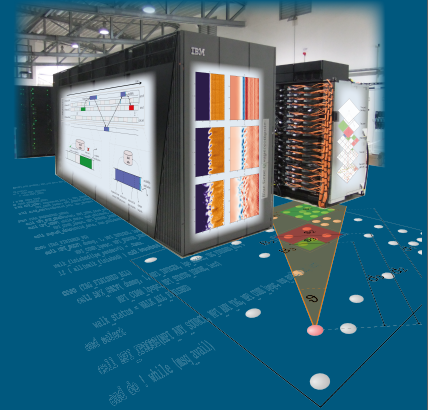
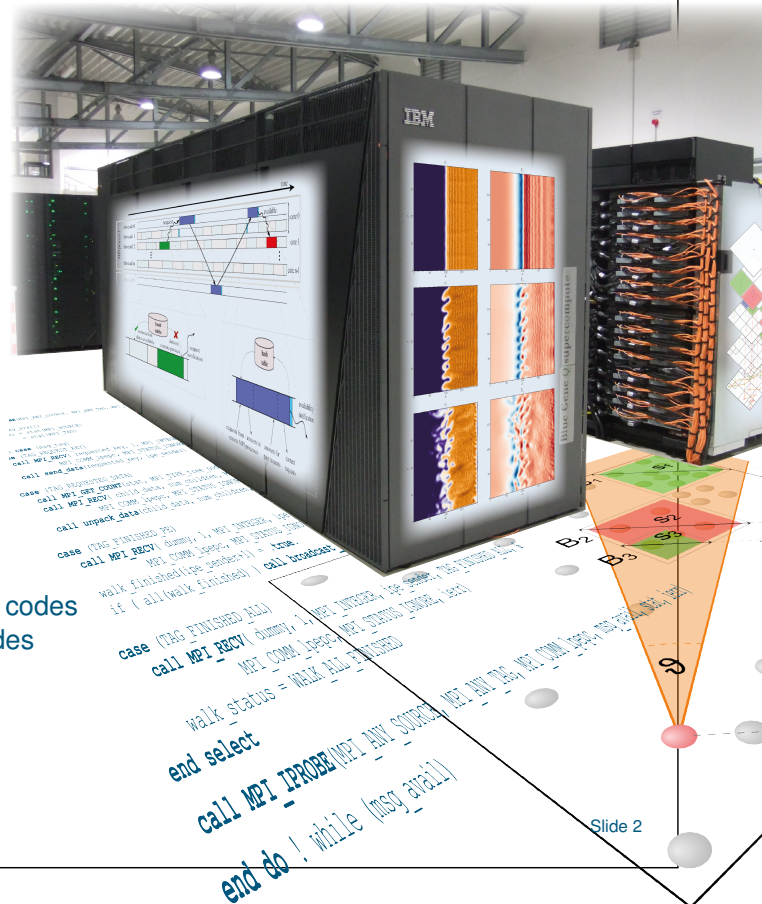


The Barnes-Hut Tree Algorithm

and its highly scalable parallel implementation PEPC



10.09.2013 | Mathias Winkel | Jülich Supercomputing Centre



- 1 – Introduction
- 2 – The Barnes-Hut tree code
- 3 – Periodic boundary conditions for tree codes
- 4 – Parallelization of Barnes-Hut tree codes
- 5 – Applications
- 6 – Outroduction ☺

Electrostatics recap – Particle representation

Potential due to general charge distribution $\rho(r)$ is given by Poisson's equation:

$$\nabla^2 \phi(\vec{r}) = -4\pi\rho(\vec{r}),$$

which has a general solution

$$\phi(\vec{r}) = \int_{V'} \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3r'.$$

Discretization of charge density with point particles

$$\rho(\vec{r}, t) = \sum_j q_j \delta(\vec{r}_j - \vec{r})$$

yields pairwise potential sum

$$\phi_i(\vec{r}_i, t) = \sum_{j \neq i} \frac{q_j}{|\vec{r}_i - \vec{r}_j|}.$$

Dynamics governed by Newton's law:

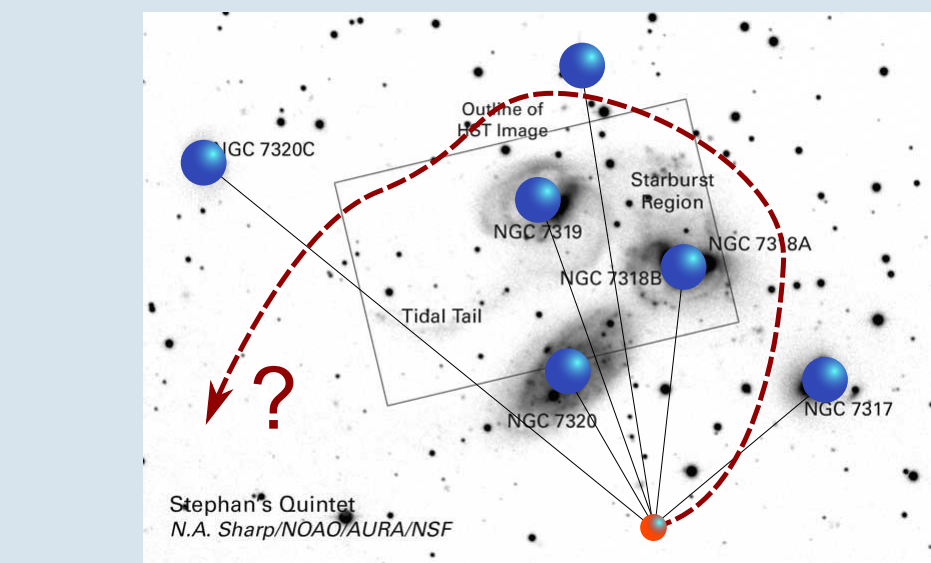
$$\frac{d}{dt} \vec{v}_i = -\frac{q_i}{m_i} \nabla \phi(\vec{r}_i) ; \quad \frac{d}{dt} \vec{r}_i = \vec{v}_i.$$

→ Basis for MD with charged particles, gravitational dynamics, ...

Multipole algorithms

The idea behind...

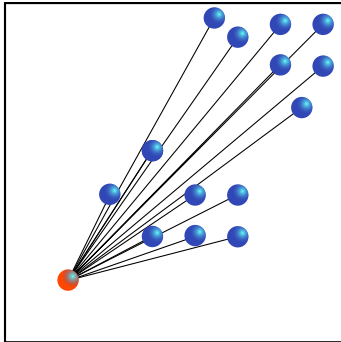
Approximate galaxies by point masses



Multipole algorithms

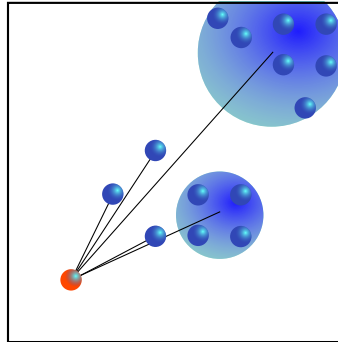
... different flavors

Direct Summation



total interactions: $\mathcal{O}(N^2)$

Treecode



$\mathcal{O}(N \log N)$



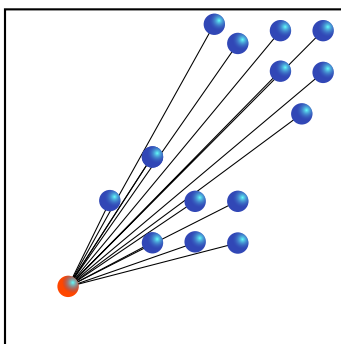
- Clustering of remote particle groups
- Interaction via their multipole expansion
- Hierarchical algorithm for finding interaction partners

[J. Barnes & P. Hut, Nature **324**, 446 (1986)]

Multipole algorithms

... different flavors

Direct Summation

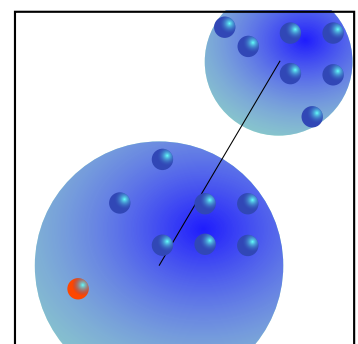


total interactions: $\mathcal{O}(N^2)$



- Clustering of remote and local particle groups
- Interaction via their multipole expansions and elaborate shifting/conversion operators
- Hierarchical algorithm for finding interaction partners

