Windows-HPC @ RWTH Aachen University

Christian Terboven
terboven@rz.rwth-aachen.de

Center for Computing and Communication
RWTH Aachen University

WinHP³C

Center for Computing and Communication

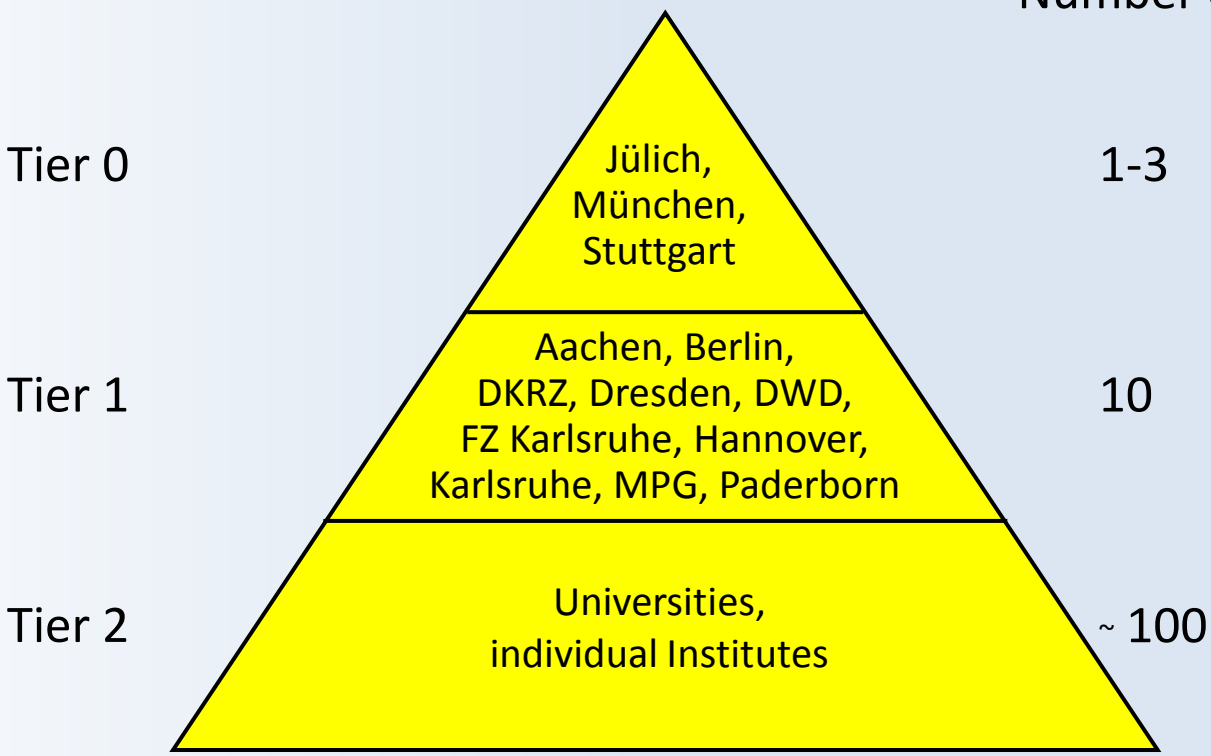
March 3rd, Redmond, WA, USA

Agenda

- Motivation
- HPC on Windows: Case Studies
- Challenges: Outlook
- Summary

HPC-Pyramid

- Proposal of the *HPC @ DE* task force:



- 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.*

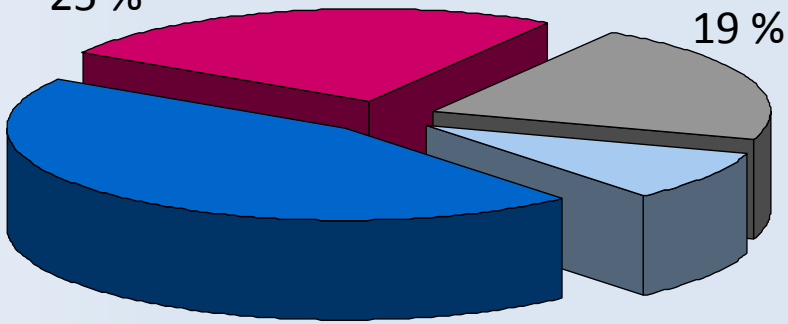
RWTH Aachen University (2007)

