HPC Centre of Excellence
Performance Optimisation and Productivity

Brian Wylie, Jülich Supercomputing Centre

HORIZON-EUROHPC-JU-2023-COE

EuroHPC Joint Undertaking

1 January 2024– 31 December 2026

Grant Agreement No 101143931

POP CoE

- A **Centre of Excellence**
  - On **Performance Optimisation and Productivity**
  - Promoting **best practices in parallel programming**
- Providing **FREE Services**
  - Precise understanding of application execution and system behaviour
  - Suggestion/support on how to refactor code in the most productive way
- **Horizontal**
  - Transversal across application areas, platforms, scales
- **For (EuroHPC) academic AND industrial codes and users!**

EuroHPC
Joint Undertaking

Partners

- **Who?**
  - BSC, ES (coordinator)
  - HLRS, DE
  - INESC-ID, PT
  - IT4I, CZ
  - JSC, DE
  - RWTH Aachen, IT Center, DE
  - TERATEC, FR
  - UVSQ, FR

A team with

- Excellence in performance tools and tuning
- Excellence in programming models and practices
- Research and development background AND
  proven commitment in application to real academic and industrial use cases

Motivation

Why?

- Complexity of parallel computers and codes
  - ⇨ Frequent lack of quantified understanding of actual behaviour
  - ⇨ Not clear most productive direction of code refactoring
- Important to maximize execution efficiency (performance, energy usage) and scaling of compute intensive applications and productivity of development efforts

What?

- Parallel programs, mainly MPI & OpenMP
  - Although also OpenACC, OpenCL, CUDA/HIP, …

EuroHPC
Joint Undertaking

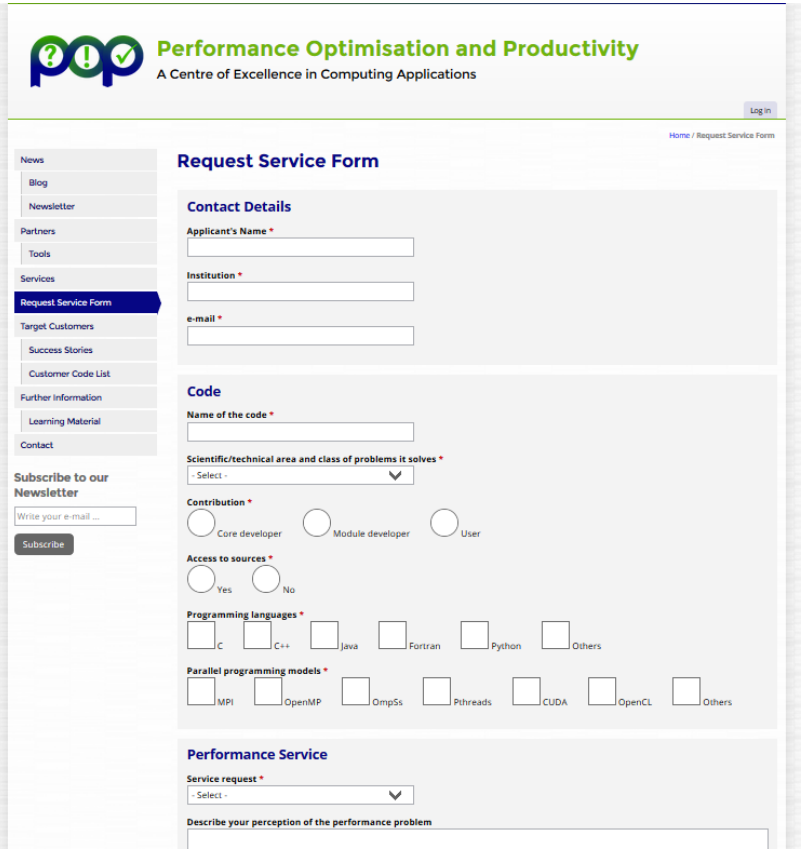
The Process …

When?

