Karlsruhe Institute of Technology

A short Overview of KIT and SCC and the Future of Grid Computing

Frank.Schmitz@kit.edu

Agenda

- Introduction
  - KIT and SCC, organisation and resources
  - HPC for the state Baden-Württemberg
  - Project CampusGrid
- Software
- Grid computing and scientists?
- I/O-Performance and benchmark results
- Usage and resources
- European projects
- D-Grid and C^3
- Cloud computing
- Conclusion

Karlsruhe Institute of Technology (KIT)

KIT...

University

11	fakulties
120	institutes
4,000	employees
18,500	students
250 mio €	budget

Forschungszentrum

10	programs (research areas)
21	large institutes
4,000	employees
310 mio €	budget

science

education

structures for the administration

institutes

services

infrastructure

KIT / SCC

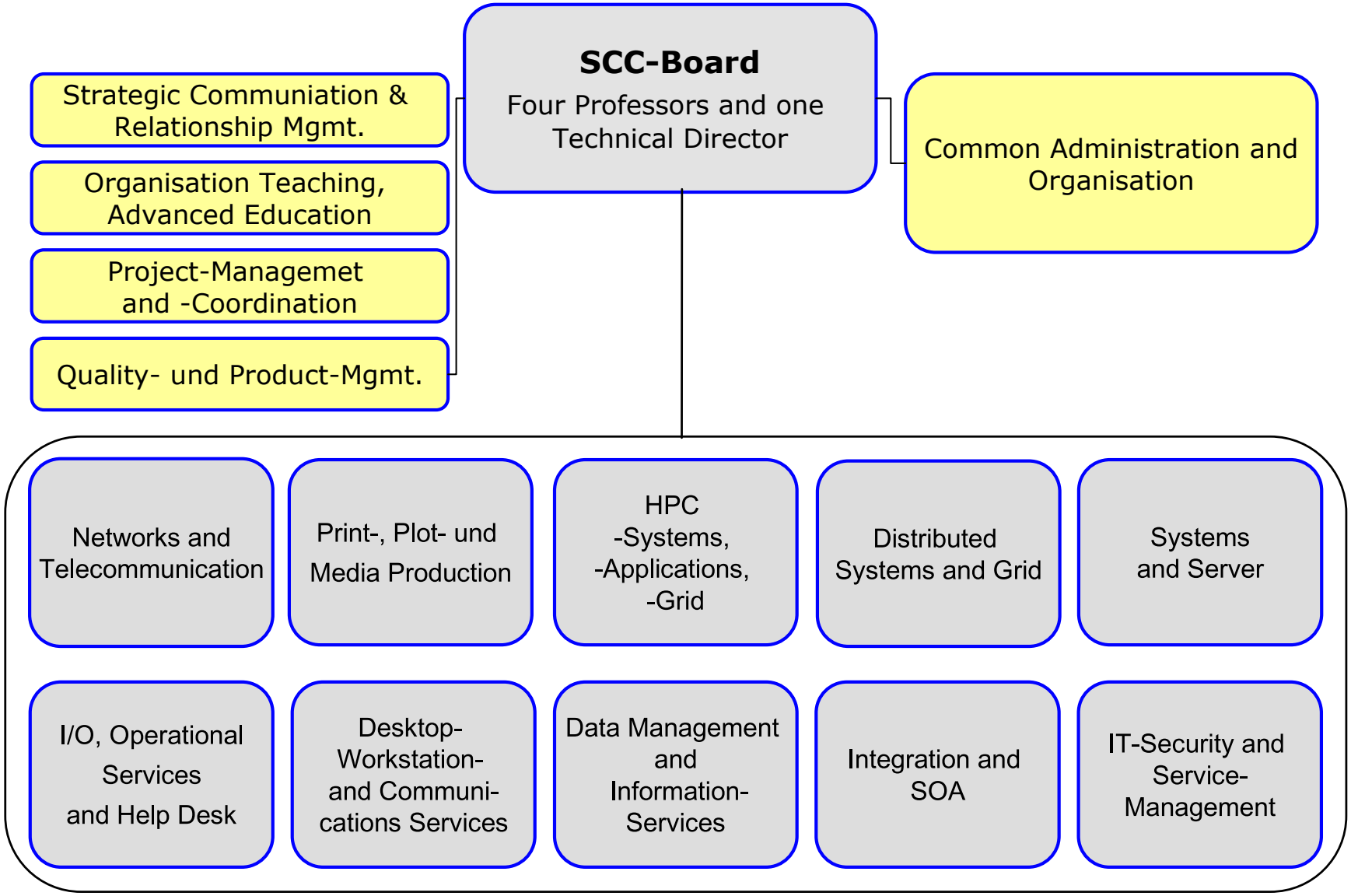
- **Karlsruhe Institute of Technology - two strong partners united**
  - Universität Karlsruhe (TH)
    - One of the most research oriented universities in Germany
    - Nominated to one of three top universities in October 2006
    - Capitalized with over 100 million € over 5 years
  - Forschungzentrum Karlsruhe
    - Member of the Helmholtz Association
    - Internationally acknowledged non-university institution
- **Under the roof of KIT – Steinbuch Centre for Computing**
  - Computing Center of the Universität Karlsruhe (TH)
  - Institute of Scientific Computing at the Forschungszentrum Karlsruhe

Virtual Computing Centre Karlsruhe since 1996

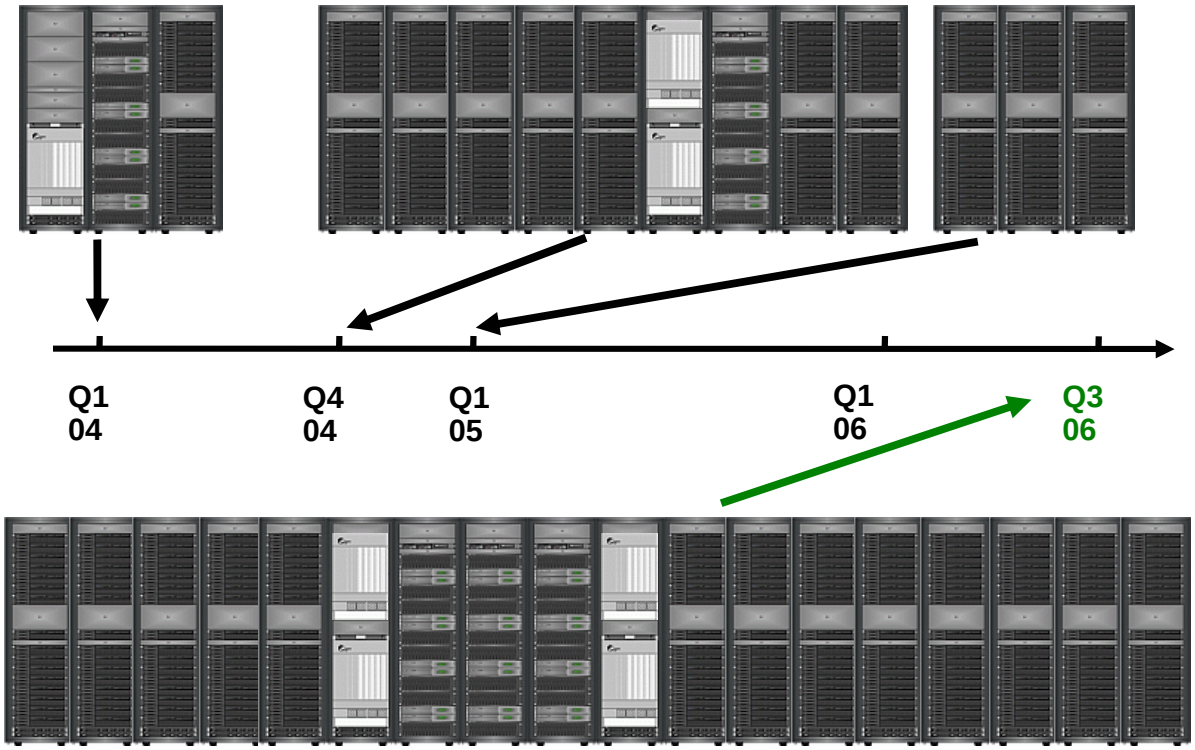
- IT-Services for UKA
- High Performance Computing
- Numerical Methods on HPC
- Integrated Information Mgmt.

