

HPC-Tools JUBE, LLview and SIONlib at JSC: Recent developments

-- Slides about SIONlib (DEEP-ER) --

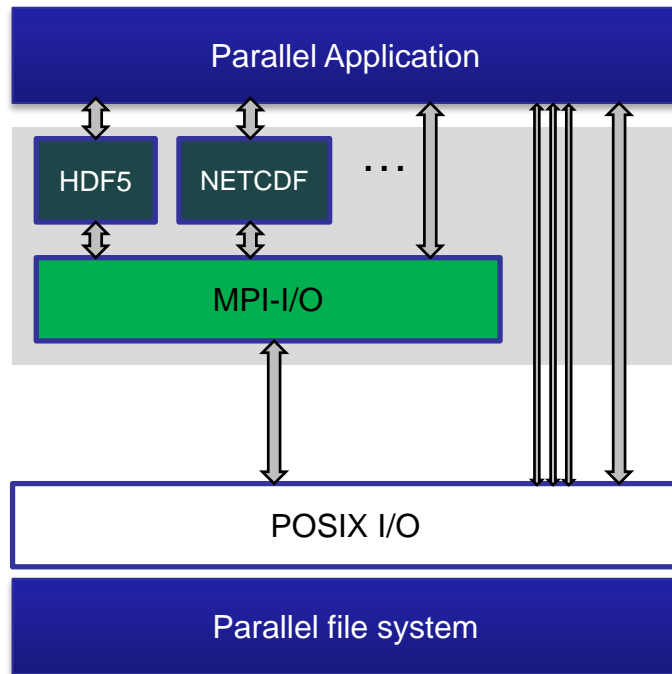
Wolfgang Frings, Sebastian Lührs, Kay Thust

