

Simulation Science on Windows

Christian Terboven <terboven@rz.rwth-aachen.de>

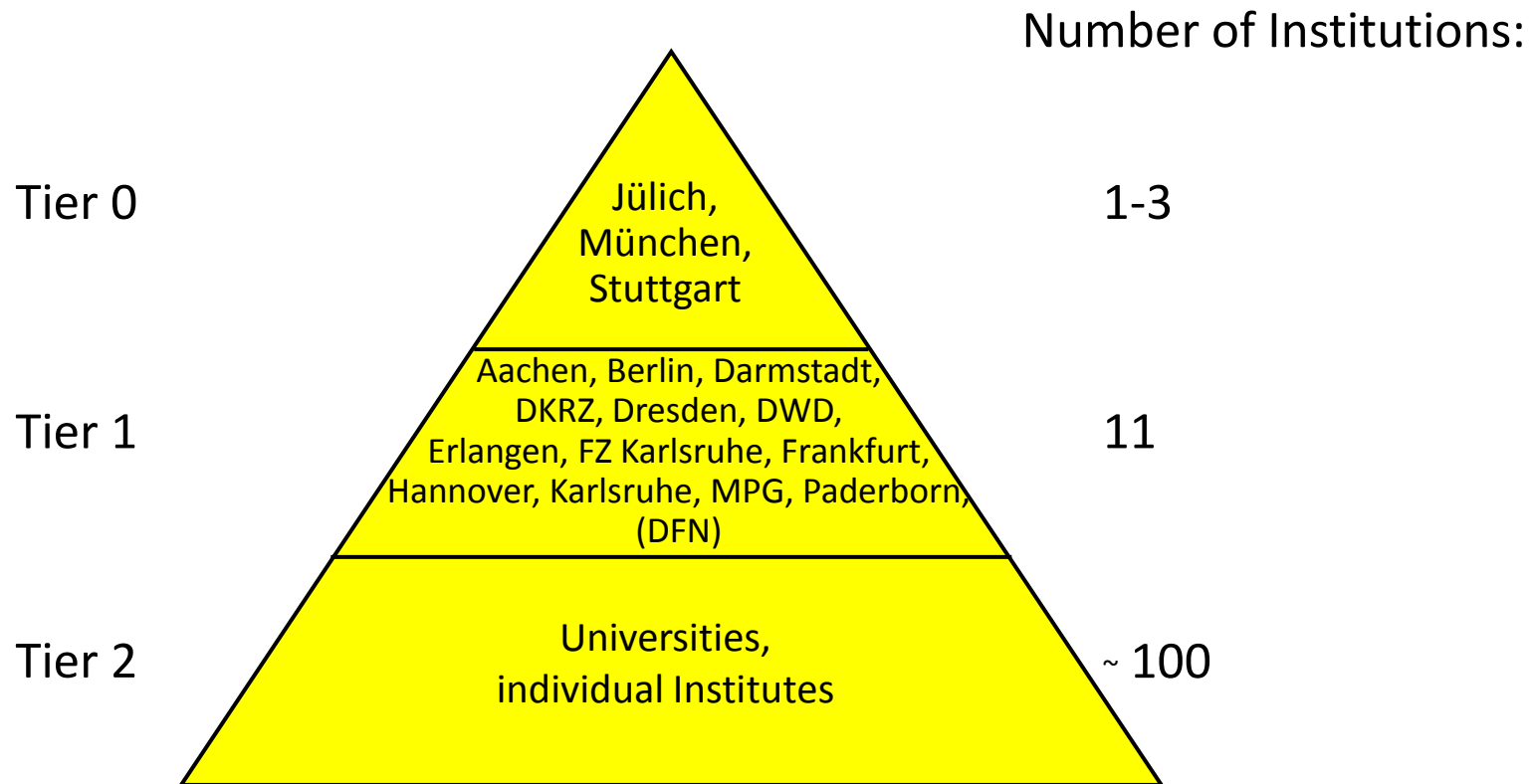
15.11.2010 / New Orleans, LA, USA

Stand: 14.11.2010

Version 2.3

- ▶ **The German Windows HPC User Group**
- ▶ **Computational Sciences Applications**
- ▶ **Looking into the Fog of Cloud Computing**
- ▶ **What is missing**
- ▶ **Summary**

The German Windows HPC User Group



- ▶ **Simulation Science: Main Pillar of Research**
- ▶ ***RWTH: The simulation of complex models lies at the core of innovative products and basic research at RWTH Aachen University.***



32943 students



454 professorships



2488 academic staff



280 institutes

Natural Sciences

24%

Medicine,
Dentistry

9%

Humanities, Social
Sciences
and Economics

16%

51%

Engineering

Students by Discipline

The three funding lines:

Graduate School

Aachen Institute of Advanced Study in Computational Engineering Science (AICES)

Cluster of Excellence

Ultra High-Speed Mobile Information and Communication (UMIC)

Integrative Production Technology for High-wage Countries

Tailor-made Fuels from Biomass

Institutional Strategy

RWTH 2020 – Meeting Global Challenges

including:

JARA | Jülich Aachen Research Alliance

JARA|BRAIN

JARA|ENERGY

JARA|FIT

JARA|HPC

The Integrated Interdisciplinary University of Technology

Further information at www.rwth-aachen.de/exzellenz

- ▶ **We do HPC on Unix for many years - why try Windows?**
 - ▶ Third party cooperations often require the Windows platform
 - ▶ Desktop-like HPC experience for usage and development