Fast Multipole Method



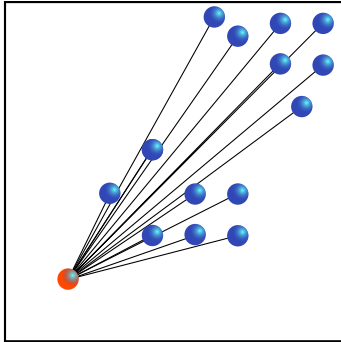
$\mathcal{O}(N)$

[L. Greengard & V. Rokhlin, J. Copmp. Phys. **73**, 325 (1987)]

Multipole algorithms

... different flavors

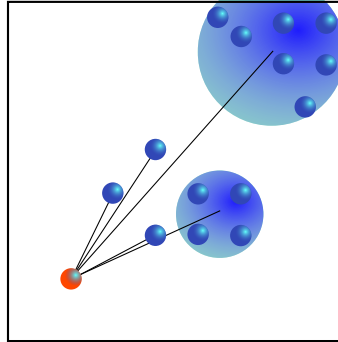
Direct Summation



total interactions: $\mathcal{O}(N^2)$

- good load balancing
- kernel flexibility
- simple parallelization

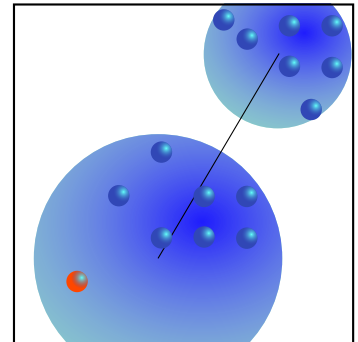
Treecode



$\mathcal{O}(N \log N)$



Fast Multipole Method



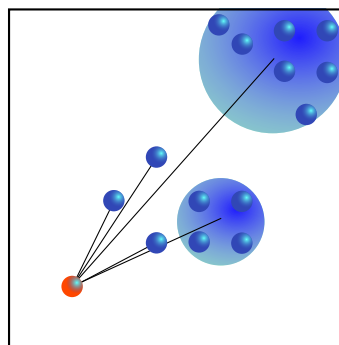
$\mathcal{O}(N)$

- less floating point ops
- more logistics
- elaborate parallelization

Multipole algorithms

... different flavors

Treecode



$\mathcal{O}(N \log N)$

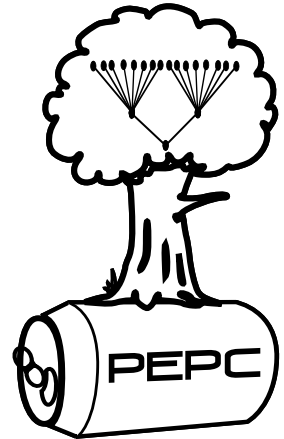
- straightforward load-balancing for dynamic simulations
- extensibility
- exchange of interaction kernel

$$\phi(\vec{r}, t) = \sum_{i=1}^N q_i \varphi_i(\vec{r} - \vec{r}_i(t))$$

- Coulomb/Gravitation:
 $\varphi(\vec{r}) \propto \frac{1}{|\vec{r}|}$
- reg. Coulomb:
 $\varphi(\vec{r}) \propto \frac{1}{\sqrt{r^2 + \alpha^2}}$
- vector-valued potentials (vortex method)
- ...

PEPC – The Pretty Efficient Parallel Coulomb Solver

- developed at JSC since 2003
- several modifications to original algorithm and parallelization
- world records in scalability and particle number for Barnes-Hut tree codes
- part of the ScaFaCoS library
(with some restrictions due to technical reasons)



[M. Winkel *et al.*, *Comp. Phys. Comm.* 187, 880–889 (2012)]

PEPC is open-source, freely available via www.fz-juelich.de/ias/jsc/pepc
mailing list for reaching all developers: pepc@fz-juelich.de

The Barnes-Hut Tree Algorithm

Part II: The Barnes-Hut tree code

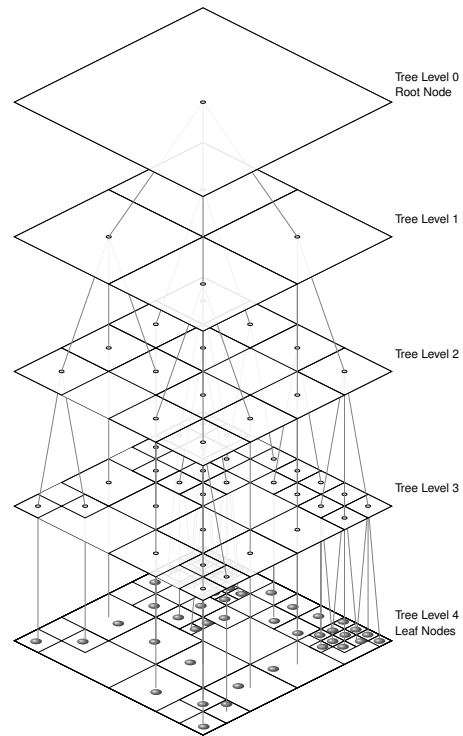
10.09.2013 | Mathias Winkel

The Barnes-Hut tree code

Tree construction

- Recursive spatial subdivision. . .
- . . . until every particle lies in its own box
 - tree leaves
- Lower-level boxes
 - tree nodes

- Hierarchy of spatially coarsened particle groups
- Potential interaction partners



Multipole shifting rules

- center-of-charge

$$\hat{\vec{r}}_{\text{coc}} = \frac{\sum_{i=1}^{N_c} |Q_i| \cdot \vec{r}_{\text{coc},i}}{\sum_{i=1}^{N_c} |Q_i|}$$

- shift vector

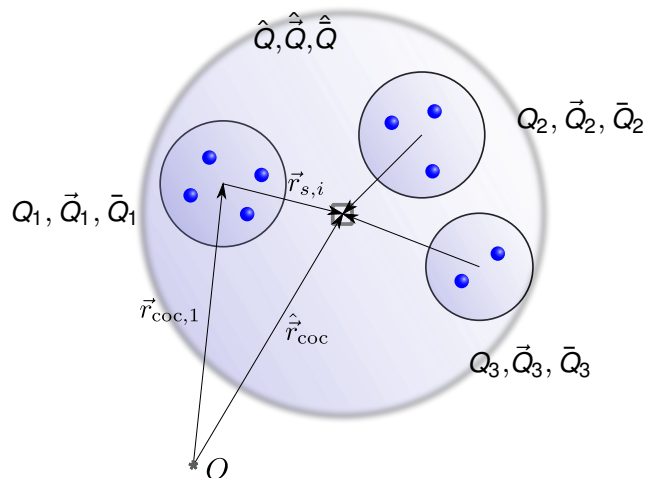
$$\vec{r}_{s,i} = \hat{\vec{r}}_{\text{coc}} - \vec{r}_{\text{coc},i}$$

- shifting rules for multipole moments

$$\hat{Q} = \sum_{i=1}^{N_c} Q_i$$

$$\hat{\vec{Q}} = \sum_{i=1}^{N_c} (\vec{Q}_i - Q_i \cdot \vec{r}_{s,i})$$

$$\hat{\vec{Q}} = \sum_{i=1}^{N_c} \left(\vec{Q}_i - \vec{r}_{s,i} \otimes \vec{Q}_i - (\vec{r}_{s,i} \otimes \vec{Q}_i)^T + \vec{r}_{s,i} \otimes \vec{r}_{s,i} \cdot Q_i \right)$$

