDR	-9% / 102% -10% / 106%	-7% / 102% -7% / 102%	-7% / 102% -9% / 108%	-6% / 102% -7% / 104%	-10% / 103% -14% / 110%	-6% / 102% -7% / 103%	-14% / 103% -20% / 107%
SPR w. HBM	-4% / 101% -6% / 105%	-9% / 94% -11% / 98%	-7% / 101% -7% / 101%	-7% / 102% -8% / 104%	-4% / 99% -4% / 99%	-4% / 100% -6% / 104%	-12% / 102% -13% / 103%
Grace CPU	-22% / 101% -35% / 122%	-13% / 103% -33% / 128%	-9% / 103% -29% / 126%	-19% / 101% -28% / 137%	-26% / 103% -36% / 109%	-8% / 102% -20% / 117%	-16% / 102% -30% / 110%
Cascade Lake D2.2*	-6% / 102% -12% / 126%	-9% / 103% -12% / 130%	-6% / 101% -13% / 115%	-7% / 102% -11% / 123%	-5% / 102% -11% / 118%	-3% / 102% -13% / 127%	30% / 102% 36% / 110%

Table 2: **Energy consumption reduction [kJ] / Runtime [s]**; 1st row - results with no runtime extension (runtime close to 100% of default); 2nd row - results for maximum energy savings. *Cascade Lake - results from deliverable D2.2 are presented for comparison.

FIL and ChaNGa). RAMSES savings are similar and OpenGADGET sees up to 11% savings. The NVidia GRACE ARM-based processor achieves the highest savings, around 30% for most codes with runtime extension, and around 20% for Pluto, RAMSES and BHAC with less than 3% runtime extension.

Overall, the ARM-based GRACE platform offers superior energy efficiency compared to Sapphire Rapids. Tuning CPU core frequency can further improve its efficiency by up to 30%, or between 8% and 26% without affecting runtime.

The results presented in this deliverable are not usable only for the A&C community but, in general, for other parallel codes running on supercomputers around the world. If one considers that a machine consumes 20 MW, energy savings of 5-10 % are in orders in megawatts, which significantly reduces operational costs.

The energy efficiency analysis of the parallel codes is done by the open-source MERIC tool developed at IT4I@VSB. Similar services, as presented in this deliverable, are also offered under the POP3 Center of Excellence to the general public. We want to highlight that for the static tuning approach, there is no need to change the code as the optimal settings, and it can be applied job-wide using a job scheduler.

Contents

1	Introduction	8
1.1	Revising the Deliverable 2.2	8
1.2	New hardware platforms	9
2	Energy efficiency analysis methodology	10
2.1	Use-cases	11
2.2	Report component description	12
3	Results	15
3.1	Pluto	15
3.2	OpenGadget	19
3.3	iPIC3D	23
3.4	RAMSES	27
3.5	BHAC	31
3.6	FIL	35
3.7	ChaNGa	39
4	Conclusion	44
4.1	Recommendations outside of the scope of the project	45
4.2	Future work	45

List of Figures

1	Heatmap example #1: Impact of the static tuning to run time in the Sapphire Rapids system . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run. (results for Pluto code)	12
2	Heatmap example #2: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	13
3	Example of an energy consumption and savings table for the Nvidia A100 system (Karolina GPU)	14
4	Example of an energy consumption and savings table for the Grace CPU system	14
5	Pluto: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).	15
6	Pluto: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	16
7	Pluto: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	16
8	Pluto: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	17
9	Pluto: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	17
10	Pluto: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	18
11	OpenGadget: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).	19
12	OpenGadget: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	20
13	OpenGadget: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	20
14	OpenGadget: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	21
15	OpenGadget: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	21
16	OpenGadget: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	22
17	iPIC3D: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).	23
18	iPIC3D: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	24
19	iPIC3D: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	24
20	iPIC3D: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	25
21	iPIC3D: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	25

22	iPIC3D: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	26
23	Ramses: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	28
24	Ramses: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	28
25	Ramses: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	29
26	Ramses: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	29
27	RAMSES: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	30
28	BHAC: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	32
29	BHAC: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	32
30	BHAC: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	33
31	BHAC: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	33
32	BHAC: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	34
33	FIL: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	36
34	FIL: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	36
35	FIL: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	37
36	FIL: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	37
37	FIL: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	38
38	ChaNGa: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).	40
39	ChaNGa: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	41
40	ChaNGa: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	41
41	ChaNGa: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.	42

42	ChaNGa: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory . The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.	42
43	ChaNGa: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.	43

List of Tables

1	The energy efficiency of the CPU platforms. Results are shown for hardware configurations generating very small runtime extensions (less than extra 3% of default runtime).	1
2	Energy consumption reduction [kJ] / Runtime [s] ; 1st row - results with no runtime extension (runtime close to 100% of default); 2nd row - results for maximum energy savings. *Cascade Lake - results from deliverable D2.2 are presented for comparison.	1
3	Summary of the results of deliverable D2.2. Energy savings achieved on the Barbora cluster under different maximum allowed runtime extension. The results are for static tuning of the CPU core and uncore frequency.	8
4	Hardware architectures, available monitoring systems, and tuned power knobs	10
5	Node hardware details for individual architectures. (note: CS stands for IT4I's Complementary Systems partition).	10
6	Summary of experiments to be presented in this deliverable	11
7	An example of a summary table for a particular SPACE application	14
8	Pluto: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	18
9	OpenGadget: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	22
10	iPIC3D: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	26
11	Ramses: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	27
12	BHAC: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	31
13	FIL: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	35
14	ChaNGa: Summary of the energy consumption and static frequency tuning measurements on all HW platforms	39
15	The energy efficiency of the CPU platforms. Results are shown for hardware configurations generating very small runtime extensions (less than 105% of default).	44
16	The energy efficiency of the CPU platforms. Results are shown for hardware configurations generating maximum energy savings and noticeable runtime extension.	44
17	Energy consumption reduction [kJ] / Runtime [s] ; 1st row - results with no runtime extension (runtime close to 100% of default); 2nd row - results for maximum energy savings. *Cascade Lake - results from deliverable D2.2 are presented for comparison.	45

List of Acronyms

CAPEX	Capital Expenditure (for supercomputing cluster)
CF	Frequency of CPU cores
CoE	Centre of Excellence
DDR	Double Data Rate
FLOP	Floating-point arithmetic operation
HBM	High Bandwidth Memory
MHD	MagnetoHydroDynamics
OPEX	Operating Expenditure (for supercomputing cluster)
RAPL	Running Average Power Limit
SPACE	Scalable Parallel Astrophysical Codes for Exascale
TDP	Thermal Design Power
UCF	Frequency of CPU uncore subsystem

1 Introduction

This deliverable reports and evaluates the activities implemented to enhance the energy efficiency of all SPACE applications. Energy efficiency is measured as the ratio of performance (floating-point operations per second – FLOPs) to power consumption (Watts) of the given application. FLOPs are measured using performance counters that are platform-dependent and measured from the complete application execution. We investigated the energy consumption and efficiency by adjusting specific hardware power settings. Our approach used static frequency tuning, meaning that a single hardware configuration (CPU/GPU frequency) was applied at the start of each application run and remained constant throughout. Execution time, energy consumption, and FLOPs were monitored and evaluated.

1.1 Revising the Deliverable 2.2

In D2.2 we have performed an initial analysis of the SPACE CoE applications on the Barbora cluster CPU partition at IT4Innovations. Barbora’s compute nodes were equipped with 2x Intel Xeon Gold 6240 (18-core, 2.6 GHz nominal frequency, Cascade Lake microarchitecture).

The system was selected for the energy efficiency analysis because computational nodes accommodate on board the Atos High Definition Energy Efficient Monitoring (HDEEM) system [1], which reads the power consumption from the system hardware sensors and stores the data to an integrated memory. The sensor that monitors the consumption of the entire node measures 1000 power consumption samples per second.

Also, the Intel CPUs allows to manage several knobs that influence its power consumption: 1.) CPU Core frequency, 2.) CPU uncore frequency 3.) the number of active cores and 4.) power limit. In D2.2 we used core and uncore frequency tuning (we use the same knobs in this deliverable for newer generation on Intel CPUs codenamed Sapphire Rapids).

The summary of the results of D2.2 is presented in Table 3. It shows three scenarios, each with a different run-time penalty allowed. The runtime penalty stands for runtime extension when the processor runs on lower core and uncore frequencies. The results show that it is possible to save between 3–6 % of energy without impacting the runtime of the application (max. runtime extension of 2 %). When runtime can be extended by 5 % the savings are in the range of 6.6–9.4 %. The best results were achieved for the ChaNGa code, which can achieve 30% energy savings with almost no impact on its performance.

We remind the reader of these results so that it is not needed to read the D2.2 again to compare the results of this deliverable with the previous results. In D2.2, only one CPU architecture was evaluated. In this deliverable, we cover 3 more CPU architectures (Intel Sapphire Rapids CPU with DDR memory, Intel Sapphire Rapids CPU with HBM memory, NVidia Grace ARM-based CPU, and 1 GPU architecture (NVidia A100, Ampere architecture). All are equally important from both the production and the co-design point of view.

We maintain the same methodology, static tuning. Therefore, current results can be directly compared to the previous ones.

It must be noted that energy efficiency savings achieved by tuning the same HW parameters is going to be different for different platforms, as the power and thermal control mechanisms on different generations of Intel CPUs are continuously evolving.

Code name	<i>Energy savings with maximum given runtime extension</i>			
	max. 2%	max. 5%	max. 10%	unlimited
Pluto	5.5 %	7.3 %	10.5 %	11.5 %
OpenGadget	5.4 %	6.8 %	7.4 %	11.8 %
iPic3D	5.7 %	9.4 %	11.0 %	12.9 %
RAMSES	7.4 %	10.1 %	11.9 %	13.4 %
BHAC	5.0 %	6.8 %	10.8 %	11.2 %
FIL	2.8 %	6.6 %	8.9 %	12.6 %
ChaNGa	30.2 %	35.8 %	36.0 %	36.0 %

Table 3: Summary of the results of deliverable D2.2. Energy savings achieved on the Barbora cluster under different maximum allowed runtime extension. The results are for static tuning of the CPU core and uncore frequency.

1.2 New hardware platforms

The SPACE codes have been executed on several hardware platforms that play vital role in the project, i.e., Nvidia A100, Intel Sapphire Rapids with HBM memory, Nvidia Grace CPU. In our measurements, a single fully occupied (all the CPU cores) compute node was used. These platforms allowed us to use one or more of the following power knobs to control the frequency:

- CF = CPU core frequency,
- U(n)CF = CPU uncore frequency,
- SMF = GPU streaming multiprocessors frequency.

Intel CPUs have a maximum achievable core frequency (CF) that varies depending on several factors. The number of active cores plays a role, as increasing core counts lower the peak frequency. The instruction set also influences the maximum CF, with vector instructions (SSE, AVX/AVX2, AVX-512) generally having lower limits than scalar instructions. Furthermore, a power capping mechanism, typically set to the Thermal Design Power (TDP), limits CPU/GPU power consumption by scaling frequencies, which can also reduce the achievable CF.

Our energy efficiency improvements focused on reducing power consumption by adjusting these hardware settings. We employed a static approach, configuring the hardware at the start of execution and maintaining it throughout the run. This method, while simple to implement, might be suboptimal for complex applications with varying resource demands. Dynamic hardware configuration adjustments during runtime could potentially yield greater energy savings without degrading performance. The dynamic hardware configuration adjustments during runtime will be explored in future phases of the project. This report presents two different static configurations for each SPACE code: (1) a configuration that aims to preserve application performance as much as possible (approx. 3% runtime extension allowed) but provides fewer energy savings, and (2) a setting that maximizes energy savings no matter how large the overhead (runtime extension) it produces.

Several power monitoring systems exist for our purposes, each with its own sampling frequency and monitored power domains. Some, like Intel RAPL, perf counters [2], and NVML[3], focus on compute units, while others, such as PDUs and IPMI, monitor the entire compute node. To ensure comparable results across all evaluated platforms, we report node-level energy consumption.

For systems that only monitor compute unit power, we established a "node power baseline" to account for the energy used by other on-node components. This baseline represents the power consumption of these unmonitored components. We determined this baseline using Intelligent Platform Management Interface (IPMI) data, which report complete node power consumption but due to its low sampling frequency (less than 1 Hz), it is unsuitable for the energy measurements of an application execution. To differentiate the power consumption of the compute and other on-node components reliably, we ran standardized workload for approximately one minute.

The power monitoring systems evaluate energy consumption based on power consumption samples that the power monitoring system took in time colored in time? during runtime? , working at a specific sampling frequency. RAPL samples at 1 kHz frequency. Consumed energy (in Joules) is calculated from power samples (in Watts), that are measured at a sampling frequency (in Hertz):

$$Energy(t) = \int_0^t Power(x) dx \approx \frac{\sum_{i=0}^n PowerSample_i}{SamplingFrequency} . \quad (1)$$

In energy consumption measurement and controlling power management knobs, we have relied on the *MERIC* runtime system [4, 5, 6] which has been developed by IT4Innovations colored add references for MERIC . Details about the target hardware platforms can be found in Tables 4 and 5.

Hardware	Power Monitoring	Node Baseline Power	Tuned Power Knob
Nvidia A100-SXM4 GPU - 400 W TDP	NVML	600 W	SMF (1.41 - 0.21 GHz)
Intel Xeon Max 9468 CPU (Sapphire Rapids) + HBM - 48 cores - 2.1 GHz nominal frequency - 3.5 GHz turbo frequency (2.6 GHz with all cores utilized) - 350 W TDP	Intel RAPL	190 W	CF (2.6 - 0.8 GHz), UCF (2.5 - 0.8 GHz)
Nvidia Grace CPU superchip - 2x 72 cores (2x Grace CPUs) - 3.1 GHz nominal frequency - 500 W TDP	PDU	-	CF (3.3 - 0.9 GHz)

Table 4: Hardware architectures, available monitoring systems, and tuned power knobs

Hardware	Node hardware	System name
Nvidia A100	2× AMD 7763 CPU, 16× DDR4, 8× A100 GPU (with 40 GB HBM2 per GPU)	IT4I Karolina
Intel Sapphire Rapids	2× CPU, 2× 64GB HBM, 2× 128 GB (DDR5)	IT4I CS
Nvidia Grace CPU	1× superchip (2× 72 CPU cores) with 2× 120 GB LPDDR5X	IT4I CS

Table 5: Node hardware details for individual architectures. (note: CS stands for IT4I’s Complementary Systems partition).

2 Energy efficiency analysis methodology

We have performed the second energy-efficiency analysis of the SPACE applications by executing the codes on a new set of hardware platforms: (a) NVidia A100 GPUs, (a) Intel Sapphire Rapids CPU with HBM memory, and a NVidia ARM-based Grace CPU. In all cases, a single server has been used, as we do not have access to clusters of such hardware technologies that are equipped with an external power monitoring system.

The code owners contributed to this deliverable by installing their codes on the given partition and verifying that it runs and performs as expected. The codes were then instrumented (at least the main function was instrumented), compiled, and linked with the MERIC runtime systems. A set of submission scripts (one per platform) was prepared. Each script executed the code with the correct use case on all platforms. Because ChaNGa is code based on a Charm++ runtime that is not supported directly by MERIC, external measurements are performed similarly to the D2.2 deliverable.

General requirements are that (a) each case runs for at least one minute and no more than 5 minutes as we run a large number of experiments, and (b) the runs can use MPI, OpenMP or a hybrid version of the code. The CPU use case size should be defined to fit into 128 GB of Sapphire Rapids (2 CPUs, each 64 GB of HBM) memory and Grace-Grace (2 CPUs, each with 240 GB of RAM). The GPU version of the benchmark is designed to efficiently utilize all 8 A100 GPUs, each with 40 GB of memory, of Karolina’s GPU-accelerated nodes.

For this deliverable, we have executed static application tuning of CPU core (CF) and uncore (UCF) frequencies that influence the power consumption of the CPU, to identify available energy savings. To present the applications’ behavior in more detail, we executed an exhaustive state-space search; in other words, we executed the code for all possible HW configurations. This is a long process involving running many experiments, which can be shortened for production by using a more optimal search space algorithm converging from the initial to the optimal configuration in a shorter time. Resource consumption is visualized in heatmaps of absolute values (energy consumption in Joules, runtime in seconds) and relative values (in %) with respect to resource consumption in the default hardware configuration (i.e. when no tuning is performed).

Since energy efficiency is defined as performance per Watt, we have evaluated both (a) the number of

Hardware Platform	Nvidia A100 GPU	Sapphire Rapids w. DDR and HBM	Nvidia Grace CPU
Where	IT4I Karolina production system	IT4I complementary system	IT4I complementary system
Key features	Evaluation of accelerated platform	evaluation of modern x86 CPU and effect of DDR and HBM memory on energy efficiency	Evaluation of energy efficiency of modern ARM-based CPU to have comparison with x86
Tuning	static tuning of GPU streaming multiprocessor frequency	static tuning of CPU core frequency and CPU uncore frequency	static tuning of CPU core frequency
OpenGadget	yes	yes	yes
Pluto	yes	yes	yes
ChaNGa	yes	yes	yes
iPIC3D	yes	yes	yes
RAMSES	no	yes	yes
FIL	no	yes	yes
BHAC	no	yes	yes

Table 6: Summary of experiments to be presented in this deliverable

FLOP instructions executed by the application running on the CPU (both Intel and Grace CPUs provide this information through performance counters), and (b) energy consumption. For the latter, we use Intel RAPL (Sapphire Rapids CPU) and NVML (NVIDIA A100 GPU) in-band power monitoring features of the given platform, which allows us to measure chip energy consumption. In addition, complementary systems provide server power and energy consumption monitoring using a Power Distribution Unit (PDU) equipped with power monitoring.

The deliverable contains experiments and measurements shown in Table 6.

2.1 Use-cases

For the **Pluto** (version: 0.72) CPU code of the application, the MHD Orszag Tang (Configuration #11) benchmark was used with the iteration count limited to 20 and gridsize $468 \times 468 \times 468$. The GPU version of Pluto was tested with MHD Orszag Tang (Configuration #11) benchmark also limited to 150 iterations and gridsize $576 \times 576 \times 576$.

For the **OpenGadget** (version 0.42) CPU code of the application, the Cosmological box 2×256^3 120Mpc/h benchmark was used, with the running time limited to 5 minutes. The GPU version of OpenGadget was tested with a Cosmological box 2×256^3 120Mpc/h benchmark also limited to 5 min.

For **iPIC3D** (commit ID: 1c558b2) CPU code of the application, the GEM3D benchmark was used with the iteration count limited to 30 and a cell grid size $24 \times 24 \times 24$. The GPU version of iPIC3D (commit ID: 71a0823) was also tested with the GEM3D benchmark limited to 500 iterations and with a cell grid size $64 \times 48 \times 40$.

For the **RAMSES** (version 3) CPU code of the application, the Sedov blast benchmark was used with the iteration count (global time steps) limited to 250 and grid size $512 \times 512 \times 512$.

For both the CPU and GPU version of the **ChaNGa** code, we use a 70 Mpc/h cosmological box with $N = 80 \cdot 10^6$ particles, representing the final state of a dark matter simulation. We run the benchmark for 10 global timesteps.

For the CPU version of **FIL**, the benchmark that has been used to profile the code is a head-on collision of two Neutron stars with the iteration count limited to 5. The size of the grid that was used for the profiling is $128 \times 128 \times 64$ cells where symmetry of the problem in the z direction is used to halve the number of cells in the Z direction.