October 2015 – March 2018
December 2018 – November 2021
January 2024 – December 2026

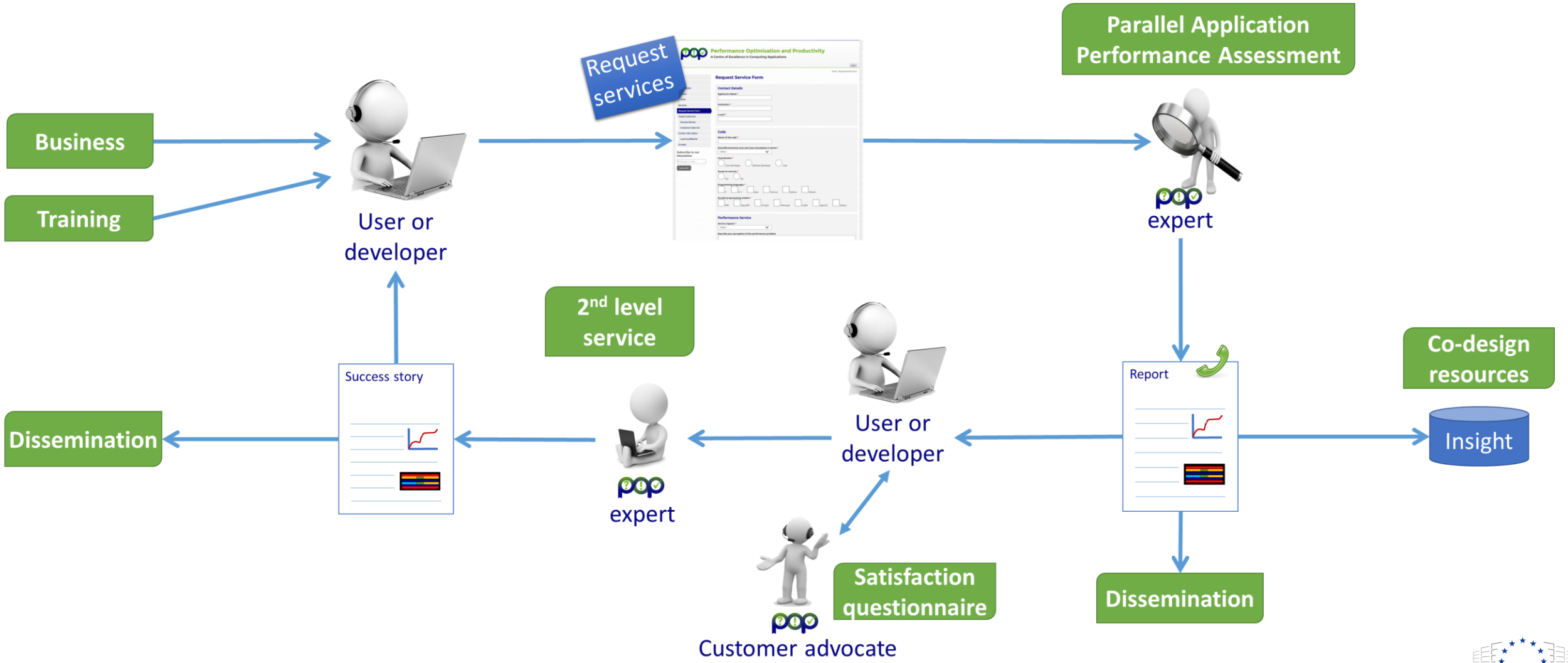
How?