- IT-Services for FZK
- Scientific Computing
- GridKa, CampusGrid, D-Grid
- Nat. & Intl. Grid Projects

10 SCC Departments



HPC platforms at SCC: HP XC6000 / XC4000



Phase 0 (Q1 2004)
- »12 2-way nodes (Itanium 2)
- »4 file server nodes
  - 2 TB shared storage
- »Single rail Quadrics interconnect

Phase 1 (Q4 2004)
- »108 2-way nodes
  - Intel Madison 6M, 1.5 Ghz, 12 GB
- »8 file server nodes
  - Approx. 11 TB storage system
- »Single rail Quadrics interconnect

Phase 1 (Q1 2005)
- »6 16-way nodes
  - Intel Madison 6M, 1.6 Ghz
  - 2 partions, each 8 CPUs, 64 GB
- »Single rail Quadrics interconnect

Phase 2:
- » Total of 3000 processor cores
- » Total of 15.6 TFlop/s peak performance
- » Total of 12 TB of main memory

Phase 2 (Q3 2006)
- »750 4-way nodes (two sockets)
  - dual core AMD Opteron 2.6 Ghz
  - 16 GB main memory
- »10 server nodes
- »10 file server nodes
  - Approx. 56 TB storage system
- »Infiniband DDR interconnect

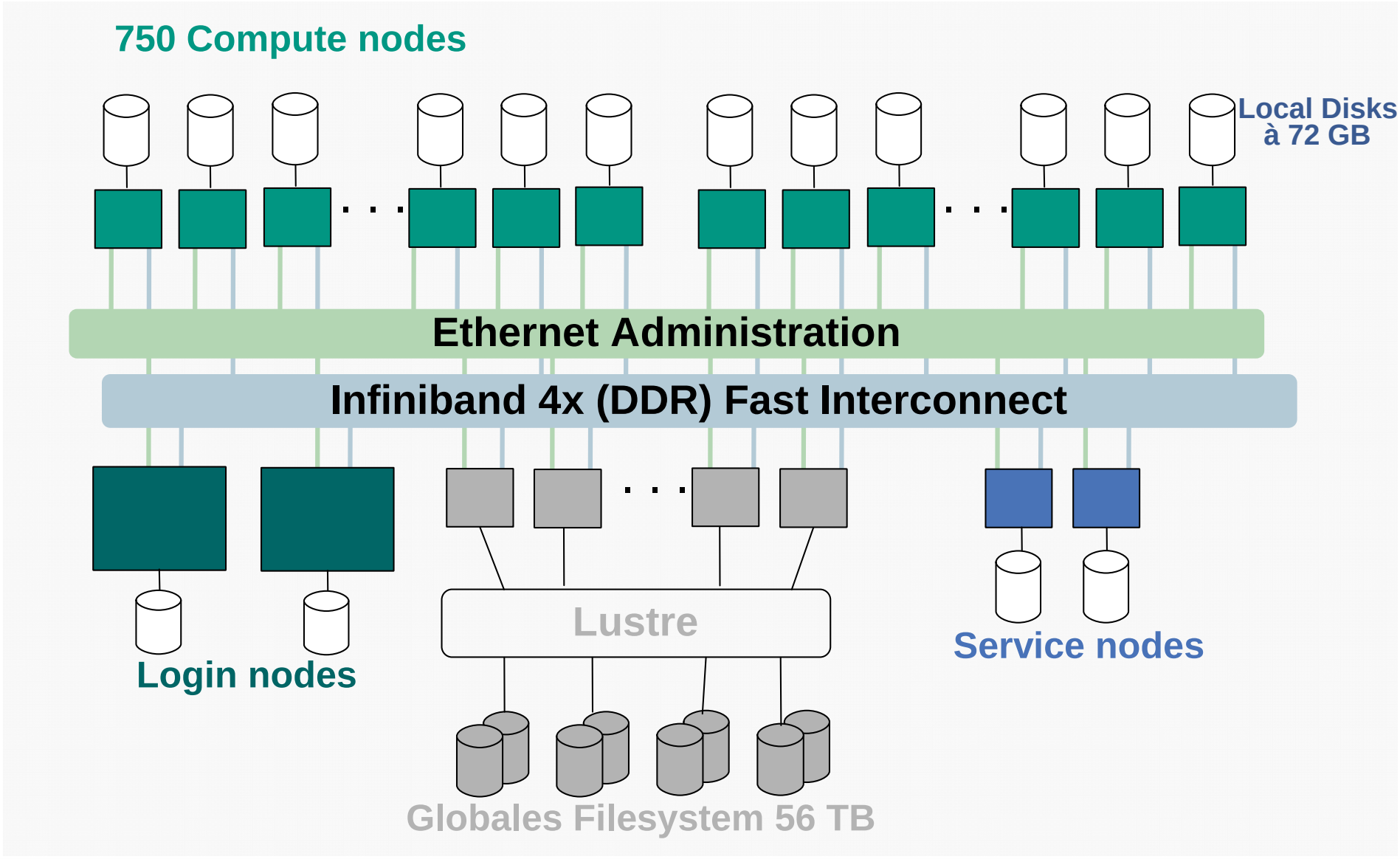
HP XC4000 – Baden-Württemberg-State-Supercomputer

Phase 2 (Q3 2006)

- »750 4-way nodes (two sockets)
  - dual core AMD Opteron 2.6 Ghz
  - 16 GB main memory
- »10 server nodes
- »10 file server nodes
  - Approx. 56 TB storage system
- »Infiniband DDR interconnect
- »20 water cooled racks (HP MCS)

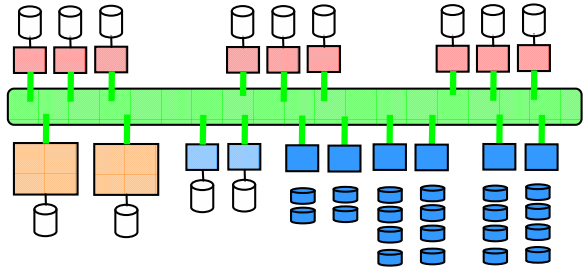
- » Total of **3000 processor cores**
- » Total of **15.6 TFlop/s peak** performance
- » Total of **12 TB** of main memory
- » Total **110 TB** of local disk space (scratch)

Configuration of the HP XC4000



Parallel File System HP SFS on XC4000

- Shared storage for all nodes of XC system
- 10 file server nodes
  - 2 MDS / Admin
  - 2 OSS for $HOME
  - 6 OSS for $WORK
- 56 TB file space
  - 8 TB $HOME
  - 48 TB $WORK
- Bandwidth
  - Read / write from one node: 340 MB/s / 340 MB/s
  - Total read / write bandwidth of $HOME: 600 MB/s / 360 MB/s
  - Total read / write bandwidth of $WORK: 3600 MB/s / 2200 MB/s