For **BHAC** (Black Hole Accretion Code), the relativistic magnetized accretion torus benchmark was run with a grid size of $480 \times 192 \times 192$, 50 iterations (~ 3 min runtime), and ~ 4 blocks per process using a 3D modified Kerr-Schild (MKS) coordinate system. The simulation solved the ideal RMHD equations with a gamma-law

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	203.1	202.5	203.1	201.0	203.4	204.5	206.0	207.4	210.6	216.6	229.3
2.6	203.4	203.3	203.3	203.2	201.6	203.4	206.0	207.6	211.0	217.0	231.6
2.5	205.5	205.8	206.5	206.1	207.1	208.5	207.9	211.5	215.2	221.0	227.3
2.3	227.8	215.3	215.1	215.3	214.4	217.6	219.1	220.7	224.1	231.5	243.8
2.1	228.3	228.2	226.0	226.6	227.6	242.3	232.0	229.7	233.2	241.2	254.8
1.9	252.0	239.8	239.7	241.0	241.6	242.8	244.3	246.4	252.1	257.3	267.6
1.7	268.3	260.2	257.8	258.3	258.9	260.2	277.1	266.0	266.5	272.6	287.4
1.5	292.0	286.1	280.8	281.8	283.4	283.5	292.2	287.3	291.0	295.0	312.3
1.3	324.1	311.4	312.5	316.2	313.7	314.3	320.9	322.8	319.8	327.1	343.7
1.1	369.4	356.5	364.1	356.9	362.5	359.0	364.9	363.2	366.7	375.3	380.9
0.9	439.7	426.1	426.3	422.8	429.5	427.2	424.9	433.1	436.0	435.1	454.5
0.8	488.1	474.8	474.6	469.2	476.9	476.2	478.2	479.5	481.3	486.6	498.6

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.7%	100.0%	98.9%	100.1%	100.7%	101.4%	102.1%	103.7%	106.6%	112.9%
2.6	100.1%	100.1%	100.1%	100.0%	99.3%	100.1%	101.4%	102.2%	103.9%	106.8%	114.0%
2.5	101.1%	101.3%	101.7%	101.4%	101.9%	102.7%	102.4%	104.1%	105.9%	108.8%	111.9%
2.3	112.1%	106.0%	105.9%	106.0%	105.5%	107.1%	107.9%	108.7%	110.3%	113.9%	120.0%
2.1	112.4%	112.3%	111.3%	111.6%	112.1%	119.3%	114.2%	113.1%	114.8%	118.7%	125.4%
1.9	124.0%	118.0%	118.0%	118.6%	118.9%	119.5%	120.3%	121.3%	124.1%	126.6%	131.7%
1.7	132.1%	128.1%	126.9%	127.1%	127.4%	128.1%	136.4%	130.9%	131.2%	134.2%	141.5%
1.5	143.7%	140.8%	138.2%	138.7%	139.5%	139.5%	143.8%	141.4%	143.2%	145.2%	153.7%
1.3	159.5%	153.3%	153.8%	155.7%	154.4%	154.7%	158.0%	158.9%	157.4%	161.0%	169.2%
1.1	181.9%	175.5%	179.2%	175.7%	178.4%	176.7%	179.6%	178.8%	180.5%	184.7%	187.5%
0.9	216.5%	209.7%	209.8%	208.1%	211.4%	210.3%	209.2%	213.2%	214.6%	214.2%	223.7%
0.8	240.3%	233.7%	233.6%	231.0%	234.7%	234.4%	235.4%	236.0%	236.9%	239.5%	245.5%

(b)

Figure 1: Heatmap example #1: Impact of the static tuning to **run time in the Sapphire Rapids system**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run. (results for Pluto code)

EOS, employing finite-volume discretization, HRSC methods, HLL Riemann solvers, and constrained transport. The total number of iterations was 10.

2.2 Report component description

Each system has allowed us to collect different data. These differences influence the way we visualize them. In the following paragraphs, we report the different formats we have used. Within this section we use results for the Pluto code as example.

2.2.1 Heatmaps

Data measured with Intel Sapphire Rapids are represented by *heatmaps*. A heatmap has core frequencies on the Y-axis and uncore frequencies on the X-axis. A color gradient is used to graphically distinguish the values inside the table. We provide two sets of heatmaps. One (Figure 1) shows the run-time duration of an individual CF+UCF configuration (Figure 1a) and its comparison with the default run 0/0 (Figure 1b). The second set (Figure 2) shows the energy consumption of a node for each CF+UCF combination (Figure 2a) and a comparison of the particular combination and the default configuration 0/0 (Figure 2b).

The **0/0 configuration** represents the default behavior of the CPU when no tuning is performed. In all measurements, this serves as a baseline against which we evaluate the runtime extension and energy savings.

In the case of the heatmap with the time differences (Figure 1b), a number greater than 100% signals a run-time extension i.e., the execution with particular frequency configuration was slower in comparison with the default run. In Figure 2b, a number less than 100% stands for a case in which less energy was consumed with the particular frequency configuration than in the default run, i.e., energy savings were achieved.

2.2.2 Energy consumption and savings tables

For Nvidia A100 and Grace CPU systems, we provide energy consumption and savings tables. Figure 3 shows a table example for A100 and Figure 4 for Grace CPU. We added a color gradient for an individual column

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	183.5	183.1	183.6	181.5	179.9	174.5	170.7	167.1	165.5	166.5	171.4
2.6	183.7	183.8	183.8	183.5	178.5	173.3	170.4	167.3	165.8	166.9	173.3
2.5	185.7	186.1	186.7	185.8	180.3	174.7	168.8	167.2	165.9	166.7	168.7
2.3	205.7	194.5	194.2	188.7	179.6	175.5	171.5	168.1	166.4	167.7	172.4
2.1	205.4	205.5	201.9	192.6	184.9	188.2	175.3	169.0	166.9	169.0	174.2
1.9	184.4	214.8	209.2	199.6	191.2	184.8	180.0	176.5	175.1	175.1	178.1
1.7	189.5	227.0	219.3	208.3	199.3	192.5	198.6	185.2	180.4	180.6	185.8
1.5	198.2	229.0	232.5	221.2	212.3	203.9	203.4	194.2	191.6	189.9	196.5
1.3	211.1	258.2	252.1	241.4	228.4	219.6	216.8	211.6	204.1	204.2	210.0
1.1	229.5	287.5	284.9	264.7	256.2	243.2	239.1	230.8	226.7	227.2	225.1
0.9	261.7	333.7	323.9	304.2	294.2	280.3	269.4	266.4	261.0	254.3	260.2
0.8	283.9	366.5	355.5	332.4	321.0	307.5	298.5	289.9	283.1	279.8	280.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.8%	100.0%	98.9%	98.1%	95.1%	93.0%	91.1%	90.2%	90.7%	93.4%
2.6	100.1%	100.2%	100.2%	100.0%	97.3%	94.4%	92.9%	91.2%	90.3%	90.9%	94.5%
2.5	101.2%	101.4%	101.7%	101.2%	98.3%	95.2%	92.0%	91.1%	90.4%	90.9%	91.9%
2.3	112.1%	106.0%	105.8%	102.8%	97.9%	95.7%	93.5%	91.6%	90.7%	91.4%	94.0%
2.1	112.0%	112.0%	110.0%	105.0%	100.8%	102.6%	95.5%	92.1%	91.0%	92.1%	94.9%
1.9	100.5%	117.0%	114.0%	108.7%	104.2%	100.7%	98.1%	96.2%	95.4%	95.4%	97.1%
1.7	103.3%	123.7%	119.5%	113.5%	108.6%	104.9%	108.2%	100.9%	98.3%	98.4%	101.3%
1.5	108.0%	124.8%	126.7%	120.5%	115.7%	111.1%	110.8%	105.9%	104.4%	103.5%	107.1%
1.3	115.0%	140.7%	137.4%	131.6%	124.5%	119.6%	118.1%	115.3%	111.2%	111.3%	114.4%
1.1	125.1%	156.7%	155.2%	144.2%	139.6%	132.5%	130.3%	125.8%	123.6%	123.8%	122.7%
0.9	142.6%	181.8%	176.5%	165.7%	160.3%	152.8%	146.8%	145.2%	142.3%	138.6%	141.8%
0.8	154.7%	199.7%	193.7%	181.1%	174.9%	167.6%	162.7%	158.0%	154.3%	152.4%	153.0%

(b)

Figure 2: Heatmap example #2: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids** system. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

to distinguish values graphically. Depending on the applied power monitoring system, we were able to collect different data. In the A100 system, we could obtain energy data separately for CPUs (AMD RAPL), GPUs (NVML), and complete node. For the Grace CPU system, we read the data from PDU which measures the energy consumption of the complete node. The columns “Runtime extension”, “Energy savings for GPUs+CPUs”, and “Compute node energy savings” compare the corresponding quantity (run time or energy consumption) in each row with the one from the default run. In the A100 tables, the default configuration is represented by the first row with 1410 MHz. In case of the Grace CPU, it is the first row starting with “zero” frequency. In the energy savings columns, a positive number signals energy savings, while a negative value represents energy overhead. Typically, this occurs with low frequencies when the run is slowed down too much, so that it consumes more energy in the end due to the long run time.

2.2.3 Summary tables

All SPACE codes are provided with a measurement summary table. An example is shown in Table 7. For every HW platform (architecture), there are three rows. The first row contains the default run, i.e., a run without any frequency tuning. In the second row, we show a configuration that provides energy savings but has the smallest possible impact on run time (up to approx. 3 % overhead). The third row shows a configuration with the maximum possible energy savings.

Set GPU frequency [MHz]	Runtime [s]	Runtime extension	Average power consumption of the GPU [W]	Average power consumption of the CPU [W]	CPUs + GPUs energy consumption [kJ]	Energy savings for GPUs + CPUs	Node energy consumption [kJ]	Compute node energy savings
1410	187.4	100.0%	185.8	118.4	322.9	0.0%	435.3	0.0%
1350	189.9	101.3%	172.7	119.0	307.5	4.8%	421.5	3.2%
1290	192.1	102.5%	160.0	119.6	291.8	9.6%	407.1	6.5%
1230	197.3	105.3%	149.0	119.7	282.4	12.5%	400.8	7.9%
1170	204.3	109.0%	138.9	120.1	276.1	14.5%	398.7	8.4%
1110	211.7	113.0%	129.5	119.4	270.0	16.4%	397.0	8.8%
1050	217.8	116.2%	122.1	119.8	264.8	18.0%	395.5	9.2%
990	225.7	120.4%	118.6	119.5	268.1	17.0%	403.5	7.3%
930	233.8	124.8%	115.9	119.4	272.8	15.5%	413.1	5.1%
870	235.1	125.5%	115.0	119.4	272.6	15.6%	413.7	5.0%
810	250.0	133.4%	111.1	119.1	281.7	12.7%	431.7	0.8%
750	265.3	141.6%	107.6	119.3	291.6	9.7%	450.8	-3.5%
690	280.5	149.7%	104.4	119.2	301.1	6.8%	469.4	-7.8%

Figure 3: Example of an energy consumption and savings table for the Nvidia A100 system (Karolina GPU)

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	111.0	100.0%	686.6	76.2	0.0%	628.2
3.30	107.0	96.4%	707.7	75.7	0.6%	632.2
3.10	114.0	102.7%	600.9	68.5	10.1%	698.8
2.90	112.0	100.9%	530.6	59.4	22.0%	805.4
2.70	122.0	109.9%	474.6	57.9	24.0%	826.7
2.50	130.0	117.1%	435.9	56.7	25.6%	844.7
2.30	134.0	120.7%	398.7	53.4	29.9%	895.8
2.10	135.0	121.6%	367.6	49.6	34.9%	964.4
1.90	143.0	128.8%	346.3	49.5	35.0%	966.5
1.70	161.0	145.0%	331.4	53.4	30.0%	896.9
1.50	168.0	151.4%	320.4	53.8	29.4%	888.9
1.30	187.0	168.5%	306.6	57.3	24.8%	834.6
1.10	222.0	200.0%	291.7	64.7	15.0%	739.0
0.90	262.0	236.0%	277.2	72.6	4.7%	658.9

Figure 4: Example of an energy consumption and savings table for the Grace CPU system

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Nvidia A100	default	-	435.3 kJ	187.4 s (100 %)
	1290 MHz	-	407.1 kJ (-6 %)	192.1 s (103 %)
	1110 MHz	-	397.0 kJ (-9 %)	211.7 s (113 %)
Intel Sapphire Rapids with DDR	default	304.8 MFLOPs/W	183.5 kJ	203.1 s (100 %)
	CF 2.6 GHz, UCF 1.4 GHz	345.5 MFLOPs/W	167.3 kJ (-9 %)	207.6 s (102 %)
	CF 2.5 GHz, UCF 1.2 GHz	353.2 MFLOPs/W	165.9 kJ (-10 %)	215.2 s (106 %)
Intel Sapphire Rapids with HBM	default	376.6 MFLOPs/W	149.0 kJ	166.8 s (100 %)
	CF 2.6 GHz, UCF 1.4 GHz	398.5 MFLOPs/W	142.9 kJ (-4 %)	168.9 s (101 %)
	CF 2.6 GHz, UCF 1.0 GHz	411.0 MFLOPs/W	140.6 kJ (-6 %)	174.4 s (105 %)
Nvidia Grace CPU	default	628.2 MFLOPs/W	76.2 kJ	111 s (100 %)
	CF 2.9 GHz	805.4 MFLOPs/W	59.4 kJ (-22 %)	112 s (101 %)
	CF 2.1 GHz	964.4 MFLOPs/W	49.6 kJ (-35 %)	135 s (122 %)

Table 7: An example of a summary table for a particular SPACE application

3 Results

This section provides the performance and energy measurements of all SPACE codes. For each code, we provide a short description of the results.

3.1 Pluto

PLUTO [7] is designed to integrate a general system of conservation laws that we write as

$$\frac{\partial U}{\partial t} = -\nabla \cdot T(U) + S(U). \quad (2)$$

Here U denotes a state vector of conservative quantities, $T(U)$ is a rank 2 tensor (the rows of which are the fluxes of each component) and $S(U)$ defines the source terms. Additional source terms may implicitly arise when taking the divergence of $T(U)$ in a curvilinear system of coordinates. An arbitrary number of advection equations may be added to the original conservation law.

Table 8 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 5 (A100), 6 and 7 (Sapphire Rapids with DDR), 8 and 9 (Sapphire Rapids with HBM), and 10 (Grace CPU).

For the first three architectures (A100, Sapphire Rapids with DDR and HBM), the maximum energy savings achieved reach up to 10 % (see Table 8). We may notice that with a small run time extension ($\leq 3\%$) we can already save 9 % of the energy in the case of Sapphire Rapids with DDR and 6 % for Nvidia A100. When comparing the results of the DDR and HBM memories (the size of test case was the same for both memory types), we may notice that despite the relative energy savings achieved with DDR, HBM leads to faster computation with less energy consumed and better overall energy efficiency. The greatest energy savings (and energy efficiency) were achieved with Grace CPU. By extending the run time by only 1 %, we could save up to 22 % of the energy and achieve an interesting energy efficiency. The maximum energy savings with both Sapphire Rapids configurations (i.e., 10 % for DDR and 6 % for HBM) was achieved with still a moderate run-time penalty, i.e., 5–6 %. With the Nvidia A100 and Grace CPU systems, the runtime penalty is higher, i.e., 13 % and 22 % respectively.

Set GPU frequency [MHz]	Runtime [s]	Runtime extension	Average power consumption of the GPU [W]	Average power consumption of the CPU [W]	CPUs + GPUs energy consumption [kJ]	Energy savings for GPUs + CPUs	Node energy consumption [kJ]	Compute node energy savings
1410	187.4	100.0%	185.8	118.4	322.9	0.0%	435.3	0.0%
1350	189.9	101.3%	172.7	119.0	307.5	4.8%	421.5	3.2%
1290	192.1	102.5%	160.0	119.6	291.8	9.6%	407.1	6.5%
1230	197.3	105.3%	149.0	119.7	282.4	12.5%	400.8	7.9%
1170	204.3	109.0%	138.9	120.1	276.1	14.5%	398.7	8.4%
1110	211.7	113.0%	129.5	119.4	270.0	16.4%	397.0	8.8%
1050	217.8	116.2%	122.1	119.8	264.8	18.0%	395.5	9.2%
990	225.7	120.4%	118.6	119.5	268.1	17.0%	403.5	7.3%
930	233.8	124.8%	115.9	119.4	272.8	15.5%	413.1	5.1%
870	235.1	125.5%	115.0	119.4	272.6	15.6%	413.7	5.0%
810	250.0	133.4%	111.1	119.1	281.7	12.7%	431.7	0.8%
750	265.3	141.6%	107.6	119.3	291.6	9.7%	450.8	-3.5%
690	280.5	149.7%	104.4	119.2	301.1	6.8%	469.4	-7.8%

Figure 5: Pluto: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	203.1	202.5	203.1	201.0	203.4	204.5	206.0	207.4	210.6	216.6	229.3
2.6	203.4	203.3	203.3	203.2	201.6	203.4	206.0	207.6	211.0	217.0	231.6
2.5	205.5	205.8	206.5	206.1	207.1	208.5	207.9	211.5	215.2	221.0	227.3
2.3	227.8	215.3	215.1	215.3	214.4	217.6	219.1	220.7	224.1	231.5	243.8
2.1	228.3	228.2	226.0	226.6	227.6	242.3	232.0	229.7	233.2	241.2	254.8
1.9	252.0	239.8	239.7	241.0	241.6	242.8	244.3	246.4	252.1	257.3	267.6
1.7	268.3	260.2	257.8	258.3	258.9	260.2	277.1	266.0	266.5	272.6	287.4
1.5	292.0	286.1	280.8	281.8	283.4	283.5	292.2	287.3	291.0	295.0	312.3
1.3	324.1	311.4	312.5	316.2	313.7	314.3	320.9	322.8	319.8	327.1	343.7
1.1	369.4	356.5	364.1	356.9	362.5	359.0	364.9	363.2	366.7	375.3	380.9
0.9	439.7	426.1	426.3	422.8	429.5	427.2	424.9	433.1	436.0	435.1	454.5
0.8	488.1	474.8	474.6	469.2	476.9	476.2	478.2	479.5	481.3	486.6	498.6

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.7%	100.0%	98.9%	100.1%	100.7%	101.4%	102.1%	103.7%	106.6%	112.9%
2.6	100.1%	100.1%	100.1%	100.0%	99.3%	100.1%	101.4%	102.2%	103.9%	106.8%	114.0%
2.5	101.1%	101.3%	101.7%	101.4%	101.9%	102.7%	102.4%	104.1%	105.9%	108.8%	111.9%
2.3	112.1%	106.0%	105.9%	106.0%	105.5%	107.1%	107.9%	108.7%	110.3%	113.9%	120.0%
2.1	112.4%	112.3%	111.3%	111.6%	112.1%	119.3%	114.2%	113.1%	114.8%	118.7%	125.4%
1.9	124.0%	118.0%	118.0%	118.6%	118.9%	119.5%	120.3%	121.3%	124.1%	126.6%	131.7%
1.7	132.1%	128.1%	126.9%	127.1%	127.4%	128.1%	136.4%	130.9%	131.2%	134.2%	141.5%
1.5	143.7%	140.8%	138.2%	138.7%	139.5%	139.5%	143.8%	141.4%	143.2%	145.2%	153.7%
1.3	159.5%	153.3%	153.8%	155.7%	154.4%	154.7%	158.0%	158.9%	157.4%	161.0%	169.2%
1.1	181.9%	175.5%	179.2%	175.7%	178.4%	176.7%	179.6%	178.8%	180.5%	184.7%	187.5%
0.9	216.5%	209.7%	209.8%	208.1%	211.4%	210.3%	209.2%	213.2%	214.6%	214.2%	223.7%
0.8	240.3%	233.7%	233.6%	231.0%	234.7%	234.4%	235.4%	236.0%	236.9%	239.5%	245.5%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	183.5	183.1	183.6	181.5	179.9	174.5	170.7	167.1	165.5	166.5	171.4
2.6	183.7	183.8	183.8	183.5	178.5	173.3	170.4	167.3	165.8	166.9	173.3
2.5	185.7	186.1	186.7	185.8	180.3	174.7	168.8	167.2	165.9	166.7	168.7
2.3	205.7	194.5	194.2	188.7	179.6	175.5	171.5	168.1	166.4	167.7	172.4
2.1	205.4	205.5	201.9	192.6	184.9	188.2	175.3	169.0	166.9	169.0	174.2
1.9	184.4	214.8	209.2	199.6	191.2	184.8	180.0	176.5	175.1	175.1	178.1
1.7	189.5	227.0	219.3	208.3	199.3	192.5	198.6	185.2	180.4	180.6	185.8
1.5	198.2	229.0	232.5	221.2	212.3	203.9	203.4	194.2	191.6	189.9	196.5
1.3	211.1	258.2	252.1	241.4	228.4	219.6	216.8	211.6	204.1	204.2	210.0
1.1	229.5	287.5	284.9	264.7	256.2	243.2	239.1	230.8	226.7	227.2	225.1
0.9	261.7	333.7	323.9	304.2	294.2	280.3	269.4	266.4	261.0	254.3	260.2
0.8	283.9	366.5	355.5	332.4	321.0	307.5	298.5	289.9	283.1	279.8	280.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.8%	100.0%	98.9%	98.1%	95.1%	93.0%	91.1%	90.2%	90.7%	93.4%
2.6	100.1%	100.2%	100.2%	100.0%	97.3%	94.4%	92.9%	91.2%	90.3%	90.9%	94.5%
2.5	101.2%	101.4%	101.7%	101.2%	98.3%	95.2%	92.0%	91.1%	90.4%	90.9%	91.9%
2.3	112.1%	106.0%	105.8%	102.8%	97.9%	95.7%	93.5%	91.6%	90.7%	91.4%	94.0%
2.1	112.0%	112.0%	110.0%	105.0%	100.8%	102.6%	95.5%	92.1%	91.0%	92.1%	94.9%
1.9	100.5%	117.0%	114.0%	108.7%	104.2%	100.7%	98.1%	96.2%	95.4%	95.4%	97.1%
1.7	103.3%	123.7%	119.5%	113.5%	108.6%	104.9%	108.2%	100.9%	98.3%	98.4%	101.3%
1.5	108.0%	124.8%	126.7%	120.5%	115.7%	111.1%	110.8%	105.9%	104.4%	103.5%	107.1%
1.3	115.0%	140.7%	137.4%	131.6%	124.5%	119.6%	118.1%	115.3%	111.2%	111.3%	114.4%
1.1	125.1%	156.7%	155.2%	144.2%	139.6%	132.5%	130.3%	125.8%	123.6%	123.8%	122.7%
0.9	142.6%	181.8%	176.5%	165.7%	160.3%	152.8%	146.8%	145.2%	142.3%	138.6%	141.8%
0.8	154.7%	199.7%	193.7%	181.1%	174.9%	167.6%	162.7%	158.0%	154.3%	152.4%	153.0%

