

Master Thesis

Concurrent I/O

from

Xeon Phi Accelerator Cards

Romain Monnard

Professor: Frédéric Bapst
Supervisors: Paolo Calafiura
Wim Lavrijsen
Expert: Michel Yerly

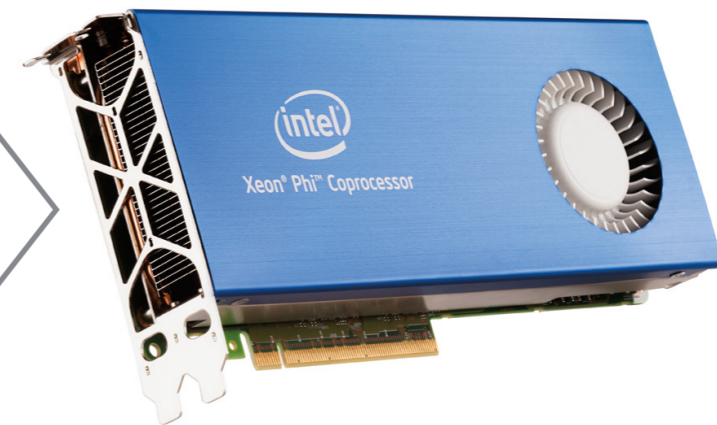


01/26/2015

Objectives

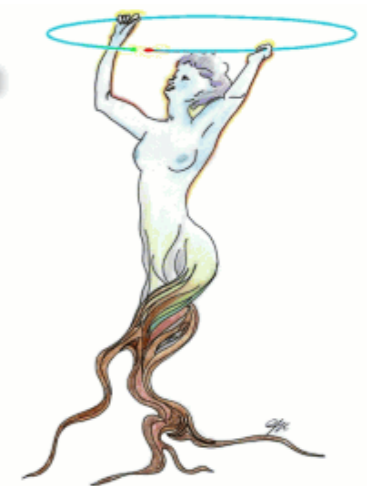


Geant 4



ROOT

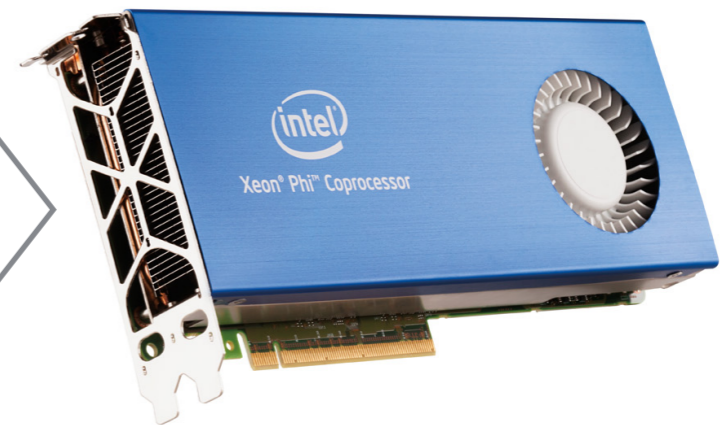
An Object-Oriented
Data Analysis Framework