- ▶ **We started early:**
 - ▶ 03/04: Intel ThreadingTools GUI only available on Windows
 - ▶ 2004: Visual Studio 2005 with OpenMP support (beta program)
 - ▶ 2005: Windows Compute Cluster Server (beta program)
 - ▶ 2007: Windows (HPC) Server 2008 (beta program)
 - ▶ 2009: Windows (HPC) Server 2008 R2 (beta program)

- ▶ **Cooperation with Microsoft Germany since 2008:**
WinHP3C = Windows High Performance Cluster Competence Center



▶ **Subjects of this Cooperation:**

- ▶ Top500 project for Cluster installed in 06/2008 (done)
- ▶ Open Source network + HPC projects (see, e.g., Virtual Reality)
- ▶ Foundation of independent User Group for HPC on Windows
 - ▶ Focus on Germany, moderated by RWTH, supported by Microsoft
- ▶ RWTH: User training „HPC on Windows“ open for everyone
- ▶ RWTH: Provision of benchmarking resources on request
- ▶ Microsoft: Technical support for RWTH (and User Group)

▶ **Participation in**

- ▶ TAP: Technology Adoption Program for HPC Server v3
- ▶ CAB: Customer Advisory Board

3rd meeting of German HPC user group

- 11-12 March 2010 at Fraunhofer-Gesellschaft
- Birlinghoven Castle near Bonn



Activities: Emphasis on User Training

Date	Event	Location
05/2006	Windows-HPC training (de), 1 day, 36 participants	Aachen
04/2007	Windows-HPC training (en), 1 day, 42 participants	
05/2007	Intel Multi-Core Programming for Academia (w/ LfBS), 45 p.	Aachen
04/2008	1st Windows-HPC User Group Meeting, 1.5 days, 110 part.	Aachen
04/2008	Windows-HPC Deep Dive Training (en), 3 days, 20 part.	Aachen
07/2008	Windows-HPC + Deployment training (de), 2 days, 55 part.	Aachen
09/2008	Windows-HPC + Deployment training (en), 2 days, 30 part.	Aachen
02/2009	Parallel Programming in CES (en), 1 week, 50 part.	Aachen
03/2009	2nd Windows-HPC User Group Meeting, 1.5 days, 90 part.	Dresden
06/2009	Visual Studio 2008 + Windows HPC Server workshop, 20 p.	Kingston, CA
06/2009	Windows-HPC demo station at the Microsoft ISC booth	Hamburg
03/2010	Windows-HPC Deep Dive Training (en), 3 days, 20 part.	St. Augustin
03/2010	3rd Windows-HPC User Group Meeting, 1.5 days, 75 part.	St. Augustin
03/2010	Windows-HPC Deep Dive Training (en), 3 days, 30 part.	St. Augustin
03/2010	Parallel Programming in CES (en), 1 week, 35 to 80 part.	Aachen
03/2011	Planned: 4th Windows-HPC User Group Meeting, 1.5 days	Karlsruhe

Porting Open Source HPC Software to Microsoft Windows Platforms

- Portal for open source software developed and ported by Fraunhofer-SCAI (Elmer, OpenFOAM)
- In the future: uploads and downloads of YOUR open source software
- Best practises



Open MS HPC

URL: <http://www.scai.fraunhofer.de/openmshpc.html>

- **gromacs (3.3.1)**
 - Molecular dynamics software, C code with Fortran/Assembler subroutines, MPI parallelized
- **metis (4.0.0)**
 - Set of serial graph partition algorithms, used by many codes to find a (communication minimizing) mesh partition
- **Elmer**
 - Open Source Finite Element Software for
- **samgp – parallel algebraic multigrid for systems**
 - High performance solver for linear systems, potentially $O(N)$, developed at SCAI, Fortran Code , MPI parallelized
- **Openfoam**
 - Open Field Operation and Manipulation, Includes solver for many CFD problems

Computational Sciences Applications

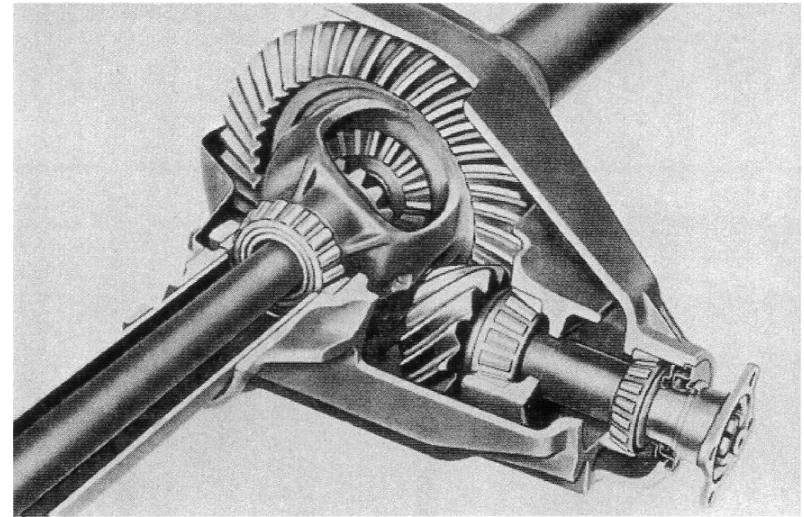
▶ Contact analysis simulation of Bevel Gears

