TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Center for Information Services and High Performance Computing (ZIH)

Interactive Trace Analysis with Vampir

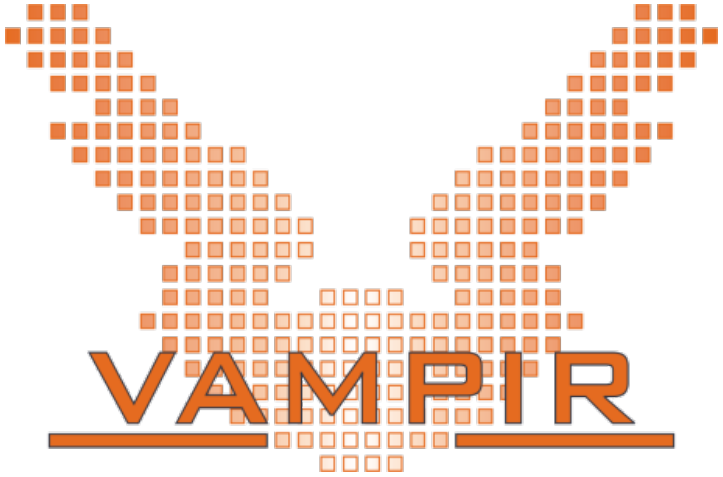
ORNL – Tools Workshop

19-20 Jan 2022

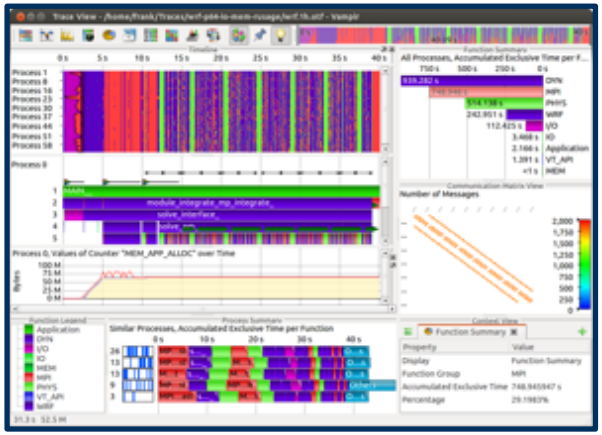
Outline

Welcome to the Vampir Tool Suite

- Mission
- Event Trace Visualization
- Vampir & VampirServer
- The Vampir Displays
- What's New
- Showcases



Mission



Visualization of dynamics
of complex parallel processes

Requires two components

- Monitor/Collector (Score-P)
- Charts/Browser (Vampir)

Typical questions that Vampir helps to answer:

- What happens in my application execution during a given time in a given process or thread?
- How do the communication patterns of my application execute on a real system?
- Are there any imbalances in computation, I/O or memory usage and how do they affect the parallel execution of my application?

Event Trace Visualization with Vampir

- Alternative and supplement to automatic analysis
- Show dynamic run-time behavior graphically at any level of detail
- Provide statistics and performance metrics

Timeline charts

- Show application activities and communication along a time axis, which can be zoomed and scrolled
- Master timeline showing all parallel processes/threads
- Process timeline focusing on a single process/thread

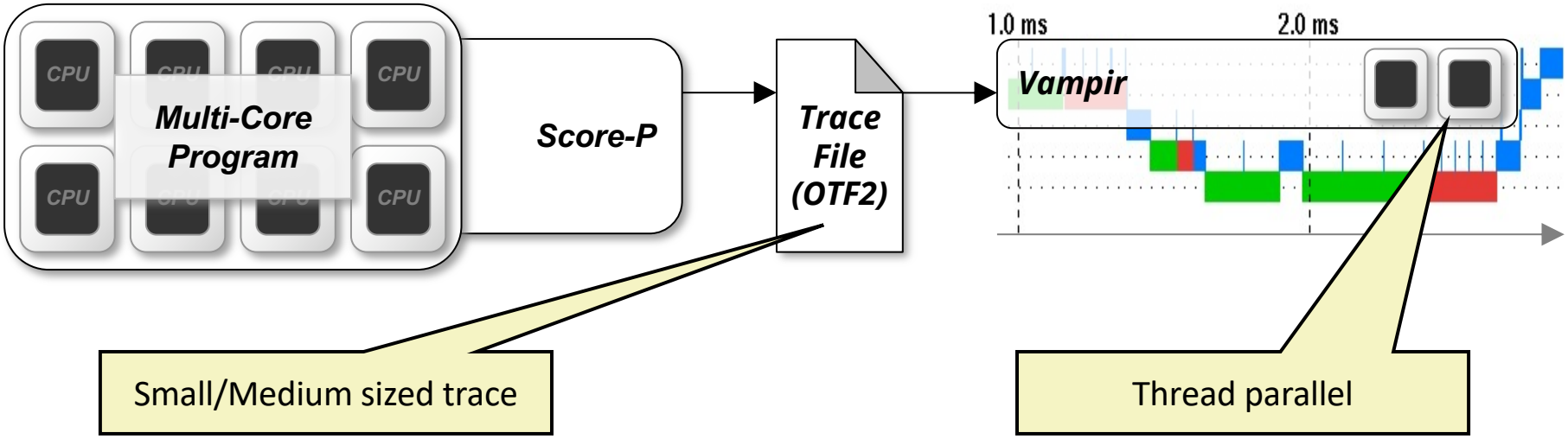
Summary charts

- Provide quantitative results for the currently selected time interval (e.g., Message Summary)

Vampir – Visualization Modes (1)

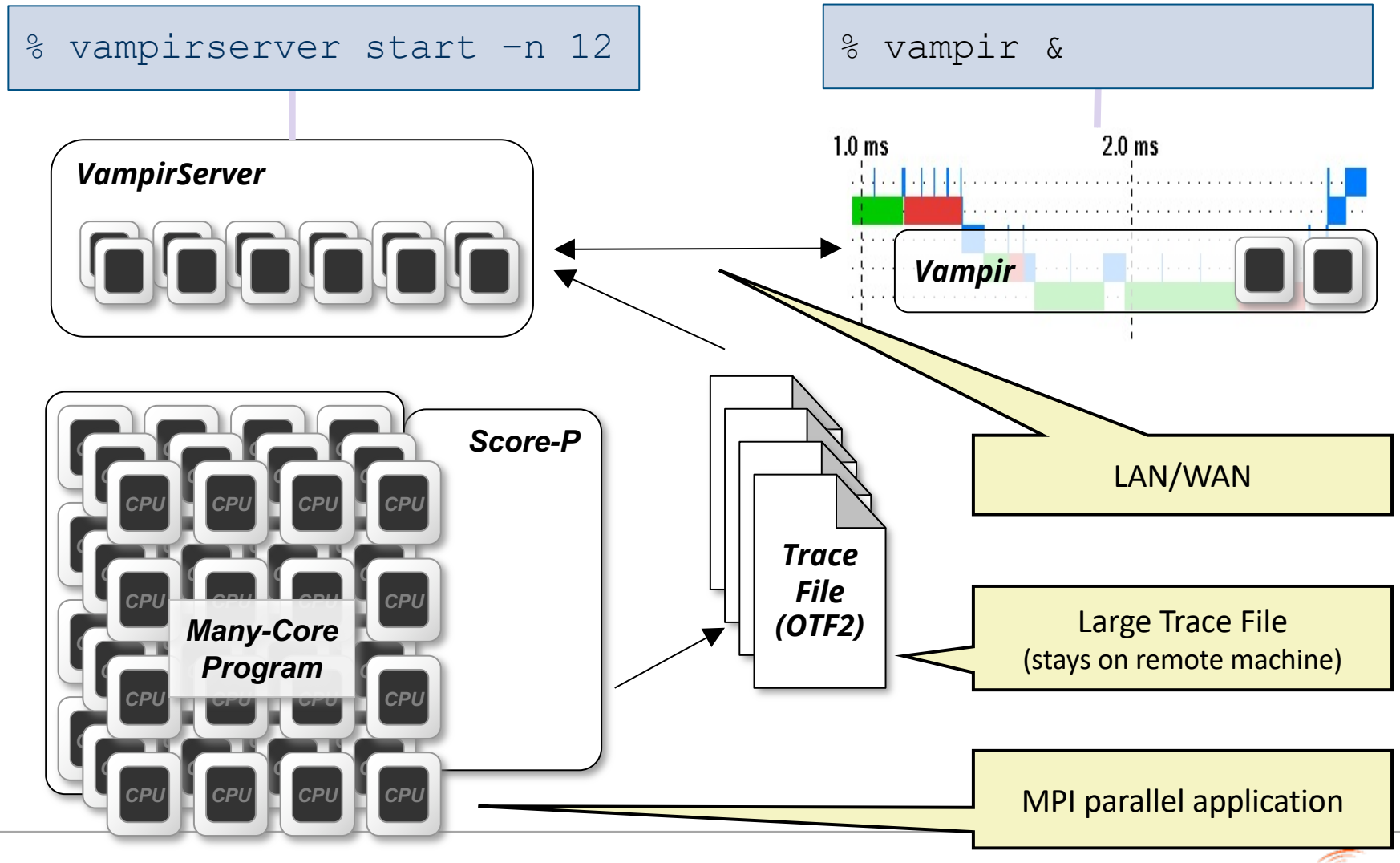
Directly on front end or local machine

```
% vampir &
```



Vampir – Visualization Modes (2)

On local machine with remote VampirServer



Vampir on OLCF Resources

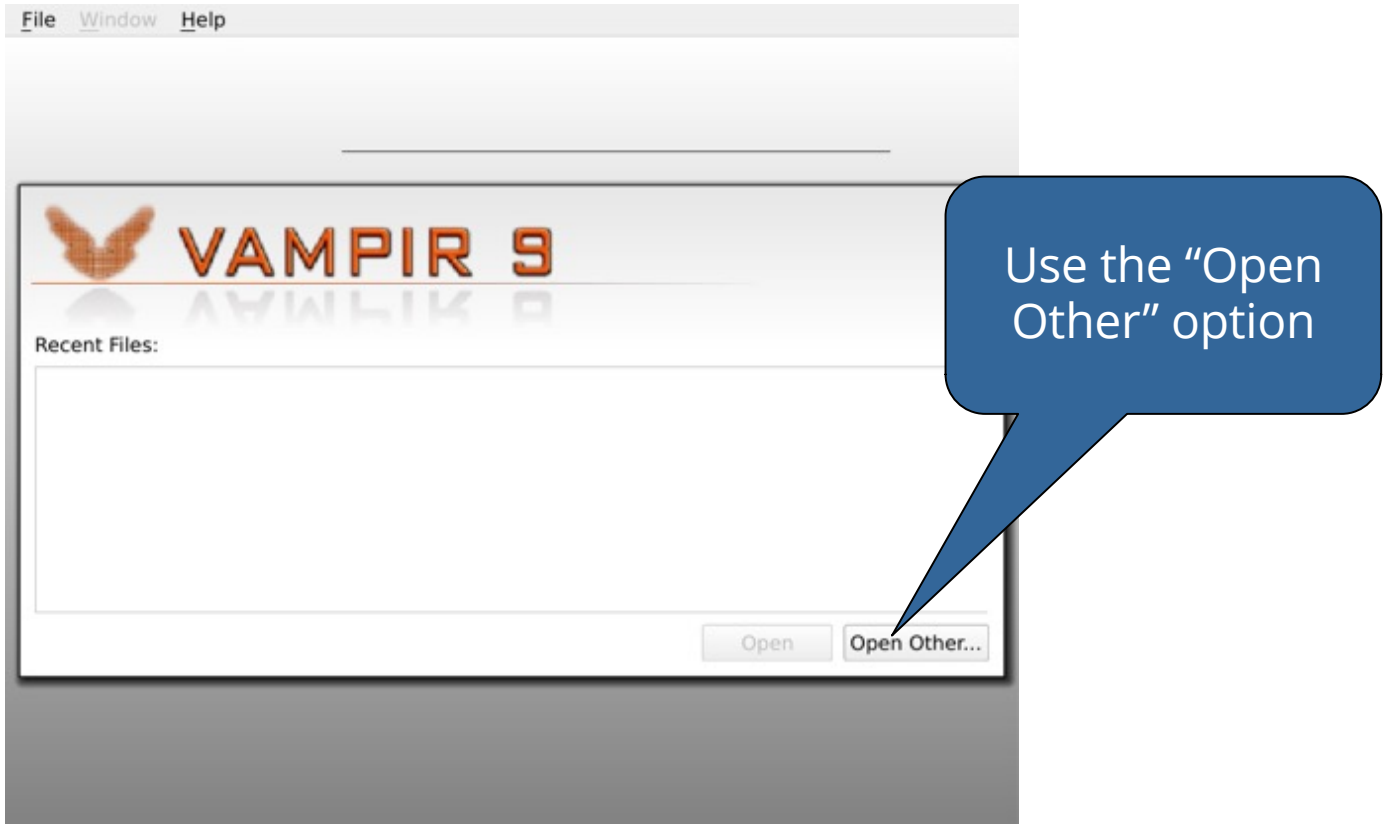
Please follow instructions on

- https://docs.olcf.ornl.gov/software/profiling/Vampir.html
- Note that a running VampirServer is protected by a password

```
$ vampirserver start -- -P <project>
Launching VampirServer...
Submitting LSF batch job (this might take a while)...
Warning: more than 1 task/rank assigned to a core
VampirServer 9.11.1 OLCF (4626dba5)
Licensed to ORNL
Running 4 analysis processes... (abort with vampirserver stop 29200)
User: <user>
Password: <password>
VampirServer <29200> listens on: h19n03:30035
```

