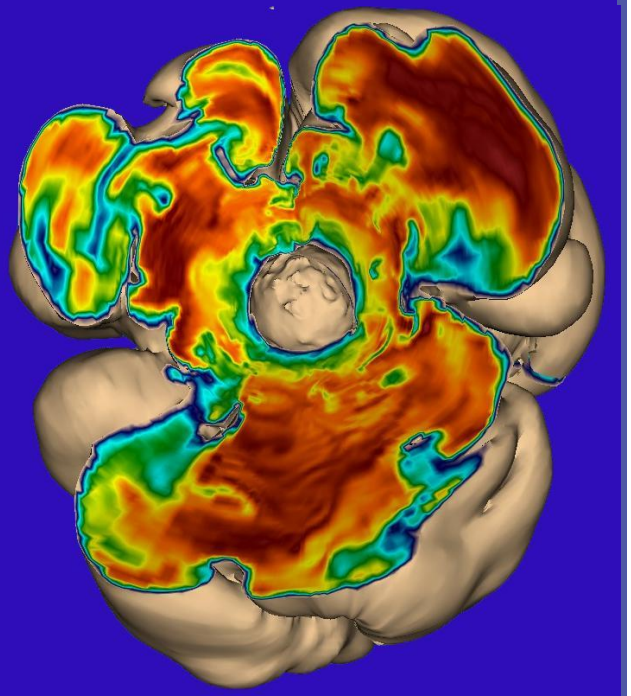


Towards Petaflop simulations of core collapse supernovae

ORAP Forum
October 10, 2013

Andreas Marek[1]

together with Markus Rampp[1] , Florian Hanke[2], and Thomas Janka[2]

[1] Rechenzentrum der Max-Planck-Gesellschaft
[2] Max-Planck Institut für Astrophysik

rzg RECHEN-ZENTRUM GARCHING

Content

- Supernova modelling: the challenges
- The VERTEX code
- Scaling
- Performance
- Conclusions

How to understand them…

Theorists' Tools for Verification

CURIE at CEA/Saclay,
JUROPA at NIC

SuperMUC at LRZ

Supercomputers are **"microscopes" and slow-motion cameras** for looking deep into the cores of exploding stars!

Supernova modeling

Support by Computer Time Grants through:

DEISA

John von Neumann
Institut für Computing

HLRS

Support by Research Grants from:

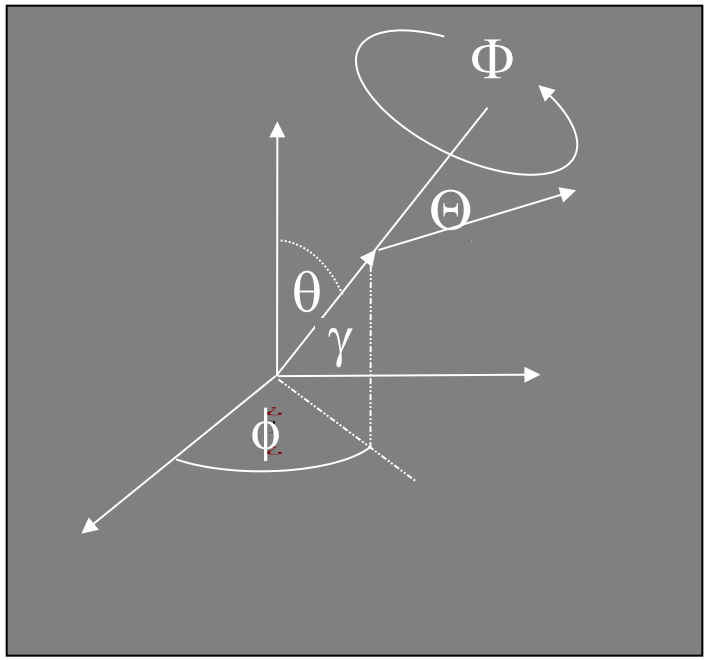
The challenges: The curse of dimensions

- **Boltzmann equation determines the neutrino distribution function**

$$f(r,\theta,\phi,\Theta,\Phi,\varepsilon,t)$$

- **Integration in momentum space: source terms for hydrodynamics**

$$Q(r,\theta,\phi,t),\ dY_e(r,\theta,\phi,t)/dt$$



Solution approach

- **3D** hydro + **6D** direct discretization of Boltzmann Eq. (code development by Sumiyoshi & Yamada 2012)
- **3D** hydro + two-moment closure of Boltzmann Eq. (next feasible step to full 3D; cf. Kuroda et al. 2012)
- **2D or 3D** hydro + **"ray-by-ray-plus"** variable Eddington factor method (MPA, RZG)

Required resources

>= 10 - 100 PFlop/s (sustained!)