Xeon Phi



PCI-Express 2
8GB/s



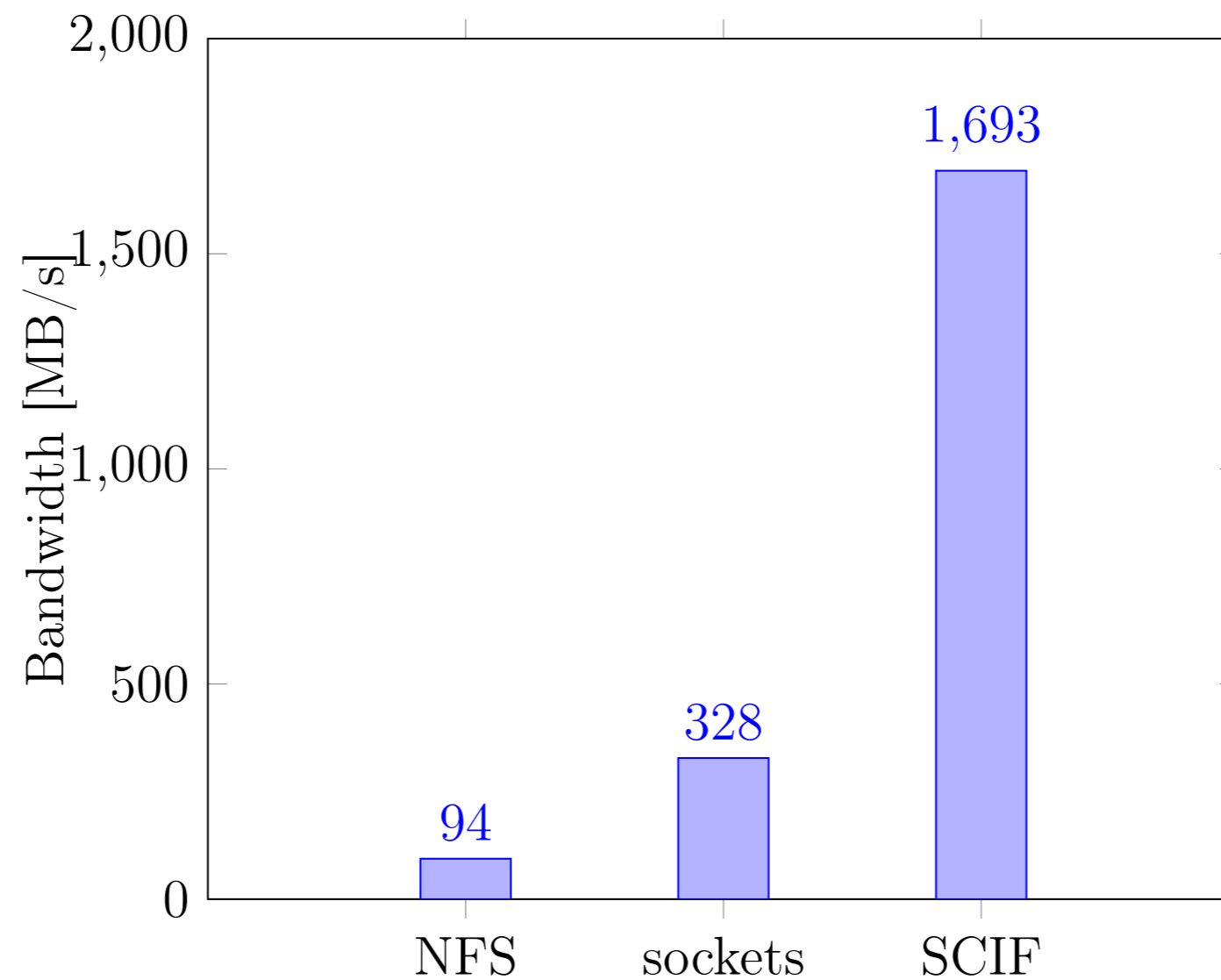
Technologies

Sockets

NFS

Intel SCIF

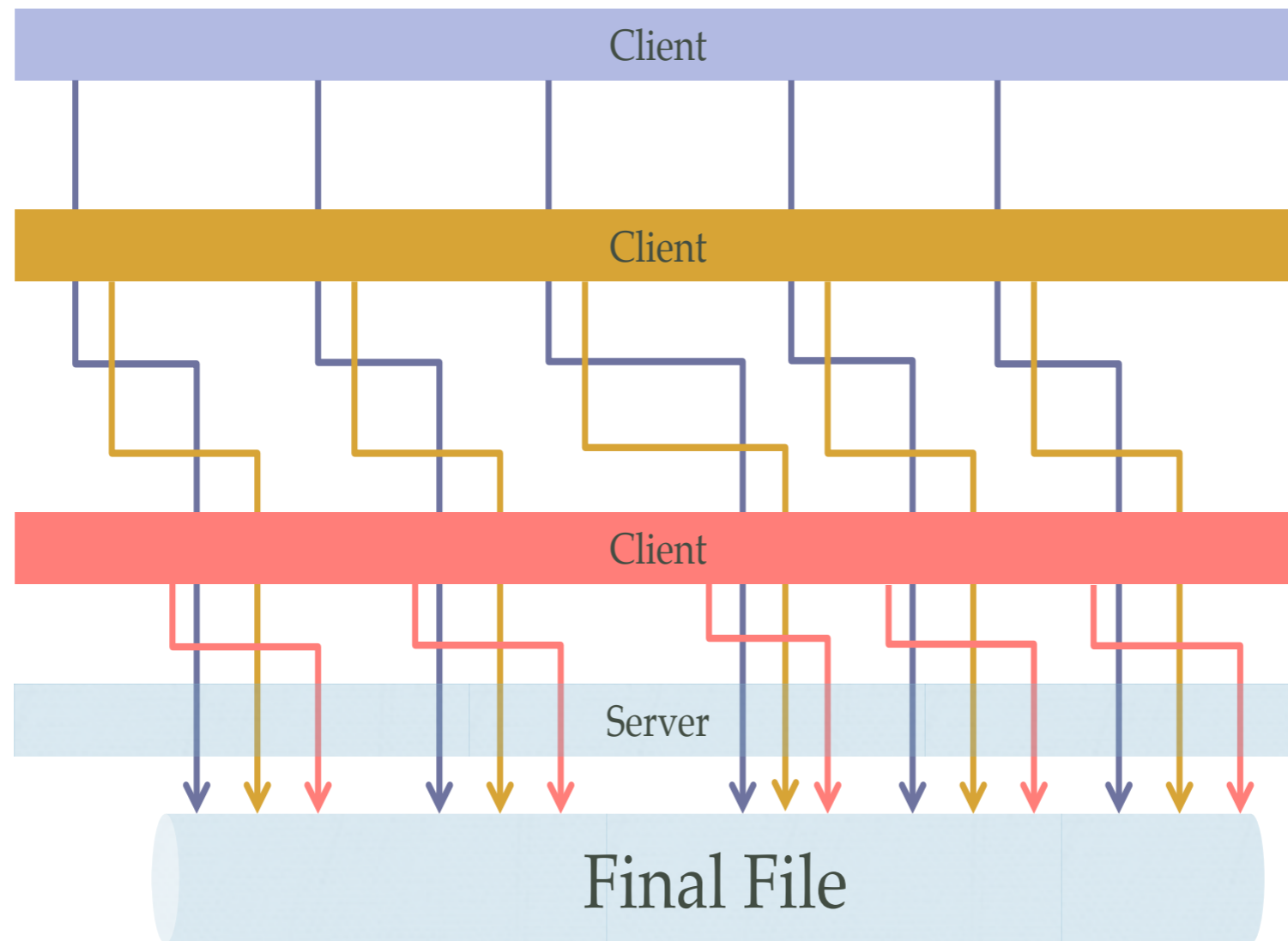
Technologies



ROOT - Sockets

- TSocket - TServerSocket
- TMemFile
 - A completely in-memory version of TFile
- TParallelMergingFile
 - A TMemFile that on a call to Write will upload its content and reset the TTree objects.

