



TALP (Live) Demo

Valentin Seitz, BSC

48th VI-HPS Tuning Workshop 09.02.2026

HORIZON-EUROHPC-JU-2023-COE



EuroHPC
Joint Undertaking

1 January 2024– 31 December 2026

Grant Agreement No 101143931

- Hybrid TALP without regions
- Hybrid TALP with regions
- TALP with GPUs

- TALP Pages
- TALP Pages in a GitLab Repo

TALP without regions MPI+OpenMP



- Most basic use case
- Nothing needed except installed DLB

```
export DLB_ARGS="--talp --ompt --talp-openmp --talp-papi"
```

https://dlb-docs.readthedocs.io/en/stable/how_to_run_talp.html#hybrid-openmp-mpi-executions

TALP without regions MPI+OpenMP

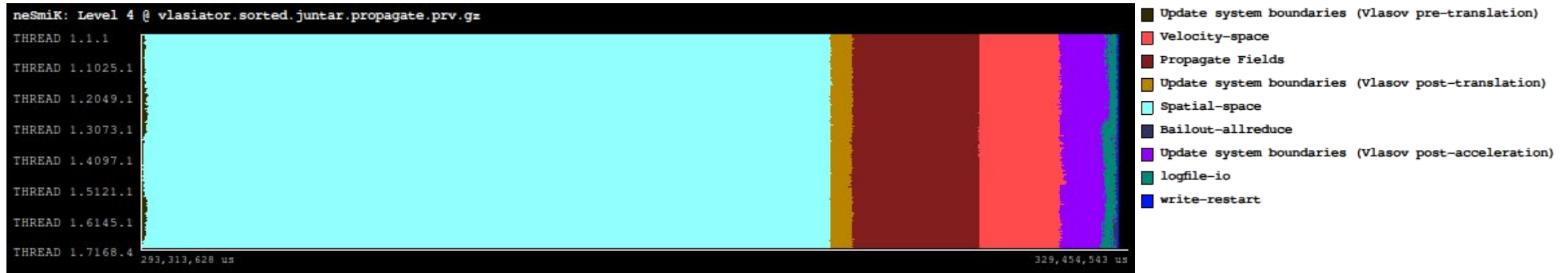
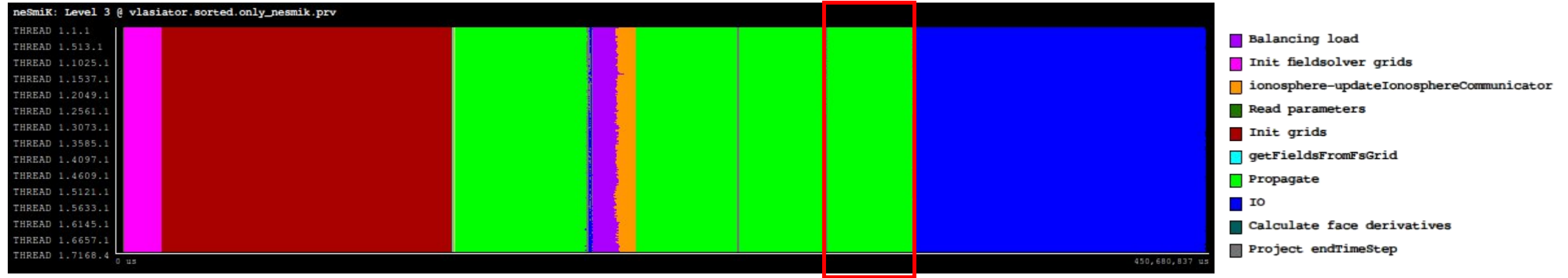


```
DLB[gs09r3b44:36727]: dlb 3.5.3
Maximum refinement level is 3
Finished first level of refinement
Finished second level of refinement
Finished third level of refinement
(MAIN): Completed grid initialization.
(MAIN): Starting main simulation loop.
(MAIN): Completed requested simulation. Exiting.
DLB[gs09r3b44:36727]: ##### Monitoring Region POP Metrics #####
DLB[gs09r3b44:36727]: ### Name: Global
DLB[gs09r3b44:36727]: ### Elapsed Time: 85.86 s
DLB[gs09r3b44:36727]: ### Parallel efficiency: 0.22
DLB[gs09r3b44:36727]: ### - MPI Parallel efficiency: 0.48
DLB[gs09r3b44:36727]: ### - Communication efficiency: 0.67
DLB[gs09r3b44:36727]: ### - Load Balance: 0.72
DLB[gs09r3b44:36727]: ### - In: 0.75
DLB[gs09r3b44:36727]: ### - Out: 0.96
DLB[gs09r3b44:36727]: ### - OpenMP Parallel efficiency: 0.40
DLB[gs09r3b44:36727]: ### - Load Balance: 0.97
DLB[gs09r3b44:36727]: ### - Scheduling efficiency: 1.00
DLB[gs09r3b44:36727]: ### - Serialization efficiency: 0.42
DLB[gs09r3b44:36727]: ### Computational metrics:
DLB[gs09r3b44:36727]: ### - Average useful IPC: 2.70
DLB[gs09r3b44:36727]: ### - Average useful frequency: 2.90 GHz
DLB[gs09r3b44:36727]: ### - Number of instructions: 1.35E+14
```