>= 1 – 10 PFlop/s

>= 0.1 TFlop/s – 1 TFlop/s (2D)

>= 0.1 Pflop/s – 1 PFlop/s (3D)

Computing requirements for current simulations

2D and 3D supernova modeling

Time dependent simulations: t ~ 1 second, ~ 10^6 timesteps

CPU-time requirements for one model run

- **In 2D with 600 radial zones, 1 degree lateral resolution**
  ~ 3 *10^{18} Flops, about 10^6 processor-core hours

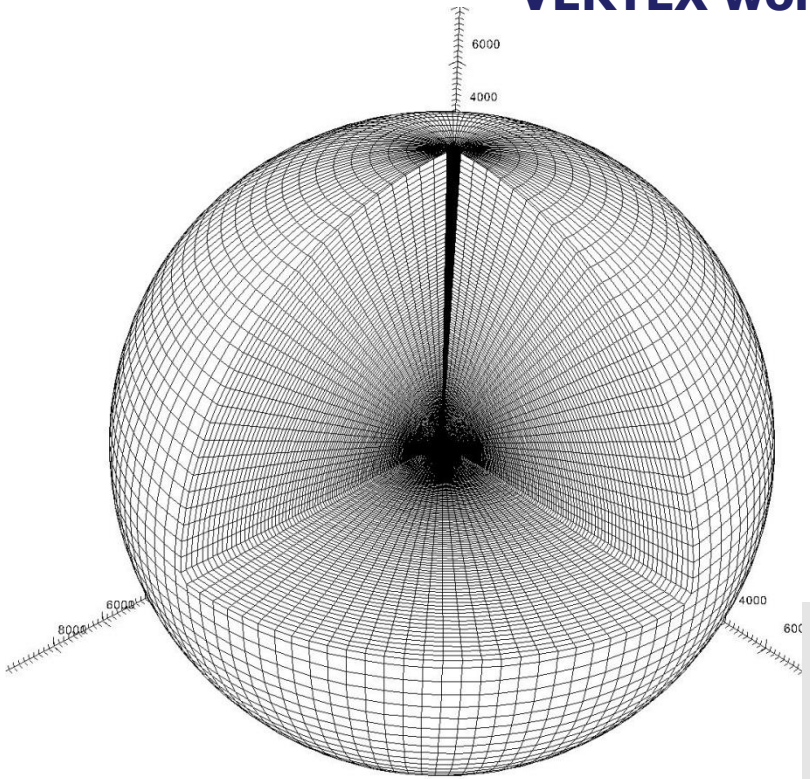
- **In 3D with 600 radial zones, 1.5 degree angular resolution**
  ~ 3 *10^{20} Flops, about 10^8 processor-core hours

and in the near future with increased resultion this will be at least 3 times more

The VERTEX code : Setup

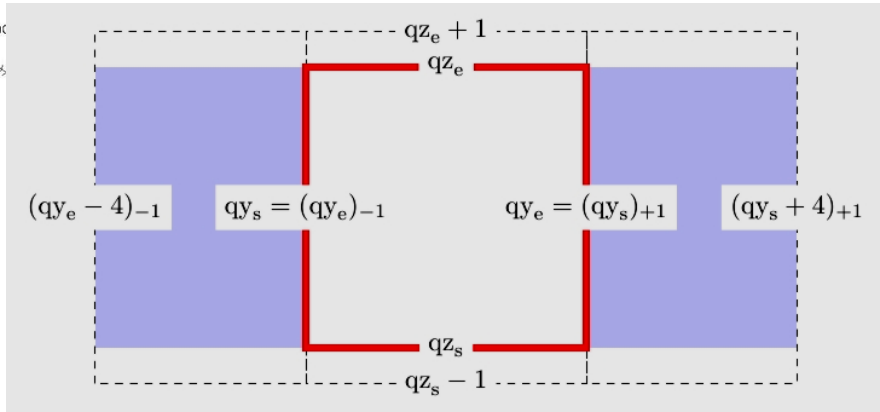
Hybrid MPI/OpenMP parallelized version of VERTEX

