Performance Analysis of Large-Scale OpenMP and Hybrid MPI/OpenMP Applications with Vampir NG

Holger Brunst

Center for High Performance Computing
Dresden University, Germany

June 1st, 2005

ZHR

TECHNISCHE
UNIVERSITÄT
DRESDEN

Overview

- Motivation
- Analysis of large hybrid parallel applications
  - Integration of existing monitoring systems
  - Scalable overall concept
  - Parallelization of analysis
  - Separation of visualization and analysis
- Performance results und experiences
- Conclusion

Motivation

- OpenMP most commonly used standard for shared-memory based parallel computing
- MPI well established in distributed computing with respect to large problem and system sizes
- Most applications are either MPI or OpenMP
- Large clusters of SMPs
  - hybrid applications are one way to go
  - no automatic parallelization
- Most tools support either MPI or OpenMP
- Available for dedicated platforms of certain vendors only

OpenMP Performance-Analysis Framework

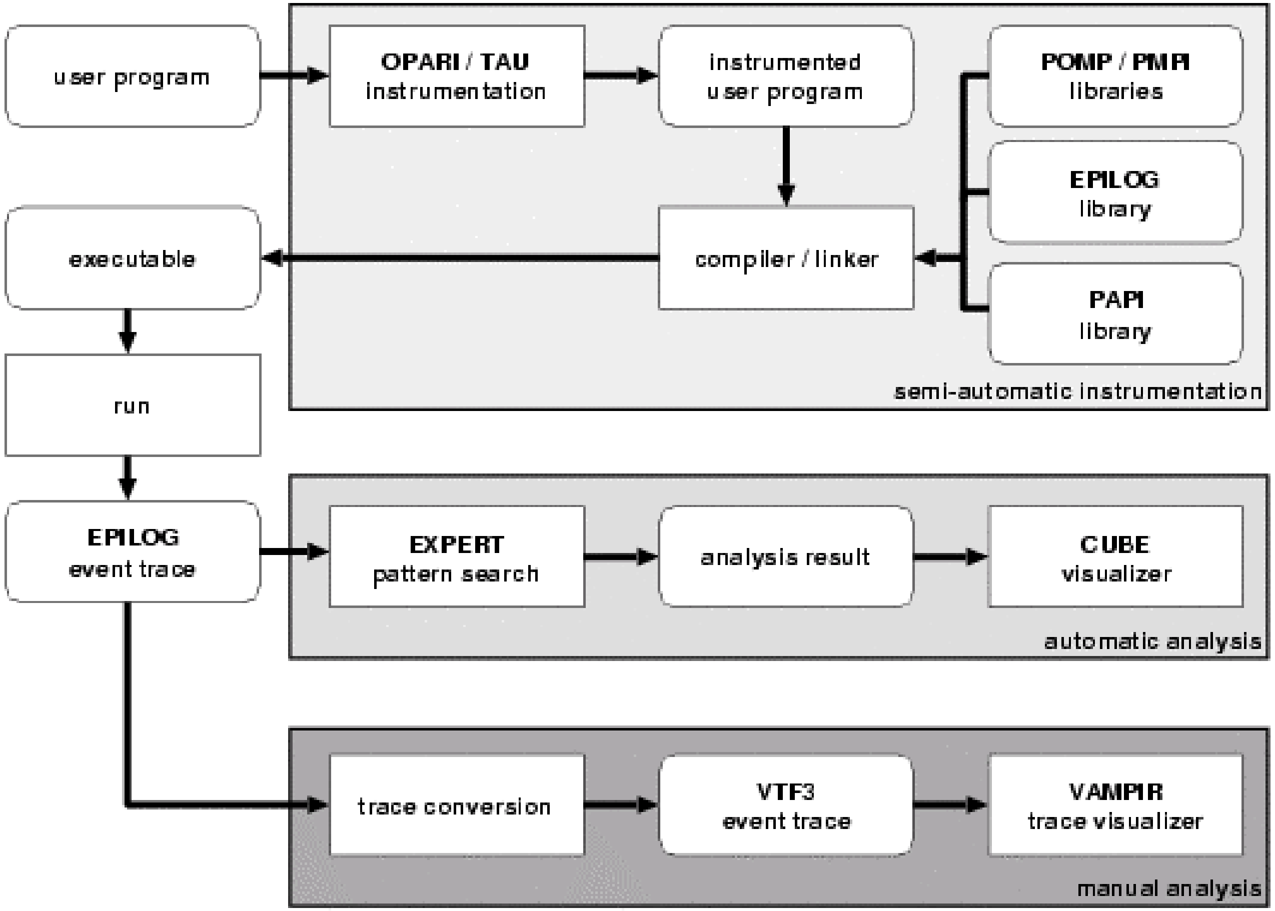
- Instrumentation
  - insert/append monitoring infrastructure
  - manual-, source-, compiler-, binary- and dynamic binary-instrumentation
  - OPARI source translation (see KOJAK project)
- Trace generation
  - KOJAK measurement system
  - EPILOG to VAMPIR mapping
- Visualization
  - Vampir NG (parallel) / Vampir (sequential)
  - scalable parallel analysis and visualization