Motivation for the project CampusGrid

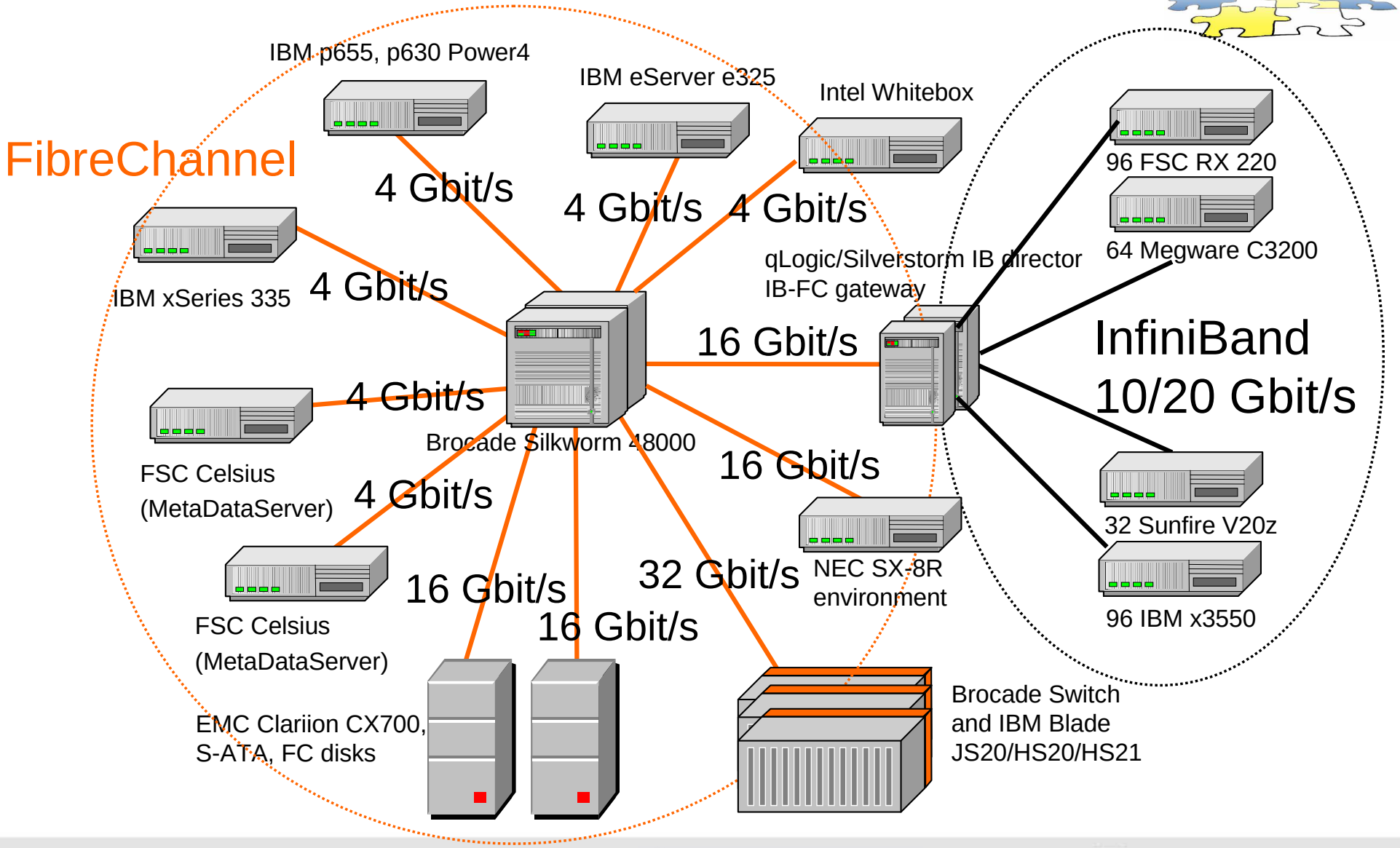
- heterogeneous IT-environment: vector-, SMP-, cluster-, blade-systems, SAN, NAS, Unix, Linux, Windows, Solaris, SuperUX, ….
- global view by the user
- only one user management (ADS from Microsoft)
- one job management
- metacomputing (MPI, ..), “real-time” applications
- access data for visualization at the local workstation
- global accounting
- seamless integration into different projects and middleware concepts like gLite, LCG, D-Grid, Unicore, GT4, ….

Hardware running in the project

• 1500 Intel and AMD based compute cores for the project CampusGrid (www.CampusGrid.de) and D-Grid (www.d-grid.de)

- 50 TB online storage, 10 TB mirrored storage
- 52 Power4 cores
- 16 vector processors each 32 GB main memory

Infrastructure of the project CampusGrid



Parts of the CampusGrid Project

a SX-9 node will be added in autumn 2008

New software

- ddt from Allinea is the standard debugger for the KIT, the Intel and PGI compilers are used
- UNICORE , GT4 is used as a standard middleware solution for Grid-based access
- gLite is only used for special user Groups

Is Grid computing user friendly?

- if users using UNICORE, yes!
- if users want to use GT4 or gLite, it‘s only for high-energie physicists or you have to write a web based interface
- later on I will explain the new buzzword ‘cloud‘ computing

I/O performance

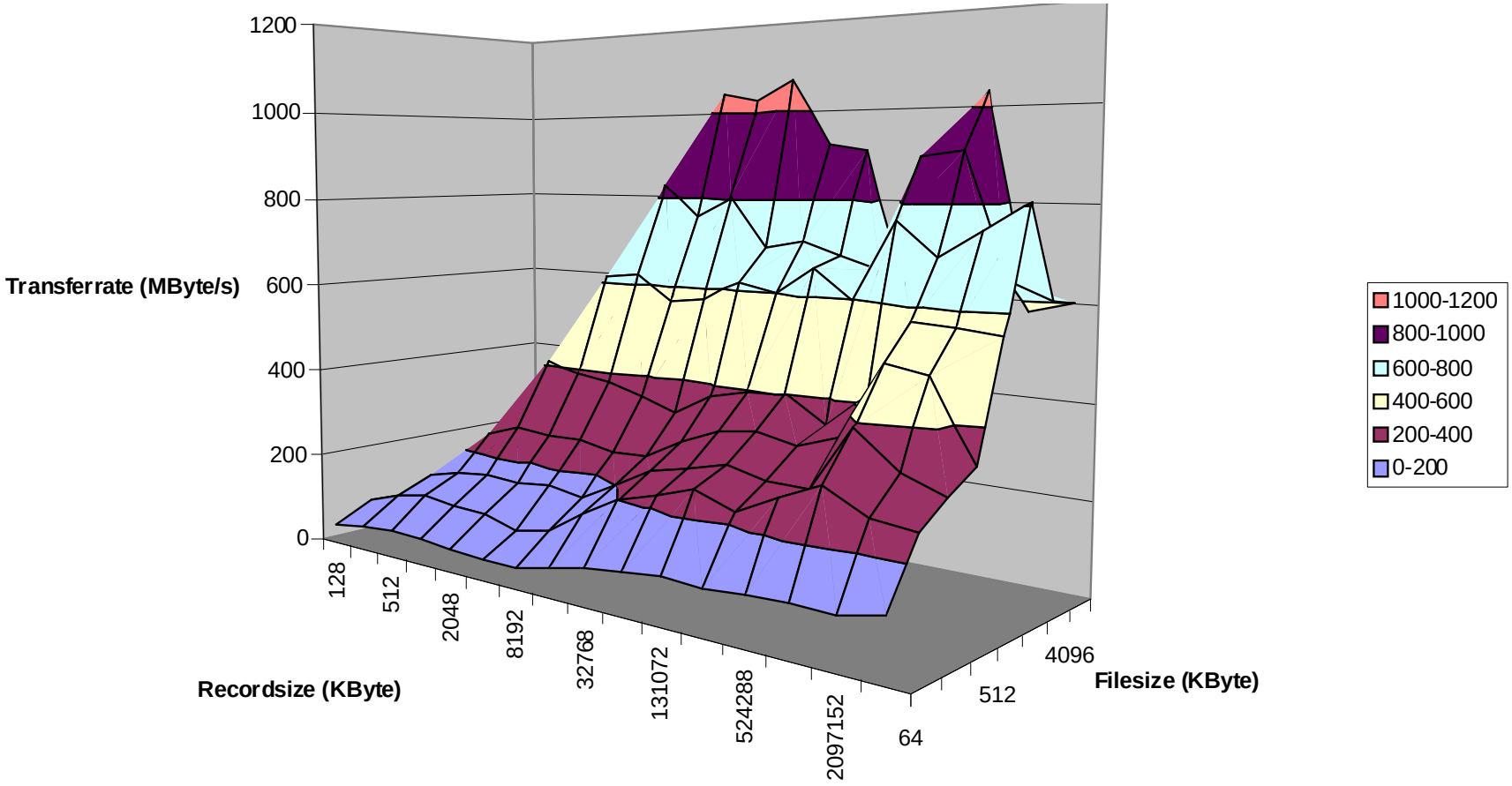
- Why? because CPU is waiting to finish the input or output operation
- But no problem in a only Grid-based environment, CPU is cheap!
- But a problem in the HPC environment, the high-speed network is 30-40% of the node costs depending on the main memory size

Benchmark results

- A factor between 20-25 for an the application LM from DWD for a given set of data.
- It means, from the price-point, a SX-8R (8 processors, 128 GB main memory) node prize should not exceed T€ 150 in summer 2007 to be comparable in prize and performance for LM as given application.