- Apply via submitting short questionnaire describing application and needs https://pop-coe.eu/request-service-form
  - Questions? Ask pop@bsc.es
- Selection/assignment process to analysts
- Install tools on your production machine (local, EuroHPC, …)
- Interactively: Gather data → Analysis → Report



EuroHPC Joint Undertaking

The Process …


Business
Training
User or developer
Request services
Performance Optimisation and Productivity
A Centre of Excellence in Computing Applications
Request Service Form
Contact Details
Code
Performance Service
Parallel Application Performance Assessment
expert
Report
Co-design resources
Insight
Dissemination
Satisfaction questionnaire
Customer advocate
User or developer
2nd level service
expert
Success story
Dissemination

FREE Services provided by the CoE

- **Parallel Application Performance Assessment**
  - Initial analysis measuring a range of performance metrics to assess quality of execution and identify the issues affecting performance (at customer site)
  - If needed, undertakes further performance evaluations to identify the root causes of the issues found and qualify and quantify approaches to address them (recommendations)
- **Second Level Services**
  - Second level services may follow after conclusion of an initial performance assessment:
    - **Proof-of-concept prototype**: explore the potential benefit of proposed optimisations by applying them to selected regions of the applications
    - **Correctness-check**: evaluate the correctness of hybrid MPI + OpenMP applications
    - **Energy-efficiency study**: investigate improvements of energy consumption or efficiency
    - **Advisory study**: ongoing follow-up consultancy and assessments for customers that choose to implement proposed optimisations on their own
- Note: Effort shared between our expert analysts and customer!

EuroHPC
Joint Undertaking

Training provided by the CoE

- Contributions to hands-on workshops (hackathons)
  - joint work on participants' own application code(s) on host's HPC system(s)
  - primarily VI-HPS Tuning Workshops
    - Virtual Institute - High Productivity Supercomputing
  - special events organized *by & for women*$^{+}$
    - exclusively female instructors (*instructrices*) & assistants
    - targeted and prioritized towards under-represented groups in HPC
    - promoted by the Women in HPC organisation
- Tutorials delivered at HPC events/conferences
  - participants' work on their notebook computers with provided measurements
  - ISC21 & ISC24: Determining parallel application execution efficiency & scaling using the POP methodology
    - Judit Giménez, Marta García-Gasulla, Sandra Mendez, Anke Visser, Brian Wylie

Target Customers

- **Code developers**
  - Assessment of detailed actual execution behaviour
  - Suggestion of most productive directions to refactor code
- **Users**
  - Assessment of achieved performance in specific production conditions
  - Possible improvements from modifying environment setup
  - Evidence to interact with code provider
- **Infrastructure operators**
  - Assessment of achieved performance in specific production conditions
  - Possible improvements from modifying environment setup
  - Information for computation time allocation processes
  - Training of support staff
- **Vendors**
  - Benchmarking
  - Customer support
  - System dimensioning/design