Goal

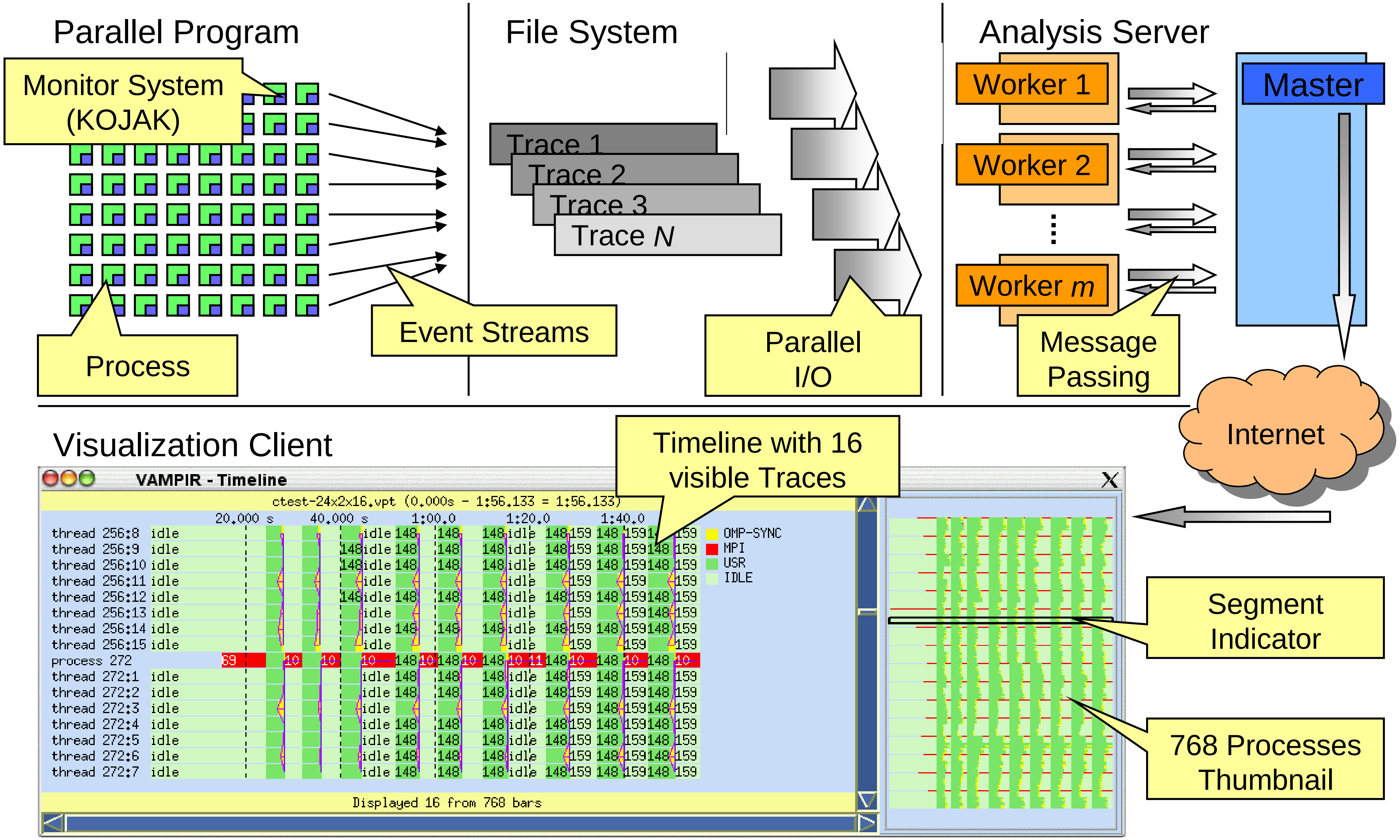
- Hybrid Performance-Analysis off large applications and systems
  - MPI, OpenMP, also pthreads
- Support
  - many thousand threads of execution
  - at least 10^9 performance events
- Distributed/shared memory
- Interactive analysis with short response times
- Seamless integration in production environments
  - high requirements regarding portability
- Extensible with analysis plugins

OpenMP Monitoring: KOJAK

- Tracing based
- OpenMP, MPI or both
- Source translation (POMP)
- Wrapper (PMPI)
- User functions (TAU)
- Hardware Counter (PAPI)
- Automatic analysis with EXPERT
- Manual visualization and analysis with Vampir NG