(b)

Figure 6: Pluto: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 7: Pluto: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	166.8	165.9	168.2	168.2	165.5	167.2	167.0	166.6	172.6	177.0	183.5
2.6	167.1	166.3	168.4	168.7	165.3	166.0	164.4	168.9	172.7	174.4	184.8
2.5	167.8	171.0	171.0	177.1	169.8	170.0	169.6	173.1	176.6	181.8	187.7
2.3	179.5	179.8	177.6	179.2	180.4	181.4	179.5	185.1	188.8	192.6	198.0
2.1	192.4	192.2	192.2	189.8	193.2	197.9	196.0	195.3	202.0	206.3	214.5
1.9	273.7	209.3	207.1	206.3	209.5	211.5	213.0	214.6	219.2	219.6	230.5
1.7	243.2	229.1	229.5	229.6	231.0	232.3	233.2	235.7	238.9	240.9	247.4
1.5	273.7	253.0	253.4	254.9	255.0	258.8	259.6	261.5	264.8	267.0	278.0
1.3	303.7	290.9	289.3	291.7	292.3	297.5	294.1	296.7	313.5	303.2	314.7
1.1	353.2	339.0	344.3	339.5	340.4	348.5	342.6	344.9	352.0	350.5	356.8
0.9	443.5	409.8	416.3	416.6	410.9	415.5	419.0	412.8	429.7	420.5	431.2
0.8	478.7	471.0	458.9	481.8	466.2	467.3	468.8	470.8	472.2	467.7	487.3

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.5%	100.8%	100.8%	99.2%	100.2%	100.1%	99.9%	103.5%	106.1%	110.0%
2.6	100.1%	99.7%	101.0%	101.1%	99.1%	99.5%	98.6%	101.2%	103.5%	104.5%	110.8%
2.5	100.6%	102.5%	102.5%	106.2%	101.8%	101.9%	101.7%	103.8%	105.9%	109.0%	112.5%
2.3	107.6%	107.8%	106.5%	107.4%	108.2%	108.7%	107.6%	111.0%	113.2%	115.5%	118.7%
2.1	115.3%	115.2%	115.2%	113.8%	115.8%	118.6%	117.5%	117.1%	121.1%	123.7%	128.6%
1.9	164.1%	125.5%	124.1%	123.6%	125.6%	126.8%	127.7%	128.6%	131.4%	131.7%	138.2%
1.7	145.8%	137.3%	137.6%	137.6%	138.5%	139.3%	139.8%	141.3%	143.2%	144.4%	148.3%
1.5	164.1%	151.6%	151.9%	152.8%	152.9%	155.1%	155.6%	156.8%	158.7%	160.0%	166.7%
1.3	182.0%	174.4%	173.5%	174.8%	175.2%	178.3%	176.3%	177.9%	187.9%	181.8%	188.7%
1.1	211.7%	203.2%	206.4%	203.5%	204.1%	208.9%	205.4%	206.7%	211.0%	210.1%	213.9%
0.9	265.9%	245.6%	249.6%	249.7%	246.3%	249.1%	251.2%	247.4%	257.6%	252.1%	258.5%
0.8	286.9%	282.3%	275.1%	288.8%	279.5%	280.1%	281.0%	282.2%	283.0%	280.4%	292.2%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	149.0	148.6	150.6	150.6	148.1	149.1	145.2	140.9	142.3	142.8	144.3
2.6	149.3	148.9	150.8	151.1	147.9	148.1	142.9	142.9	142.4	140.6	145.4
2.5	150.2	153.1	153.1	158.5	151.7	148.9	144.4	143.4	142.4	143.5	144.3
2.3	160.7	161.0	159.0	160.1	157.3	152.3	146.3	146.6	145.7	145.3	145.7
2.1	171.9	172.0	171.8	166.1	162.2	159.6	153.4	148.5	149.4	149.4	151.7
1.9	192.1	187.0	184.1	175.7	170.6	165.3	161.5	158.0	157.0	154.1	158.1
1.7	165.0	203.2	199.0	189.3	182.0	175.7	171.2	168.0	165.7	163.5	164.1
1.5	179.1	218.8	213.7	203.9	194.7	189.6	184.2	180.3	177.6	175.2	178.5
1.3	191.4	244.4	236.2	225.8	215.9	210.7	201.8	197.5	202.7	192.2	195.2
1.1	214.1	276.3	272.4	254.3	243.2	238.4	227.1	221.6	219.9	214.6	213.7
0.9	258.2	323.3	318.5	301.8	283.8	274.7	267.5	256.0	259.1	247.9	249.2
0.8	273.4	365.3	345.7	342.8	316.0	303.7	294.1	286.6	279.6	270.7	276.3

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.7%	101.1%	101.1%	99.4%	100.1%	97.5%	94.6%	95.5%	95.9%	96.9%
2.6	100.2%	99.9%	101.2%	101.4%	99.3%	99.4%	95.9%	95.9%	95.6%	94.4%	97.6%
2.5	100.8%	102.8%	102.8%	106.4%	101.8%	100.0%	96.9%	96.2%	95.6%	96.3%	96.9%
2.3	107.9%	108.0%	106.7%	107.5%	105.6%	102.2%	98.2%	98.4%	97.8%	97.5%	97.8%
2.1	115.4%	115.5%	115.3%	111.5%	108.8%	107.1%	103.0%	99.7%	100.3%	100.3%	101.8%
1.9	128.9%	125.5%	123.5%	117.9%	114.5%	110.9%	108.4%	106.1%	105.4%	103.4%	106.1%
1.7	110.7%	136.4%	133.6%	127.1%	122.2%	117.9%	114.9%	112.8%	111.2%	109.7%	110.1%
1.5	120.2%	146.9%	143.4%	136.9%	130.7%	127.3%	123.6%	121.0%	119.2%	117.6%	119.8%
1.3	128.5%	164.1%	158.6%	151.5%	144.9%	141.4%	135.4%	132.6%	136.1%	129.0%	131.0%
1.1	143.7%	185.4%	182.8%	170.7%	163.2%	160.0%	152.4%	148.7%	147.6%	144.0%	143.4%
0.9	173.3%	217.0%	213.8%	202.5%	190.5%	184.4%	179.5%	171.8%	173.9%	166.4%	167.2%
0.8	183.5%	245.2%	232.1%	230.1%	212.1%	203.8%	197.4%	192.4%	187.7%	181.7%	185.4%

(b)

Figure 8: Pluto: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 9: Pluto: Impact of the static tuning to **node energy consumption [kJ]** in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	111.0	100.0%	686.6	76.2	0.0%	628.2
3.30	107.0	96.4%	707.7	75.7	0.6%	632.2
3.10	114.0	102.7%	600.9	68.5	10.1%	698.8
2.90	112.0	100.9%	530.6	59.4	22.0%	805.4
2.70	122.0	109.9%	474.6	57.9	24.0%	826.7
2.50	130.0	117.1%	435.9	56.7	25.6%	844.7
2.30	134.0	120.7%	398.7	53.4	29.9%	895.8
2.10	135.0	121.6%	367.6	49.6	34.9%	964.4
1.90	143.0	128.8%	346.3	49.5	35.0%	966.5
1.70	161.0	145.0%	331.4	53.4	30.0%	896.9
1.50	168.0	151.4%	320.4	53.8	29.4%	888.9
1.30	187.0	168.5%	306.6	57.3	24.8%	834.6
1.10	222.0	200.0%	291.7	64.7	15.0%	739.0
0.90	262.0	236.0%	277.2	72.6	4.7%	658.9

Figure 10: Pluto: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Nvidia A100	default	-	435.3 kJ	187.4 s (100 %)
	1290 MHz	-	407.1 kJ (-6 %)	192.1 s (103 %)
	1110 MHz	-	397.0 kJ (-9 %)	211.7 s (113 %)
Intel Sapphire Rapids with DDR	default	240.7 MFLOPs/W	183.5 kJ	203.1 s (100 %)
	CF 2.6 GHz, UCF 1.4 GHz	264.0 MFLOPs/W	167.3 kJ (-9 %)	207.6 s (102 %)
	CF 2.5 GHz, UCF 1.2 GHz	266.2 MFLOPs/W	165.9 kJ (-10 %)	215.2 s (106 %)
Intel Sapphire Rapids with HBM	default	296.4 MFLOPs/W	149.0 kJ	166.8 s (100 %)
	CF 2.6 GHz, UCF 1.4 GHz	309.0 MFLOPs/W	142.9 kJ (-4 %)	168.9 s (101 %)
	CF 2.6 GHz, UCF 1.0 GHz	314.1 MFLOPs/W	140.6 kJ (-6 %)	174.4 s (105 %)
Nvidia Grace CPU	default	628.2 MFLOPs/W	76.2 kJ	111 s (100 %)
	CF 2.9 GHz	805.4 MFLOPs/W	59.4 kJ (-22 %)	112 s (101 %)
	CF 2.1 GHz	964.4 MFLOPs/W	49.6 kJ (-35 %)	135 s (122 %)

Table 8: Pluto: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

3.2 OpenGadget

OpenGadget [8] is a code for cosmological simulations. It solves the gravitational and hydrodynamical equations that rule the formation and evolution of cosmic structures. It computes the gravitational forces with a hierarchical tree algorithm in combination with a particle-mesh scheme for long-range gravitational forces. Then the fluid flows are computed using smoothed particle hydrodynamics (SPH) or Meshless Finite Mass (MFM). In addition, OpenGadget contains multiple sub-modules to describe various physical processes (e.g. radiative cooling, star formation, stellar feedback, magnetic fields, black holes, just to name the major ones) that shape the fundamental properties of baryons in the Universe.

Table 9 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 11 (A100), 12 and 13 (Sapphire Rapids with DDR), 14 and 15 (Sapphire Rapids with HBM), and 16 (Grace CPU).

The static frequency tuning has low impact on the run time in the case of Nvidia A100. We may notice it in Figure 11 where the maximum run-time extension is 3.8 %. The most promising configuration (930 MHz) saves 7 % of the energy and prolongs the run time by only 1.6 %. Similar energy savings were achieved with Sapphire Rapids and DDR memory. The run time was extended by 2 %. An interesting behavior may be observed with Sapphire Rapids and HBM memory. By reducing the uncore frequency we may achieve energy savings and, unexpectedly, also a speedup. We found two interesting configurations (see Table 9): (1) CF 2.6 GHz + UCF 1.6 GHz and (2) 2.6 GHz + UCF 1.0 GHz. Both are faster than the default one, but differ mutually in speedup and energy savings. The reason why this behavior is possible is related to the maximum TDP of the processor and its frequency in this state. Theoretically, the Sapphire Rapids CPU allows all cores to achieve 2.6 GHz (CF) when they are fully utilized. However, when the uncore frequency is also high (i.e., 2.0 - 2.5 GHz), the real core frequency is reduced to approx. 2550 MHz in order to maintain the maximum allowed level of TDP. When we reduce the uncore frequency (i.e., 1.8 GHz and below), the CPU allows us to increase the core frequency back to 2.6 GHz (259X MHz in reality according to our measurements). This causes the effect that we see in our results, i.e., the speedup when the uncore frequency is reduced. The results of the measurement with Grace CPU (Table 9) show that with the 3 % run-time penalty 13 % of the energy can be saved. It is possible to save up to 33 % of the energy but this would prolong the run time by 28 %.

Set GPU frequency [MHz]	Runtime [s]	Runtime extension	Average power consumption of the GPU [W]	Average power consumption of the CPU [W]	CPUs + GPUs energy consumption [kJ]	Energy savings for GPUs + CPUs	Node energy consumption [kJ]	Compute node energy savings
1410	190.3	100.0%	73.2	196.9	186.4	0.0%	300.6	0.0%
1350	191.6	100.7%	68.8	198.5	181.6	2.6%	296.5	1.3%
1290	189.6	99.6%	65.6	200.6	175.5	5.8%	289.3	3.7%
1230	190.5	100.1%	62.8	200.1	171.9	7.8%	286.2	4.8%
1170	192.9	101.4%	60.6	199.2	170.4	8.6%	286.2	4.8%
1110	190.9	100.3%	58.7	200.3	166.1	10.9%	280.7	6.6%
1050	192.2	101.0%	56.9	200.1	164.4	11.8%	279.7	6.9%
990	193.1	101.5%	56.1	199.4	163.6	12.2%	279.4	7.0%
930	193.3	101.6%	55.6	199.2	163.0	12.6%	279.0	7.2%
870	196.6	103.3%	55.2	198.1	164.7	11.6%	282.7	5.9%
810	195.8	102.9%	54.8	199.0	163.7	12.2%	281.2	6.4%
750	197.2	103.6%	54.3	197.4	163.5	12.3%	281.9	6.2%
690	197.5	103.8%	53.9	197.9	163.3	12.4%	281.8	6.2%

Figure 11: OpenGadget: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	107.0	106.3	108.4	106.5	106.7	107.2	106.0	108.7	108.9	113.7	117.5
2.6	106.8	106.3	108.0	106.3	108.3	106.0	106.0	108.2	108.9	113.7	117.6
2.5	107.8	109.3	108.0	109.1	108.2	108.6	111.2	110.3	114.3	115.5	122.6
2.3	115.0	117.0	114.9	116.3	115.5	116.1	119.0	117.7	121.7	122.6	129.6
2.1	124.0	125.2	123.7	126.5	124.4	126.0	125.8	128.2	128.4	133.3	137.4
1.9	141.7	134.3	136.8	136.1	135.2	137.4	136.7	139.5	139.1	143.8	147.8
1.7	155.2	149.9	147.7	149.9	150.5	148.9	151.1	152.5	152.4	158.5	160.7
1.5	173.7	181.4	165.1	165.3	168.1	166.7	166.9	169.9	169.5	175.6	180.9
1.3	196.3	186.7	189.2	190.5	191.0	188.0	191.4	196.2	194.2	194.4	203.7
1.1	226.0	220.9	221.4	220.2	218.3	220.8	221.9	222.9	224.8	228.9	232.9
0.9	273.3	264.8	265.4	267.4	265.8	267.4	266.6	268.7	267.8	274.6	278.0
0.8	306.1	302.9	301.8	297.9	297.6	298.9	300.1	305.5	303.0	304.8	310.4

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.4%	101.3%	99.5%	99.8%	100.2%	99.1%	101.6%	101.9%	106.3%	109.8%
2.6	99.9%	99.4%	101.0%	99.4%	101.3%	99.1%	99.1%	101.2%	101.8%	106.3%	109.9%
2.5	100.8%	102.2%	101.0%	102.0%	101.2%	101.5%	104.0%	103.1%	106.9%	108.0%	114.7%
2.3	107.5%	109.4%	107.5%	108.7%	108.0%	108.5%	111.2%	110.1%	113.8%	114.6%	121.1%
2.1	115.9%	117.0%	115.7%	118.3%	116.4%	117.8%	117.6%	119.9%	120.1%	124.7%	128.5%
1.9	132.5%	125.6%	127.9%	127.2%	126.4%	128.4%	127.8%	130.4%	130.1%	134.5%	138.2%
1.7	145.2%	140.1%	138.1%	140.2%	140.8%	139.3%	141.3%	142.6%	142.5%	148.2%	150.3%
1.5	162.5%	169.6%	154.4%	154.5%	157.2%	155.9%	156.1%	158.9%	158.5%	164.2%	169.2%
1.3	183.5%	174.6%	176.9%	178.1%	178.6%	175.8%	178.9%	183.5%	181.5%	181.8%	190.5%
1.1	211.3%	206.6%	207.0%	205.9%	204.1%	206.4%	207.5%	208.4%	210.2%	214.0%	217.8%
0.9	255.5%	247.6%	248.1%	250.0%	248.5%	250.0%	249.3%	251.3%	250.4%	256.8%	259.9%
0.8	286.2%	283.2%	282.2%	278.5%	278.2%	279.5%	280.5%	285.6%	283.3%	285.0%	290.2%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	95.9	95.5	97.4	95.6	95.4	94.0	91.0	91.2	89.4	91.6	92.6
2.6	95.9	95.5	97.0	95.5	96.9	93.0	91.1	90.8	89.3	91.7	92.8
2.5	96.9	98.2	97.0	98.0	95.7	93.5	93.3	90.4	91.6	90.7	94.4
2.3	103.3	105.1	103.2	102.7	98.2	95.3	95.1	91.7	92.7	91.7	94.8
2.1	111.1	112.2	109.9	107.5	101.4	99.1	96.1	95.6	93.6	95.2	96.1
1.9	98.5	119.0	118.3	112.1	106.8	104.6	101.1	100.6	97.9	99.1	99.9
1.7	104.0	129.1	123.9	119.6	115.2	109.7	108.1	106.2	103.6	105.6	104.9
1.5	111.8	151.7	134.5	127.9	124.5	118.9	115.5	114.4	111.2	112.9	114.1
1.3	121.5	151.7	149.7	142.8	137.1	129.8	127.9	127.5	123.0	120.6	124.2
1.1	134.3	174.4	170.1	160.4	151.6	147.3	143.3	139.8	137.4	137.2	136.5
0.9	156.1	203.1	198.0	188.8	178.7	172.5	166.4	162.9	158.7	158.8	157.3
0.8	171.6	228.8	221.6	206.9	197.2	189.9	184.3	182.0	175.6	172.8	172.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.5%	101.5%	99.7%	99.5%	98.0%	94.9%	95.1%	93.2%	95.5%	96.5%
2.6	100.0%	99.6%	101.1%	99.5%	101.0%	96.9%	94.9%	94.6%	93.1%	95.6%	96.7%
2.5	101.0%	102.3%	101.1%	102.1%	99.8%	97.4%	97.2%	94.2%	95.5%	94.6%	98.4%
2.3	107.6%	109.5%	107.6%	107.0%	102.3%	99.3%	99.1%	95.6%	96.6%	95.5%	98.8%
2.1	115.8%	117.0%	114.5%	112.0%	105.7%	103.3%	100.2%	99.7%	97.5%	99.2%	100.2%
1.9	102.6%	124.0%	123.3%	116.8%	111.3%	109.0%	105.4%	104.8%	102.0%	103.3%	104.1%
1.7	108.3%	134.5%	129.2%	124.6%	120.0%	114.4%	112.7%	110.7%	107.9%	110.0%	109.3%
1.5	116.5%	158.1%	140.2%	133.3%	129.8%	123.9%	120.4%	119.2%	115.9%	117.6%	118.9%
1.3	126.7%	158.2%	156.1%	148.9%	142.8%	135.3%	133.3%	132.9%	128.2%	125.7%	129.4%
1.1	139.9%	181.8%	177.3%	167.2%	158.0%	153.5%	149.3%	145.7%	143.2%	143.0%	142.3%
0.9	162.7%	211.7%	206.4%	196.8%	186.2%	179.8%	173.5%	169.8%	165.4%	165.5%	164.0%
0.8	178.8%	238.5%	230.9%	215.7%	205.5%	197.9%	192.1%	189.7%	183.0%	180.1%	180.0%

(b)

Figure 12: OpenGadget: Impact of the static tuning to **run time** in the **Sapphire Rapids** system with the **DDR memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 13: OpenGadget: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids** system with the **DDR memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.9	95.7	96.7	95.7	95.6	96.7	94.5	97.5	96.6	98.8	104.1
2.6	102.8	95.6	95.6	96.9	95.7	96.5	94.5	95.4	99.1	98.8	102.7
2.5	104.6	98.6	97.0	97.0	97.8	97.2	99.9	98.7	100.0	103.8	105.5
2.3	106.2	104.2	104.2	105.9	104.2	106.3	105.4	107.8	107.6	110.0	115.4
2.1	113.1	115.5	113.3	115.2	113.4	115.9	114.4	116.9	116.6	118.6	124.5
1.9	130.0	130.8	125.1	124.2	124.4	125.5	125.6	127.8	127.7	131.0	135.9
1.7	143.7	140.5	137.5	139.4	138.1	140.3	139.1	142.4	143.6	143.0	149.8
1.5	161.1	156.7	158.0	155.1	157.1	155.9	158.7	159.2	158.3	163.1	165.5
1.3	183.7	179.8	179.7	181.3	178.2	183.7	183.3	180.1	182.8	187.4	188.9
1.1	219.7	208.9	213.6	213.5	211.1	213.7	213.6	221.4	212.1	217.4	219.7
0.9	263.8	258.7	258.1	258.9	258.3	260.6	260.5	261.4	261.4	262.2	265.6
0.8	302.5	298.2	290.4	294.4	292.7	291.4	292.3	292.3	292.8	295.4	301.2

