# WiP: Simulating Application and System Interaction with PIOsimHD

Julian M. Kunkel, Thomas Ludwig

julian.martin.kunkel@informatik.uni-hamburg.de

Scientific Computing
Department of Informatics
University of Hamburg

2012-02-16

informatik
die zukunft

informatik
die zukunft

# Motivation

## Goals of the project

- Analyze performance of MPI-IO applications in-silicon
- Localize bottlenecks
- Evaluate collective calls in MPI
- Foster understanding of performance factors
- Evaluate MPI algorithms in arbitrary cluster environments
- Extrapolate system performance for future systems

informatik die zukunft

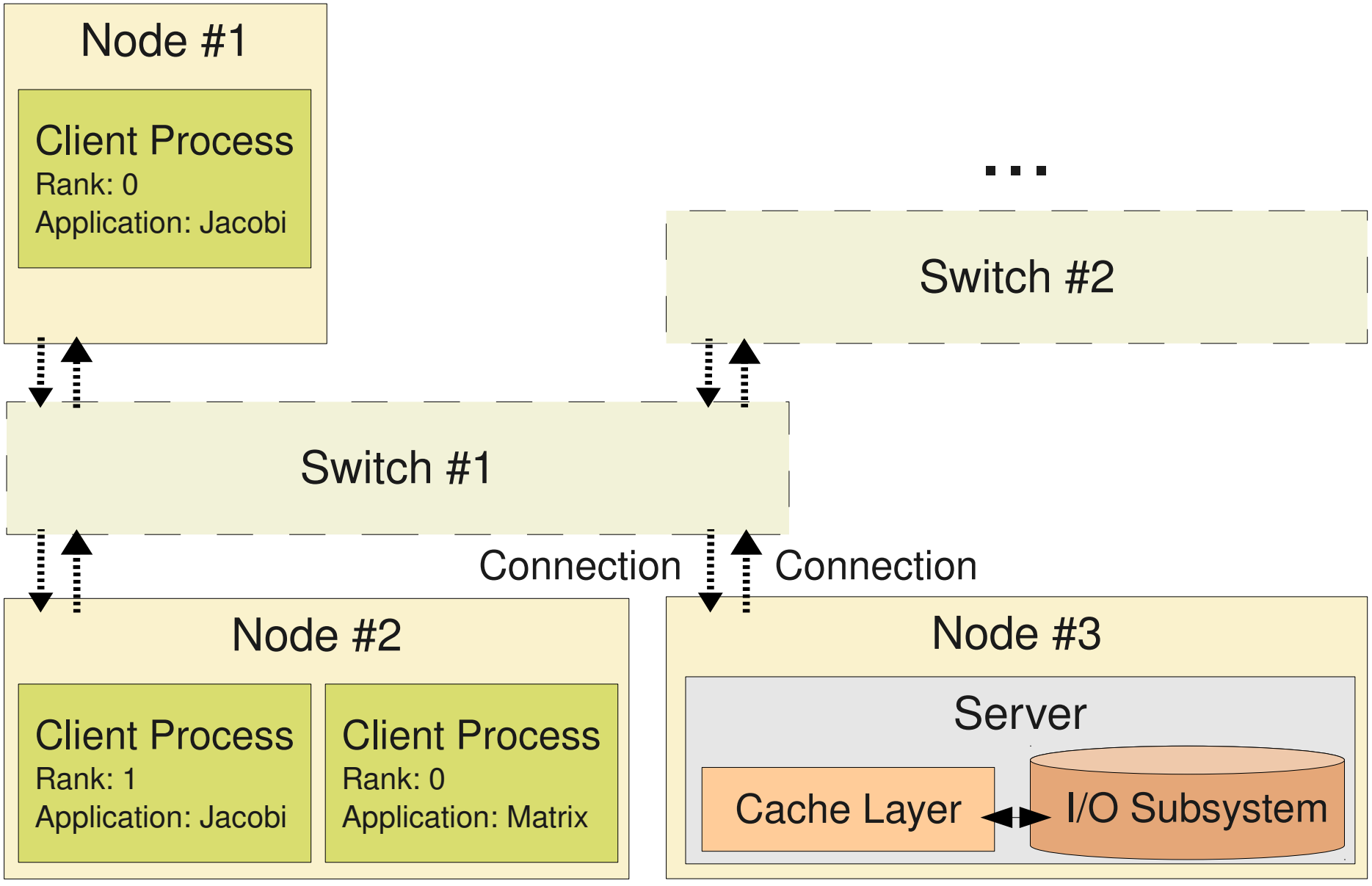
# PIOsimHD

## Selected features

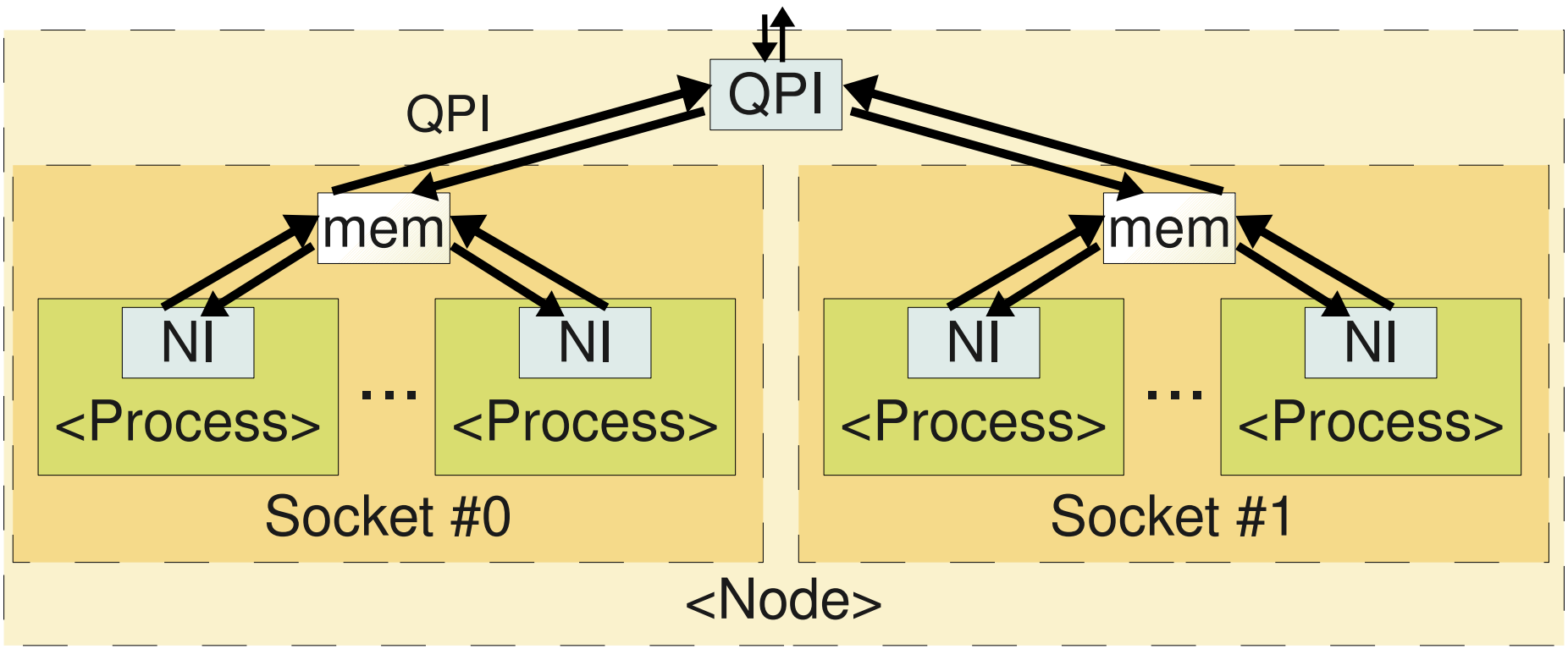
- Most important hardware characteristics are modeled:
  - Nodes, network, block device, memory access[1]
- Simulated software aspects:
  - MPI: P2P & Several collective algorithms are already implemented
  - Abstract (simple) model of parallel file system
  - Write-behind cache
- Implementation and characteristics can be selected per component
- Trace/Replay of existing applications
- Visualization of simulation results with trace viewer
- HDTrace allows to compare existing runs

[1] For SMP-communication.

