

Collaboration with Numascale On Current and Future Parallel Computing Platforms

Raphael HIRSCHI

Why Collaborate Between Industry And Academia?

At first sight, difficult BUT rewarding:

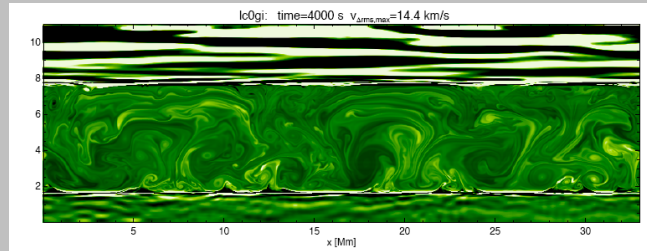
For Industry:

- Boost R&D for free:
exploratory studies
done by PhD/MSc students
- Help with recruitment

For academia:

- Having the edge over
academic competitors
- Improve career prospects/training
for group members (PhD students)
- Personal interest

The ERC encourages in particular proposals that cross disciplinary boundaries. ...



SHYNE:

Stellar HYdrodynamics Nucleosynthesis & Evolution

Key facts:

ERC starting grant

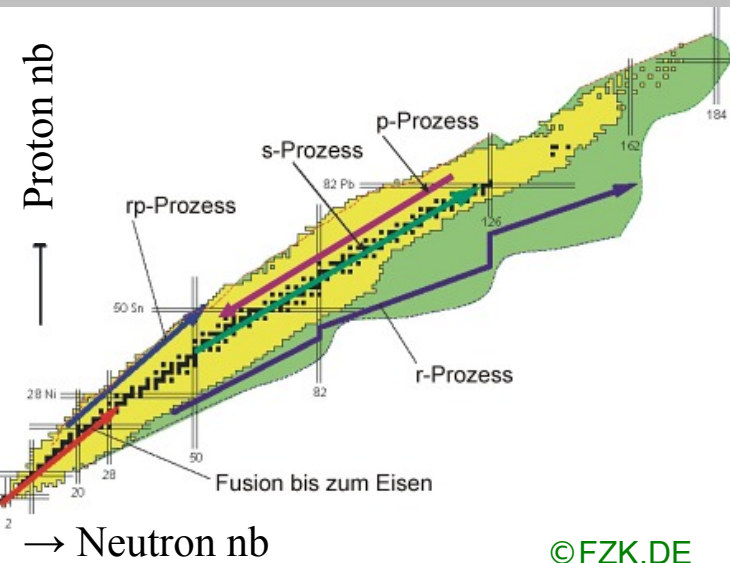
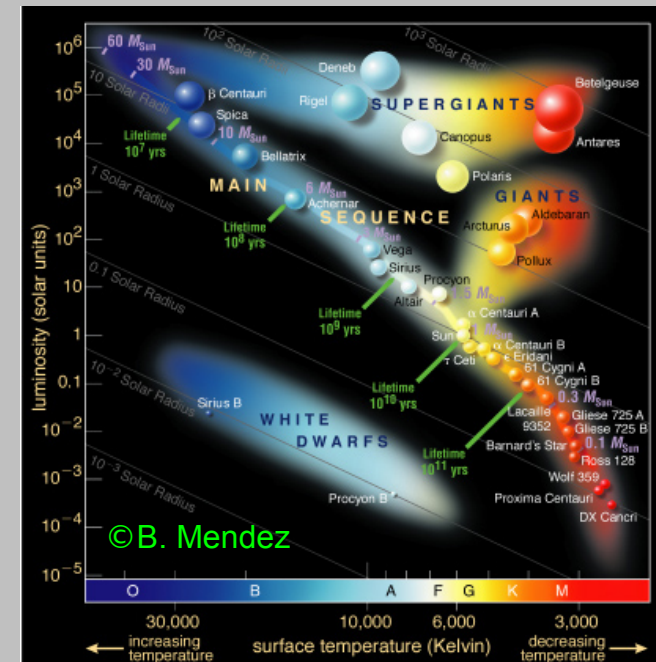
PI: R. Hirschi