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	94.8%	95.8%	94.8%	94.7%	95.8%	93.6%	96.6%	95.8%	97.9%	103.2%
2.6	101.8%	94.7%	94.7%	96.1%	94.9%	95.6%	93.7%	94.5%	98.2%	97.9%	101.8%
2.5	103.7%	97.7%	96.1%	96.1%	96.9%	96.3%	99.0%	97.8%	99.1%	102.9%	104.5%
2.3	105.2%	103.2%	103.2%	104.9%	103.3%	105.3%	104.4%	106.8%	106.6%	109.0%	114.4%
2.1	112.0%	114.4%	112.3%	114.1%	112.4%	114.8%	113.4%	115.8%	115.6%	117.5%	123.3%
1.9	128.8%	129.6%	124.0%	123.1%	123.3%	124.3%	124.4%	126.6%	126.6%	129.8%	134.7%
1.7	142.4%	139.2%	136.3%	138.2%	136.9%	139.1%	137.8%	141.1%	142.2%	141.7%	148.4%
1.5	159.6%	155.2%	156.6%	153.7%	155.7%	154.5%	157.2%	157.8%	156.9%	161.6%	164.0%
1.3	182.0%	178.1%	178.1%	179.6%	176.6%	182.0%	181.7%	178.5%	181.2%	185.6%	187.2%
1.1	217.7%	207.0%	211.6%	211.6%	209.2%	211.7%	211.7%	219.4%	210.2%	215.4%	217.7%
0.9	261.4%	256.3%	255.7%	256.6%	256.0%	258.2%	258.1%	259.0%	259.0%	259.8%	263.2%
0.8	299.8%	295.5%	287.8%	291.7%	290.0%	288.8%	289.6%	289.6%	290.1%	292.7%	298.4%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	90.3	85.7	86.6	85.7	85.5	86.0	82.8	83.5	80.9	81.3	84.1
2.6	92.0	85.6	85.7	86.8	85.7	85.9	82.7	81.6	83.0	81.4	83.1
2.5	93.7	88.3	86.9	86.9	87.2	85.2	85.2	82.3	81.7	83.2	82.9
2.3	95.1	93.3	93.3	94.1	89.9	88.5	85.4	85.2	83.2	83.6	86.0
2.1	100.8	103.2	100.8	98.6	93.4	92.1	88.5	88.2	86.0	85.8	88.5
1.9	89.7	115.8	108.5	103.1	99.1	96.4	93.7	93.0	90.8	91.3	93.0
1.7	95.4	120.9	115.7	111.9	106.2	104.1	100.2	99.8	98.2	96.1	98.7
1.5	103.0	130.9	129.0	120.5	116.7	111.7	110.3	107.8	104.5	105.5	105.2
1.3	113.2	146.1	142.3	136.4	128.3	127.1	123.0	117.5	116.2	116.8	115.8
1.1	129.9	165.0	164.2	155.7	146.9	142.8	138.3	139.2	130.1	130.8	129.4
0.9	150.1	198.4	192.3	182.9	173.9	168.4	162.9	158.6	154.5	151.9	150.7
0.8	168.8	225.4	213.4	204.7	194.0	185.1	179.6	174.3	169.8	167.7	167.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	95.0%	95.9%	95.0%	94.8%	95.3%	91.7%	92.5%	89.7%	90.1%	93.2%
2.6	101.9%	94.8%	94.9%	96.2%	95.0%	95.1%	91.7%	90.4%	91.9%	90.2%	92.0%
2.5	103.8%	97.8%	96.3%	96.3%	96.6%	94.4%	94.4%	91.2%	90.5%	92.2%	91.9%
2.3	105.3%	103.3%	103.4%	104.3%	99.6%	98.0%	94.6%	94.4%	92.2%	92.6%	95.3%
2.1	111.7%	114.4%	111.7%	109.3%	103.5%	102.1%	98.1%	97.8%	95.3%	95.1%	98.0%
1.9	99.4%	128.3%	120.2%	114.2%	109.8%	106.8%	103.8%	103.0%	100.6%	101.2%	103.1%
1.7	105.7%	133.9%	128.2%	124.0%	117.7%	115.3%	111.1%	110.6%	108.8%	106.5%	109.4%
1.5	114.1%	145.0%	142.9%	133.5%	129.3%	123.8%	122.2%	119.4%	115.8%	116.9%	116.5%
1.3	125.4%	161.9%	157.7%	151.1%	142.1%	140.9%	136.3%	130.2%	128.8%	129.4%	128.3%
1.1	143.9%	182.8%	181.9%	172.5%	162.8%	158.2%	153.2%	154.2%	144.1%	144.9%	143.3%
0.9	166.3%	219.8%	213.1%	202.6%	192.7%	186.5%	180.5%	175.7%	171.1%	168.3%	167.0%
0.8	187.1%	249.8%	236.5%	226.8%	214.9%	205.1%	199.0%	193.1%	188.2%	185.8%	185.9%

(b)

Figure 14: OpenGadget: Impact of the static tuning to **run time** in the **Sapphire Rapids** system with the **HBM memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 15: OpenGadget: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids** system with the **HBM memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	78.0	100.0%	662.2	51.6	0.0%	626.7
3.30	77.0	98.7%	678.5	52.2	-1.2%	619.6
3.10	80.0	102.6%	565.0	45.2	12.5%	716.2
2.90	82.0	105.1%	494.0	40.5	21.6%	799.2
2.70	90.0	115.4%	450.9	40.6	21.4%	797.7
2.50	90.0	115.4%	411.3	37.0	28.3%	874.5
2.30	99.0	126.9%	374.4	37.1	28.2%	873.3
2.10	100.0	128.2%	344.1	34.4	33.4%	940.7
1.90	113.0	144.9%	323.6	36.6	29.2%	885.5
1.70	114.0	146.2%	310.9	35.4	31.4%	913.6
1.50	131.0	167.9%	299.1	39.2	24.1%	826.3
1.30	135.0	173.1%	287.2	38.8	24.9%	835.3
1.10	152.0	194.9%	276.2	42.0	18.7%	771.4
0.90	188.0	241.0%	263.1	49.5	4.2%	655.0

Figure 16: OpenGadget: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Nvidia A100	default	-	300.6 kJ	190.3 s (100 %)
	930 MHz	-	279.0 kJ (-7 %)	193.3 s (101.6 %)
	930 MHz	-	279.0 kJ (-7 %)	193.3 s (101.6 %)
Intel Sapphire Rapids with DDR	default	128.7 MFLOPs/W	95.9 kJ	107.0 s (100 %)
	CF 2.6 GHz, UCF 1.2 GHz	138.2 MFLOPs/W	89.3 kJ (-7 %)	108.9 s (102 %)
	CF 2.6 GHz, UCF 1.2 GHz	138.2 MFLOPs/W	89.3 kJ (-7 %)	108.9 s (102 %)
Intel Sapphire Rapids with HBM	default	136.8 MFLOPs/W	90.3 kJ	100.9 s (100 %)
	CF 2.6 GHz, UCF 1.6 GHz	149.2 MFLOPs/W	82.7 kJ (-9 %)	94.5 s (94 %)
	CF 2.6 GHz, UCF 1.0 GHz	151.7 MFLOPs/W	81.4 kJ (-11 %)	98.8 s (98 %)
Nvidia Grace CPU	default	626.7 MFLOPs/W	51.6 kJ	78 s (100 %)
	CF 3.1 GHz	716.2 MFLOPs/W	45.2 kJ (-13 %)	80 s (103 %)
	CF 2.1 GHz	940.7 MFLOPs/W	34.4 kJ (-33 %)	100 s (128 %)

Table 9: OpenGadget: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

3.3 iPIC3D

iPIC3D [9] is a **PIC!** (**PIC!**) code developed to study plasma dynamics at the kinetic scale, where the individual (macro)particles of a plasma are evolved in a Lagrangian framework, whereas the moments (electric current, density, etc.) and the electric and magnetic fields are tracked on an Eulerian grid. It solves the Maxwell-Vlasov equation self-consistently using the implicit moment method. The three main kernels of iPIC3D are (a) Particle Mover, (b) Moment Gatherer, and (c) Field Solver.

Owing to the “implicitness” of the method involved, unresolved scales do not result in any numerical instabilities, as opposed to explicit **PIC!** methods, which are restricted to the Courant-Friedrichs-Lewy constraint. This allows us to choose time-step sizes and grid sizes that can be 10 to 100 times larger than those used in explicit PIC codes.

Table 10 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 17 (A100), 18 and 19 (Sapphire Rapids with DDR), 20 and 21 (Sapphire Rapids with HBM), and 22 (Grace CPU).

The results (Table 10) show that in the case of configurations with the smallest possible run-time penalty ($\leq 3.5\%$) we may achieve energy savings between 3% to 9% depending on the architecture (A100: 3%, SPR+DDR: 7%, SPR+HBM: 7%, and Grace: 9%). The Grace CPU reaches the most interesting savings here. For Sapphire Rapids with HBM memory, the configuration with the smallest possible impact on run time (CF 2.6 GHz, UCF 1.6 GHz) also represents the one with the highest possible energy savings. The configurations with the highest possible energy savings for A100 and SPR+DDR do not bring additional significant savings (5% and 9%) compared to those with the smallest possible impact on the run time (3% and 7%). Grace CPU allowed us to obtain interesting maximum savings (29%) but with a run-time penalty of 26%.

Set GPU frequency [MHz]	Runtime [s]	Runtime extension	Average power consumption of the GPU [W]	Average power consumption of the CPU [W]	CPUs + GPUs energy consumption [kJ]	Energy savings for GPUs + CPUs	Node energy consumption [kJ]	Compute node energy savings
1410	239.0	100.0%	158.7	126.5	363.9	0.0%	507.3	0.0%
1350	247.4	103.5%	142.2	127.5	344.5	5.3%	492.9	2.8%
1290	256.7	107.4%	129.5	128.0	331.7	8.9%	485.7	4.3%
1230	265.4	111.1%	119.5	127.9	321.6	11.6%	480.8	5.2%
1170	280.8	117.5%	109.1	127.8	316.9	12.9%	485.4	4.3%
1110	288.4	120.7%	102.1	127.9	309.4	15.0%	482.4	4.9%
1050	296.7	124.2%	96.7	127.9	305.4	16.1%	483.4	4.7%
990	314.6	131.7%	92.3	128.1	313.0	14.0%	501.8	1.1%
930	330.3	138.2%	90.4	127.9	323.5	11.1%	521.6	-2.8%
870	352.2	147.4%	88.0	128.0	338.0	7.1%	549.3	-8.3%
810	372.5	155.9%	85.9	128.0	351.4	3.4%	574.9	-13.3%
750	396.3	165.8%	83.7	128.1	366.8	-0.8%	604.5	-19.2%
690	424.9	177.8%	81.2	128.2	385.0	-5.8%	639.9	-26.1%

Figure 17: iPIC3D: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	153.4	153.9	153.2	154.2	154.5	155.8	158.8	162.4	168.2	181.7	202.0
2.6	153.3	153.7	153.6	153.1	154.4	156.4	158.9	162.6	168.3	181.7	202.3
2.5	156.5	156.5	155.8	157.0	157.8	159.8	162.6	165.6	172.7	185.5	205.9
2.3	163.1	163.2	168.5	164.5	166.6	168.6	170.6	174.6	181.6	196.1	214.4
2.1	173.2	173.3	180.4	175.1	181.1	179.0	181.0	184.9	191.9	204.7	225.1
1.9	192.4	185.9	186.7	186.9	188.7	191.8	194.2	198.0	214.4	216.7	237.3
1.7	209.0	201.2	202.0	202.2	204.9	207.0	209.7	213.3	220.1	231.9	252.3
1.5	228.5	220.3	221.3	223.0	225.1	227.0	229.5	233.5	239.0	253.2	272.9
1.3	257.7	246.1	246.8	248.3	251.4	252.9	255.0	261.1	266.3	277.8	298.8
1.1	292.2	283.0	283.5	285.1	286.8	289.7	291.3	295.1	302.2	314.5	336.3
0.9	344.1	335.3	338.2	337.7	338.9	341.6	346.2	347.9	353.8	367.5	384.9
0.8	382.8	371.4	372.4	372.3	378.9	375.5	379.7	386.4	389.9	402.9	427.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	100.3%	99.9%	100.5%	100.7%	101.6%	103.5%	105.9%	109.7%	118.5%	131.7%
2.6	100.0%	100.2%	100.1%	99.8%	100.6%	101.9%	103.6%	106.0%	109.7%	118.5%	131.9%
2.5	102.0%	102.0%	101.6%	102.3%	102.8%	104.2%	106.0%	107.9%	112.6%	121.0%	134.2%
2.3	106.3%	106.4%	109.8%	107.2%	108.6%	109.9%	111.2%	113.8%	118.4%	127.8%	139.8%
2.1	112.9%	113.0%	117.6%	114.2%	118.1%	116.7%	118.0%	120.5%	125.1%	133.4%	146.7%
1.9	125.4%	121.2%	121.7%	121.8%	123.0%	125.0%	126.6%	129.1%	139.8%	141.3%	154.7%
1.7	136.2%	131.2%	131.7%	131.8%	133.6%	135.0%	136.7%	139.1%	143.5%	151.2%	164.5%
1.5	149.0%	143.6%	144.3%	145.3%	146.8%	148.0%	149.6%	152.2%	155.8%	165.0%	177.9%
1.3	168.0%	160.5%	160.9%	161.9%	163.9%	164.9%	166.3%	170.2%	173.6%	181.1%	194.8%
1.1	190.5%	184.5%	184.8%	185.9%	186.9%	188.8%	189.9%	192.4%	197.0%	205.0%	219.3%
0.9	224.3%	218.6%	220.5%	220.1%	220.9%	222.7%	225.7%	226.8%	230.6%	239.5%	250.9%
0.8	249.5%	242.1%	242.8%	242.7%	247.0%	244.8%	247.5%	251.9%	254.2%	262.6%	278.9%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	138.0	138.6	137.8	136.5	132.5	128.7	127.5	127.1	128.3	135.9	147.5
2.6	137.9	138.5	138.1	135.6	132.3	129.2	127.7	127.3	128.4	135.9	147.9
2.5	140.9	140.9	139.5	137.6	132.5	129.5	127.9	126.9	129.1	135.8	147.3
2.3	145.6	145.9	148.8	139.0	134.8	131.3	128.9	128.5	130.3	137.9	147.0
2.1	152.2	152.5	155.2	142.9	141.7	134.6	131.9	131.0	132.5	138.3	148.6
1.9	147.0	160.0	156.2	148.4	143.4	140.2	137.6	136.4	140.1	142.2	152.3
1.7	141.2	168.8	165.0	156.7	151.6	147.1	144.3	142.7	143.3	147.6	157.1
1.5	149.3	180.2	176.2	168.2	162.0	157.0	153.5	151.8	151.2	156.6	165.3
1.3	162.4	196.6	191.2	182.4	176.1	170.0	165.9	164.9	163.8	166.8	175.7
1.1	177.5	220.3	214.2	204.1	195.5	189.5	184.1	181.2	180.3	183.2	192.0
0.9	201.3	254.6	249.0	235.3	224.9	217.2	212.5	207.1	204.7	208.2	213.2
0.8	219.7	278.2	270.8	255.9	247.9	235.6	229.8	226.7	222.3	224.7	233.6

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	100.5%	99.8%	98.9%	96.0%	93.2%	92.4%	92.1%	93.0%	98.5%	106.9%
2.6	100.0%	100.3%	100.1%	98.2%	95.9%	93.6%	92.5%	92.2%	93.0%	98.4%	107.1%
2.5	102.1%	102.1%	101.1%	99.7%	96.0%	93.8%	92.7%	91.9%	93.6%	98.4%	106.7%
2.3	105.5%	105.7%	107.9%	100.7%	97.7%	95.1%	93.4%	93.1%	94.4%	99.9%	106.5%
2.1	110.3%	110.5%	112.4%	103.5%	102.7%	97.5%	95.6%	94.9%	96.0%	100.2%	107.7%
1.9	106.5%	115.9%	113.2%	107.5%	103.9%	101.6%	99.7%	98.8%	101.5%	103.0%	110.4%
1.7	102.3%	122.3%	119.6%	113.5%	109.8%	106.6%	104.6%	103.4%	103.9%	106.9%	113.8%
1.5	108.2%	130.6%	127.7%	121.9%	117.4%	113.8%	111.2%	110.0%	109.5%	113.5%	119.8%
1.3	117.7%	142.5%	138.5%	132.2%	127.6%	123.2%	120.2%	119.5%	118.7%	120.9%	127.3%
1.1	128.7%	159.6%	155.2%	147.9%	141.7%	137.3%	133.4%	131.3%	130.7%	132.8%	139.1%
0.9	145.8%	184.5%	180.4%	170.5%	163.0%	157.4%	154.0%	150.1%	148.4%	150.9%	154.5%
0.8	159.2%	201.6%	196.3%	185.4%	179.7%	170.7%	166.5%	164.3%	161.1%	162.8%	169.3%

(b)

Figure 18: iPIC3D: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 19: iPIC3D: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	115.2	116.4	115.2	116.1	113.9	114.3	116.6	119.8	123.5	127.7	137.1
2.6	115.0	116.5	115.2	115.5	114.3	115.2	116.0	119.4	122.4	128.6	135.8
2.5	119.5	118.3	119.5	117.3	118.5	118.1	121.1	125.7	130.1	131.5	140.6
2.3	126.1	126.8	127.2	126.0	125.4	128.2	129.1	132.0	137.0	141.7	152.3
2.1	135.5	136.8	135.4	137.0	136.2	138.2	140.3	144.7	146.9	150.8	159.6
1.9	152.3	148.2	149.2	148.2	150.2	151.7	151.9	156.5	158.4	164.4	172.4
1.7	178.8	163.4	165.2	166.6	164.6	167.6	168.7	170.4	176.1	180.6	186.5
1.5	188.7	185.2	185.9	184.0	187.0	188.2	190.3	190.1	195.5	200.3	208.1
1.3	215.2	210.3	211.5	218.6	219.0	214.5	216.2	218.2	219.8	227.2	235.0
1.1	252.3	248.9	249.5	250.0	251.9	249.4	252.0	260.7	257.7	262.3	269.6
0.9	307.4	302.9	303.0	307.5	304.2	310.4	306.9	314.9	312.7	314.7	322.2
0.8	345.7	343.8	343.0	341.1	366.5	355.0	343.5	344.1	349.7	350.8	363.0

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	101.0%	100.0%	100.8%	98.9%	99.2%	101.2%	104.0%	107.2%	110.8%	119.0%
2.6	99.8%	101.2%	100.0%	100.2%	99.2%	100.0%	100.7%	103.7%	106.3%	111.7%	117.9%
2.5	103.8%	102.7%	103.7%	101.8%	102.9%	102.6%	105.1%	109.1%	112.9%	114.2%	122.1%
2.3	109.5%	110.1%	110.5%	109.3%	108.9%	111.3%	112.1%	114.6%	118.9%	123.0%	132.2%
2.1	117.6%	118.7%	117.5%	118.9%	118.2%	119.9%	121.8%	125.6%	127.5%	130.9%	138.5%
1.9	132.2%	128.7%	129.5%	128.7%	130.4%	131.7%	131.9%	135.8%	137.5%	142.7%	149.7%
1.7	155.2%	141.8%	143.4%	144.6%	142.9%	145.5%	146.5%	147.9%	152.9%	156.7%	161.9%
1.5	163.8%	160.7%	161.4%	159.7%	162.3%	163.3%	165.2%	165.0%	169.7%	173.9%	180.7%
1.3	186.8%	182.6%	183.6%	189.8%	190.1%	186.2%	187.7%	189.4%	190.8%	197.2%	204.0%
1.1	219.0%	216.1%	216.6%	217.0%	218.6%	216.5%	218.8%	226.3%	223.7%	227.7%	234.0%
0.9	266.8%	262.9%	263.1%	266.9%	264.0%	269.4%	266.4%	273.3%	271.4%	273.2%	279.7%
0.8	300.1%	298.5%	297.7%	296.1%	318.1%	308.2%	298.2%	298.7%	303.6%	304.5%	315.1%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	102.8	104.2	103.1	102.8	99.1	97.1	96.3	96.5	97.0	98.2	103.4
2.6	102.9	104.4	103.1	102.3	99.5	97.9	95.8	96.2	96.1	99.1	102.3
2.5	107.0	105.9	106.7	103.0	101.9	98.2	97.9	98.4	100.0	98.8	103.6
2.3	112.2	112.9	112.1	108.2	103.7	102.1	99.9	99.4	100.5	101.9	107.5
2.1	118.8	120.0	117.0	113.7	108.4	105.8	104.4	104.7	103.5	104.0	107.8
1.9	117.6	128.0	126.0	119.5	116.0	112.7	109.4	109.7	108.1	109.9	113.0
1.7	118.3	137.6	135.9	130.5	123.4	120.7	117.7	115.5	116.4	116.9	118.2
1.5	120.3	152.2	148.9	140.2	136.1	131.6	128.7	125.0	125.2	125.6	128.0
1.3	133.1	168.4	164.9	161.1	153.8	145.4	141.8	139.0	136.2	137.9	140.4
1.1	150.2	194.2	189.4	179.7	172.6	164.0	160.1	160.6	154.9	154.1	155.4
0.9	175.8	230.2	223.5	214.4	202.5	197.5	188.9	188.3	181.8	179.1	179.5
0.8	194.0	257.7	249.8	235.0	240.3	224.3	208.4	202.5	200.2	196.4	199.2