{W.Frings,S.Luehrs,K.Thust}@fz-juelich.de

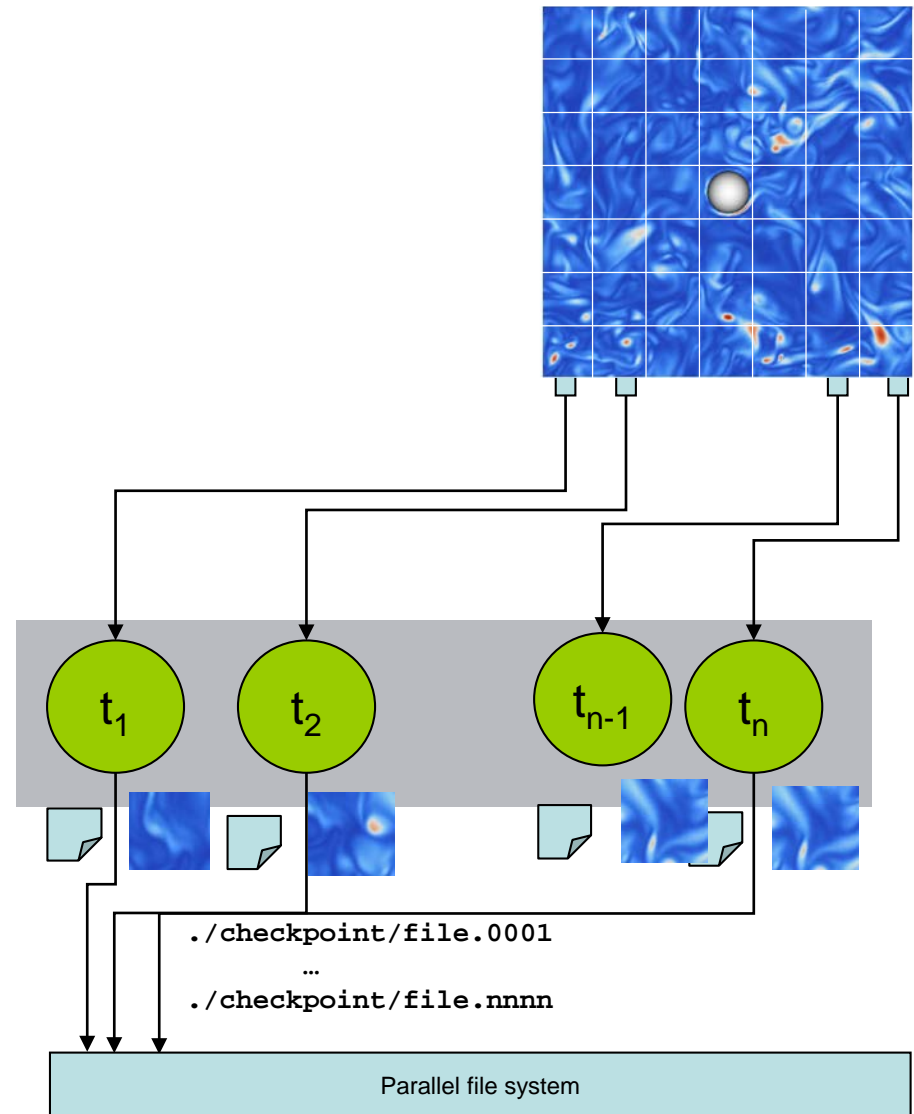
Jülich Supercomputing Centre

JLESC Workshop, Kobe, December 1st, 2016

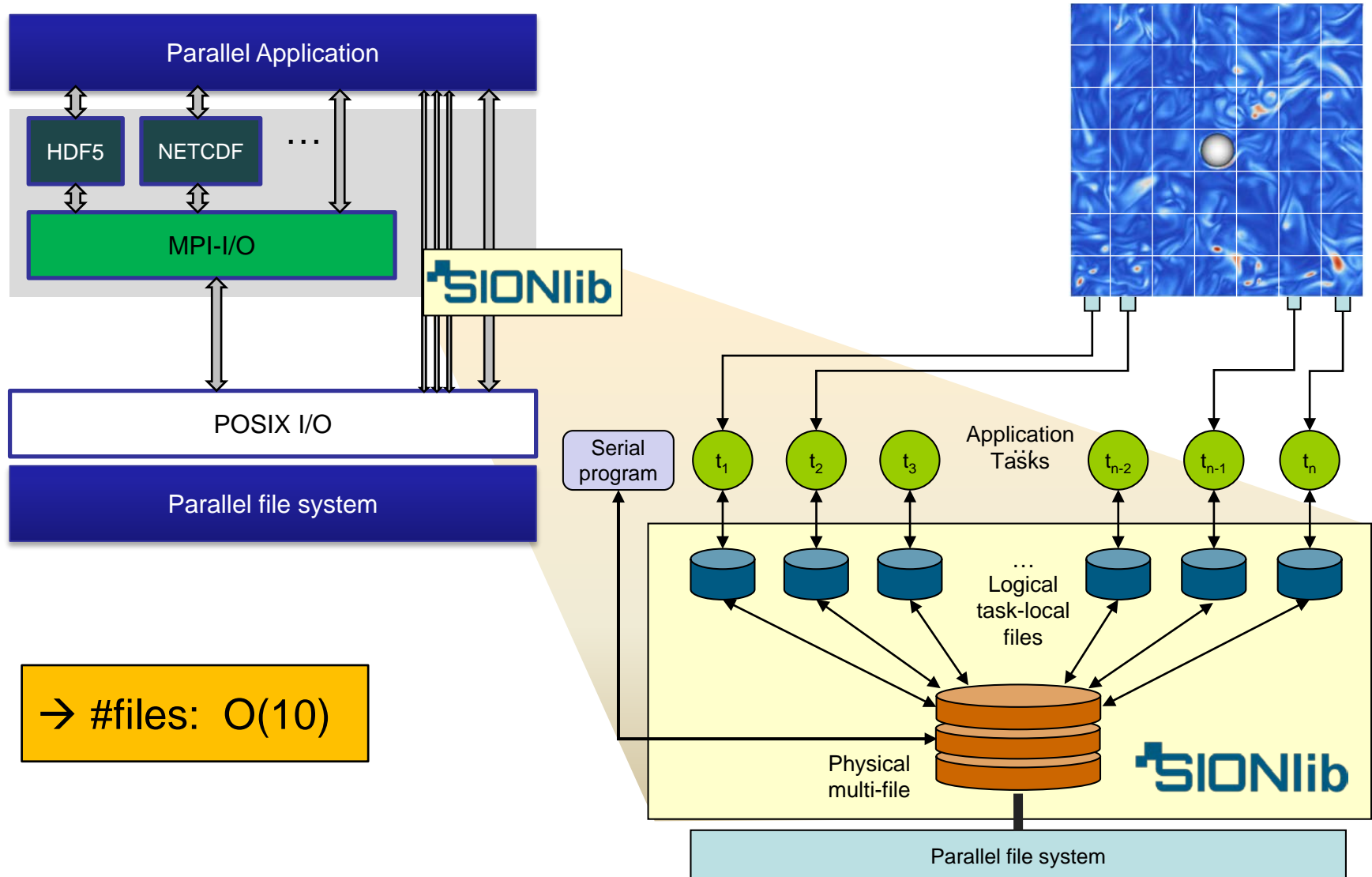
SIONlib: Shared Files for Task-local Data