30260 students

450 professorships

4000 academic staff

280 institutes



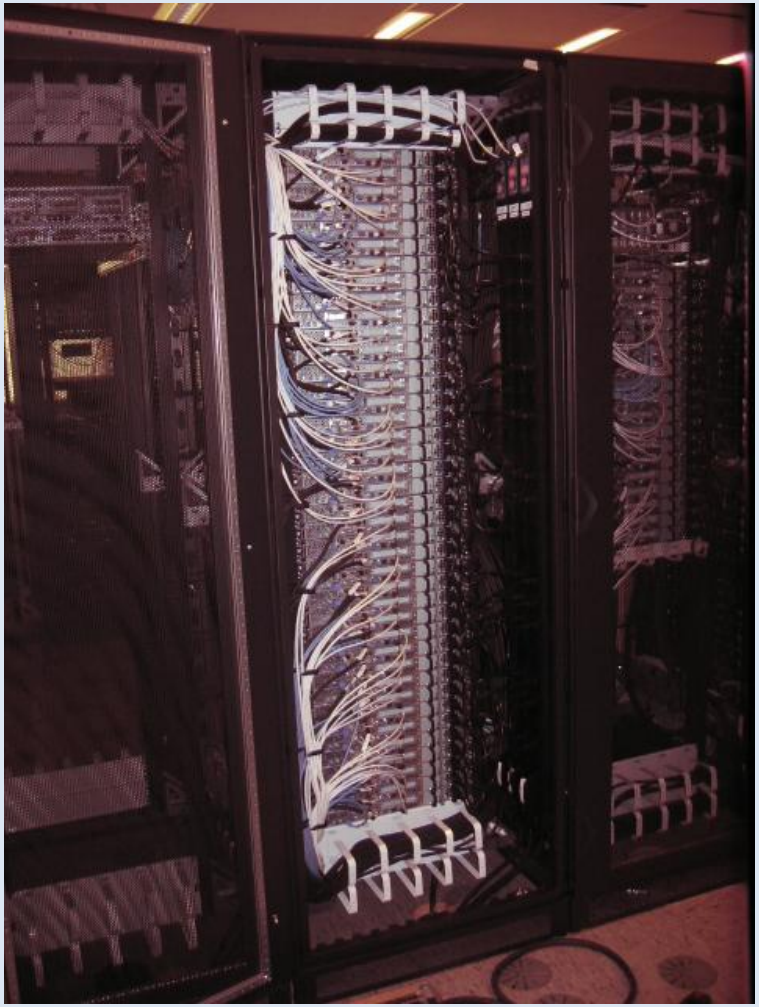
HPC on Windows @ RWTH Aachen

- We do HPC on Unix for many years - why try Windows?
  - Third party cooperations sometimes depend on Windows
  - Some users look out for the Desktop-like HPC experience
  - Windows is a great development platform
- 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)
- Cooperation with Microsoft Germany since 2008:

  WinHP3C = Windows High Performance Cluster Competence Center
  - Participation in CAB and TAP programs

Top500 run (LINPACK) on Windows

- Cluster installed in Q1/2008:
  - Fujitsu-Siemens
    Primergy RS 200 S4 Server
    - 2x Intel Xeon 5450
      (quad-core, 3.0 GHz)
    - 16 / 32 GB memory per node
    - 4x DDR InfiniBand:
      MPI latency: ~5 us
      MPI bandwidth: 1250 MB/s
- LINPACK-Tuning in Excel ☺
- Rank 100 in Top500 Liste in June 2008: 18,81 Tflop/s.
  - Windows HPC Server 2008 beta



WinHP3C activities in 2008

- Windows-HPC activities as part of the WinHP3C are open (at no cost) for external participants:
  - 04/2008: Windows-HPC user group meeting in Germany
    - 1.5 days, 110 participants from Industry and Academia
  - 04/2008: Windows HPC Deep Dive Training in Aachen
    - 3 days, 20 invited participants
  - 07/2008: Windows-HPC 2008 Workshop
    - 2 days, 55 participants
    - Opportunity to tune and parallelize your own codes!
  - 09/2008: Windows-HPC 2008 Workshop in english
  - 09/2008: Windows HPC Server 2008 Deployment Workshop
- Materials (Slides + Videos) are available online.