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	101.3%	100.3%	100.0%	96.4%	94.4%	93.6%	93.8%	94.3%	95.5%	100.5%
2.6	100.1%	101.5%	100.3%	99.4%	96.7%	95.2%	93.1%	93.5%	93.4%	96.4%	99.4%
2.5	104.0%	103.0%	103.8%	100.2%	99.1%	95.4%	95.2%	95.6%	97.2%	96.1%	100.7%
2.3	109.1%	109.8%	109.0%	105.2%	100.8%	99.3%	97.1%	96.6%	97.8%	99.1%	104.5%
2.1	115.5%	116.6%	113.8%	110.5%	105.4%	102.9%	101.5%	101.8%	100.6%	101.1%	104.9%
1.9	114.3%	124.5%	122.5%	116.2%	112.8%	109.6%	106.4%	106.7%	105.1%	106.9%	109.9%
1.7	115.0%	133.8%	132.1%	126.9%	120.0%	117.4%	114.5%	112.3%	113.2%	113.7%	114.9%
1.5	117.0%	148.0%	144.8%	136.3%	132.3%	127.9%	125.1%	121.5%	121.8%	122.1%	124.4%
1.3	129.4%	163.7%	160.3%	156.7%	149.6%	141.4%	137.8%	135.2%	132.5%	134.1%	136.5%
1.1	146.0%	188.8%	184.1%	174.8%	167.8%	159.5%	155.7%	156.1%	150.6%	149.9%	151.1%
0.9	171.0%	223.8%	217.4%	208.5%	196.9%	192.1%	183.7%	183.1%	176.8%	174.2%	174.5%
0.8	188.7%	250.5%	242.9%	228.5%	233.6%	218.1%	202.6%	196.9%	194.6%	191.0%	193.6%

(b)

Figure 20: iPIC3D: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 21: iPIC3D: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	77.0	100.0%	601.6	46.3	0.0%	722.9
3.30	78.0	101.3%	593.4	46.3	0.1%	723.5
3.10	79.0	102.6%	535.8	42.3	8.6%	791.3
2.90	87.0	113.0%	479.2	41.7	10.0%	803.3
2.70	84.0	109.1%	428.4	36.0	22.3%	930.7
2.50	88.0	114.3%	394.3	34.7	25.1%	965.1
2.30	92.0	119.5%	362.7	33.4	28.0%	1003.5
2.10	97.0	126.0%	338.1	32.8	29.2%	1021.0
1.90	111.0	144.2%	318.2	35.3	23.7%	948.0
1.70	115.0	149.4%	305.4	35.1	24.2%	953.6
1.50	122.0	158.4%	298.5	36.4	21.4%	919.6
1.30	141.0	183.1%	287.1	40.5	12.6%	827.4
1.10	158.0	205.2%	276.6	43.7	5.7%	766.3
0.90	173.0	224.7%	266.5	46.1	0.5%	726.3

Figure 22: iPIC3D: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Nvidia A100	default	-	507.3 kJ	239.0 s (100 %)
	1350 MHz	-	492.9 kJ (-3 %)	247.4 s (103.5 %)
	1230 MHz	-	480.8 kJ (-5 %)	265.4 s (111 %)
Intel Sapphire Rapids with DDR	default	223.0 MFLOPs/W	138.0 kJ	153.4 s (100 %)
	CF 2.6 GHz, UCF 1.8 GHz	238.2 MFLOPs/W	129.2 kJ (-7 %)	156.4 s (102 %)
	CF 2.5 GHz, UCF 1.4 GHz	242.6 MFLOPs/W	126.9 kJ (-9 %)	165.6 s (108 %)
Intel Sapphire Rapids with HBM	default	299.2 MFLOPs/W	102.8 kJ	115.2 s (100 %)
	CF 2.6 GHz, UCF 1.6 GHz	321.4 MFLOPs/W	95.8 kJ (-7 %)	116.0 s (101 %)
	CF 2.6 GHz, UCF 1.6 GHz	321.4 MFLOPs/W	95.8 kJ (-7 %)	116.0 s (101 %)
Nvidia Grace CPU	default	722.9 MFLOPs/W	46.3 kJ	77 s (100 %)
	CF 3.1 GHz	791.3 MFLOPs/W	42.3 kJ (-9 %)	79 s (103 %)
	CF 2.1 GHz	964.4 MFLOPs/W	32.8 kJ (-29 %)	97 s (126 %)

Table 10: iPIC3D: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

3.4 RAMSES

RAMSES [10, 11] is an Adaptive-Mesh-Refinement (AMR) code which is used to study astrophysical fluid dynamics and the formation of structures in the Universe. It is based on an oct-tree structure, where parent cells are refined into children cells on a cell-by-cell basis following some user-defined criteria. RAMSES can deal with 1D, 2D and 3D Cartesian grids.

Table 11 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 23 and 24 (Sapphire Rapids with DDR), 25 and 26 (Sapphire Rapids with HBM), and 27 (Grace CPU). The Sapphire Rapids machine can achieve maximum energy savings with 4 % runtime increase for both DDR and HBM memory. With DDR the savings are 7 % and for HBM 8 % of the energy. The energy savings are reached by lowering the uncore frequency. Lowering core frequency increases the overall energy consumption. Using HBM memory instead of DDR improves runtime by 3 %. Significant energy saving can be reached on Grace machine. With only 1 % runtime penalty we can save 19 % of energy by lowering the core frequency to 2.9 GHz. With 37 % of addition to the original runtime we can achieve maximum 28 % of energy saving by using 2.3 GHz core frequency.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Intel Sapphire Rapids with DDR	default	374.1 MFLOPs/W	110.6 kJ	123.6 s (100 %)
	CF 2.6 GHz, UCF 1.4 GHz	399.7 MFLOPs/W	103.5 kJ (-6 %)	125.7 s (102 %)
	CF 2.6 GHz, UCF 1.2 GHz	401.6 MFLOPs/W	103.1 kJ (-7 %)	128.7 s (104 %)
Intel Sapphire Rapids with HBM	default	386.8 MFLOPs/W	107.0 kJ	120.0 s (100 %)
	CF 2.6 GHz, UCF 1.2 GHz	417.8 MFLOPs/W	99.1 kJ (-7 %)	122.2 s (102 %)
	CF default, UCF 1.0 GHz	418.5 MFLOPs/W	98.9 kJ (-8 %)	124.3 s (104 %)
Nvidia Grace CPU	default	695.3 MFLOPs/W	67.1 kJ	117 s (100 %)
	CF 2.9 GHz	854.9 MFLOPs/W	54.5 kJ (-19 %)	118 s (101 %)
	CF 2.3 GHz	968.6 MFLOPs/W	48.1 kJ (-28 %)	160 s (137 %)

Table 11: Ramses: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	123.6	125.1	124.5	124.9	124.7	124.6	125.5	127.8	128.6	133.0	138.0
2.6	124.4	124.6	124.8	124.1	124.2	124.2	125.5	125.7	128.7	132.7	137.9
2.5	127.4	127.8	127.8	128.1	128.1	129.1	129.9	131.9	134.9	137.8	142.9
2.3	137.0	138.0	136.9	138.9	139.0	139.0	139.7	140.5	144.1	145.3	153.8
2.1	148.0	148.0	149.0	147.4	149.7	149.2	151.3	153.9	155.0	157.4	161.4
1.9	174.2	161.7	162.5	161.0	163.1	164.3	165.9	166.7	167.6	171.8	182.4
1.7	195.3	178.7	179.6	182.0	180.1	178.9	182.8	185.3	185.2	185.6	194.4
1.5	213.9	205.7	201.4	200.1	203.4	203.9	205.6	207.0	208.7	207.9	216.5
1.3	243.6	231.3	234.2	234.7	235.3	234.4	235.5	237.4	240.3	240.9	247.9
1.1	283.6	273.3	272.7	275.1	275.6	278.9	278.8	281.3	277.8	280.8	287.5
0.9	341.3	336.8	332.6	333.8	340.1	333.7	335.5	339.3	338.0	339.3	351.8
0.8	386.4	380.8	377.0	378.9	375.0	377.8	375.8	379.6	380.1	385.2	385.3

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	101.2%	100.7%	101.0%	100.8%	100.8%	101.5%	103.4%	104.0%	107.5%	111.6%
2.6	100.7%	100.7%	100.9%	100.3%	100.4%	100.5%	101.5%	101.7%	104.1%	107.3%	111.5%
2.5	103.0%	103.4%	103.4%	103.6%	103.6%	104.4%	105.0%	106.7%	109.1%	111.4%	115.6%
2.3	110.8%	111.6%	110.7%	112.3%	112.4%	112.4%	113.0%	113.7%	116.5%	117.5%	124.4%
2.1	119.7%	119.7%	120.5%	119.2%	121.1%	120.7%	122.4%	124.5%	125.4%	127.3%	130.5%
1.9	140.9%	130.8%	131.5%	130.2%	131.9%	132.9%	134.2%	134.9%	135.6%	138.9%	147.6%
1.7	158.0%	144.5%	145.2%	147.2%	145.7%	144.7%	147.8%	149.8%	149.8%	150.1%	157.2%
1.5	173.0%	166.3%	162.9%	161.9%	164.5%	165.0%	166.3%	167.4%	168.8%	168.2%	175.1%
1.3	197.0%	187.1%	189.4%	189.9%	190.3%	189.6%	190.4%	192.0%	194.3%	194.9%	200.5%
1.1	229.3%	221.0%	220.6%	222.5%	222.9%	225.6%	225.5%	227.5%	224.7%	227.1%	232.5%
0.9	276.1%	272.4%	269.0%	270.0%	275.1%	269.9%	271.3%	274.4%	273.4%	274.4%	284.5%
0.8	312.5%	308.0%	304.9%	306.5%	303.3%	305.6%	304.0%	307.1%	307.4%	311.5%	311.6%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	110.6	112.2	111.7	111.7	110.9	107.8	105.6	104.7	103.1	104.3	105.7
2.6	111.3	111.7	111.9	111.0	110.6	107.5	105.6	103.5	103.1	103.9	105.7
2.5	113.8	114.4	114.3	114.1	112.0	108.9	106.6	105.5	105.1	105.1	106.5
2.3	122.3	123.1	122.1	120.8	116.2	112.1	109.5	107.3	107.0	106.2	109.1
2.1	131.5	132.0	130.9	123.9	120.3	115.8	113.6	112.4	110.4	109.8	110.3
1.9	115.5	141.8	138.9	131.3	127.0	123.0	120.4	117.7	115.8	115.5	118.8
1.7	124.3	152.2	148.7	143.2	135.6	130.2	128.1	126.4	122.8	121.3	123.3
1.5	132.4	169.0	162.0	153.2	147.6	142.6	139.4	136.5	133.5	131.1	132.6
1.3	144.9	184.9	182.0	172.9	165.5	158.6	154.2	151.0	148.5	146.3	146.9
1.1	163.4	212.7	206.4	196.9	187.9	182.6	176.0	172.3	166.6	164.6	165.0
0.9	190.1	254.0	244.5	231.9	224.1	212.2	205.9	201.2	195.9	192.4	194.8
0.8	211.1	283.2	272.7	258.5	244.7	235.6	226.9	221.9	216.5	213.7	210.7

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100%	101%	101%	101%	100%	97%	96%	95%	93%	94%	96%
2.6	101%	101%	101%	100%	100%	97%	95%	94%	93%	94%	96%
2.5	103%	103%	103%	103%	101%	98%	96%	95%	95%	95%	96%
2.3	111%	111%	110%	109%	105%	101%	99%	97%	97%	96%	99%
2.1	119%	119%	118%	112%	109%	105%	103%	102%	100%	99%	100%
1.9	104%	128%	126%	119%	115%	111%	109%	106%	105%	104%	107%
1.7	112%	138%	134%	129%	123%	118%	116%	114%	111%	110%	111%
1.5	120%	153%	146%	138%	133%	129%	126%	123%	121%	119%	120%
1.3	131%	167%	165%	156%	150%	143%	139%	136%	134%	132%	133%
1.1	148%	192%	187%	178%	170%	165%	159%	156%	151%	149%	149%
0.9	172%	230%	221%	210%	203%	192%	186%	182%	177%	174%	176%
0.8	191%	256%	247%	234%	221%	213%	205%	201%	196%	193%	190%

(b)

Figure 23: Ramses: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 24: Ramses: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	120.0	120.5	119.5	119.8	119.1	119.0	120.7	120.7	124.0	124.3	131.8
2.6	119.9	120.3	120.0	120.2	120.7	120.5	120.7	120.6	122.2	124.8	128.6
2.5	123.8	122.8	123.9	123.0	124.0	123.3	125.5	125.5	128.9	128.8	137.5
2.3	132.0	135.2	132.1	133.2	133.0	134.4	133.6	137.0	138.4	138.6	149.0
2.1	143.9	145.0	145.3	144.4	147.9	145.5	146.6	146.6	150.0	152.3	154.9
1.9	169.6	160.8	157.3	159.2	158.4	160.4	163.7	161.1	164.9	167.0	167.9
1.7	187.6	178.0	174.3	177.9	177.5	177.2	176.7	179.8	183.8	183.8	185.7
1.5	209.9	198.0	198.1	200.8	196.8	200.1	203.9	202.2	207.4	205.0	208.6
1.3	243.3	227.9	229.8	232.0	232.2	235.0	235.5	231.3	239.2	242.2	246.6
1.1	287.5	273.9	273.4	274.4	274.5	276.0	277.3	277.7	280.6	283.9	289.3
0.9	352.8	333.8	335.3	336.4	335.0	336.8	338.1	340.5	342.0	349.1	349.3
0.8	392.1	377.9	381.3	378.4	383.4	381.1	385.5	384.7	388.5	392.9	405.5

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	100.4%	99.6%	99.8%	99.3%	99.1%	100.6%	100.5%	103.3%	103.5%	109.8%
2.6	99.9%	100.2%	100.0%	100.2%	100.5%	100.4%	100.5%	100.4%	101.8%	104.0%	107.2%
2.5	103.2%	102.3%	103.2%	102.4%	103.3%	102.7%	104.6%	104.5%	107.4%	107.3%	114.5%
2.3	110.0%	112.6%	110.1%	111.0%	110.8%	112.0%	111.3%	114.1%	115.3%	115.5%	124.1%
2.1	119.9%	120.8%	121.0%	120.3%	123.2%	121.2%	122.2%	122.1%	124.9%	126.9%	129.1%
1.9	141.3%	134.0%	131.1%	132.7%	132.0%	133.6%	136.4%	134.2%	137.4%	139.1%	139.9%
1.7	156.3%	148.3%	145.2%	148.2%	147.8%	147.6%	147.2%	149.8%	153.1%	153.2%	154.7%
1.5	174.9%	165.0%	165.0%	167.3%	163.9%	166.7%	169.9%	168.5%	172.7%	170.8%	173.8%
1.3	202.7%	189.9%	191.5%	193.2%	193.5%	195.8%	196.2%	192.7%	199.2%	201.8%	205.4%
1.1	239.5%	228.2%	227.8%	228.6%	228.7%	229.9%	231.0%	231.4%	233.8%	236.5%	241.0%
0.9	293.9%	278.1%	279.3%	280.2%	279.1%	280.5%	281.7%	283.7%	284.9%	290.8%	291.0%
0.8	326.6%	314.8%	317.6%	315.2%	319.4%	317.5%	321.1%	320.5%	323.7%	327.3%	337.8%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	107.0	107.8	106.9	106.9	106.1	103.9	102.3	100.1	100.0	98.9	101.9
2.6	106.8	107.6	107.1	107.4	107.2	104.9	102.3	100.1	99.1	99.2	100.2
2.5	110.4	109.6	110.5	109.4	109.0	104.9	103.7	101.3	101.2	99.7	103.3
2.3	117.3	120.2	117.7	116.7	112.1	109.0	105.7	105.0	103.7	102.1	106.0
2.1	127.8	129.0	128.0	121.5	118.9	113.3	110.7	107.9	107.3	106.9	106.6
1.9	112.6	140.5	134.9	129.6	123.9	120.4	118.9	114.5	113.6	112.9	111.8
1.7	120.1	151.3	145.0	140.1	133.8	128.8	124.7	123.0	122.3	119.9	119.2
1.5	129.7	163.6	159.4	152.9	144.3	140.1	138.1	133.3	133.1	129.3	129.0
1.3	144.4	182.3	178.7	170.9	163.6	158.4	153.9	147.3	147.6	146.3	145.9
1.1	164.9	212.6	206.5	196.1	187.0	180.2	175.3	170.3	167.3	165.9	165.3
0.9	195.0	252.0	245.7	233.3	221.2	213.3	206.8	202.0	197.4	197.0	193.1
0.8	213.5	281.4	275.4	258.5	249.0	237.6	231.7	224.5	220.1	217.5	219.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100%	101%	100%	100%	99%	97%	96%	94%	93%	92%	95%
2.6	100%	101%	100%	100%	100%	98%	96%	94%	93%	93%	94%
2.5	103%	102%	103%	102%	102%	98%	97%	95%	95%	93%	97%
2.3	110%	112%	110%	109%	105%	102%	99%	98%	97%	95%	99%
2.1	119%	121%	120%	114%	111%	106%	104%	101%	100%	100%	100%
1.9	105%	131%	126%	121%	116%	112%	111%	107%	106%	106%	105%
1.7	112%	141%	136%	131%	125%	120%	117%	115%	114%	112%	111%
1.5	121%	153%	149%	143%	135%	131%	129%	125%	124%	121%	121%
1.3	135%	170%	167%	160%	153%	148%	144%	138%	138%	137%	136%
1.1	154%	199%	193%	183%	175%	168%	164%	159%	156%	155%	154%
0.9	182%	235%	230%	218%	207%	199%	193%	189%	184%	184%	180%
0.8	200%	263%	257%	242%	233%	222%	217%	210%	206%	203%	205%

(b)

Figure 25: Ramses: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 26: Ramses: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	117.0	100.0%	573.8	67.1	0.0%	695.3
3.30	118.0	100.9%	574.6	67.8	-1.0%	688.3
3.10	115.0	98.3%	523.4	60.2	10.3%	774.6
2.90	118.0	100.9%	462.1	54.5	18.8%	854.9
2.70	127.0	108.5%	407.7	51.8	22.9%	900.4
2.50	141.0	120.5%	365.6	51.6	23.2%	904.8
2.30	147.0	125.6%	327.4	48.1	28.3%	968.6
2.10	160.0	136.8%	304.6	48.7	27.4%	956.8
1.90	173.0	147.9%	285.6	49.4	26.4%	943.3
1.70	201.0	171.8%	270.6	54.4	19.0%	857.1
1.50	214.0	182.9%	263.5	56.4	16.0%	826.4
1.30	244.0	208.5%	254.4	62.1	7.5%	750.7
1.10	271.0	231.6%	247.0	66.9	0.3%	696.0
0.90	327.0	279.5%	237.1	77.5	-15.5%	600.7

Figure 27: RAMSES: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

3.5 BHAC

The Black Hole Accretion Code (BHAC) [12, 13, 14, 15] is a multidimensional General Relativistic Magneto-hydrodynamics (GRMHD) code that solves the equations of ideal GRMHD in one, two or three dimensions in order to perform (magneto)hydrodynamical simulations of accretion flows onto compact objects in arbitrary stationary space-times (Cowling approximation) using an efficient block based approach. BHAC is build upon the MPI-Adaptive Mesh Refinement-Versatile Advection Code (MPI-AMRVAC). MPI-AMRVAC [16, 17] is a parallel adaptive mesh refinement framework aimed at solving (primarily hyperbolic) partial differential equations (PDEs) by a number of different numerical schemes. The framework supports 1D to 3D simulations, in a number of different geometries (Cartesian, cylindrical, spherical). MPI-AMRVAC is written in Fortran 90 and uses MPI for parallelization.

