CHALLENGES AND OPPORTUNITIES OF HYBRID COMPUTING SYSTEMS

S. Roller[1], H. Klimach[1], S. Wesner[1]

[1] High Performance Computing Center Stuttgart (HLRS)
Universität Stuttgart
70550 Stuttgart, GERMANY
{roller, klimach, wesner}@hlrs.de

Future computing systems and their usage in software development are the big challenges for the next future. We observed a growing heterogeneity in both, hardware and software, in the past years. The idea is to make use of that by mapping each application or each module of an application to the appropriate piece of hardware. This talk shortly summarizes the hybrid prototype which is currently being installed at HLRS. With respect to future architectural systems, HLRS is currently investigating a hybrid computing systems approach aiming for a tight coupling of different architectures within one single system. The concept is currently evaluated in collaboration with NEC with an innovative PRACE prototype integrating vector processors, x86 based processors and General Purpose Graphical Processing Units (GPGPUs). The underlying question is: How can the application development keep track with the hardware development?

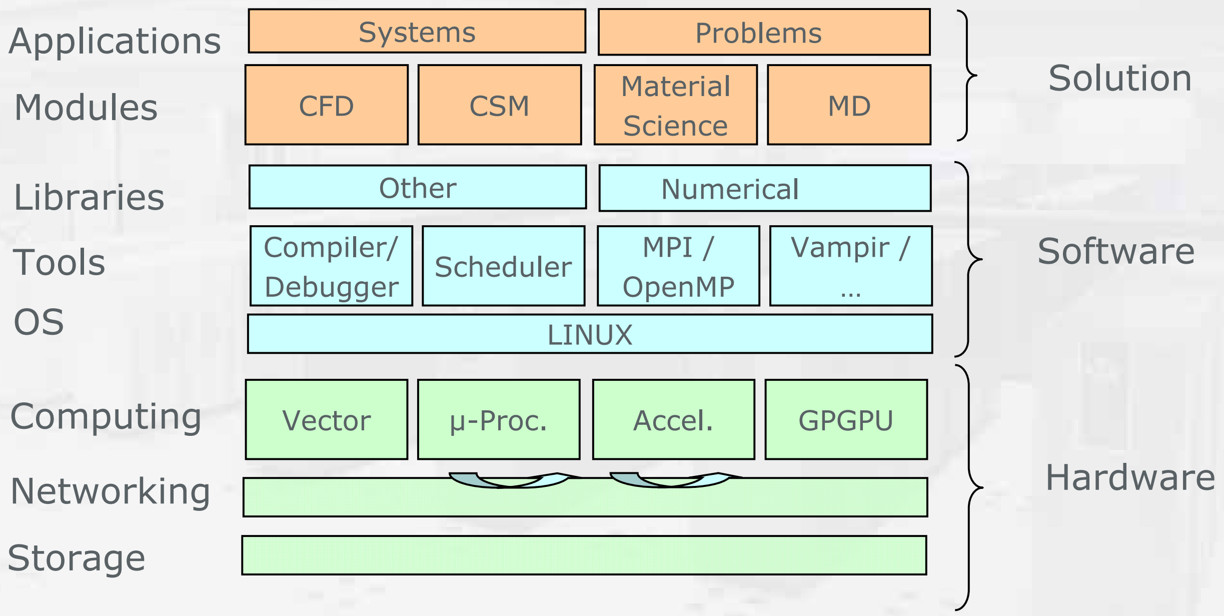
The main part of the talk will focus on the European and German context for HPC. The European Roadmap for Research Infrastructures (ESFRI) is the first comprehensive definition at the European level. Research Infrastructures are seen to be one of the crucial pillars of the European Research Area. A European HPC service foresees the impact on: strategic competitiveness, attractiveness for researchers, supporting industrial development. The HPC ecosystem in Europe follows a layered approach, which is supplemented by an again layered approach on the national level within Germany. The talk will focus on the political and strategic plans in Europe and Germany.