Framework • Scalable OpenMP Analysis

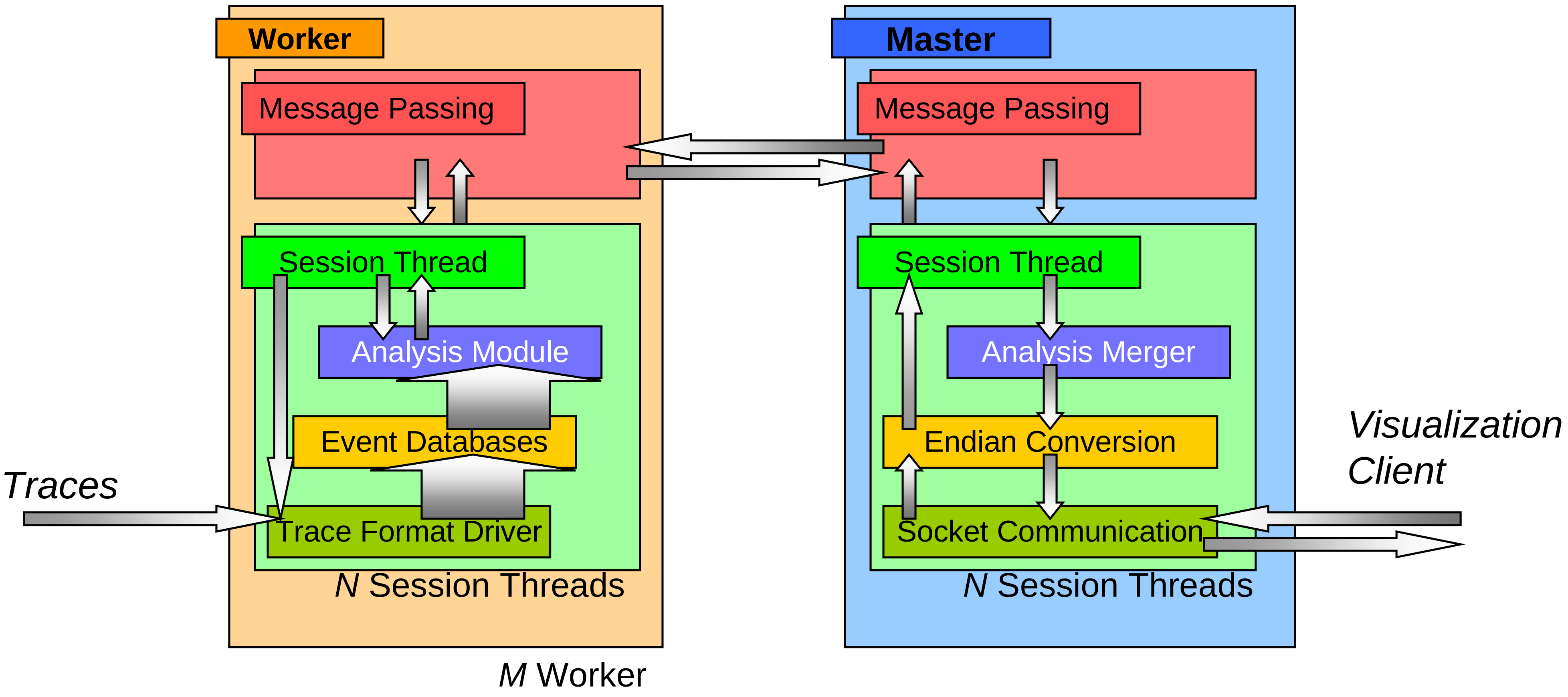


VAMPIR - Timeline

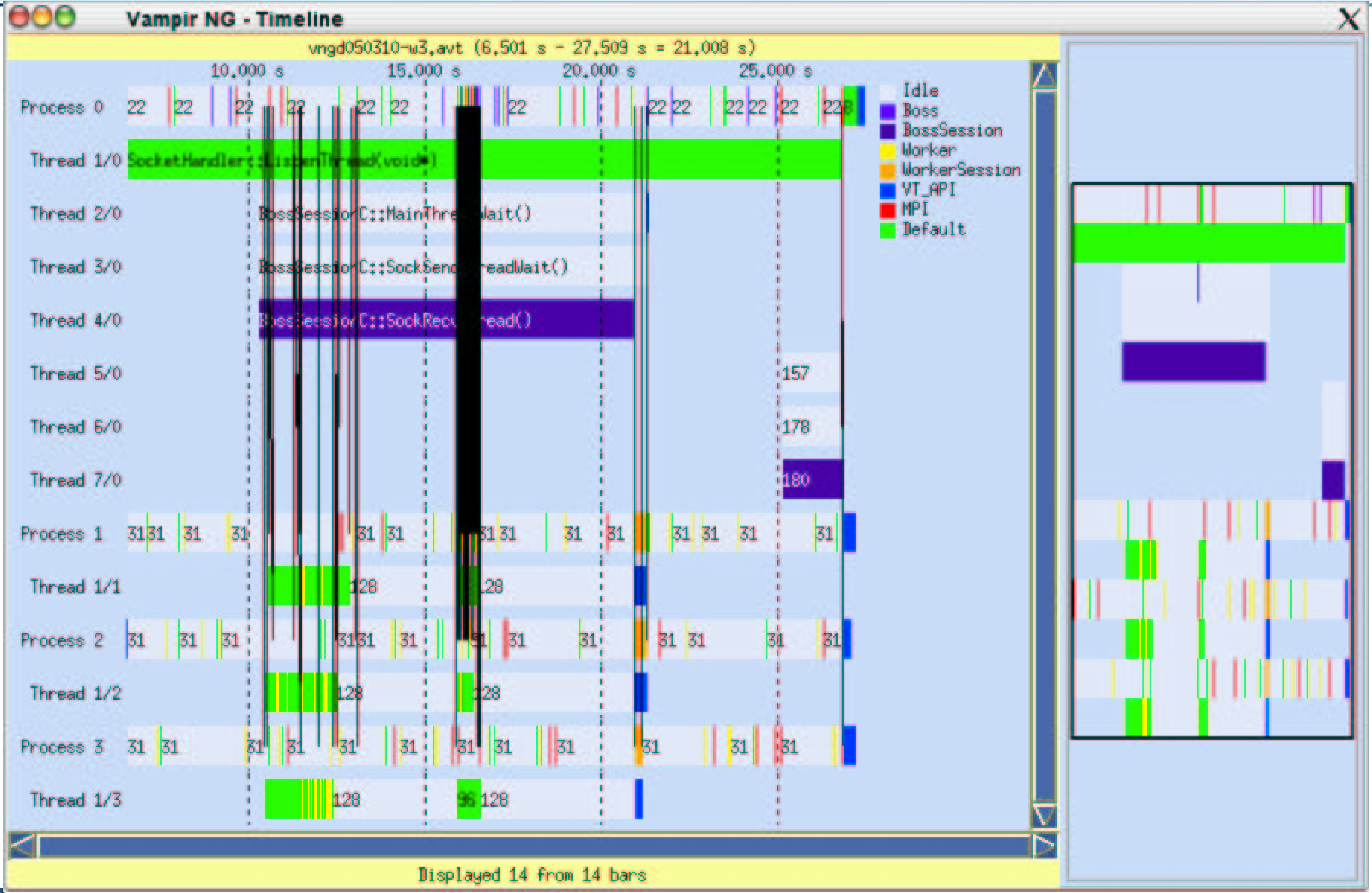
```
ctest-24x2x16.vpt (0.000s - 1:56.133 = 1:56.133)
                20.000 s    40.000 s    1:00.0    1:20.0    1:40.0
thread 256:8  idle
thread 256:9  idle
thread 256:10 idle
thread 256:11 idle
thread 256:12 idle
thread 256:13 idle
thread 256:14 idle
thread 256:15 idle
process 272
thread 272:1  idle
thread 272:2  idle
thread 272:3  idle
thread 272:4  idle
thread 272:5  idle
thread 272:6  idle
thread 272:7  idle
Displayed 16 from 768 bars
```

