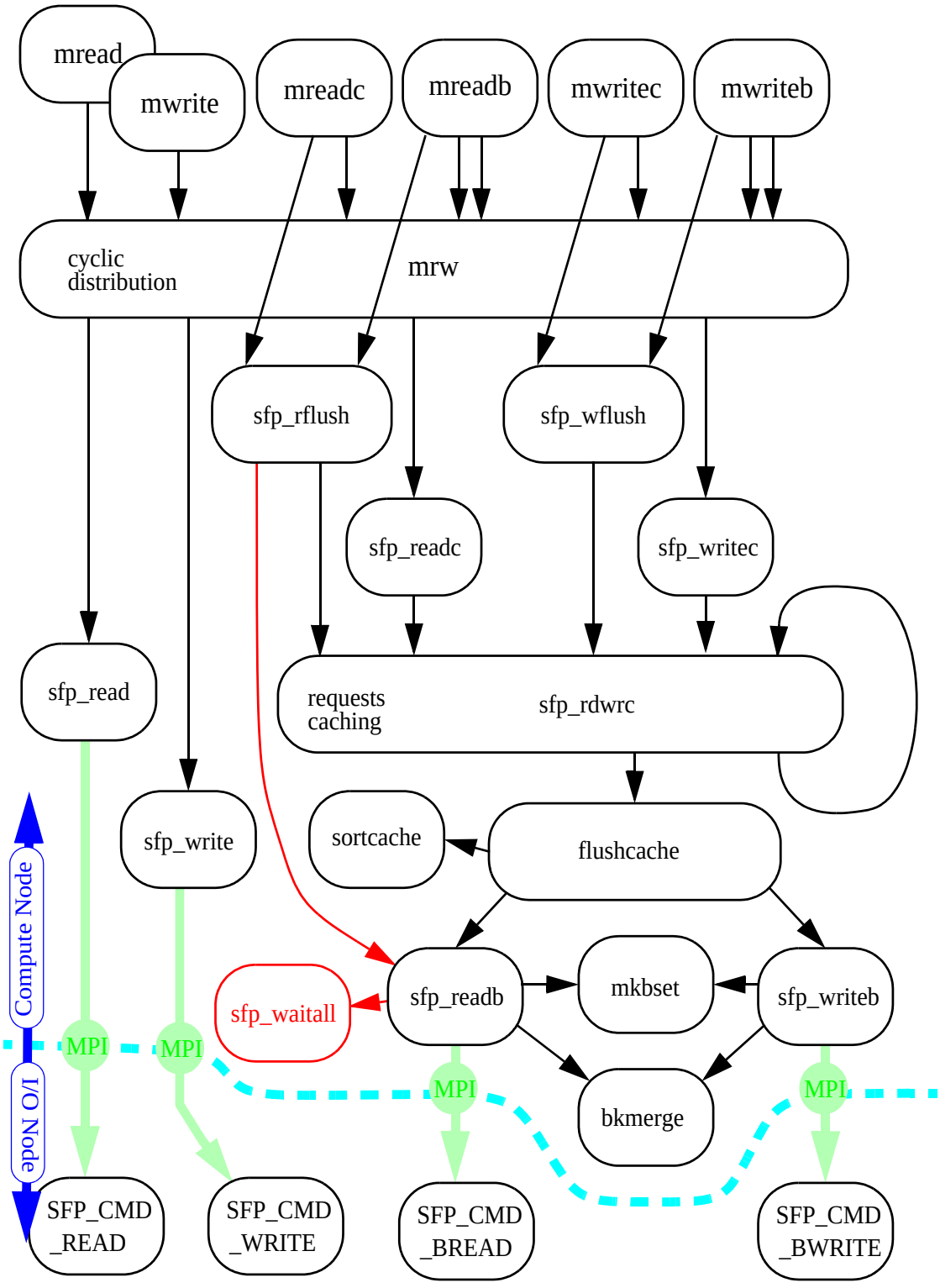
SFIO progress on Swiss-Tx

SCS meeting on Frangipani: a scalable distributed file system to Linux December 1, 2000