Challenges and opportunities of hybrid computing systems

Dr.-Ing. Sabine Roller, Dipl.-Ing. Harald Klimach, Dr.-Ing. Stefan Wesner
High Performance Computing Centre Stuttgart
http://www.hlrs.de

THE HLRS VIEW ON FUTURE COMPUTING SYSTEMS

The HLRS Hybrid Prototype (Conceptual View)



HLRS Hybrid Prototype Facts

- Unique "System of Systems" concept
  - Multi-physics / multi-scale apps on optimized hardware
  - Hybrid configuration: Vector (SX9) + Scalar (Nehalem) + GPU (Tesla)
  - Highly innovative configuration
  - Expandable (e.g. with Cell, GPU, FPGA, …)
  - Shared filesystem and heterogeneous network

NEC SX-9 vector part
12 nodes
192 processors

19.2 TF

X86-64 scalar part
Dual Socket Intel Nehalem
700 nodes
5600 cores

62,7 TF

GPGPU part
16 units
64 accelerator cards
15360 cores

64 TF

HLRS Hybrid Prototype Facts

- Concept enables industry-related applications
- Allow for continuous system update using an open architecture integrating new and updated architecture types continuously
- Energy efficient
  - Applications can use the best suited architecture for their problem, even distinct for distinct parts
    - No waste of CPU time and no waste of energy
  - Vectorization as programming model is energy efficient
- Partially funded by PRACE
- Availability January 2009 – April 2009

NEC SX-9 vector part 12 nodes 192 processors 19.2 TF
X86-64 scalar part Dual Socket Intel Nehalem 700 nodes 5600 cores 62,7 TF
GPGPU part 16 units 64 accelerator cards 15360 cores 64 TF

Challenges and opportunities

- System is unique in the world
- Coupled multi-physics, multi-scale simulations increasingly important
  - Simultaneous simulation of several aspects at once (e.g. Fluid-Structure, Aero-Acoustics, Combustion Procs)
  - Different aspects require different architectures for optimal use → Hybrid systems
- Examples:
  - Optimization of power plants
  - Emission reduction in production
  - Climate ocean-atmosphere
  - Medical apps / surgery planning

Challenges and opportunities

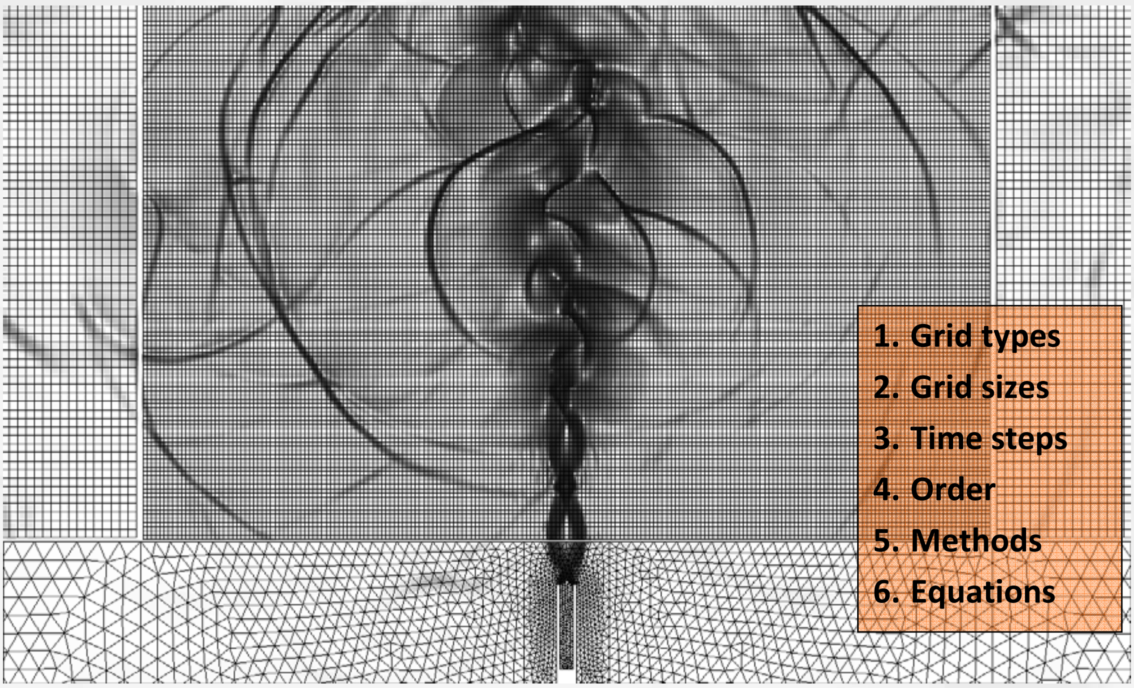
- Hybrid systems imposed challenges
  - Alignment of operating systems
  - Integrated scheduling systems
  - Heterogeneous networks
  - Hybrid compilers & analysis tools
- Programming models/ specific benchmarks for future integrated hybrid systems