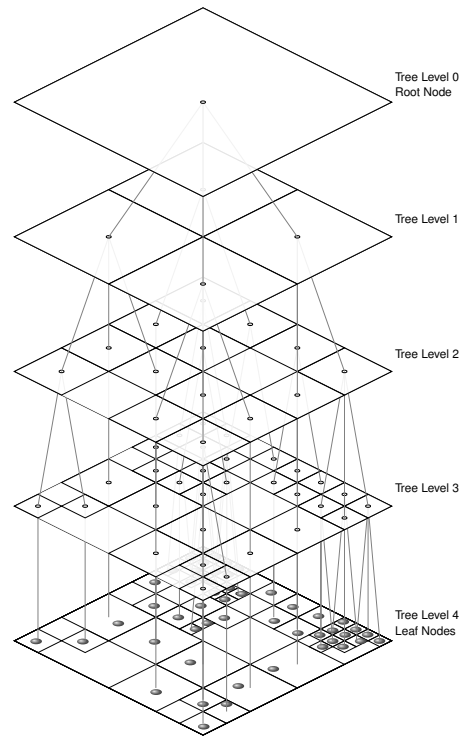


The Barnes-Hut tree code

Tree construction

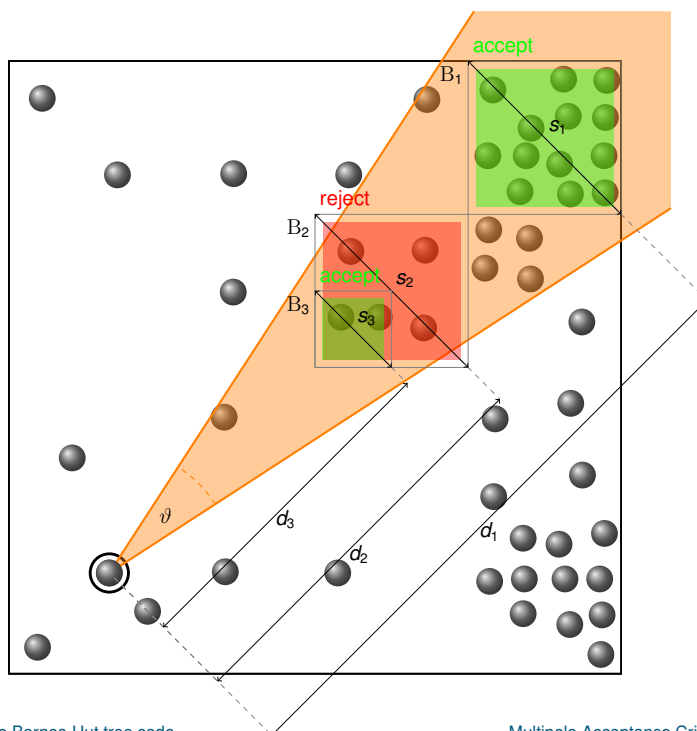
- Recursive spatial subdivision. . .
- . . . until every particle lies in its own box
 - tree leaves
- Lower-level boxes
 - tree nodes

How to systematically identify interaction partners ??



The Barnes-Hut tree code

Multipole Acceptance Criterion (MAC)



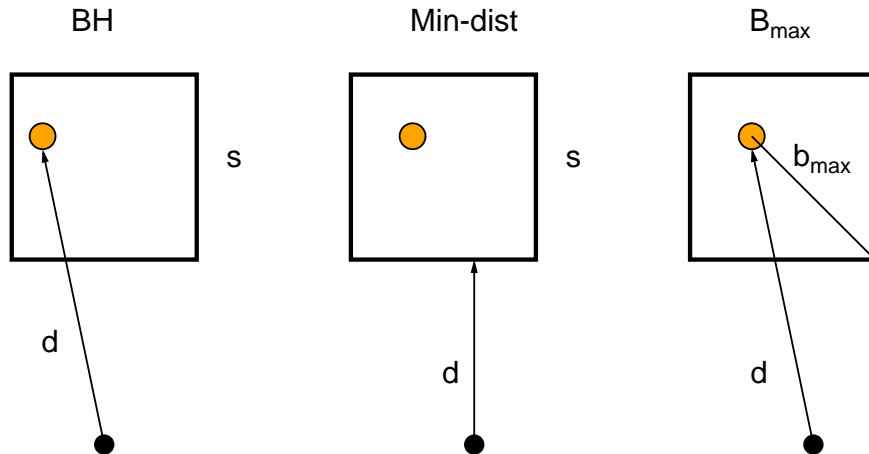
- cluster substructure ignored if

$$\frac{\text{clustersize}}{\text{clusterdistance}} = \frac{s}{d} \leq \vartheta$$

$$\mathcal{O}(N \log N) : 0.0 < \vartheta$$

$$\mathcal{O}(N^2) : 0.0 = \vartheta$$

MAC variations



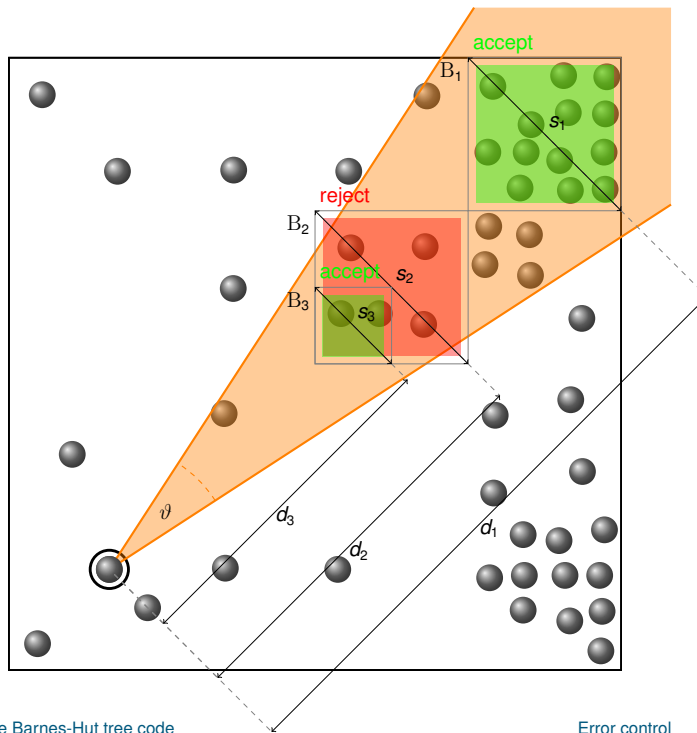
Tree traversal - identification of interaction partners

```

for all particle in particle_list do
  todo_stack.clear()
  node ← root traversal for the particle starts at root node
  repeat
    if (MAC_OK(particle, node)) and not (particle ∈ node) then MAC evaluation
      interaction allowed
      call INTERACT(particle, node)
      due to interacting with this node, its children do not have to be considered
    else
      the MAC requires the node to be further resolved
      todo_stack.push(node.children) proceed vertically in tree
    end if
  until IS_INVALID(node ← todo_stack.pop())
  todo_stack is empty for this particle – its traversal is complete
end for
  
```

The Barnes-Hut tree code

Multipole Acceptance Criterion (MAC)



- cluster substructure ignored if

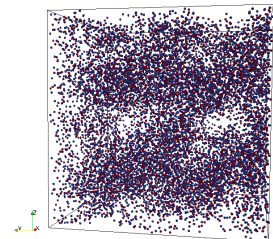
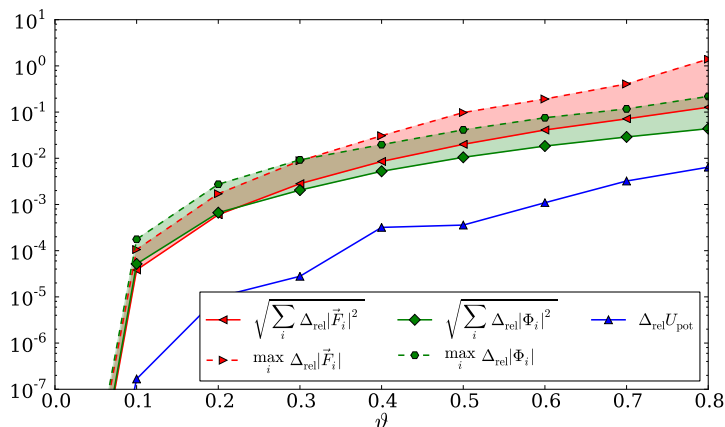
$$\frac{\text{clustersize}}{\text{clusterdistance}} = \frac{s}{d} \leq \vartheta$$

$$\mathcal{O}(N \log N) : 0.0 < \vartheta$$

$$\mathcal{O}(N^2) : 0.0 = \vartheta$$

What about precision / error ??

Precision of the tree code depending on ϑ



Example system: 12,960 particles open-boundary extract of a melting NaCl crystal that includes some density variations

↗ Hands-On Session (Wednesday)

Proper error control

Generalized algebraic kernel (Coulomb, Plummer, r^{-n} , etc.):

$$g_\lambda(r) = \sum_{l=0}^{\lambda} \frac{a_l \cdot r^{2l}}{(r^2 + \sigma^2)^{\lambda + \frac{1}{2}}}$$

Plummer:

$$g_0(r) = \frac{a_0}{(r^2 + \sigma^2)^{\frac{1}{2}}}$$

High order (vortex):

$$g_3(r) = a_3 \cdot \frac{r^2 + \frac{3}{2}\sigma^2}{(r^2 + \sigma^2)^{\frac{3}{2}}}$$

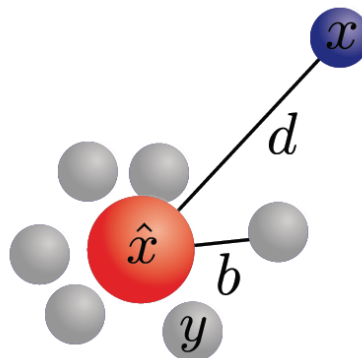
[R. Speck, PhD Thesis, U. Wuppertal (2011)]

Error bound

Can show that error per interaction has upper bound

$$\varepsilon_{(p+1)}^\tau \leq \mathcal{D}_{2\tau-1}(p+1) \cdot \frac{M_0}{(d-b)^{2\tau}} \cdot \left(\frac{b}{d-b}\right)^{p+1}$$

with monopole term M_0 , $b = \sup |y - \hat{x}|$, polynomial \mathcal{D} of rank $2\tau - 1$.