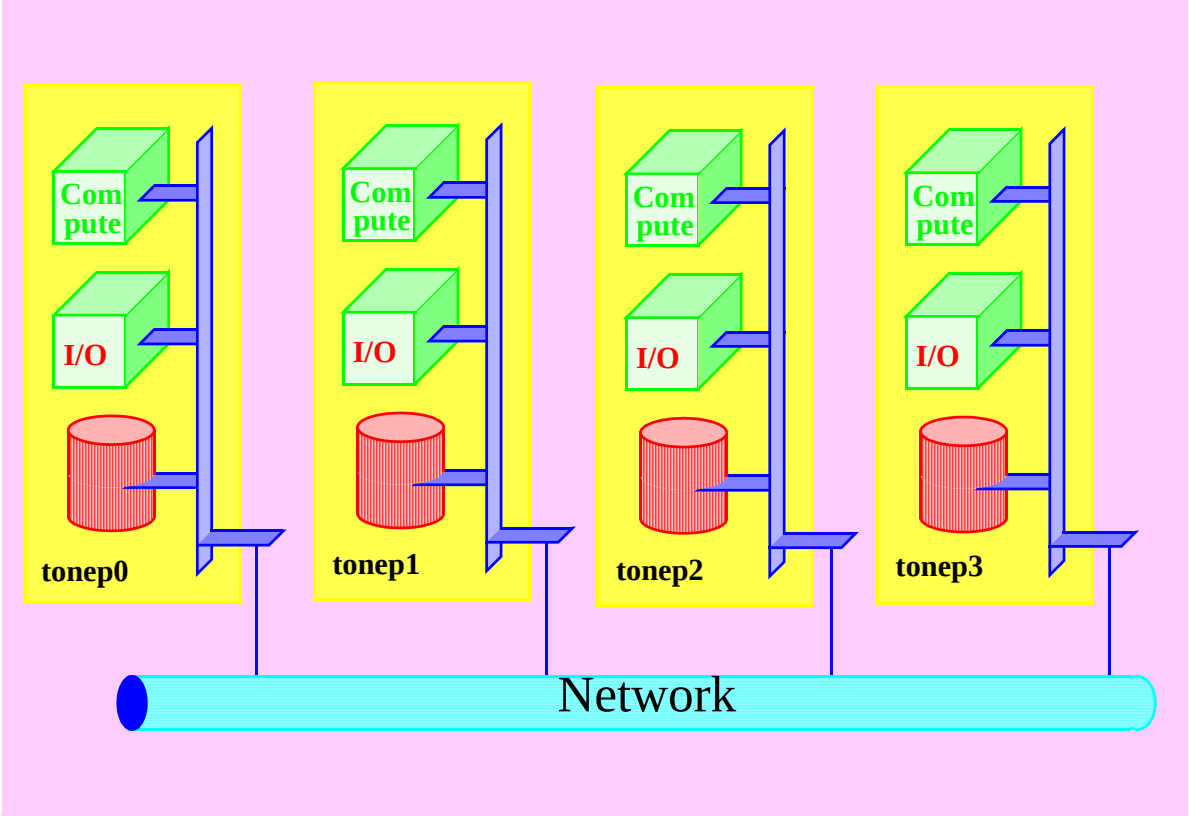
Emin Gabrielyan

EPFL, Computer Science Dept.
Peripheral Systems Lab.
Emin.Gabrielyan@epfl.ch

- SFIO library architecture
- SFIO on top of MPICH and on top of MPIFCI, performance on T1
- performance of SFIO on top of MPIFCI on T1. Very large files, no cache effect.
- Swiss-T1's topology. Possible influence to the SFIO performance
- Conclusion
- Future work


mread
mwrite
mreadc
mreadb
mwritec
mwriteb
cyclic
distribution
mrw
sfp_rflush
sfp_wflush
sfp_readc
sfp_writec
sfp_read
requests
caching
sfp_rdwrc
sfp_write
sortcache
flushcache
Compute Node
I/O Node
sfp_waitall
sfp_readb
mkbset
sfp_writeb
MPI
MPI
MPI
MPI
bkmerge
SFP_CMD
_READ
SFP_CMD
_WRITE
SFP_CMD
_BREAD
SFP_CMD
_BWRITE

