



The BSC performance tools and their role in POP3 services

Marta Garcia-Gasulla, BSC

HORIZON-EUROHPC-JU-2023-COE



EuroHPC
Joint Undertaking

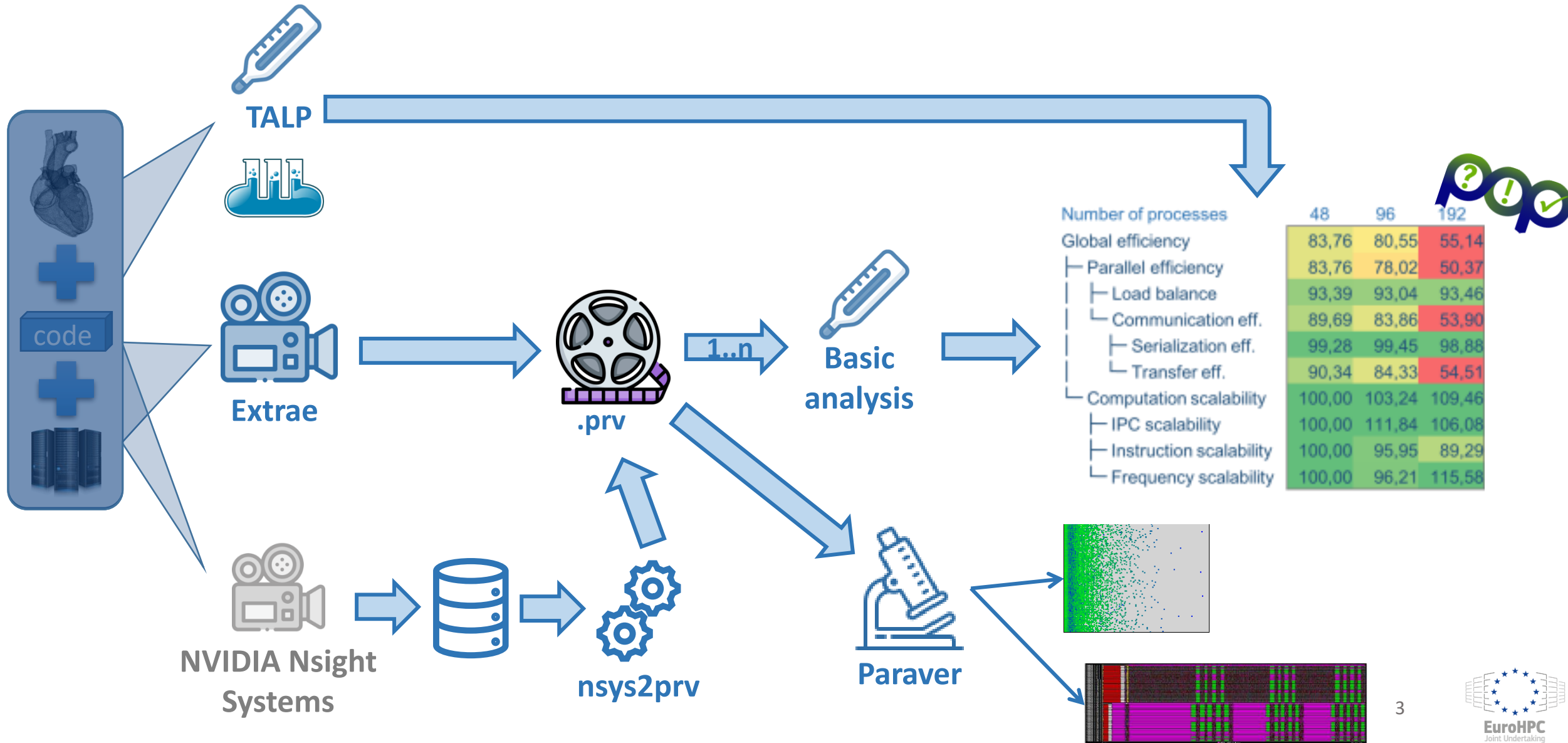
1 January 2024– 31 December 2026

Grant Agreement No 101143931

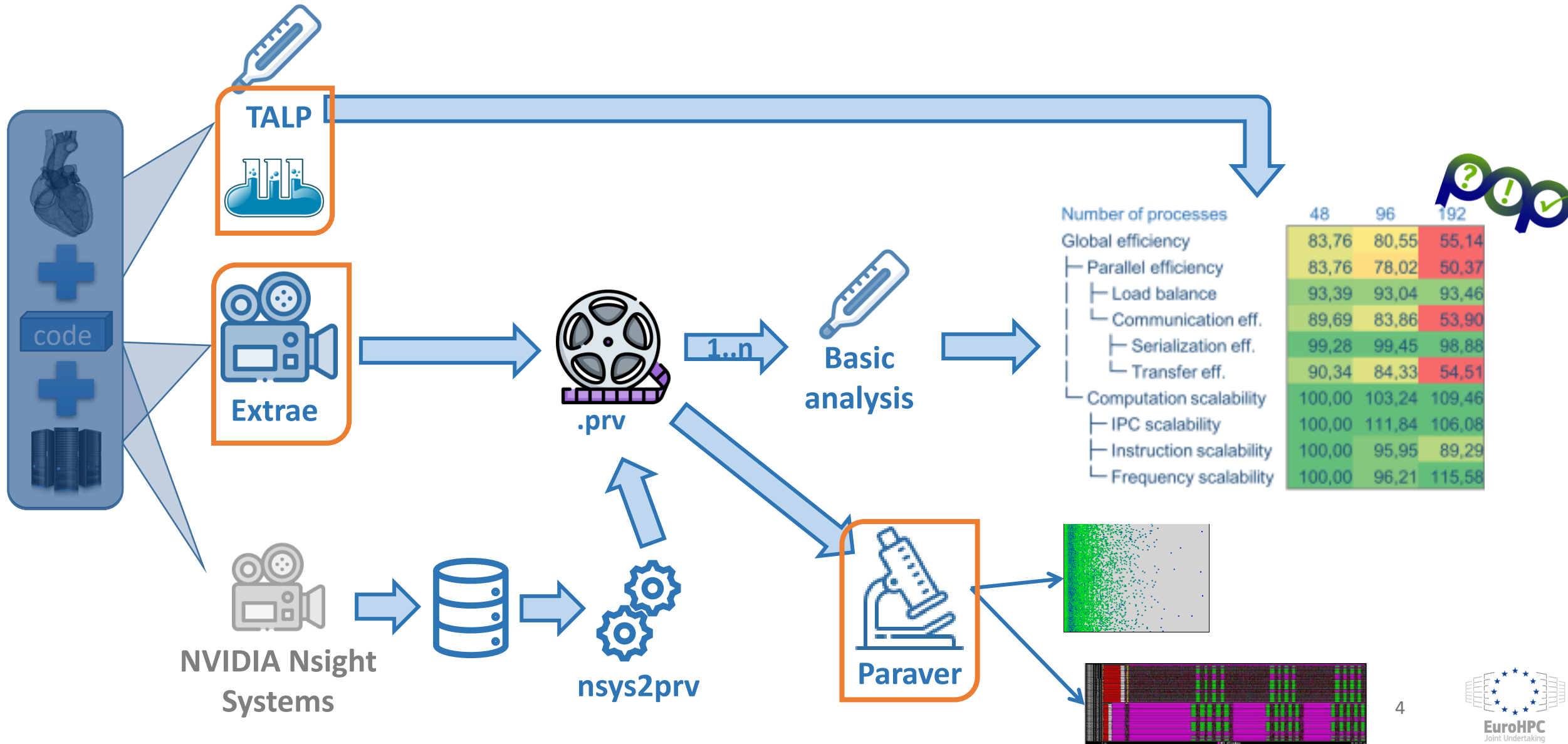


The ecosystem of BSC performance tools

Tracing and profiling tools



Tracing and profiling tools





TALP + Extrae + Paraver in action

Metrics for Propagate slice



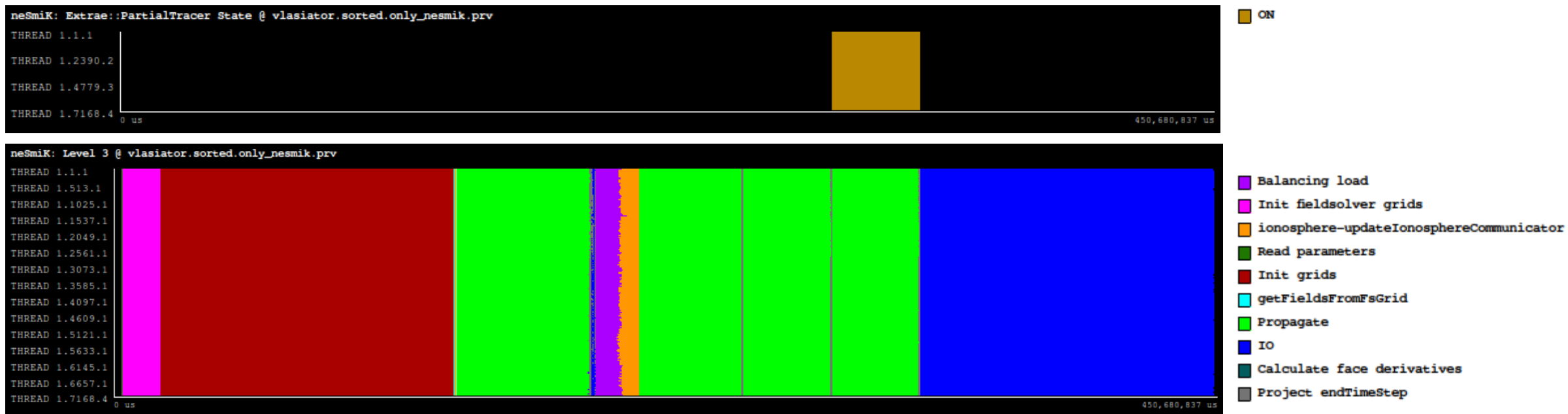
Metrics	7168xMPI 4xOpenMP
Global Efficiency	0.15
- Parallel efficiency	0.15
-- MPI Parallel efficiency	0.25
--- MPI Communication efficiency	0.79
--- MPI Load balance	0.32
---- MPI In-node load balance	0.38
---- MPI Inter-node load balance	0.83
-- OpenMP Parallel efficiency	0.36
--- OpenMP Scheduling efficiency	1
--- OpenMP Load balance	1
--- OpenMP Serialization efficiency	0.36
- Computation Scalability	1
-- Instructions scaling	1
-- IPC scaling	1
-- Frequency scaling	1
Useful IPC	1.43
Frequency [GHz]	2.91
Elapsed time [s]	226.79