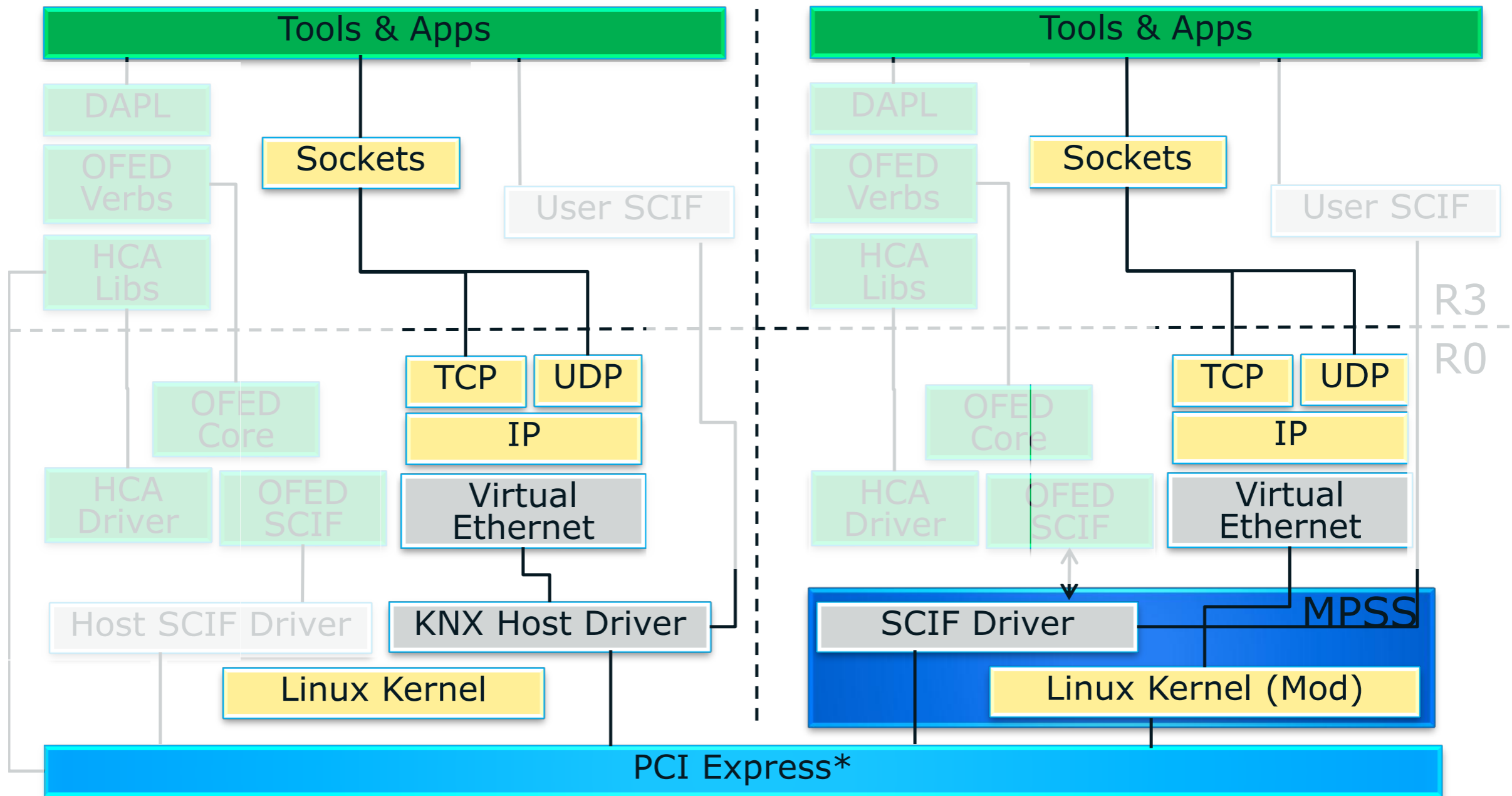
TParallelMergingFile



Sockets vs SCIF

Host

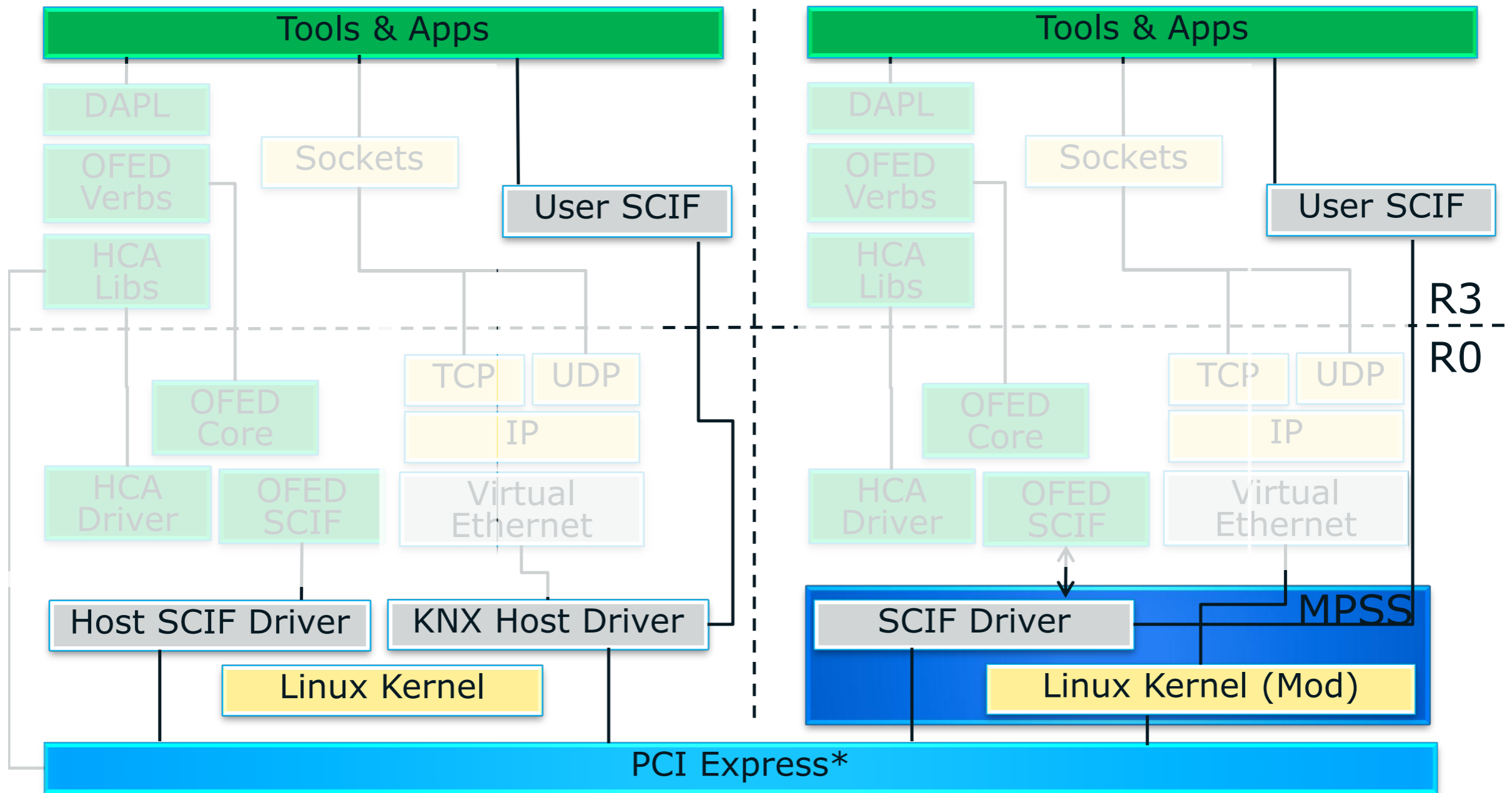
Xeon Phi



Sockets vs SCIF

Host

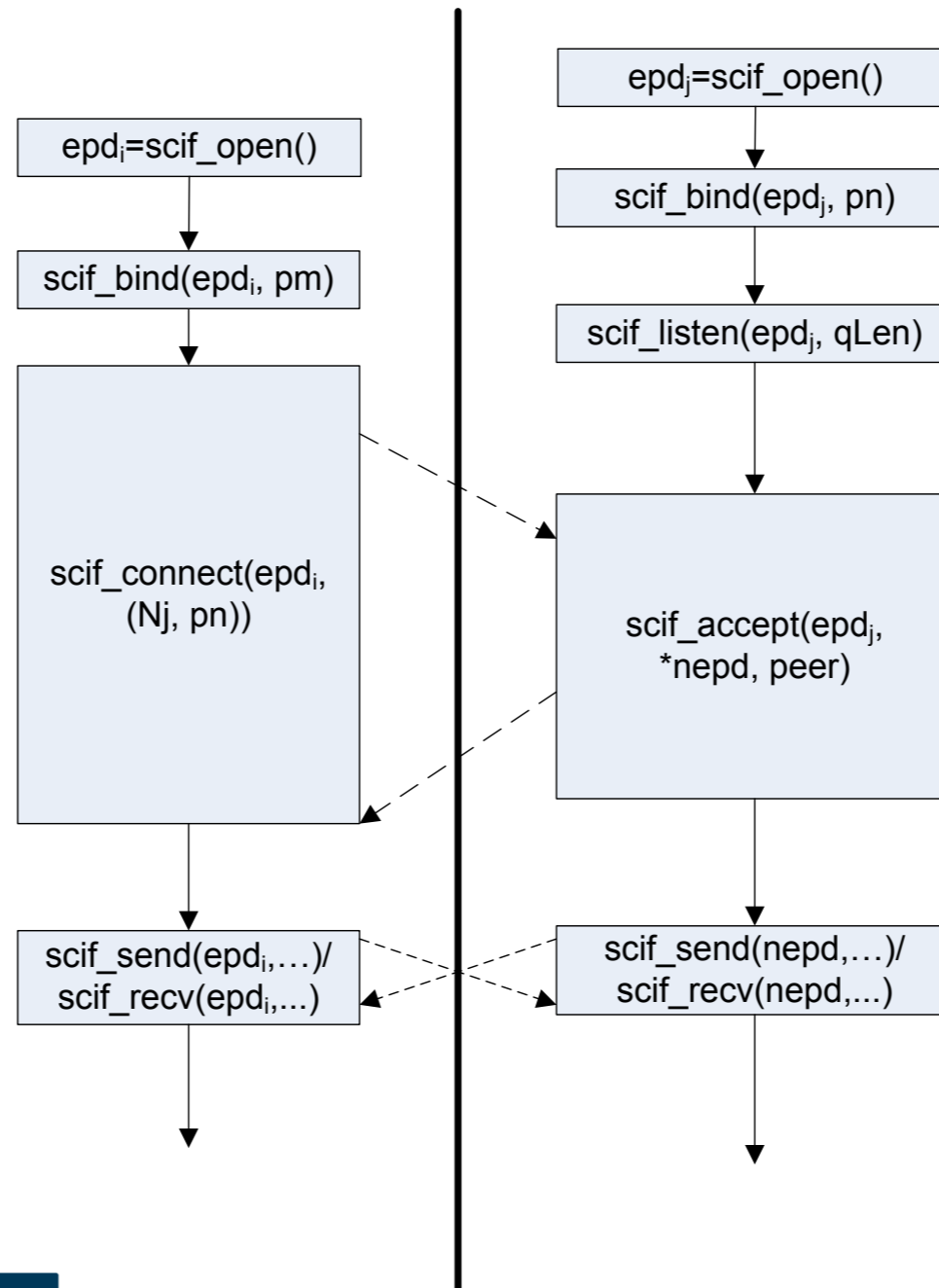
Xeon Phi



Intel SCIF