DETAILED EXAMPLE MOTIVATING THE HYBRID APPROACH

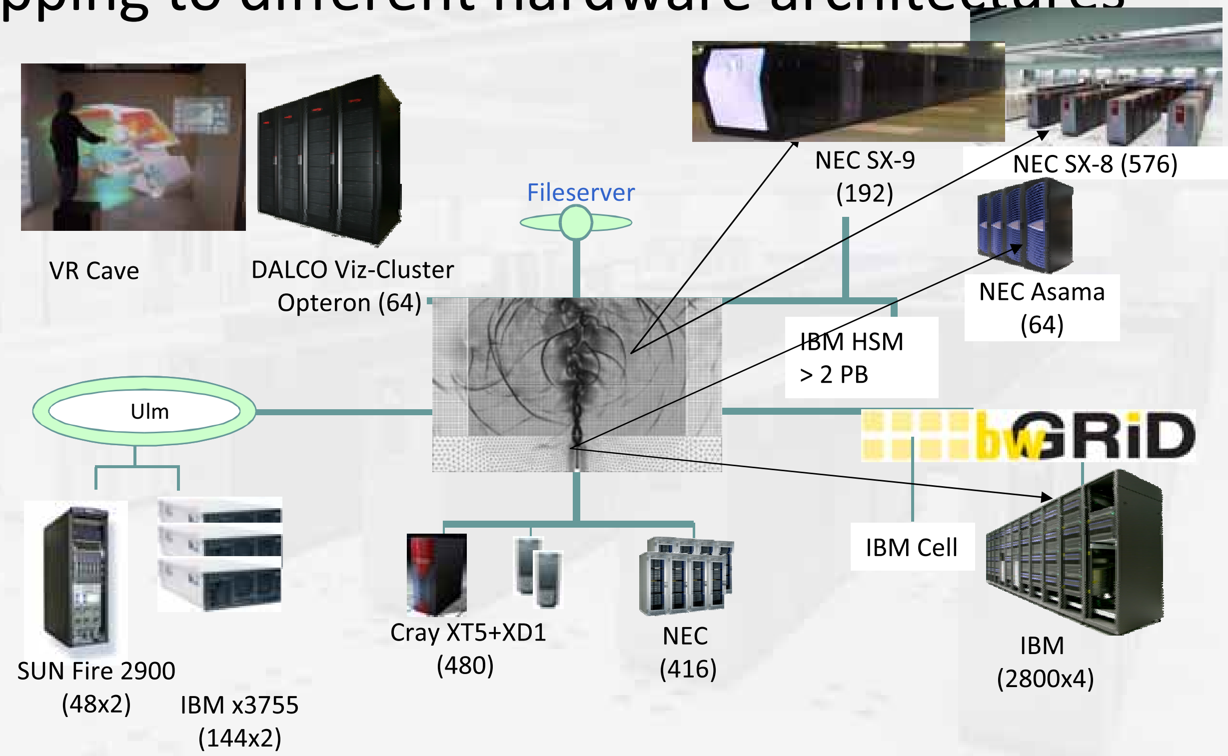
Motivation / Example

- Coupled applications: multi-scale, multi-physics
- Aero-acoustics: different scales between flow and acoustics
- Different approximations:
  - Flow: fine grid, unstructured, small time step, DG, non-linear Navier-Stokes
  - Acoustics: coarser grid, structured, larger time step, FV / FD, non-linear / linearized Euler equations
- Idea: Reduce the computational effort by adapting the
  - numerical methods (DG,FV,FD)
  - grids (structured, unstructured)
  - equations (N.S., Euler, LEE)
  - time steps (explicit method)
- Heterogeneous domain decomposition