- We check the metrics obtained with TALP for the relevant region

Macro-Structure



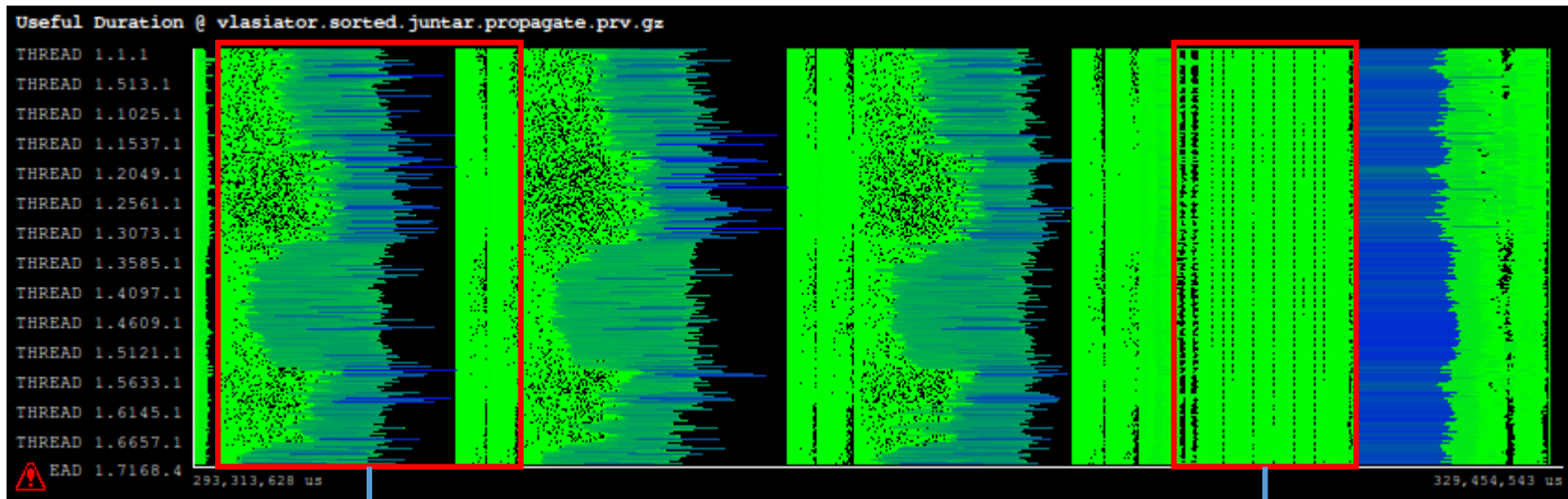
- Partial-tracing set to trace fourth iteration of time-step region: Propagate



Propagate



- Update system boundaries (Vlasov pre-translation)
- Velocity-space
- Propagate Fields
- Update system boundaries (Vlasov post-translation)
- Spatial-space
- Bailout-allreduce
- Update system boundaries (Vlasov post-acceleration)
- logfile-io
- write-restart