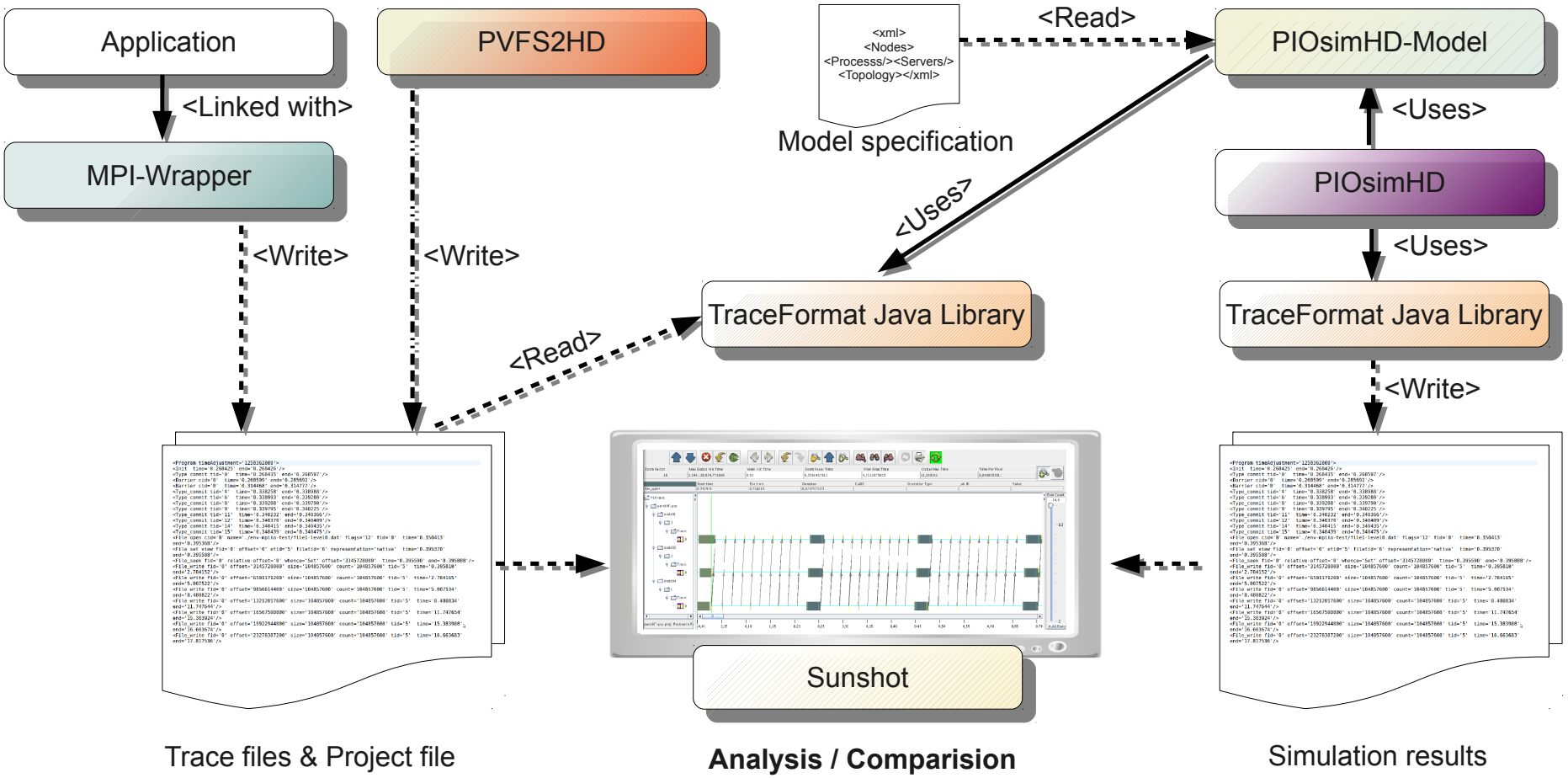
# Cluster model



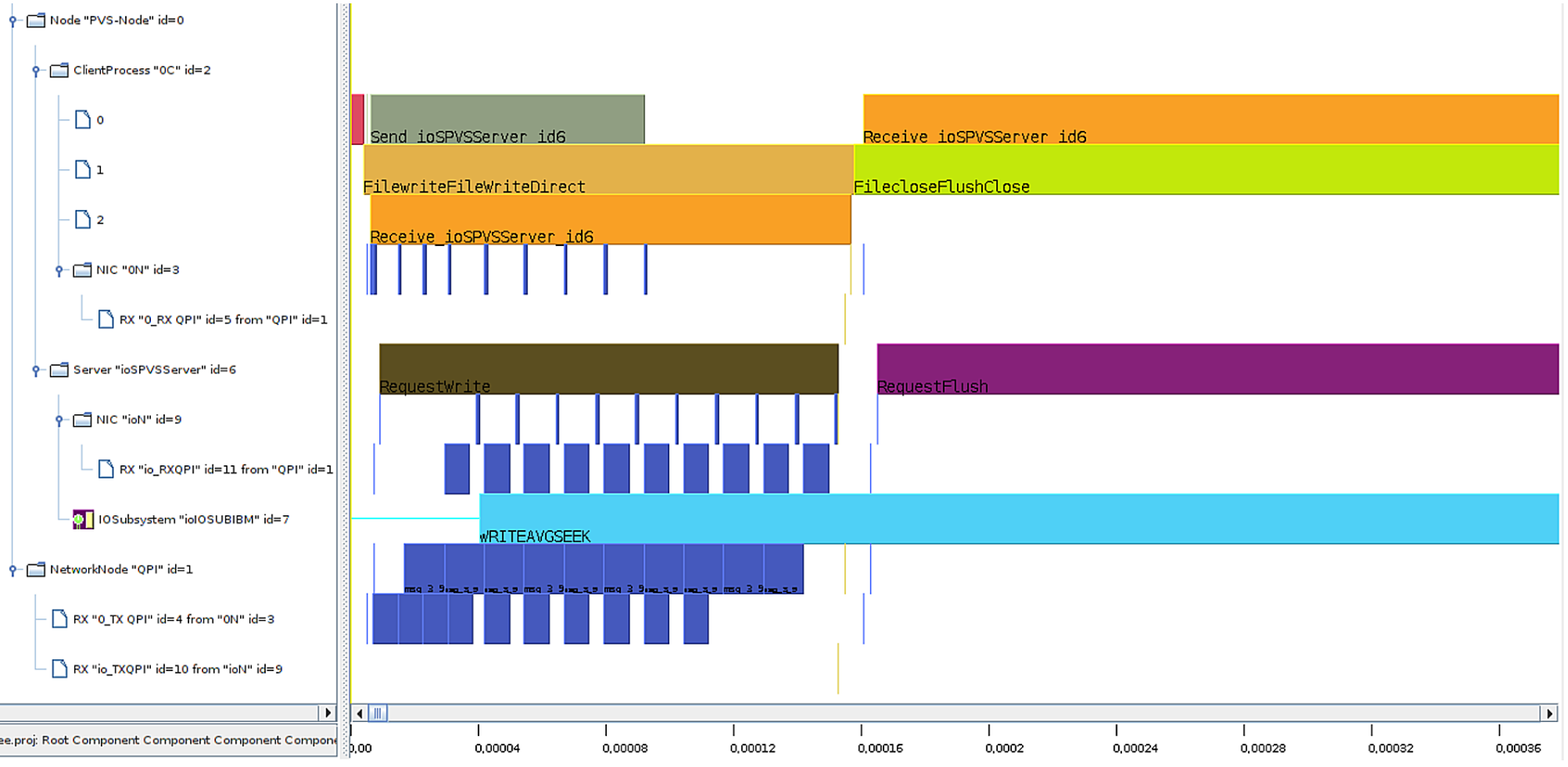
# Model for a dual-socket node



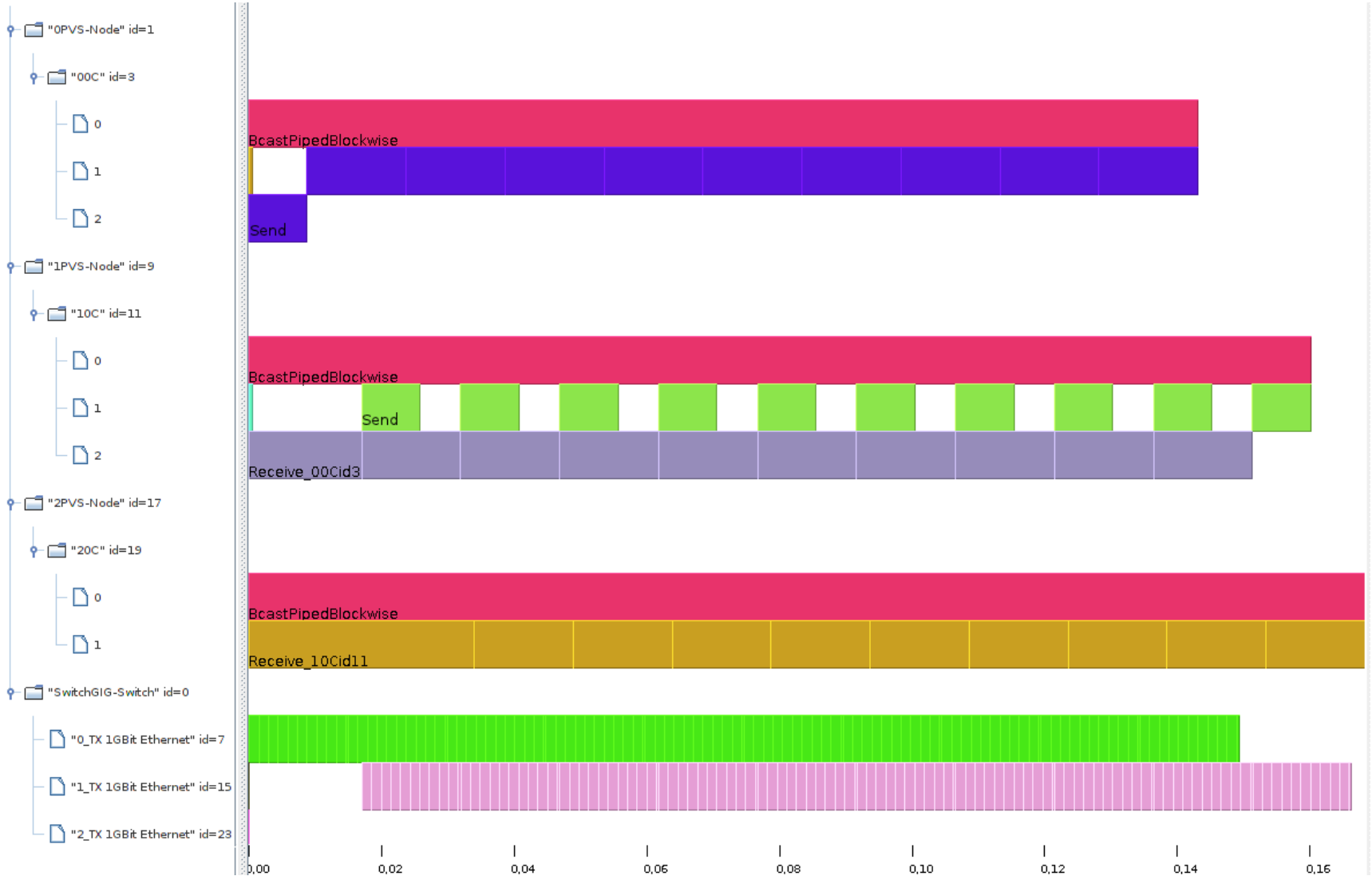
# Simulation workflow



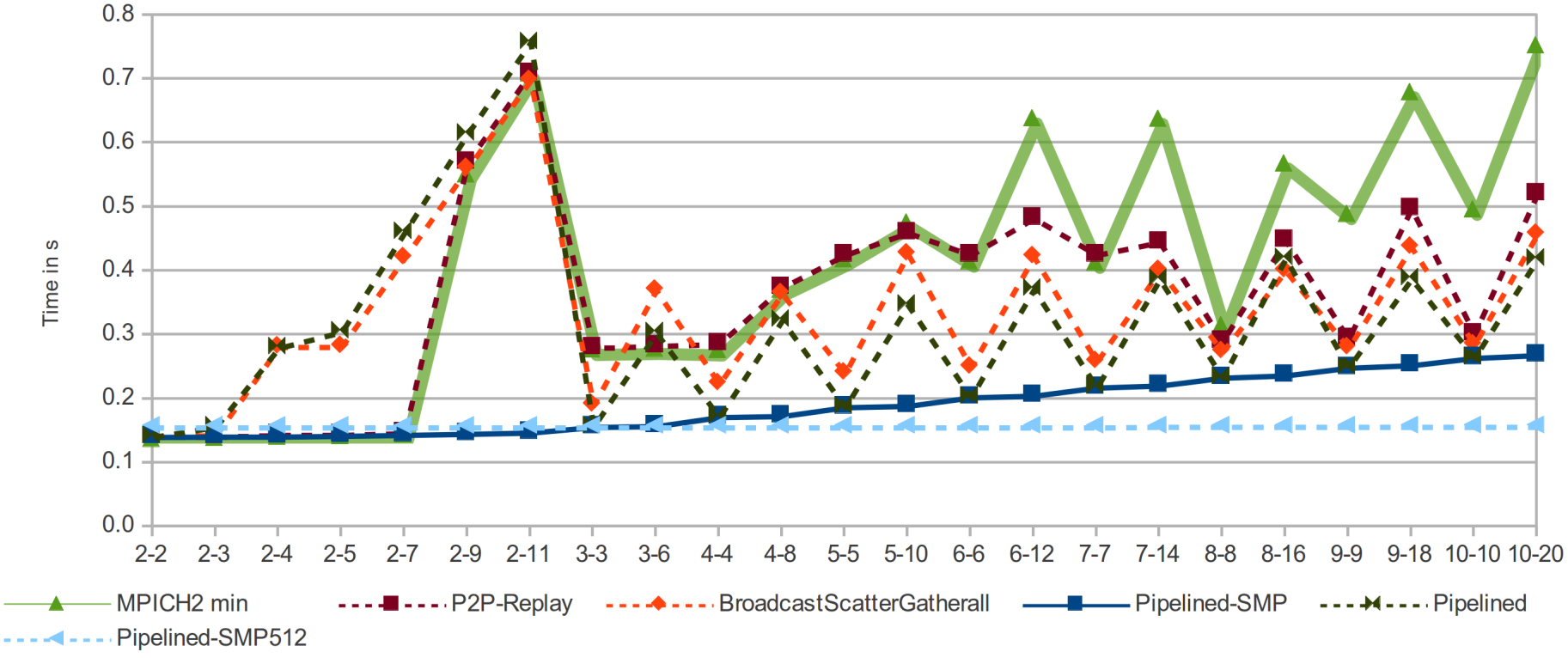
# Screenshot of one simulated I/O operation



# Screenshot of the Pipelined `MPI_Bcast()` implementation



# Performance comparison of several implementations for `MPI_Bcast()`



# Current status / Roadmap

## Current status

- Validation of the cluster (and file system) model has been done
- Some algorithms for Bcast() have been implemented and evaluated for cluster systems
- PhD thesis will be published this year

## Future work

- Conduct more experiments with HDTrace/PIOsimHD e.g:
  - Evaluate relaxed MPI collective calls (e.g. `MPI_Reduce()`)
  - Evaluate and implement improved collective calls for clusters