Table 12 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 28 and 29 (Sapphire Rapids with DDR), 30 and 31 (Sapphire Rapids with HBM), and 32 (Grace CPU).

All the I/O during calculation was turned of. We were not able to turn off the I/O to write result at the end of calculation. The HBM memory improved runtime by 8 % compared to DDR memory. But for the settings with highest energy savings the run with DDR has 2 % higher savings than the run with HBM.

In this configuration, the Sapphire Rapids node can save up to 14 % energy with 10 % runtime increase with use of DDR memory. Using HBM memory the energy saving is only 4 % with no runtime penalty. The Grace machine can achieve 26 % energy savings with runtime increase 3 % and a maximum of 36 % energy savings with an additional 9 % runtime increase.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Intel Sapphire Rapids with DDR	default	109.4 MFLOPs/W	39.0 kJ	43.4 s (100 %)
	CF 2.5 GHz, UCF 1.6 GHz	121.9 MFLOPs/W	34.9 kJ (-10 %)	44.6 s (103 %)
	CF 2.1 GHz, UCF 1.2 GHz	127.8 MFLOPs/W	33.3 kJ (-14 %)	47.9 s (110 %)
Intel Sapphire Rapids with HBM	default	120.6 MFLOPs/W	35.3 kJ	40.0 s (100 %)
	CF 2.3 GHz, UCF 2.2 GHz	125.5 MFLOPs/W	33.9 kJ (-4 %)	39.4 s (99 %)
	CF 2.3 GHz, UCF 2.2 GHz	125.5 MFLOPs/W	33.9 kJ (-4 %)	39.4 s (99 %)
Nvidia Grace CPU	default	218.1 MFLOPs/W	42.6 kJ	69 s (100 %)
	CF 2.7 GHz	292.9 MFLOPs/W	31.7 kJ (-26 %)	71 s (103 %)
	CF 2.1 GHz	340.7 MFLOPs/W	27.3 kJ (-36 %)	75 s (109 %)

Table 12: BHAC: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	43.4	43.5	43.4	43.6	43.7	43.9	44.5	45.1	46.3	48.7	54.8
2.6	43.4	43.6	43.8	43.5	43.8	44.0	44.5	45.3	46.3	48.8	54.7
2.5	43.7	43.7	43.9	43.7	44.2	44.3	44.6	45.3	46.7	49.0	54.9
2.3	44.4	44.6	44.3	44.4	44.8	44.8	45.3	46.0	47.1	49.7	55.7
2.1	45.1	45.1	45.1	45.4	45.6	45.7	46.0	46.7	47.9	50.8	56.3
1.9	47.8	46.3	46.1	46.3	46.3	46.8	47.1	48.3	49.0	51.3	57.2
1.7	49.1	47.5	47.3	47.8	47.9	48.2	48.4	49.0	50.2	52.7	58.1
1.5	50.7	49.1	49.1	49.8	49.7	49.7	49.9	51.2	52.4	54.5	60.1
1.3	52.8	51.4	52.2	52.5	52.4	52.8	53.4	53.4	54.2	57.0	62.2
1.1	58.8	55.3	55.0	55.3	55.7	55.7	57.4	58.9	58.8	61.0	65.8
0.9	63.9	59.7	61.0	62.6	61.0	62.0	63.0	63.6	66.1	67.9	71.0
0.8	67.2	68.1	65.7	65.5	68.2	67.9	65.3	68.4	75.4	70.7	76.7

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	100.0%	100.0%	100.3%	100.6%	101.2%	102.5%	103.8%	106.6%	112.1%	126.2%
2.6	100.0%	100.4%	100.8%	100.2%	100.9%	101.2%	102.5%	104.2%	106.7%	112.3%	125.8%
2.5	100.7%	100.5%	101.1%	100.7%	101.7%	102.0%	102.8%	104.3%	107.5%	112.7%	126.5%
2.3	102.3%	102.7%	102.1%	102.3%	103.2%	103.2%	104.3%	105.9%	108.5%	114.4%	128.3%
2.1	103.8%	103.7%	103.8%	104.5%	105.1%	105.3%	105.9%	107.5%	110.3%	116.9%	129.7%
1.9	110.0%	106.5%	106.2%	106.6%	106.6%	107.7%	108.4%	111.2%	112.8%	118.1%	131.6%
1.7	113.0%	109.4%	108.8%	110.1%	110.2%	111.0%	111.4%	112.9%	115.6%	121.2%	133.8%
1.5	116.7%	113.1%	113.1%	114.6%	114.5%	114.3%	114.9%	118.0%	120.7%	125.4%	138.5%
1.3	121.5%	118.4%	120.1%	120.9%	120.7%	121.5%	122.9%	123.1%	124.7%	131.3%	143.2%
1.1	135.4%	127.3%	126.6%	127.3%	128.2%	128.3%	132.1%	135.5%	135.3%	140.5%	151.4%
0.9	147.2%	137.4%	140.4%	144.0%	140.4%	142.7%	145.1%	146.5%	152.2%	156.4%	163.5%
0.8	154.7%	156.9%	151.2%	150.8%	157.1%	156.4%	150.4%	157.6%	173.7%	162.8%	176.6%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	39.0	39.0	38.9	38.7	37.4	36.0	35.2	34.7	34.5	35.4	38.7
2.6	39.0	39.1	39.1	38.7	37.4	36.0	35.2	34.7	34.6	35.5	38.5
2.5	39.2	39.1	39.2	38.7	37.3	35.8	34.9	34.4	34.4	35.2	38.2
2.3	39.6	39.8	39.4	38.4	36.8	35.3	34.4	33.9	33.7	34.6	37.5
2.1	39.9	40.0	39.8	38.3	36.6	35.1	34.1	33.5	33.3	34.4	37.0
1.9	39.8	40.8	40.4	38.3	36.5	35.2	34.2	33.8	33.4	34.0	36.7
1.7	40.3	41.6	40.8	38.8	37.0	35.6	34.5	33.8	33.5	34.2	36.6
1.5	40.8	42.6	41.5	39.7	37.7	35.9	34.9	34.5	34.2	34.7	37.2
1.3	41.4	43.9	43.2	41.0	38.9	37.3	36.4	35.3	34.7	35.5	37.6
1.1	43.9	46.1	44.6	42.3	40.4	38.6	38.2	37.8	36.6	37.1	38.9
0.9	45.5	48.6	48.1	46.4	43.1	41.7	40.9	39.9	40.1	40.2	41.0
0.8	46.5	54.3	51.0	48.0	47.2	44.9	41.9	42.3	44.8	41.3	43.5

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100%	100%	100%	99%	96%	92%	90%	89%	89%	91%	99%
2.6	100%	100%	100%	99%	96%	92%	90%	89%	89%	91%	99%
2.5	101%	100%	101%	99%	96%	92%	90%	88%	88%	90%	98%
2.3	102%	102%	101%	99%	94%	91%	88%	87%	86%	89%	96%
2.1	103%	103%	102%	98%	94%	90%	87%	86%	86%	88%	95%
1.9	102%	105%	104%	98%	94%	90%	88%	87%	86%	87%	94%
1.7	103%	107%	105%	100%	95%	91%	88%	87%	86%	88%	94%
1.5	105%	109%	107%	102%	97%	92%	89%	89%	88%	89%	95%
1.3	106%	113%	111%	105%	100%	96%	93%	91%	89%	91%	97%
1.1	113%	118%	114%	108%	104%	99%	98%	97%	94%	95%	100%
0.9	117%	125%	124%	119%	111%	107%	105%	102%	103%	103%	105%
0.8	119%	139%	131%	123%	121%	115%	107%	109%	115%	106%	112%

(b)

Figure 28: BHAC: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 29: BHAC: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	40.0	40.6	40.0	39.3	41.2	45.8	51.7	56.9	65.1	77.1	96.4
2.6	40.0	40.7	40.3	39.0	41.0	45.8	50.9	57.0	65.4	77.2	95.8
2.5	40.1	40.7	40.2	40.0	41.6	46.0	51.0	57.1	65.4	77.3	95.3
2.3	40.6	41.4	40.8	39.4	42.3	46.6	51.6	57.9	66.1	78.2	98.5
2.1	41.9	41.8	41.4	40.1	43.4	47.6	52.3	58.5	66.9	80.5	96.7
1.9	42.7	42.4	42.5	41.2	44.5	48.3	53.1	59.8	70.1	81.0	97.4
1.7	43.0	43.7	43.0	43.5	45.6	49.7	55.7	61.5	69.6	83.7	98.8
1.5	44.8	45.0	46.2	45.6	48.5	53.2	57.8	67.3	73.1	84.4	101.5
1.3	48.1	49.4	48.5	49.4	50.7	57.7	59.8	67.6	79.9	93.9	103.2
1.1	56.9	55.5	58.2	57.3	58.8	62.7	71.3	72.9	83.9	93.4	112.9
0.9	65.7	68.8	67.2	72.5	69.6	77.5	86.8	92.3	105.9	107.1	124.8
0.8	92.6	84.0	84.5	94.2	88.5	98.0	110.2	104.7	118.8	113.0	134.0

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	101.7%	100.2%	98.4%	103.1%	114.5%	129.4%	142.4%	163.0%	192.9%	241.3%
2.6	100.0%	101.8%	100.9%	97.6%	102.7%	114.6%	127.4%	142.7%	163.7%	193.3%	239.8%
2.5	100.4%	101.8%	100.6%	100.2%	104.0%	115.2%	127.6%	142.9%	163.6%	193.6%	238.6%
2.3	101.5%	103.7%	102.1%	98.6%	106.0%	116.7%	129.2%	144.8%	165.4%	195.7%	246.5%
2.1	104.9%	104.7%	103.6%	100.3%	108.7%	119.2%	130.9%	146.5%	167.4%	201.5%	242.0%
1.9	106.8%	106.2%	106.4%	103.0%	111.4%	121.0%	132.8%	149.7%	175.5%	202.6%	243.9%
1.7	107.7%	109.4%	107.5%	108.8%	114.1%	124.4%	139.5%	153.8%	174.3%	209.5%	247.2%
1.5	112.2%	112.8%	115.7%	114.1%	121.4%	133.1%	144.7%	168.5%	182.9%	211.3%	254.0%
1.3	120.4%	123.7%	121.4%	123.7%	126.9%	144.3%	149.7%	169.1%	200.0%	235.0%	258.2%
1.1	142.4%	138.9%	145.7%	143.3%	147.1%	157.0%	178.5%	182.4%	210.1%	233.7%	282.7%
0.9	164.3%	172.2%	168.3%	181.4%	174.2%	193.9%	217.3%	231.0%	265.0%	268.2%	312.3%
0.8	231.9%	210.3%	211.5%	235.9%	221.5%	245.2%	275.9%	262.0%	297.2%	282.9%	335.3%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	35.3	36.0	35.4	34.4	35.6	38.9	42.5	45.2	49.9	57.2	69.3
2.6	35.4	36.1	35.6	34.1	35.5	38.9	41.9	45.2	50.0	57.3	68.9
2.5	35.4	36.0	35.4	34.8	35.8	38.7	41.4	44.6	49.2	56.4	67.3
2.3	35.6	36.4	35.7	33.9	35.9	38.1	40.5	43.6	47.9	54.9	66.7
2.1	36.5	36.5	35.9	34.1	36.1	37.7	39.7	42.7	47.0	54.6	63.5
1.9	34.8	36.8	36.5	34.7	36.3	37.4	39.3	42.6	47.8	53.5	62.4
1.7	34.7	37.5	36.5	36.1	36.3	37.5	40.2	42.6	46.5	54.0	61.8
1.5	35.3	38.2	38.6	37.1	37.6	39.0	40.7	45.2	47.5	53.2	61.9
1.3	36.7	41.1	39.9	39.2	38.3	41.1	41.2	44.4	50.3	57.2	61.5
1.1	41.4	45.3	46.4	43.8	42.7	43.3	47.0	46.5	51.5	55.9	65.3
0.9	45.4	54.9	52.6	53.3	48.9	51.5	55.1	56.5	62.5	62.0	70.0
0.8	60.7	65.7	64.3	67.2	60.2	63.3	68.1	62.9	68.9	64.5	74.1

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100%	102%	100%	97%	101%	110%	120%	128%	141%	162%	196%
2.6	100%	102%	101%	97%	100%	110%	119%	128%	142%	162%	195%
2.5	100%	102%	100%	98%	101%	110%	117%	126%	139%	160%	191%
2.3	101%	103%	101%	96%	102%	108%	115%	123%	136%	155%	189%
2.1	103%	103%	101%	97%	102%	107%	112%	121%	133%	155%	180%
1.9	99%	104%	103%	98%	103%	106%	111%	120%	135%	152%	177%
1.7	98%	106%	103%	102%	103%	106%	114%	121%	132%	153%	175%
1.5	100%	108%	109%	105%	106%	110%	115%	128%	134%	151%	175%
1.3	104%	116%	113%	111%	108%	116%	117%	126%	142%	162%	174%
1.1	117%	128%	131%	124%	121%	123%	133%	132%	146%	158%	185%
0.9	128%	155%	149%	151%	138%	146%	156%	160%	177%	175%	198%
0.8	172%	186%	182%	190%	170%	179%	193%	178%	195%	183%	210%

(b)

Figure 30: BHAC: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 31: BHAC: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	69.0	100.0%	617.7	42.6	0.0%	218.1
3.30	71.0	102.9%	613.8	43.6	-2.3%	213.3
3.10	71.0	102.9%	528.8	37.5	11.9%	247.6
2.90	71.0	102.9%	481.4	34.2	19.8%	272.0
2.70	71.0	102.9%	447.0	31.7	25.5%	292.9
2.50	72.0	104.3%	417.1	30.0	29.5%	309.6
2.30	75.0	108.7%	388.0	29.1	31.7%	319.4
2.10	75.0	108.7%	363.7	27.3	36.0%	340.7
1.90	81.0	117.4%	345.1	28.0	34.4%	332.5
1.70	82.0	118.8%	337.3	27.7	35.1%	336.1
1.50	83.0	120.3%	328.0	27.2	36.1%	341.4
1.30	86.0	124.6%	320.4	27.6	35.3%	337.2
1.10	90.0	130.4%	314.1	28.3	33.7%	328.8
0.90	99.0	143.5%	300.6	29.8	30.2%	312.3

Figure 32: BHAC: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

3.6 FIL

FIL is a module of the Einstein Toolkit (ET). The ET is a publicly available evolution framework that is designed, maintained and extended to enable numerical relativity simulations. Modules of the ET are known as Thorns. FIL is used in conjunction with other Thorns of the ET to solve the General Relativistic Magneto-Hydrodynamic (GRMHD) equations in 3-dimensions. FIL can therefore simulate magnetic fields in dynamical space-times allowing for the study of binary neutron star and black hole-neutron star mergers. Further information on both FIL and the ET can be found here [18] and here [19] respectively.

Table 13 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 33 and 34 (Sapphire Rapids with DDR), 35 and 36 (Sapphire Rapids with HBM), and 37 (Grace CPU). On the Sapphire Rapids machine with DDR memory we can have up to 7 % energy savings with only 3 % runtime increase. Using HBM memory instead of DDR reduces runtime by 11 %. Also, for the run with HBM memory the change of uncore frequency impacts the runtime less significantly compared to the run with DDR memory. Without hindering the performance we can achieve 4 % energy saving just by lowering uncore frequency to 1.6 GHz. If the frequency is reduced even further to 1.2 GHz, we can have 6 % energy savings with a 4 % increase in runtime. The best energy savings can be reached with Grace CPU where the compute node energy consumption can be reduced by 20 %, with a run time penalty of 17 %. Without a negligible performance penalty (2 %), we can have 8 % energy savings.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Intel Sapphire Rapids with DDR	default	211.5 MFLOPs/W	111.0 kJ	123.4 s (100 %)
	CF 2.6 GHz, UCF 1.6 GHz	223.9 MFLOPs/W	104.9 kJ (-6 %)	125.7 s (102 %)
	CF default, UCF 1.4 GHz	228.1 MFLOPs/W	102.9 kJ (-7 %)	126.6 s (103 %)
Intel Sapphire Rapids with HBM	default	238.4 MFLOPs/W	98.5 kJ	110.3 s (100 %)
	CF 2.6 GHz, UCF 1.6 GHz	248.9 MFLOPs/W	94.3 kJ (-4 %)	110.1 s (100 %)
	CF 2.6 GHz, UCF 1.2 GHz	253.0 MFLOPs/W	92.8 kJ (-6 %)	114.2 s (104 %)
Nvidia Grace CPU	default	482.0 MFLOPs/W	45.9 kJ	98 s (100 %)
	CF 3.1 GHz	522.2 MFLOPs/W	42.4 kJ (-8 %)	100 s (102 %)
	CF 2.5 GHz	600.0 MFLOPs/W	36.9 kJ (-20 %)	115 s (117 %)

Table 13: FIL: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	123.4	121.8	124.3	123.4	123.8	123.2	126.7	126.6	131.9	136.1	145.7
2.6	124.4	121.9	123.9	121.8	123.4	123.2	125.7	127.1	131.1	134.9	145.7
2.5	128.0	126.9	126.8	124.9	126.8	126.7	129.2	129.9	135.1	137.7	149.1
2.3	132.4	131.2	133.2	131.9	134.6	134.1	137.1	138.0	141.1	146.2	155.8
2.1	140.4	140.9	140.8	140.9	143.4	143.2	145.7	146.4	153.7	156.6	163.4
1.9	159.9	152.4	151.7	153.6	152.9	157.0	159.0	157.1	162.4	167.1	179.2
1.7	182.3	170.9	168.3	173.7	175.3	172.7	175.9	177.2	177.5	178.4	189.9
1.5	192.8	185.3	186.0	186.2	186.0	187.3	188.6	190.5	193.8	198.3	205.4
1.3	218.5	210.8	209.6	210.3	209.3	210.1	214.5	214.6	216.6	222.2	231.6
1.1	266.3	242.9	242.2	242.1	246.2	245.7	256.4	250.9	254.5	260.1	265.3
0.9	306.2	294.6	293.1	302.8	297.2	296.7	295.6	301.3	303.3	306.9	317.6
0.8	348.6	329.7	330.4	334.9	328.2	333.0	335.1	336.4	341.8	344.3	367.7

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	98.7%	100.7%	100.0%	100.3%	99.8%	102.6%	102.6%	106.8%	110.3%	118.1%
2.6	100.8%	98.8%	100.4%	98.7%	100.0%	99.8%	101.8%	103.0%	106.2%	109.3%	118.1%
2.5	103.7%	102.8%	102.7%	101.2%	102.7%	102.6%	104.7%	105.3%	109.5%	111.6%	120.8%
2.3	107.3%	106.3%	108.0%	106.9%	109.0%	108.7%	111.1%	111.8%	114.3%	118.5%	126.2%
2.1	113.7%	114.2%	114.1%	114.2%	116.2%	116.1%	118.0%	118.7%	124.5%	126.9%	132.4%
1.9	129.5%	123.5%	122.9%	124.5%	123.9%	127.2%	128.8%	127.3%	131.6%	135.4%	145.2%
1.7	147.7%	138.5%	136.4%	140.7%	142.1%	140.0%	142.5%	143.6%	143.9%	144.6%	153.9%
1.5	156.2%	150.2%	150.7%	150.9%	150.8%	151.8%	152.8%	154.3%	157.0%	160.7%	166.4%
1.3	177.0%	170.9%	169.9%	170.4%	169.6%	170.2%	173.8%	173.9%	175.5%	180.1%	187.6%
1.1	215.8%	196.8%	196.3%	196.2%	199.5%	199.1%	207.8%	203.3%	206.2%	210.8%	215.0%
0.9	248.1%	238.7%	237.5%	245.3%	240.8%	240.4%	239.5%	244.1%	245.8%	248.7%	257.4%
0.8	282.5%	267.2%	267.8%	271.4%	266.0%	269.8%	271.5%	272.6%	277.0%	279.0%	297.9%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	111.0	109.7	111.9	110.9	110.3	106.0	105.6	102.9	104.2	105.3	109.8
2.6	112.0	109.8	111.6	109.5	110.1	106.0	104.9	103.2	103.7	104.4	109.8
2.5	115.1	114.1	114.0	111.9	110.7	106.6	105.5	103.2	104.4	104.2	109.8
2.3	119.0	117.9	119.5	115.6	112.7	108.3	107.4	105.1	104.6	106.0	110.1
2.1	125.8	126.3	124.9	119.0	115.7	111.3	109.8	107.3	109.4	109.1	111.1
1.9	117.3	135.0	131.0	126.1	120.1	118.3	116.2	111.7	112.2	112.8	117.9
1.7	128.5	146.7	140.9	137.6	132.9	125.9	124.0	121.5	118.8	117.0	121.4
1.5	129.5	155.5	151.8	144.0	137.6	133.1	129.6	127.1	125.7	125.9	127.4
1.3	139.9	171.9	166.2	158.0	150.2	144.6	142.6	138.7	136.2	136.5	139.0
1.1	161.3	192.5	186.7	176.7	171.0	163.6	164.8	156.7	154.6	154.4	154.2
0.9	177.4	226.9	219.1	213.6	200.0	191.4	184.4	181.9	178.0	176.2	178.3
0.8	196.6	250.1	243.2	232.9	217.4	211.2	205.2	199.8	197.3	194.0	202.7