Performance of a small InfiniBand cluster using a FC-Gateway from qLogic (StorNextFS)

iozone (StorNextFS, write using fsync, 8 nodes, 8 processes, peak of a process is 260 MByte/s, three years old EMC² disk system)

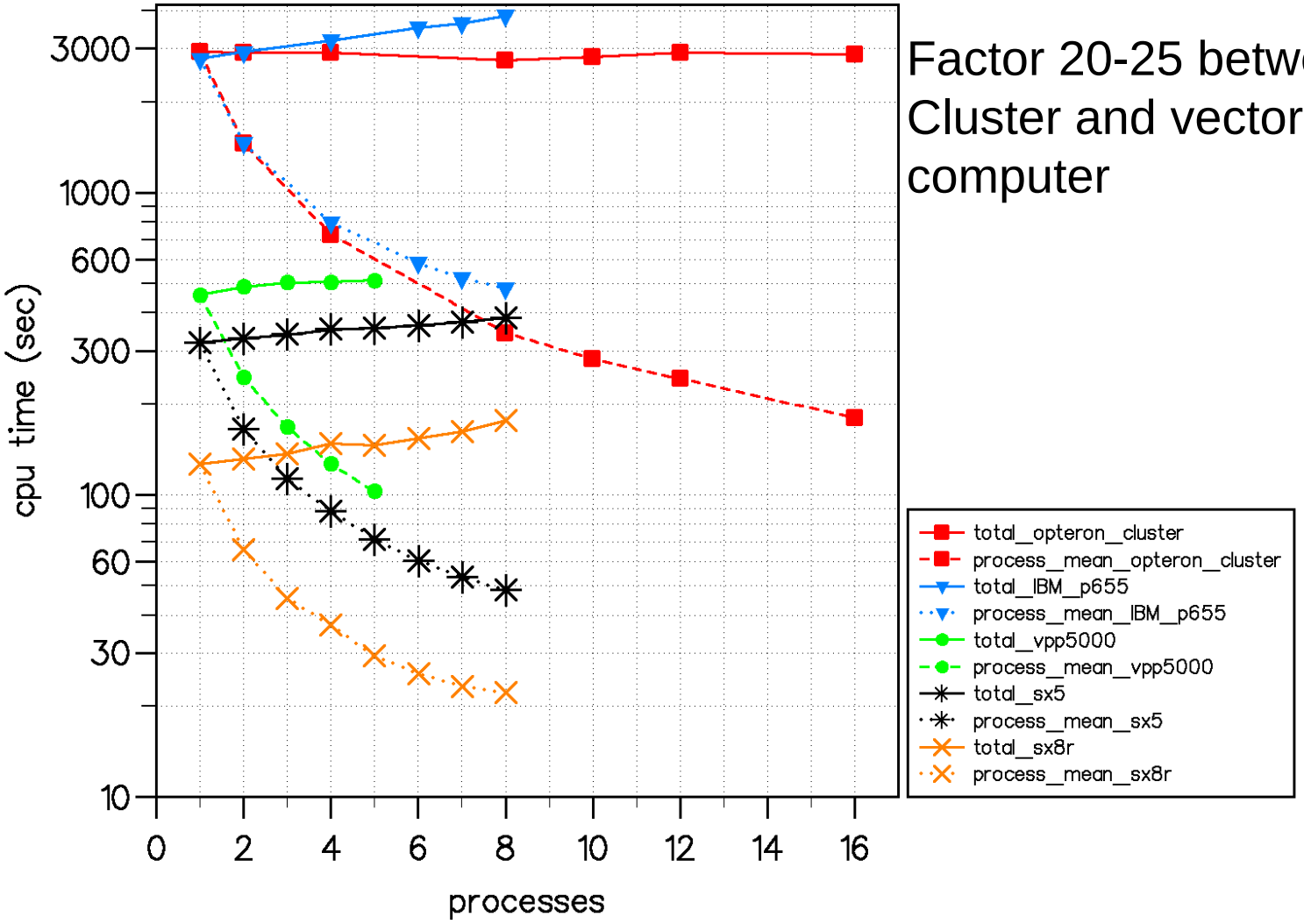


I/O performance

- Why? because CPU is waiting to finish the input or output operation
- But no problem in a only Grid-based environment, CPU is cheap!
- But a problem in the HPC environment, the high-speed network is 30-40% of the node costs depending on the main memory size

Benchmark results

- A factor between 20-25 for an the application LM from DWD for a given set of data.
- It means, from the price-point, a SX-8R (8 processors, 128 GB main memory) node prize should not exceed T€ 150 in summer 2007 to be comparable in prize and performance for LM as given application.



Factor 20-25 between Cluster and vector computer

Usage of CampusGrid

- at the moment 150 Linux jobs, 12 running in the vector environment, 25 running under AIX, both parallel and serial, 150 Linux jobs waiting for CPU resources
- MPI and OpenMP applications
- more than 50 processors Power4 of IBM using AIX as operating system
- 16 vector processors having 0.5 TByte of main memory
- some applications are using more than one week of CPU-time
- more than 50 TByte of global accessible disk space is available for CampusGrid users
- write performance to disk for global access reaches more than 1.3 GByte/s