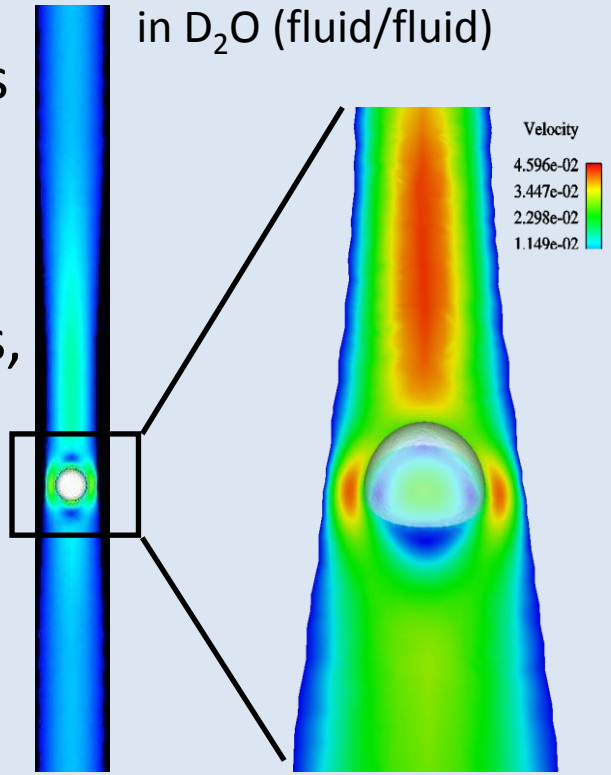
Agenda

- HPC in Germany & Aachen
- HPC on Windows: Case Studies
  - DROPS: C++, Hybrid (MPI + OpenMP)
  - KegelToleranzen: Fortan90, OpenMP
- Challenges: Outlook
- Summary

DROPS: A Navier-Stokes Solver in C++ (1/4)

- 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 D_2O (fluid/fluid)



DROPS: A Navier-Stokes Solver in C++ (2/4)

```
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 = rho / (p*q);
    x += alpha * p;
    r -= alpha * q;
    […]
```

```
y_Ax_par(&q.raw()[0],
   A.num_rows(), A.raw_val(),
   A.raw_row(), A.raw_col(),
   Addr( p.raw()));
```

```
#pragma omp for reduction
                   (+:alpha_sum)
for (long j=0; j<n; j++)
   alpha_sum += p[j]*q[j];

#pragma omp single {
   alpha = rho/alpha_sum;
}
```

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

- This programming style asks for a solid integration of compiler and debugger!

DROPS: A Navier-Stokes Solver in C++ (3/4)

```
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
    […]
    q = A
    double
    x += a
    r -= a
    […]
```

```
y_Ax_par(&q.raw()[0],
   A.num_rows(), A.raw_val(),
   A.raw_row(), A.raw_col(),
   Addr( p.raw()));
```

```
p for reduction
        (+:alpha_sum)
j=0; j<n; j++)
um += p[j]*q[j];

p single {
 rho/alpha_sum;
```

Parallelization is often a hard task!

We need IDE-support for (at least):
- Design + Development
- Parallelization
- Parallel Debugging
- Parallel Performance Analysis

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

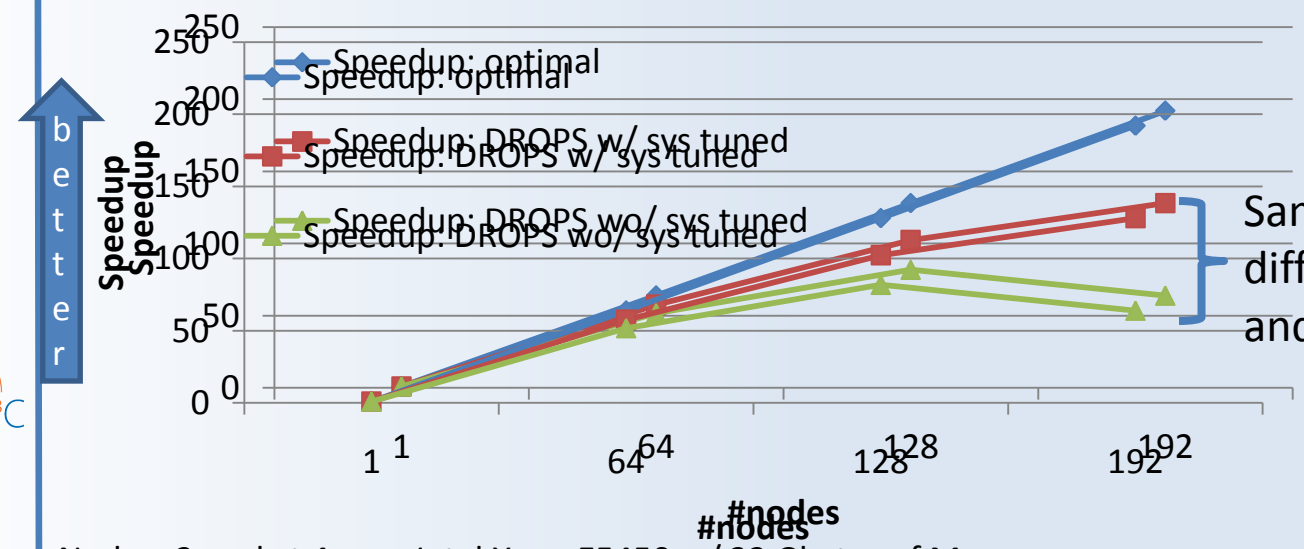
- This programming style asks for a solid integration of compiler and debugger!