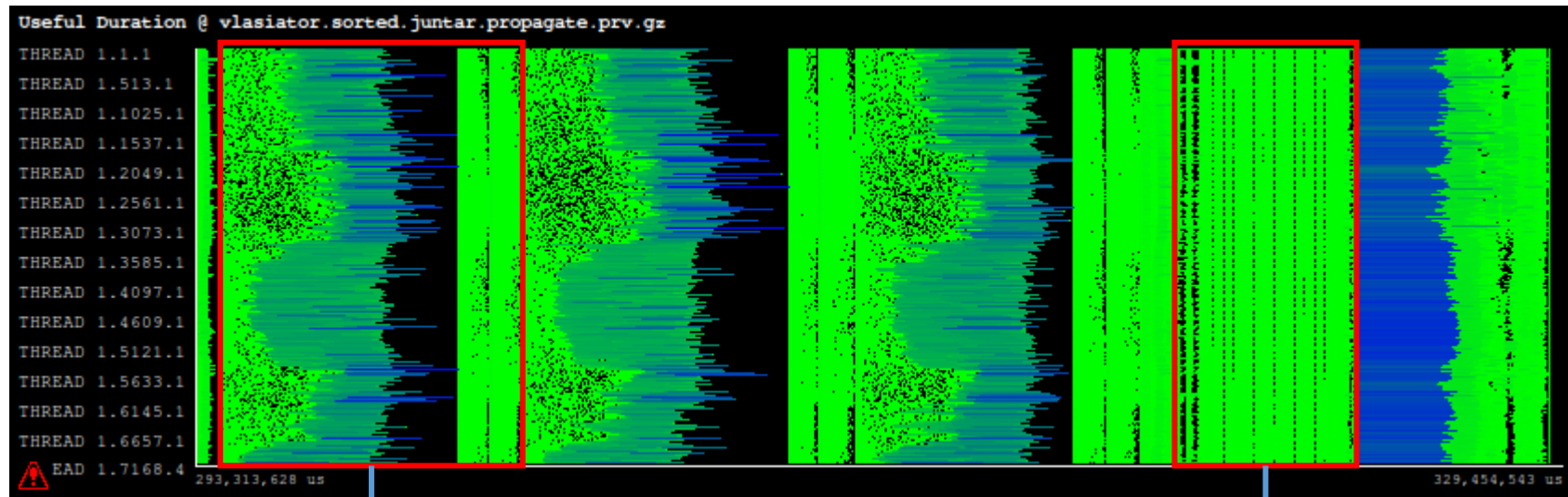
→ Spatial-space

→ Propagate Fields

Propagate



- Update system boundaries (Vlasov pre-translation)
- Velocity-space
- Propagate Fields
- Update system boundaries (Vlasov post-translation)
- Spatial-space
- Bailout-allreduce
- Update system boundaries (Vlasov post-acceleration)
- logfile-io
- write-restart