- ▶ Written in Fortran, using Intel Fortran 10.1 compiler
- ▶ Very cache-friendly → runs at high Mflop/s rates

Bevel Gear Pair



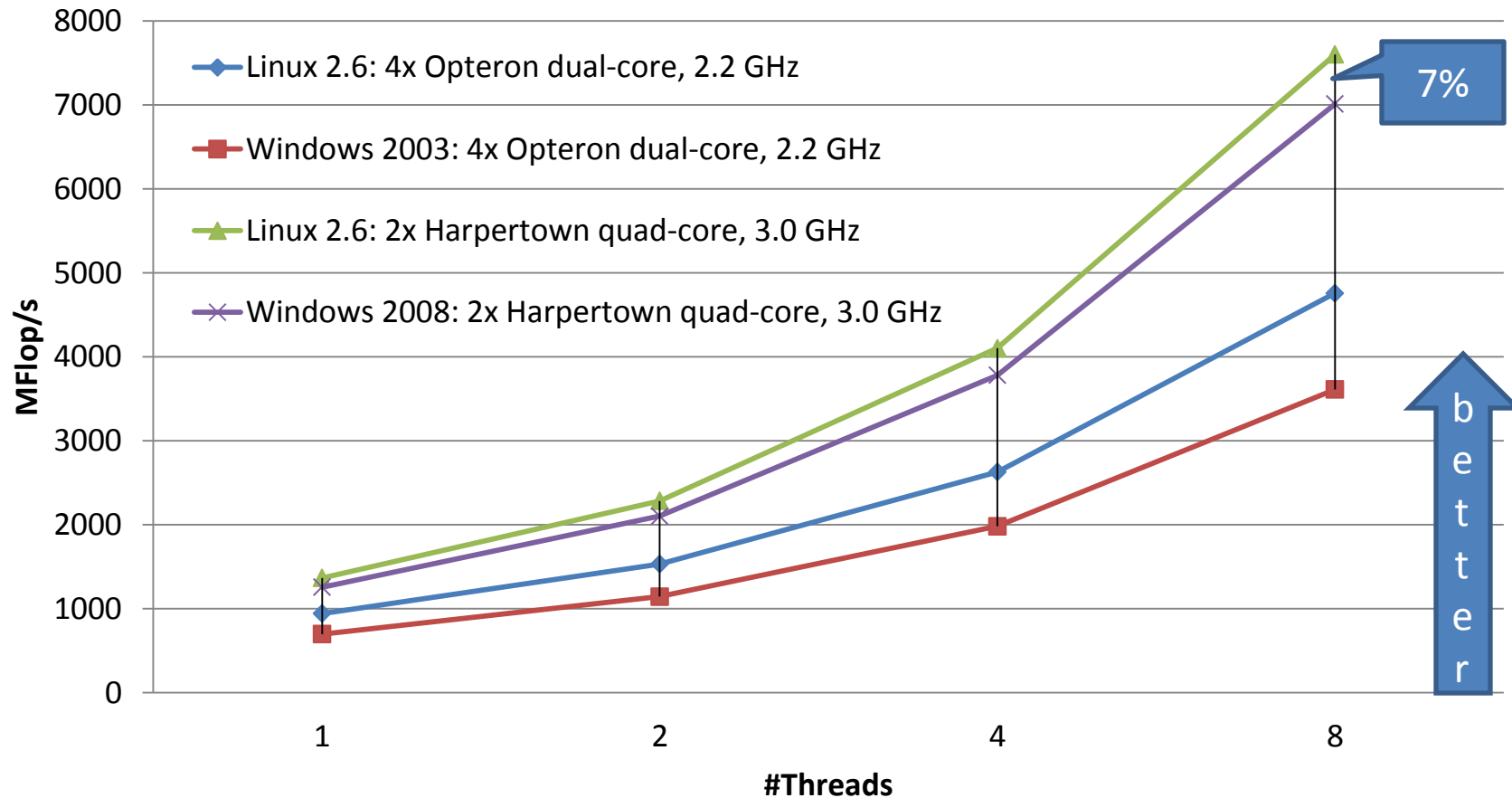
Differential Gear



- ▶ **Development is done purely on Windows, using the Intel Fortran integration in Visual Studio**
 - ▶ MATA / MATSE apprentices for long-term SW support
- ▶ **Parallelization for Multicores and GPGPUs in work**
 - ▶ Clusters are often not an option for small to medium-sized companies interested in using the software
 - ▶ Parallelization for Multicores with OpenMP, for GPGPUs with OpenCL → code is hardware-agnostic
- ▶ **Software is used by 40+ project partners from industry**
 - ▶ Need for compute power only during certain project phases
→ Technical Computing in the Cloud would be an option

▶ Comparing Linux and Windows HPC Server 2008:

Performance of KegelToleranzen

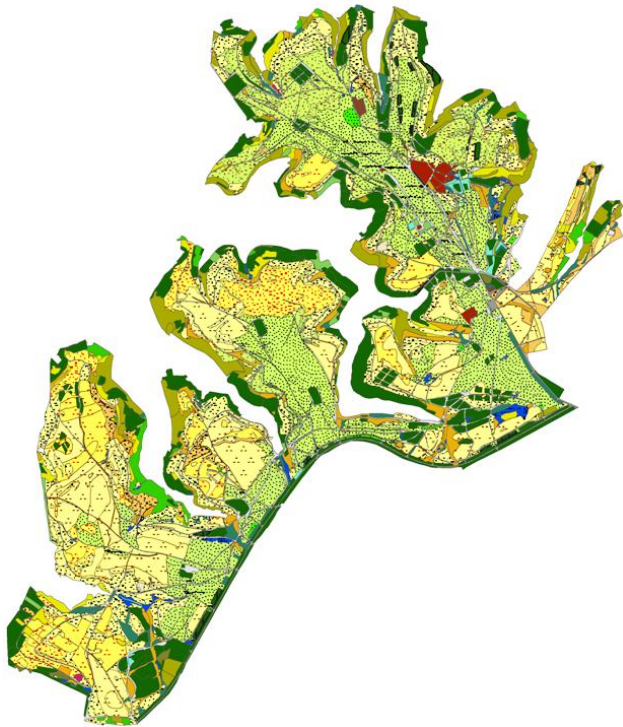


Performance gain for the user: Speedup of 5.6 on one node.
Compared to the desktop (220 MFlop/s): Speedup of 32!

▶ **Institute for Environmental Research (Bio5) and Research Institute for Ecosystem Analysis and Assessment (GAIAC)**

- ▶ Grassland development under various land use forms

→ decision support for grassland management on the landscape scale



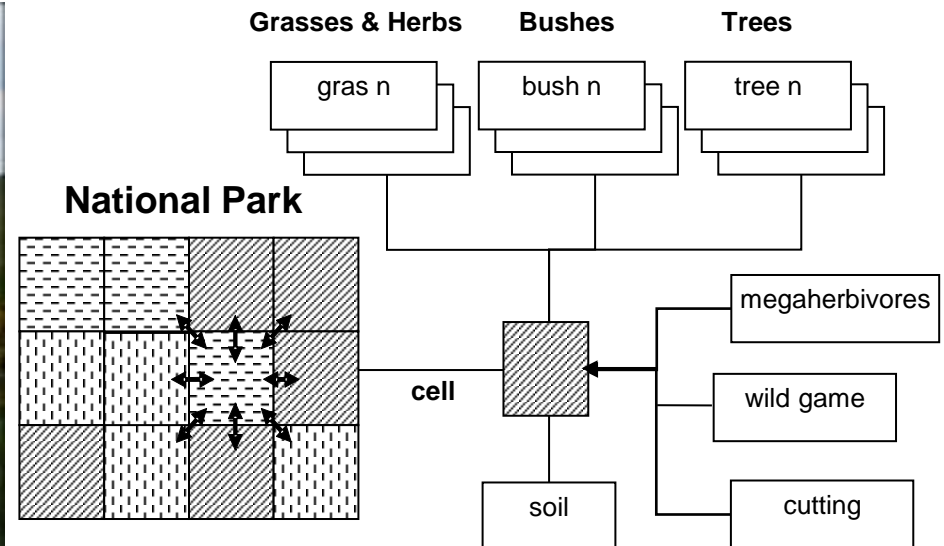
Dataset: Eifel National Park

established in 2004, includes open grasslands of former military training site

Burg Vogelsang

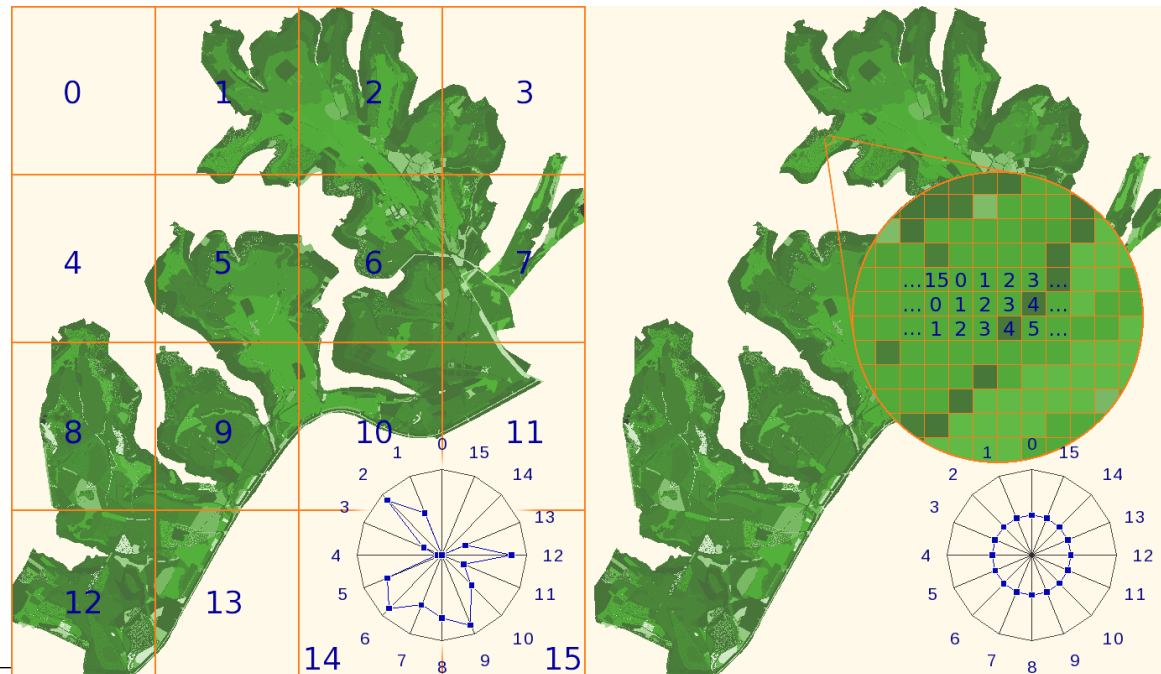
Simulation tool developed at Bio5 / GAIAC, developed in Delphi on and for Windows.