Total budget: 1.4M€

5 years: 2012-2017

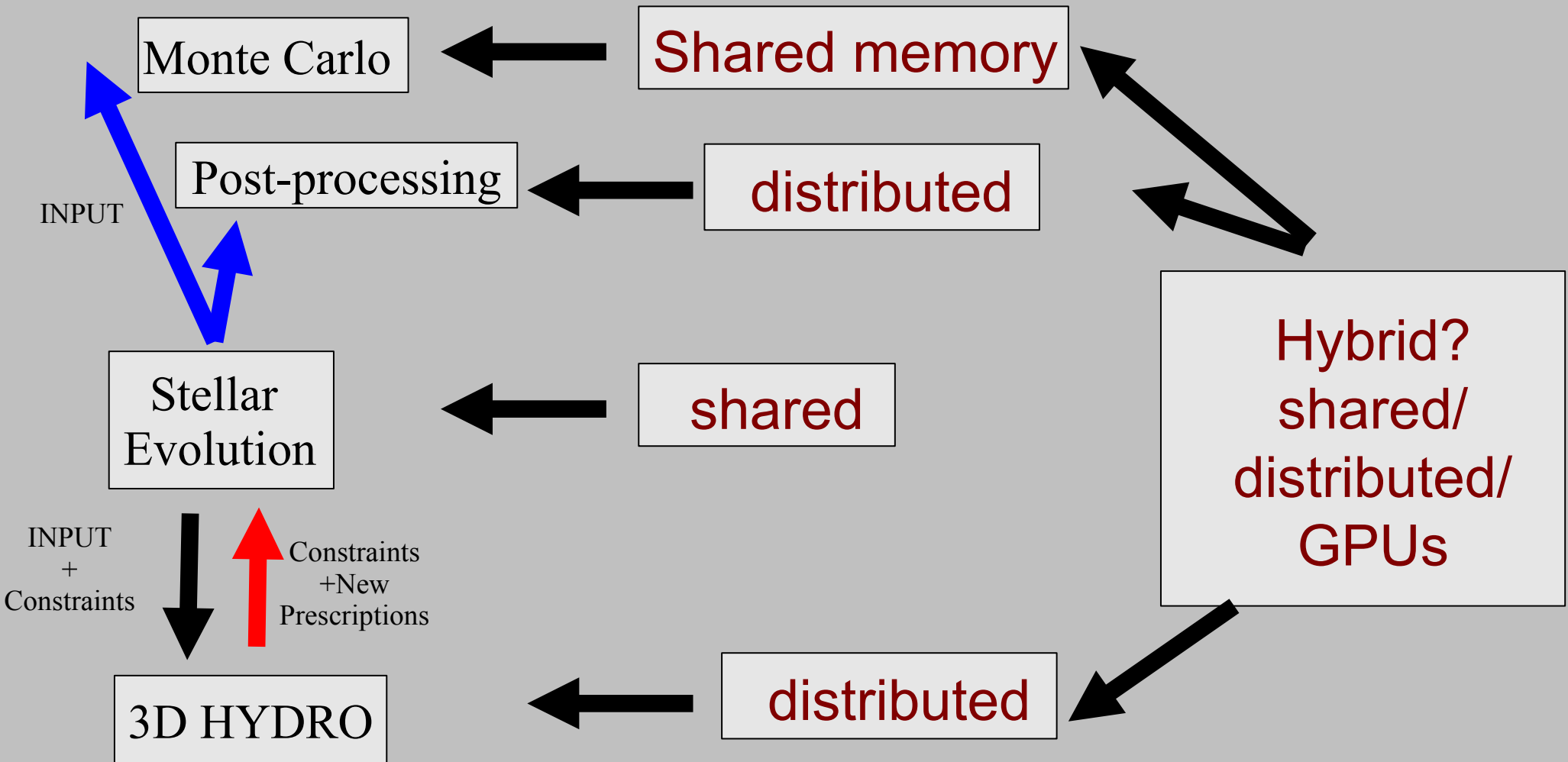
2 Postdocs + 2PhDs

1000+ core cluster



Tool Suite: Parallel Programming Platform

TOOL SUITE: Current Platform → Future Platform