→ #files: $O(100000)$



SIONlib: Shared Files for Task-local Data



→ #files: $O(10)$

SIONlib Update

→ Version 1.7.1 available

- Including Buddy Checkpointing *



The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n° 287530 and n° 610476

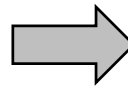
→ Next developments

- Buddy Checkpointing to NAM (xor) *

→ Planned version 2.0 of SIONlib

- Simplifying API, e.g.

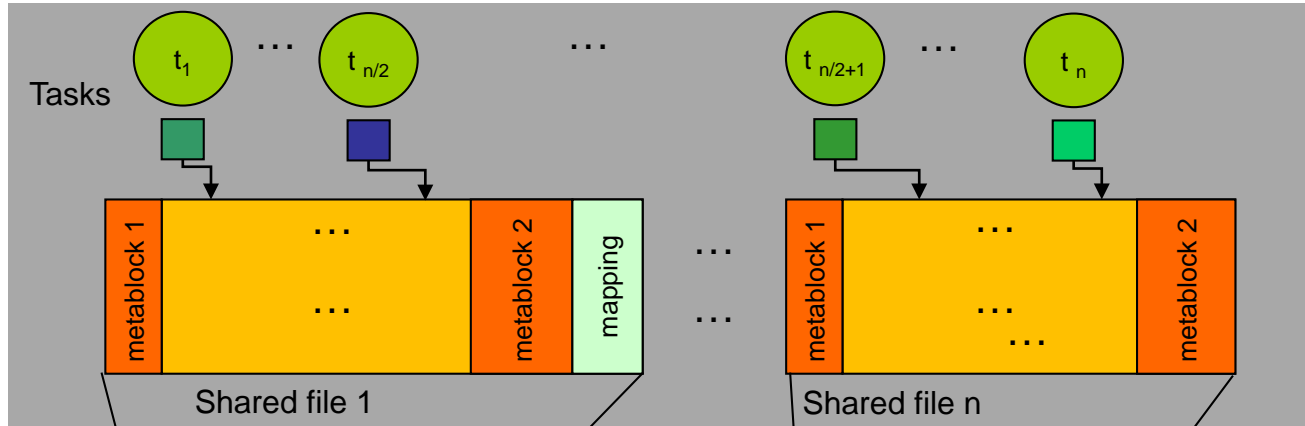
```
int sion_paropen_mpi( const char *fname,
                    const char *file_mode,
                    int *numFiles,
                    MPI_Comm gComm,
                    const MPI_Comm *lComm,
                    sion_int64 *chunksize,
                    sion_int32 *fsblksize,
                    int *globalrank,
                    FILE **fileptr,
                    char **newfname );
```