VERTEX works on a spherical grid

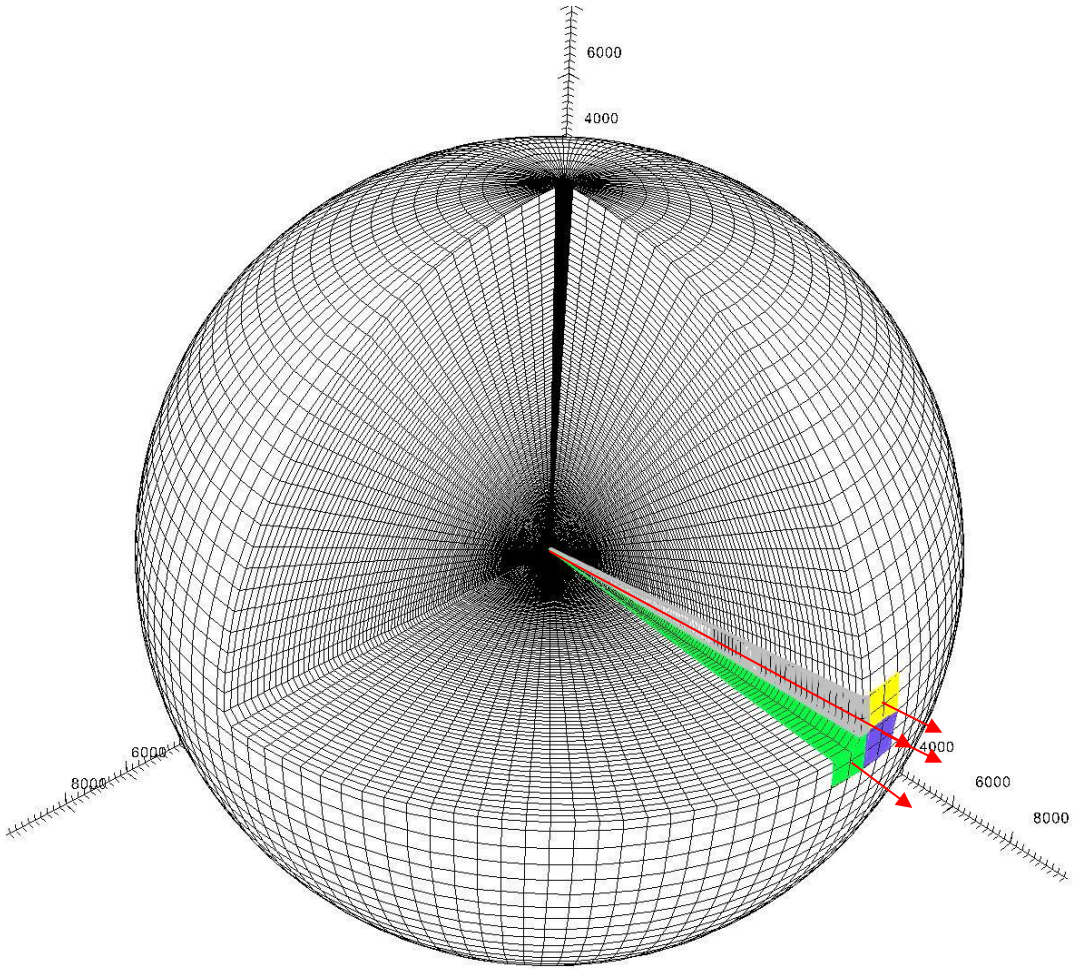


- **Hydro module**
  fully MPI parallelized
- **Transport module**
  MPI + OpenMP coarse grain parallelization along ``angular rays´´