🟨 OMP-SYNC
🟥 MPI
🟩 USR
▭ IDLE

Organization of Parallel Analysis



Framework • Parallel Analysis • Example • Self Analysis



Vampir NG - Timeline

vngd050310-w3.avt (6.501 s - 27.509 s = 21.008 s)

Displayed 14 from 14 bars

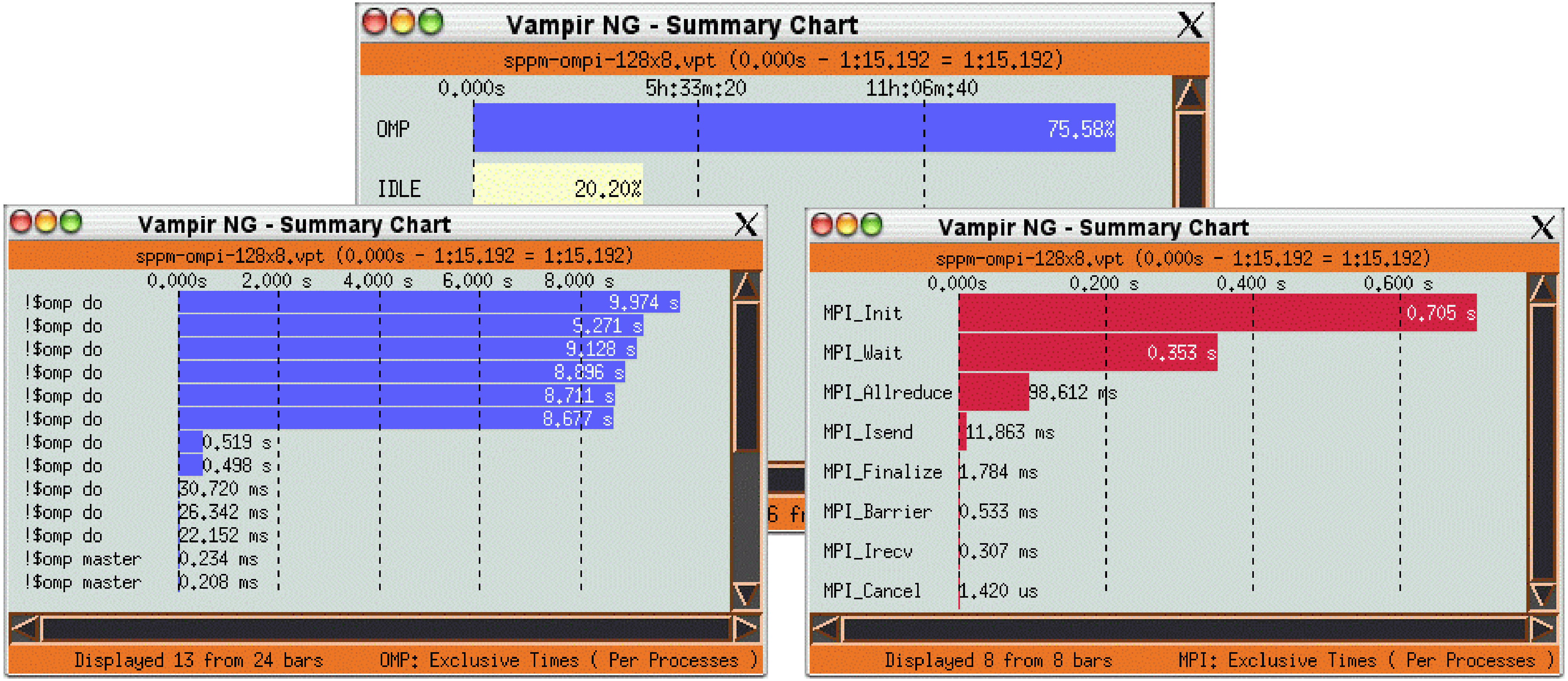
Parallel Analysis – Supported Request Types