DROPS: A Navier-Stokes Solver in C++ (4/4)

- 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



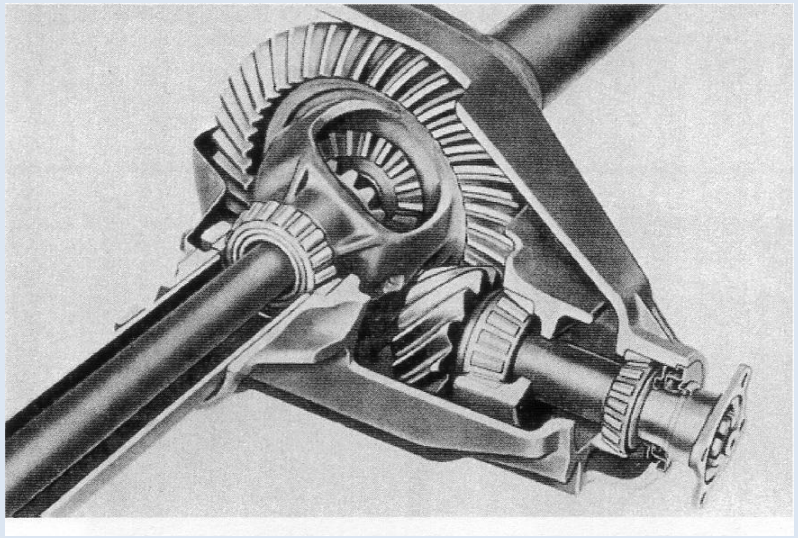
Node = 2-socket 4-core Intel Xeon E5450 w/ 32 Gbytes of Memory.

Case Study: KegelToleranzen (1/5)