→ Spatial-space

→ Propagate Fields

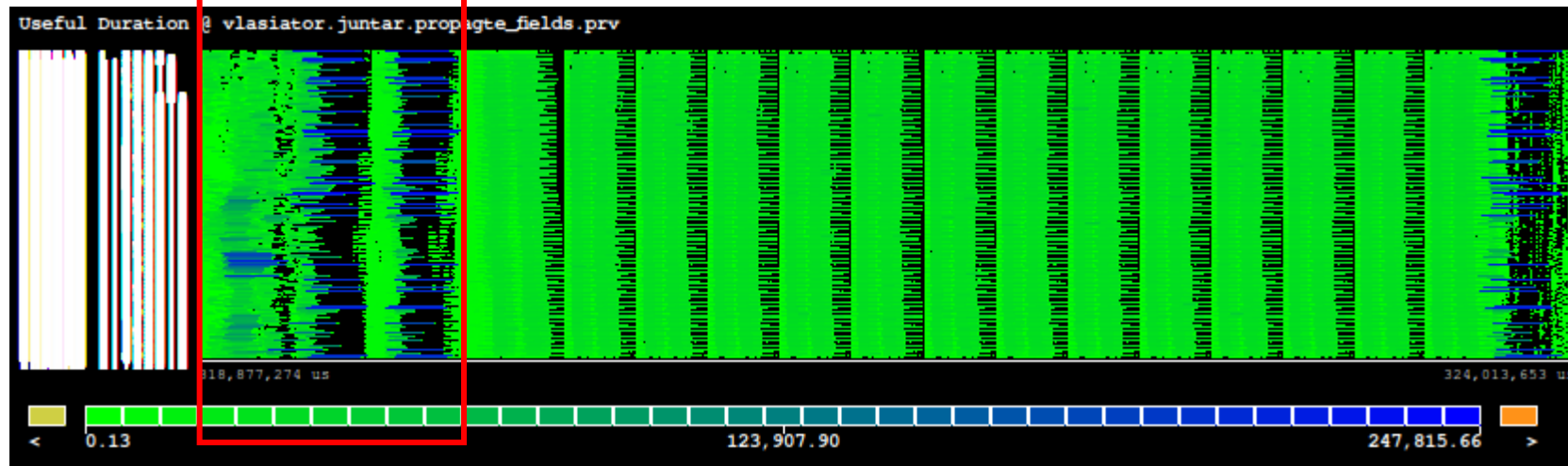
POP metrics Propagate-fields



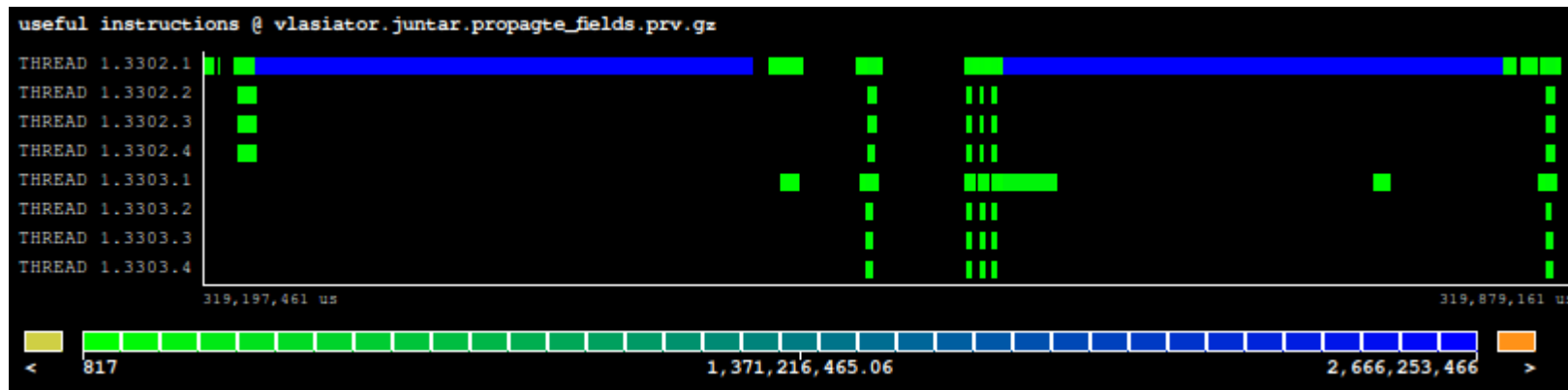
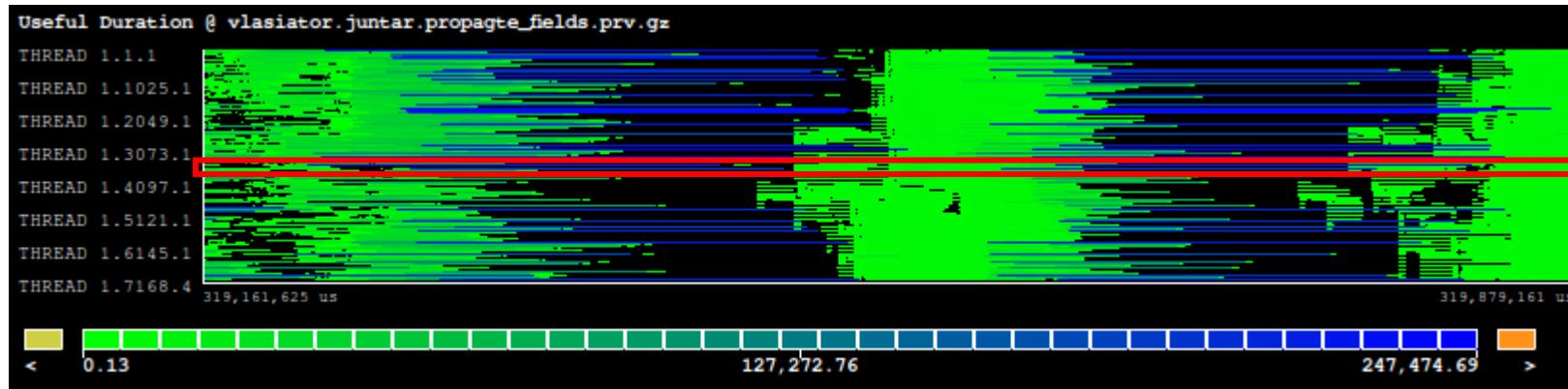
Metrics	Propagate Fields
Global Efficiency	0.15
- Parallel efficiency	0.15
-- MPI Parallel efficiency	0.27
--- MPI Communication efficiency	0.85
--- MPI Load balance	0.32
---- MPI In-node load balance	0.41
---- MPI Inter-node load balance	0.77
-- OpenMP Parallel efficiency	0.36
--- OpenMP Scheduling efficiency	1
--- OpenMP Load balance	1
--- OpenMP Serialization efficiency	0.36
- Computation Scalability	1
-- Instructions scaling	1
-- IPC scaling	1
-- Frequency scaling	1
Useful IPC	2.84
Frequency [GHz]	2.93
Elapsed time [s]	56.27

- Metrics for the Propagate-fields
- Main issues:
 - MPI Load imbalance
 - OpenMP serialization