- ▶ **Spatial Explicit Dynamic model for Grassland Succession**
 - ▶ Modeling based on Geographic Information System (GIS)
- ▶ **Several issues with respect to software engineering**
 - ▶ GraS-Model is written in Delphi which is 32-bit only
 - ▶ Parallelization with MPI funded by JARA (project start in Apr 09)
 - ▶ JARA = Jülich Aachen Research Alliance (www.jara.org)

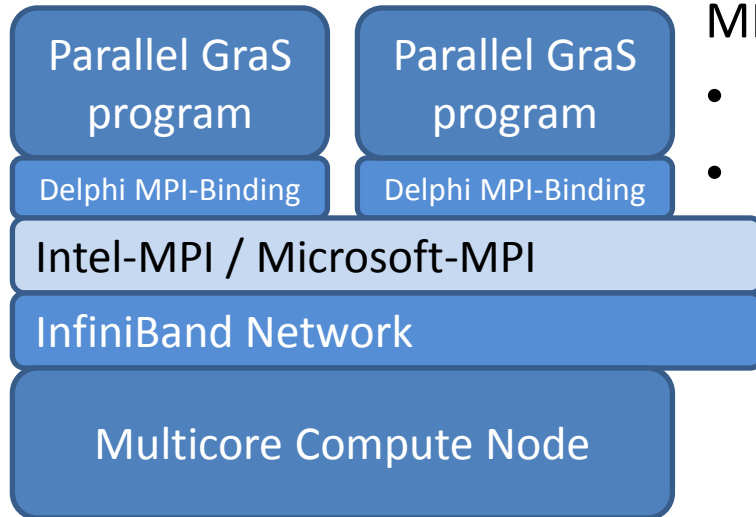


- ▶ Landscape map is flattened and separated in square cells of 10 meters edge length → rectangular grid.
- ▶ A simulated day consists of the following steps, performed independently for each cell:

1. Growth simulation for every plant contained in the cell
2. Calculation of the amount of seeds to be exchanged with other cells
3. Performing the seed exchange
4. Calculation of the impact of seed received by other cells



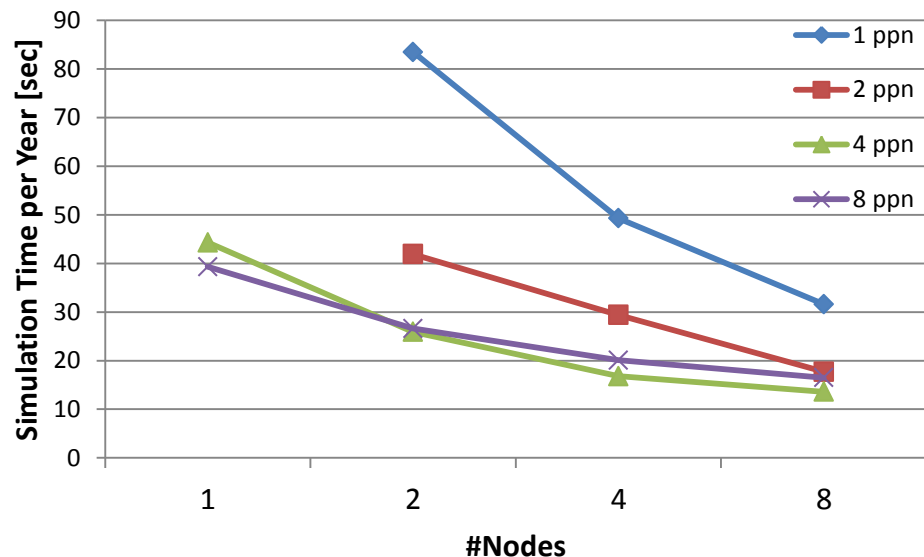
- ▶ **Delphi comes with virtually no support for parallel programming, so we created / modified a suitable environment:**



MPI Performance Analysis:

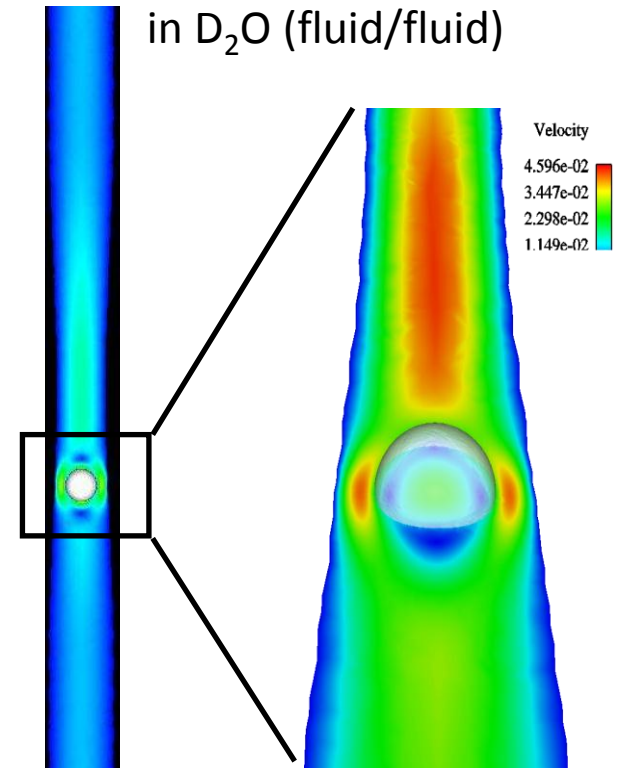
- I-MPI: Intel Trace Analyzer and Collector
- MS-MPI (w/ Tracing): Vampir on Windows