-

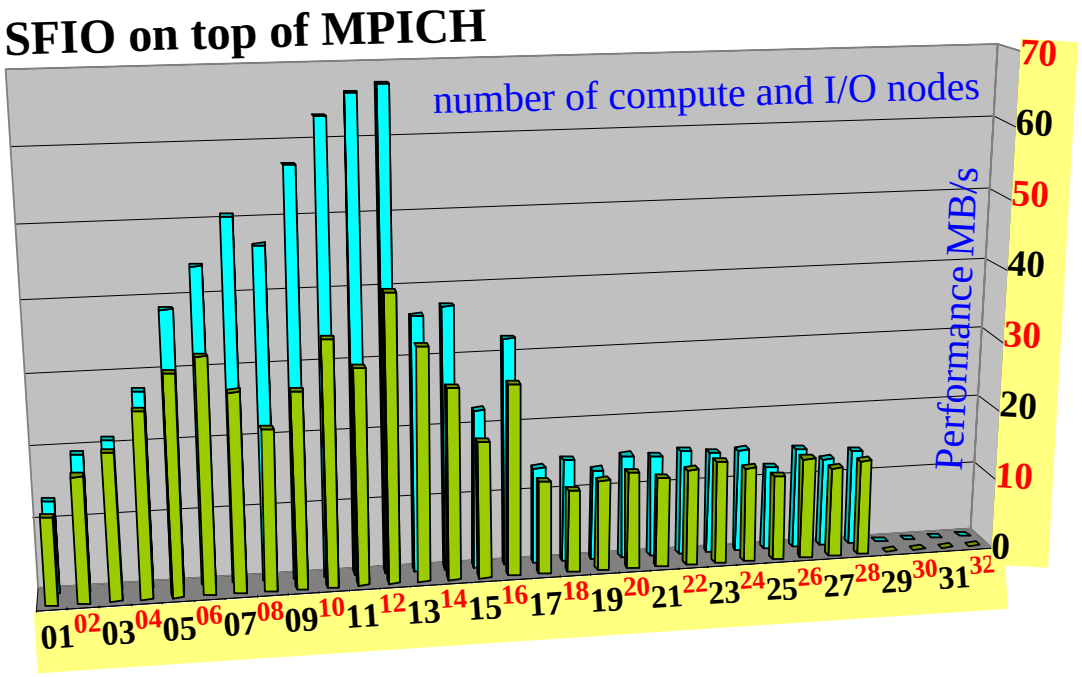


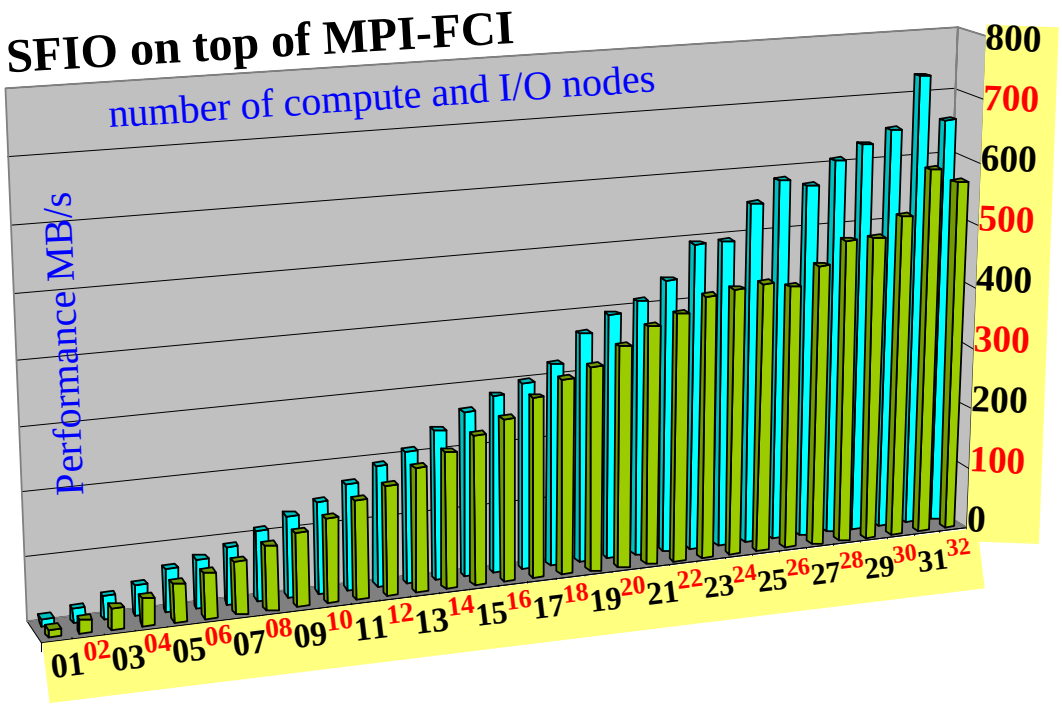
- SFIO All-to-All concurent write access from all compute nodes to all I/O nodes

- Global File size is 2000MByte for MPICH and MPIFCI.