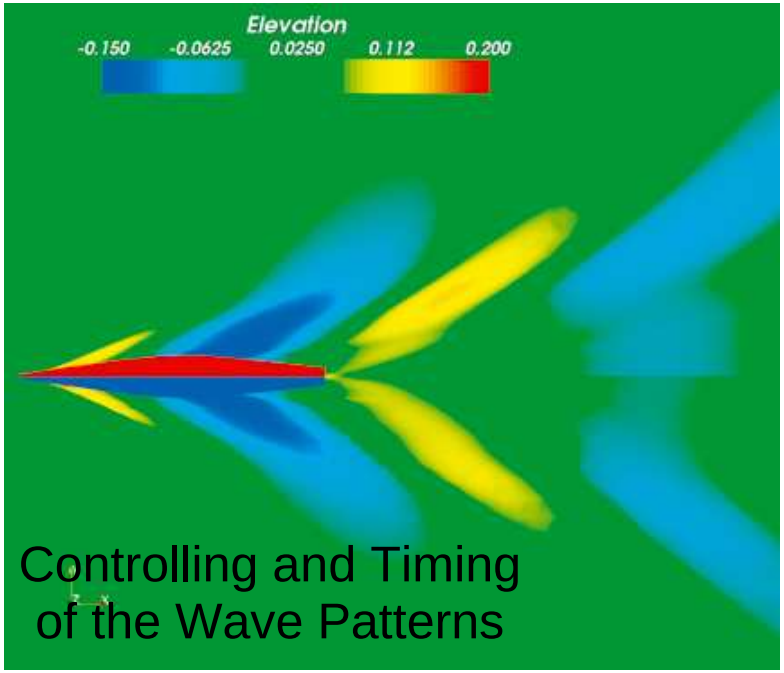
Resources for users wordwide

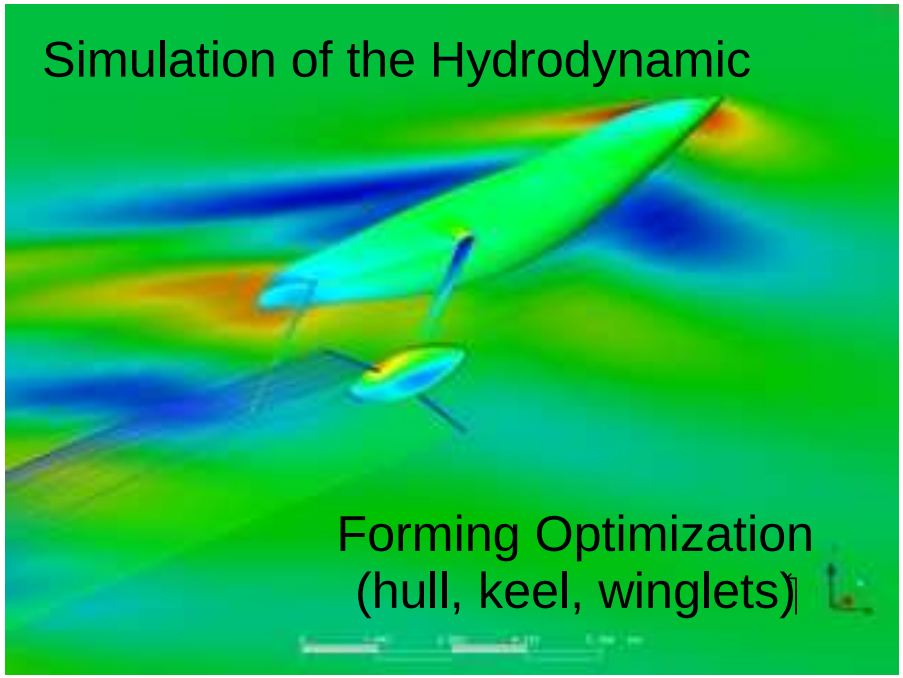
- more than 11000 cores based on Intel and AMD at KIT in principal for the projects GridKa (www.GridKa.de), CampusGrid, BW-Grid, D-Grid and the HPC system for the state Baden-Württemberg
- 5 PByte online disk storage mainly for GridKa
- additional more than 17 PByte in the next 3 years
- the number of cores for HPC is growing, InfiniBand is used as state of the art network
- the number of 10 Gbit Ethernet connections is increasing from 3 to 5 in 2008

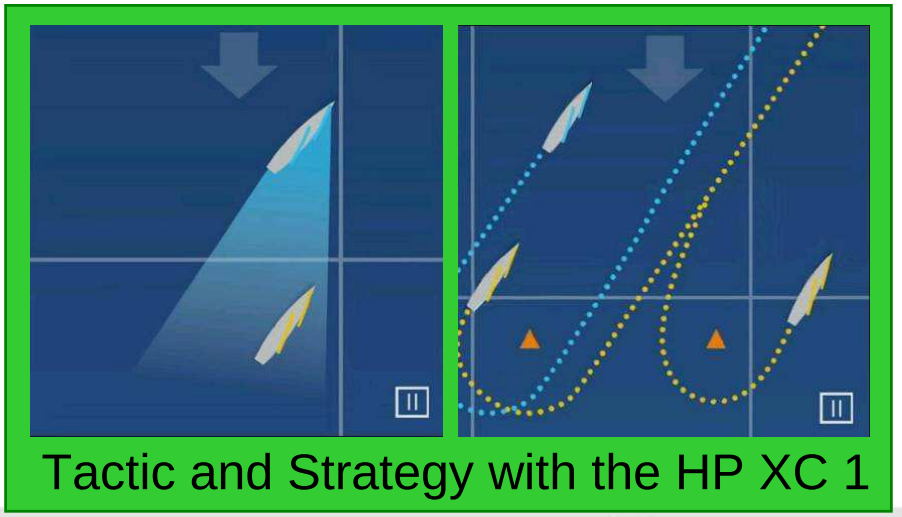
America's Cup: Shosholoza (1)

- Shosholoza Team (South Africa)
- Main sponsor: T-Systems
- Head of the Scientific Advisory Board

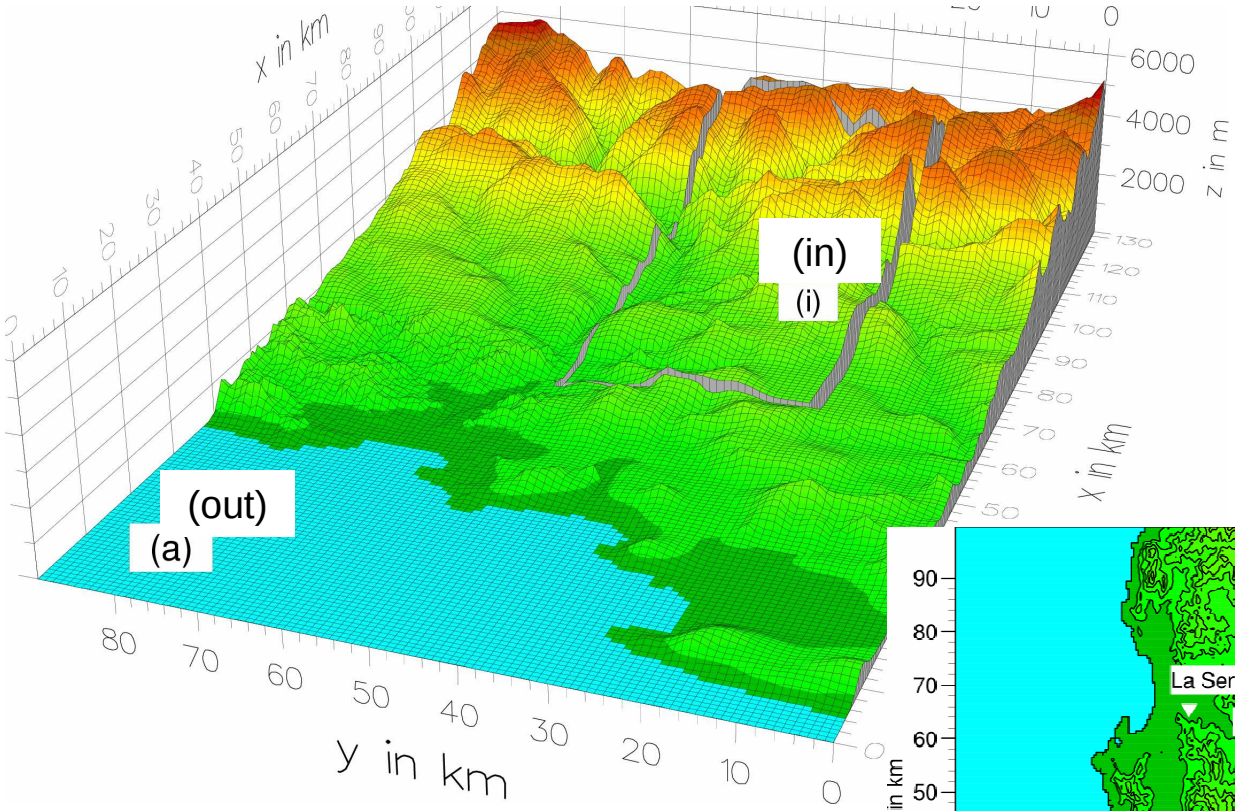
America's Cup: Shosholoza (2)