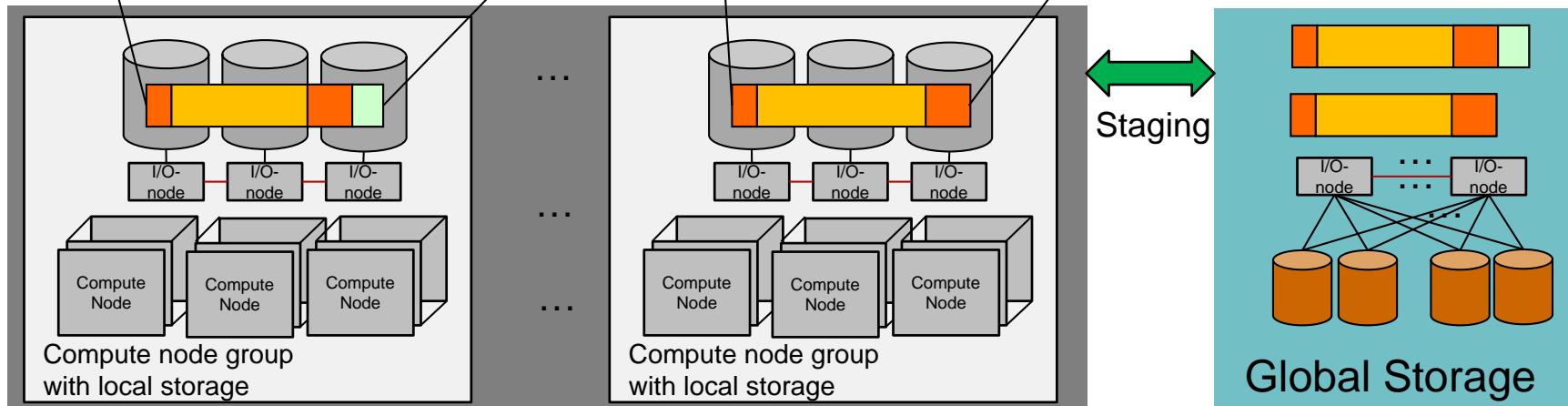
```
options = ...;
int sion_paropen_mpi( const char *fname,
                    const char *file_mode,
                    int *numFiles,
                    MPI_Comm gComm,
                    const sion_mpi_options *options );
```

similar to MPI-I/O hints

SIONlib: Strategy for Local Storage



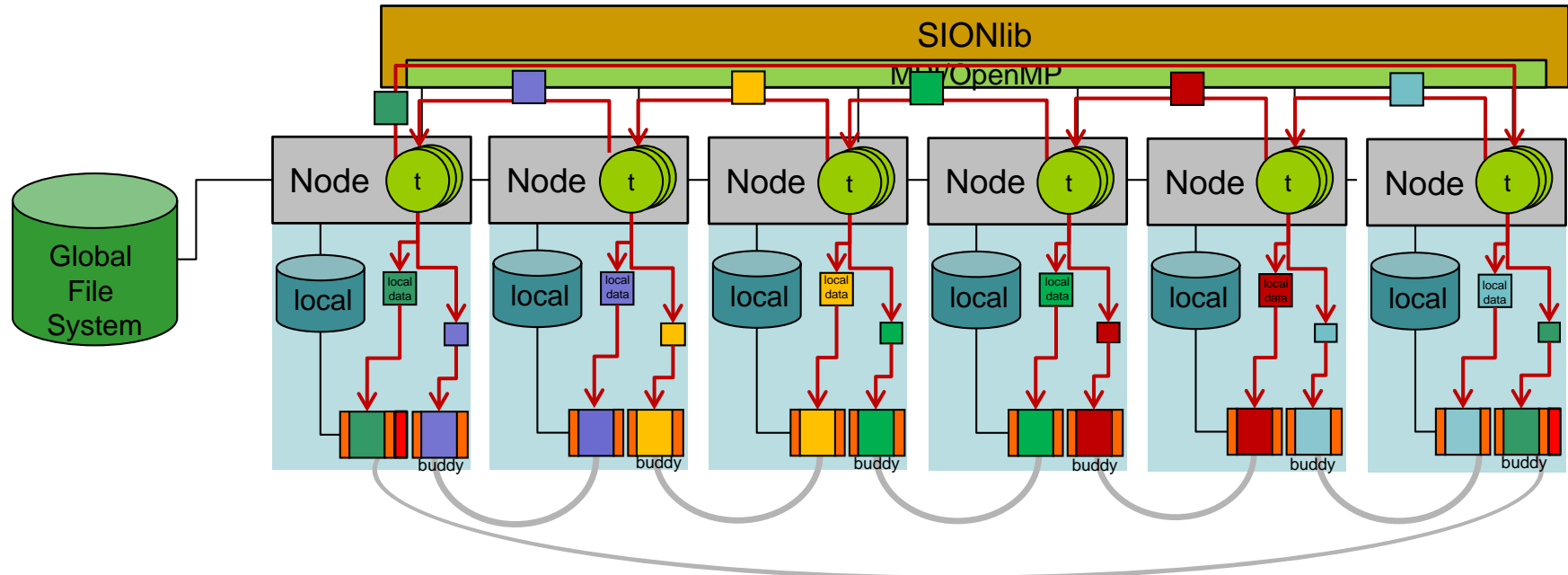
SIONlib file container
→ multi-file approach