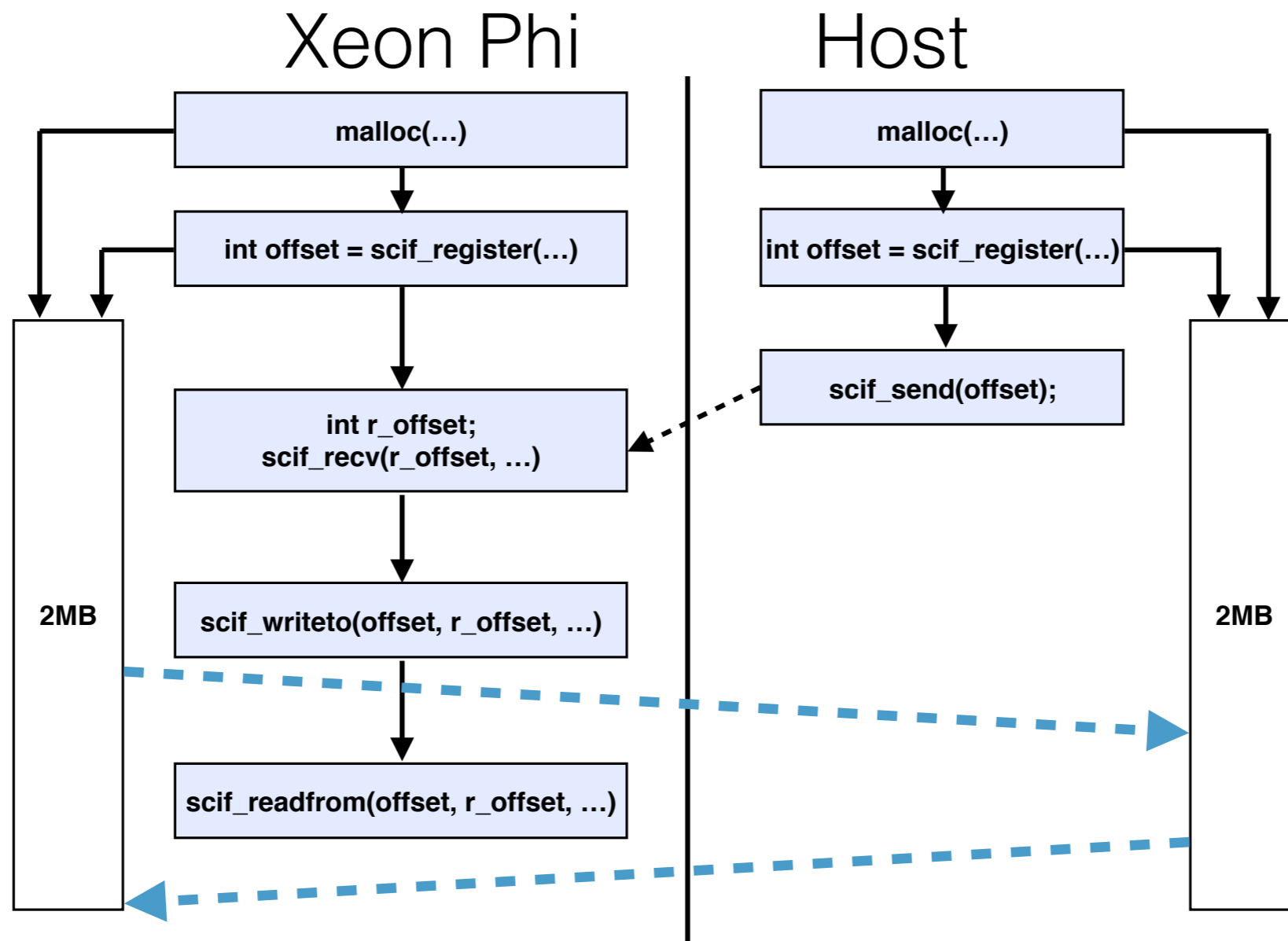
Xeon Phi

Host



Intel SCIF RMA

- Remote Memory Access

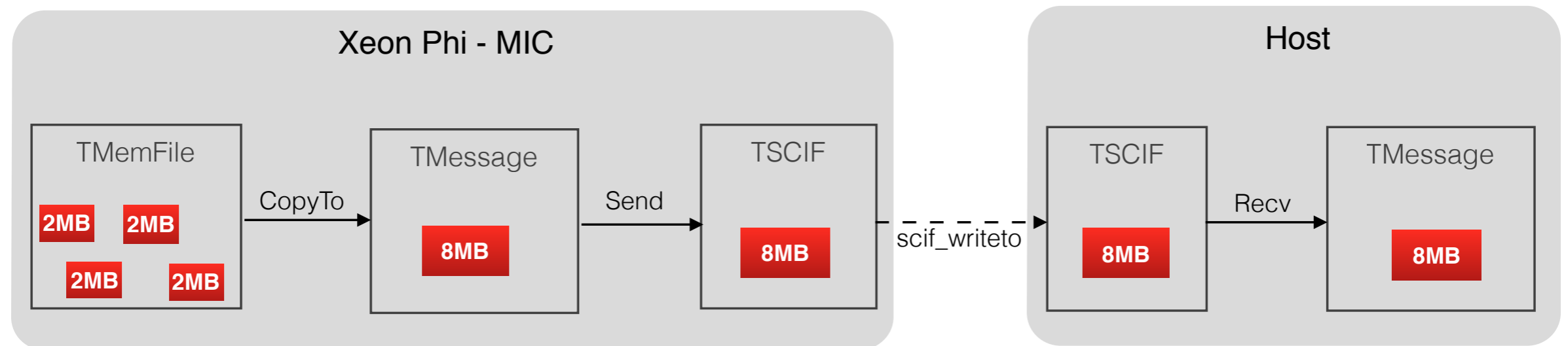


TSCIF

- Implementation of SCIF in ROOT
 - TSCIF / TServerSCIF
 - Offer the same API as TSocket / TServerSocket