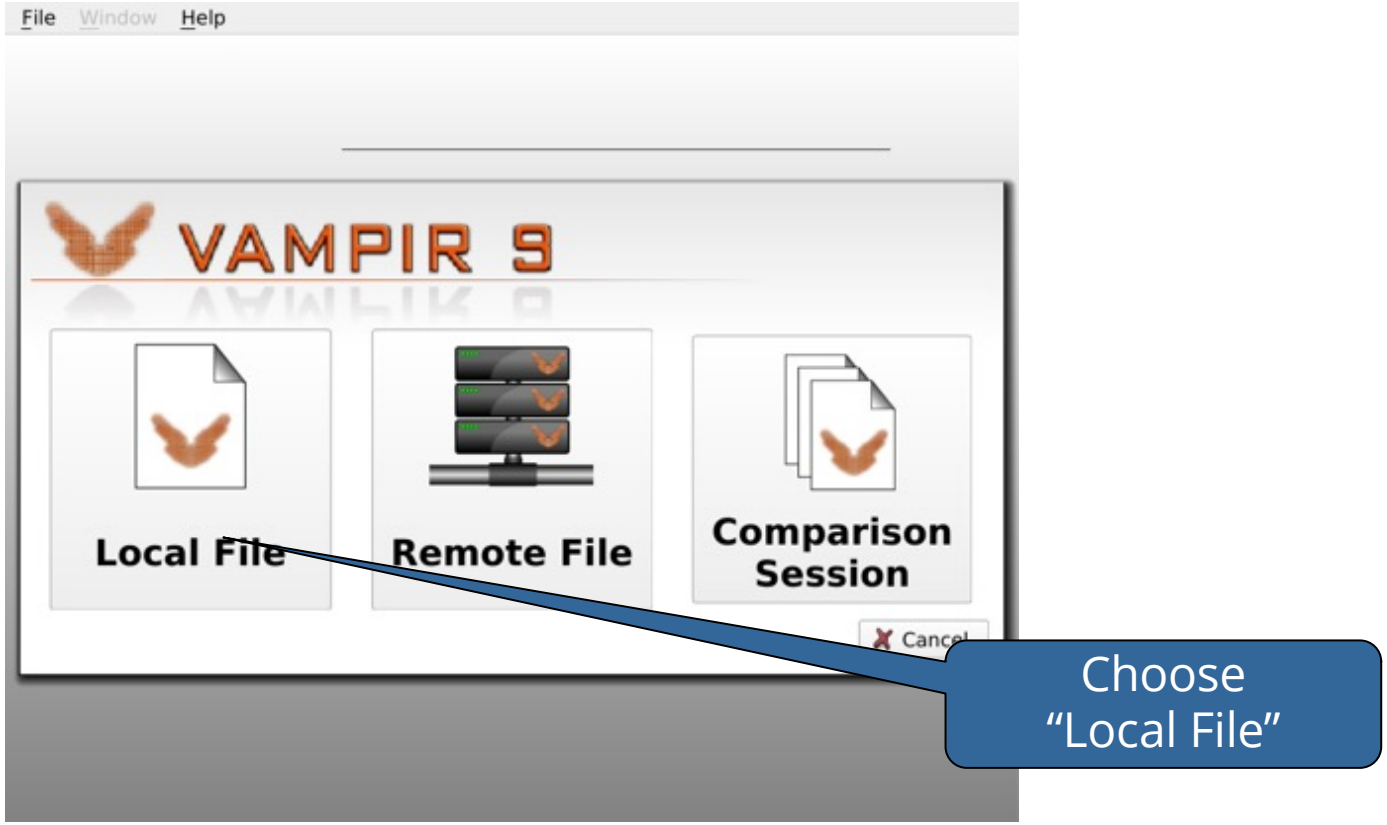
Vampir Hands on

Connect Client to Vampir Server:



Vampir Hands on

Connect Client to Vampir Server:



The main displays of Vampir

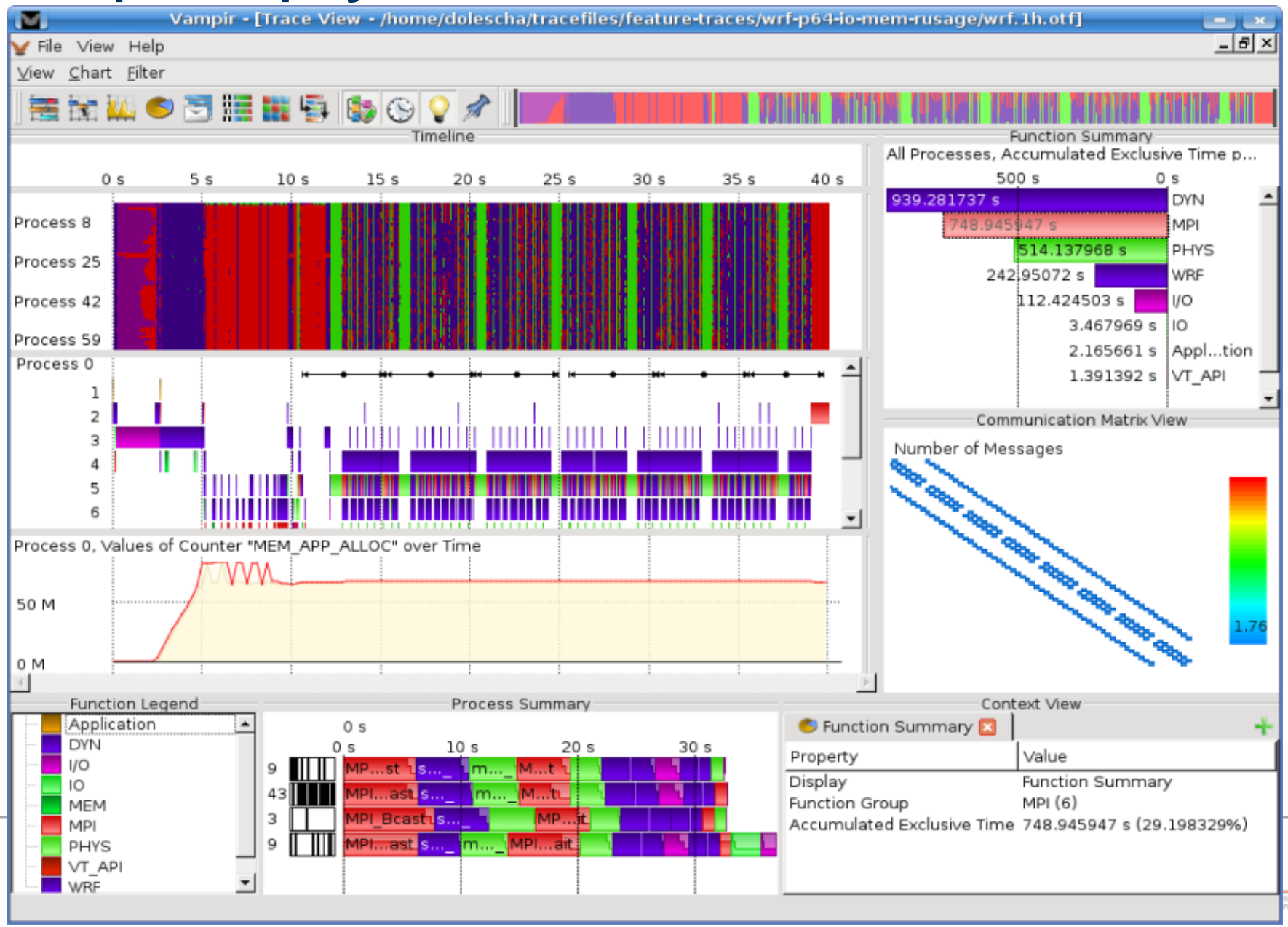
Timeline Charts:

- Master Timeline
- Process Timeline
- Counter Data Timeline
- Performance Radar

Summary Charts:

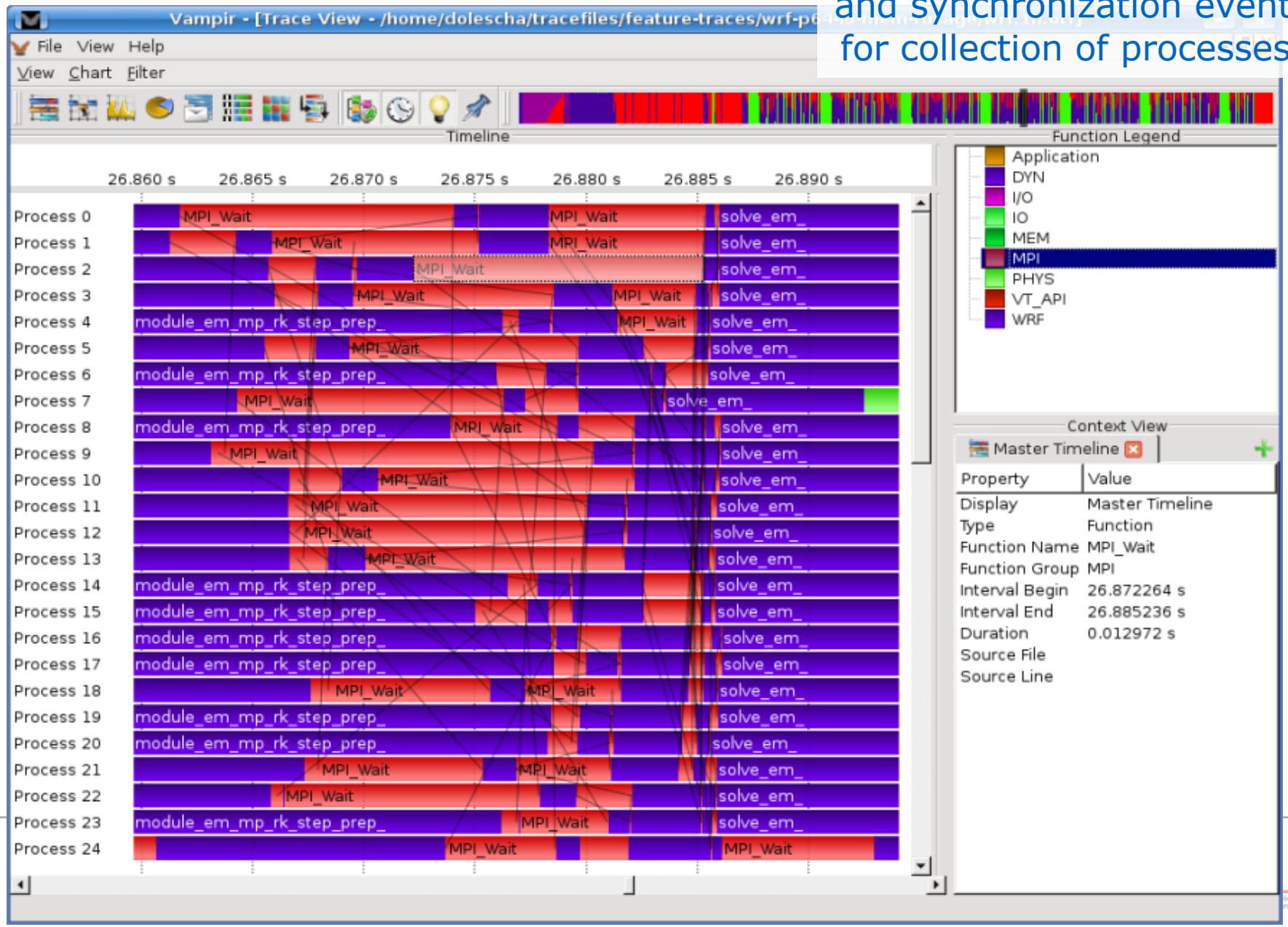
- Function Summary
- Message Summary
- Process Summary
- Communication Matrix View

Vampir: Displays for a WRF Trace with 64 Processes



Master Timeline

Detailed information about functions, communication and synchronization events for collection of processes.



Process and Counter Timeline

Detailed information about different levels of function calls in a stacked bar chart for an individual process.

Vampir - [Trace View - /home/dolescha/tracefiles/feature-traces/wrf-p64-...]

File View Help

View Chart Filter

Timeline

18.15 s 18.16 s 18.17 s 18.18 s 18.19 s 18.20 s

Process 0

2 3 4 5 6

MPI_Wait

Process 63

2 3 4 5 6

MPI_Wait

Process 0, Values of Counter "MEM_APP_ALLOC" over Time

50 M

0 M

Process 63, Values of Counter "ru_utime" over Time

750 k

0 k

Function Legend

- Application
- DYN
- I/O
- IO
- MEM
- MPI
- PHYS
- VT_API
- WRF

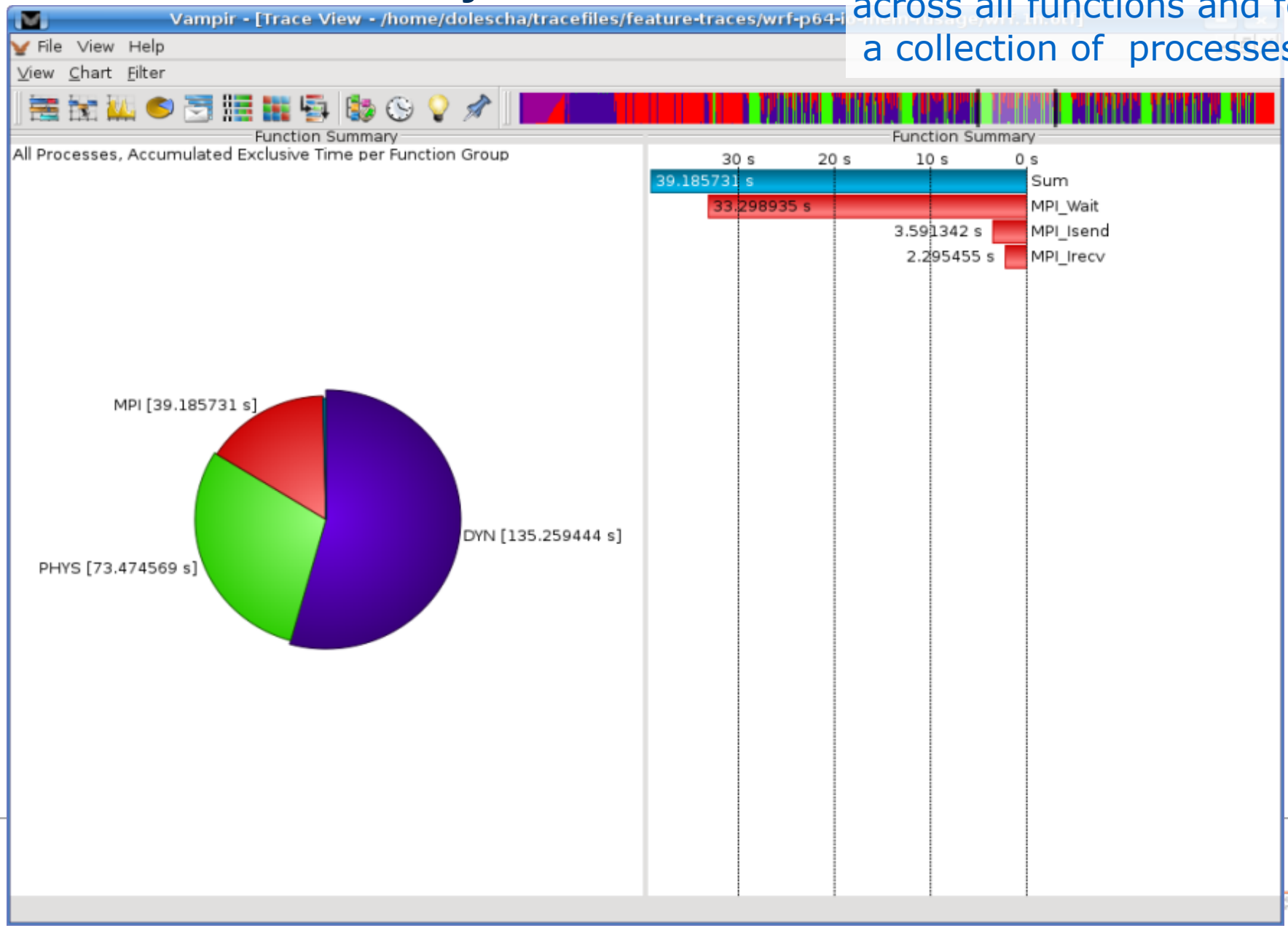
Context View

Process Timeline

Property	Value
Display	Process Timeline
Type	Function
Function Name	MPI_Wait
Function Group	MPI
Interval Begin	18.17866 s
Interval End	18.19058 s
Duration	0.01192 s
Source File	
Source Line	