- Approx. 35 Requests:
  - Stack-Tree
  - Timeline
  - Accumulative Timeline
  - Profiles
  - Thumbnails
- Process Global/Local
- Event Types: Functions, Messages, MPI/OpenMP Collectives, I/O, Hardware Counter

Scalable Visualization

- Performance-Analysis becomes more complex
  - Different/multiple communication layers
  - Combination of shared- und distributed memory
  - New information sources
- Grouping of data streams depending on the problem to be analyzed
- Hierarchical grouping
  - Static: Physical structure e.g. nodes, processes, and OpenMP threads
  - Dynamic: During analysis, to look at results from different angles

Configurable OpenMP and MPI Profiles



Vampir NG - Summary Chart

sppm-ompi-128x8.vpt (0.000s - 1:15.192 = 1:15.192)

Vampir NG - Summary Chart

sppm-ompi-128x8.vpt (0.000s - 1:15.192 = 1:15.192)

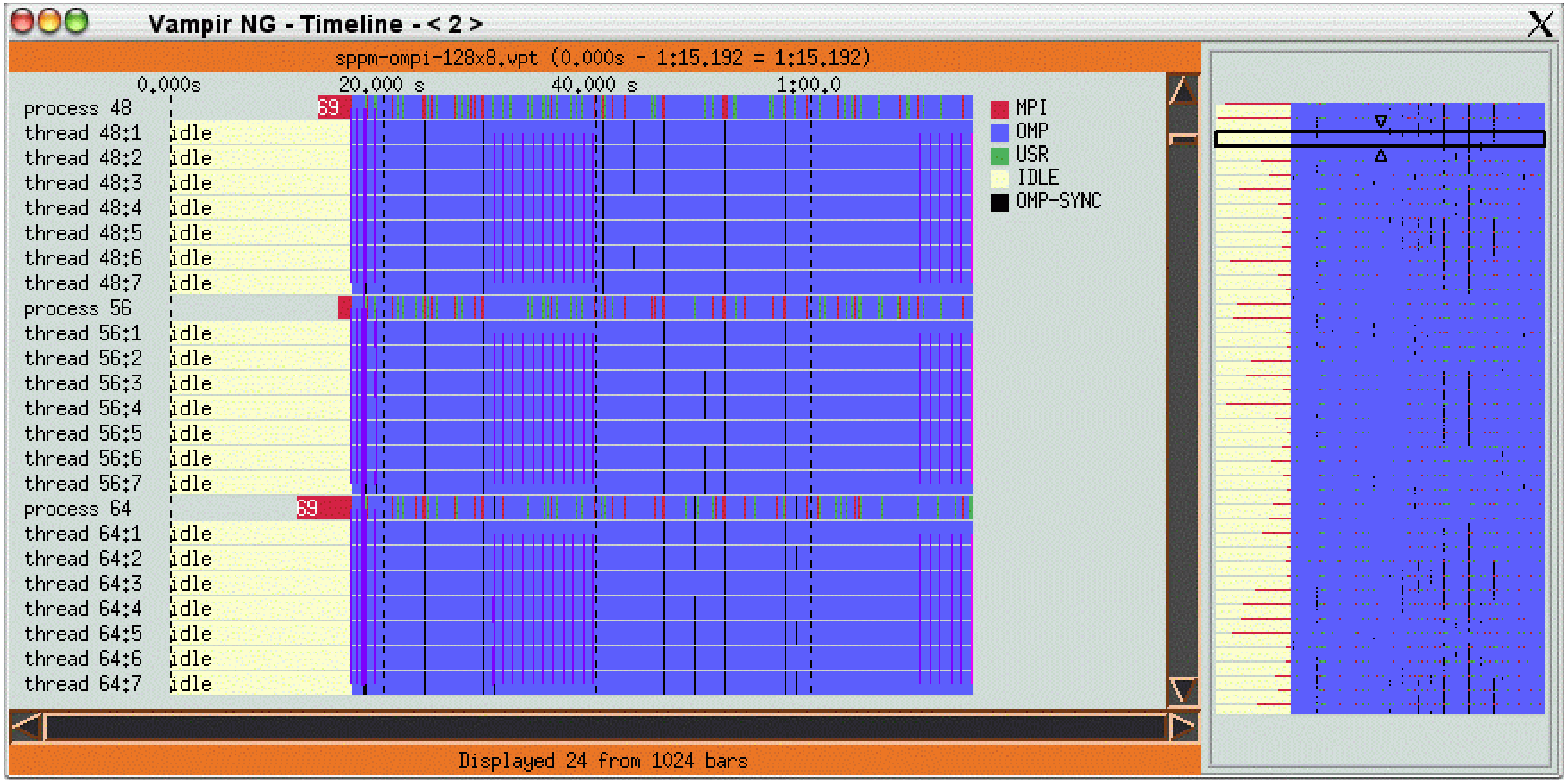
Displayed 13 from 24 bars   OMP: Exclusive Times (Per Processes)