Propagate-fields

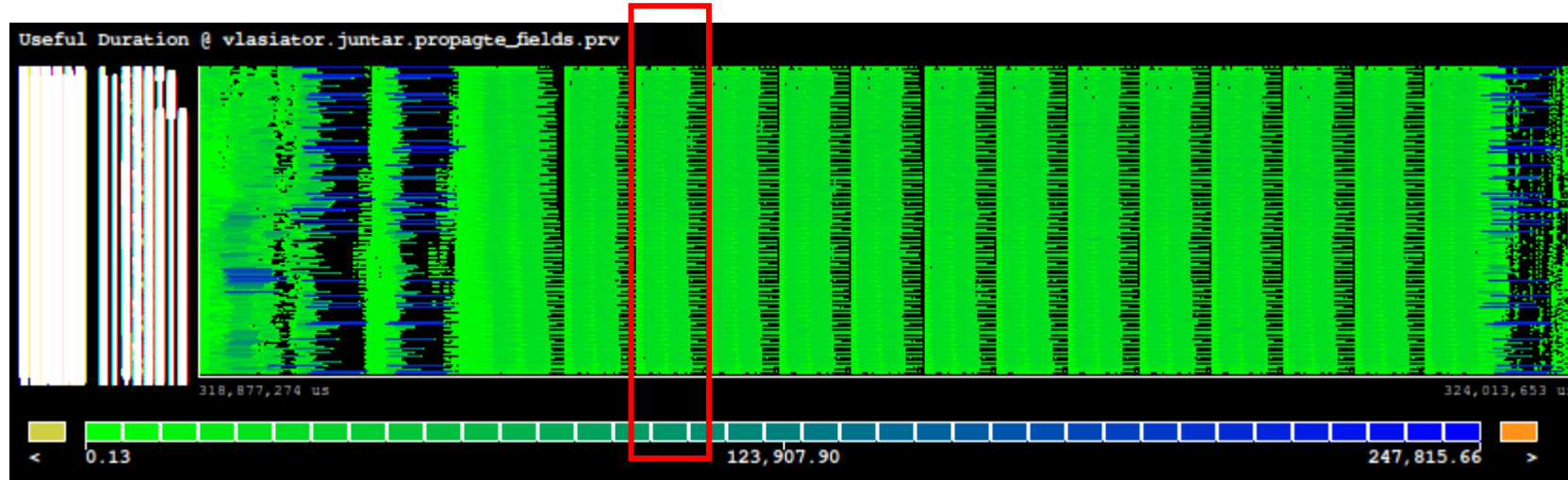


Propagate-fields

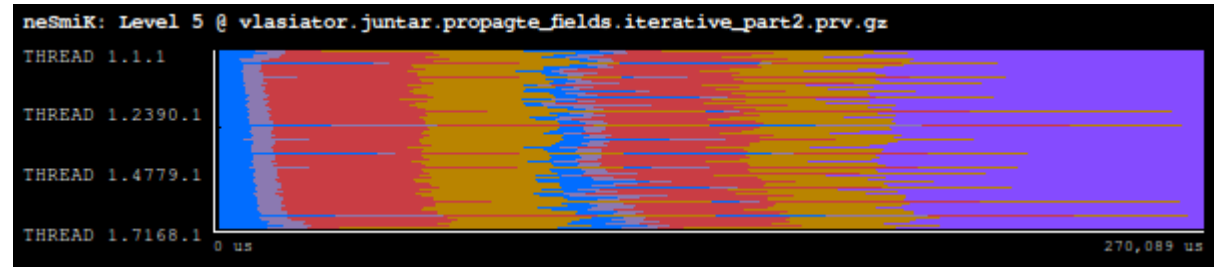
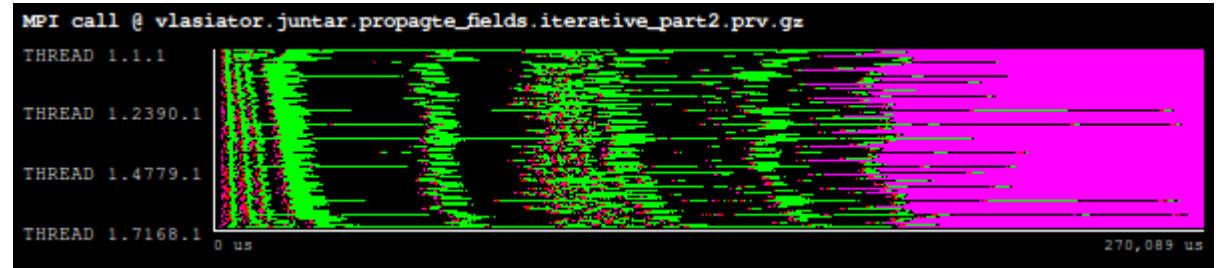
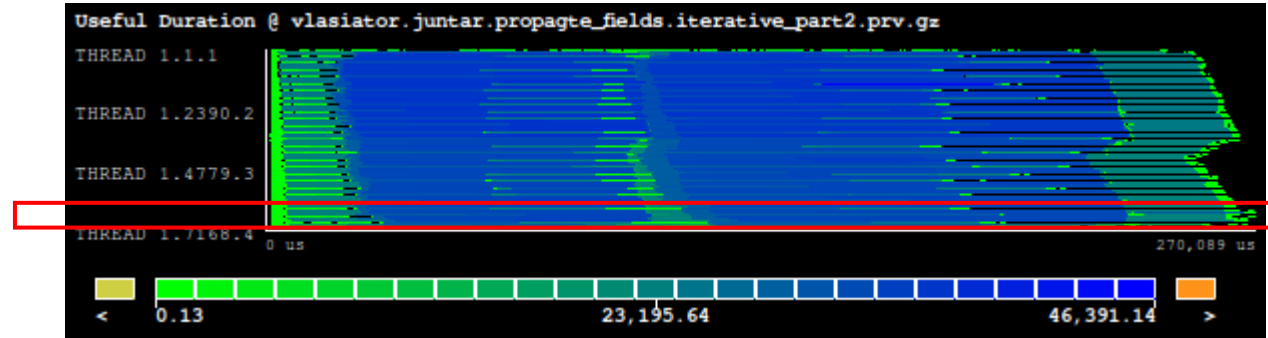