EuroHPC
Joint Undertaking

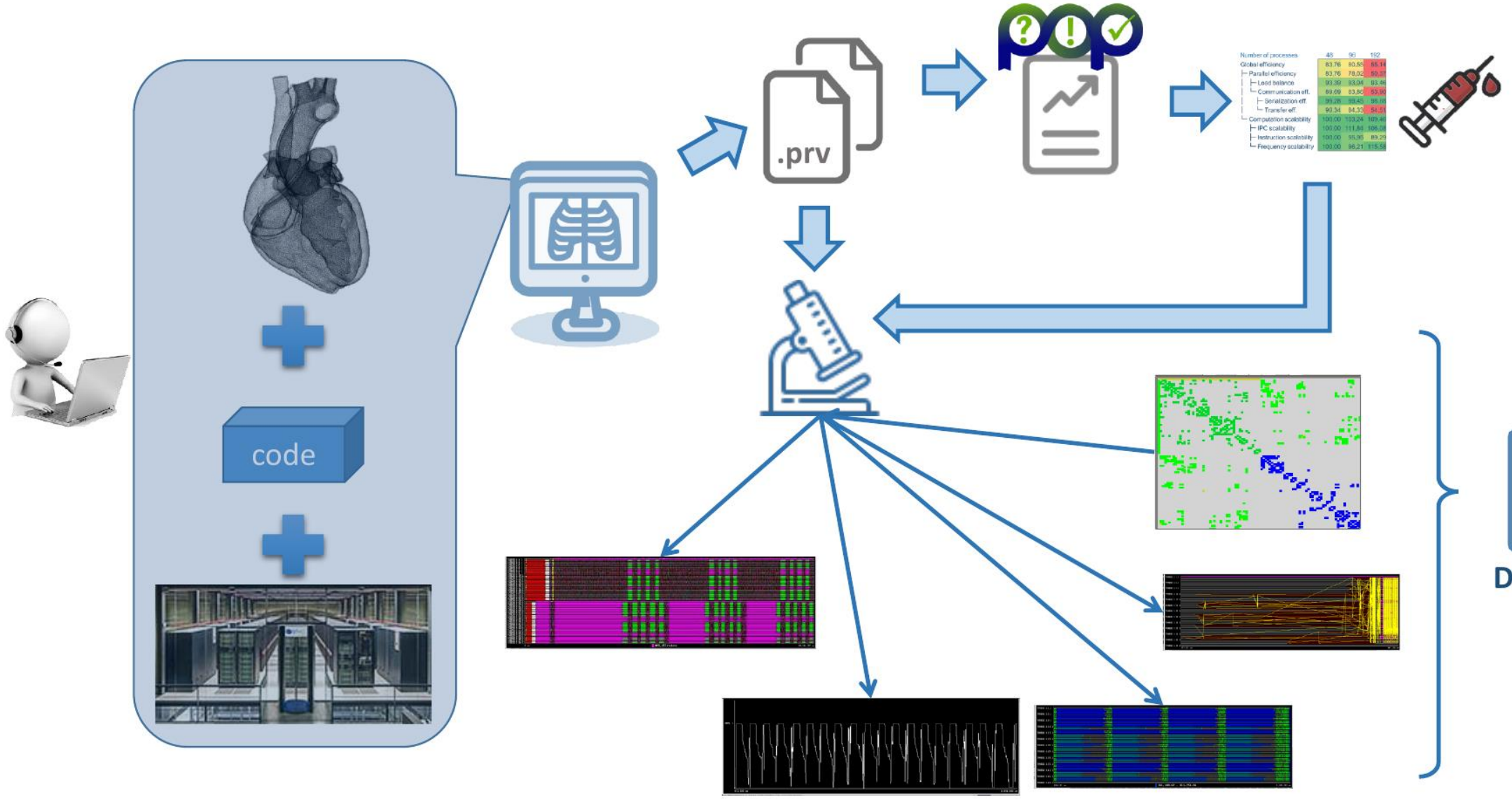
Tools

- **Install and use already available measurement & analysis technology**
  - Analysis and predictive capabilities
  - Delivering insight
    - Highly detailed (CPU + GPU)
    - Up to extreme scale
- **Open-source toolsets**
  - **Extrae + Paraver + Dimemas**
  - **Score-P + CUBE + Scalasca**
  - Archer + MUST + OTF-CPT
  - MAQAO
  - CARM
  - MERIC + RADAR
- **ISV & System Vendor toolsets**

  (when available on HPC system)
  - Vampir
  - TAU
  - FORGE (Linaro)
  - AMD, Intel & Nvidia tools
  - Cray/HPE tools