Vampir NG - Summary Chart

sppm-ompi-128x8.vpt (0.000s - 1:15.192 = 1:15.192)

Displayed 8 from 8 bars   MPI: Exclusive Times (Per Processes)

Timeline with OpenMP Activities

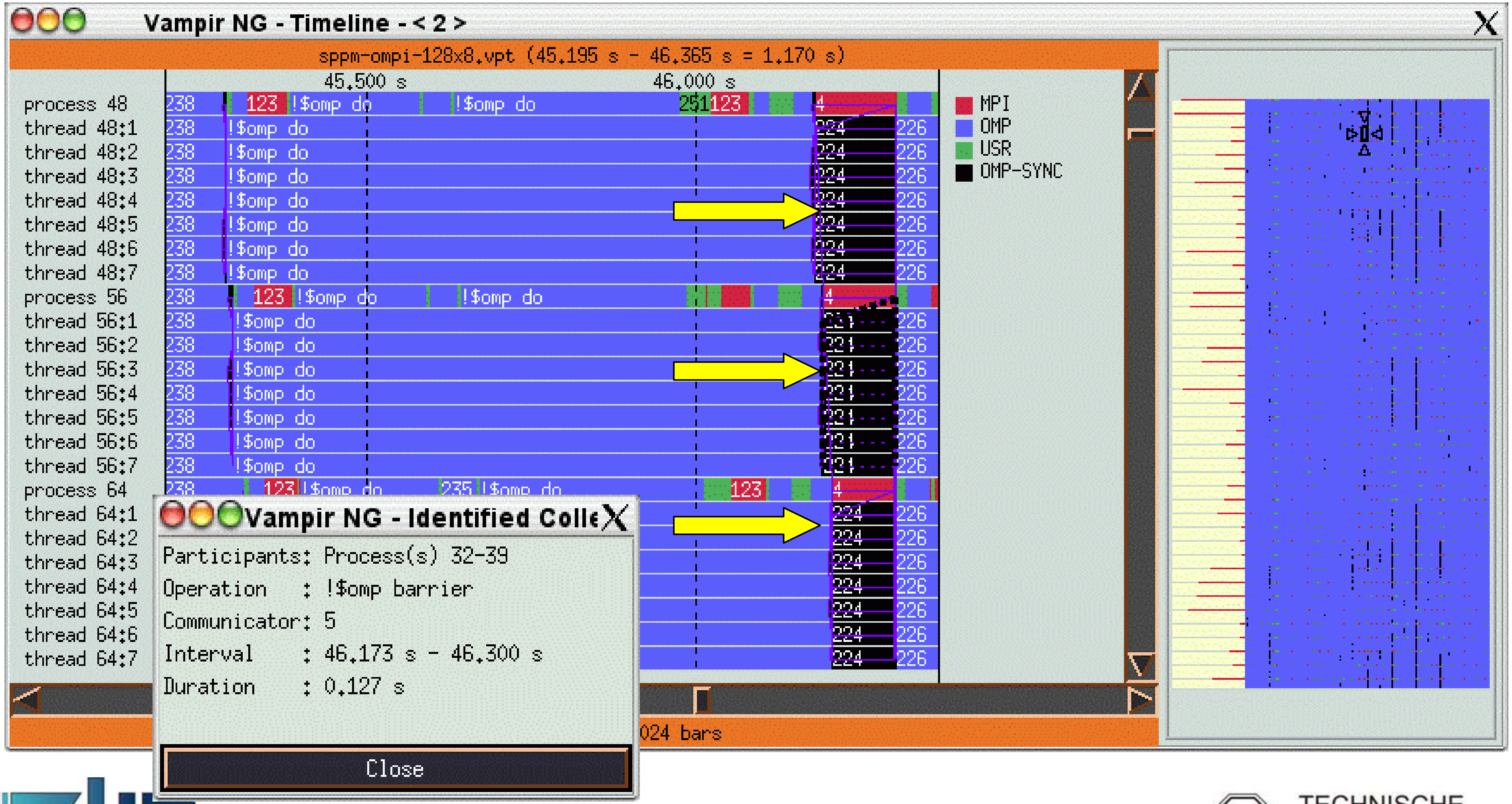


Vampir NG - Timeline - < 2 >

sppm-ompi-128x8.vpt (0.000s - 1:15.192 = 1:15.192)