- 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



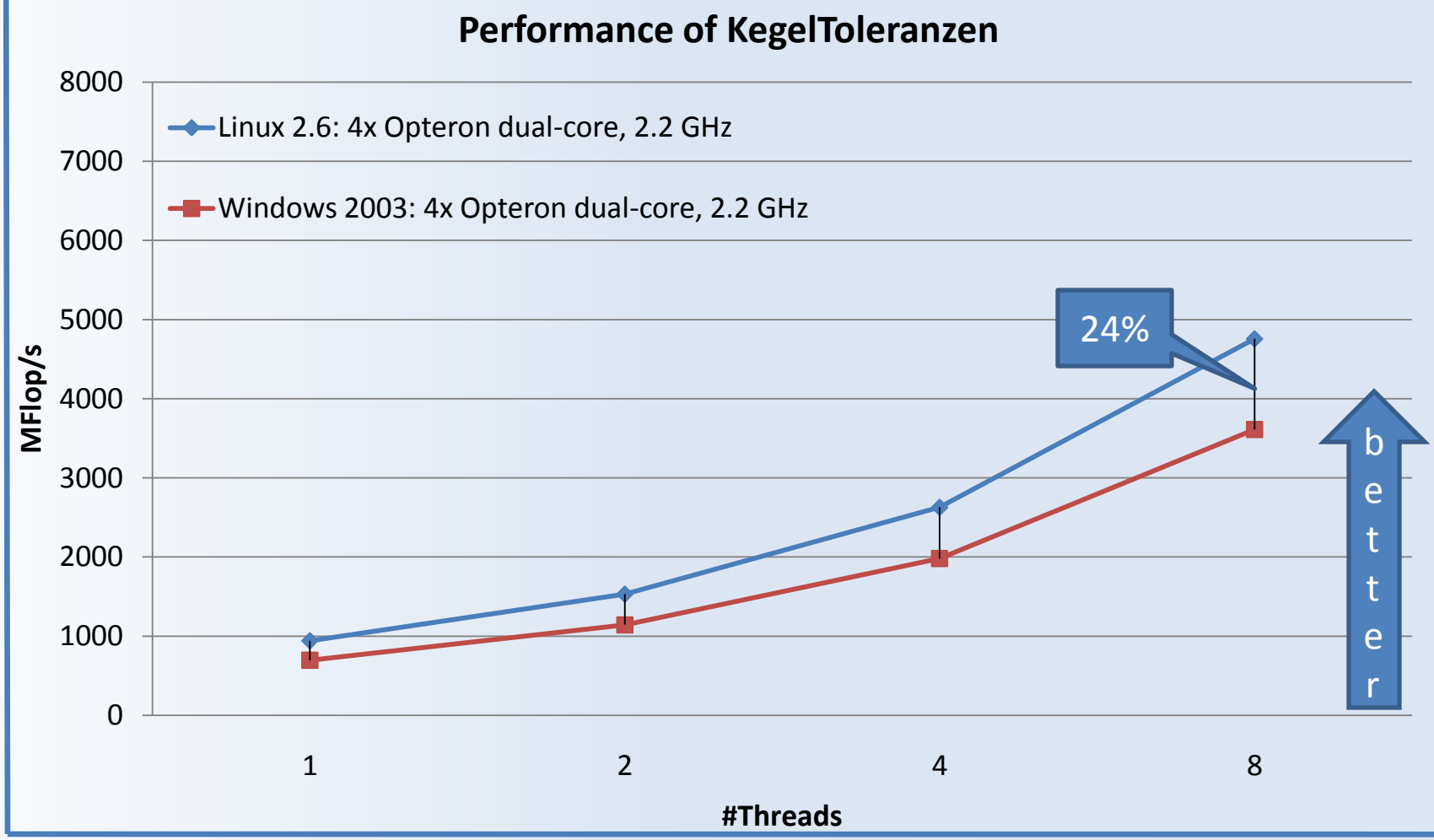
Laboratory for Machine Tools and Production Engineering, RWTH Aachen

Case Study: KegelToleranzen (2/5)

- Target
  - Pentium/Windows/Intel → Xeon/Windows/Intel
  - Serial Tuning + Parallelization with OpenMP
- Tuning and Parallelization Procedure
  - Get the tools: Porting to UltraSparc IV/Solaris/Sun Studio
  - Simulog Foresys: Convert to Fortran 90
    - 77000 Fortran77 lines → 91000 Fortran 90 lines
  - Sun Analyzer: Runtime Analysis with different datasets
    - Deduce targets for Serial Tuning and OpenMP Parallelization
  - OpenMP Parallelization: 5 Parallel Regions, 70 Directives
  - Get the tools: Porting new code to Xeon/Linux/Intel
  - Intel Thread Checker: Verification of OpenMP Parallelization
- Put new code in production on Xeon/Windows/Intel