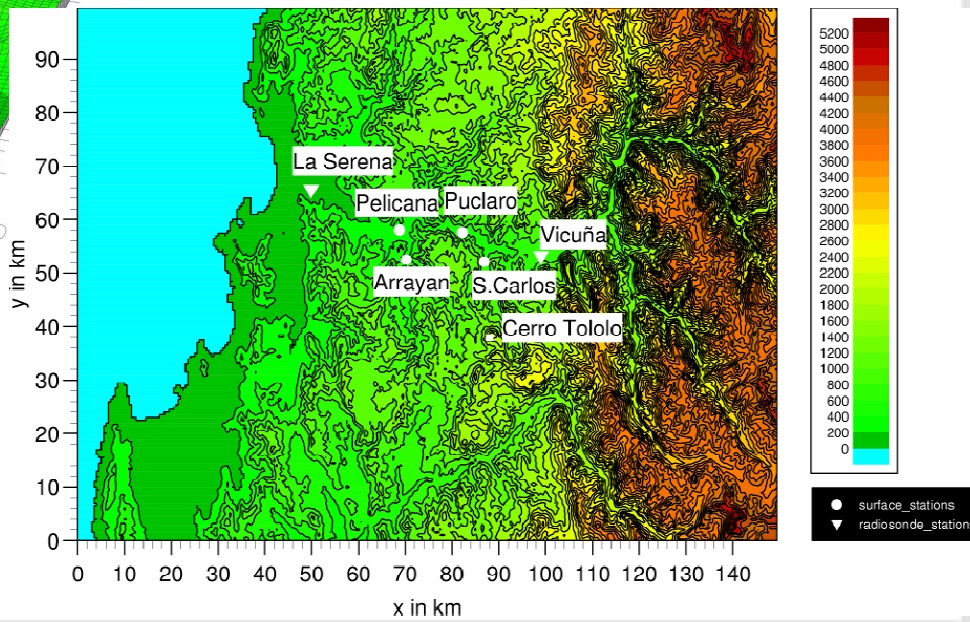




The influence of a Storage Lake in the Arid Elqui Valley in Chile on the Local Climate



225 x 135 horizontal gridpoints
100 vertical levels

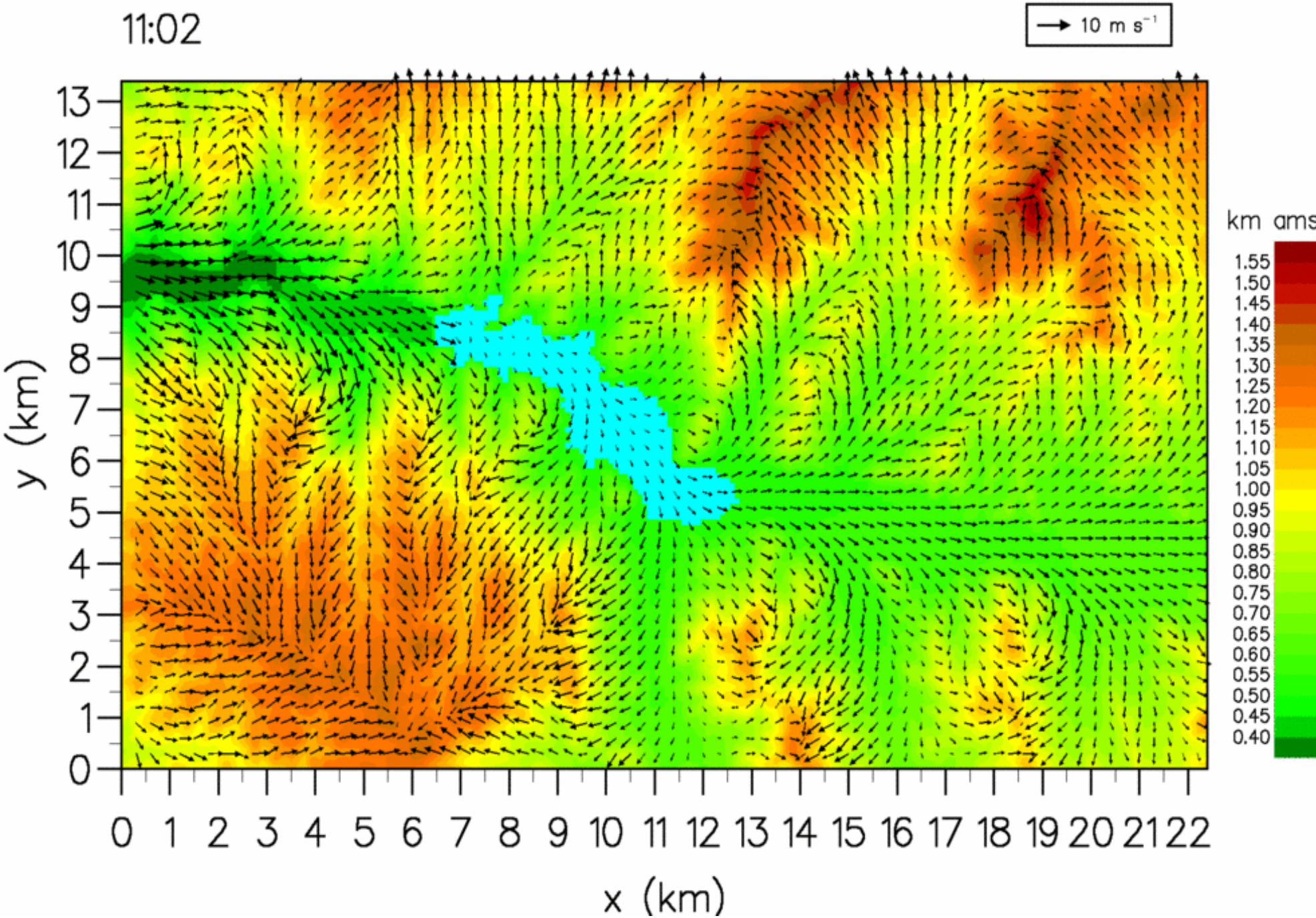


out: Δx = 800 m, Δy = 800m

in(1): Δx = 400 m, Δy = 400 m

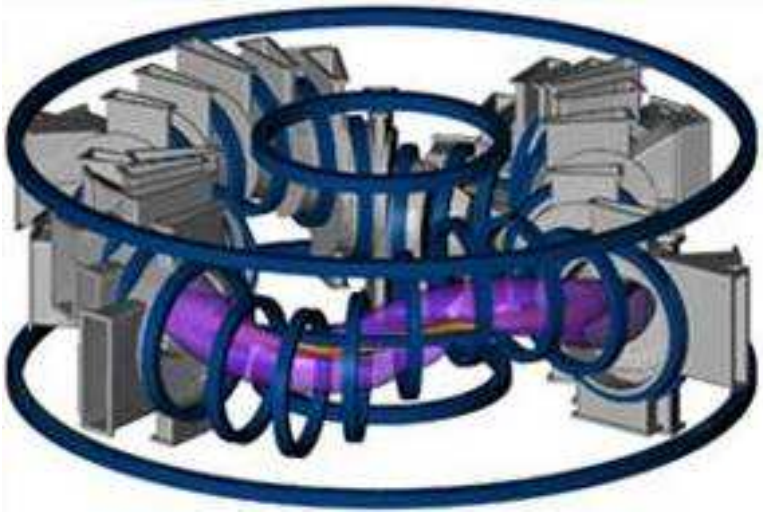
in(2): Δx = 200 m, Δy = 200 m

in(3): Δx = 100 m, Δy = 100 m


11:02
10 m s⁻¹
y (km)
x (km)
0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
km ams
1.55
1.50
1.45
1.40
1.35
1.30
1.25
1.20
1.15
1.10
1.05
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40

HPC Grid-projects in the EU Framework Program

- Euforia: 2008-2010 timeframe, the main objectives of the FP 7 Euforia EU funded project are the deployment of Grid and HPC infrastructure, the adaptation and optimisation of Fusion Codes for the ITER-project and the and the development of advanced tools for Workflow management, Visualization tools and Data mining. This project is already heavily involved in the use of Globus and UNICORE standards. → gLite of the moment only on Linux based clusters at KIT, but ……..