Example: Nozzle flow with a free jet



Mapping to different hardware architectures



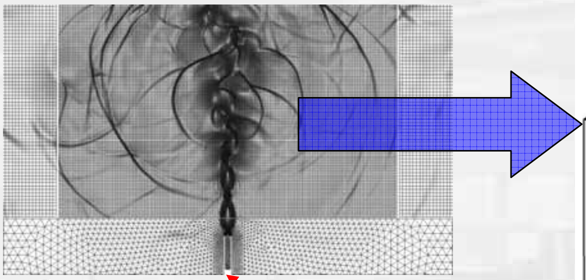
Performance and costs

	1xIA64	1xSX	1xSX + 1xIA64
UNSTRUCT:	2.993,67	**7.745,82**	3.019,24
STRUCT:	**23.887,32**	2.870,66	2.869,46
KOP:	1.012,37	321,15	554,21
waiting:	0,00	0,00	164,20
KOP calculating time:	1.012,37	321,15	390,02
Total CPU time:	27.893,35	10.937,62	6.278,72
Total elapsed (sec):	27.924,78	10.966,23	3.207,09
Total elapsed (h):	7:45'	3:03'	0:53'
Relative Price	1	1,57	0,58

Coupled simulation 4x faster and 40% cheaper !

Combined strength of both architectures by coupling with PACX-MPI
Equivalent to 8.7 IA64 CPUs or 3.42 NEC-SX8 CPUs

Striving for more



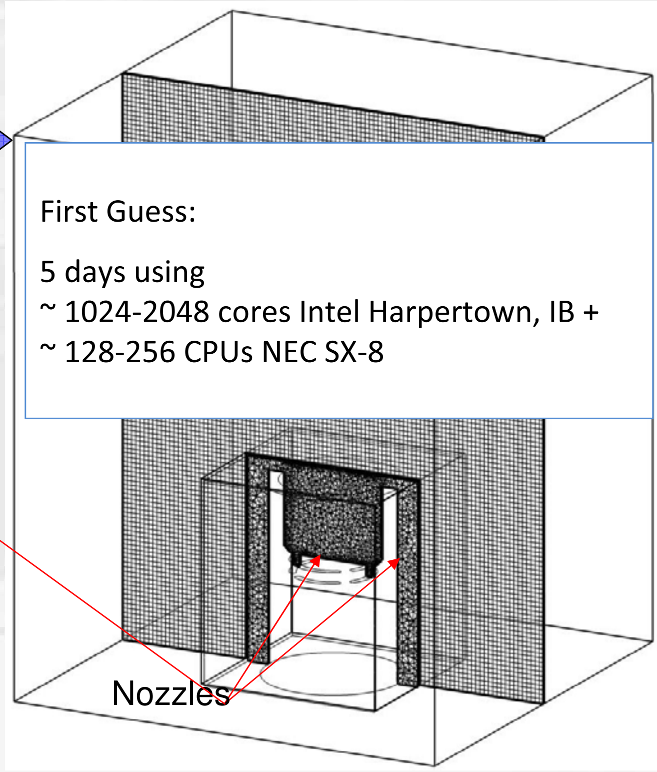
Nozzle outlet: 4 x 0.15 mm

Unstructured Domain:
8x8x10 mm³: **15 Million Elements**

Structured Domain:
2x2x5 cm³: **1 Billion Cells**

Far Field:
1x1x1 m³: **350 000 Cells**

M = 1.4; Re = 17000 – 30000



First Guess:

5 days using
~ 1024-2048 cores Intel Harpertown, IB +
~ 128-256 CPUs NEC SX-8

THE PROTOTYPE IN THE EUROPEAN CONTEXT

Computational science infrastructure in Europe

The European Roadmap for Research Infrastructures (ESFRI) is the first comprehensive definition at the European level