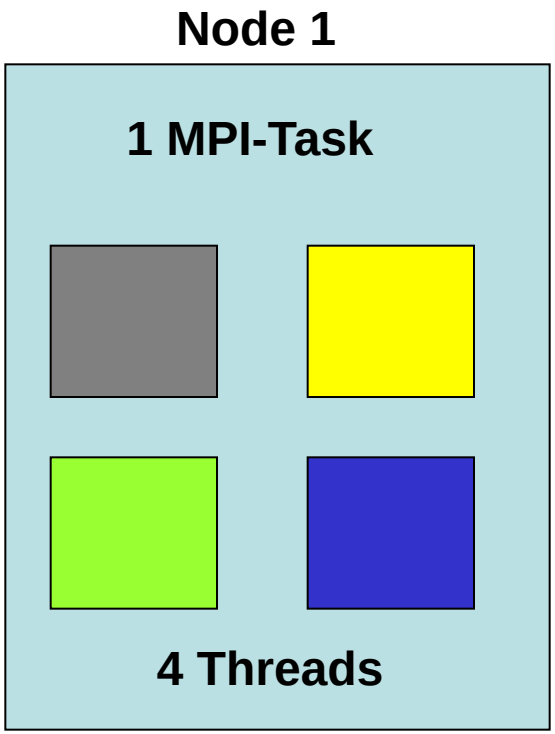
MPI couples rays



Mapping onto the processor grid

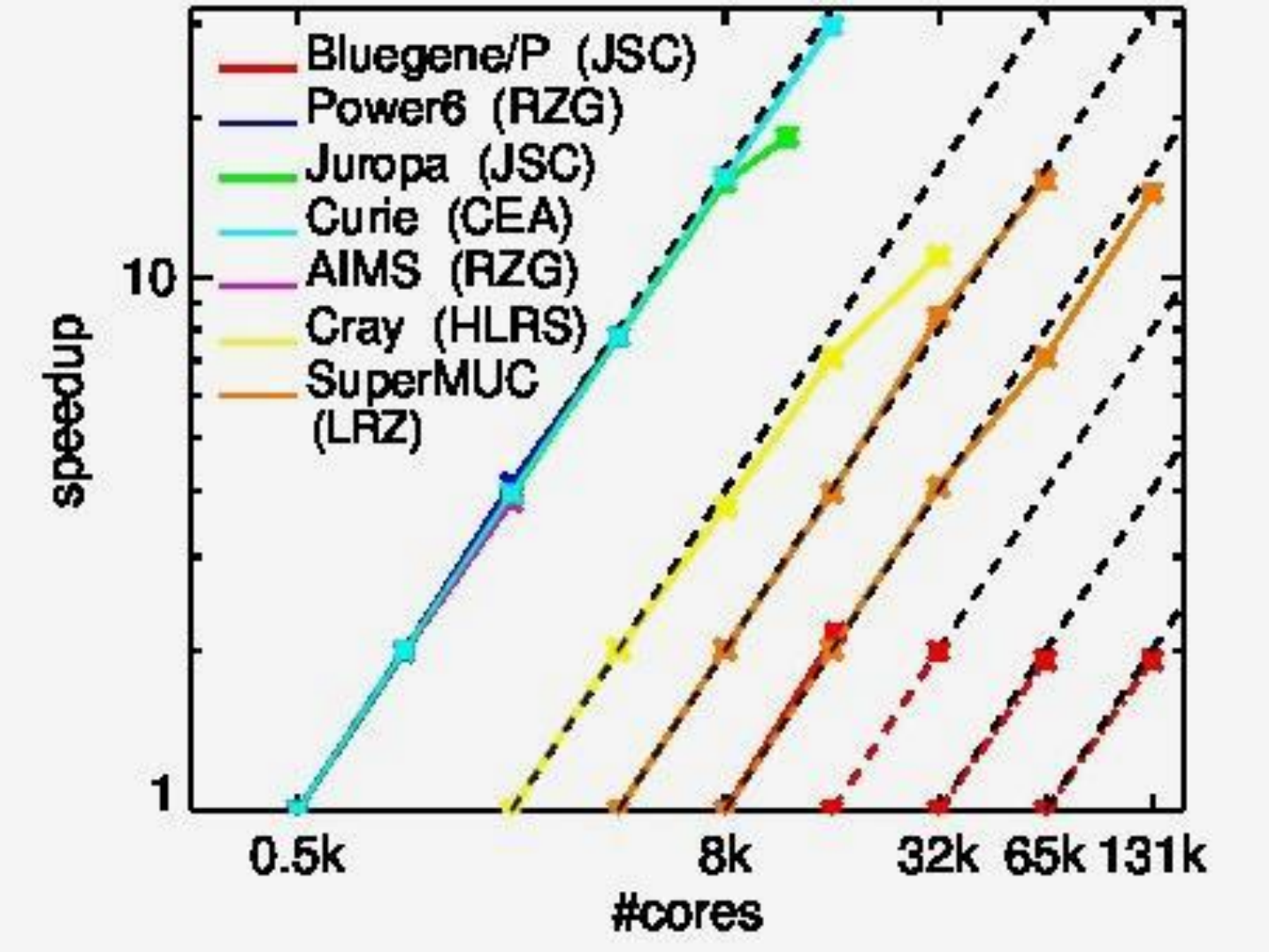


Example: BG/P

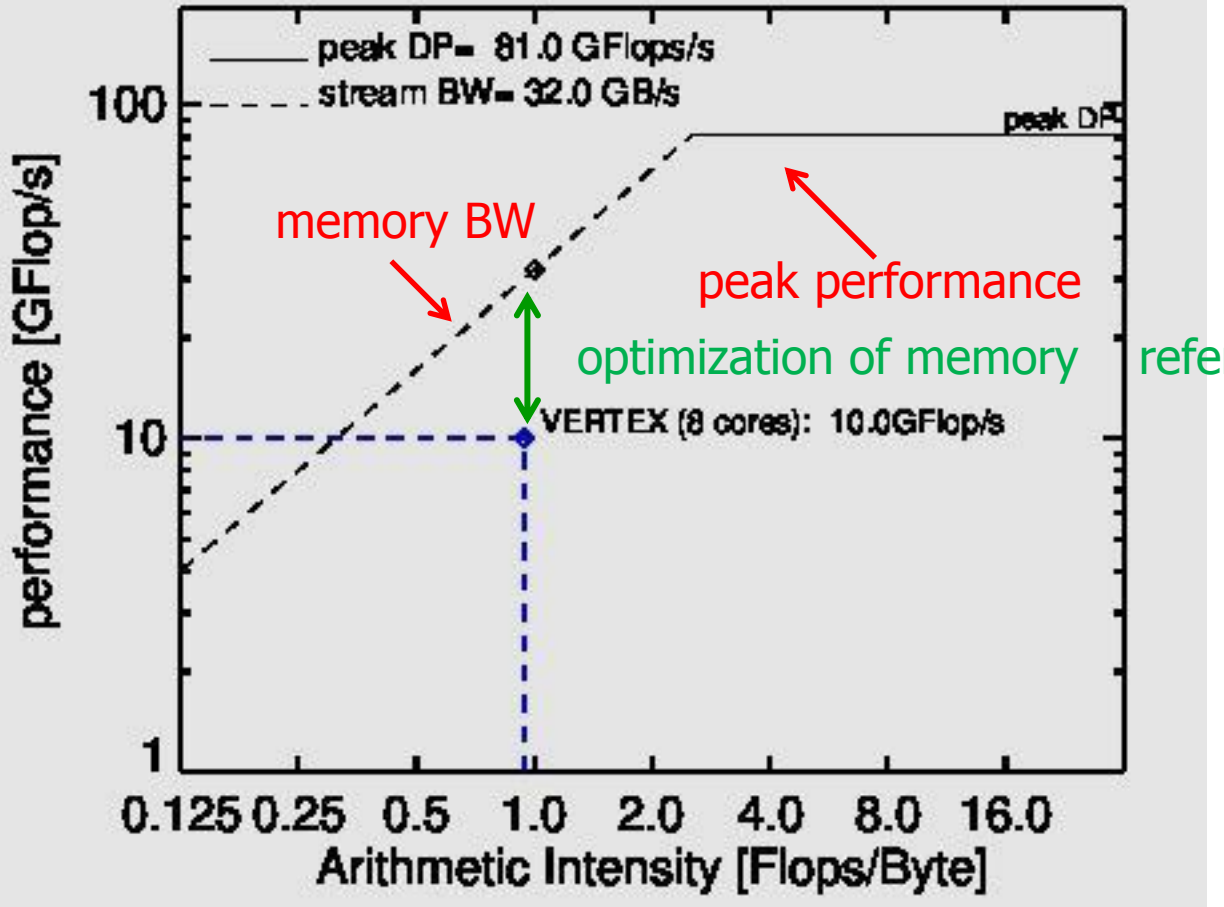


in the node: 1 angular ray per Core

Scaling


Bluegene/P (JSC)
Power6 (RZG)
Juropa (JSC)
Curie (CEA)
AIMS (RZG)
Cray (HLRS)
SuperMUC (LRZ)
speedup
10
1
0.5k
8k
32k
65k
131k
#cores

Performance Estimate: Roofline model



Measurement: Intel Nehalem node, 8 cores @ 2.53 GHz, SSE4.2
done with the ``likwid´´,``perflib´´, and ``PAPI´´
=> VERTEX is ``memory bound´´ runs ~ 12.5% of peak performance

Performance: SandyBridge cores

Performance on 131k cores of SuperMuc:

- **8192 nodes**

  => ~ 0.245 PFlop/s on LRZ`s SuperMUC

  ~ 9 % of peak performance

Scaling: development

Ken Batcher:

“ A supercomputer is a device for turning compute-bound problems into data-transfer-bound problems”

Challenges of for the future

Ken Batcher:
" A supercomputer is a device for turning compute-bound problems into data-transfer-bound problems"