- Provision as a Web Service is the goal of a project with the University of Paderborn:
→ Technical Computing in the Cloud as HPC overflow resource



- ▶ Numerical Simulation of two-phase flow
- ▶ Modeled by instationary and non-linear Navier-Stokes equation
- ▶ Level Set function is used to describe the interface between the two phases
- ▶ Written in C++: is object-oriented, uses nested templates, uses STL types, uses compile-time polymorphism, ...
- ▶ (Adaptive) Tetrahedral Grid Hierarchy
- ▶ Finite Element Method (FEM)

Example: Silicon oil drop in D₂O (fluid/fluid)



```
PCG(const MatrixCL& A, VectorCL& x, const VectorCL& b,  
    const PreCon& M, int& max_iter,  
    double& tol)  
{  
    VectorCL p(n), z(n), q(n), r(n);  
    [...]  
    for (int i=1; i<max_iter; i++)  
    {  
        [...]  
        q = A * p;  
        double alpha = ...;  
        x += alpha * q;  
        r -= alpha * q;  
        [...]  
    }  
}
```

```
y_Ax_par(&q.row()[0],  
         A.num_rows(), A.row_val(),  
         A.row_row(), A.row_col(),  
         Addr(p.row()));
```

Parallelization is often a hard task!

We need IDE-support for (at least):

- Design + Development
- Parallelization
- Parallel Debugging
- Parallel Performance Analysis

```
omp for reduction  
    (+:alpha_sum)  
    for (long j=0; j<n; j++)  
        alpha_sum += p[j]*q[j];  
omp single {  
    rho/alpha_sum;
```

► **This programming style asks for solid integration of compiler and debugger!**

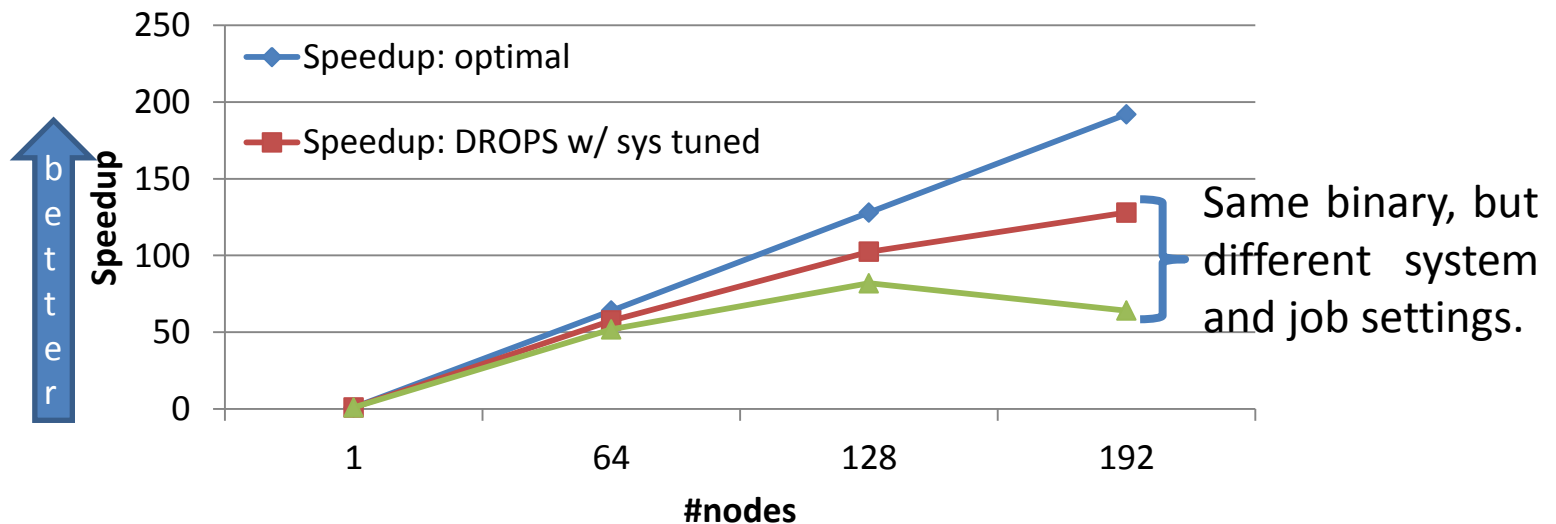
```
#pragma omp for  
for (long j=0; j<n; j++){  
    x[j] += alpha * p[j];  
    r[j] -= alpha * q[j];  
}
```

► Strength of Windows Development Platform:

- Visual Studio is a powerful C++ development environment
- Exceptional support for parallel Shared-Memory debugging
- DROPS: Development of parallel versions shifts to Windows