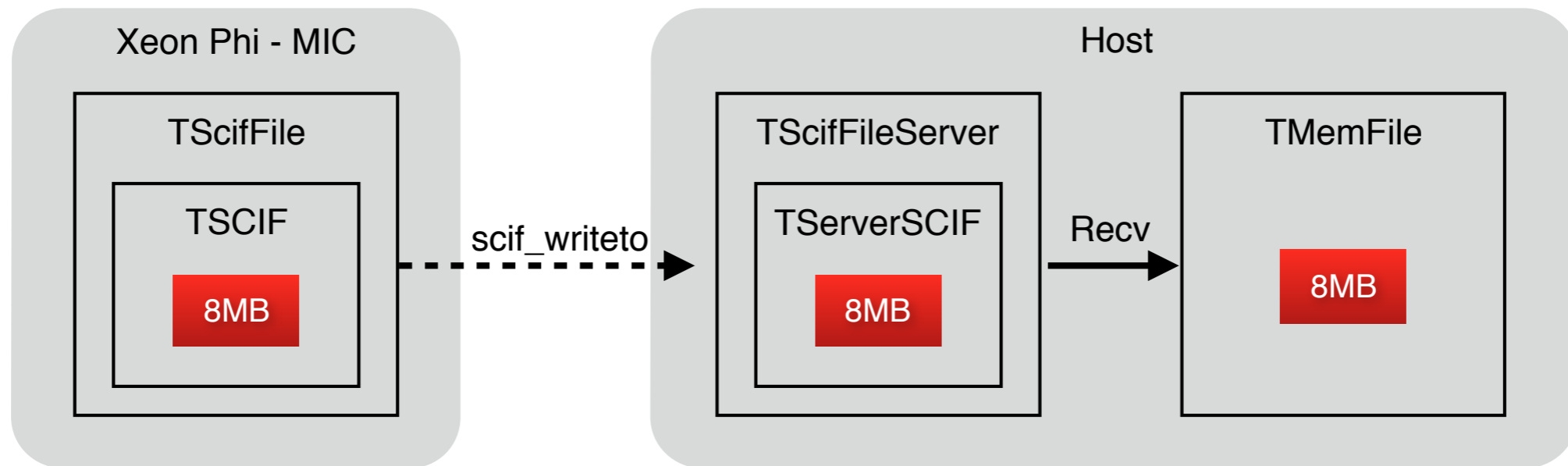
TSCIF

- TParallelMergingFile based on TSCIF

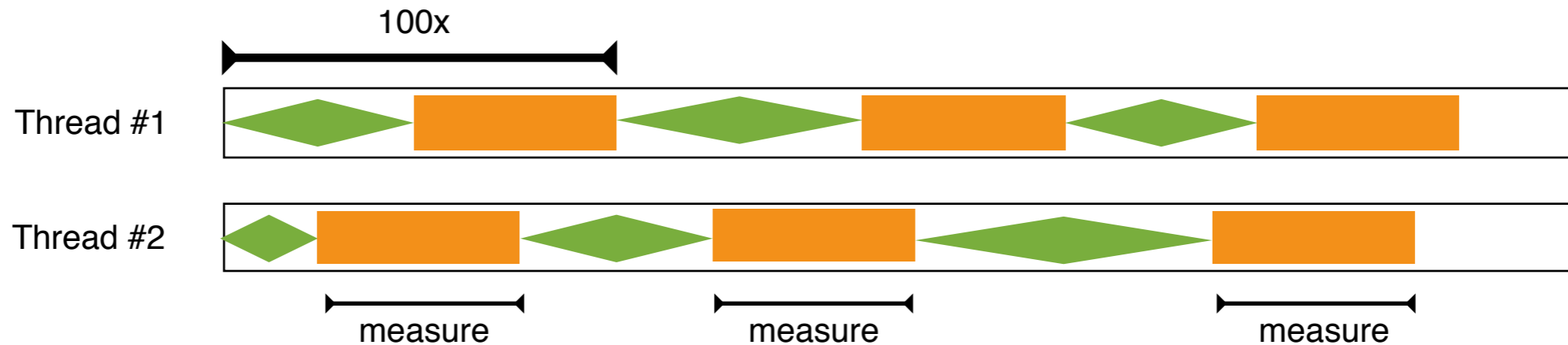



- $240 \text{ threads} * 8\text{MB} * 3 = 5760\text{MB}$
- Only 8GB of memory on the card

TSCIFFile



Measurements

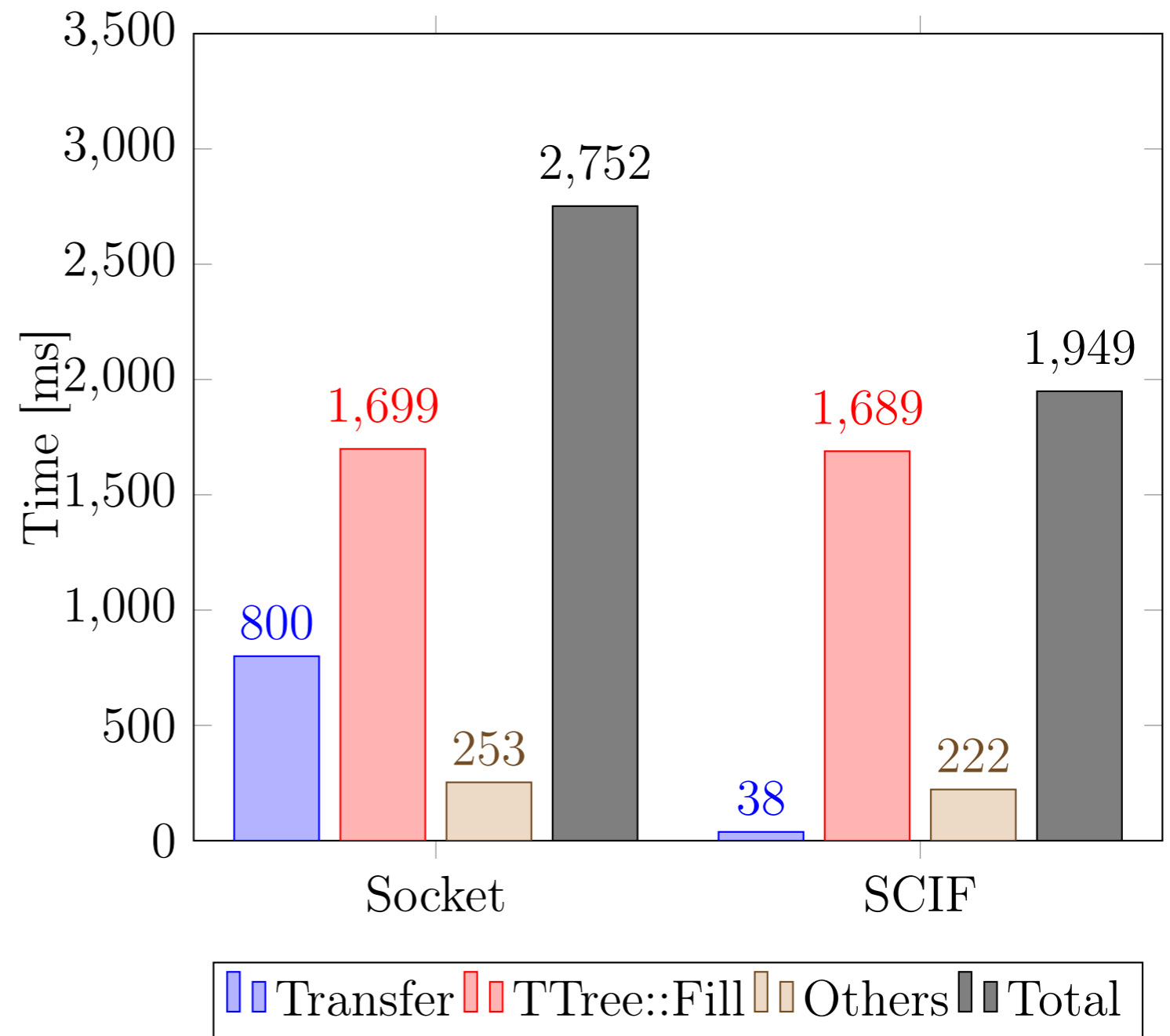


 Simulate computations of Geant4.
Busy waiting of 10 to 15s (random)



Results

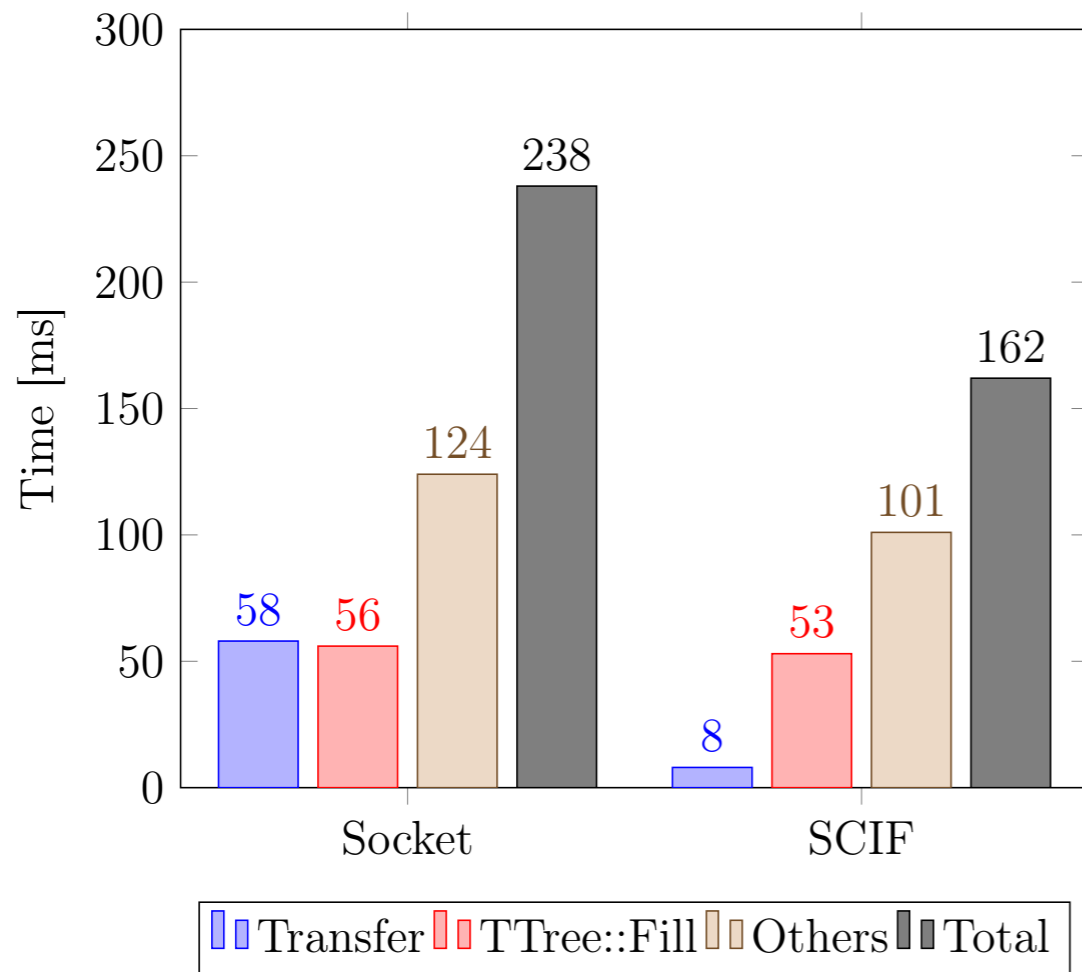
- 60 threads
- 100 x 2MB files
- Throughput:
 - Socket: 250 MB/s
 - SCIF: 5263 MB/s