[R. Speck, PhD Thesis, U. Wuppertal (2011)]

The Barnes-Hut Tree Algorithm

Part III: Periodic boundary conditions for tree codes

10.09.2013 | Mathias Winkel

The Barnes-Hut tree code

Tree construction

- Recursive spatial subdivision. . .
- . . . until every particle lies in its own box
 - tree leaves
- Lower-level boxes
 - tree nodes

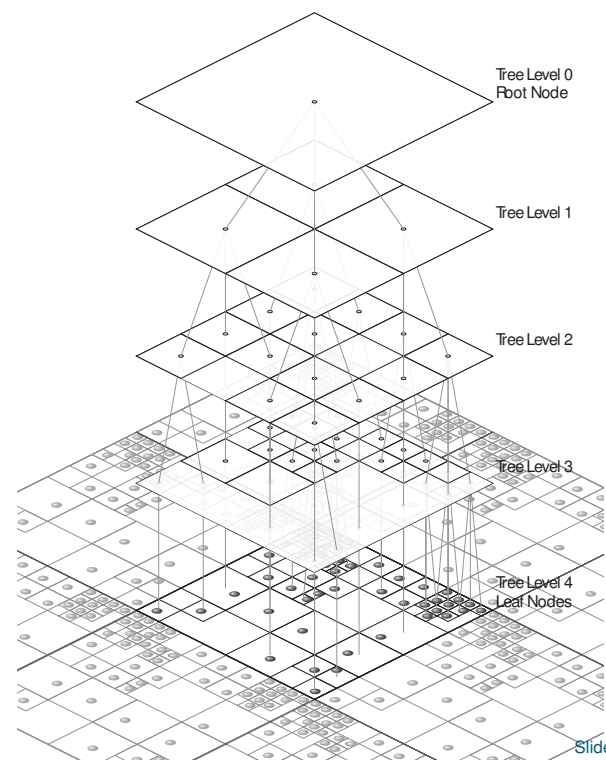
→ Hierarchy of spatially coarsened particle groups

→ Potential interaction partners
but

- $\mathcal{O}(10^{23})$ particles still infeasible
- Artificial surfaces

→ Periodic boundaries

Periodic boundary conditions for tree codes



Periodic boundary conditions for tree codes

An adaptation from the Fast Multipole Method

Bipolar expansion of the inverse distance

$$\frac{1}{|\vec{r}_1 - (\vec{r}_2 + \vec{n})|} = \sum_{l=0}^{\infty} \sum_{m=-l}^l \sum_{j=0}^{\infty} \sum_{k=-j}^j (-1)^j \mathcal{O}_l^m(\vec{r}_1) \mathcal{M}_{l+j}^{m+k}(\vec{n}) \mathcal{O}_j^k(\vec{r}_2)$$

- Multipole coefficients

$$\mathcal{O}_l^m(\vec{r} = [r, \theta, \varphi]) = \frac{r^l}{(l+m)!} P_{lm}(\cos \theta) e^{-im\varphi}$$

- Taylor coefficients

$$\mathcal{M}_l^m(\vec{r} = [r, \theta, \varphi]) = \frac{(l-m)!}{r^{l+1}} P_{lm}(\cos \theta) e^{im\varphi}$$

- associated Legendre polynomials $P_{lm}(z)$

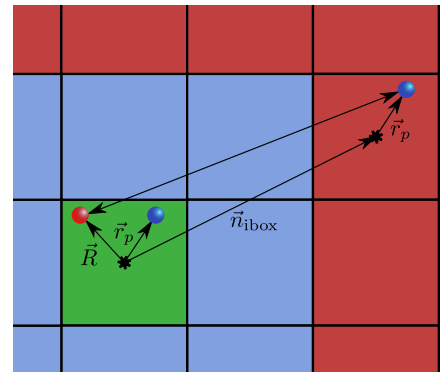
↗ FMM-Lecture by Ivo Kabadshow (Thursday)

Periodic boundary conditions for tree codes

An adaptation from the Fast Multipole Method

- Periodic continuation of the interaction potential

$$\Phi^{\text{lat}}(\vec{R}) = \sum_{\vec{n} \in \mathbb{Z}^3} \sum_{p=1}^N \frac{q_p}{|\vec{R} - (\vec{r}_p + \vec{n})|}$$



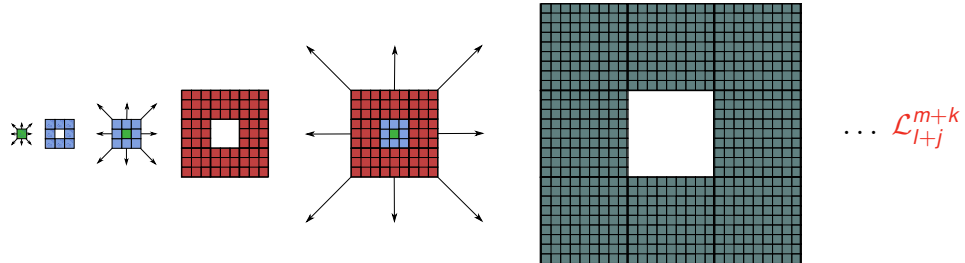
$$\Phi^{\text{lat}}(\vec{R}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathcal{O}_l^m(\vec{R}) \underbrace{\sum_{j=0}^{\infty} \sum_{k=-j}^j (-1)^j \sum_{\vec{n} \in \mathbb{Z}^3} \mathcal{M}_{l+j}^{m+k}(\vec{n})}_{\mathcal{L}_{l+j}^{m+k}} \underbrace{\sum_{p=1}^N q_p \mathcal{O}_j^k(\vec{r}_p)}_{\omega_j^k}$$

$|\vec{n}| > |\vec{R}| + |\vec{r}_p|$ $\mu_l^{m, \text{cent}}$

[M. Challacombe *et al.*, J. Chem. Phys. **107**, 10131 (1997)]

Periodic boundary conditions for tree codes

A renormalization approach for the real space sum



Number of included boxes after n iterations

$$N_n^{\text{FF}, \text{dim}} \propto 3^{n \cdot \text{dim}}$$

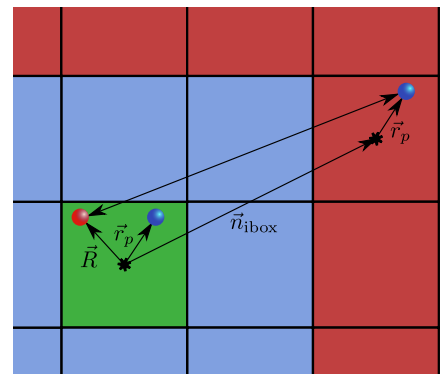
[G. Kudin & G. Scuseria, J. Chem. Phys. **121**, 2886 (2004)]

Periodic boundary conditions for tree codes

An adaptation from the Fast Multipole Method

- Periodic continuation of the interaction potential

$$\phi^{\text{lat}}(\vec{R}) = \sum_{\vec{n} \in \mathbb{Z}^3} \sum_{p=1}^N \frac{q_p}{|\vec{R} - (\vec{r}_p + \vec{n})|}$$



$$\phi^{\text{lat}}(\vec{R}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathcal{O}_l^m(\vec{R}) \underbrace{\sum_{j=0}^{\infty} \sum_{k=-j}^j (-1)^j \sum_{\vec{n} \in \mathbb{Z}^3} \mathcal{M}_{l+j}^{m+k}(\vec{n})}_{\mathcal{L}_{l+j}^{m+k}} \underbrace{\sum_{p=1}^N q_p \mathcal{O}_j^k(\vec{r}_p)}_{\omega_j^k}$$

$|\vec{n}| > |\vec{R}| + |\vec{r}_p|$

$\mu_l^{m, \text{cent}}$

[M. Challacombe *et al.*, J. Chem. Phys. **107**, 10131 (1997)]