Local Storage

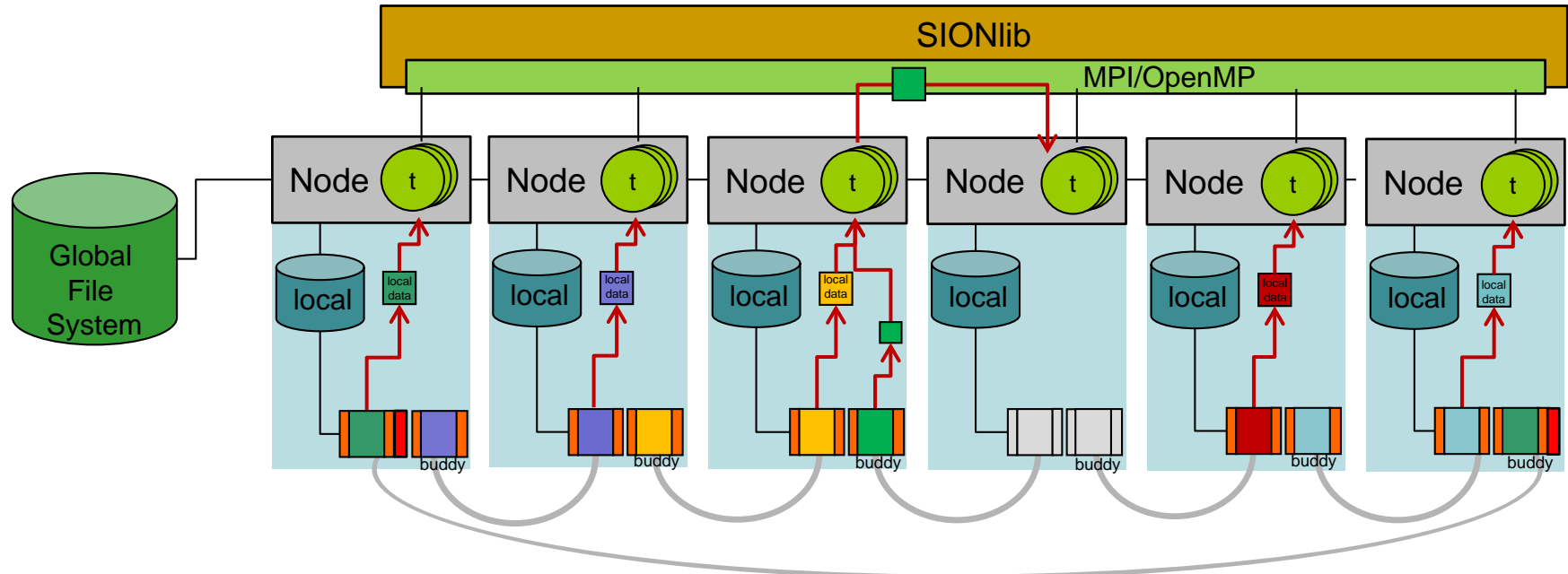
Transparent access
from application to
SIONlib container