Results

60 threads

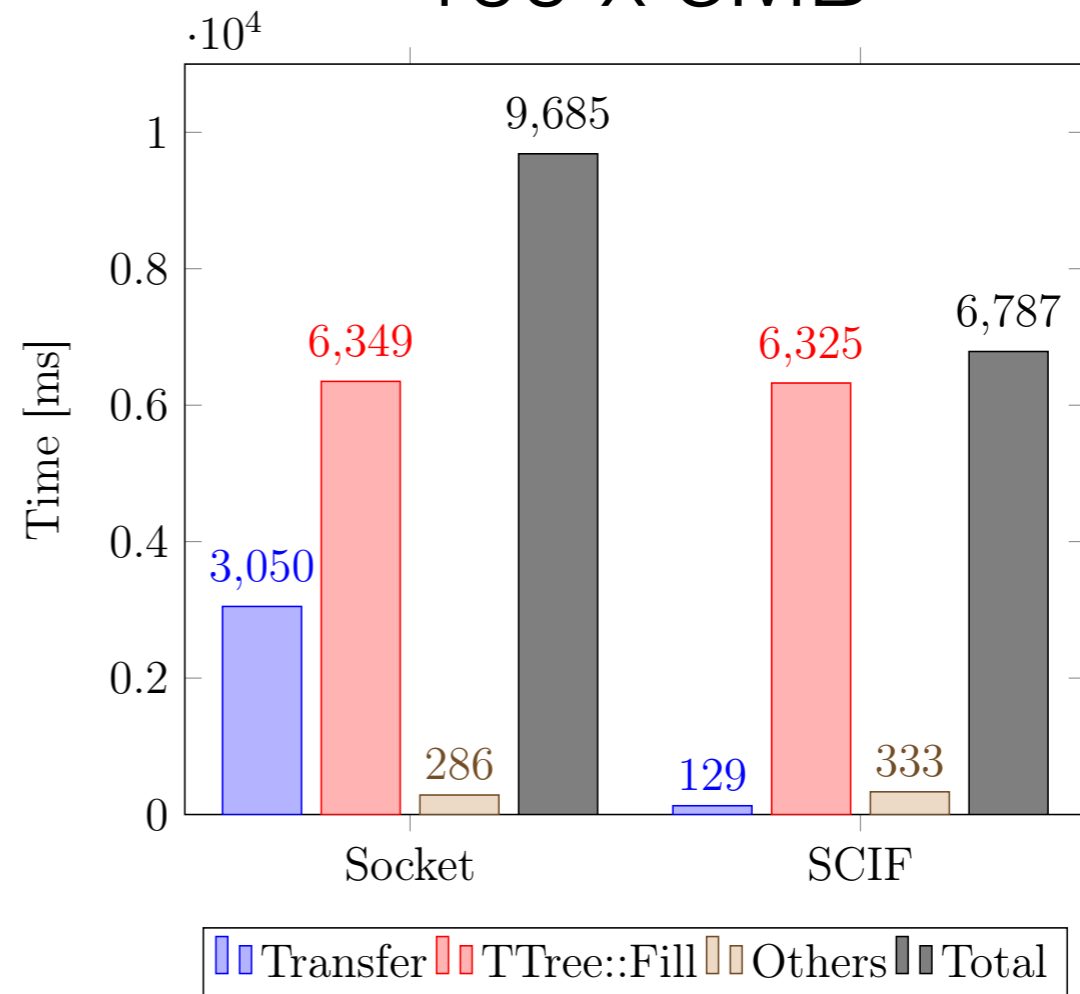
100 x 64kB



107MB/s

780MB/s

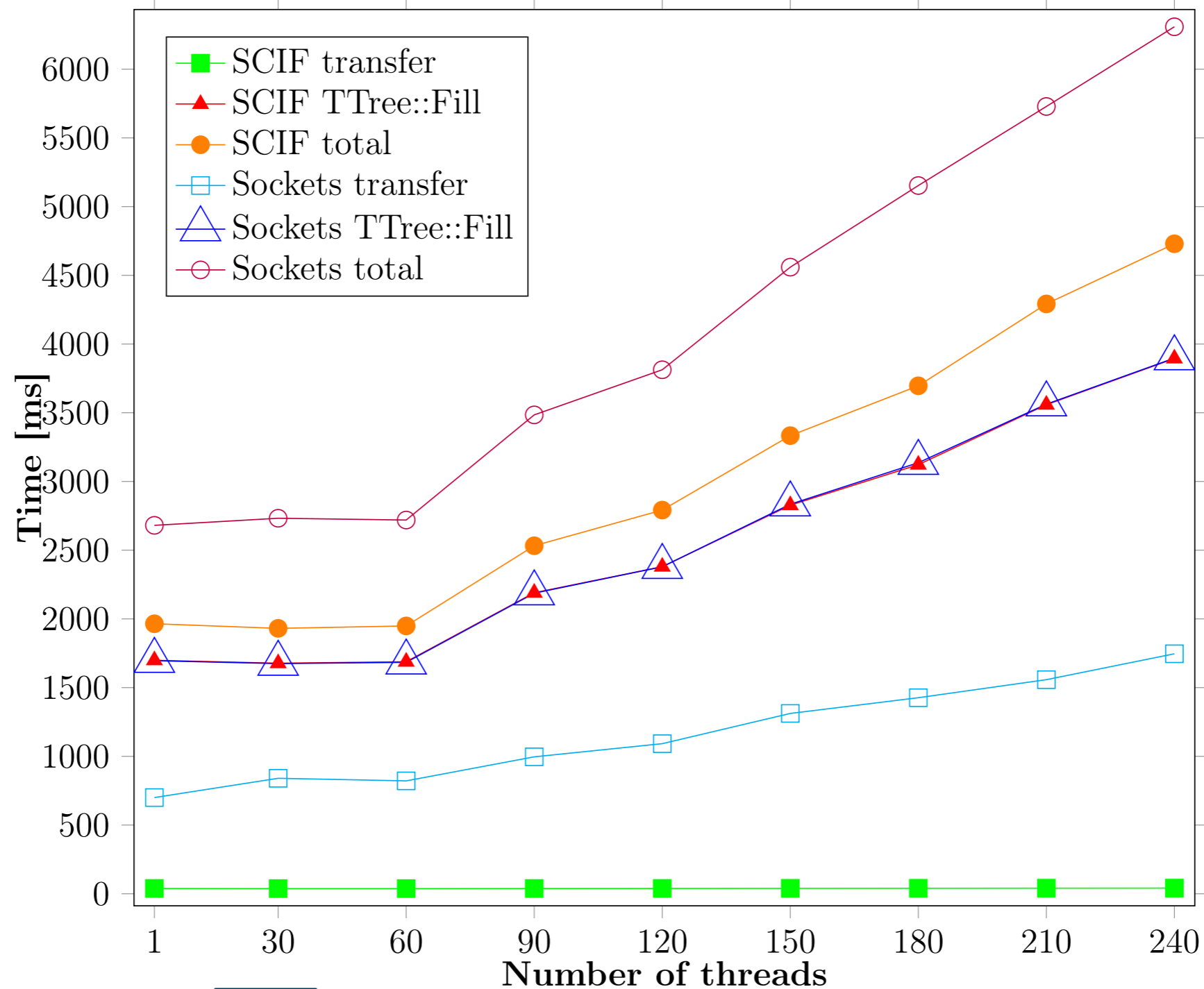
100 x 8MB



262MB/s

6200MB/s

Results - 1 to 240 threads

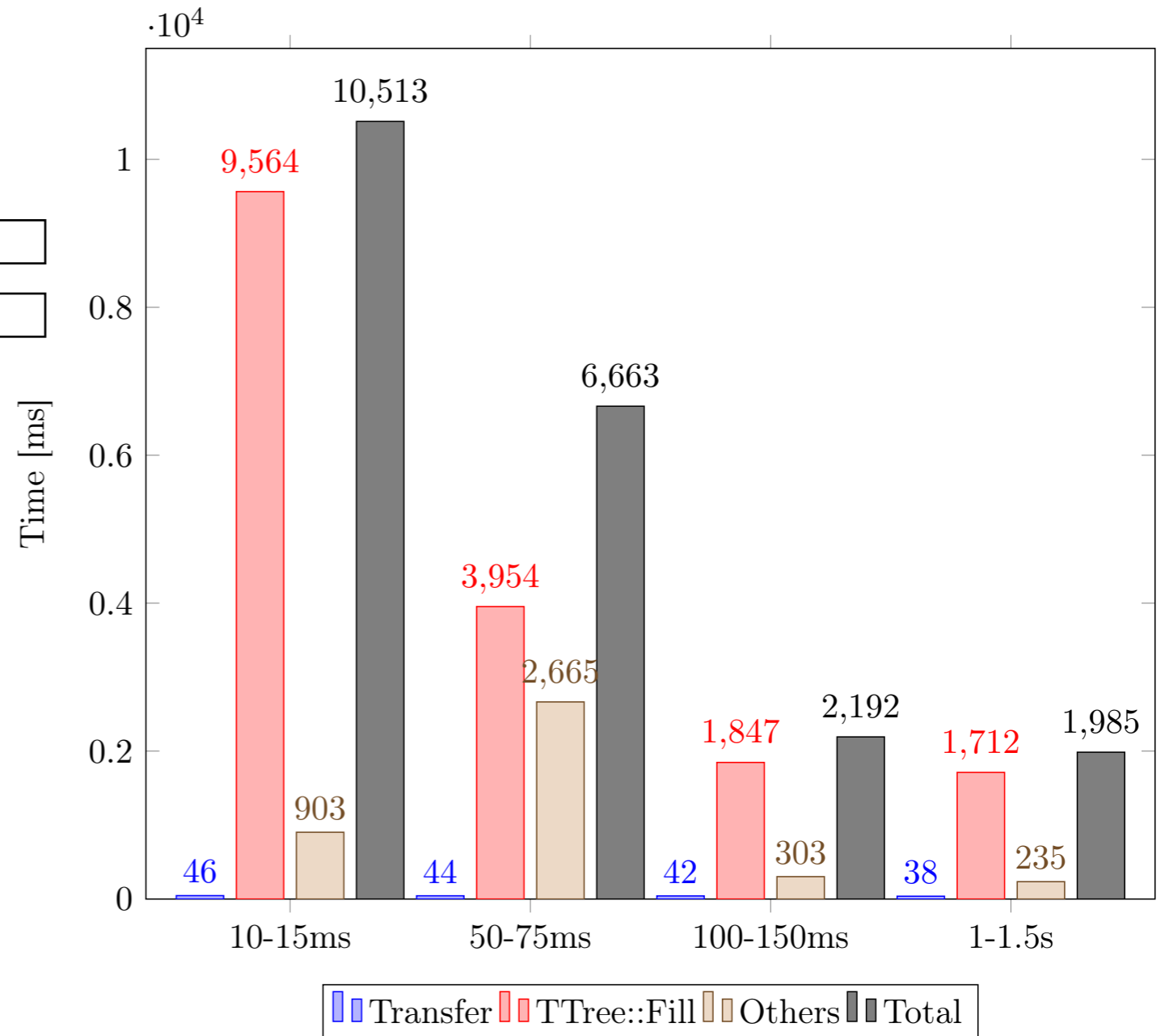