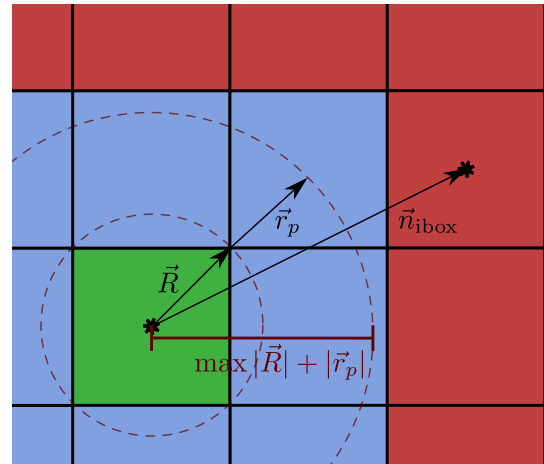
Near field contribution

- near-field treated separately to satisfy convergence criterion

$$|\vec{R}| + |\vec{r}_p| < |\vec{n}|$$

- simple loop through all NF cells
 - worst case: $N^{\text{eff}} = 27 \cdot N$
 - usually significantly less additional interactions
- too non-cubic cell violates convergence criterion

→ increase size of near field



Dipole (Extrinsic-to-Intrinsic) Correction

- electrostatic potential in a lattice (C, L : shape, volume, ... of unit cell/full crystal)

$$\underbrace{\Phi(\vec{r}, C, L)}_{\text{Real space}} = \underbrace{\Phi^{\text{int}}(\vec{r})}_{\text{Ewald}} + \Phi^{\text{ext}}(\vec{r}, C, L)$$

- for cubic unit cell:

$$\Phi^{\text{ext}}(\vec{r}, C, L) = \frac{4\pi}{3V}(\vec{r} - \vec{r}_0) \sum_{k=1}^N q_k \vec{r}_k - \frac{2\pi}{3V} \sum_{k=1}^N q_k |\vec{r}_k|^2$$

(result of infinite summation of conditionally convergent dipole contribution)

$$\Phi^{\text{Ewald}} = \Phi^{\text{far-field}} - \Phi^{\text{ext}}$$

- alternatively: correction of unit cell dipole moment \vec{d} with fictitious charges

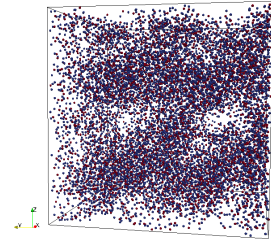
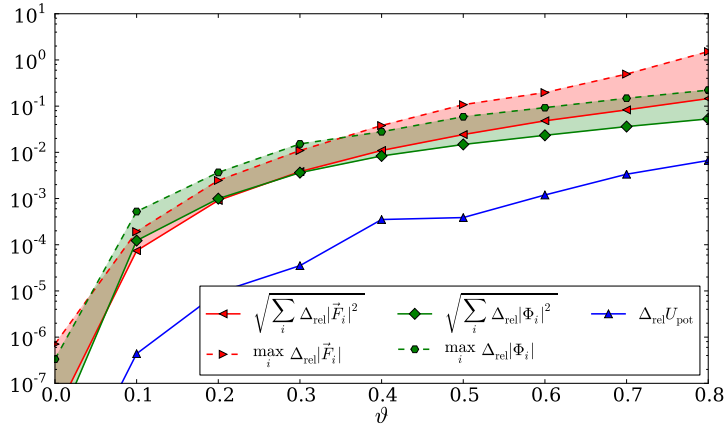
$$\begin{aligned} \vec{p}_1 &= (1,0,0)^T & \vec{p}_2 &= (0,1,0)^T & \vec{p}_3 &= (0,0,1)^T & \vec{p}_4 &= (0,0,0)^T \\ q_1 &= \vec{d}^{(1)}/L^{(1)} & q_2 &= \vec{d}^{(2)}/L^{(2)} & q_3 &= \vec{d}^{(3)}/L^{(2)} & q_4 &= -(q_1 + q_2 + q_3) \end{aligned}$$

[Redlack & Grindlay, J. Chem. Phys. **101**, 5024 (1994)]

[Kudin, Chem. Phys. Lett. **283**, 61 (1998)]

Precision of the periodic tree code depending on ϑ

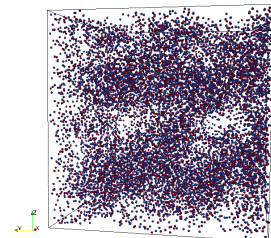
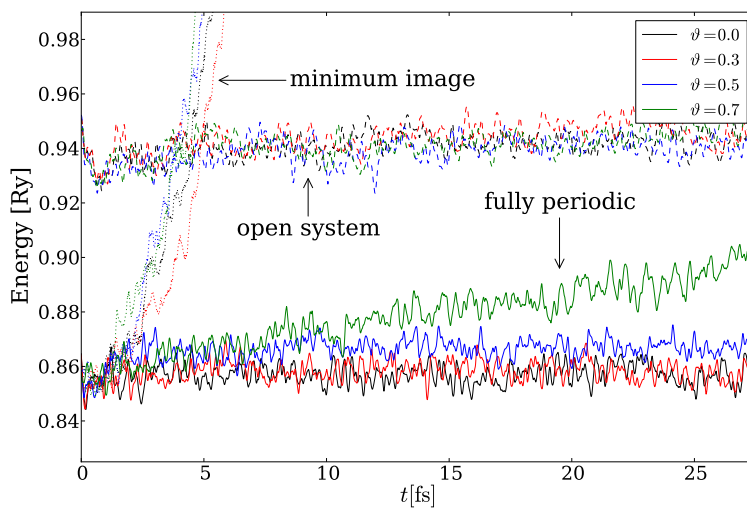
Comparison with standard Ewald method



Example system: 12,960 particles melting NaCl crystal that includes some density variations, periodically continued

Precision of the periodic tree code depending on ϑ

Energy conservation with leap-frog integration



Example system: 12,960 particles melting NaCl crystal that includes some density variations, periodically continued

The Barnes-Hut Tree Algorithm

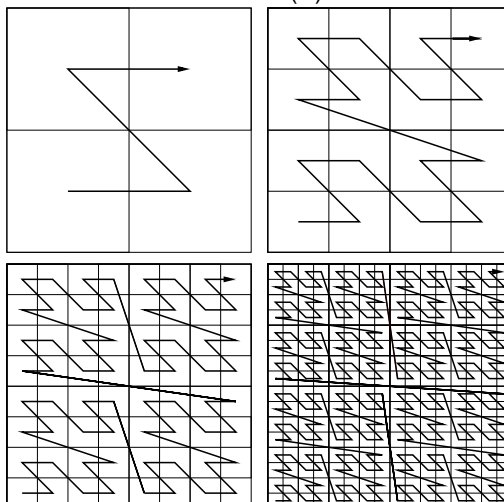
Part IV: Parallelization of Barnes-Hut tree codes

10.09.2013 | Mathias Winkel

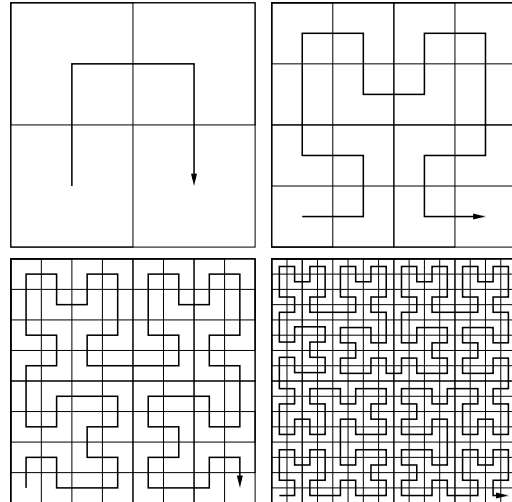
Space Filling Curve

A continuous function $f : \mathcal{I} \rightarrow \mathbb{R}^n$ of the compact set $\mathcal{I} \subset \mathbb{R}$ into \mathbb{R}^n ($n \geq 2$) is called **space filling curve** if its image $f_*(\mathcal{I})$ has a Jordan content (area, volume, ...) greater 0.