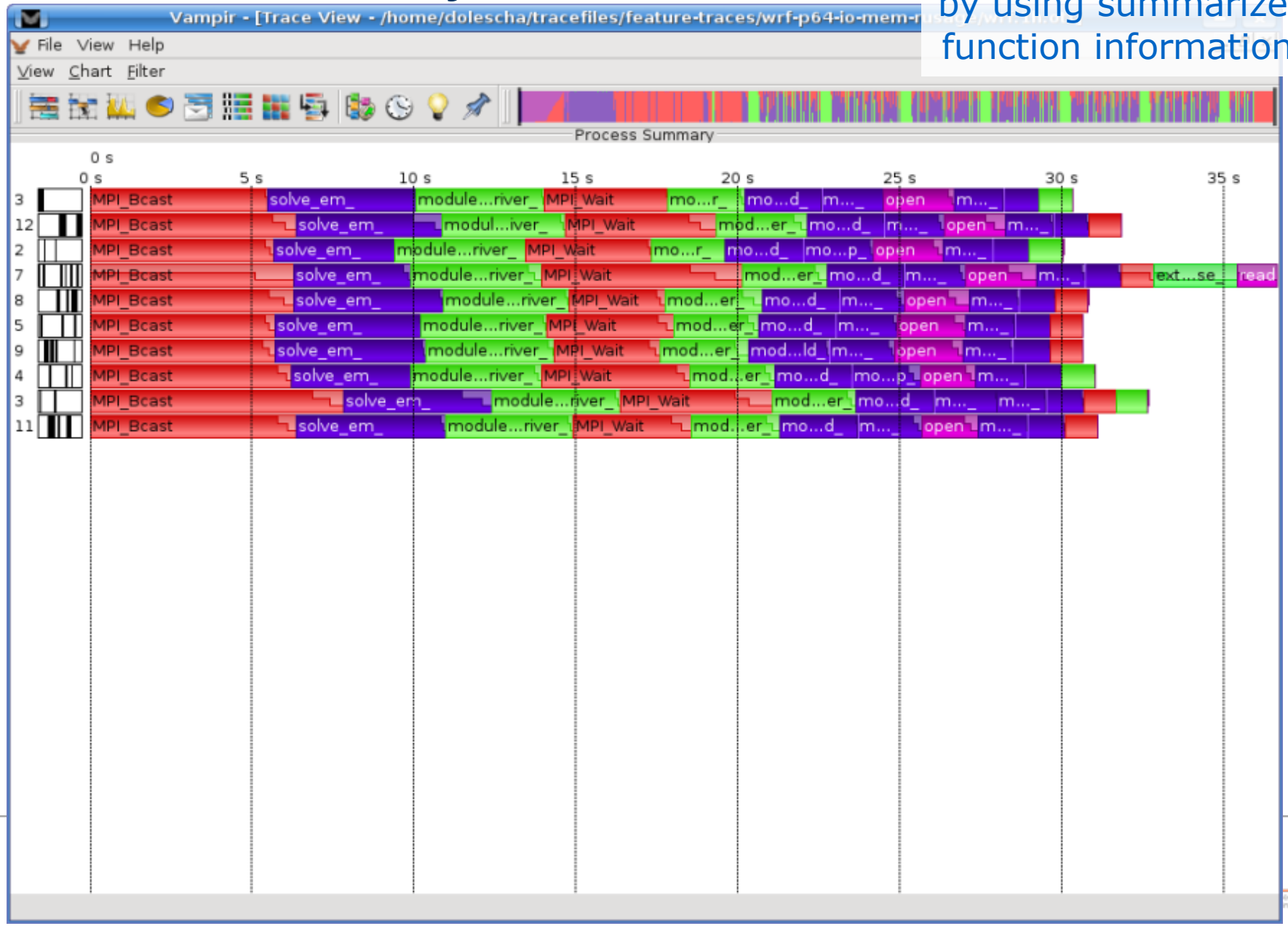
Function Summary

Overview of the accumulated information across all functions and for a collection of processes.



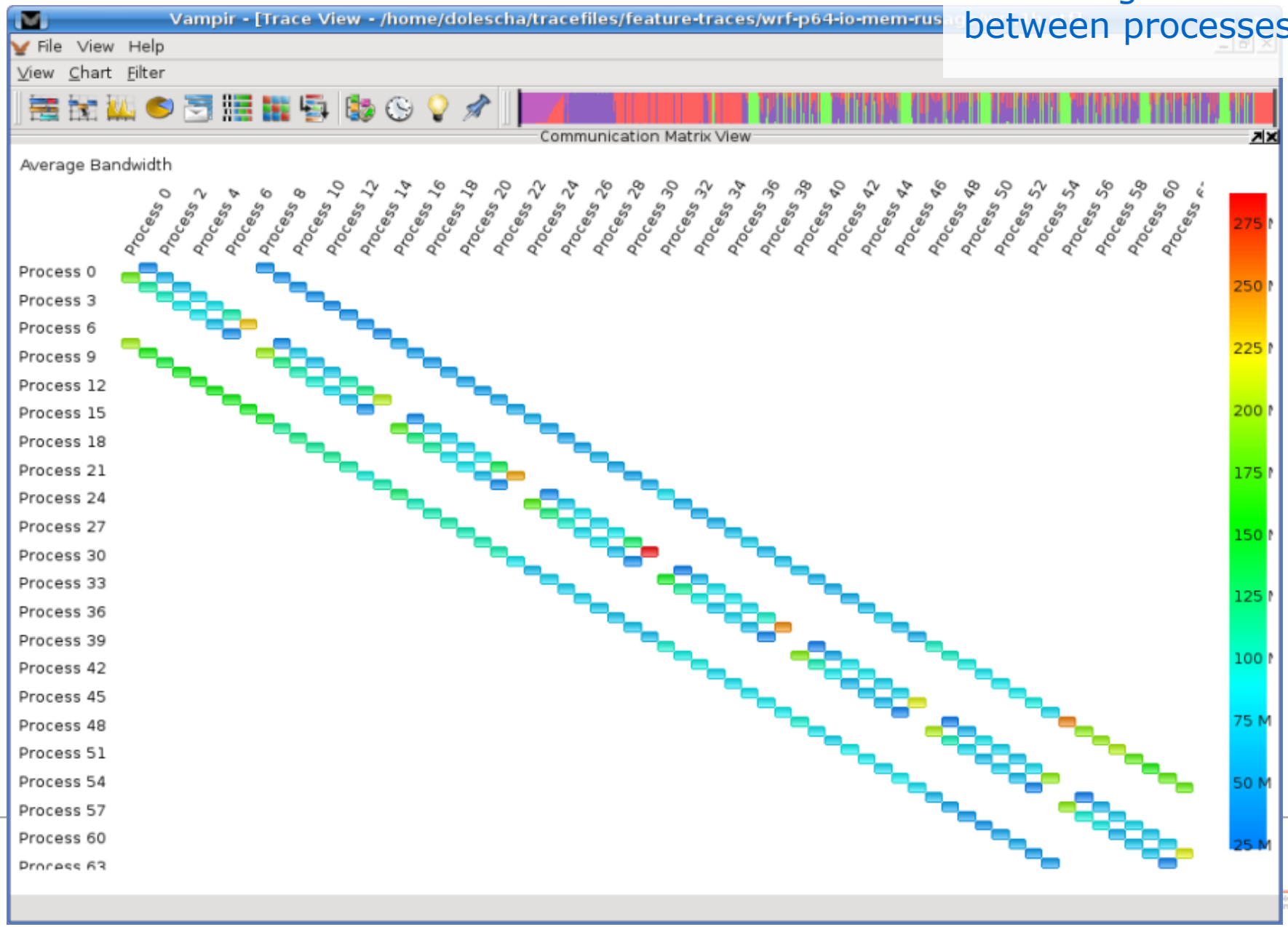
Process Summary

Find groups of similar processes and threads by using summarized function information.

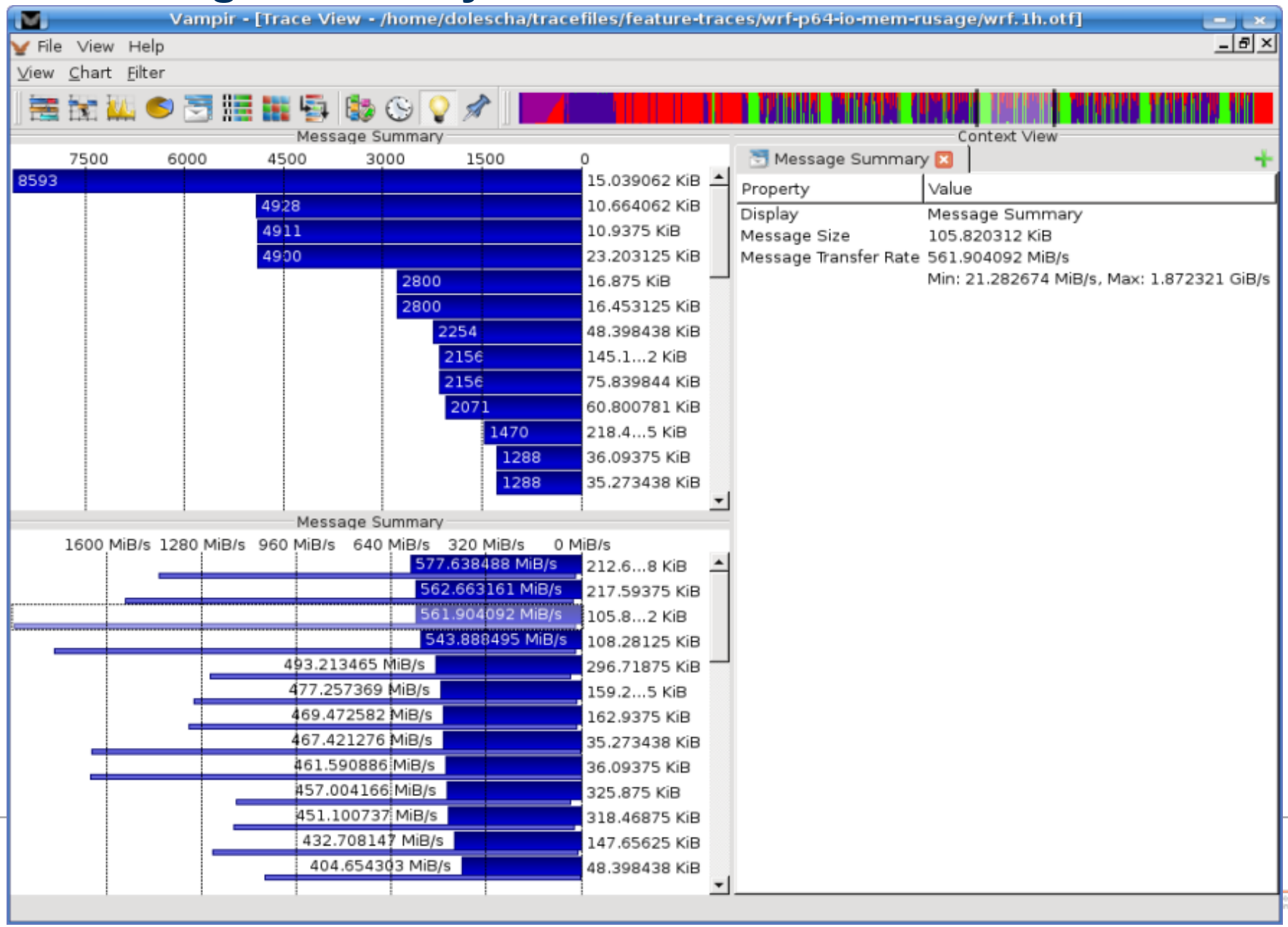


Communication Matrix

Information about messages sent between processes.