Porting of ParMETIS: Graph partitioning for load balancing

Porting of DDD: Library for high-level network communication



Looking into the Fog of Cloud Computing

- ▶ **The IKA (Institut für Kraftfahrzeuge) of RWTH Aachen University is a partner in euroFOT: European Field Operational Test on Active Safety Systems (<http://www.eurofot-ip.eu/>)**
 - ▶ 240 vehicles continuously collect data to be sent to a central server for analysis and long-term storage in a database
 - ▶ Expected data volume: 12 TB

 - ▶ **Services provided by the RZ:**
 - ▶ Central Server to collect the data sent from the vehicles
 - ▶ Database storage with long-term support
 - ▶ Hosting and development of an .NET component to automatically start MATLAB jobs for data analysis as soon as packages have been transferred
- **HPC is *just* one part in the Workflow**

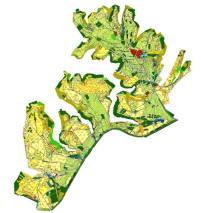
▶ **Science Cloud: GraS-as-a-Service**

- ▶ Evaluation of a Cloud Environment: Eucalyptus + ESX-Server
- ▶ Application = Web-Portal + VM image + Mgmt. services
- ▶ GraS: Grassland development under various land use forms

Institute for Environmental Research (Bio5) and GAIAC

▶ **Goal: Find answers to the following questions:**

- ▶ Which kind of (HPC) applications can be run in a Cloud?
- ▶ How do customers like the Portal + Cloud approach?



▶ **Current state of the project:**

- ▶ Application has been ported, Portal has been implemented (HPC)
- ▶ Waiting for network and machines to be set up (RuD + UPB)
- ▶ Poster + Demo at our booth at the ISC in June 2010, Hamburg

▶ Web-Portal for submitting GraS jobs into the Cloud:



(Tip: you can [submit](#) here a new job.)

ID	Owner	Comments	Status	Internal State	runtime	Cloud: Computing Instances
1	ca190236		aborted	[]	0:02:53.546617	0
2	ca190236		aborted	[]	0:00:47.979302	0
3	ca190236		finished	[]	0:01:58.685092	0
4	ca190236		preperation	[]	n/a	0

super user can see all jobs

Submitting a new Job

please, fill out:

config file *

Comments

Amount of Nodes

Amount of Processes per Node

▶ Open question:

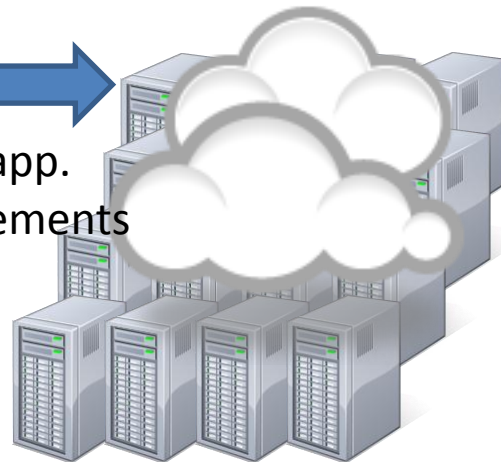
- ▶ What is the right level of abstraction and interaction?
- ▶ How to specify set of resources / time requirements / cost limits?
- ▶ User wants to run an application (not care about number of nodes / cores / ...)

- ▶ User develop model at a workstations and uses a GUI to setup and submit jobs.



- KegelToleranzen app.
- Ressource requirements
- Job configuration
- Misc. input files

Cloud



- User uses a GUI to download simulation results and perform analysis.



- Simulation result



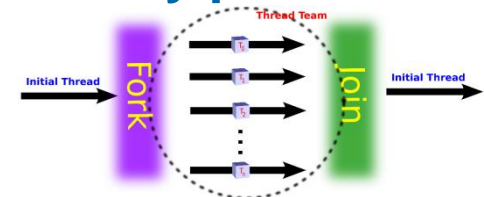
- ▶ Geometry files (large) may live in the Cloud.



- ▶ *Worker Role* just starts the native application

- ▶ Reacts to abort-signal
- ▶ May report application progress

XL instances needed to support shared memory parallelism



KegelToleranzen on MS Azure (2/2)

- ▶ User develop model at a workstations and uses a GUI to setup and submit jobs.

GUI: Web-Portal or Silverlight-app
Cloud

- ▶ User uses a GUI to download simulation results and perform analysis.



- KegelToleranzen app.
- Ressource requirements
- Job configuration
- Misc. input files



MPI support is missing even on the node!