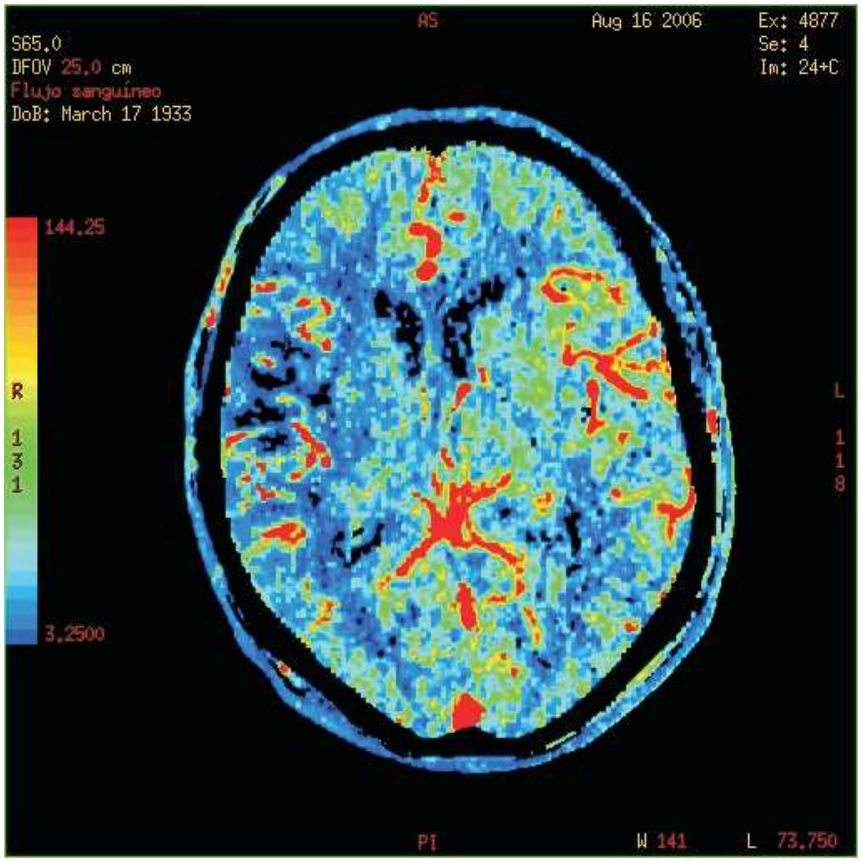
New computing power →
new powerful results

HPC Grid-projects

- int.eu.grid: 2006-2008 (FP 6), the interactive European Grid projec will deal with MPI applications based on Linux-clusters without fast network devices.

int.eu.grid

Medical applications → (computer tomography, brain perfusion, blood flow)



C³-Grid: Collaborative Climate Community Grid in the D-Grid project (www.C3Grid.de)

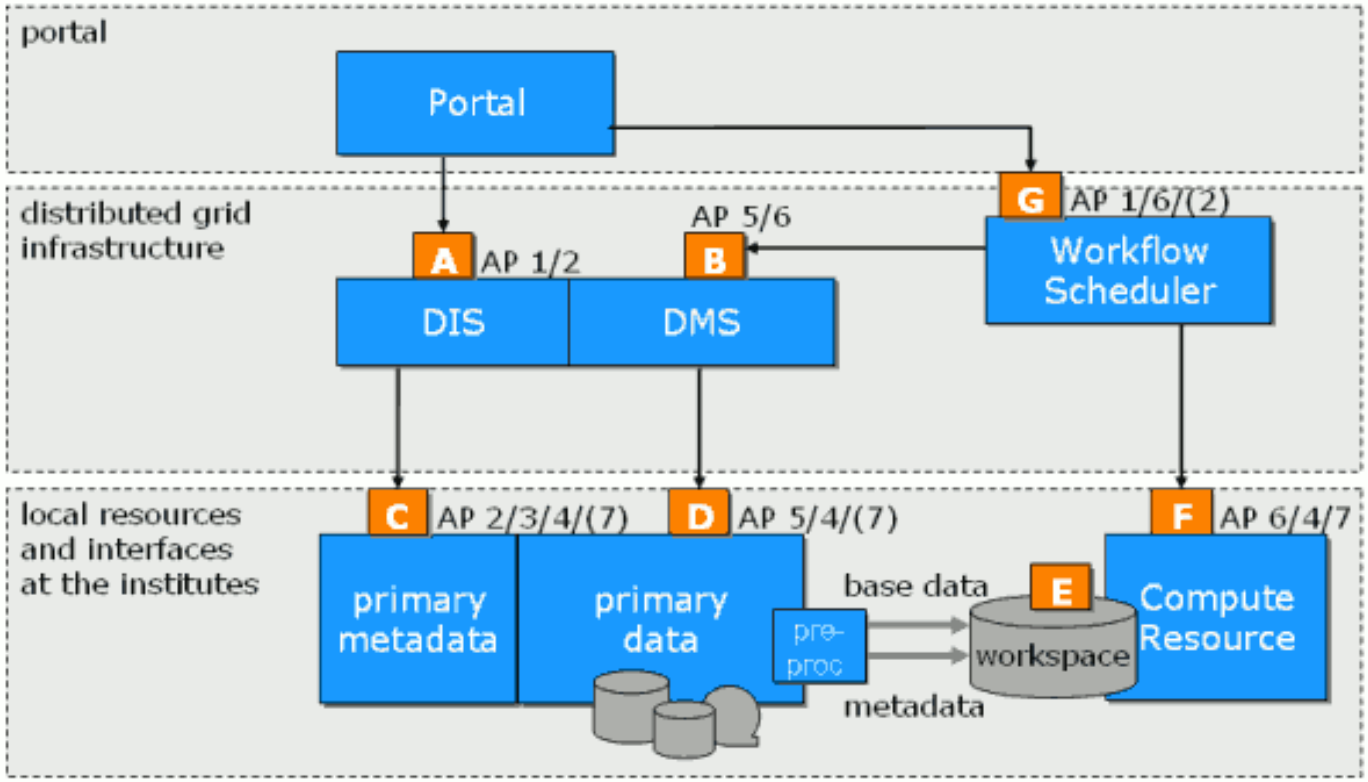
- Why?
  - research activities to investigate the interaction of the climate system with the socio-economical system
  - runs in the past (20 years or more!) at different departments in Germany
  - more complex and high-resolution model simulations
  - the availability of increasing satellite data is relevant for monitoring the earth system
- Goal
  - Connection of distributed climate data archives
  - Uniform access and central metadata catalogue
  - it is not a prior aim to participate in compute Grid resources!

C³-Grid

- Explanation
  - Access to data collections of different organizations is extremely difficult because of local storage architecture and inhomogeneous data information.
  - To link various data sets together is an essential part of a scientist's workflow for the research. Thus, the missing uniform access is a critical shortage for the earth system research community.
  - Historically, if data volume is high, it is decentralized stored at distributed archives at various research departments.