EuroHPC
Joint Undertaking

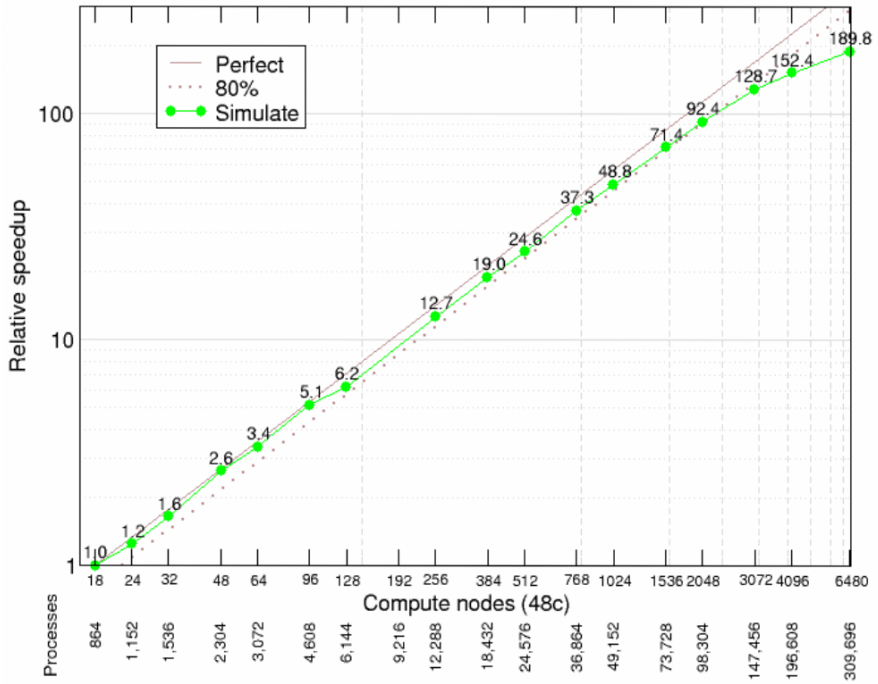
Assessment methodology



Execution efficiency & scaling

Compute nodes	24	32	48	64	96	128	192	256	384	512	768	1024	1536	2048	3072	4096	6452
Processes	**1152**	**1536**	**2304**	**3072**	**4608**	**6144**	**9216**	**12288**	**18432**	**24576**	**36864**	**49152**	**73728**	**98304**	**147456**	**196608**	**309696**
Global scaling efficiency	**0.79**	**0.79**	**0.84**	**0.80**	**0.82**	**0.75**		**0.73**	**0.72**	**0.73**	**0.74**	**0.68**	**0.68**	**0.65**	**0.62**	**0.57**	**0.45**
- Parallel efficiency	0.79	0.80	0.87	0.83	0.86	0.80		0.75	0.74	0.74	0.77	0.71	0.72	0.70	0.72	0.70	0.73
- - Load balance efficiency	0.79	0.80	0.88	0.84	0.86	0.80		0.75	0.74	0.75	0.78	0.72	0.74	0.72	0.74	0.73	0.80
- - Communication efficiency	1.00	1.00	1.00	1.00	1.00	1.00		1.00	1.00	0.99	0.99	0.99	0.98	0.98	0.97	0.96	0.92
- Computation scaling	*1.00*	0.99	0.96	0.96	0.95	0.93		0.98	0.98	0.98	0.96	0.96	0.94	0.93	0.87	0.81	0.61
- - Instructions scaling	*1.00*	1.00	1.00	1.00	1.00	1.00		1.00	1.00	1.00	0.99	0.97	0.94	0.89	0.79	0.67	0.45
- - IPC scaling	*1.00*	0.99	0.96	0.96	0.95	0.93		0.98	0.98	0.99	0.98	0.99	1.00	1.04	1.11	1.21	1.36
IPC	1.411	1.395	1.353	1.355	1.342	1.316		1.377	1.387	1.396	1.383	1.390	1.417	1.473	1.566	1.704	1.919