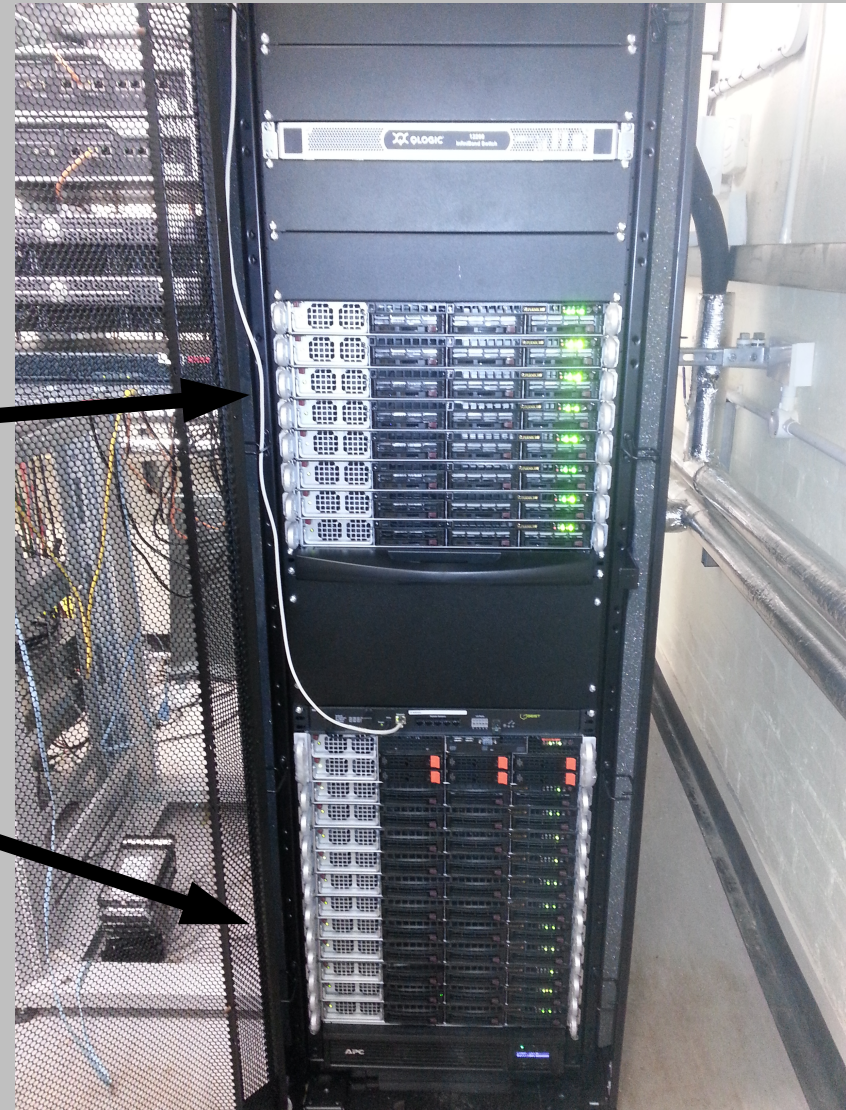
(shared memory: OpenMP / Distributed mem.: MPI)

SHYNE Computer Cluster

- The cluster comprises a total of 1056 AMD-based CPU-cores with 2GB RAM per core.