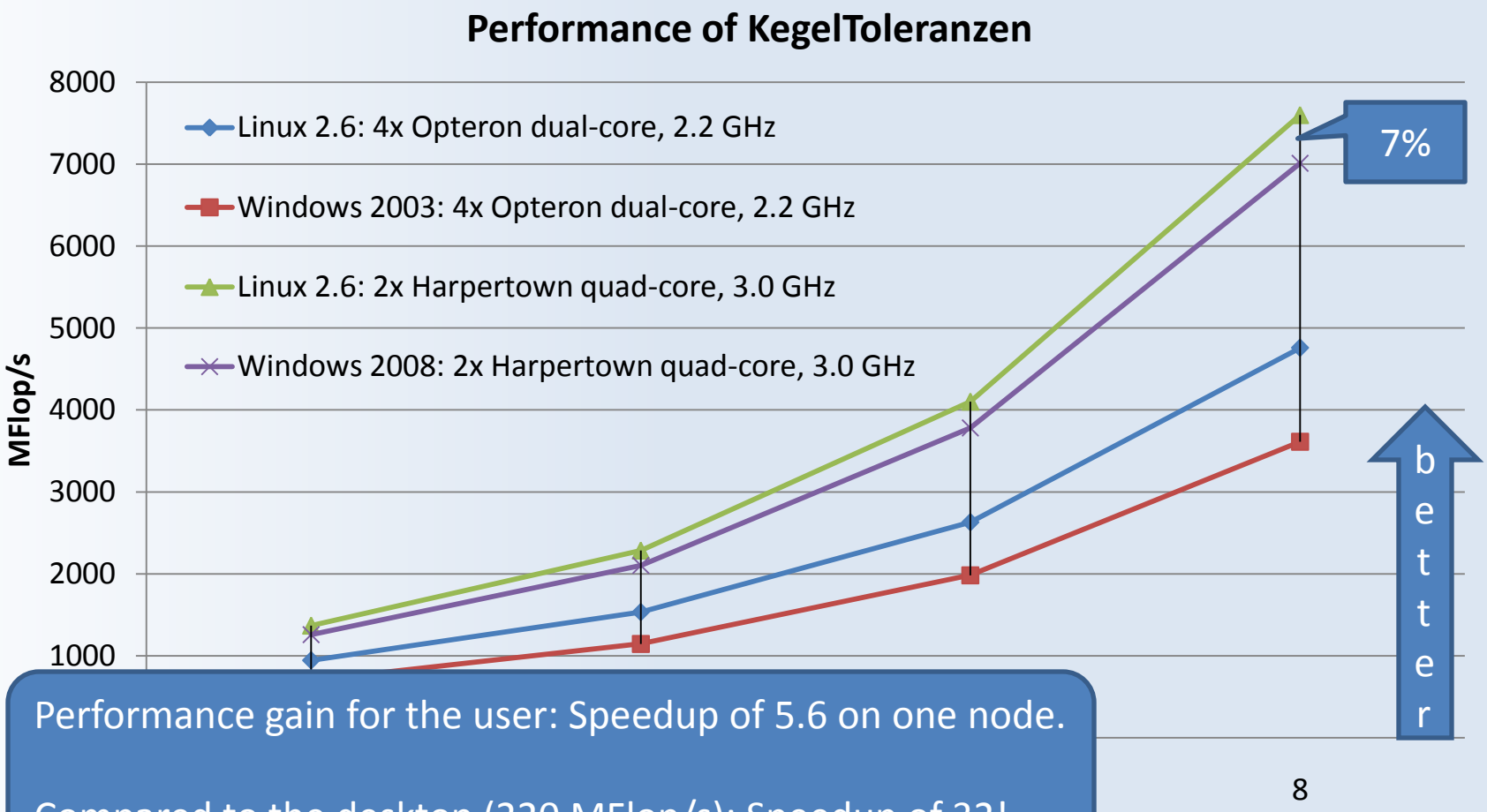
Case Study: KegelToleranzen (3/5)

- Comparing Linux and Windows Server **2003**:



Case Study: KegelToleranzen (4/5)

- Comparing Linux and Windows Server **2008**:

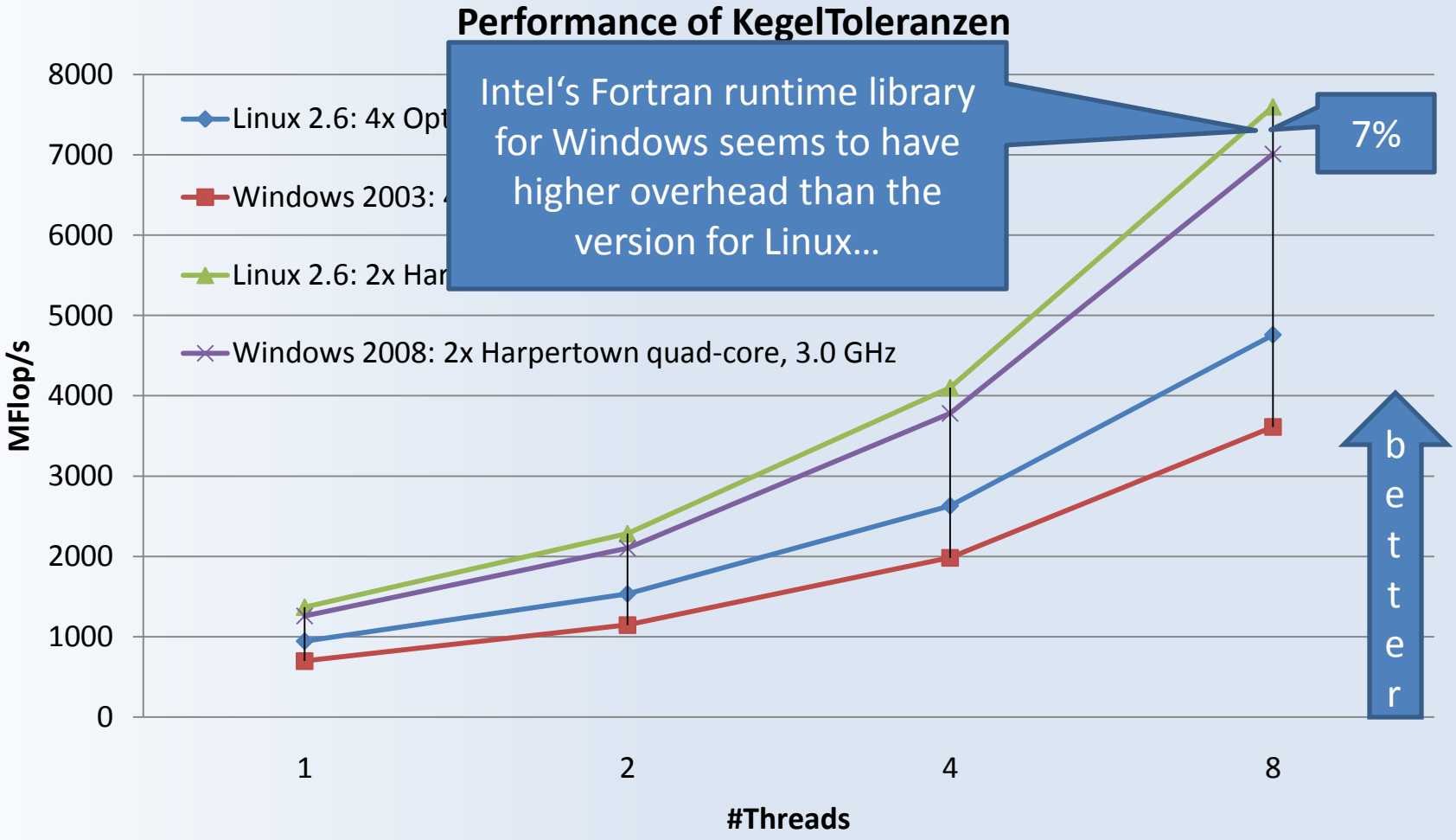


Performance gain for the user: Speedup of 5.6 on one node.

Compared to the desktop (220 MFlop/s): Speedup of 32!

Case Study: KegelToleranzen (5/5)