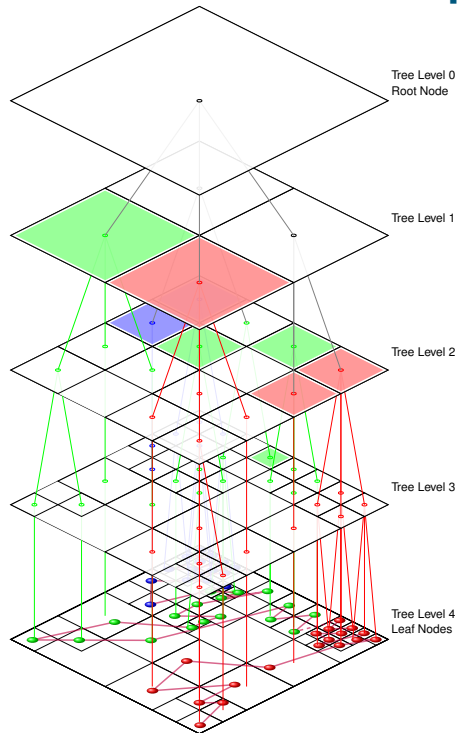
Morton (Z)



Hilbert



Parallel tree buildup



- thread particles onto space-filling curve
- partition space-filling curve (i.e. particles and tree branches above them) across MPI ranks
- distribute local roots ('branch nodes') across all MPI ranks
- build global tree above branch nodes

[J. Barnes & P. Hut, Nature **324**, 446-449 (1986)]

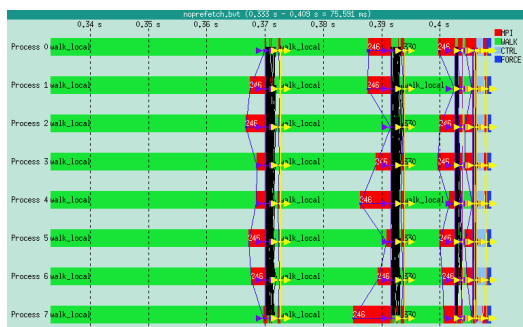
[M. S. Warren & J. K. Salmon, Proc. of the 1993 ACM/IEEE Conf. on Supercomputing, 12-21 (1993)]

Parallelization of Barnes-Hut tree codes

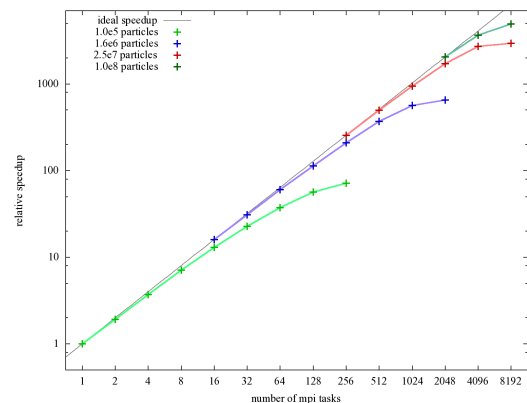
Parallel tree construction

Slide 35

Scaling of original pure MPI scheme



Vampir trace with 8 cores



- reason for scalability problems beyond 8k cores: load-balancing issues during multi-pass tree-walks

[R. Speck, L. Arnold, P. Gibbon, J. Comp. Sci. 2, 138 (2011)]

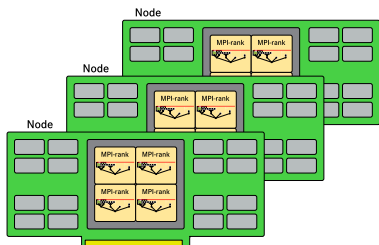
Parallelization of Barnes-Hut tree codes

Parallel tree traversal

Slide 36

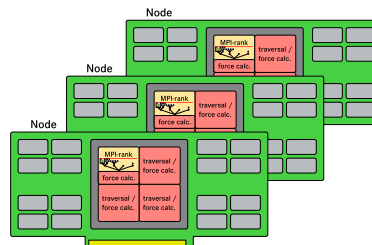
Hybrid parallelization: What do we gain?

previously:



one MPI-rank per processor core

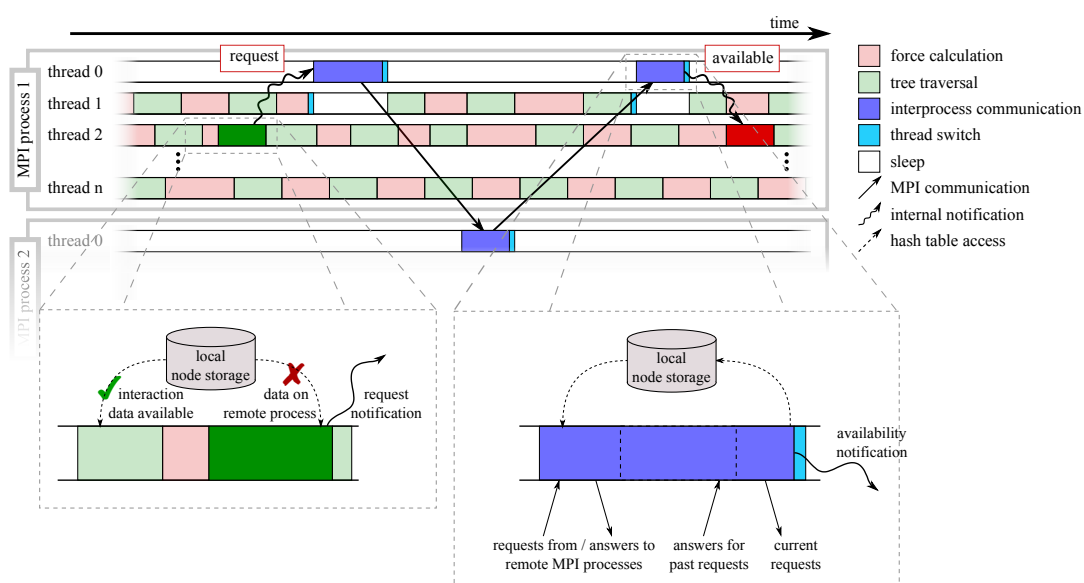
now:



one MPI-rank and several threads per compute node

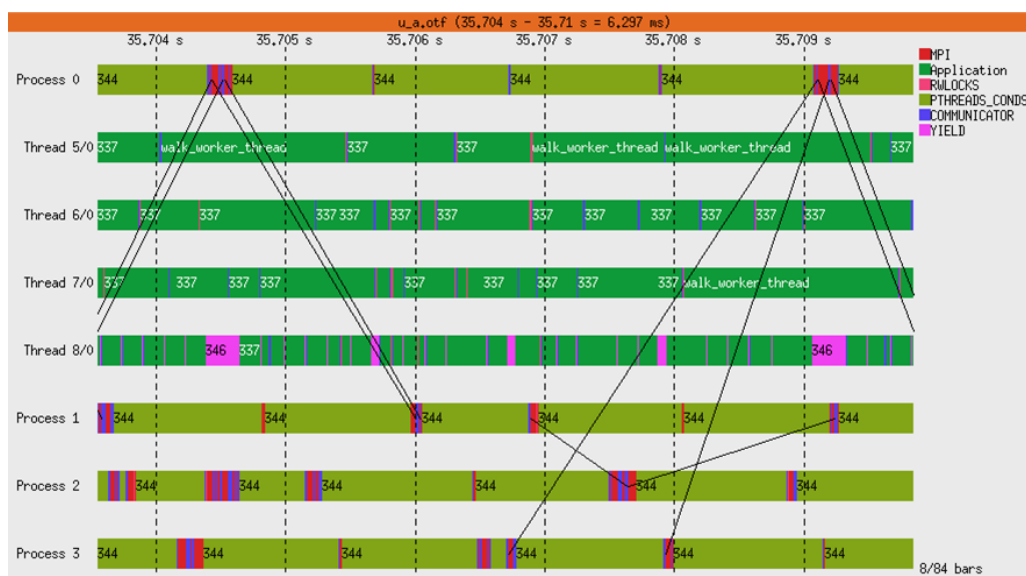
- only one tree per compute node instead of per core
- reduction of number of MPI ranks per node leads to:
 - fewer multipole/tree-node copies = less MPI communication and lower storage needed
 - reduced effective answer latency for individual tree node requests
- potential for overlap of computation and communication via dedicated communicator thread

Hybrid parallelization: MPI & PThreads in detail



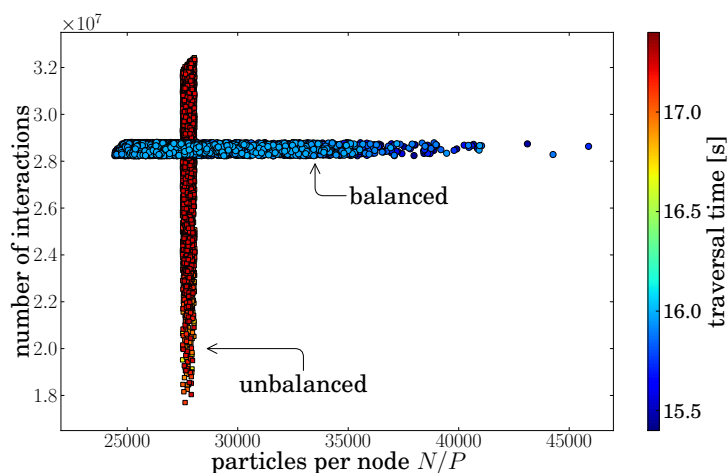
[M. Winkel *et al.*, *Comp. Phys. Comm.* 187, 880–889 (2012)]