- **in the future data movement will be the problem for good scaling**
  - **problem of MPI at large scales ?**
  - **memory per cores will still decrease**
  - **data transfer within a node (scaling within a "traditional node")**
  - **OpenMP scaling (e.g. Xeon PHI)**
  - **data transfer to accelerator cards (Xeon PHI, GPU)**

=> think of (and rewrite) memory layout that it scales on many core nodes as well as 10000 -100000 s of nodes

- **Other issue: robustness of machines**

Further Code development

Even stronger scaling capability needed:

- **The next or next-next supernova models will need even more computing power!**

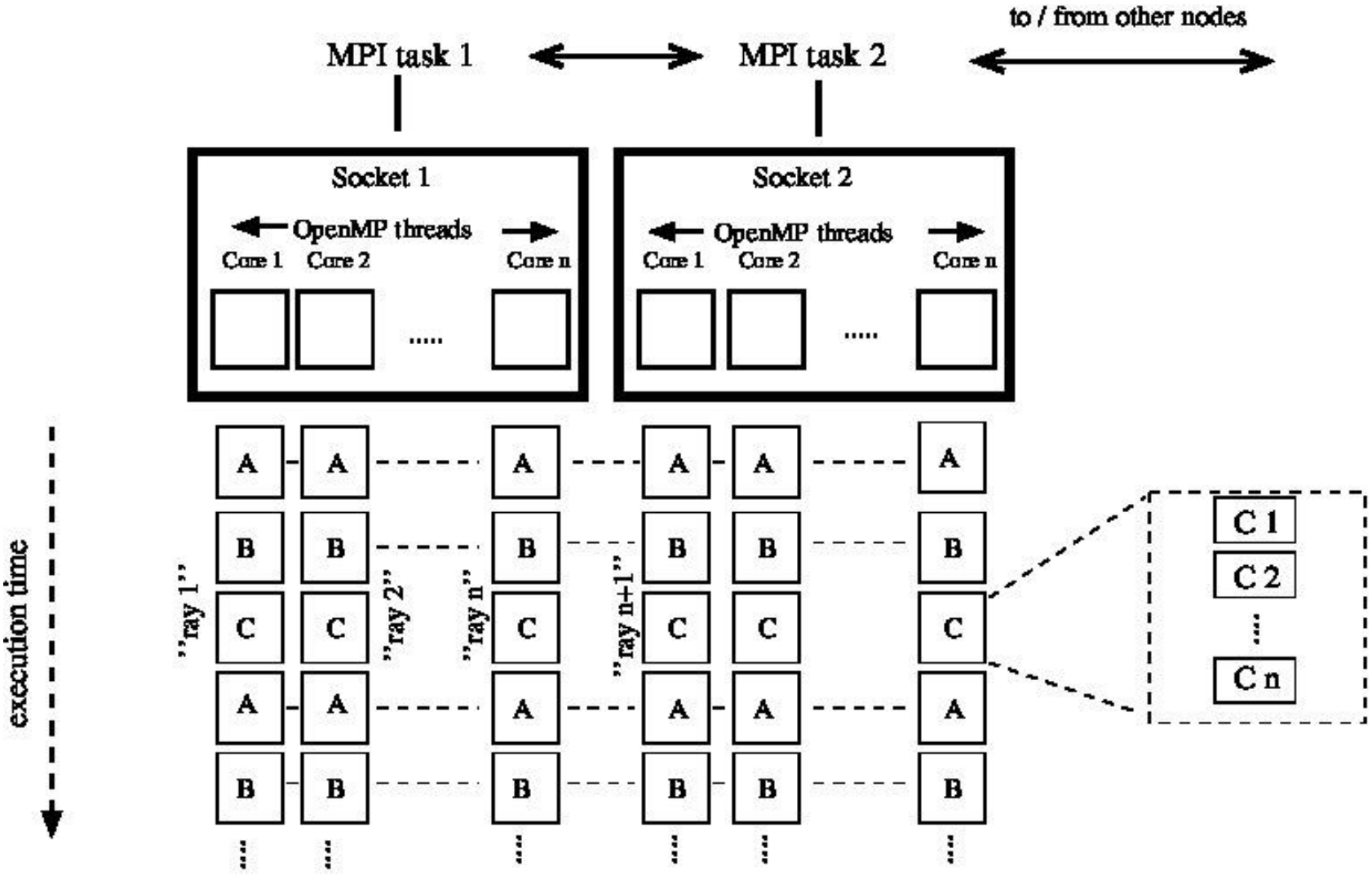
``Traditional´´ development: usage of more cores