The main specifications are the following:

- **288 cores Numascale**
 - 8 x 3 socket servers (each socket has 12-core CPUs, AMD opteron 6172, 2.1 MHz)
 - Single memory image 576GB
 - Single operating system image
 - Numascale inter-connect
- **768 cores QDR Infiniband**
 - 12 x 4 socket servers (each socket has 16-core CPUs, AMD opteron 6272, 2.1MHz)
 - 1 O/S per server
 - Distributed memory image 128 GB per server
- Unified cluster management for both architectures
 - IBM Platform HPC
 - 2 LSF queues Numa and IB
 - Numa nodes visible as single machine with 288 cores and 576 GB RAM
- Dedicated water cooled environment up to 30kW in 1 rack

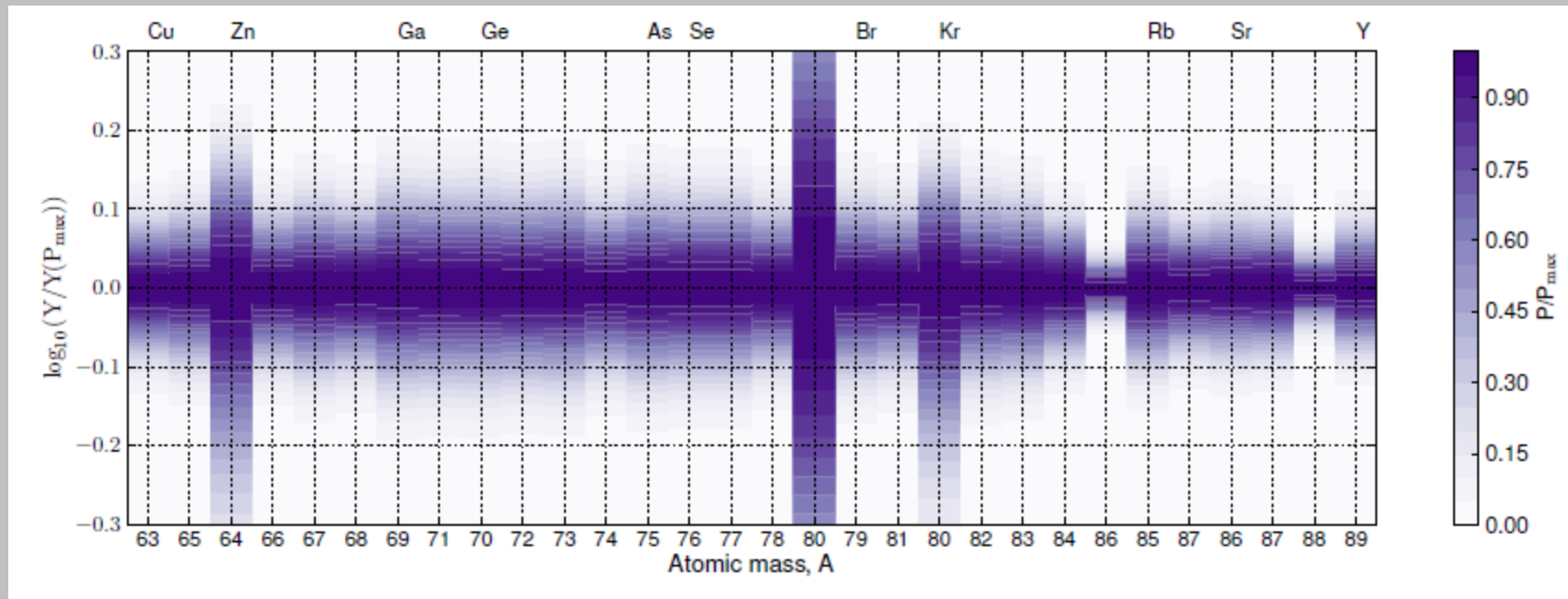


Monte Carlo Sensitivity Studies

PizBuin Monte Carlo wrapper (T. Rauscher)