SIONlib: Buddy-CP, Logical mapping



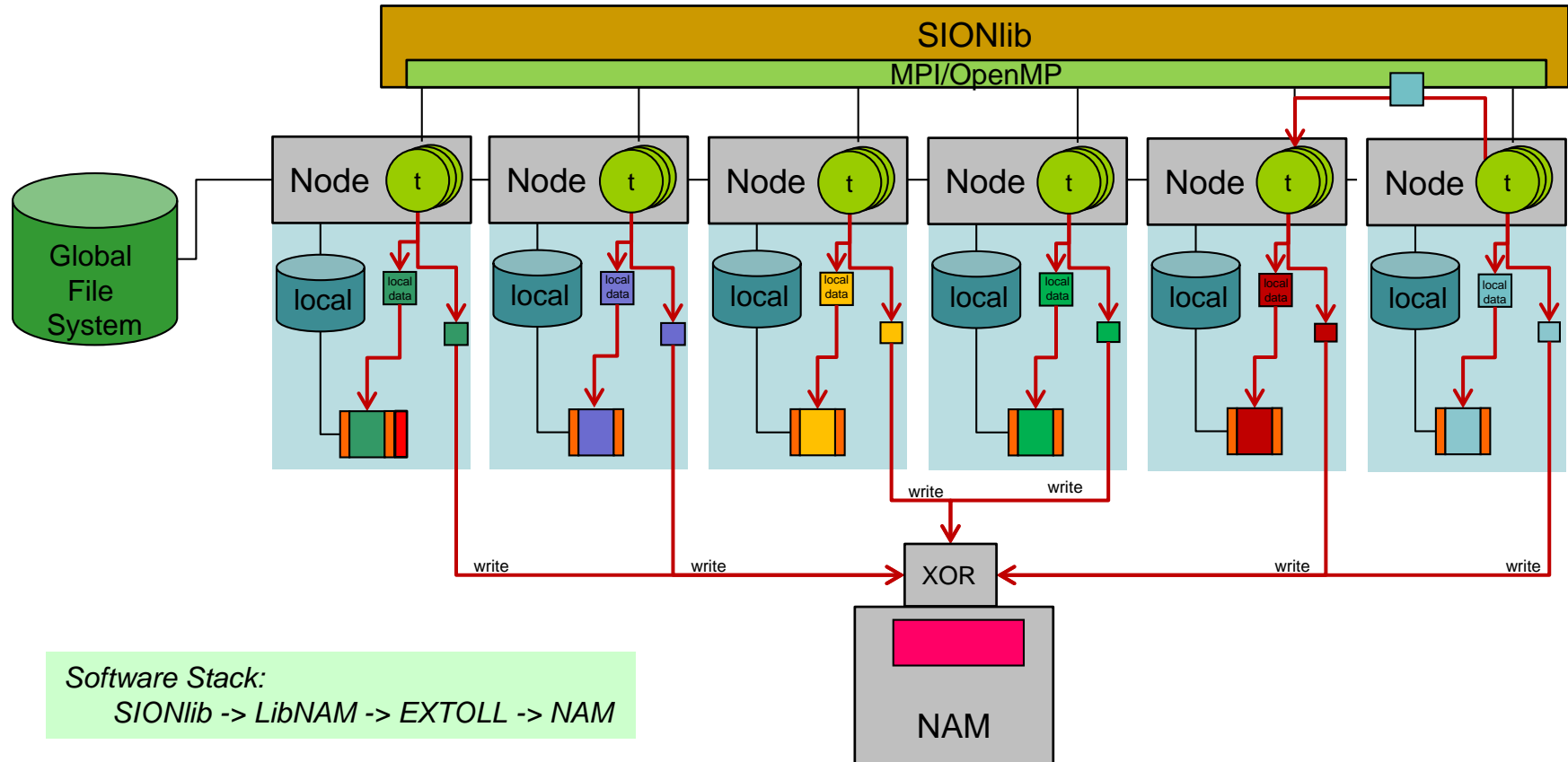
```
Open: sid=sion_paropen_mpi(...,"bw,buddy=x",MPI_COMM_WORLD, ... )  
Write: sion_coll_write_mpi(data,size,n,sid)  
Close: sion_parclose(sid)
```

Restore checkpoint after failure



```
Open: sid=sion_paropen_mpi(..., "br,buddy=x", MPI_COMM_WORLD, ... )  
Read: sion_coll_read_mpi(data,size,n,sid)  
Close: sion_parclose(sid)
```