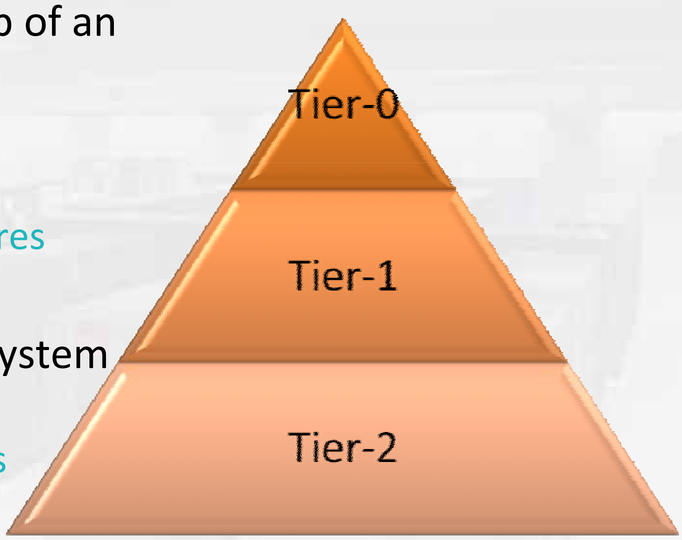
Research Infrastructures are one of the crucial pillars of the European Research Area

A European HPC service – impact foreseen:

- strategic competitiveness
- attractiveness for researchers
- supporting industrial development

The ESFRI Vision for a European HPC service

- European HPC-facilities at the top of an HPC provisioning pyramid
  - Tier-0: 3-5 European Centres
  - Tier-1: National Centres
  - Tier-2: Regional/University Centres
- Creation of a European HPC ecosystem involving all stakeholders
  - HPC service providers on all tiers
  - Grid Infrastructures
  - Scientific and industrial user communities
  - The European HPC hard- and software industry



ESFRI – Estimated costs

- Unlike other European Research Infrastructures:
  - Tier-0 resources have to be renewed every 2-3 years
  - Construction cost 200 – 400 Mio. € every 2-3 years
  - Annual running cost 100 – 200 Mio. €
- A truly European challenge – also in terms of funding
- PRACE – The Partnership for Advanced Computing in Europe
  - An Initiative created to implement the ESFRI vision of a European HPC service

PRACE – Project Facts

- Objectives of the PRACE Project:
  - Prepare the contracts to establish the PRACE permanent Research Infrastructure as a single Legal Entity in 2010 including governance, funding, procurement, and usage strategies.
  - Perform the technical work to prepare operation of the Tier-0 systems in 2009/2010 including deployment and benchmarking of prototypes for Petaflops systems and porting, optimising, peta-scaling of applications
- Project facts:
  - Partners: 16 Legal Entities from 14 countries
  - Project duration: January 2008 – December 2009
  - Project budget: 20 M € , EC funding: 10 M €
- PRACE is funded in part by the EC under the FP7 Capacities programme grant agreement INFSO-RI-211528

CAPACITIES e-infrastructure

PRACE – Project Consortium

New Partners of the PRACE Initiative:

HPC Provisioning concept in Germany

National High Performance Computing Centres

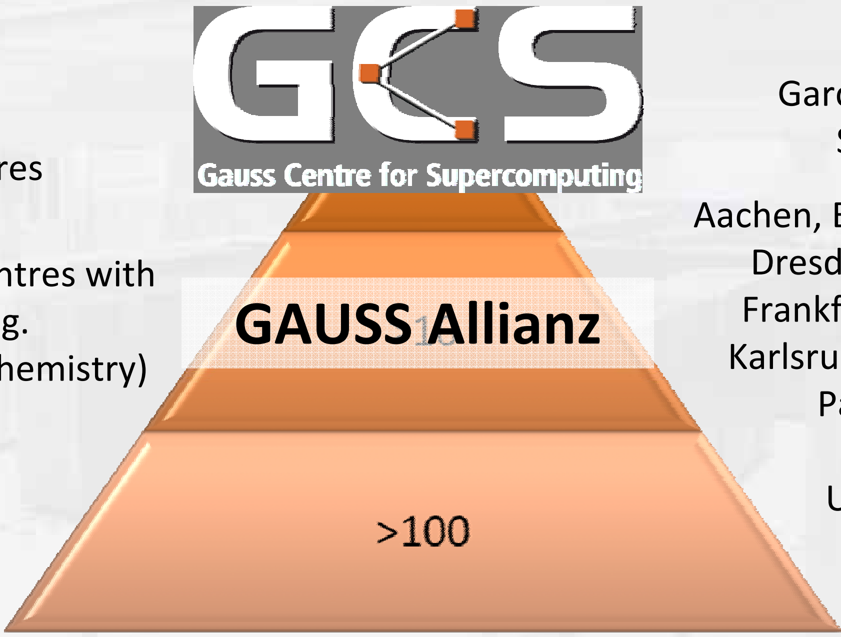
Garching, Jülich, Stuttgart

Regional HPC centres with domain focus (e.g. computational chemistry)

Aachen, Berlin, Hannover, Dresden, Erlangen, Frankfurt, Garching, Karlsruhe, Offenbach, Paderborn

HPC-Server

Universities



HLRS profile within GCS

- HLRS has since its inauguration a close connection to the engineering community and industrial users in particular from the automotive domain
- HLRS is performing research in many areas in order to offer advanced services to its customers
  - Programming Models and tools (e.g. PACX-MPI, MARMOT, …)
  - Technical and Scientific Visualization (COVISE)
  - Coupled Applications and Scalable Computing
  - Grid Computing
- Close relationship to the Mechanical Engineering faculty in Stuttgart (via the Institute for High Performance Computing)

Role of Prototypes in PRACE

- The HLRS hybrid prototype is one of the 6 PRACE prototypes
- Prototypes will be used to evaluate the architectures in near-production situation with regard to
  - application performance and scalability
  - total cost of ownership
  - energy consumption
- Prototypes will enable
  - evaluation of software for managing distributed infrastructures
  - preparation of benchmarks for future Petascale systems
  - Progress in scaling and optimization of libraries and codes
  - Definition of technical requirements and procurement procedures for the PRACE Petaflop/s production systems for 2009/2010.

RELATED RUNNING AND PLANNED RESEARCH ACTIVITIES

New architectures impose new research challenges

- New Programming Models?
  - How to efficiently use the new architectures?
  - How to achieve an acceptable development time?
- How to cope with the increasingly complex and large computing infrastructure?
  - Towards knowledge based and policy driven management concepts
  - More intelligent model driven solutions for the management of large systems

New architectures impose new research challenges

- Address new user communities
  - More open for new programming languages (UPC, CAF, Chapel, X10, CUDA, …)
  - Multi-/Manycore widened the user community significantly beyond "MPI experts"
- Faster development cycle
  - Speed up the development of users from tier-2 towards tier-0
  - Reduce time for industrial take up of tier-0 methods and solutions

Related Research Programmes

German National HPC Software Iniative

- STEDG
  - High order methods for turbulent flow simulation
- SKALB
  - Highly scalable Lattice-Boltzmann code
- IMEMO
  - Molekular simulations in material science and bio/nanotechnologie
- VISPME
  - a flexible, parallel and scalable integration environment for the interactive visualization of very large data sets
- ASIL
  - Advanced Solvers for very large equation systems
- TIMaCS
  - Management tools and methods for very large computing systems

Related Research Programmes

European Programmes

- INFRA
  - Make PRACE a sustainable activity from project to a Research Infrastructure legal entity
  - DEISA2
    - Distributed HPC infrastructure operation
  - HPC-EUROPA2
    - Training and Exchange program for new user communities
- ICT
  - Computing Systems
    - Models and approaches for multi-/manycore from embedded systems up to HPC
  - ICT for energy efficiency
    - energy efficient service provision
- FET (Future, Emerging Technologies)
  - Tera-Device Computing
    - Service Oriented Operating Systems
    - New Programming Models

Thanks! Any questions?

Dr. Sabine Roller,
Dipl.-Ing. Harald Klimach,
Dr. Stefan Wesner
High Performance Computing Centre Stuttgart
roller@hlrs.de

Contains material produced by Harald Klimach (HLRS) and Claus-Dieter Munz, Jens Utzmann Institute for Aerodynamics and Gasdynamics (IAG)