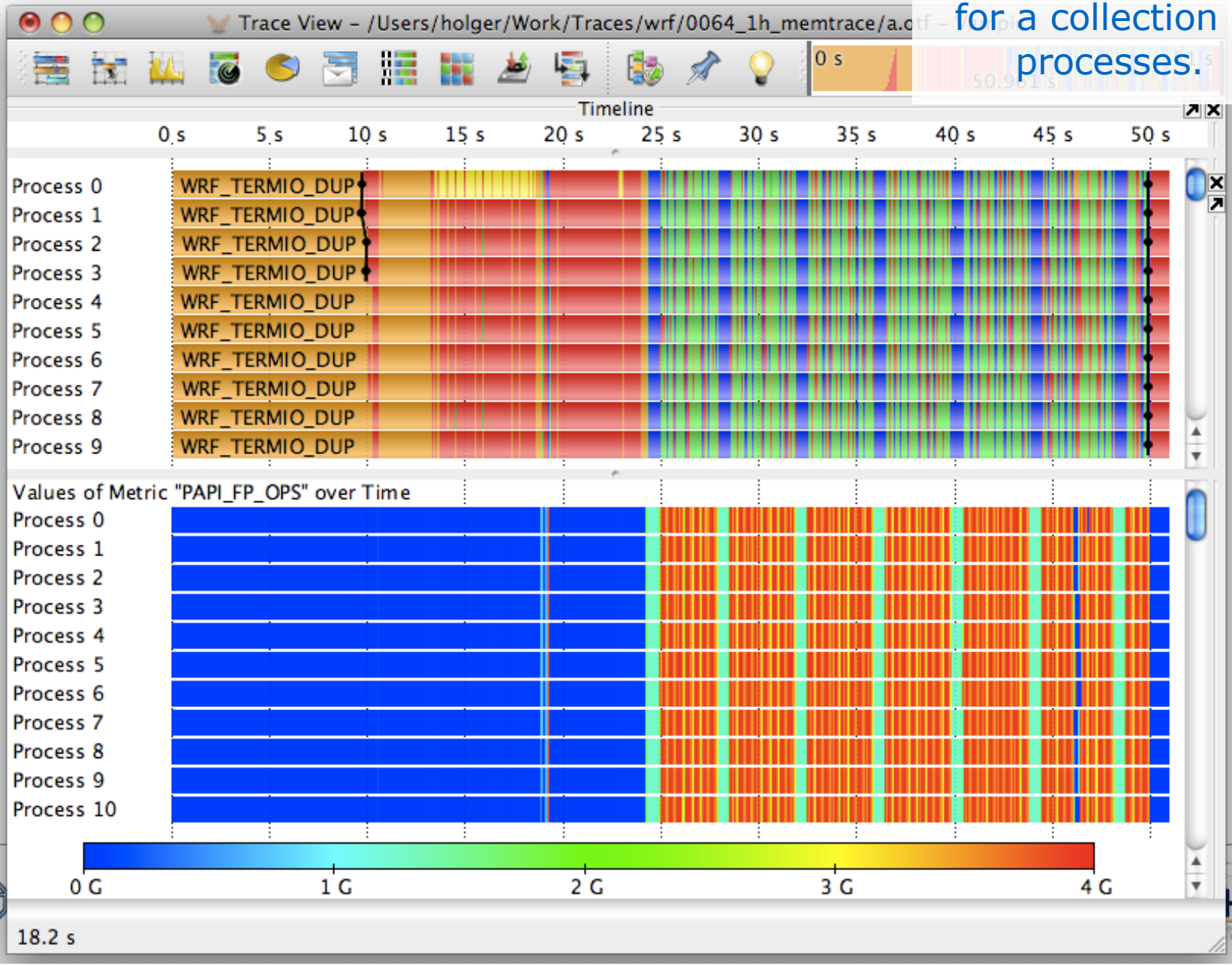


Message Summary



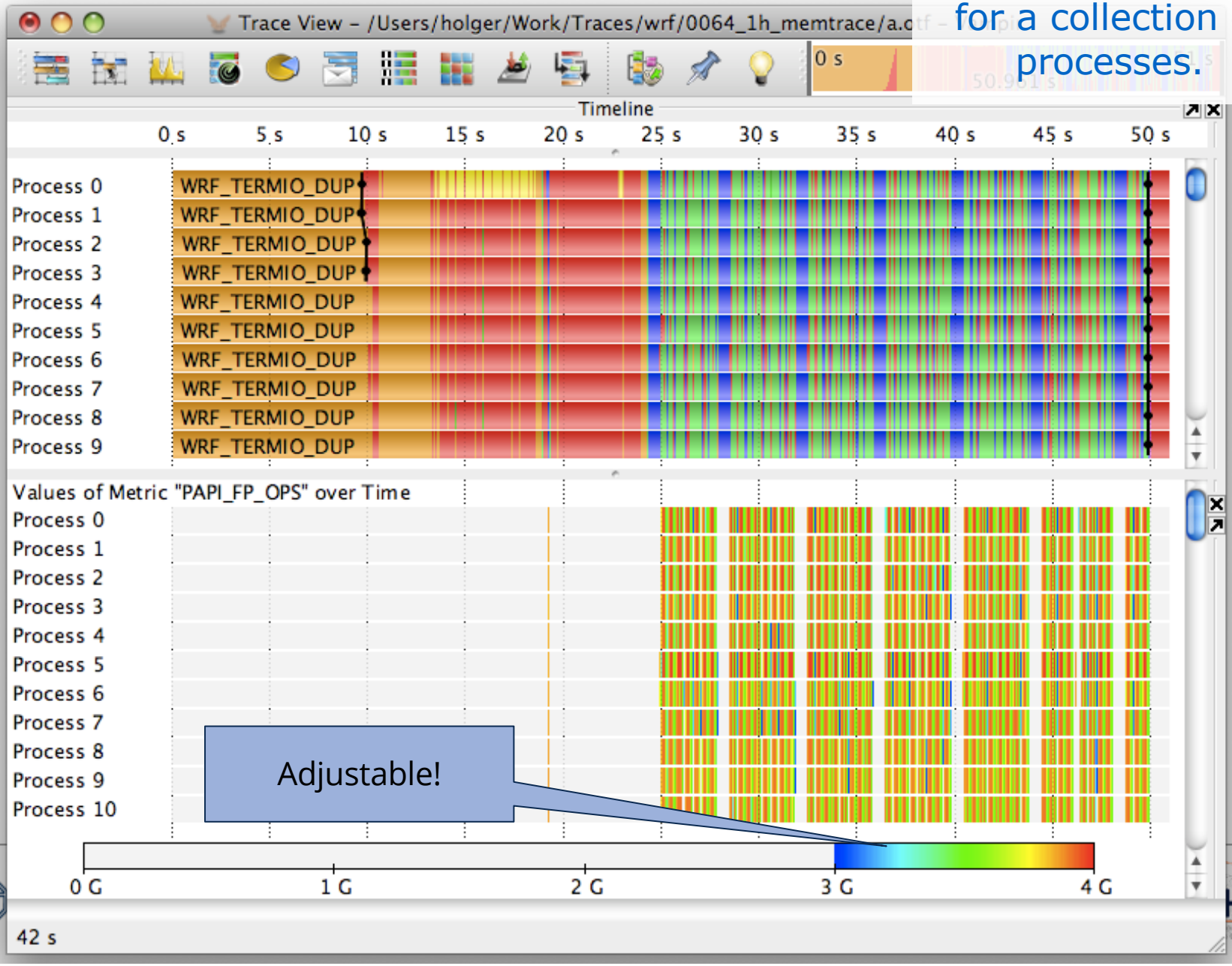
Performance Radar

Detailed counter information over time for a collection of processes.



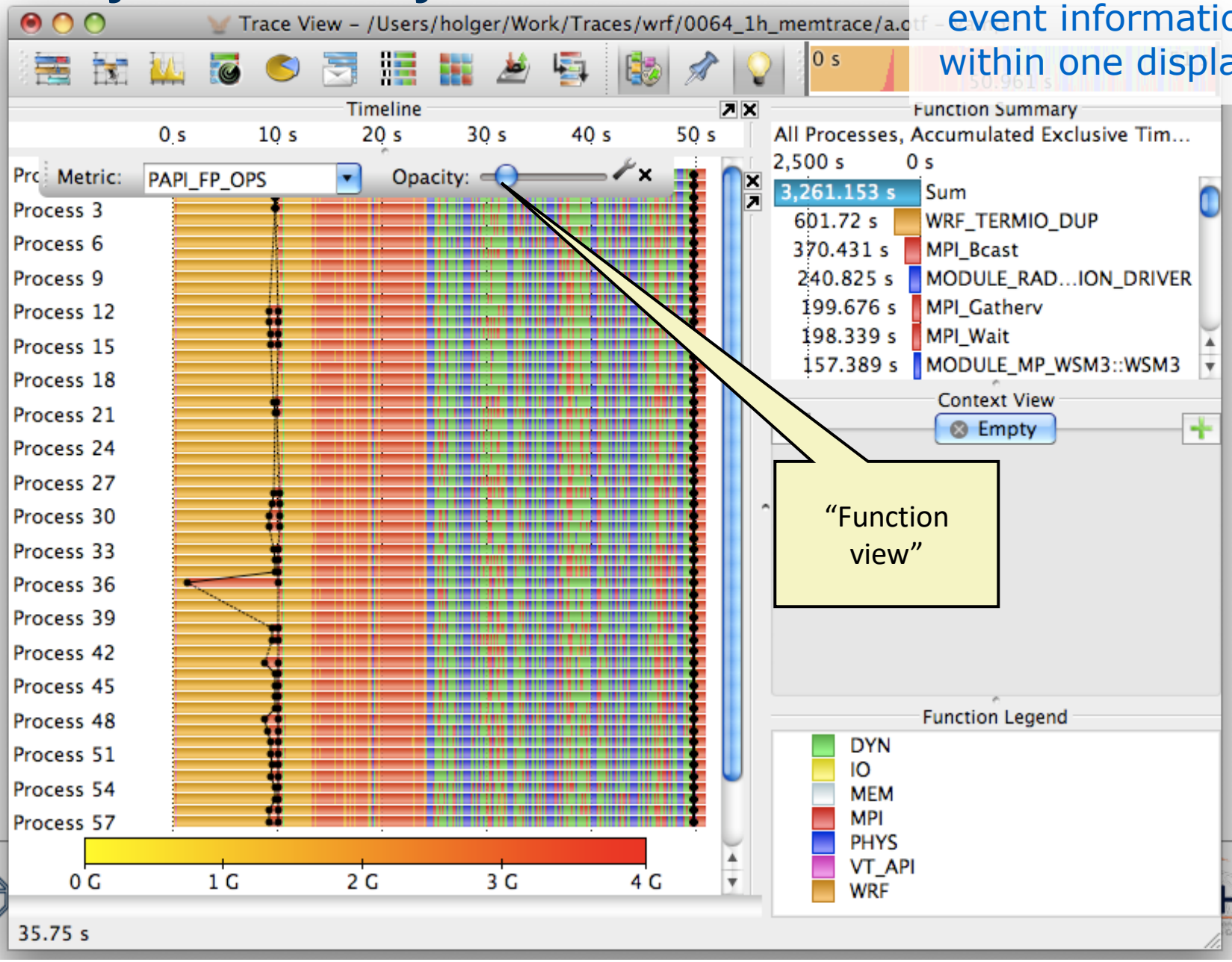
Performance Radar

Detailed counter information over time for a collection of processes.



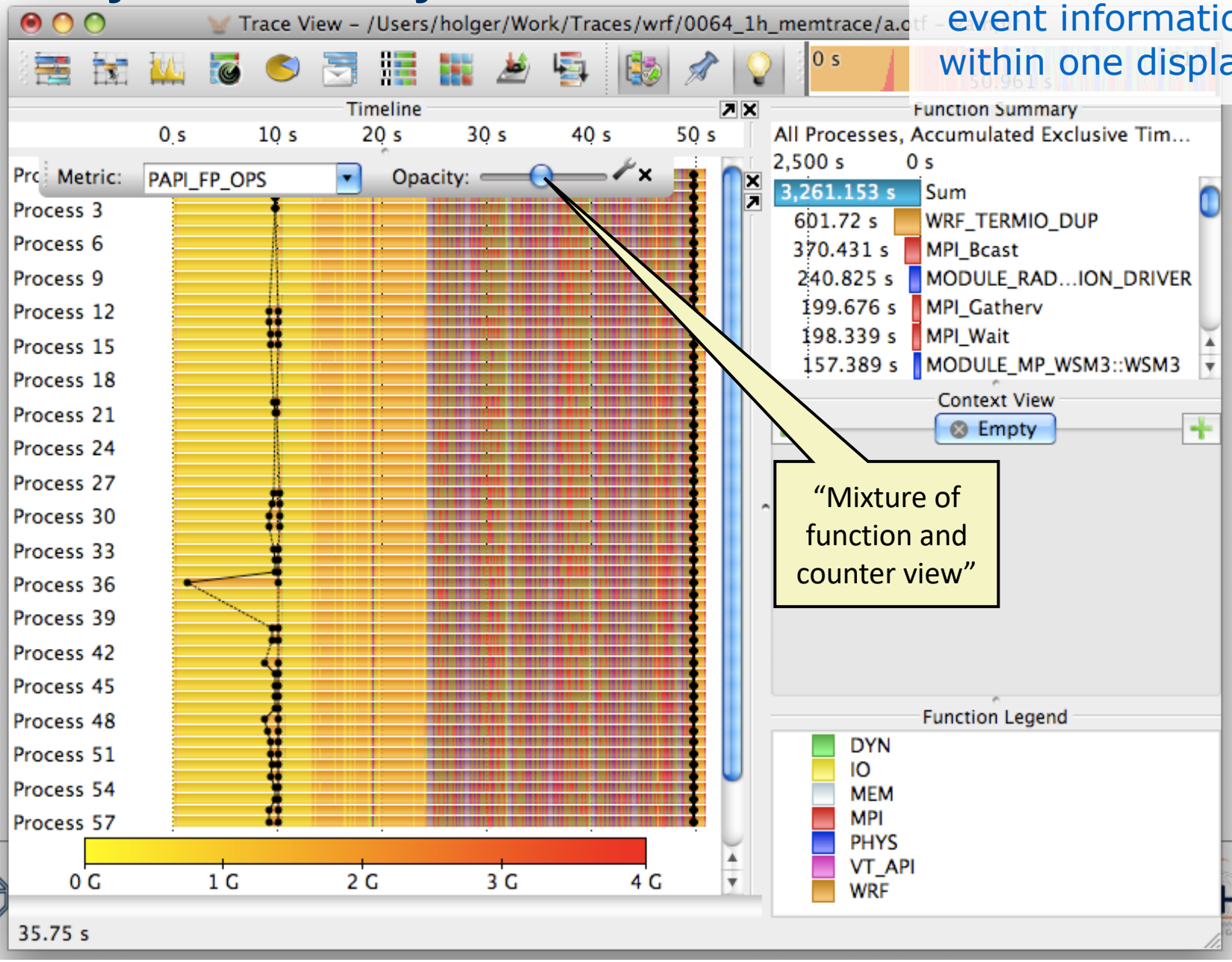
Overlay Functionality of Master Timeline

Correlate counter information with event information within one display.



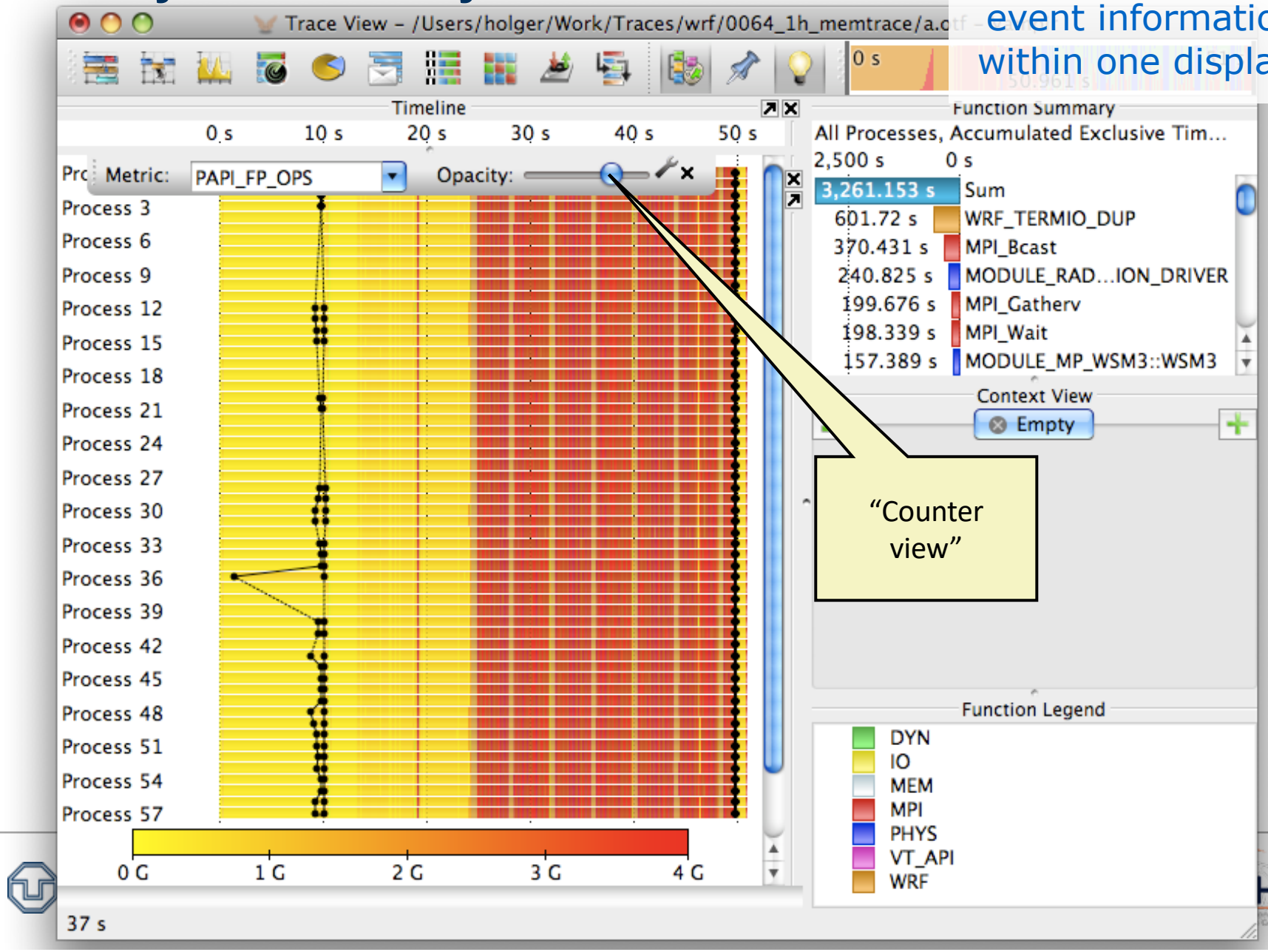
Overlay Functionality of Master Timeline

Correlate counter information with event information within one display.