- Stripe unit size is 200Byte only

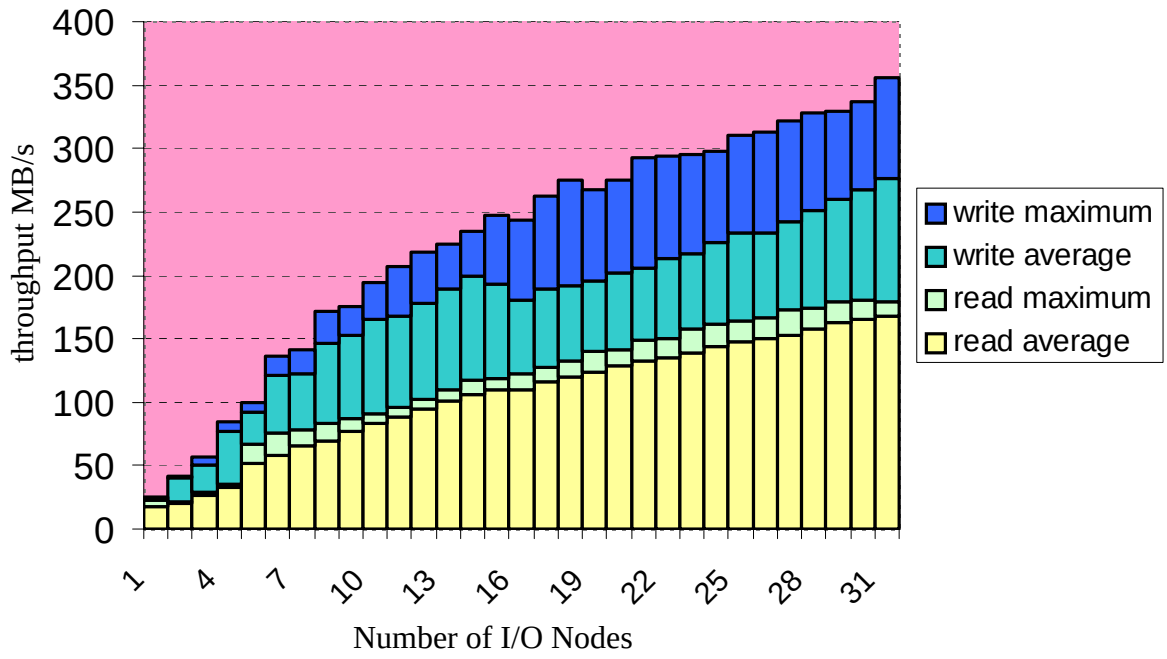
SFIO all-to-all I/O performance on Swiss-T1's Fast Ethernet and Tnet