- Simulation result



- ▶ Geometry files (large) may live in the Cloud



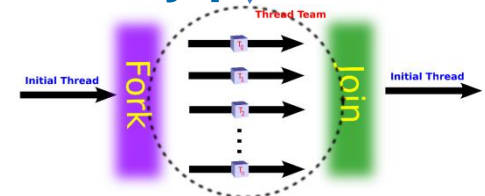
- ▶ Worker Role just start

Shared storage has several shortcomings

- ▶ Reacts to abort-signal
- ▶ May report application progress

OpenMP works well if node is large enough

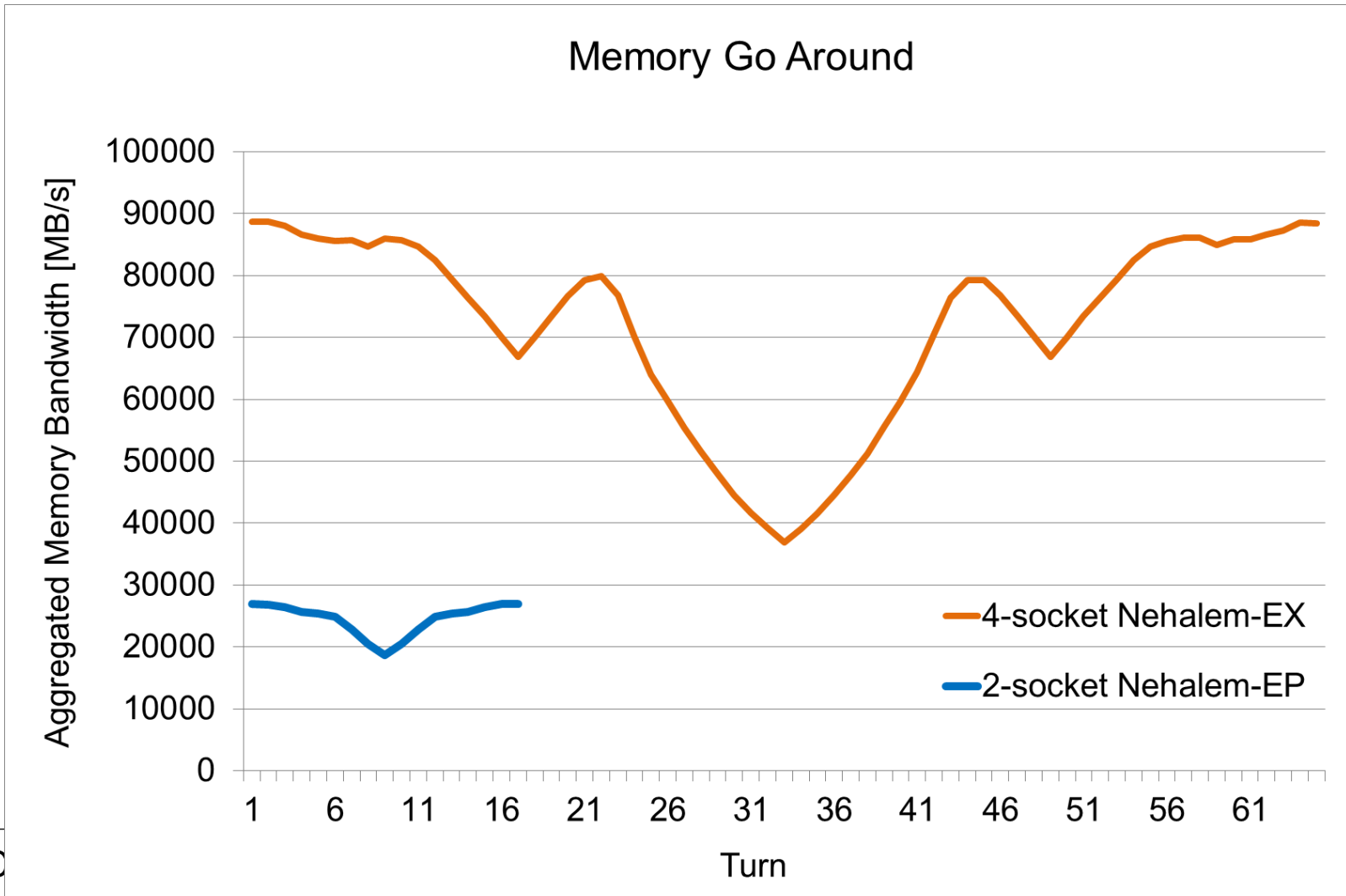
XL instances needed to support shared memory parallelism



What is missing

▶ **Programmers (and Operating Systems) have to deal with that.**

- ▶ Windows HPC Server 2008 is missing memory page migration functionality.



- ▶ **I really love that Microsoft's parallel programming offerings are similar in spirit and functionality for both the Native and the Managed languages!**
- ▶ **Microsoft (and Intel) are providing plenty of choices of paradigms addressing multicore:**
 - ▶ New Functionality?
- ▶ **There are many legacy codes still *alive* and the support of MPI with OpenMP is crucial for many user communities**
 - ▶ OpenMP 3.0 support is missing in Visual Studio 2010
- ▶ **Visual Studio 2010 provides exceptional performance analysis facilities for the node → MPI-support for the cluster has to follow.**

▶ **User Group events**

- ▶ Involvement of Microsoft experts from the US
- ▶ HPC hardware vendors sometimes have a *Beat the Vendor* session where users complain and developers / engineers listen

▶ **Microsoft**

- ▶ Presence on Pre-SC activities of HPC vendors?

▶ **Please care about the *Missing Middle in HPC ...***

Summary

- ▶ **The German Windows HPC User Group started in 2008 and provides a platform for users and vendors to address issues and exchange ideas.**
- ▶ **Windows HPC Server enabled HPC off the beaten track as well as it supports traditional MPI-based workloads.**
- ▶ **HPC is moving from an expert-only tool to become a regular part of the scientific workflow. Cloud Computing facilitates this trend. The connectivity of the Microsoft solution clearly helps.**
- ▶ **Buying machines is simple. Programming is much harder.**

The End

Thank you for your attention.

Website of the German Windows HPC User Group:
<http://www.rz.rwth-aachen.de/winhpcug>