- **at the moment another third parallelization level is implemented**
  => within one ``ray´´ a nested OpenMP workload is possible
  => use 10 to 20 times more cores

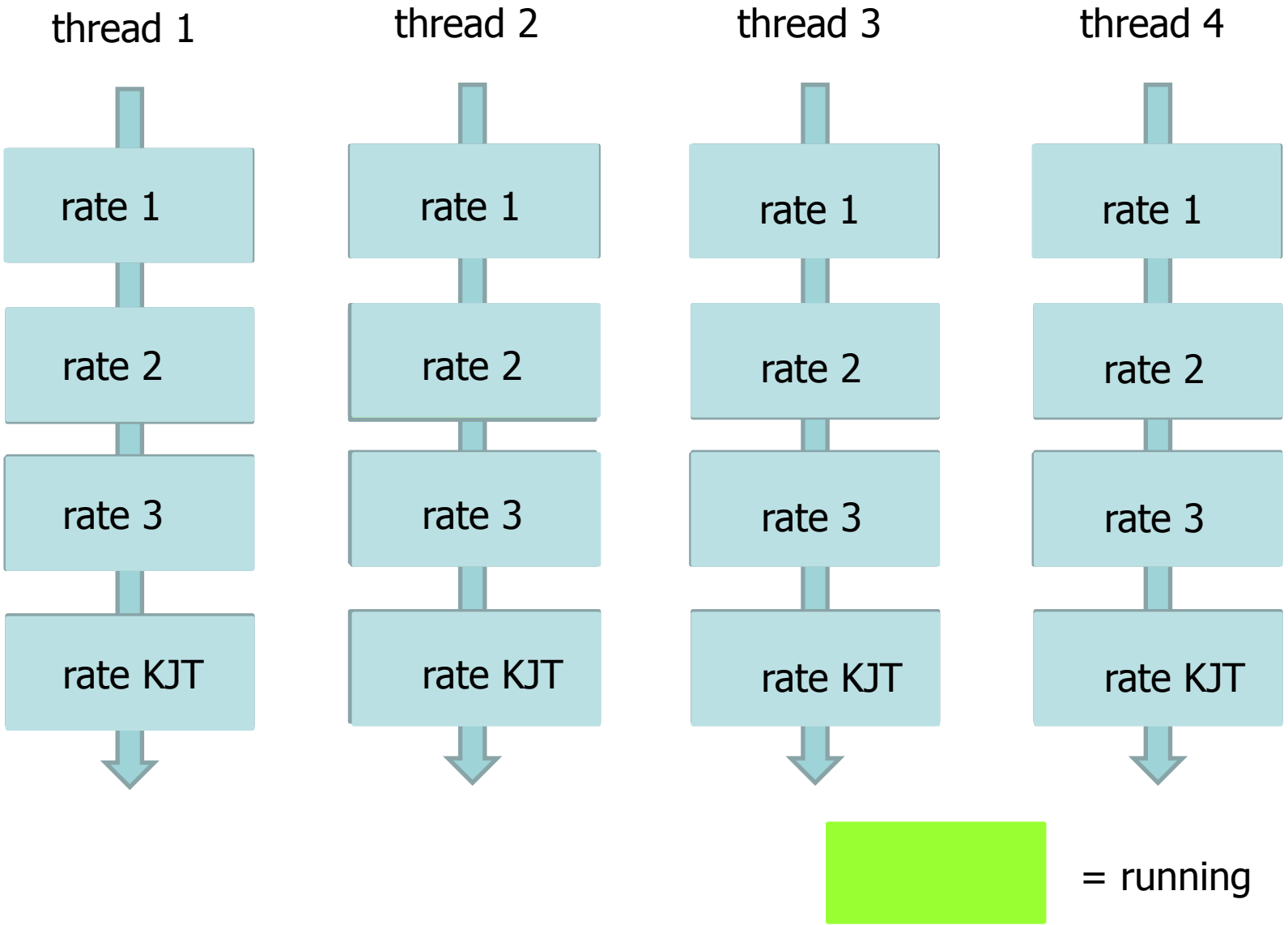
Accelerator cards development:

- **the usage of GPUs and Xeon Phi is tested**
  => GPU implementation already shows promising results
  => factor 2 achieved, 3 seems possible