+ WinNet (Winteler+ 12) < Reaclib

= McWinNet

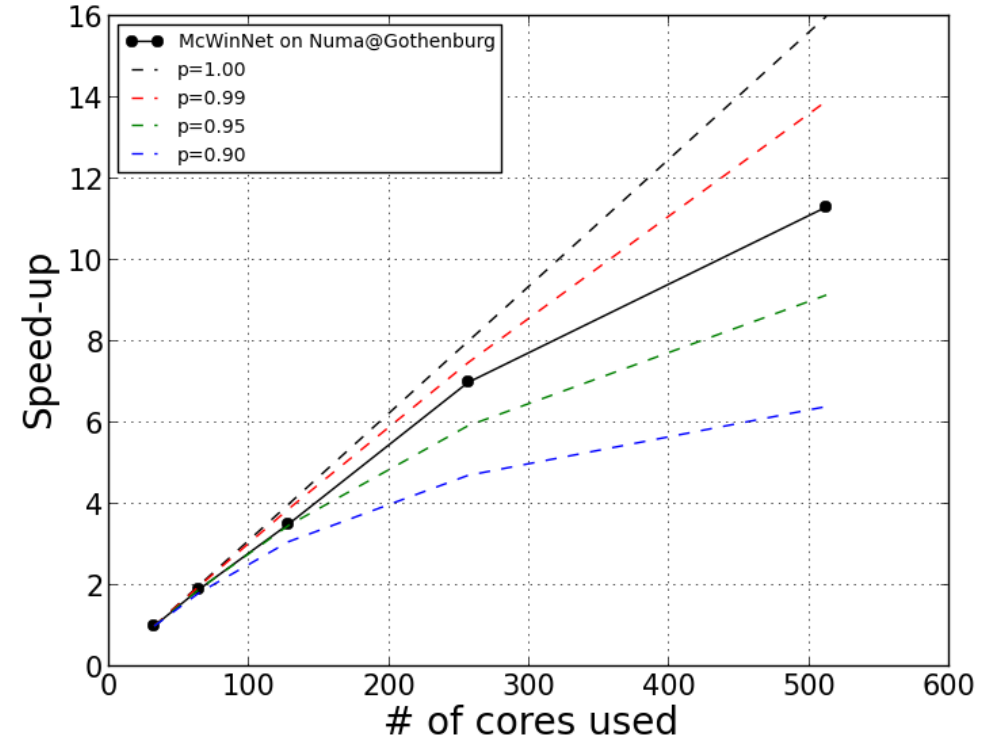
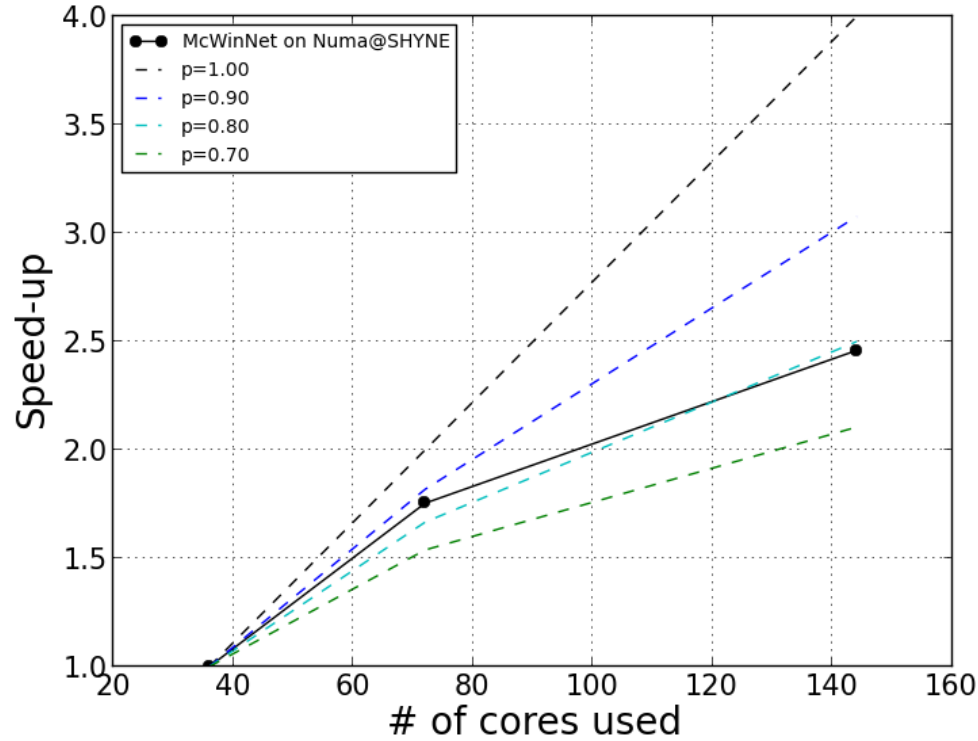