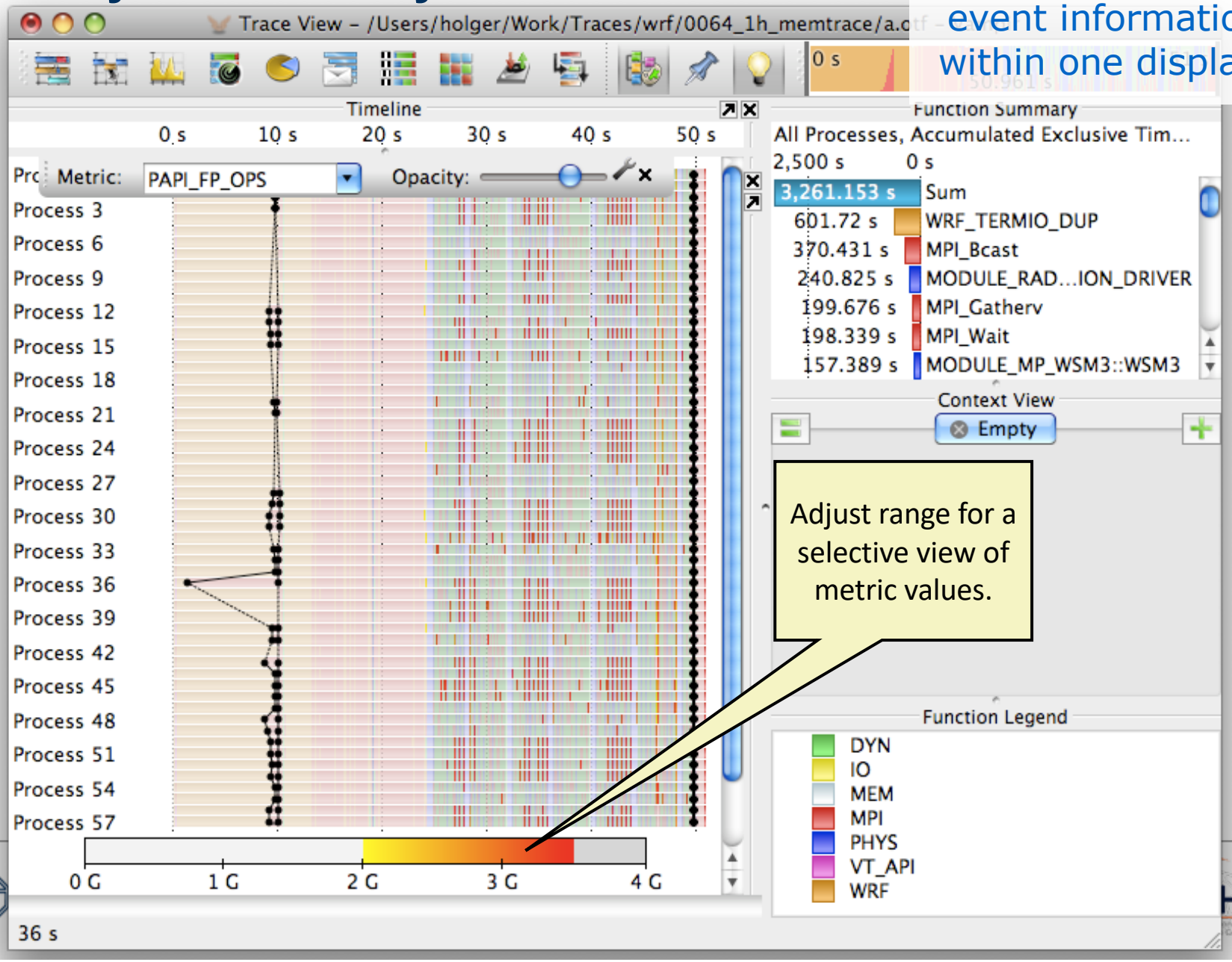
Overlay Functionality of Master Timeline

Correlate counter information with event information within one display.



Overlay Functionality of Master Timeline

Correlate counter information with event information within one display.



TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Center for Information Services and High Performance Computing (ZIH)

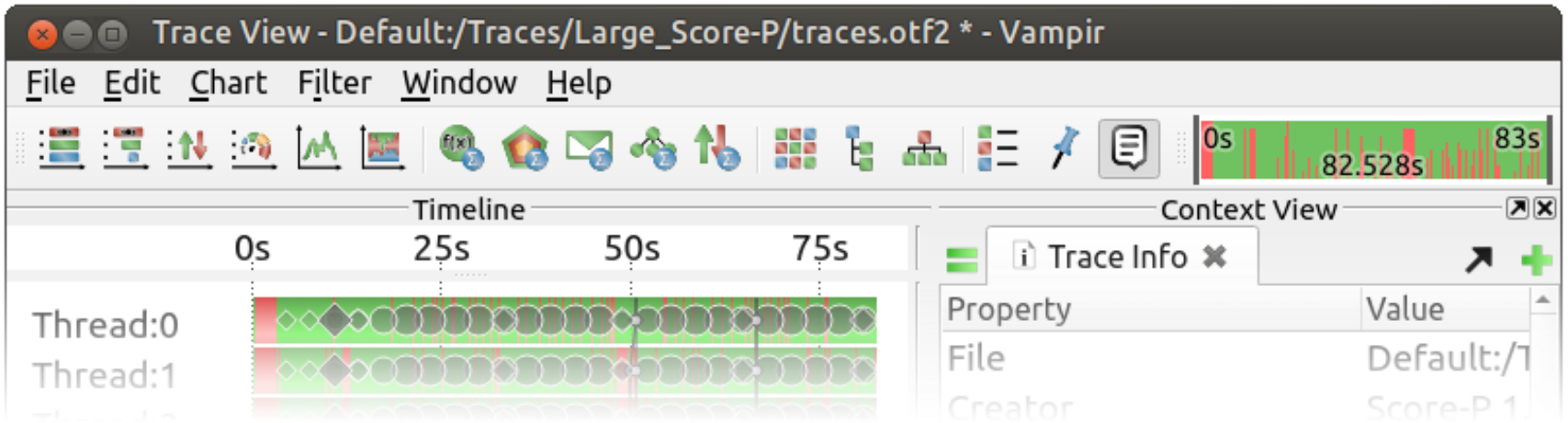
Vampir: What's New

What‘s New in Vampir 10

(released in Nov ’21)

What’s New Dialog when starting new version the first time

New Chart Icons



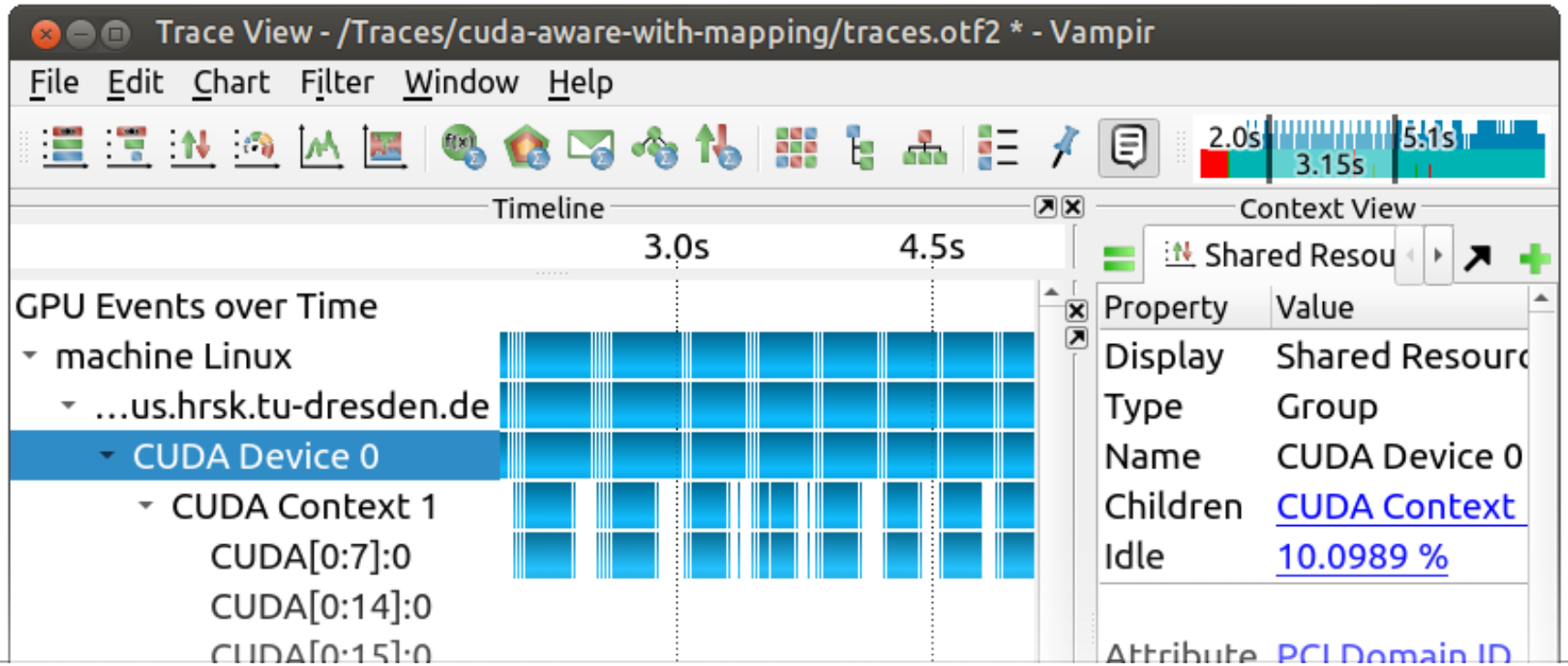
What's New in Vampir 10

(released in Nov '21)

Support new OTF2 3.0 features

- Accelerator devices and contexts
- Non-blocking collective operations

No Score-P release yet

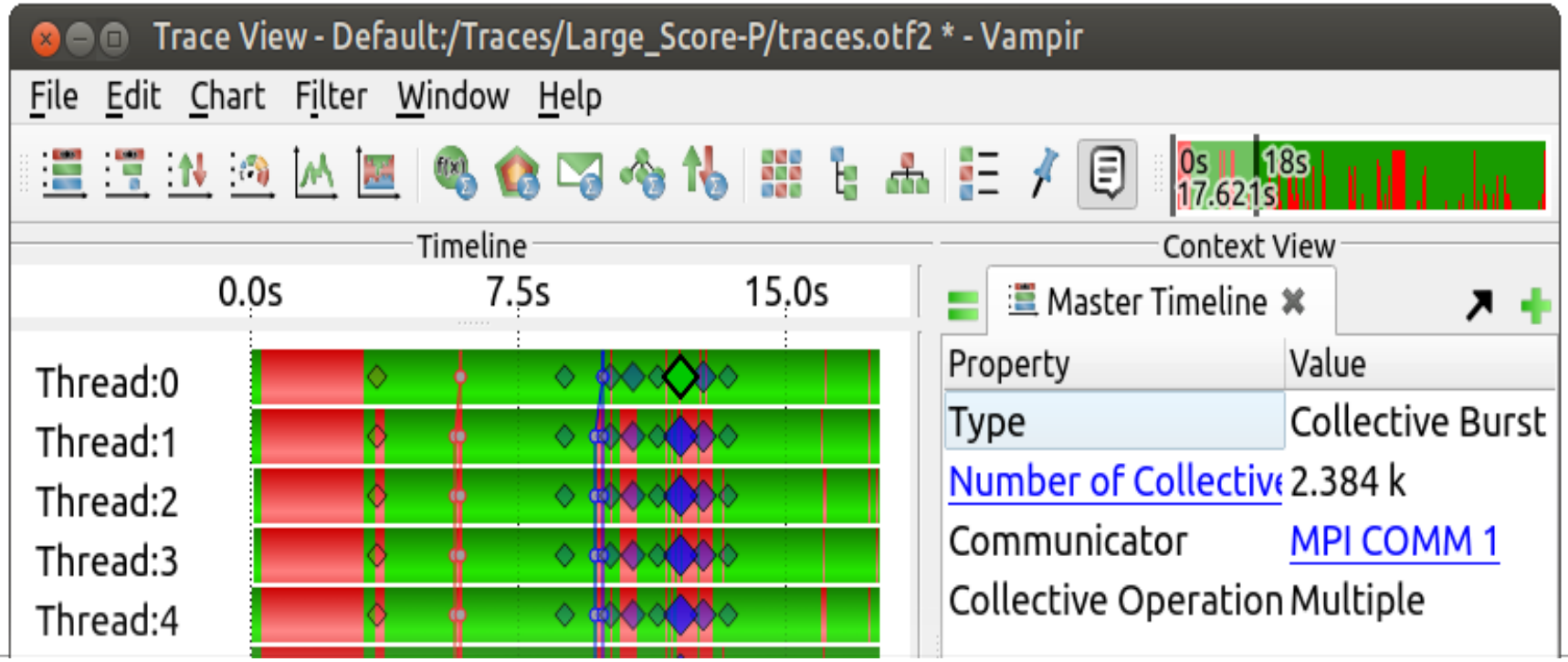


What's New in Vampir 10

(released in Nov '21)

Collective Operation Bursts

- Collated into bursts, when too many of them occur in a specific interval
- Collectives also visible if beginning is not in the zoom interval

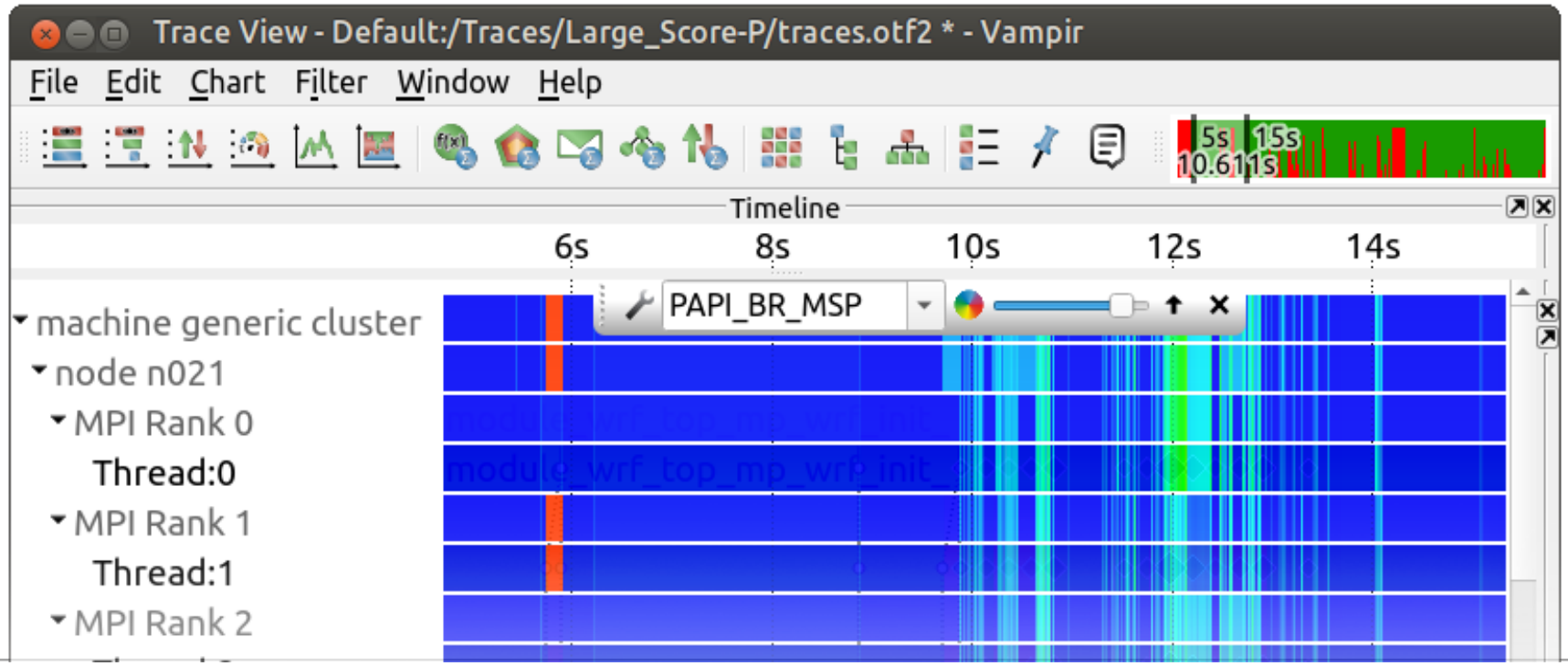


What's New in Vampir 10

(released in Nov '21)