Let's have regions

Vlasiator: regions revisited



TALP with regions MPI+OpenMP



- Most basic use case
- Nothing needed except installed DLB

```
export DLB_ARGS="--talp --ompt --talp-openmp --talp-papi  
--talp-output-file=talp.json"
```

https://dlb-docs.readthedocs.io/en/stable/how_to_run_talp.html#hybrid-openmp-mpi-executions

TALP with regions MPI+OpenMP



```
DLB[gs17r2b28:1656366]: ##### Monitoring Region POP Metrics
DLB[gs17r2b28:1656366]: ### Name: Initialization
DLB[gs17r2b28:1656366]: ### Elapsed Time: 31.98 s
DLB[gs17r2b28:1656366]: ### Parallel efficiency: 0.20
DLB[gs17r2b28:1656366]: ### - MPI Parallel efficiency: 0.51
DLB[gs17r2b28:1656366]: ### - Communication efficiency: 0.80
DLB[gs17r2b28:1656366]: ### - Load Balance: 0.64
DLB[gs17r2b28:1656366]: ### - In: 0.67
DLB[gs17r2b28:1656366]: ### - Out: 0.95
DLB[gs17r2b28:1656366]: ### - OpenMP Parallel efficiency: 0.38
DLB[gs17r2b28:1656366]: ### - Load Balance: 0.99
DLB[gs17r2b28:1656366]: ### - Scheduling efficiency: 1.00
DLB[gs17r2b28:1656366]: ### - Serialization efficiency: 0.39
DLB[gs17r2b28:1656366]: ### Computational metrics:
DLB[gs17r2b28:1656366]: ### - Average useful IPC: 3.39
DLB[gs17r2b28:1656366]: ### - Average useful frequency: 2.84 GHz
DLB[gs17r2b28:1656366]: ### - Number of instructions: 5.42E+13
DLB[gs17r2b28:1656366]: ##### Monitoring Region POP Metrics
```

TALP with regions MPI+OpenMP



```
DLB[gs17r2b28:1656366]: ### Name: Propagate
DLB[gs17r2b28:1656366]: ### Elapsed Time: 50.27 s
DLB[gs17r2b28:1656366]: ### Parallel efficiency: 0.26
DLB[gs17r2b28:1656366]: ### - MPI Parallel efficiency: 0.48
DLB[gs17r2b28:1656366]: ### - Communication efficiency: 0.72
DLB[gs17r2b28:1656366]: ### - Load Balance: 0.67
DLB[gs17r2b28:1656366]: ### - In: 0.70
DLB[gs17r2b28:1656366]: ### - Out: 0.96
DLB[gs17r2b28:1656366]: ### - OpenMP Parallel efficiency: 0.43
DLB[gs17r2b28:1656366]: ### - Load Balance: 0.96
DLB[gs17r2b28:1656366]: ### - Scheduling efficiency: 1.00
DLB[gs17r2b28:1656366]: ### - Serialization efficiency: 0.45
DLB[gs17r2b28:1656366]: ### Computational metrics:
DLB[gs17r2b28:1656366]: ### - Average useful IPC: 2.38
DLB[gs17r2b28:1656366]: ### - Average useful frequency: 2.93 GHz
DLB[gs17r2b28:1656366]: ### - Number of instructions: 8.00E+13
DLB[gs17r2b28:1656366]: ##### Monitoring Region POP Metrics #####
```



GPUs?