Hybrid parallelization: MPI & PThreads in detail



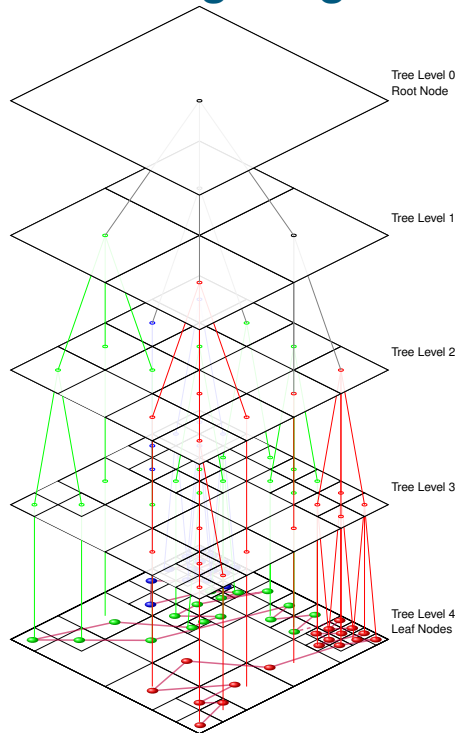
[M. Winkel *et al.*, *Comp. Phys. Comm.* 187, 880–889 (2012)]

Load balancing



- number of interactions per MPI rank significant, **not** number of particles
- total runtime reduces significantly
- based on workload of previous simulation timestep

Prefetching / Eager sending



- requests for tree nodes processed by remote communicator thread
 - additional payload information in request: particle position
 - provisional traversal by remote communicator
 - identification of all necessary tree nodes (not just one level)
 - answer with child, grandchild, ... nodes
- less requests, avoids latency for sequential requests

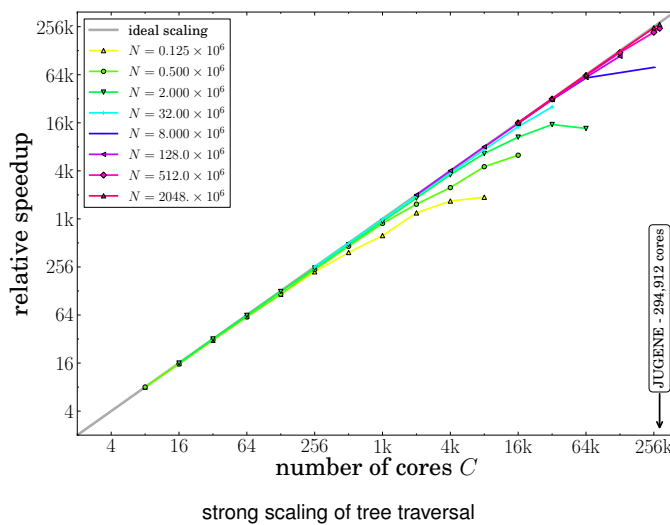
Parallelization of Barnes-Hut tree codes

Additional notes on parallelization

Slide 40

Hybrid parallelization: MPI & PThreads in action

Scaling things up – Homogeneous, quasi-neutral cube



- up to 2×10^9 particles on 294 912 processors
- load-balanced

Parallelization of Barnes-Hut tree codes

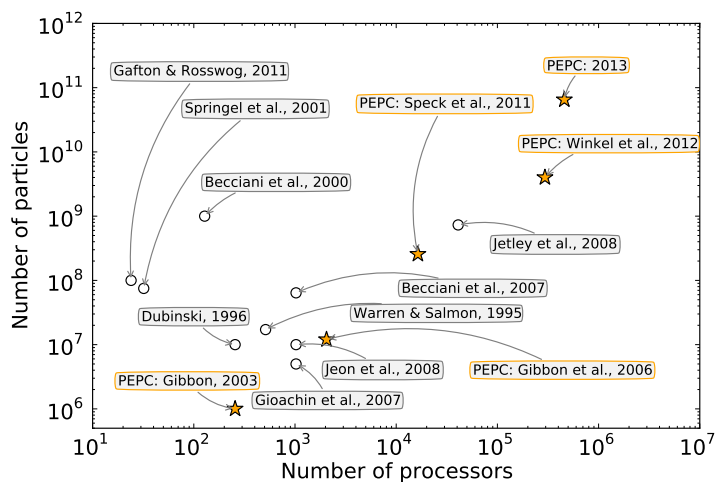
Hybrid scaling

Slide 41

PEPC – The Pretty Efficient Parallel Coulomb Solver

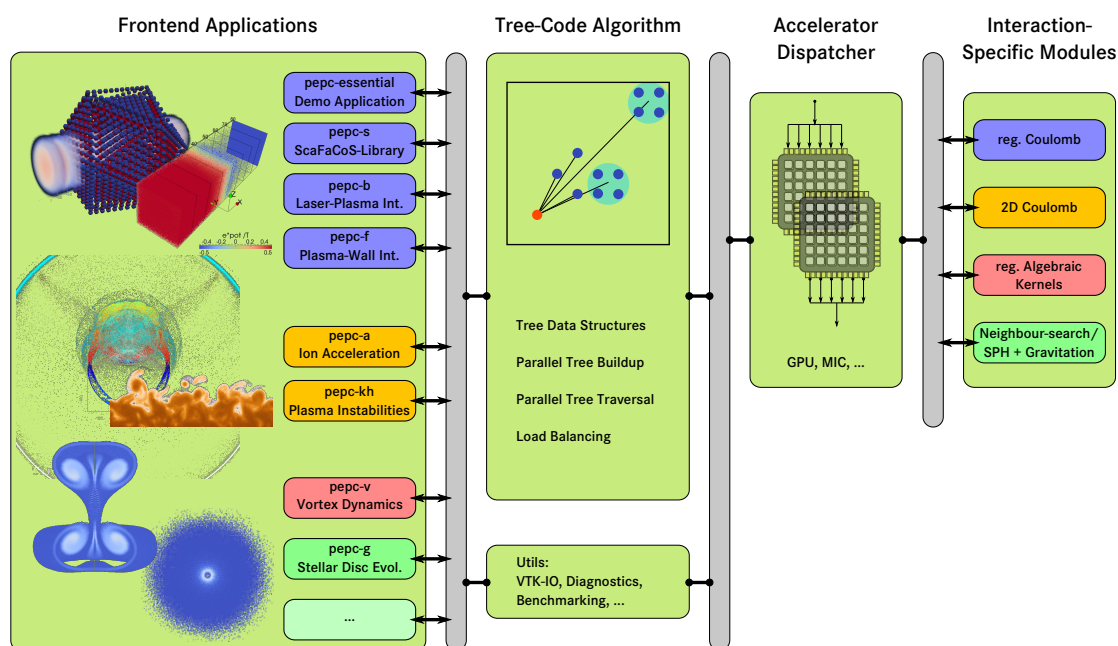
World records in scalability and particle number for Barnes-Hut tree codes

- **Parallelization:**
MPI + pthreads
- **Target Architectures:**
 - GNU, OSX, ARM, ...
 - JuGene (BG/P)
2 billion particles on 294,912 processors
 - JUQUEEN (BG/Q)
65 billion particles on 458,752 processors
 - JuRoPa (Nehalem cluster)
- sophisticated inter- & intra-node load balancing
- **unique modular concept**



PEPC is open-source, freely available via www.fz-juelich.de/ias/jsc/pepc
 mailing list for reaching all developers: pepc@fz-juelich.de

PEPC – The Pretty Efficient Parallel Coulomb Solver



[M. Winkel et al., Comp. Phys. Comm. 187, 880 (2012)]