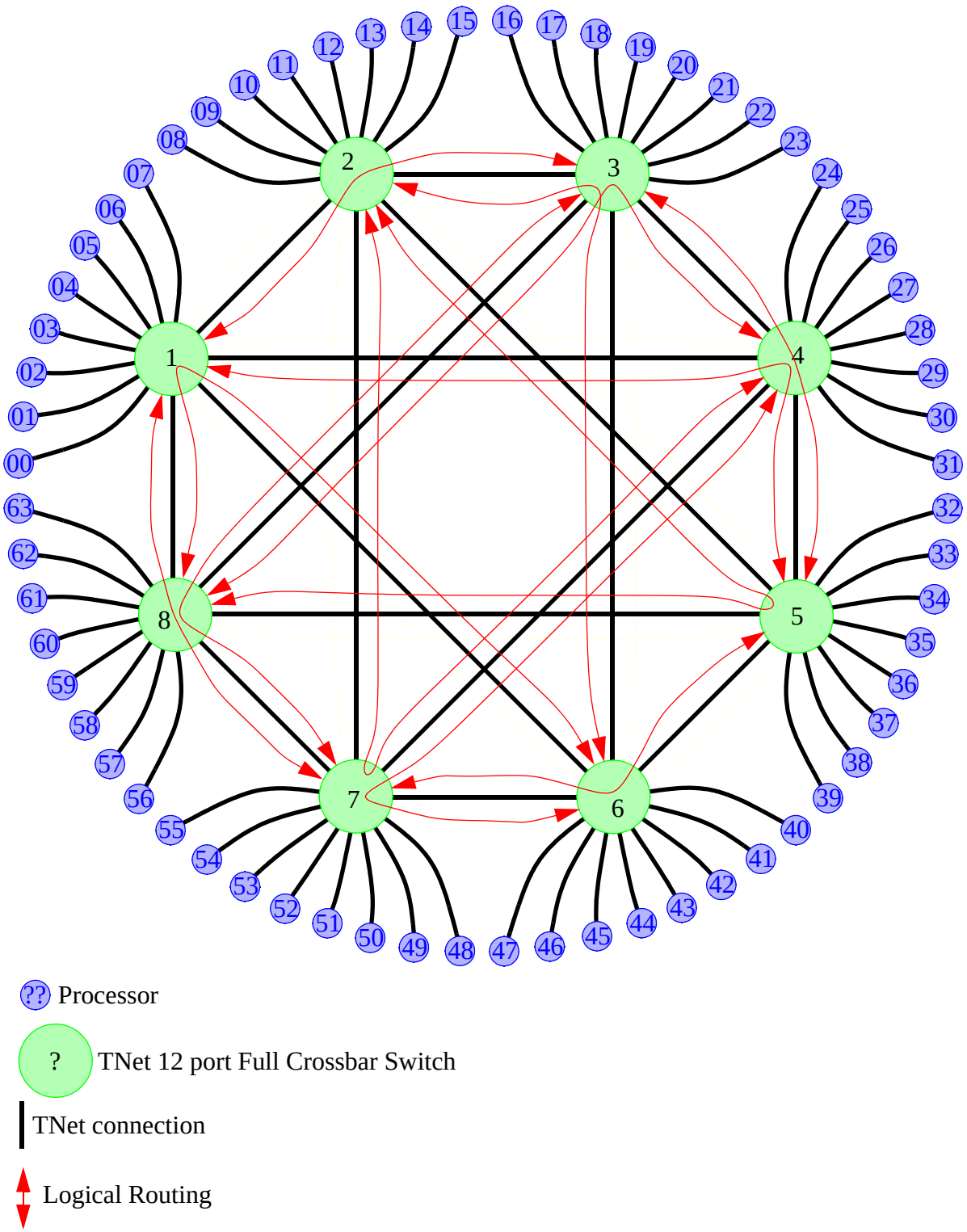


- Superlinear speedup of SFIO/FCI due to augmentation of cache effect when increasing the number of I/O nodes.