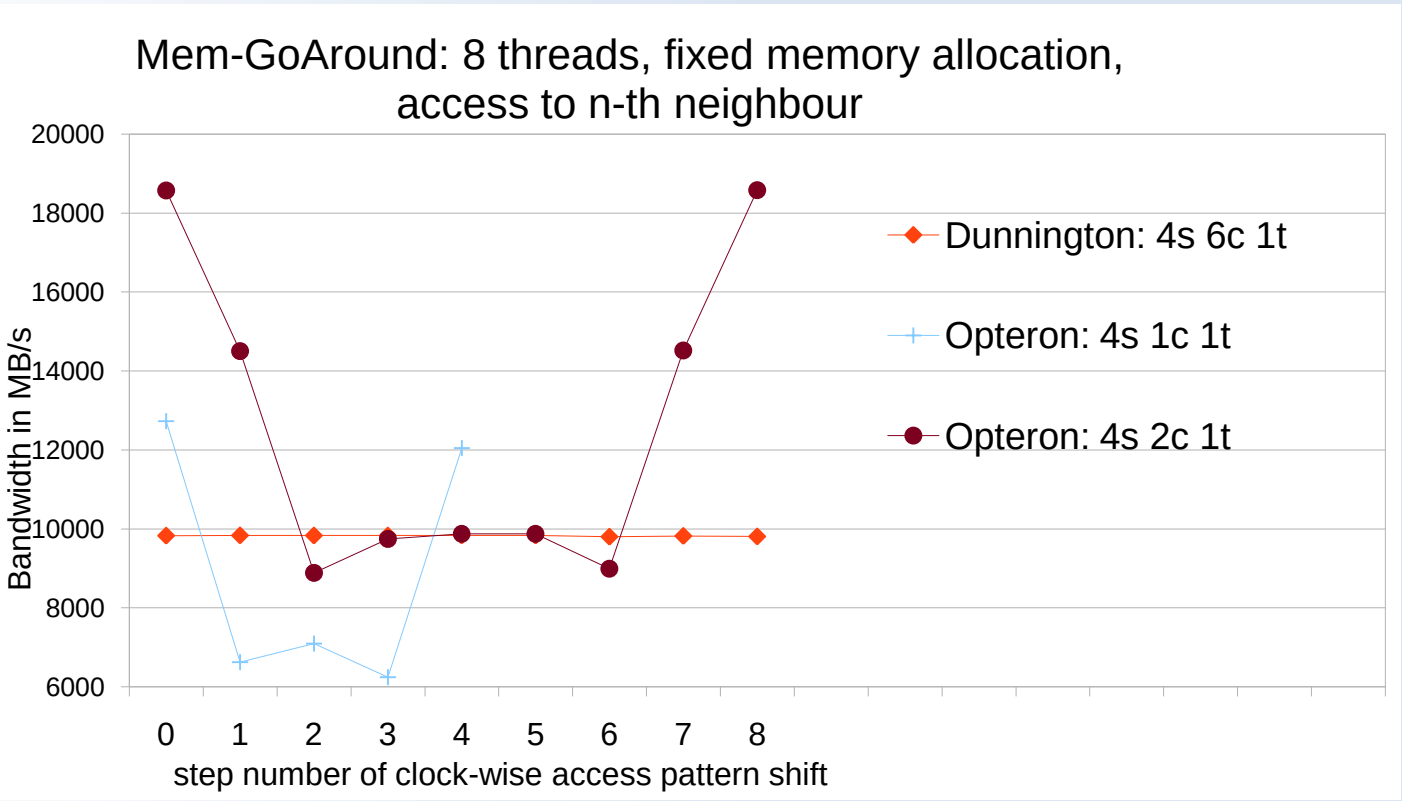
- Comparing Linux and Windows Server **2008**:



Agenda

- HPC in Germany & Aachen
- HPC on Windows: Case Studies
- Challenges: Outlook
- Summary

Shared-Memory systems go cc-NUMA!



- Programmers (and Operating Systems) have to deal with that.
  - Windows HPC Server 2008 is cc-NUMA aware to some extent,
  - but is still missing memory page migration functionality.

The Crystal Ball is showing many, many cores

- SUN Microsystems won our recent HPC procurement:
  - 2008: Dunnington-based systems with
    - 4s 6c 1t → 24 threads per node
  - 2009/2010: Nehalem-based systems with
    - 8s 8c 2t → 128 threads per node
  - In total about 220 TFLOP/s (limited by energy cost budget)

Buying is simple.

Programming is much harder.

- .NET is nice, but MSFT should not forget about native code:
  - We hope for soon-to-be-found OpenMP 3.0 support!
- Today's severest issue of Windows-HPC in mixed env.: I/O
  - How far does CIFS scale (i.e. delivered by NetApp)?
  - We need clients for parallel filesystems, such as Lustre!
    Lustre: Intended cooperation between SUN and RWTH Aachen.

Agenda

- HPC in Germany & Aachen
- HPC on Windows: Case Studies
- Challenges: Outlook
- Summary

Summary

- RWTH: *The simulation of complex models lies at the core of innovative products and basic research at RWTH Aachen.*
  - Simulation Science (HPC) is main pillar of Research.
- Windows HPC Server 2008 has proven to be a solid solution for HPC on Windows:
  - We experience a growing interest in Industry and Academia.
- The future will bring many, many cores plus GPUs
  - HPC market happily lives at the forefront of technology.
  - Bringing HPC to the mass market means providing an integrated solution to program and manage HPC systems.

The End

Thank you for

your attention!

WinHP3C:

http://www.rz.rwth-aachen.de/winhp3c

2nd German Windows-HPC User Group Meeting:

http://www.rz.rwth-aachen.de/go/id/siu

March 30.-31. 2009, Dresden, Germany