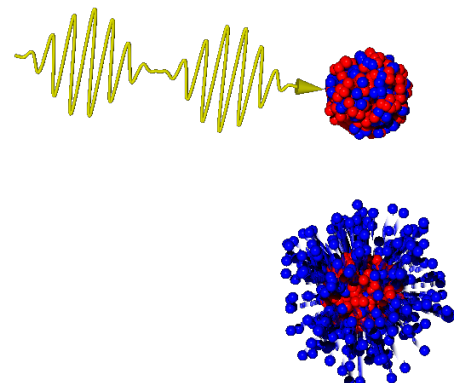
The Barnes-Hut Tree Algorithm

Part V: Applications

10.09.2013 | Mathias Winkel

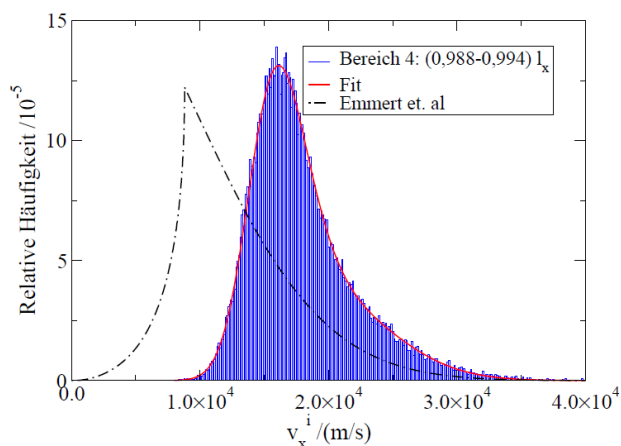
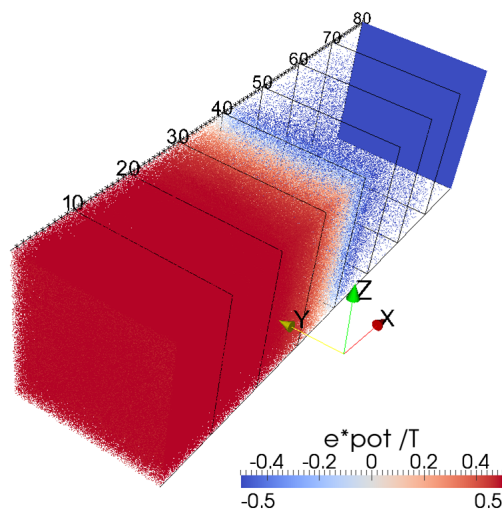
Strongly coupled plasmas

- Transport coefficients of warm, dense matter
- collective effects in nano clusters
- Strong coupling + non-equilibrium
- Laser-solid interactions, stellar interior, inertial fusion



[M. Winkel, PhD thesis, RWTH Aachen (2013)]

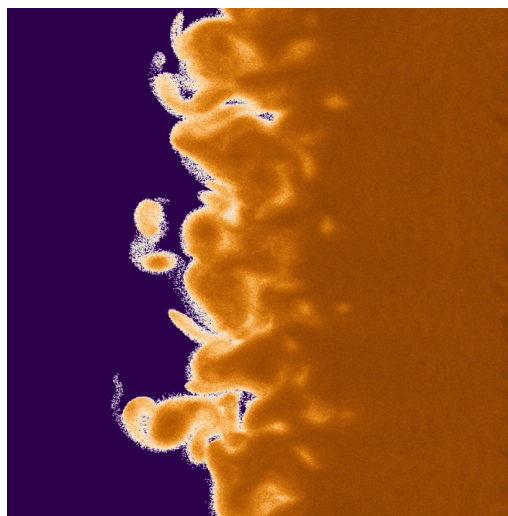
Plasma-wall interaction in tokamaks



[B. Berberich, PhD thesis (2011); C. Salmagne (IEK-4), in progress]

Magnetized Plasmas

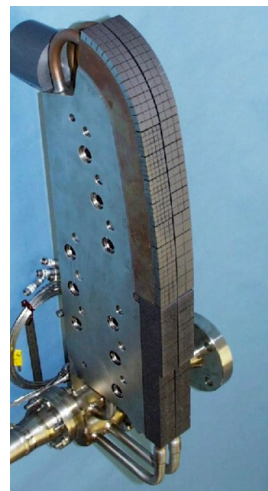
- Kelvin-Helmholtz instabilities at plasma-vacuum interfaces
- Driving mechanism:
 - 1 magnetised plasma slab, $n = \text{const.}$, in contact with vacuum
 - 2 different Larmor radii lead to different density gradient scales
 - 3 charge separation creates a sheared electric field
 - 4 sheared $\vec{E} \times \vec{B}$ flow feeds KH instability



[B. Steinbusch (JSC), in progress]

Boundary element method

- plasma wall interaction in fusion vessels
- complex geometry: castellated tiles
- metal walls, free charges
- real world geometries can be modeled in a CAD program
- Supported boundary conditions:
 - (mixed) Dirichlet and Neumann
 - periodic
 - metal wall with floating potential

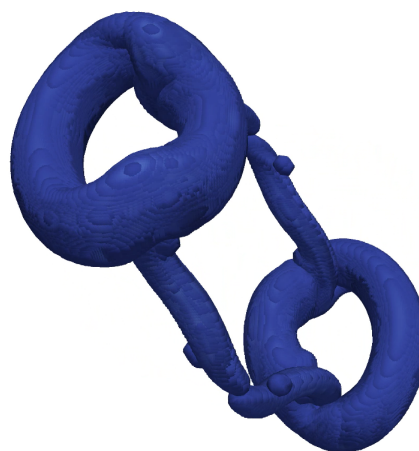


M. Hellwig (IEK-4)

[B. Steinbusch (JSC), in progress]

Vortex fluids

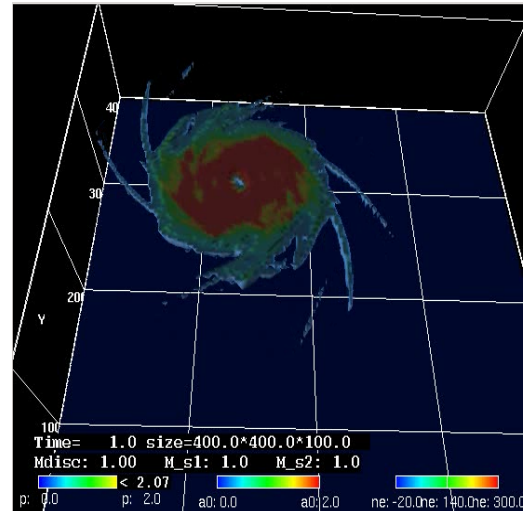
- Mathematical equivalence between Navier-Stokes vortex equations and magnetostatics
- Vortex particles must overlap → remeshing
- Multipole expansion of *smoothed* vorticity kernel



[Robert Speck, PhD thesis, U. Wuppertal]

Self-gravitating accretion discs

- Planet formation & disc dynamics with first principles particle simulation (SPH + gravity)
- nearest neighbour search kernel
- Here: $M_{\text{disc}} = 0.5M_{\odot}$
- Observed: $M_{\text{disc}} = 0.01 - 0.1M_{\odot}$;
 $M_E = 10^{-4}M_{\text{disc}}$



[Andreas Breslau, Susanne Pfalzner, MPI Radioastronomie Bonn]

Contributors

- Paul Gibbon (Head of group; JSC): first parallel implementation at JSC (2003)
- Michael Hofmann, (U. Chemnitz): parallel key sort
- M. W. (JSC): hybrid MPI-Pthread algorithm; periodic boundaries; strongly-coupled plasmas
- Benedikt Steinbusch (JSC): hybrid scaling, performance, boundary element method, wall-plasma interactions
- Dirk Brömmel (SLPP; JSC): novel architectures, accelerators, GPU
- Christian Salmagne (IEK-4): wall-plasma interactions
- Andreas Breslau (MPIfR Bonn): near-neighbour search; SPH; protoplanetary discs
- Robert Speck (JSC, U. Lugano): vortex model
- Lukas Arnold (JSC, now U. Wuppertal): novel architectures, performance
- Helge Hübner (JSC, now U. Hamburg): Hilbert curve, memory analysis, branch node optimizations
- Benjamin Berberich (IEK-4, now Carl-Zeiss): gyrokinetics; 3rd-party collisions; wall-plasma interactions

Further reading

- M. Winkel, R. Speck, H. Hübner, L. Arnold, R. Krause, P. Gibbon:
A massively parallel, multi-disciplinary Barnes-Hut tree code for extreme-scale N-body simulations
Comp. Phys. Commun. (2012).
- R. Speck, L. Arnold, P. Gibbon:
Scaling and Efficiency of the PEPC Library
J. Comp. Sci., **2**, 138 (2011).
- P. Gibbon, R. Speck, L. Arnold, M. Winkel, H. Hübner:
Parallel Tree Codes,
in Fast Methods for Long-Range Interactions in Complex Systems, Summer School, 6-10 September 2010, Jülich.
- P. Gibbon, R. Speck, A. Karmakar, L. Arnold, W. Frings, B. Berberich, D. Reiter, M. Mašek:
Progress in Mesh-Free Plasma Simulation with Parallel Tree Codes
IEEE Trans. Plasma Sci. **38**, 2367-2376, (2010).
- S. Pfalzner, P. Gibbon:
Many Body Tree Methods in Physics,
Cambridge University Press, New York (September 2005), ISBN 0-521-01916-8.

Thank you...

... for your attention.

Questions?

www.fz-juelich.de/ias/jsc/pepc
pepc@fz-juelich.de