- Largest simulations: 1000 trajectories x 1hr run x 10,000 iterations

McWinNet Code Scaling on Numa Systems

Numa@SHYNE

Numa@Gothenburg



- Excellent scaling up to 512 cores

Conclusions

- Collaboration between industry and academia beneficial to both
- Numascale excellent and flexible platform
- Future platforms will be hybrid
- Numascale will provide large shared-mem. islands into future exascale systems → better balance

Plan

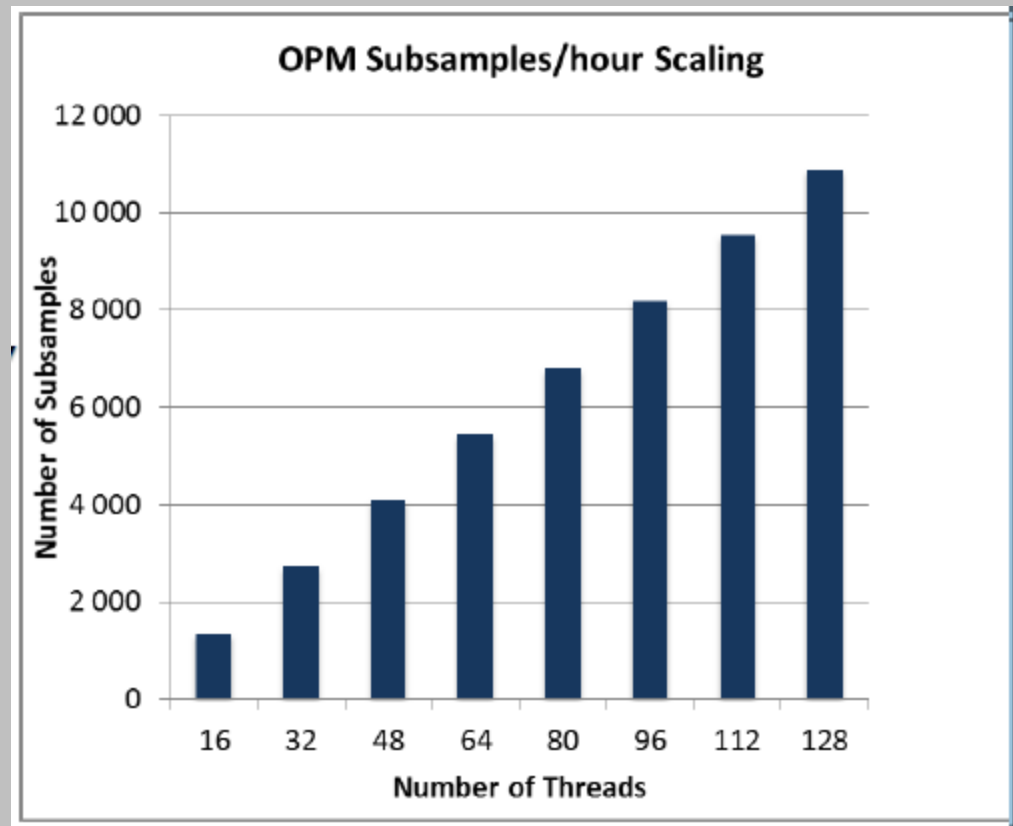
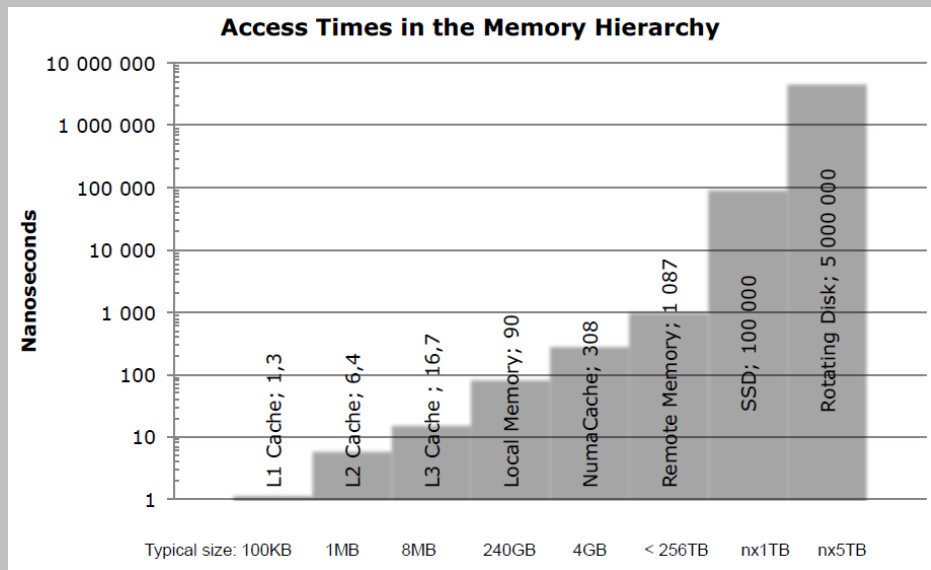
- Why collaborate between industry and academia?
- Collaboration with Numascale company within
ERC SHYNE project