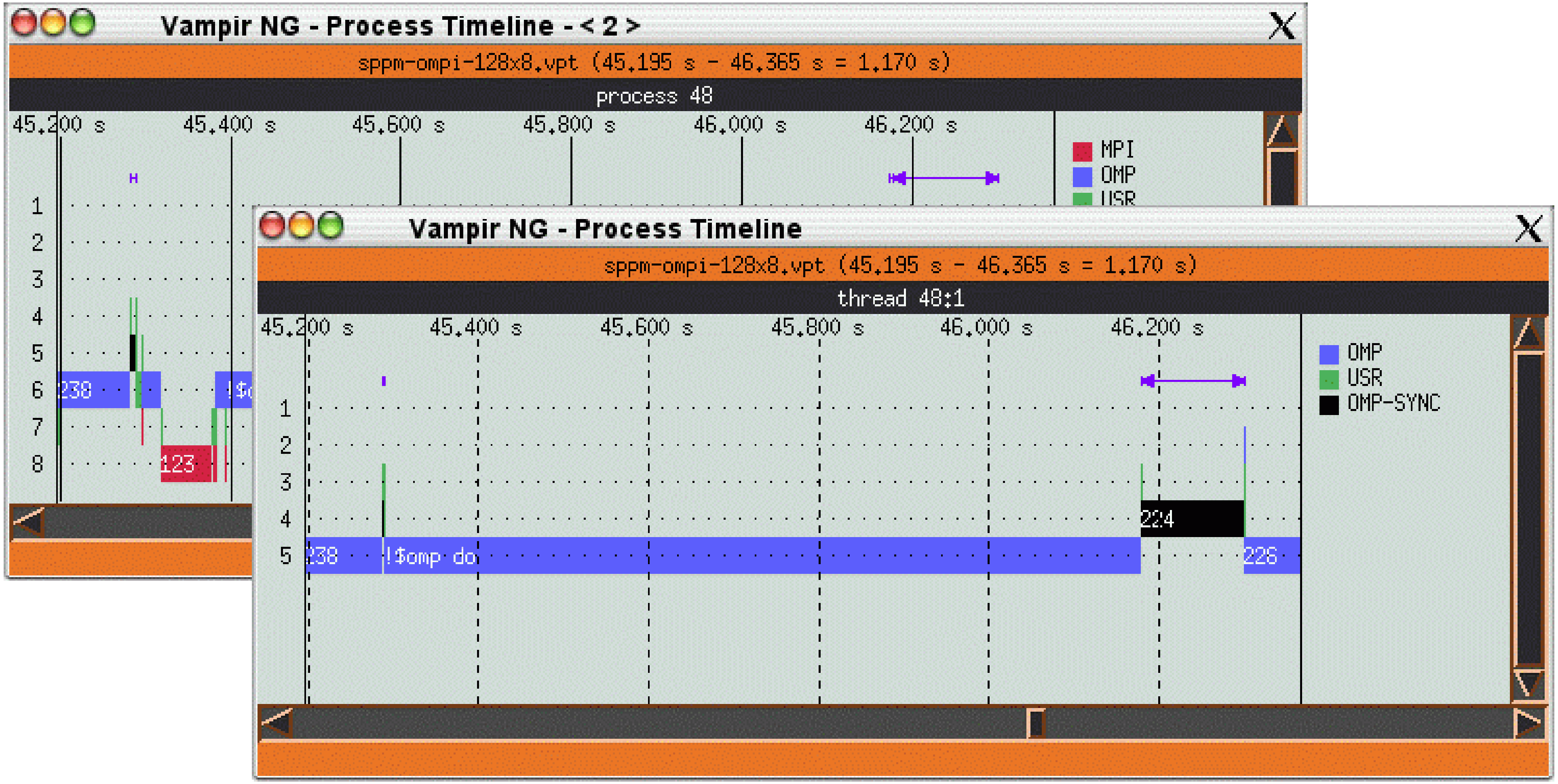
Displayed 24 from 1024 bars

OpenMP Barrier Synchronization



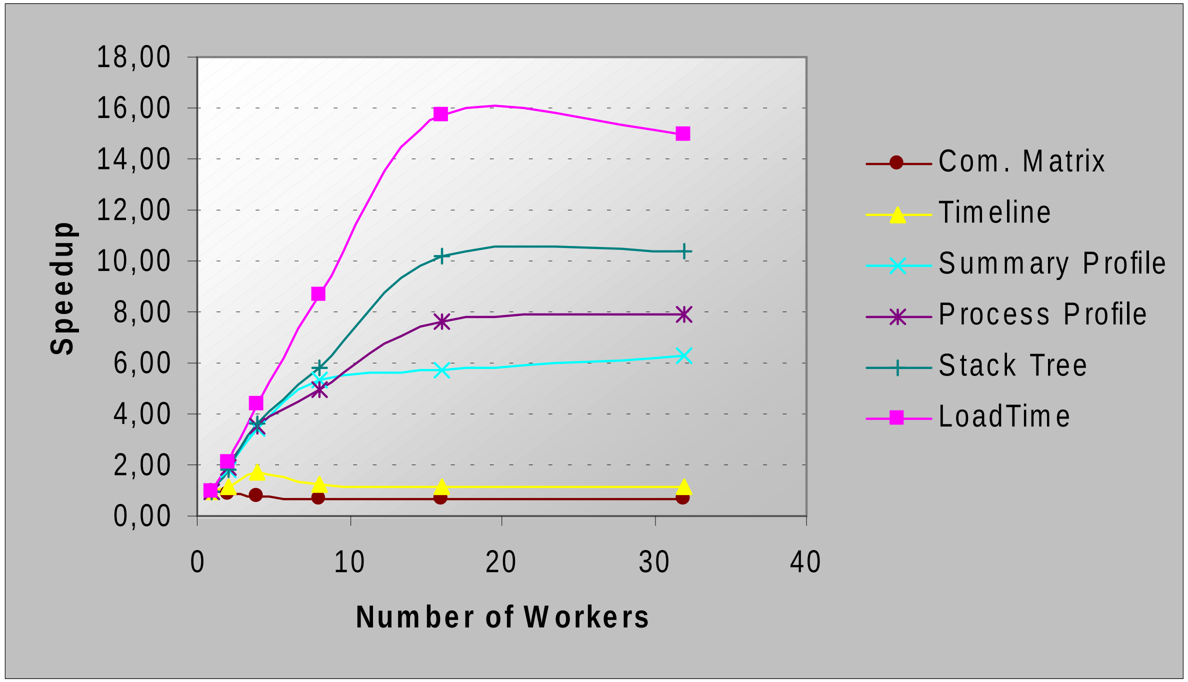
Participants: Process(s) 32-39
Operation : !$omp barrier
Communicator: 5
Interval : 46.173 s - 46.300 s
Duration : 0.127 s