- Some processes more instructions than others
 - MPI LB
- Region not parallelized with OpenMP
 - OpenMP serialization

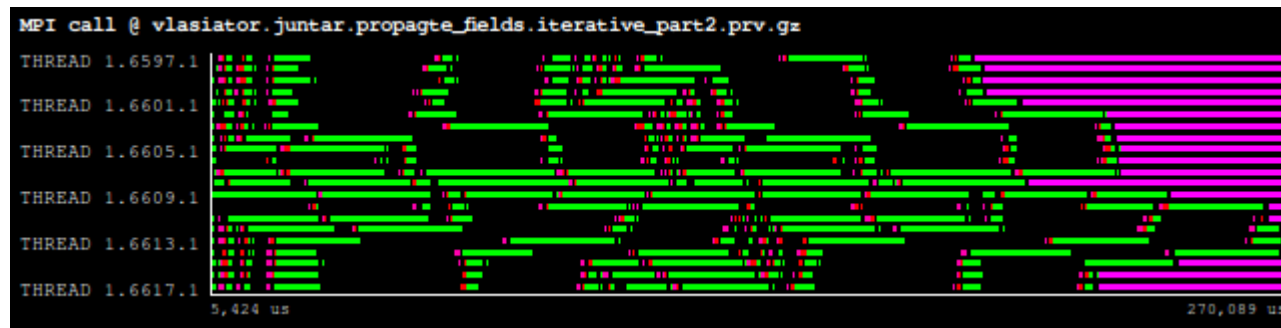
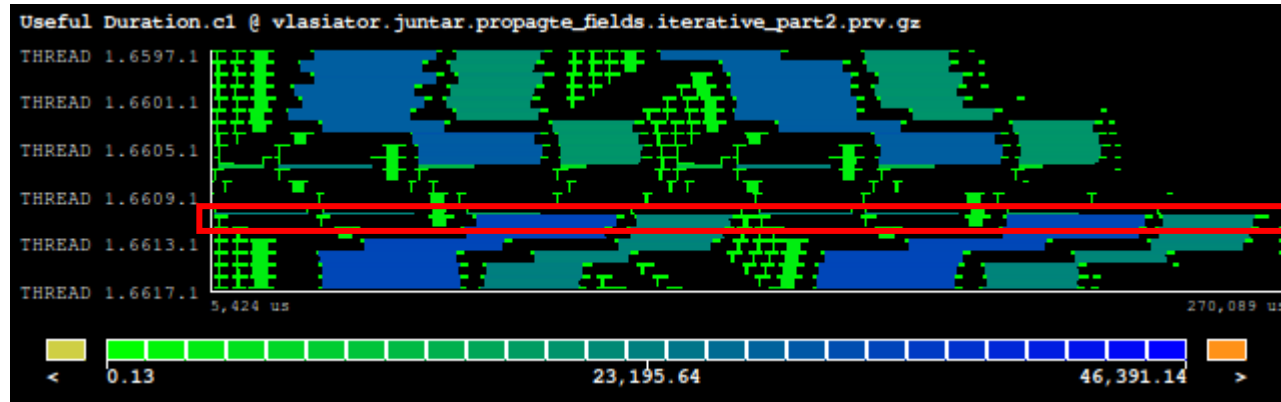
Propagate-fields