TALP with regions MPI+CUDA



- Most basic use case
- Nothing needed except installed DLB

```
export DLB_ARGS="--talp --plugin=cupti  
                --talp-output-file=talp.json"
```

https://dlb-docs.readthedocs.io/en/stable/how_to_run_talp.html#nvidia-gpu-executions

TALP with regions MPI+CUDA



```
DLB[as06r4b14:2946559]: ##### Monitoring Region POP Metrics #####
DLB[as06r4b14:2946559]: ### Name: Initialization
DLB[as06r4b14:2946559]: ### Elapsed Time: 179.46 s
DLB[as06r4b14:2946559]: ### Host
DLB[as06r4b14:2946559]: ### ----
DLB[as06r4b14:2946559]: ### Parallel efficiency: 0.32
DLB[as06r4b14:2946559]: ### - MPI Parallel efficiency: 0.82
DLB[as06r4b14:2946559]: ### - Communication efficiency: 0.95
DLB[as06r4b14:2946559]: ### - Load Balance: 0.86
DLB[as06r4b14:2946559]: ### - In: 0.86
DLB[as06r4b14:2946559]: ### - Out: 1.00
DLB[as06r4b14:2946559]: ### - Device Offload efficiency: 0.39
DLB[as06r4b14:2946559]: ###
DLB[as06r4b14:2946559]: ### NVIDIA Device
DLB[as06r4b14:2946559]: ### -----
DLB[as06r4b14:2946559]: ### Parallel efficiency: 0.03
DLB[as06r4b14:2946559]: ### - Load Balance: 0.94
DLB[as06r4b14:2946559]: ### - Communication efficiency: 0.57
DLB[as06r4b14:2946559]: ### - Orchestration efficiency: 0.06
```

TALP with regions MPI+CUDA



```
DLB[as06r4b14:2946559]: ##### Monitoring Region POP Metrics #####
DLB[as06r4b14:2946559]: ### Name: Propagate
DLB[as06r4b14:2946559]: ### Elapsed Time: 51.09 s
DLB[as06r4b14:2946559]: ### Host
DLB[as06r4b14:2946559]: ### ----
DLB[as06r4b14:2946559]: ### Parallel efficiency: 0.32
DLB[as06r4b14:2946559]: ### - MPI Parallel efficiency: 0.61
DLB[as06r4b14:2946559]: ### - Communication efficiency: 0.65
DLB[as06r4b14:2946559]: ### - Load Balance: 0.93
DLB[as06r4b14:2946559]: ### - In: 0.93
DLB[as06r4b14:2946559]: ### - Out: 1.00
DLB[as06r4b14:2946559]: ### - Device Offload efficiency: 0.52
DLB[as06r4b14:2946559]: ###
DLB[as06r4b14:2946559]: ### NVIDIA Device
DLB[as06r4b14:2946559]: ### -----
DLB[as06r4b14:2946559]: ### Parallel efficiency: 0.23
DLB[as06r4b14:2946559]: ### - Load Balance: 0.85
DLB[as06r4b14:2946559]: ### - Communication efficiency: 0.97
DLB[as06r4b14:2946559]: ### - Orchestration efficiency: 0.28
```



Post-processing the JSONs



```
talp ci-report -i talp -o talp-report --regions Propagate Initialization
```

https://dlb-docs.readthedocs.io/en/stable/talp_pages.html

TALP Pages



Metrics	112xMPI 4xOpenMP	224xMPI 4xOpenMP
Global Efficiency	0.17	0.15
- Parallel efficiency	0.17	0.15
-- MPI Parallel efficiency	0.47	0.42
--- MPI Communication efficiency	0.69	0.65
--- MPI Load balance	0.68	0.64
---- MPI In-node load balance	0.7	0.68
---- MPI Inter-node load balance	0.96	0.95
-- OpenMP Parallel efficiency	0.37	0.35
-- OpenMP Scheduling efficiency	1	1
-- OpenMP Load balance	0.99	0.99
-- OpenMP Serialization efficiency	0.37	0.35
- Computation Scalability	1	1.01
-- Instructions scaling	1	1
-- IPC scaling	1	1.01
-- Frequency scaling	1	1.01
Useful IPC	3.38	3.4
Frequency [GHz]	2.82	2.84
Elapsed time [s]	73.63	43.07



Integration into CI/CD



`https://pages.pm.bsc.es/dlb/talp-pages-example/talp/`

I wanna run the experiments as well



Go to `/gpfs/scratch/nct_362/TALP/demo/inputs`

- `input_cpu`: CPU version with example `submit.sh`
- `input_gpu`: GPU version with example `submit.sh`



Performance Optimisation and Productivity 3

A Centre of Excellence in HPC

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.