Key:	<0.65	<0.75	<0.85	<0.95	<1.00	>1.00

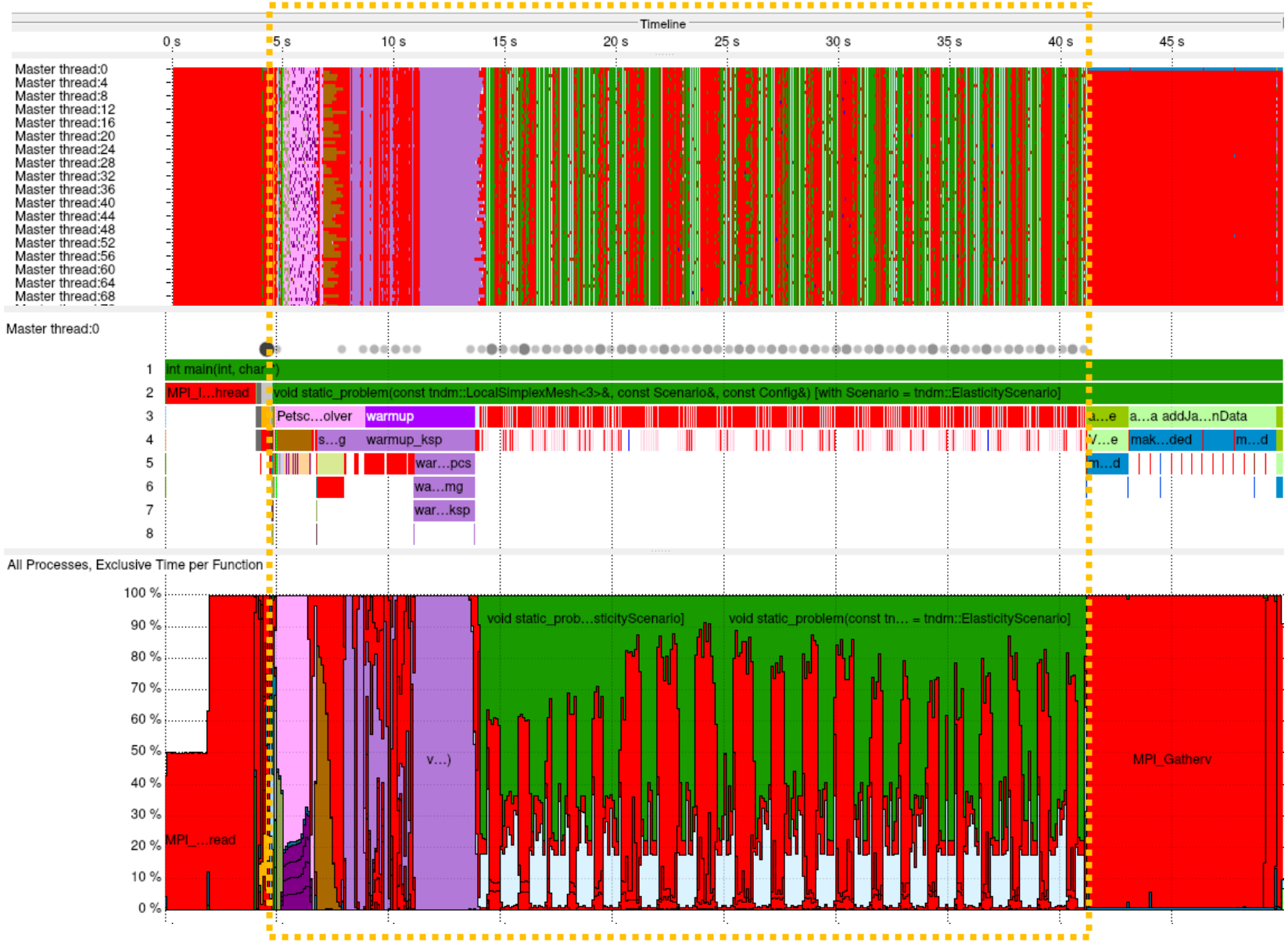


Characterisation of parallel execution

- universal model, composition of factors
- quantify efficiency factors
- their evolution with increasing scale