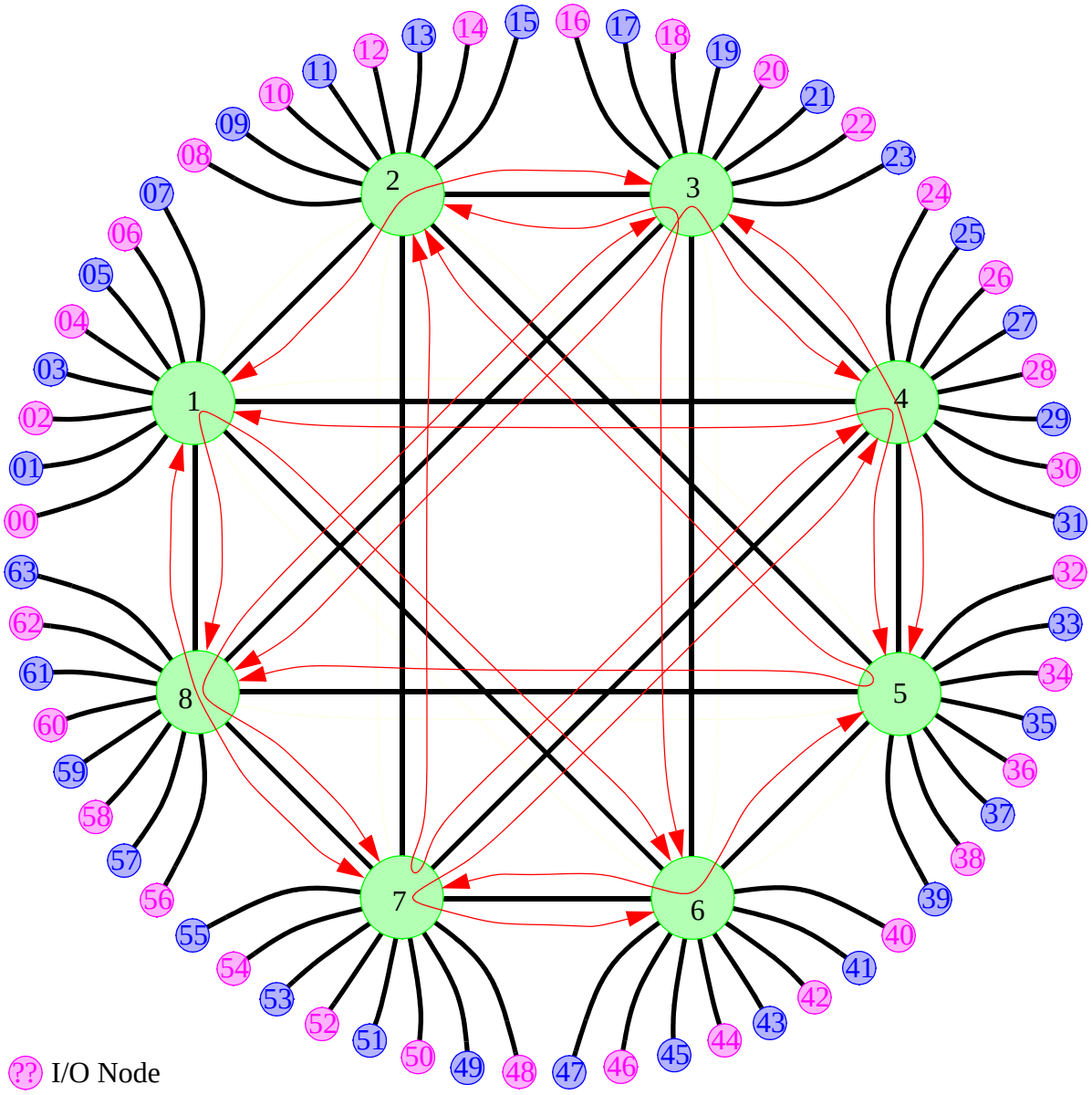
SFIO All-to-all performance on T1.
(1GB-31GB file size, 200Byte chunk, 53 measurements)