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100%	99%	101%	100%	99%	95%	95%	93%	94%	95%	99%
2.6	101%	99%	101%	99%	99%	95%	94%	93%	93%	94%	99%
2.5	104%	103%	103%	101%	100%	96%	95%	93%	94%	94%	99%
2.3	107%	106%	108%	104%	102%	98%	97%	95%	94%	95%	99%
2.1	113%	114%	113%	107%	104%	100%	99%	97%	99%	98%	100%
1.9	106%	122%	118%	114%	108%	107%	105%	101%	101%	102%	106%
1.7	116%	132%	127%	124%	120%	113%	112%	109%	107%	105%	109%
1.5	117%	140%	137%	130%	124%	120%	117%	115%	113%	113%	115%
1.3	126%	155%	150%	142%	135%	130%	128%	125%	123%	123%	125%
1.1	145%	173%	168%	159%	154%	147%	148%	141%	139%	139%	139%
0.9	160%	204%	197%	192%	180%	172%	166%	164%	160%	159%	161%
0.8	177%	225%	219%	210%	196%	190%	185%	180%	178%	175%	183%

(b)

Figure 33: FIL: Impact of the static tuning to run time in the Sapphire Rapids system with the DDR memory. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 34: FIL: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the DDR memory. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	110.3	109.5	111.5	109.4	109.8	111.0	110.2	113.7	114.3	119.1	122.5
2.6	108.8	110.4	110.5	111.1	109.7	110.7	110.1	113.1	114.2	119.2	123.0
2.5	113.1	111.6	111.5	112.8	112.0	114.0	116.6	117.4	117.3	122.4	126.5
2.3	121.0	119.7	120.7	119.1	120.8	120.4	123.9	123.1	129.0	128.5	139.1
2.1	128.2	131.2	137.2	131.2	133.4	136.7	135.3	137.8	139.7	142.4	148.1
1.9	153.8	144.4	146.2	140.3	143.9	142.4	146.0	148.2	146.1	152.8	154.5
1.7	168.0	156.0	155.4	157.6	158.6	157.1	160.0	159.7	163.5	166.9	168.2
1.5	183.6	177.8	173.4	178.6	184.8	178.5	179.6	180.9	183.8	182.1	193.9
1.3	211.4	202.8	203.0	199.6	204.0	207.2	204.9	205.9	209.2	213.2	225.5
1.1	245.9	238.8	238.3	238.3	238.7	242.4	239.9	241.4	252.0	246.3	251.2
0.9	306.2	292.0	290.7	293.6	291.1	293.1	293.8	299.1	297.4	302.3	307.2
0.8	338.0	332.1	330.7	333.4	338.4	335.1	334.2	332.3	339.7	349.0	351.8

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100.0%	99.3%	101.1%	99.2%	99.6%	100.7%	99.9%	103.1%	103.7%	108.0%	111.1%
2.6	98.6%	100.1%	100.2%	100.7%	99.5%	100.4%	99.9%	102.6%	103.5%	108.1%	111.5%
2.5	102.6%	101.2%	101.1%	102.3%	101.5%	103.4%	105.8%	106.5%	106.4%	110.9%	114.7%
2.3	109.7%	108.5%	109.4%	108.0%	109.6%	109.2%	112.3%	111.7%	117.0%	116.5%	126.1%
2.1	116.3%	119.0%	124.4%	118.9%	121.0%	123.9%	122.7%	125.0%	126.6%	129.1%	134.3%
1.9	139.5%	130.9%	132.6%	127.2%	130.5%	129.1%	132.4%	134.4%	132.4%	138.5%	140.0%
1.7	152.3%	141.5%	140.9%	142.9%	143.8%	142.5%	145.0%	144.8%	148.2%	151.3%	152.5%
1.5	166.4%	161.2%	157.3%	161.9%	167.5%	161.9%	162.8%	164.1%	166.7%	165.1%	175.8%
1.3	191.7%	183.9%	184.1%	181.0%	185.0%	187.9%	185.8%	186.7%	189.7%	193.3%	204.5%
1.1	223.0%	216.5%	216.1%	216.1%	216.4%	219.8%	217.5%	218.9%	228.5%	223.3%	227.7%
0.9	277.6%	264.8%	263.6%	266.2%	263.9%	265.8%	266.4%	271.2%	269.6%	274.1%	278.6%
0.8	306.5%	301.1%	299.8%	302.3%	306.9%	303.8%	303.0%	301.3%	308.0%	316.5%	318.9%

(b)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	98.5	98.0	99.7	97.7	97.9	97.8	94.4	94.6	93.0	94.8	95.3
2.6	97.2	98.6	98.7	99.1	97.7	97.4	94.3	94.2	92.8	94.7	95.5
2.5	101.0	99.7	99.5	100.6	99.6	98.1	97.4	95.6	93.3	95.0	96.0
2.3	108.0	106.9	107.6	105.9	103.3	99.2	98.9	95.8	97.5	95.4	100.4
2.1	114.4	117.0	121.6	112.4	108.9	107.2	103.2	102.3	101.0	100.5	102.4
1.9	108.8	127.9	127.1	116.9	114.3	109.0	108.2	106.7	102.8	104.8	103.9
1.7	115.0	136.1	132.0	126.9	122.1	116.4	114.7	111.4	110.9	110.7	109.5
1.5	120.4	150.3	143.0	139.3	137.3	127.8	124.6	121.8	120.4	117.1	121.6
1.3	132.6	166.5	162.2	151.1	147.3	143.4	137.3	134.0	132.5	132.1	136.4
1.1	147.6	190.1	184.4	174.6	166.8	162.2	155.5	151.7	153.7	147.5	147.1
0.9	175.4	225.7	218.1	208.2	196.6	189.7	183.9	181.3	175.3	174.3	173.3
0.8	189.6	252.6	243.8	232.8	224.5	213.1	205.2	198.2	196.7	197.0	195.0

(a)

CF/UnCF [GHz]	0	2.5	2.4	2.2	2	1.8	1.6	1.4	1.2	1	0.8
0	100%	99%	101%	99%	99%	99%	96%	96%	94%	96%	97%
2.6	99%	100%	100%	101%	99%	99%	96%	96%	94%	96%	97%
2.5	103%	101%	101%	102%	101%	100%	99%	97%	95%	96%	97%
2.3	110%	108%	109%	108%	105%	101%	100%	97%	99%	97%	102%
2.1	116%	119%	123%	114%	111%	109%	105%	104%	103%	102%	104%
1.9	110%	130%	129%	119%	116%	111%	110%	108%	104%	106%	106%
1.7	117%	138%	134%	129%	124%	118%	116%	113%	113%	112%	111%
1.5	122%	153%	145%	141%	139%	130%	126%	124%	122%	119%	123%
1.3	135%	169%	165%	153%	150%	146%	139%	136%	135%	134%	139%
1.1	150%	193%	187%	177%	169%	165%	158%	154%	156%	150%	149%
0.9	178%	229%	221%	211%	200%	193%	187%	184%	178%	177%	176%
0.8	193%	256%	248%	236%	228%	216%	208%	201%	200%	200%	198%

(b)

Figure 35: FIL: Impact of the static tuning to run time in the Sapphire Rapids system with the HBM memory. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 36: FIL: Impact of the static tuning to node energy consumption [kJ] in the Sapphire Rapids system with the HBM memory. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0.00	98.0	100.0%	468.7	45.9	0.0%	482.0
3.30	100.0	102.0%	466.7	46.7	-1.6%	474.4
3.10	100.0	102.0%	424.4	42.4	7.6%	522.2
2.90	104.0	106.1%	381.9	39.7	13.5%	557.8
2.70	109.0	111.2%	347.2	37.8	17.6%	584.8
2.50	115.0	117.3%	321.0	36.9	19.6%	600.0
2.30	125.0	127.6%	295.2	36.9	19.7%	599.7
2.10	131.0	133.7%	279.2	36.6	20.4%	605.5
1.90	145.0	148.0%	262.5	38.1	17.1%	581.2
1.70	155.0	158.2%	256.8	39.8	13.3%	556.3
1.50	164.0	167.3%	254.2	41.7	9.2%	531.0
1.30	182.0	185.7%	245.8	44.7	2.6%	494.7
1.10	208.0	212.2%	241.1	50.2	-9.2%	441.7
0.90	247.0	252.0%	233.2	57.6	-25.4%	384.1

Figure 37: FIL: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

3.7 ChaNGa

ChaNGa [20][21][22] is an N-body and smoothed particle magneto-hydrodynamics (SPMHD) code which is used to study a wide array of astrophysical systems. While the gravity and SPMHD algorithms are based on the gasoline [23] and pkdgrav [24] codes, the unique feature of ChaNGa is its implementation of the Charm++ framework, which enables highly efficient parallel scaling. Charm++ employs overdecomposition to achieve this. That is to divide the work into many more pieces (chares/tree pieces) than you have processors and let the Charm++ runtime system load balance by appropriately assigning pieces to real processors.

Table 14 summarizes the measurements of energy consumption and energy efficiency. Detailed information on the individual architecture can be seen in Figures 39 and 40 (Sapphire Rapids with DDR), 41 and 42 (Sapphire Rapids with HBM), and 43 (Grace CPU).

ChaNGa shows the best energy efficiency when executed using Nvidia Grace CPU (ignoring Nvidia A100, since we do not have FLOPs measurement for GPU execution), which is also the best energy efficiency from all SPACE applications. In the default Grace CPU frequency configuration, the test reached 1.24 GFLOPs/W, while reducing the frequency to 3.1 GHz extended runtime about 2.4 % but brings 16.2 % node energy savings (1.48 GFLOPs/W).

ChaNGa execution when using Intel Sapphire Rapids shows only 0.65 GFLOPs/W when using DDR, but HBM memory reduces the application runtime by about 34 % thus improving energy efficiency by about 0.22 GFLOPs/W. ChaNGa requires high CPU core frequency, but reducing uncore frequency leads to 12-20 % energy savings with none or minimal performance impact. When using DDR memory, the test case runtime showed fluctuations and is not strictly determined by the CPU frequencies settings. Therefore, measurements from the Sapphire Rapids CPU were evaluated from nine repetitions, and the mean runtime is presented. We have also manually removed outliers from the measurements to highlight the trends when core and uncore frequencies are tuned. At best we are able to save up to 20% of energy compared to 12% for runs in HBM memory. But the most important result is that running ChaNGa in HBM memory is 1.35× more energy efficient due to reduced runtime.

The GPU-accelerated execution shows no energy savings when scaling the frequency of the streaming multiprocessors. This implies that the GPUs executes a strongly compute-bound workload. However, we also see a limited GPU utilization as GPUs only consume 95 W on average when running ChaNGa while the maximum consumption is (TDP) is 400 W.

In the future, we need to investigate this behavior more thoroughly to define what type of workload exhibits this behavior. At this point, we attribute this behavior to the Charm++ runtime system and its property to launch a large number of relatively small tasks on GPU. Also, with consultation with code developers and analyzing performance traces, we identified that GPUs are active for approximately 50% of the runtime.

Hardware	Configuration	Energy efficiency	Node energy consumption	Runtime
Nvidia A100	default	-	1447.4 kJ	814.1 s (100 %)
	1290 MHz	-	1447.4 kJ (0 %)	817.8 s (100.5 %)
Intel Sapphire Rapids with DDR	default	648.4 MFLOPs/W	440.9 kJ	560 s (100 %)
	CF default, UCF 1.0 GHz	779.6 MFLOPs/W	379.7 kJ (14 %)	577 s (103.1 %)
	CF 2.3 GHz, UCF 0.8 GHz	825.5 MFLOPs/W	346.3 kJ (20 %)	588 s (107.6 %)
Intel Sapphire Rapids with HBM	default	866.4 MFLOPs/W	330.2 kJ	369 s (100 %)
	CF default, UCF 1.0 GHz	1018.3 MFLOPs/W	280.9 kJ (12.0 %)	376 s (101.9 %)
	CF default, UCF 0.8 GHz	1026.5 MFLOPs/W	278.7 kJ (12.7 %)	381 s (103.3 %)
Nvidia Grace CPU	default	1239.8 MFLOPs/W	226.8 kJ	369 s (100 %)
	CF 3.1 GHz	1478.9 MFLOPs/W	190.1 kJ (16.2 %)	378 s (102.4 %)
	CF 2.3 GHz	1774.7 MFLOPs/W	158.4 kJ (30.1 %)	499 s (135.2 %)

Table 14: ChaNGa: Summary of the energy consumption and static frequency tuning measurements on all HW platforms

Set GPU frequency [MHz]	Runtime [s]	Runtime extension	Average power consumption of the GPU [W]	Average power consumption of the CPU [W]	CPUs + GPUs energy consumption [kJ]	Energy savings for GPUs + CPUs	Node energy consumption [kJ]	Compute node energy savings
1410	814,1	100,0%	94,6	210,6	959,0	0,0%	1447,4	0,0%
1350	813,8	100,0%	94,6	210,5	958,7	0,0%	1447,0	0,0%
1290	817,8	100,5%	93,7	209,9	956,3	0,3%	1447,0	0,0%
1230	824,7	101,3%	93,1	210,6	961,3	-0,2%	1456,1	-0,6%
1170	841,3	103,3%	91,7	210,6	971,4	-1,3%	1476,2	-2,0%
1110	846,4	104,0%	90,3	210,6	968,2	-1,0%	1476,1	-2,0%
1050	861,0	105,8%	92,1	210,6	997,2	-4,0%	1513,8	-4,6%
990	886,2	108,9%	86,5	210,5	986,5	-2,9%	1518,2	-4,9%
930	904,8	111,1%	84,4	210,4	991,6	-3,4%	1534,5	-6,0%
870	913,4	112,2%	84,5	210,8	1002,3	-4,5%	1550,3	-7,1%
810	944,8	116,1%	81,4	210,6	1013,3	-5,7%	1580,2	-9,2%
750	958,9	117,8%	79,6	210,3	1014,2	-5,8%	1589,5	-9,8%
690	996,8	122,4%	76,8	210,4	1031,9	-7,6%	1630,0	-12,6%

Figure 38: ChaNGa: The results of energy consumption and static frequency tuning measurements in the Nvidia A100 system (Karolina GPU).

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	560,0	566,0	566,3	567,3	569,1	568,3	566,1	567,3	572,1	577,2	592,0
2,6	560,7	568,6	567,4	577,7	568,2	575,3	580,7	590,3	600,7	603,8	603,0
2,5	585,7	576,6	587,3	590,6	587,1	576,7	590,0	611,4	617,6	612,7	601,7
2,3	629,3	616,2	616,3	613,2	609,6	603,0	607,7	622,8	632,8	628,4	602,7
2,1	690,0	658,1	654,7	638,1	638,2	634,6	637,9	652,6	673,3	676,6	649,3
1,9	659,0	698,1	693,0	683,7	694,8	705,0	698,0	696,7	702,7	707,9	685,3
1,7	739,0	752,3	750,7	749,2	760,0	765,6	754,1	745,2	748,6	752,6	747,7
1,5	810,0	827,9	808,7	822,2	831,4	841,3	826,7	814,3	812,7	813,4	813,0
1,3	987,0	925,4	908,4	935,0	924,4	925,3	914,3	917,7	924,2	920,3	930,7
1,1	1048,0	1032,7	1025,3	1058,0	1025,0	1022,7	1015,7	1048,0	1054,3	1041,3	1049,0

(a)

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	100,0%	101,1%	101,1%	101,3%	101,6%	101,5%	101,1%	101,3%	102,2%	103,1%	105,7%
2,6	100,1%	101,5%	101,3%	103,2%	101,5%	102,7%	103,7%	105,4%	107,3%	107,8%	107,7%
2,5	104,6%	103,0%	104,9%	105,5%	104,8%	103,0%	105,4%	109,2%	110,3%	109,4%	107,4%
2,3	112,4%	110,0%	110,1%	109,5%	108,8%	107,7%	108,5%	111,2%	113,0%	112,2%	107,6%
2,1	123,2%	117,5%	116,9%	113,9%	114,0%	113,3%	113,9%	116,5%	120,2%	120,8%	116,0%
1,9	117,7%	124,7%	123,8%	122,1%	124,1%	125,9%	124,6%	124,4%	125,5%	126,4%	122,4%
1,7	132,0%	134,3%	134,0%	133,8%	135,7%	136,7%	134,7%	133,1%	133,7%	134,4%	133,5%
1,5	144,6%	147,8%	144,4%	146,8%	148,5%	150,2%	147,6%	145,4%	145,1%	145,3%	145,2%
1,3	176,3%	165,3%	162,2%	167,0%	165,1%	165,2%	163,3%	163,9%	165,0%	164,3%	166,2%
1,1	187,1%	184,4%	183,1%	188,9%	183,0%	182,6%	181,4%	187,1%	188,3%	186,0%	187,3%

(b)

Figure 39: ChaNGa: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	440,9	448,1	440,2	428,0	402,4	390,7	392,9	396,1	390,9	379,0	369,5
2,6	442,3	444,0	446,3	441,9	422,9	405,7	399,7	396,4	386,3	376,1	369,9
2,5	457,8	456,3	457,6	446,3	428,2	405,6	395,5	391,5	383,3	373,4	362,2
2,3	488,0	483,6	476,0	455,9	433,5	410,6	396,4	389,2	382,4	371,9	353,9
2,1	531,9	492,8	495,8	464,1	440,9	418,6	404,3	396,4	393,5	384,5	365,3
1,9	363,4	497,8	514,7	483,3	462,7	446,3	426,6	410,6	400,5	391,8	374,1
1,7	390,0	505,2	542,6	513,5	491,2	471,7	447,6	427,3	415,4	405,4	393,8
1,5	412,9	538,4	572,5	550,0	523,9	503,5	475,2	452,1	437,5	425,9	414,8
1,3	480,8	584,9	625,5	606,4	567,2	539,3	510,2	492,4	479,9	464,7	456,6
1,1	498,4	639,0	689,9	668,0	613,9	581,1	552,1	544,9	530,0	509,0	498,5

(a)

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	100 %	102 %	100 %	97 %	91 %	89 %	89 %	90 %	89 %	86 %	84 %
2,6	100 %	101 %	101 %	100 %	96 %	92 %	91 %	90 %	88 %	85 %	84 %
2,5	104 %	103 %	104 %	101 %	97 %	92 %	90 %	89 %	87 %	85 %	82 %
2,3	111 %	110 %	108 %	103 %	98 %	93 %	90 %	88 %	87 %	84 %	80 %
2,1	121 %	112 %	112 %	105 %	100 %	95 %	92 %	90 %	89 %	87 %	83 %
1,9	82 %	113 %	117 %	110 %	105 %	101 %	97 %	93 %	91 %	89 %	85 %
1,7	88 %	115 %	123 %	116 %	111 %	107 %	102 %	97 %	94 %	92 %	89 %
1,5	94 %	122 %	130 %	125 %	119 %	114 %	108 %	103 %	99 %	97 %	94 %
1,3	109 %	133 %	142 %	138 %	129 %	122 %	116 %	112 %	109 %	105 %	104 %
1,1	113 %	145 %	156 %	152 %	139 %	132 %	125 %	124 %	120 %	115 %	113 %

(b)

Figure 40: ChaNGa: Impact of the static tuning to **node energy consumption [kJ]** in the **Sapphire Rapids system with the DDR memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	369,0	372,0	373,0	370,0	370,0	371,0	373,0	373,0	375,0	376,0	381,0
2,6	370,0	367,0	369,0	368,0	369,0	369,0	372,0	377,0	374,0	378,0	387,0
2,5	383,0	381,0	381,0	381,0	382,0	386,0	384,0	388,0	393,0	391,0	397,0
2,3	413,0	412,0	412,0	417,0	412,0	414,0	420,0	418,0	418,0	426,0	427,0
2,1	451,0	448,0	447,0	449,0	453,0	451,0	458,0	460,0	461,0	459,0	469,0
1,9	508,0	493,0	498,0	494,0	493,0	501,0	496,0	502,0	501,0	507,0	510,0
1,7	564,0	561,0	550,0	551,0	554,0	553,0	563,0	554,0	556,0	558,0	563,0
1,5	633,0	620,0	619,0	625,0	620,0	620,0	626,0	623,0	627,0	631,0	635,0
1,3	726,0	710,0	712,0	712,0	716,0	718,0	716,0	717,0	718,0	724,0	726,0
1,1	854,0	838,0	839,0	837,0	837,0	841,0	850,0	848,0	845,0	850,0	862,0
0,9	1046,0	1026,0	1021,0	1022,0	1021,0	1029,0	1030,0	1035,0	1026,0	1032,0	1036,0
0,8	1161,0	1149,0	1152,0	1161,0	1148,0	1149,0	1152,0	1163,0	1169,0	1172,0	1165,0