The Memory Hierarchy & Scaling

OpenMP easier to implement than MPI but might be harder to scale

Scaling:

Memory access time:



Excellent scaling possible on Numascale system!

Link with Industry

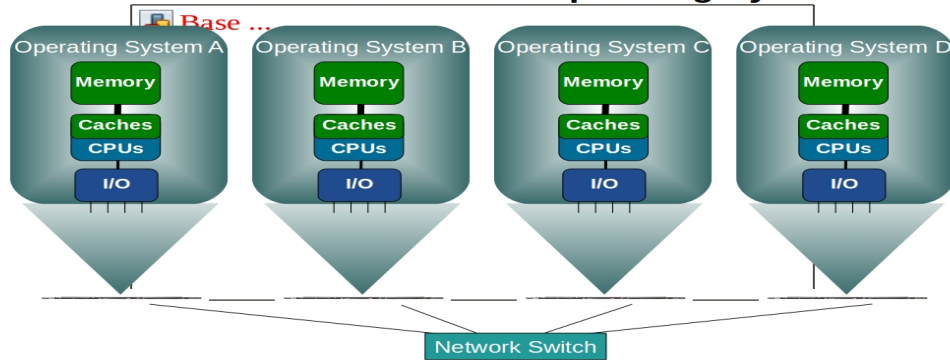
numascale

Norwegian HPC company

Clusters - NO Shared Resources

numascale

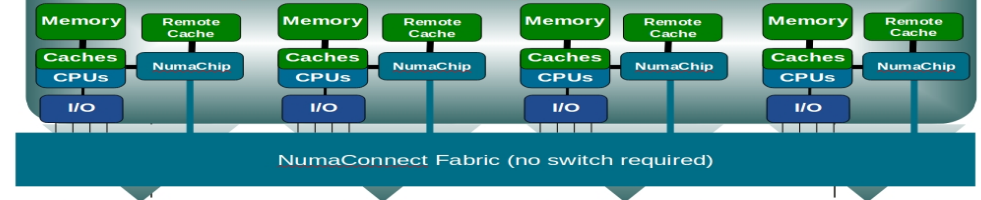
Individual Instances of the Operating System



NumaConnect - Share Everything

numascale

Shared Everything - One Single Operating System Image



Capabilities like Mainframe - Priced like Cluster

FROM: distributed memory clusters

TO: scalable shared memory clusters

For the same cost!

- Super-desktop: single OS, 288 cores, RAM 576 GB
- Large scale: better balance between shared/distributed memory