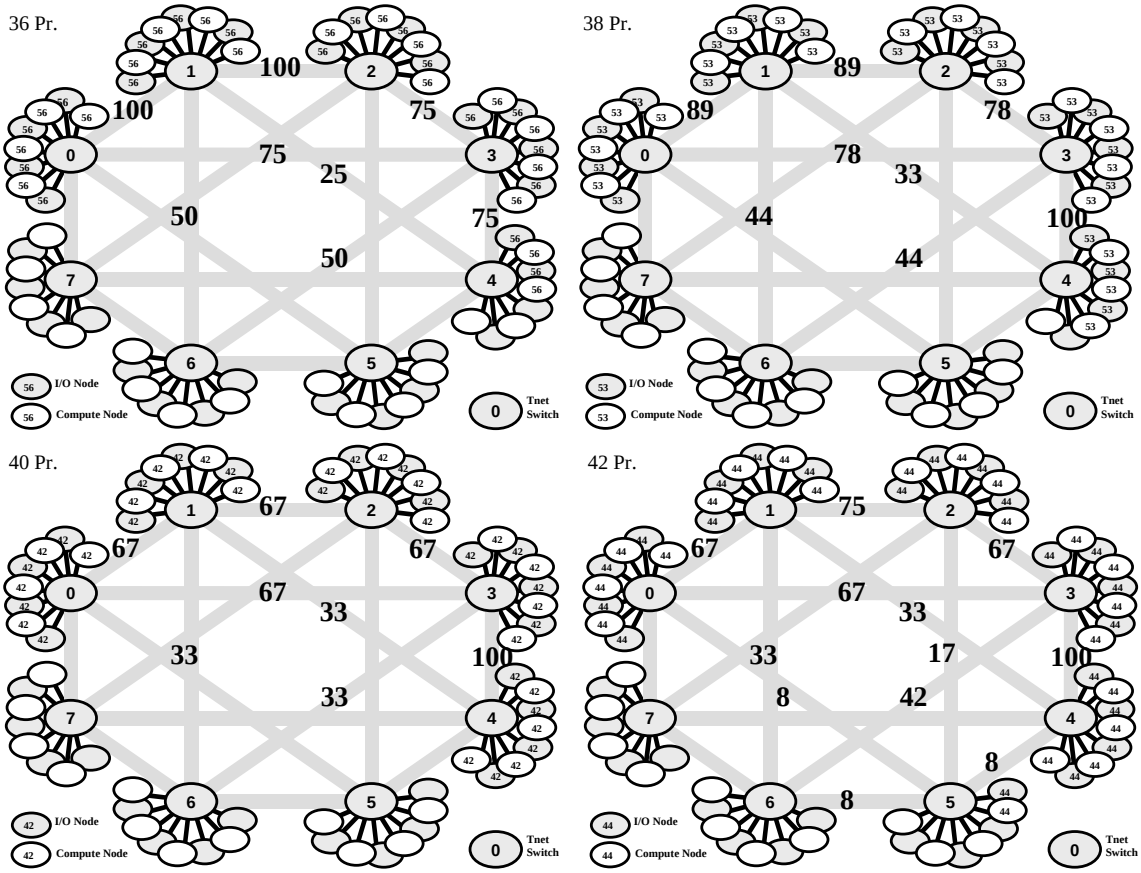
- To avoid the cache effect the total size of SFIO files is increasing when the number of I/O nodes grows.



Swiss-T1 TNet interconnection and routing topology

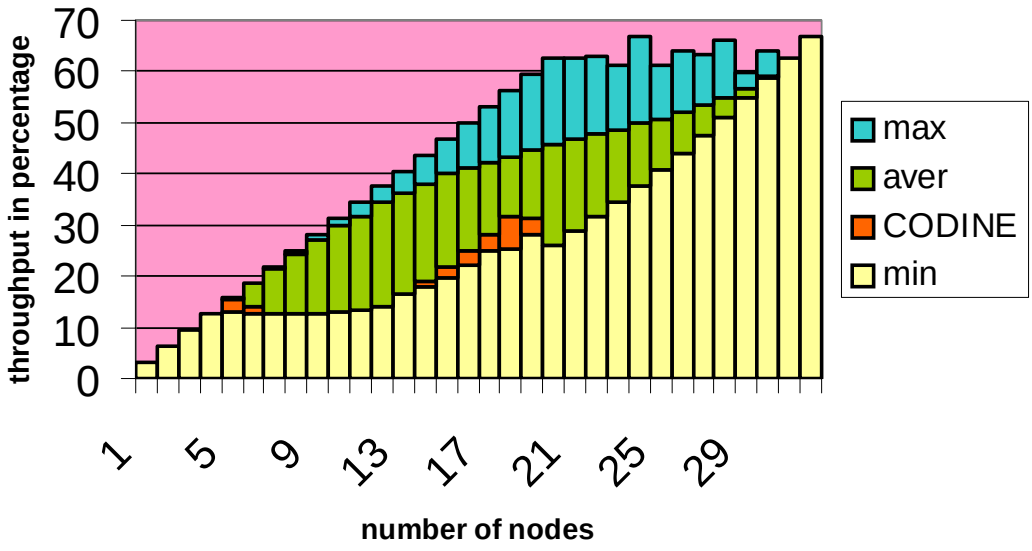


Swiss-T1 SFIO over TNet topology



connection loads

Theoretical throughput of the Swiss-T1 network as a percentage of ideal throughput of fully crossbared switch. 1008000 montecarlo events of parallel simulation. The min represent the worst topology and the max the best



Theoretical throughput

Conclusion

- SFIO is portable, highly scalable, and ready for the distribution.

Future work

- SFIO performance benchmarking on the large supercomputer of Sandia National Laboratory.

- Performance measurements of MPI-I/O interfaced to SFIO through MPICH/ADIO.

- Possibly, creation of a portable MPI-I/O interface library to SFIO.