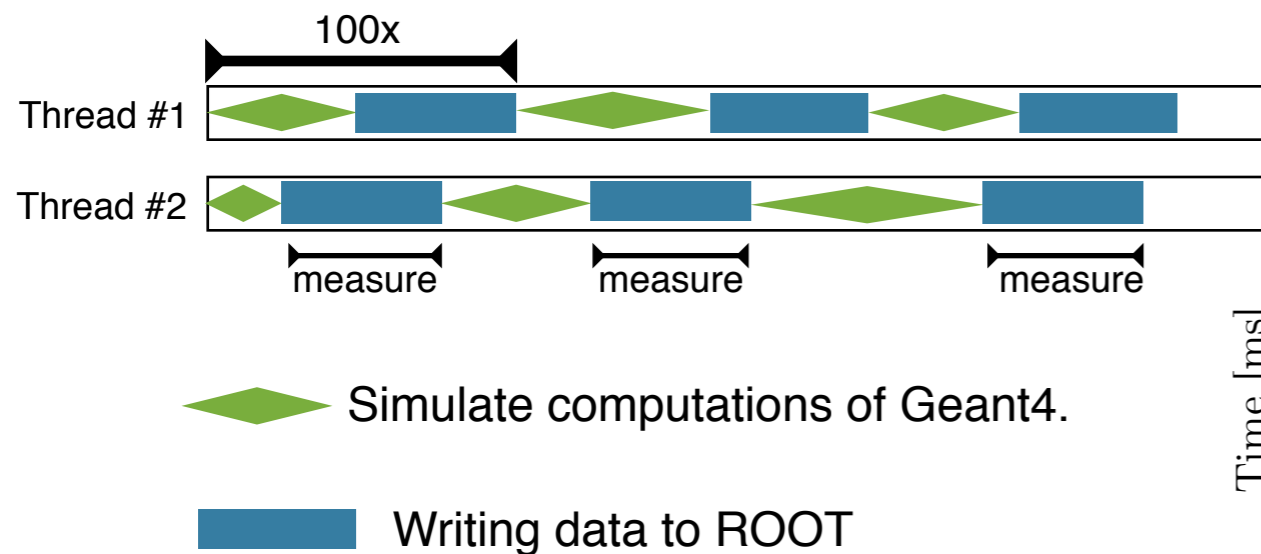


2MB files

Reduce computation time

60 threads

100 x 2MB files



Conclusion

- SCIF
 - Best throughput
 - 21 x faster than sockets
- TSCIFFile
 - Reduction of memory usage for the Xeon Phi