(a)

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	100,0%	100,8%	101,1%	100,3%	100,3%	100,5%	101,1%	101,1%	101,6%	101,9%	103,3%
2,6	100,3%	99,5%	100,0%	99,7%	100,0%	100,0%	100,8%	102,2%	101,4%	102,4%	104,9%
2,5	103,8%	103,3%	103,3%	103,3%	103,5%	104,6%	104,1%	105,1%	106,5%	106,0%	107,6%
2,3	111,9%	111,7%	111,7%	113,0%	111,7%	112,2%	113,8%	113,3%	113,3%	115,4%	115,7%
2,1	122,2%	121,4%	121,1%	121,7%	122,8%	122,2%	124,1%	124,7%	124,9%	124,4%	127,1%
1,9	137,7%	133,6%	135,0%	133,9%	133,6%	135,8%	134,4%	136,0%	135,8%	137,4%	138,2%
1,7	152,8%	152,0%	149,1%	149,3%	150,1%	149,9%	152,6%	150,1%	150,7%	151,2%	152,6%
1,5	171,5%	168,0%	167,8%	169,4%	168,0%	169,6%	168,8%	169,9%	169,9%	171,0%	172,1%
1,3	196,7%	192,4%	193,0%	193,0%	194,0%	194,6%	194,0%	194,3%	194,6%	196,2%	196,7%
1,1	231,4%	227,1%	227,4%	226,8%	226,8%	227,9%	230,4%	229,8%	229,0%	230,4%	233,6%
0,9	283,5%	278,0%	276,7%	277,0%	276,7%	278,9%	279,1%	280,5%	278,0%	279,7%	280,8%
0,8	314,6%	311,4%	312,2%	314,6%	311,1%	311,4%	312,2%	315,2%	316,8%	317,6%	315,7%

(b)

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	319,1	327,6	328,6	324,5	322,4	311,0	301,9	293,6	285,4	280,9	278,7
2,6	322,2	323,2	323,2	322,2	319,7	308,9	300,1	295,2	285,6	281,2	281,5
2,5	330,9	334,6	334,5	332,9	322,8	312,0	300,5	293,6	288,7	282,4	279,9
2,3	360,9	360,8	360,4	349,3	329,3	315,2	308,6	298,8	290,3	288,2	282,7
2,1	380,4	390,2	378,3	358,1	341,9	325,2	318,7	309,8	301,4	293,7	292,9
1,9	303,6	412,8	403,3	377,0	357,2	345,5	330,4	323,5	312,9	308,8	303,0
1,7	321,7	451,7	429,1	403,7	385,3	367,0	358,9	341,9	332,4	326,1	321,1
1,5	344,8	482,0	465,9	441,9	415,1	396,9	383,6	368,6	359,2	352,1	345,7
1,3	378,0	534,5	518,4	486,4	461,8	440,4	421,3	406,2	393,6	385,6	377,5
1,1	423,3	609,1	588,5	551,9	520,6	496,0	479,5	460,8	443,1	432,4	427,2
0,9	495,2	721,5	693,7	650,5	611,0	582,6	557,0	538,3	514,5	503,1	489,5
0,8	537,6	795,3	769,2	725,4	673,0	637,1	610,4	591,2	571,5	556,2	538,6

(a)

CF/UnCF [GHz]	0	2,5	2,4	2,2	2	1,8	1,6	1,4	1,2	1	0,8
0	100,0 %	102,7 %	103,0 %	101,7 %	101,1 %	97,5 %	94,6 %	92,0 %	89,5 %	88,0 %	87,3 %
2,6	101,0 %	101,3 %	101,3 %	101,0 %	100,2 %	96,8 %	94,1 %	92,5 %	89,5 %	88,1 %	88,2 %
2,5	103,7 %	104,9 %	104,9 %	104,3 %	101,2 %	97,8 %	94,2 %	92,0 %	90,5 %	88,5 %	87,7 %
2,3	113,1 %	113,1 %	113,0 %	109,5 %	103,2 %	98,8 %	96,7 %	93,6 %	91,0 %	90,3 %	88,6 %
2,1	119,2 %	122,3 %	118,6 %	112,2 %	107,2 %	101,9 %	99,9 %	97,1 %	94,5 %	92,1 %	91,8 %
1,9	95,2 %	129,4 %	126,4 %	118,2 %	112,0 %	108,3 %	103,6 %	101,4 %	98,1 %	96,8 %	95,0 %
1,7	100,8 %	141,6 %	134,5 %	126,5 %	120,8 %	115,0 %	112,5 %	107,2 %	104,2 %	102,2 %	100,7 %
1,5	108,1 %	151,1 %	146,0 %	138,5 %	130,1 %	124,4 %	120,2 %	115,5 %	112,6 %	110,3 %	108,4 %
1,3	118,5 %	167,5 %	162,5 %	152,4 %	144,7 %	138,0 %	132,1 %	127,3 %	123,4 %	120,8 %	118,3 %
1,1	132,7 %	190,9 %	184,5 %	173,0 %	163,2 %	155,5 %	150,3 %	144,4 %	138,9 %	135,5 %	133,9 %
0,9	155,2 %	226,1 %	217,4 %	203,9 %	191,5 %	182,6 %	174,6 %	168,7 %	161,3 %	157,7 %	153,4 %
0,8	168,5 %	249,3 %	241,1 %	227,4 %	210,9 %	199,7 %	191,3 %	185,3 %	179,1 %	174,3 %	168,8 %

(b)

Figure 41: ChaNGa: Impact of the static tuning to **run time** in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute times [s] while (b) shows differences from the default 0/0 run.

Figure 42: ChaNGa: Impact of the static tuning to **node energy consumption** [kJ] in the **Sapphire Rapids system with the HBM memory**. The figure (a) shows absolute energy consumption [kJ] while (b) shows differences from the default 0/0 run.

Set CPU frequency [GHz]	Runtime [s]	Runtime extension	Average power consumption of the node [W]	Node energy consumption [kJ]	Compute node energy savings	MFLOPS/W
0,00	369,0	100,0%	614,5	226,8	0,0%	1239,8
3,30	374,0	101,4%	597,7	223,6	1,4%	1257,6
3,10	378,0	102,4%	502,9	190,1	16,2%	1478,9
2,90	405,0	109,8%	440,9	178,6	21,3%	1574,5
2,70	429,0	116,3%	391,7	168,1	25,9%	1673,0
2,50	493,0	133,6%	343,6	169,4	25,3%	1659,7
2,30	499,0	135,2%	317,5	158,4	30,1%	1774,7
2,10	545,0	147,7%	290,6	158,4	30,1%	1774,9
1,90	620,0	168,0%	267,1	165,6	27,0%	1697,5
1,70	687,0	186,2%	257,5	176,9	22,0%	1589,4
1,50	743,0	201,4%	251,2	186,6	17,7%	1506,5
1,30	832,0	225,5%	244,5	203,4	10,3%	1382,2
1,10	984,0	266,7%	235,9	232,2	-2,4%	1211,1
0,90	1201,0	325,5%	227,2	272,9	-20,3%	1030,2

Figure 43: ChaNGa: The results of energy consumption and static frequency tuning measurements in the Grace CPU system.

4 Conclusion

This deliverable builds on the work performed and reported in the deliverable D2.2. The first report is on the evaluation of the energy efficiency of SPACE codes. This deliverable is focused on the evaluation of energy efficiency for different hardware configurations of 3 new hardware platforms: (i) NVidia A100 GPU accelerated production-grade server in the GPU partition of the IT4Innovations Karolina supercomputer; (ii) modern x86 Intel Sapphire Rapids generation processor with both DDR and HBM memory and (iii) new ARM-based GRACE CPU by NVidia.

The platforms were selected based on work carried out in SPACE CoE, as porting to GPU accelerators is undertaken by the majority of the SPACE codes, and Sapphire Rapids and NVidia GRACE are also co-design platforms. This section compares the results of all codes and all platforms that individual codes support.

Table 15 shows the most efficient platform in terms of performance per watt for the best configuration that does not allow runtime slowdown for less than 5 % from the original runtime. We do not have the FLOPS measurements for the GPU accelerator; therefore, we only compare results for CPUs. Data show that the most energy-efficient CPU platform for all codes is the NVidia GRACE CPU.

Code energy efficiency	NVidia GRACE CPU [MFLOPS/W]	Intel Sapphire Rapids with DDR [MFLOPS/W]	Intel Sapphire Rapids with HBM [MFLOPS/W]
Pluto	805.4	264.0	309.0
OpenGADGET	716.2	138.2	149.2
iPIC3D	791.3	238.2	321.4
RAMSES	854.9	399.7	417.8
BHAC	292.9	121.9	125.5
FIL	522.2	223.9	248.9
ChaNGa	1478.9	779.6	1018.3

Table 15: The energy efficiency of the CPU platforms. Results are shown for hardware configurations generating very small runtime extensions (less than 105% of default).

Table 16 also shows the energy efficiency in terms of FLOPS per Watt, but here we focus on configuration, in which we achieve the highest possible energy savings and, in most cases, also a significant extension of the runtime. In this case, also, the Grace CPU is the most efficient CPU platform.

Code energy efficiency	NVidia GRACE CPU [MFLOPS/W]	Intel Sapphire Rapids with DDR [MFLOPS/W]	Intel Sapphire Rapids with HBM [MFLOPS/W]
Pluto	964.4	266.2	314.1
OpenGADGET	940.7	138.2	151.7
iPIC3D	964.4	242.6	321.4
RAMSES	968.6	401.6	418.5
BHAC	340.7	127.8	125.5
FIL	600.0	228.1	253.0
ChaNGa	1774.7	825.5	1026.5

Table 16: The energy efficiency of the CPU platforms. Results are shown for hardware configurations generating maximum energy savings and noticeable runtime extension.

Table 17 shows achievable energy savings enabled by adjusting the hardware parameters of all platforms tested. We can see energy savings for the NVidia A100 GPU reaching up to 9% for Pluto when the runtime is extended by 13% and 6% with a very small impact on performance. OpenGADGET can save 7% and iPIC3D between 3% and 5%, depending on the runtime extension. In general, the Nvidia GPUs do not achieve as large savings as we can see for CPUs if streaming multiprocessor frequency is tuned. This means that their internal power management systems perform more efficiently. We also observe that in all cases, the GPUs consume significantly lower power than its 400W TDP: Pluto up to 185W, OpenGADGET up to 74W, iPIC3D up to 159W, and ChaNGa up to 95W.

On the Sapphire Rapids platform, we are able to save between 6% and 14% when keeping the runtime and between 7% to 20% if runtime extension is allowed when DDR memory is used for the data and application. ChaNGa and BHAC achieved the largest savings. If we use the HBM memory, the code, in general, runs faster and achieves higher performance per watt, see Table 15, but at the same time, faster memory also requires higher CPU core and uncore frequencies to utilize the higher memory bandwidth. Therefore, we see lower savings for some codes (Pluto, iPIC3D, BHAC, FIL and ChaNGa). For RAMSES the savings are similar, and higher savings are measured only for OpenGADGET (11%).

The highest energy savings can be achieved on the NVidia GRACE ARM-based processor. We can see that, for most of the codes, we can save around 30% of energy if runtime can be extended. At the same time, if one limits the runtime extension to less than 3%, energy savings can be around 20% as it is in the case of the Pluto, RAMSES and BHAC codes.

Our analysis shows that the new HPC ARM-based platform delivers superior energy efficiency for all codes compared to x86-based Sapphire Rapids. Moreover, we also see that by tuning the CPU core frequency, we can further improve the energy efficiency by up to 30% or between 8% and 26% without affecting the runtime.

	Pluto	Open GADGET	iPIC3D	RAMSES	BHAC	FIL	ChaNGa
Nvidia A100	-6% / 103% -9% / 113%	-7% / 102% -7% / 102%	-3% / 104% -5% / 111%	- - - -	- - - -	- - - -	- - - -
SPR w. DDR	-9% / 102% -10% / 106%	-7% / 102% -7% / 102%	-7% / 102% -9% / 108%	-6% / 102% -7% / 104%	-10% / 103% -14% / 110%	-6% / 102% -7% / 103%	-14% / 103% -20% / 107%
SPR w. HBM	-4% / 101% -6% / 105%	-9% / 94% -11% / 98%	-7% / 101% -7% / 101%	-7% / 102% -8% / 104%	-4% / 99% -4% / 99%	-4% / 100% -6% / 104%	-12% / 102% -13% / 103%
Grace CPU	-22% / 101% -35% / 122%	-13% / 103% -33% / 128%	-9% / 103% -29% / 126%	-19% / 101% -28% / 137%	-26% / 103% -36% / 109%	-8% / 102% -20% / 117%	-16% / 102% -30% / 135%
Cascade Lake D2.2*	-6% / 102% -12% / 126%	-9% / 103% -12% / 130%	-6% / 101% -13% / 115%	-7% / 102% -11% / 123%	-5% / 102% -11% / 118%	-3% / 102% -13% / 127%	-30% / 102% -36% / 110%

Table 17: **Energy consumption reduction [kJ] / Runtime [s]**; 1st row - results with no runtime extension (runtime close to 100% of default); 2nd row - results for maximum energy savings. *Cascade Lake - results from deliverable D2.2 are presented for comparison.

4.1 Recommendations outside of the scope of the project

We see that some energy savings can be achieved across different processing architectures, and in many cases, the runtime is not extended, which we consider the most optimal scenario. However, we still present configurations that can achieve even more significant energy savings at the cost of extended runtime because there are scenarios where this is useful: (a) a supercomputer is not fully utilized, so running the code for a longer period of time does not waste its cycles and, at the same reduces its operational costs, (b) the data center budget in terms of power consumption reaches the limit, and to be able to submit new jobs, they need to be under a limited power budget. The scheduler can apply a power cap, but our approach can identify optimal configurations to improve both energy consumption as well runtime.

When extending the application runtime, it is necessary to be aware of additional power consumption related to the execution. Thus, although we only control the power consumption of CPUs and GPUs, we present overall node energy consumption.

The results presented in this deliverable are not usable only for the A&C community but, in general, for other parallel codes running on supercomputers around the world. If one considers that a machine consumes 20 MW, energy savings of 5-10 % are in orders in megawatts, which significantly reduces operational costs.

The energy efficiency analysis of the parallel codes is done by the open-source MERIC tool developed at IT4I@VSB. Similar services, as presented in this deliverable, are also offered under the POP3 Center of Excellence to the general public. We want to highlight that for the static tuning approach, there is no need to change the code as the optimal settings, and it can be applied job-wide using a job scheduler.

4.2 Future work

This report fulfills the second deliverable, showing the results of the static tuning. By now, we have evaluated five different platforms, including NVidia A100 GPUs.

In the following two years of the project, we will take one step forward and focus on the advanced methodology of dynamic tuning. Dynamic tuning means that we find optimal configurations for different regions of the application. Then, as an application moves from one region to another, the MERIC runtime will change the hardware settings to the optimal ones for that given region. MERIC support for dynamic tuning at this point is only for CPUs.

We also plan to focus on the hardware of production EuroHPC systems and evaluate energy efficiency optimization on multiple nodes, while performing strong scaling.

References

- [1] D. Hackenberg, T. Ilsche, J. Schuchart, R. Schöne, W. Nagel, M. Simon, and Y. Georgiou, “Hdeem: High definition energy efficiency monitoring,” in *Energy Efficient Supercomputing Workshop (E2SC)*, 11 2014, pp. 1–10.
- [2] Vince Weaver, “The Unofficial Linux Perf Events Web-Page,” http://web.eece.maine.edu/~vweaver/projects/perf_events/index.html, ”[Online; accessed 2024-02-26]”.
- [3] NVIDIA Corp., “NVML API Reference,” <https://docs.nvidia.com/deploy/nvml-api/nvml-api-reference.html>, January 2023, ”[Online; accessed 2024-02-26]”.
- [4] O. Vysocky, M. Beseda, L. Riha, J. Zapletal, V. Nikl, M. Lysaght, and V. Kannan, “Evaluation of the HPC applications dynamic behavior in terms of energy consumption,” in *Proceedings of the Fifth International Conference on Parallel, Distributed, Grid and Cloud Computing for Engineering*, 2017, pp. 1–19, paper 3, 2017. doi:10.4203/ccp.111.3.
- [5] O. Vysocky, M. Beseda, L. Riha, J. Zapletal, M. Lysaght, and V. Kannan, “Meric and radar generator: Tools for energy evaluation and runtime tuning of hpc applications,” in *High Performance Computing in Science and Engineering*, T. Kozubek, M. Cermak, P. Tichy, R. Blaheta, J. Sistek, D. Lukas, and J. Jaros, Eds. Cham: Springer International Publishing, 2018, pp. 144–159.
- [6] IT4Innovations, “MERIC library,” <https://code.it4i.cz/vys0053/meric>, ”[Online; accessed 2024-02-26]”.
- [7] “The pluto code for astrophysical gasdynamics,” <http://plutocode.ph.unito.it/>, accessed: 28.10.2023.
- [8] V. Springel, “The cosmological simulation code GADGET-2,” , vol. 364, no. 4, pp. 1105–1134, Dec. 2005.
- [9] S. Markidis, G. Lapenta, and Rizwan-uddin, “Multi-scale simulations of plasma with iPIC3D,” *Math. Comput. Simul.*, vol. 80, no. 7, pp. 1509–1519, 2010.
- [10] R. Teyssier, “Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES,” , vol. 385, pp. 337–364, Apr. 2002.
- [11] “Repository of the ramses code on bitbucket,” <https://bitbucket.org/rteyssie/ramses/src/master/>, accessed: 29.10.2023.
- [12] O. Porth, H. Olivares, Y. Mizuno, Z. Younsi, L. Rezzolla, M. Moscibrodzka, H. Falcke, and M. Kramer, “The Black Hole Accretion Code,” *Computational Astrophysics and Cosmology*, vol. 4, 2017.
- [13] H. Olivares, O. Porth, J. Davelaar, E. R. Most, C. M. Fromm, Y. Mizuno, Z. Younsi, and L. Rezzolla, “Constrained transport and adaptive mesh refinement in the black hole accretion code,” *Astronomy & Astrophysics*, vol. 629, p. A61, 2019.
- [14] “The Black Hole Accretion Code – Documentation,” <https://bhac.science>, accessed: 2023-11-01.
- [15] “Repository of BHAC on GitLab,” <https://gitlab.itp.uni-frankfurt.de/BHAC-release/bhac>, accessed: 2023-11-01.
- [16] “The MPI - Adaptive Mesh Refinement - Versatile Advection Code – Documentation,” <https://amrvac.org/>, accessed: 2023-11-01.
- [17] “Repository of MPI-AMRVAC on GitHub,” <https://github.com/amrvac/amrvac>, accessed: 2023-11-01.
- [18] Z. B. Etienne, V. Paschalidis, R. Haas, P. Mösta, and S. L. Shapiro, “IllinoisGRMHD: An Open-Source, User-Friendly GRMHD Code for Dynamical Spacetimes,” *Class. Quant. Grav.*, vol. 32, p. 175009, 2015.
- [19] M. Zilhão and F. Löffler, “An Introduction to the Einstein Toolkit,” *Int. J. Mod. Phys. A*, vol. 28, p. 1340014, 2013.
- [20] “Repository of the changa code on github,” <https://github.com/N-BodyShop/changa>, accessed: 28.10.2023.

REFERENCES

- [21] P. Jetley, F. Gioachin, C. Mendes, L. V. Kale, and T. R. Quinn, “Massively parallel cosmological simulations with ChaNGa,” in *Proceedings of IEEE International Parallel and Distributed Processing Symposium 2008*, 2008.
- [22] H. Menon, L. Wesolowski, G. Zheng, P. Jetley, L. Kale, T. Quinn, and F. Governato, “Adaptive techniques for clustered N-body cosmological simulations,” *Computational Astrophysics and Cosmology*, vol. 2, p. 1, Mar. 2015.
- [23] J. W. Wadsley, B. W. Keller, and T. R. Quinn, “Gasoline2: a modern smoothed particle hydrodynamics code,” , vol. 471, no. 2, pp. 2357–2369, Oct. 2017.
- [24] J. G. Stadel, “Cosmological N-body simulations and their analysis,” Ph.D. dissertation, University of Washington, Seattle, Jan. 2001.