Summarizations in the Performance Overlay

— Values are summarized into their parent groups



TECHNISCHE UNIVERSITÄT DRESDEN

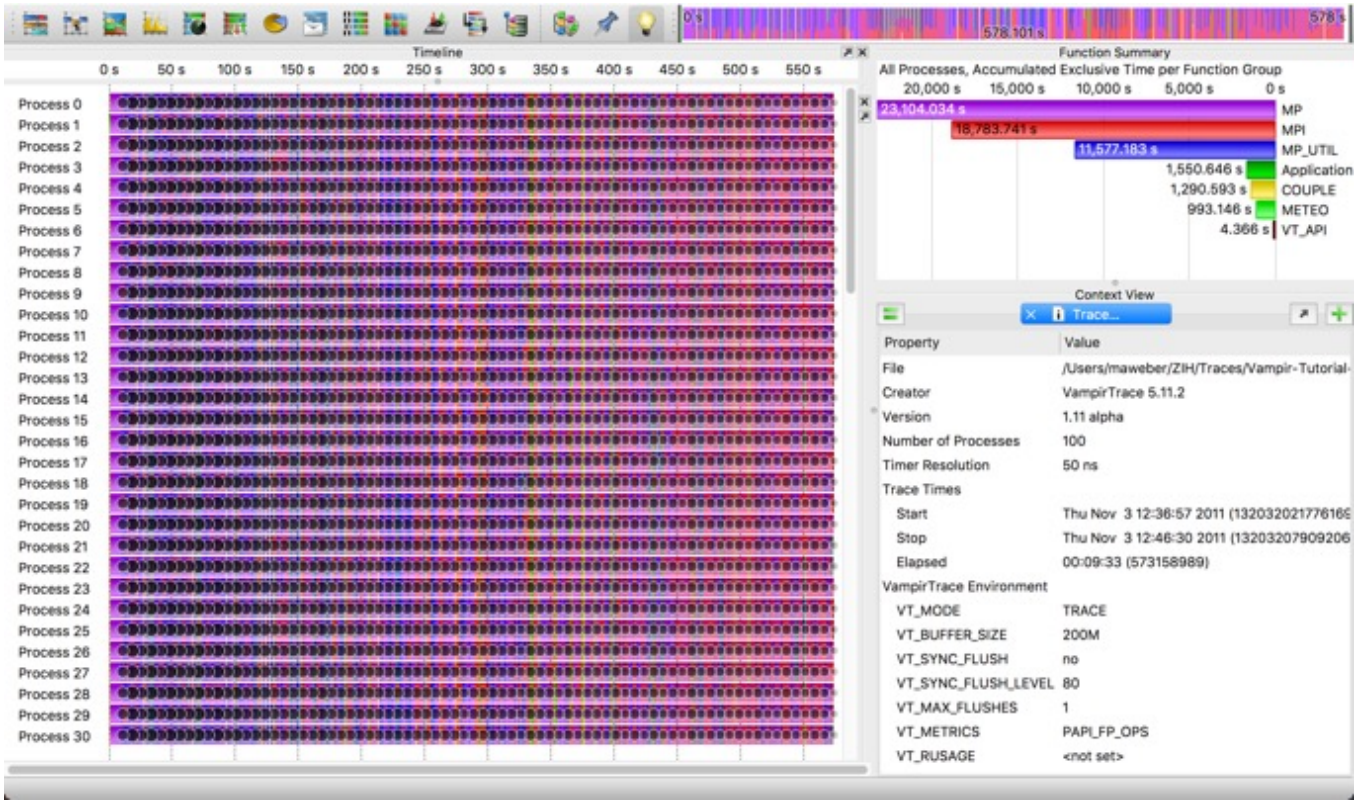
DRESDEN concept

Center for Information Services and High Performance Computing (ZIH)

Vampir Case Study: Analyzing Load Imbalance in COSMO-SPECS

ZIH
Center for Information Services & High Performance Computing

COSMO-SPECS



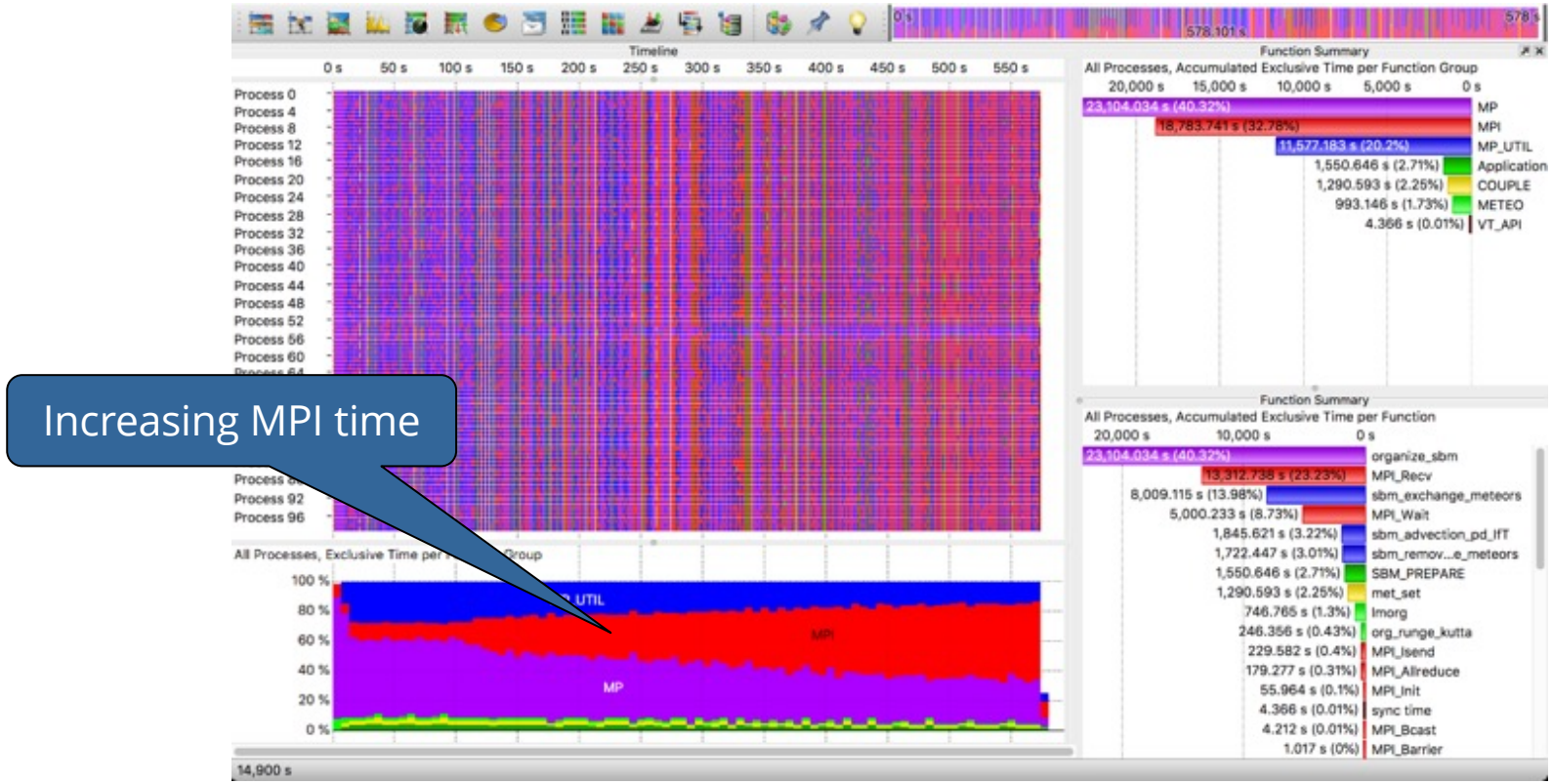
Weather forecast code `COSMO-SPECS` with 100 processes

COSMO: weather model (`METEO` group)

SPECS: microphysics for accurate cloud calculation (`MP` and `MP_UTIL` group)

Coupling of both models is done in `COUPLE` group

COSMO-SPECS

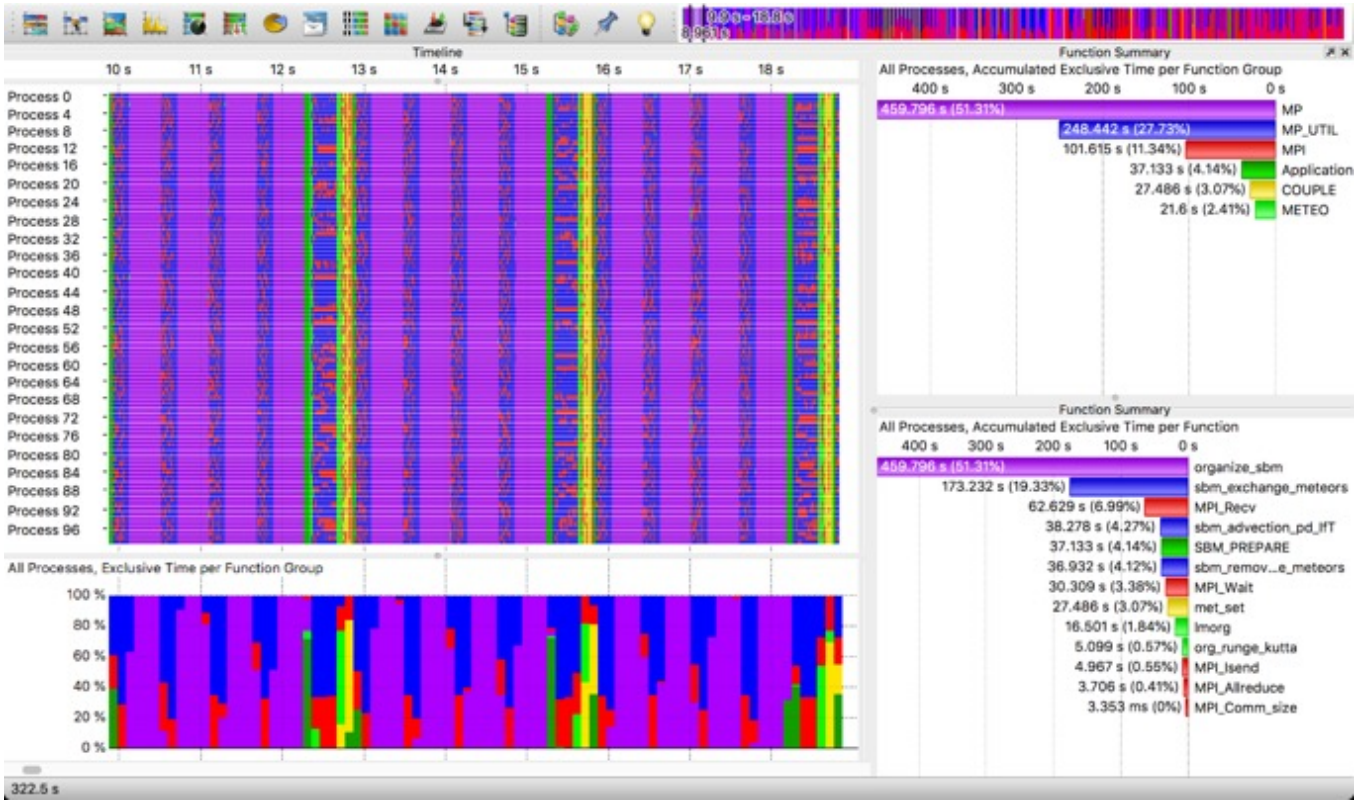


Compared to `METEO`, `MP` and `MP_UTIL` are very compute intensive, however this is due to more complex calculations and no performance issue