Single OpenMP Thread Timelines



Scalability – sPPM Analyzed on Origin 2000

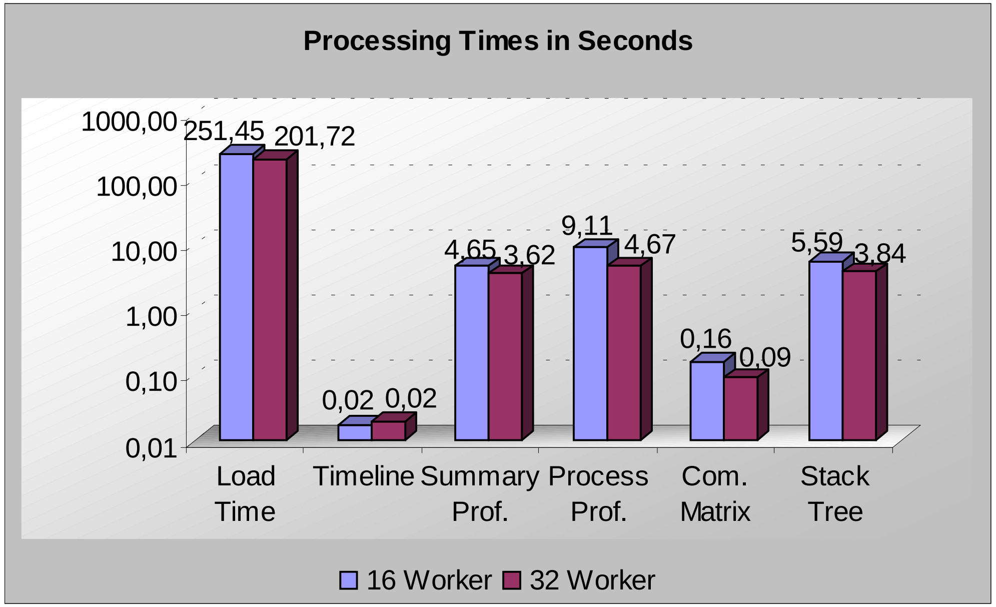
- sPPM ASCI Benchmark
  - 3D Gas Dynamic
- Data to be analyzed
  - 16 Processes
  - 200 MByte Volume



Number of Workers	1	2	4	8	16	32
Load Time	47,33	22,48	10,80	5,43	3,01	3,16
Timeline	0,10	0,09	0,06	0,08	0,09	0,09
Summary Profile	1,59	0,87	0,47	0,30	0,28	0,25
Process Profile	1,32	0,70	0,38	0,26	0,17	0,17
Com. Matrix	0,06	0,07	0,08	0,09	0,09	0,09
Stack Tree	2,57	1,39	0,70	0,44	0,25	0,25

A Fairly Large Test Case

- IRS ASCI Benchmark
  - Implicit Radiation Solver
- Data to be analyzed:
  - 64 Processes in 8 Streams
  - Approx. 800.000.000 Events
  - 40 GByte Data Volume
- Analysis Platform:
  - Jump.fz-juelich.de
  - 41 IBM p690 nodes
  - 32 processors per node
  - 128 GByte per node
- Visualization Platform:
  - Remote Laptop



Application and Experiences

- Implementation and evaluation of a prototype in the scope of an ongoing support contract with ASC Labs (LLNL, LANL, SANL)
- Machines with up to 5,000 Processors (soon: BlueGene/L with up to 130,000 Processors)
- Valuable feedback from users and developers
- Comparison to sequential approach:
  - Factor 100 regarding data volume
    (50 GByte vs. 500 MByte)
  - Analysis required at most 32 interactive processors
  - Interactive usage from remote desktop
    (even from Germany)

Summary

- Visualization and analysis of highly parallel OpenMP and hybrid OpenMP/MPI applications
  - Portable source code instrumentation with OPARI
  - Scalable monitoring with KOJAK monitoring system
  - Conception of scalable/distributed data structures, algorithms and visualization modes
  - Parallelization of analysis
  - Separation of visualization and analysis
  - Simple integration in common production environments due to portability of KOJAK and VAMPIR

Thank You!

www.fz-juelich.de/zam/kojak
icl.cs.utk.edu/kojak

www.vampir-ng.org