Propagate-fields

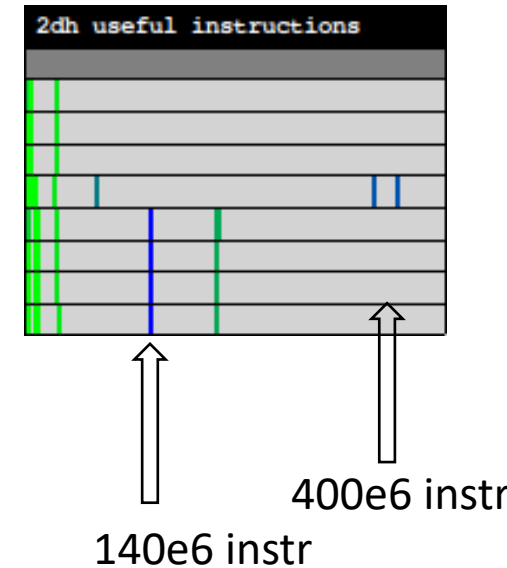
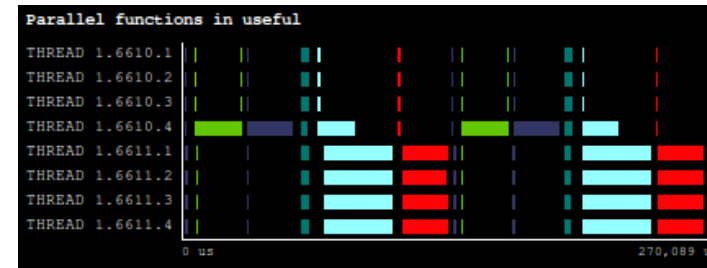
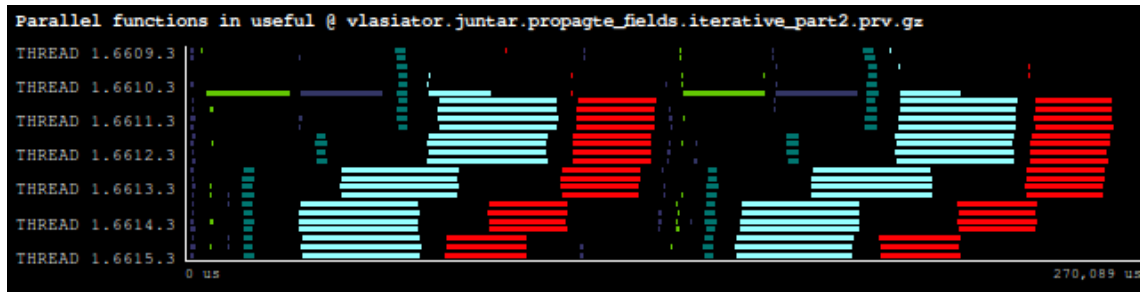
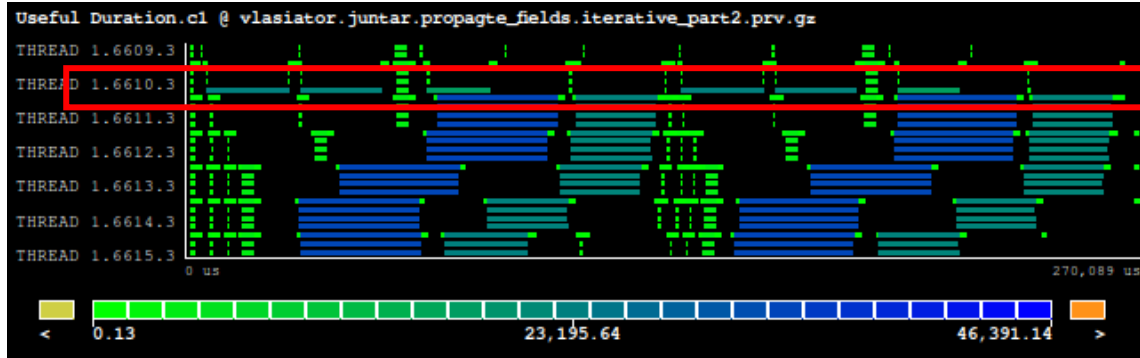


Propagate-fields



- Outside MPI
- MPI_Isend
- MPI_Irecv
- MPI_Waitall
- MPI_Allreduce

Propagate-fields



- Issues in 2 parallel regions:
 - Green/purple: MPI LB, not parallelized OpenMP
 - Blue/red: MPI and OpenMP LB



Performance Optimisation and Productivity 3

A Centre of Excellence in HPC

Contact:

 <https://www.pop-coe.eu>

 pop@bsc.es

 [@POP_HPC](#)

 [youtube.com/POPHPC](https://www.youtube.com/POPHPC)