Problem: >32% of time spent in MPI

MPI runtime share increases throughout the application run

COSMO-SPECS



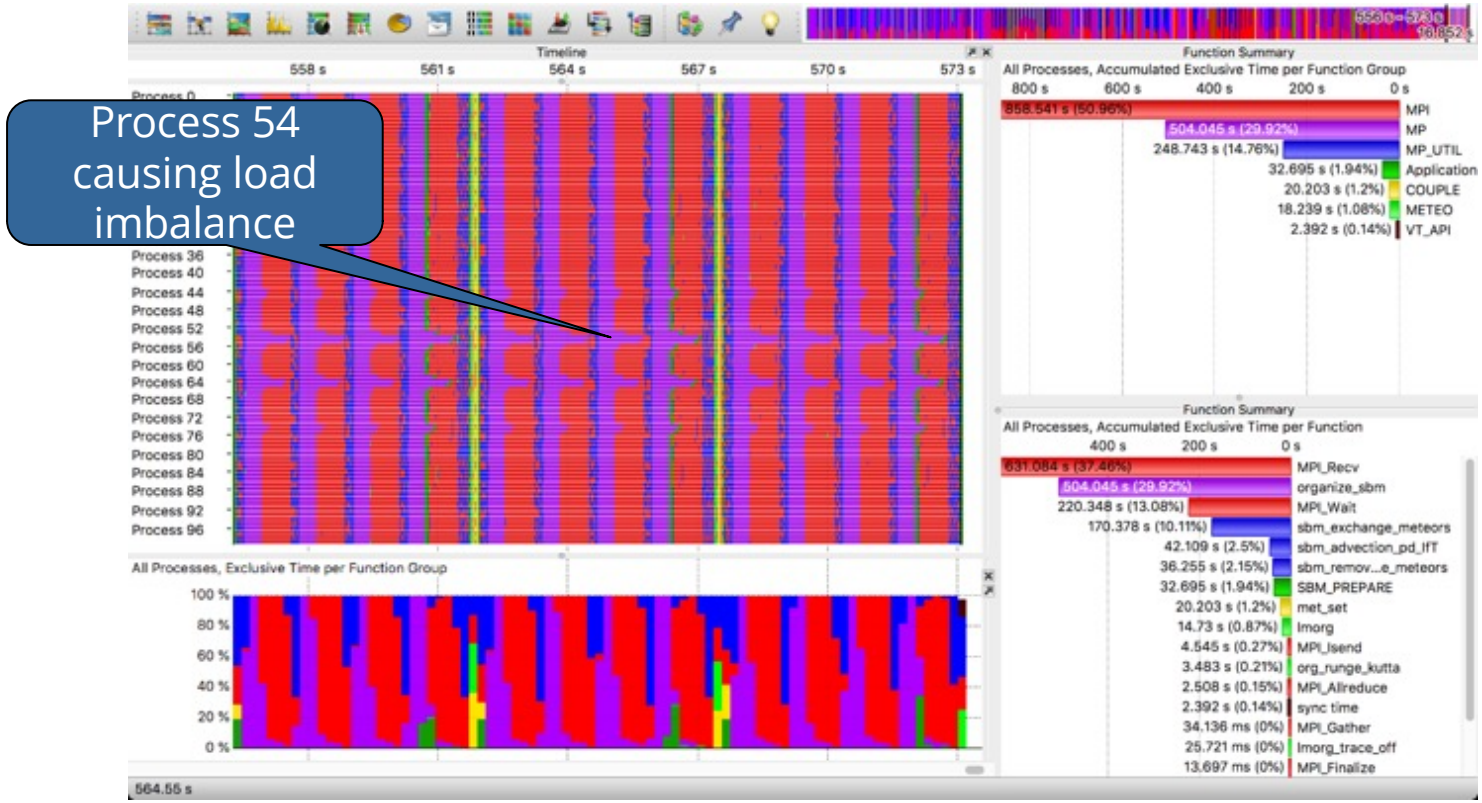
Zoom into the first three iterations

`MP`/`MP_UTIL` perform four sub-steps in one iteration

Low MPI time share

Everything is balanced and looks okay

COSMO-SPECS

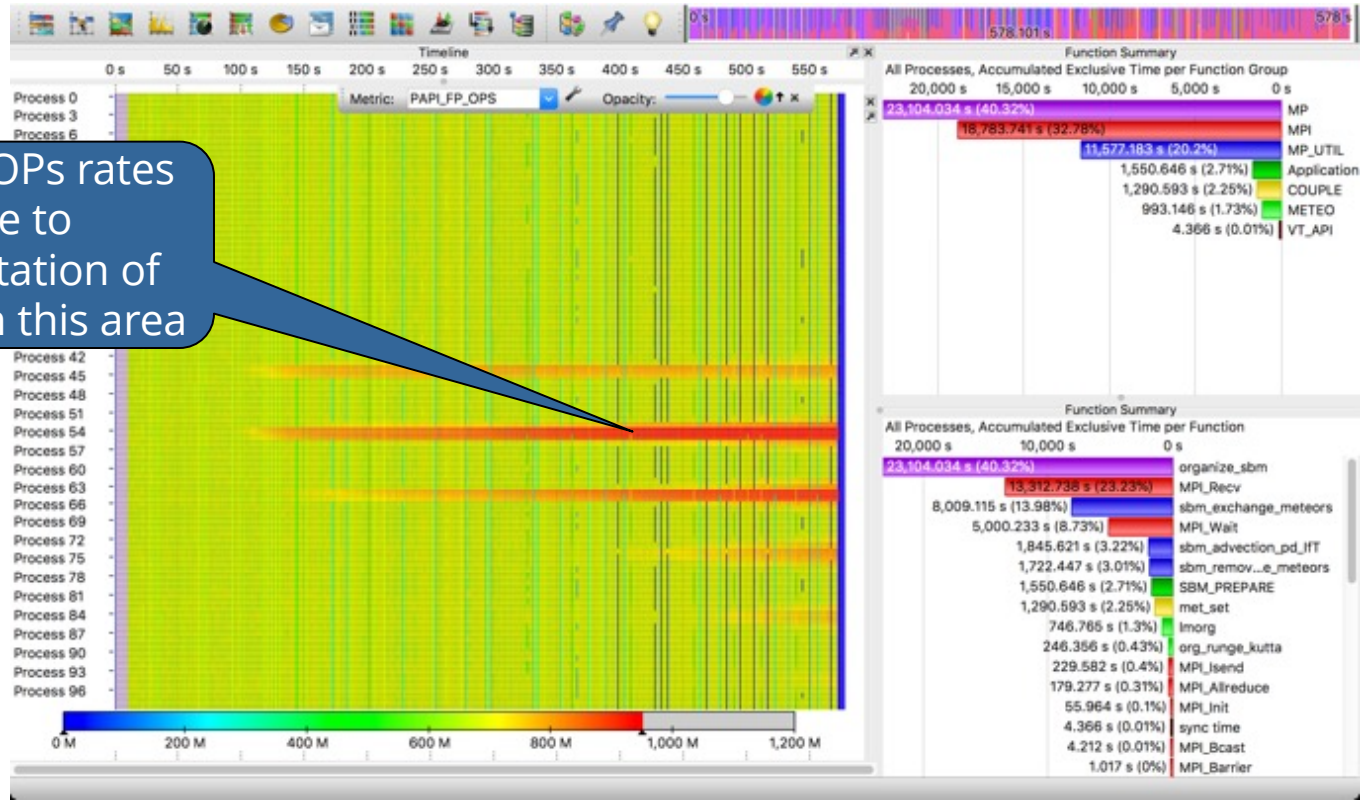


Zoom into the last three iterations

Very high MPI time share (>50%)

Large load imbalance caused by MP functions around **`Process 54`** and **`Process 64`**

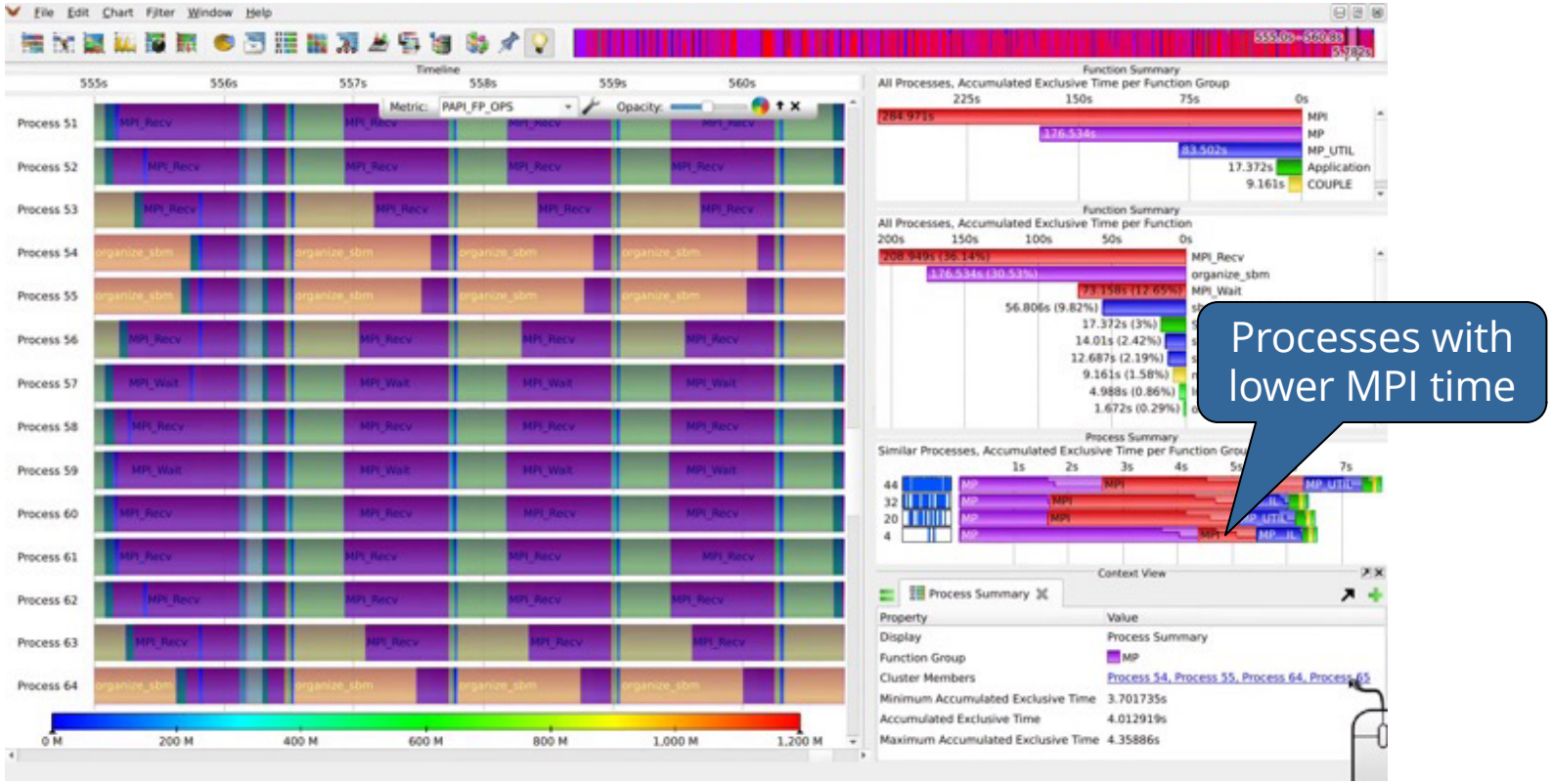
COSMO-SPECS



PAPI_FP_OPS counter showing higher FLOPs rates on processes causing the imbalance

Reason for imbalance: Static grid used for distribution of processes.
Depending on the weather, expensive cloud computations (**MP** group) may be only necessary on some processes

COSMO-SPECS



Process Summary helps finding outliers

Groups processes by their behavior (similar call/duration profile)

Number of expected groups is variable

In this case 4 yields the best results

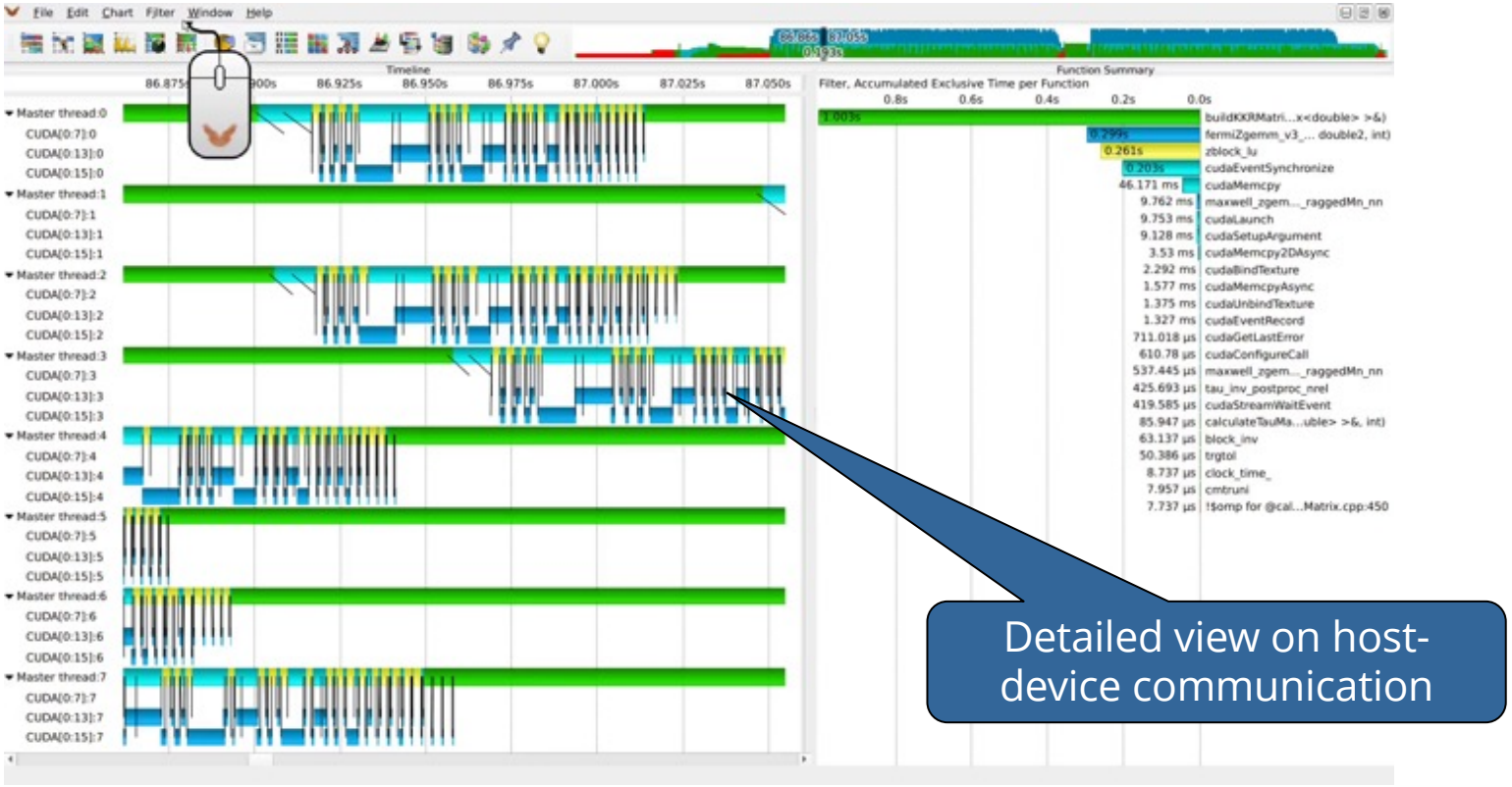
TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Center for Information Services and High Performance Computing (ZIH)

Vampir Showcase: Analyzing CUDA Applications

CUDA Application



Material science code LSMS

CUDA is utilized for heavy computations

CUDA streams are children of the owning process

Allows an in-depth analysis of host-device communication

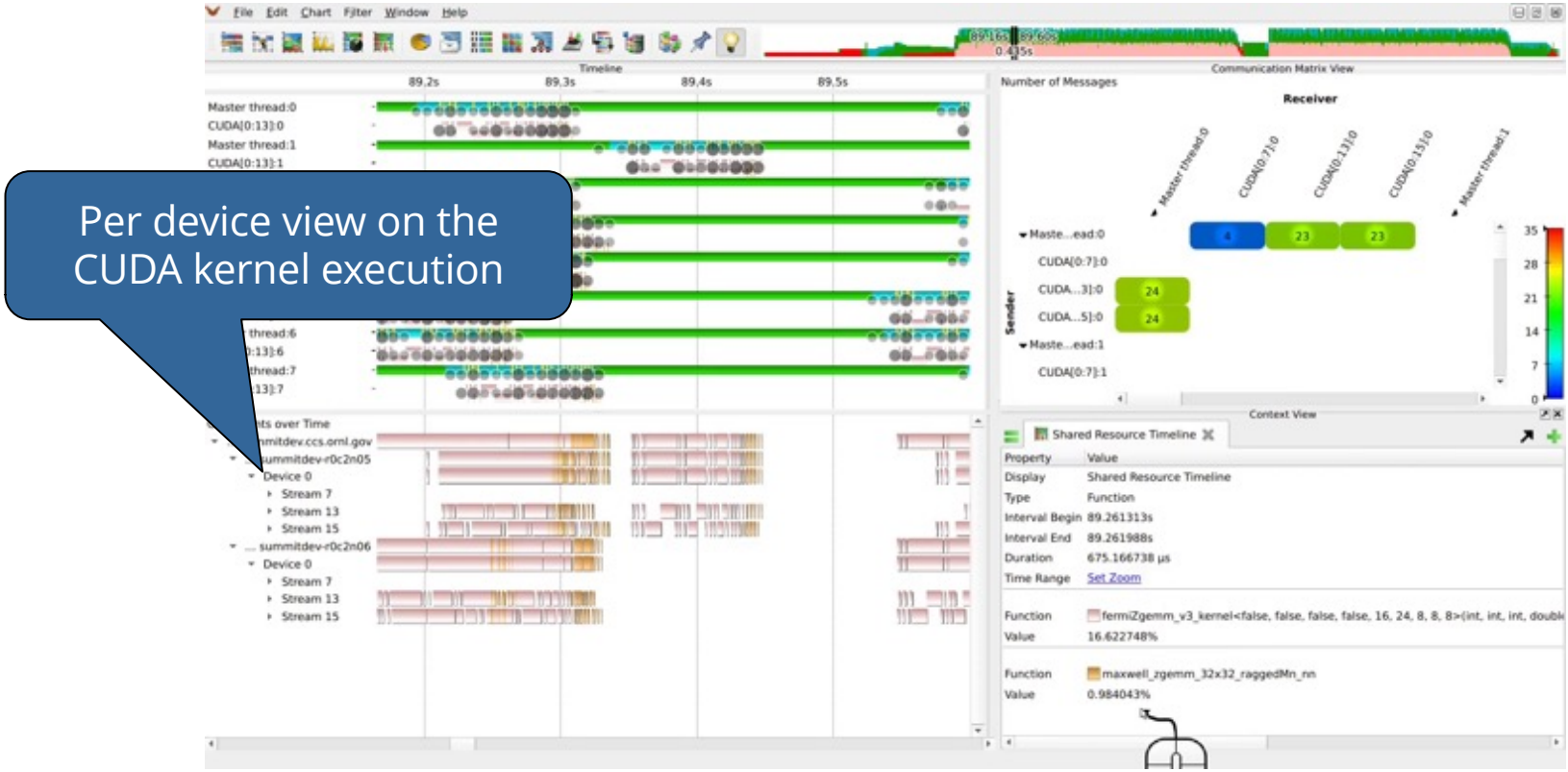
CUDA Application

Communication Matrix best for analyzing the general communication pattern

Expectation: balanced communication, represented by a symmetric matrix

Problem: communication with stream 7 is different

CUDA Application



Shared Resource Timeline offers a per device view on kernel executions

Best suited for analyzing multi-GPUs per node scenarios

Allows a dedicated analysis of kernel execution patterns

Yields insights of the actual hardware usage

TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Center for Information Services and High Performance Computing (ZIH)