SIONlib: Buddy-CP, NAM

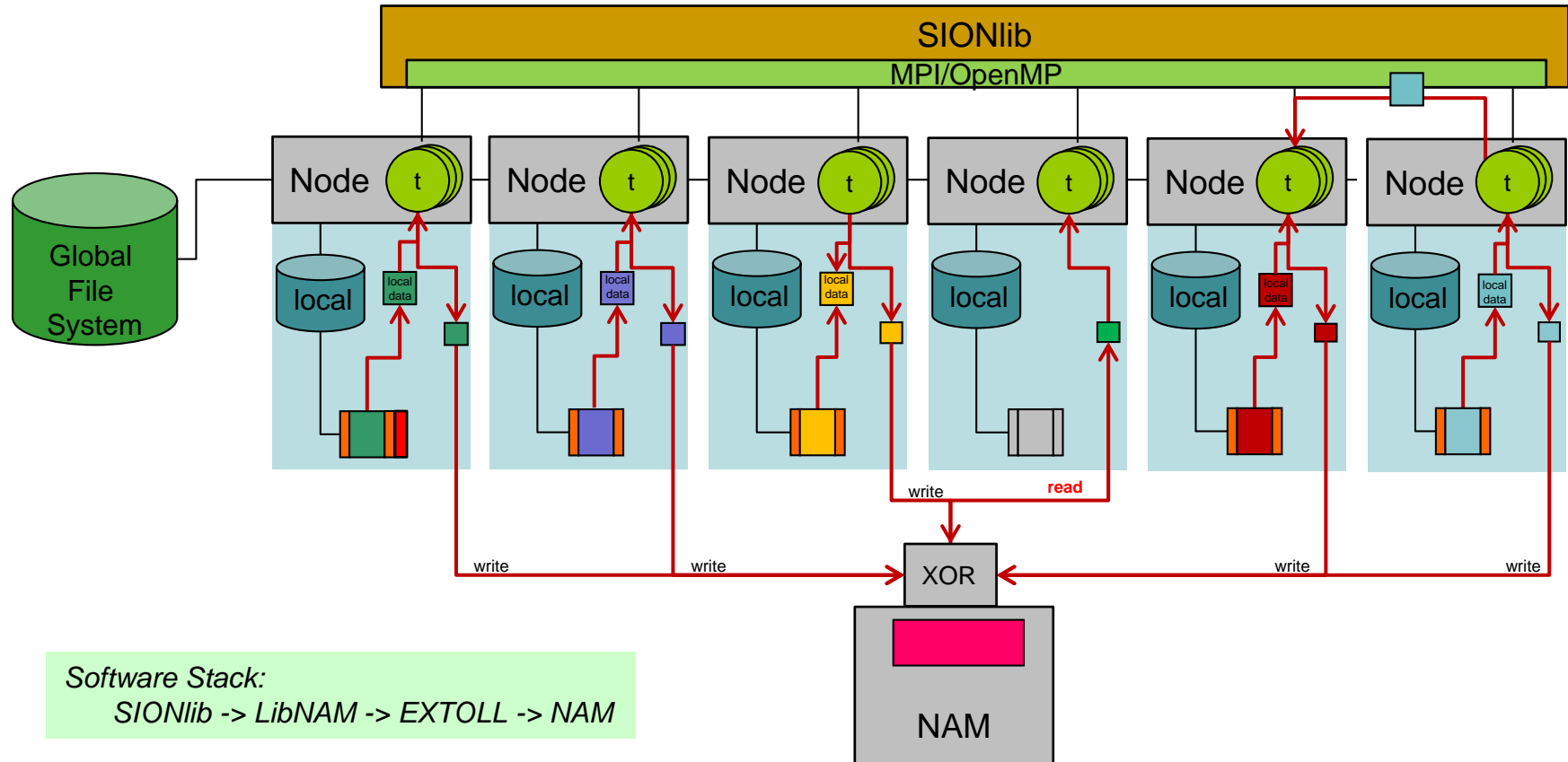


Software Stack:

SIONlib -> LibNAM -> EXTOLL -> NAM

```
Open: sid=sion_paropen_mpi(..., "bw,buddy=nam", MPI_COMM_WORLD, ... )  
Write: sion_coll_write_mpi(data,size,n,sid)  
Close: sion_parclose(sid)
```


SIONlib: Buddy-CP, NAM, after failure



Software Stack:

SIONlib -> LibNAM -> EXTOLL -> NAM

```
Open: sid=sion_paropen_mpi(..., "br,buddy=nam", MPI_COMM_WORLD, ... )  
Read: sion_coll_read_mpi(data,size,n,sid)  
Close: sion_parclose(sid)
```