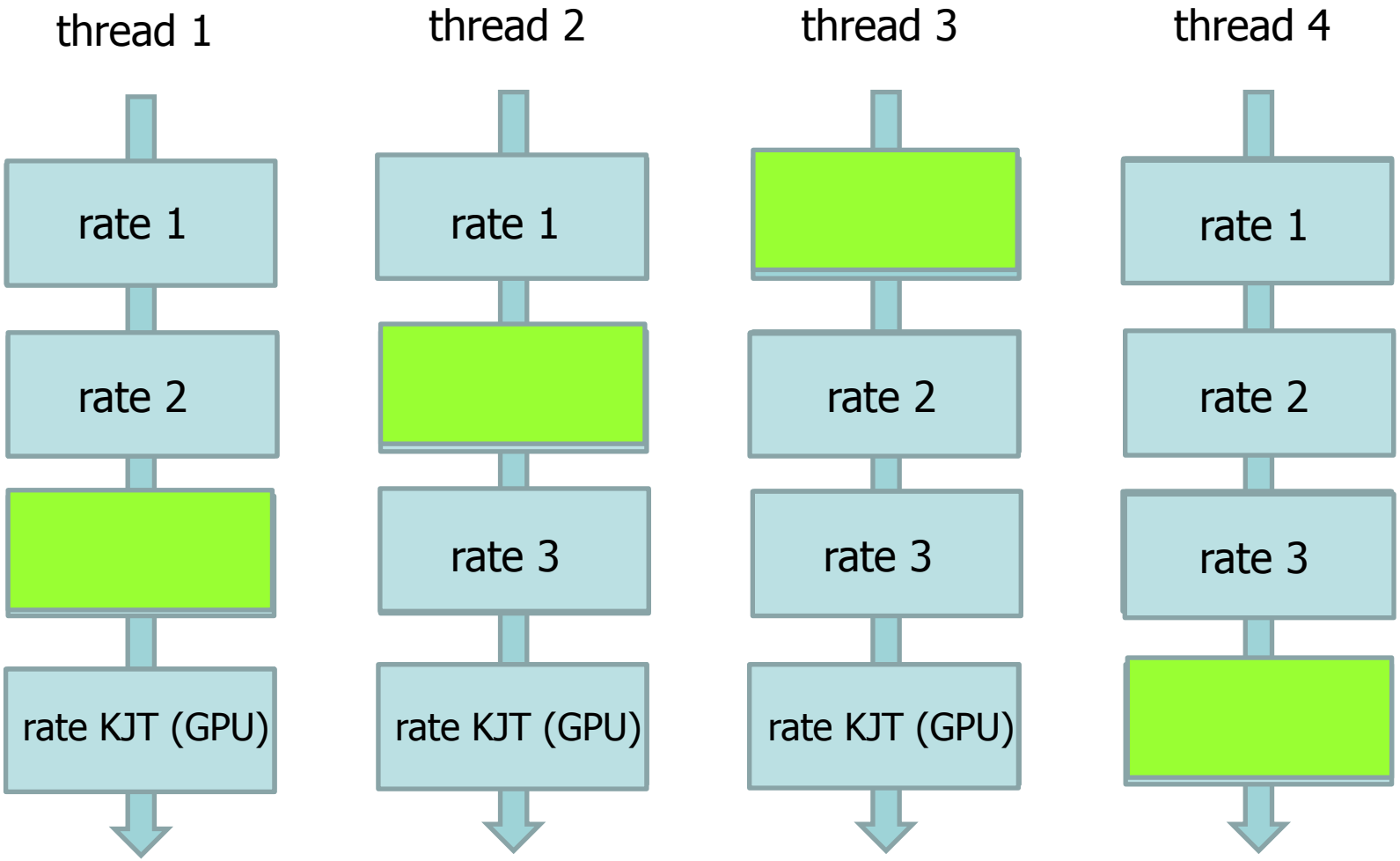
Current Code improvements



Load-Balancing (Host only)



Load-Balancing (HOST + GPU)



Summary

- We have demonstrated the need of supernova research for high end HPC computing
- at the moment production runs are done on 16k -32k cores, the next generation of models will need about 65k – 100k cores
- However, we have demonstrated that VERTEX scales already now on 131k cores of LRZ`s SuperMuc (and on Bluegene/P)
- VERTEX operates with ~ 10% of the peak performance on Intel Sandybridge nodes
- The next generation of VERTEX is already under development
  - nested OpenMP level: about 1 to 2 Mcores might be used
  - usage of accelerator cards:
    * GPU acceleration: speedup of 2 achieved 3 possible
    * Xeon PHI: still under investigation

Questions ?