Vampir Showcase: Analyzing Multilayer File I/O Applications

File I/O recording

Omnipresent in todays HPC applications

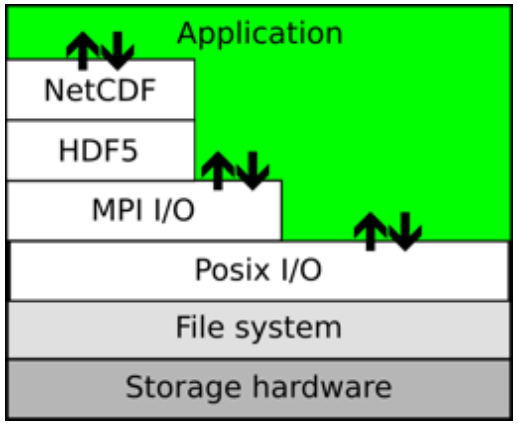
Record interaction between multiple layers

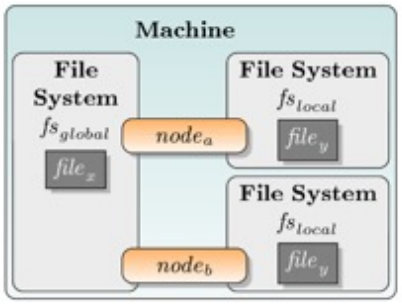
- MPI I/O (`MPI_File_open`)
- ISO C I/O (`fopen`)
- POSIX I/O (`open`, interface to OS)

System tree information determine whether file resides in a shared filesystem

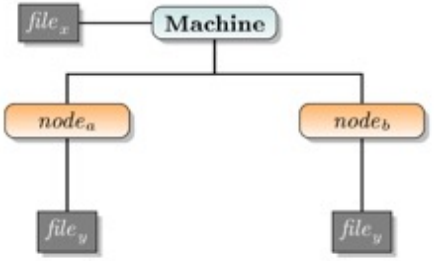
High level of detail

=> Trace data might increase dramatically





(a) Hardware topology



(b) System tree representation

NetCDF & HDF5 not supported yet

B_EFF I/O

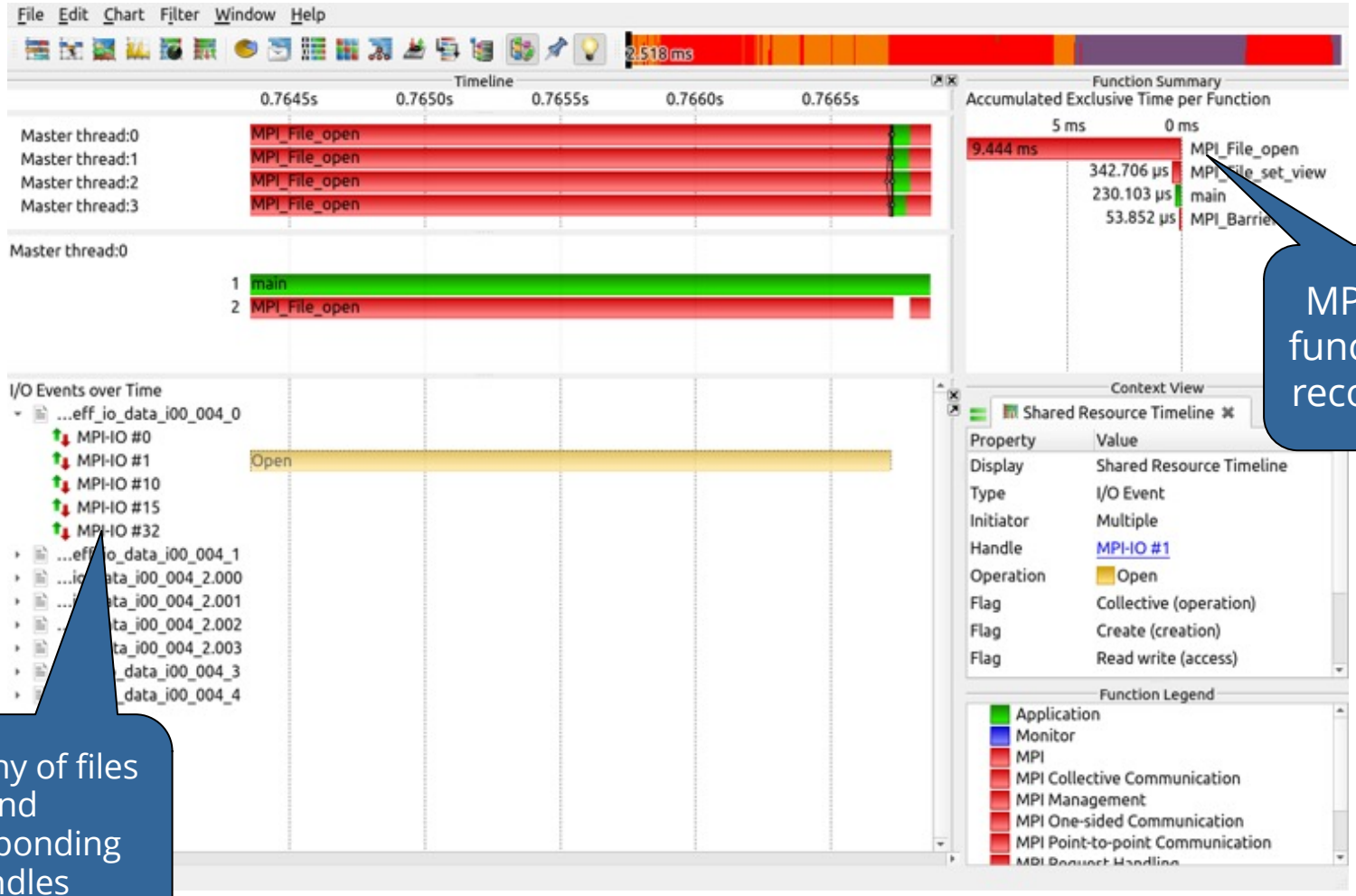
MPI I/O benchmark

fs.hlrs.de/projects/par/mpi//b_eff_io/

MPI I/O instrumentation is enabled by default by Score-P

```
% scorep-mpicc -o b_eff_io b_eff_io.c
% export SCOREP_EXPERIMENT_DIRECTORY=scorep-b_eff_io-4-profile
% mpirun -n 4 -c 6 ./b_eff_io -MB 2048 -MT 98304 -rewrite -N 4 -T 60
% scorep-scorep –g scorep-b_eff_io-4-profile/profile.cubex
% export SCOREP_EXPERIMENT_DIRECTORY=scorep-b_eff_io-4-tracing
% export SCOREP_FILTERING_FILE=initial_scorep.filter
% export SCOREP_ENABLE_TRACING=true
% export SCOREP_TOTAL_MEMORY=31MB
% mpirun -n 4 -c 6 ./b_eff_io -MB 2048 -MT 98304 -rewrite -N 4 -T 60
```

Result visualization



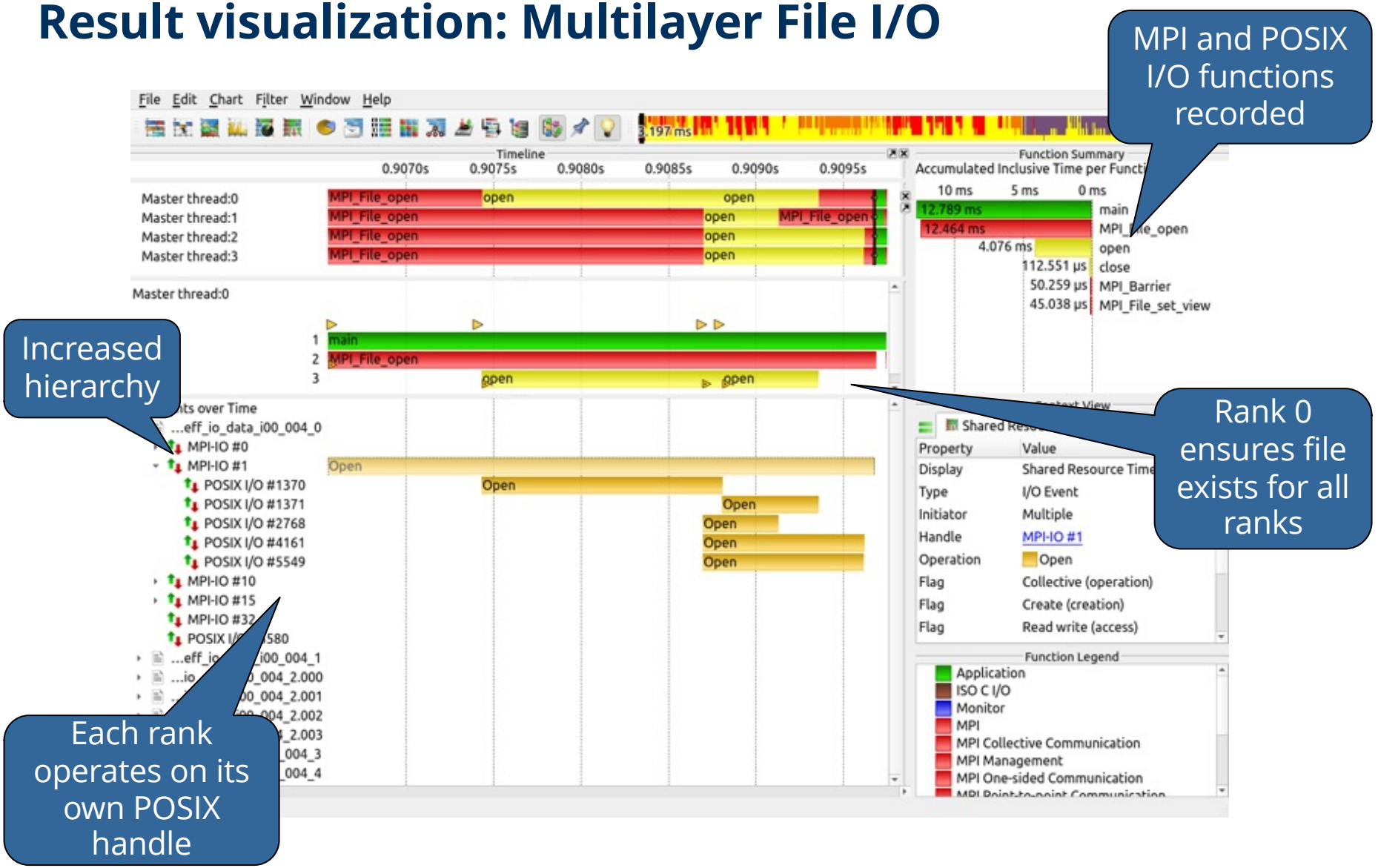
B_EFF I/O

Enabling ISO C and POSIX I/O instrumentation

Instrumentation might require threading support

```
% export SCOREP_WRAPPER_INSTRUMENTER_FLAGS=\
    '--io=runtime:posix --thread=pthread'
% scorep-mpicc -o b_eff_io b_eff_io.c
% export SCOREP_EXPERIMENT_DIRECTORY=scorep-b_eff_io-4-profile+posix
% export SCOREP_ENABLE_TRACING=false
% scorep-scorep scorep-b_eff_io-4-profile+posix/profile.cubex
% export SCOREP_EXPERIMENT_DIRECTORY=scorep-b_eff_io-4-tracing+posix
% export SCOREP_ENABLE_TRACING=true
% export SCOREP_TOTAL_MEMORY=61MB
% mpirun -n 4 -c 6 ./b_eff_io -MB 2048 -MT 98304 -rewrite -N 4 -T 60
```

Result visualization: Multilayer File I/O



TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Center for Information Services and High Performance Computing (ZIH)

Summary & Conclusion

ORNL/OLCF Vampir and Score-P Training
2022 Jan 19

Summary

Vampir & VampirServer

- Interactive trace visualization and analysis
- Intuitive browsing and zooming
- Scalable to large trace data sizes (20 TByte)
- Scalable to high parallelism (200000 processes)

Vampir for Linux, Windows and Mac OS X

Note: Vampir neither solves your problems automatically nor points you directly at them. It does, however, give you FULL insight into the execution of your application.

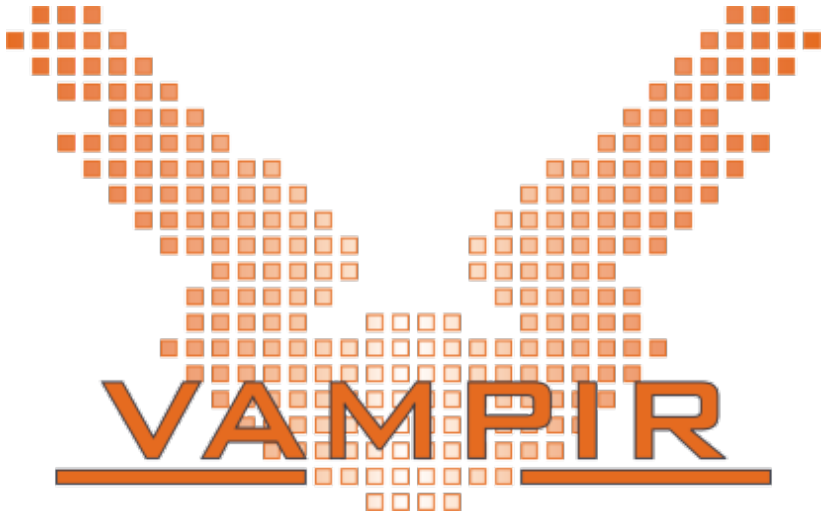
Conclusion

Performance analysis very important in HPC

Use performance analysis tools for profiling and tracing

Use tracing tools with some precautions
- Overhead
- Data volume

Contact the Vampir Team for help on usage: service@vampir.eu



https://www.vampir.eu