EuroHPC Joint Undertaking

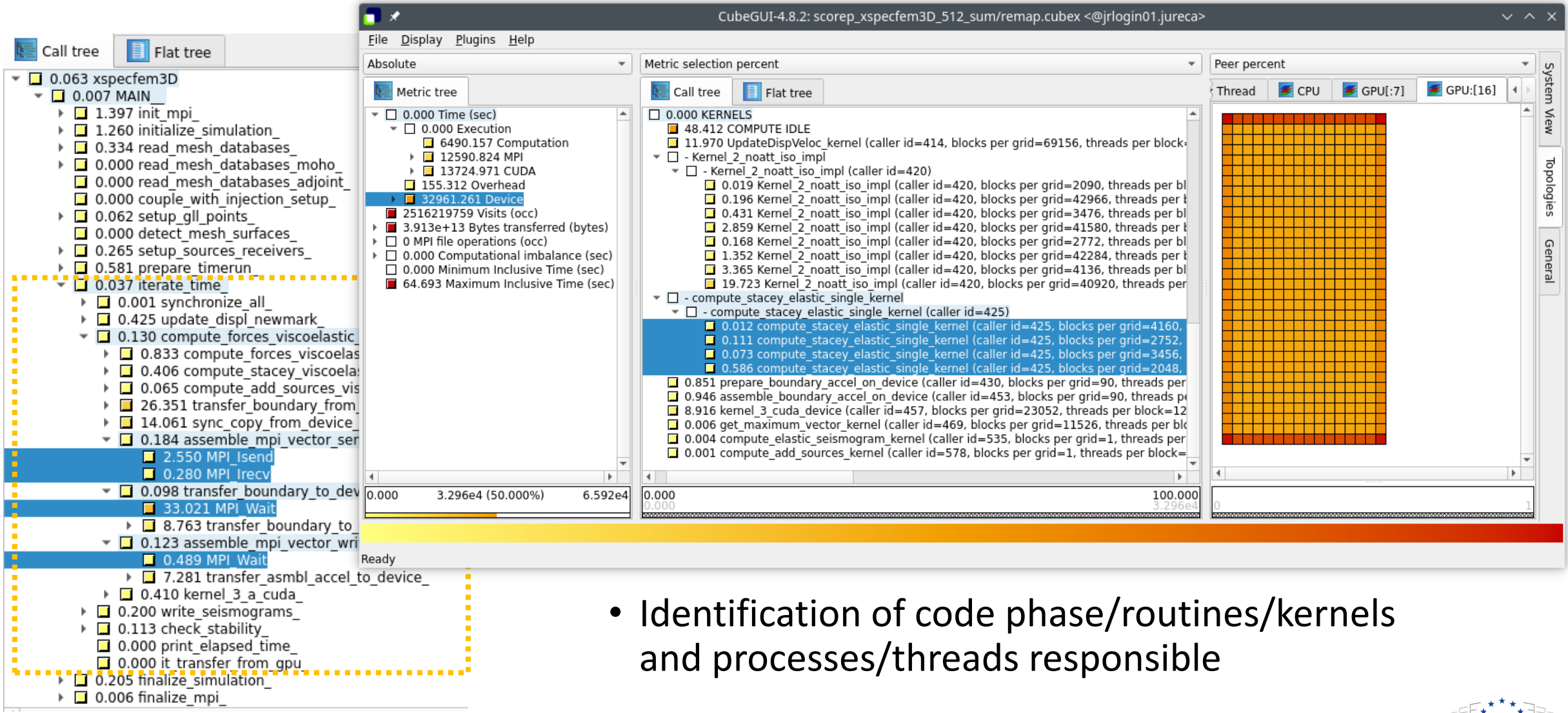
Detailed investigation of causes



- Identification of phases with distinct execution characteristics
  - Initialization
    - MPI, file read & partition
  - Computation
    - assembly
    - warmup
    - solve (20 iterations)
  - Finalization
    - collation & output
- Determine appropriate *Focus of Analysis* (FoA)

EuroHPC
Joint Undertaking

Detailed investigation of causes



- Identification of code phase/routines/kernels and processes/threads responsible

EuroHPC
Joint Undertaking

Performance Optimisation and Productivity

A Centre of Excellence in HPC

Contact:

- https://www.pop-coe.eu
- pop@bsc.es
- @POP_HPC
- youtube.com/POPHPC

This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101143931. The JU receives support from the European Union's Horizon Europe research and innovation programme and Spain, Germany, France, Portugal and the Czech Republic.