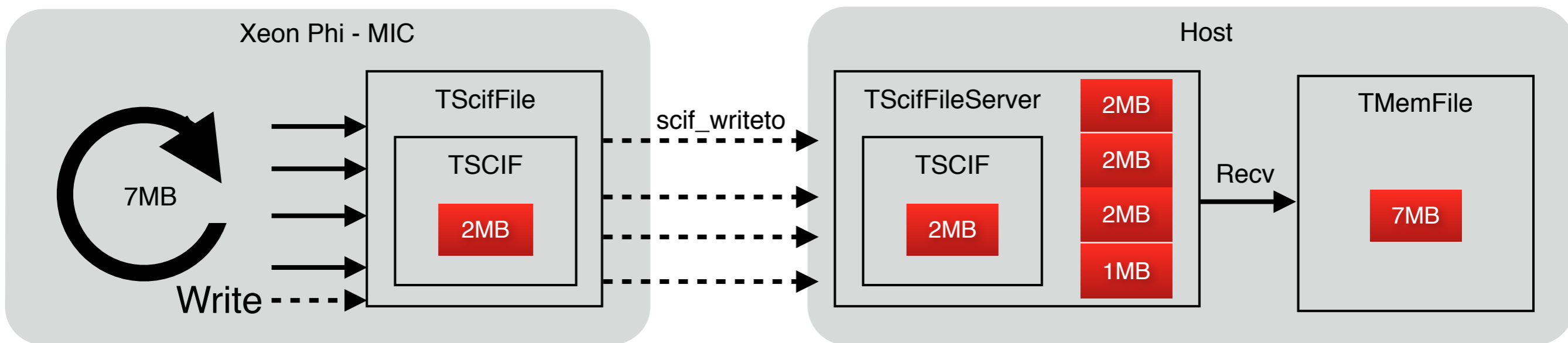
Future work

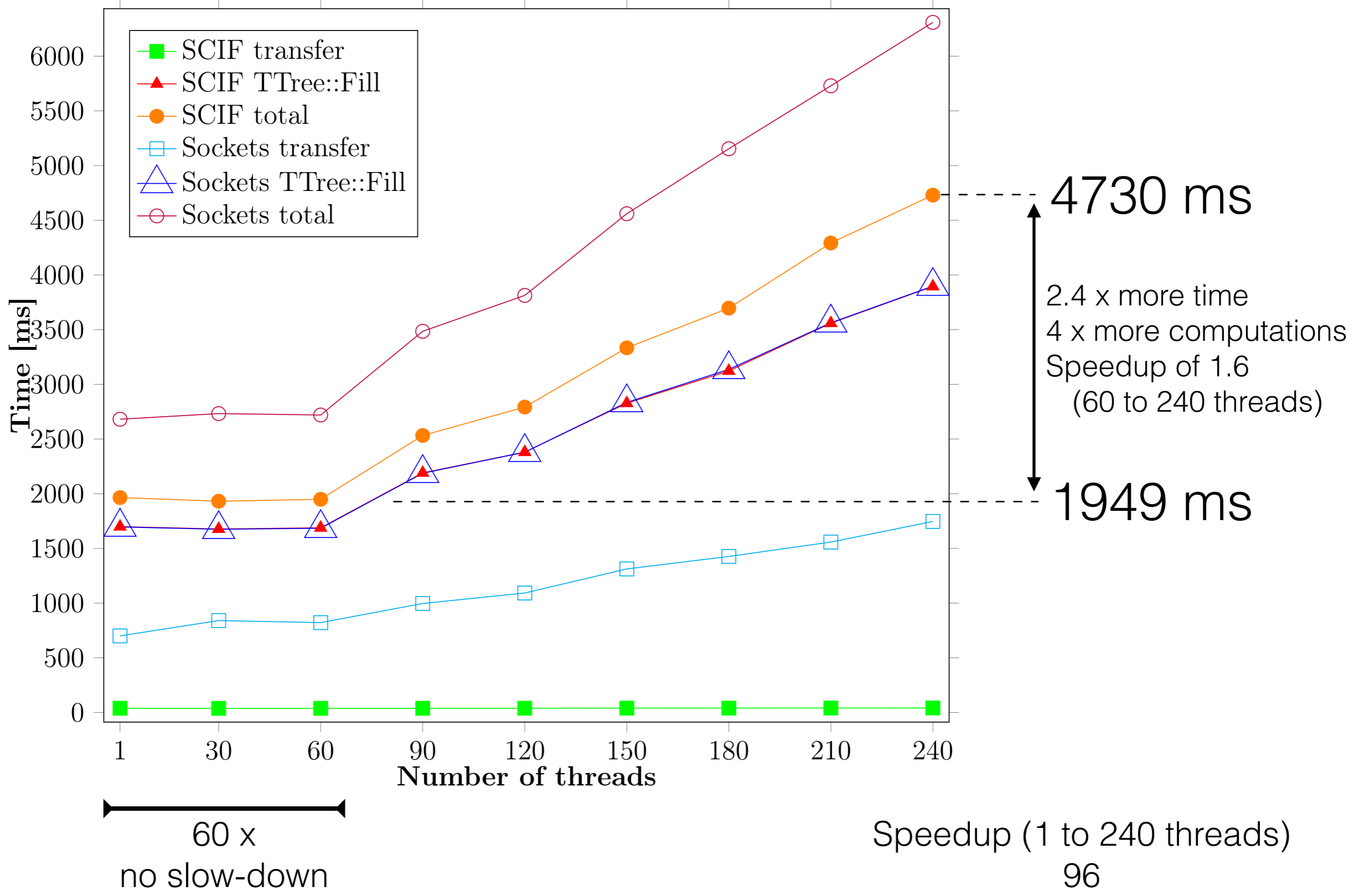
- Optimize file writing in ROOT
 - TTree::Fill
 - Concurrency

Questions



Transfer





60 x
no slow-down

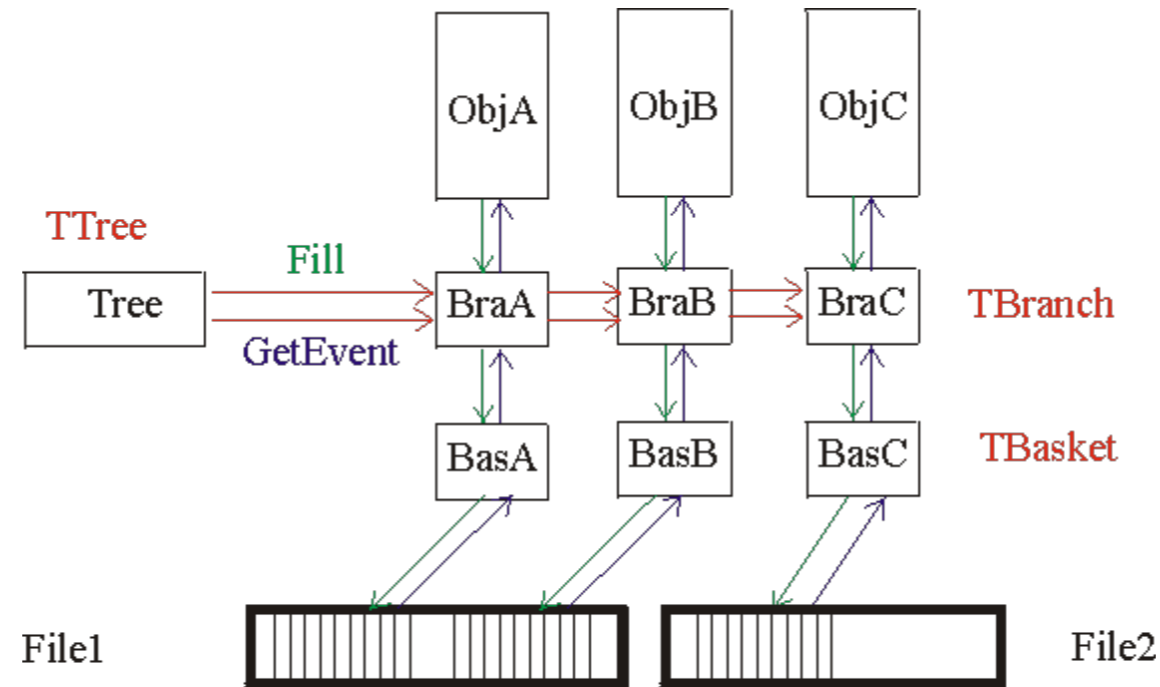
Speedup (1 to 240 threads)
96

4730 ms
2.4 x more time
4 x more computations
Speedup of 1.6
(60 to 240 threads)
1949 ms

Connection

- TSocket
 - 2.23ms
- TSCIF
 - 0.75ms
 - Memory registration (with TSCIF)
 - 2MB: 4.88ms
 - 512kB: 2.88ms
 - 64kB: 2.00ms

TFile writing



Fill writes a copy of the objects to file(s)
GetEvent reads a copy of the objects from file(s)

```
Int_t TScifFile::SysWrite(Int_t /* fd */, const void *buf, Int_t len)
```

memcpy Host vs MIC

- 1000 x memcpy of 2MB
 - Host: 237ms
 - Xeon Phi 873ms 3.68 x slower

Writing file on Xeon Phi

- Speed of tmpfs on Xeon Phi
 - Measure with dd (from BusyBox): 195MB/s
 - Measure with Bonnie: peak to 500MB/s