C³-Grid

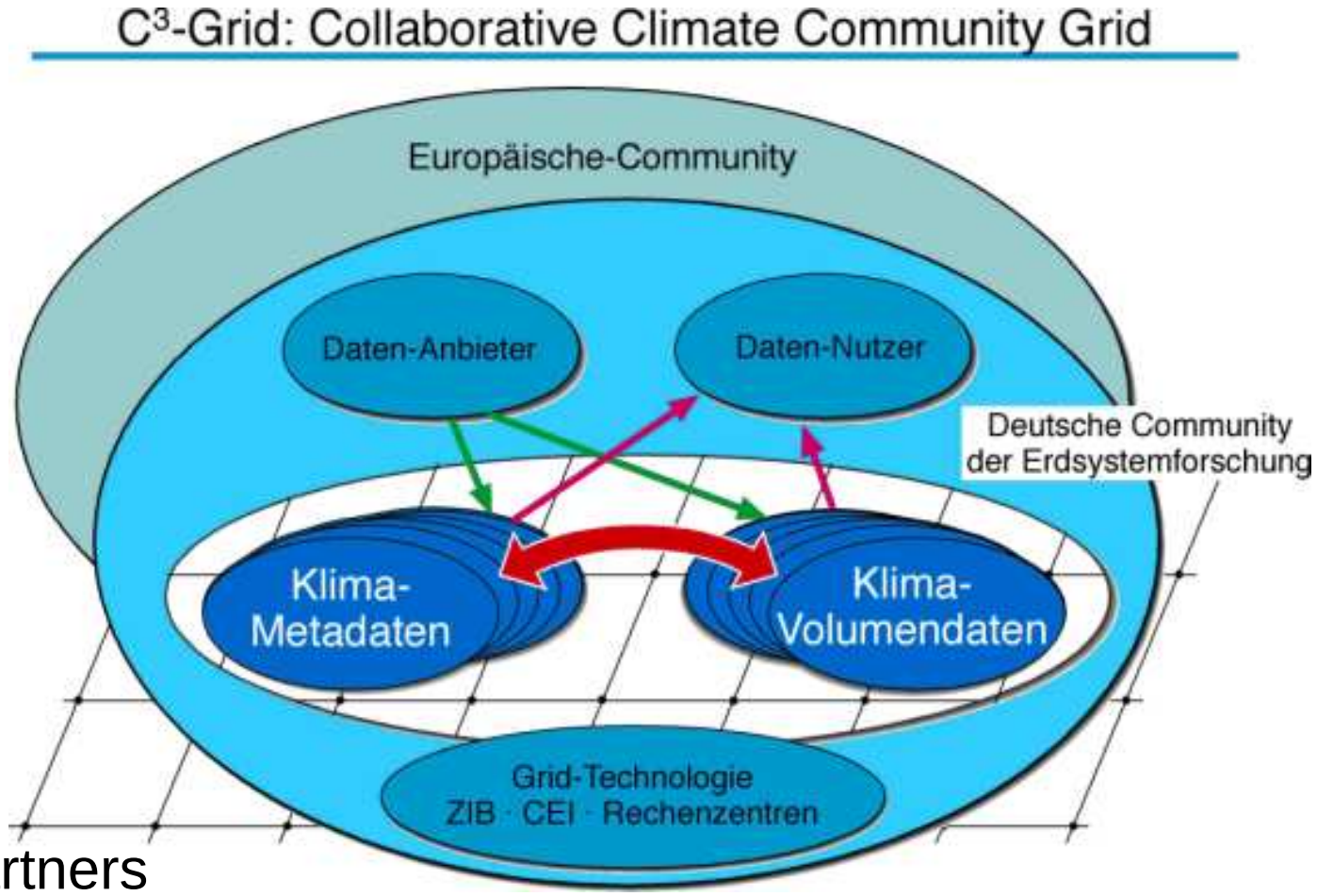
- Solution



DIS: Data Information Service
DMS: Data Management Service

- Access rules to the distributed data storage will be defined by the research department holding the data.
- A Grid-based access must be available.

C³-Grid



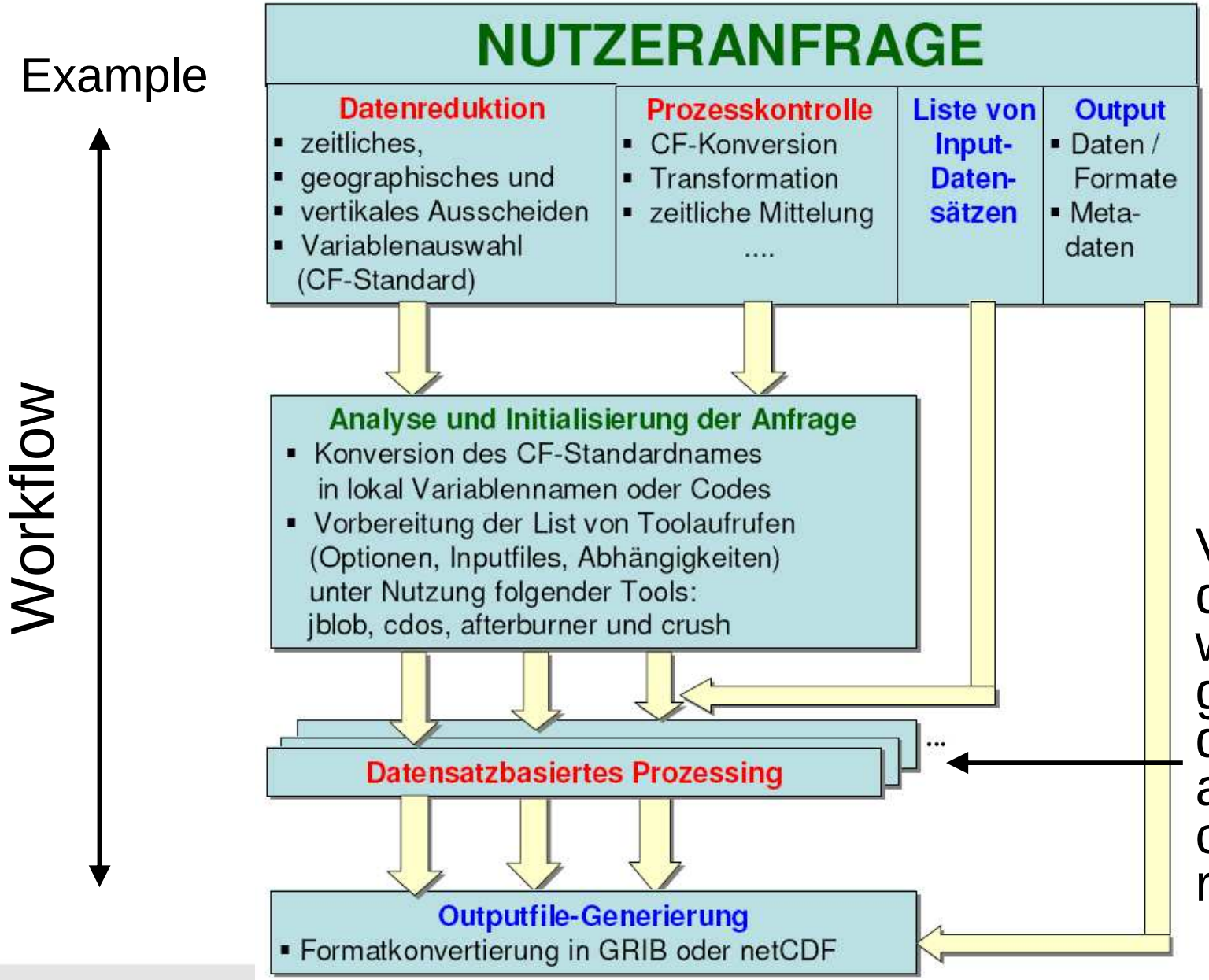
- Partners
  - Official Partners
  - KIT with the Institute of Meteorology and Climate Research and the SCC are associated to the project

Cloud Computing

- What does it means?
  - Partly funded by the European Union (M€17 grant) IBM's Haifa lab in Israel and 17 European organizations will look at managing IT services across domains, IT platforms and geographies. The project is called RESERVOIR - Resources and Services Virtualization without Barriers.
  - In the United States of America IBM collaborates with the Georgia Institute of Technology and Ohio State University to develop new autonomic technology on self-managing features for virtualized data centers in a cloud computing environment.
  - The coordinated use of hardware, operating systems, middleware and applications is a must.
  - Applications must rely on Web-based programs.

Cloud Computing for C³

- Example



Virtualized data center with guarantied data lifetime and compute resources

Cloud Computing

- Companies
  - Microsoft will expand its Microsoft Online Services utilizing Microsoft Exchange Server 2007 as well as Microsoft Office SharePoint Service 2007. Services such as e-mail, contacts, calendars and conferencing will all be available to use and sync online.
  - Google's online office utility is working in the same direction.
  - IBM and Sun are offering a compute cloud for HPC.
  - Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute and storage capacity in the cloud. It is designed to make web-scale computing easier for developers. But it's not HPC!
- Amazon EC2's web service interface
  allows you to obtain and configure capacity with minimal friction. It lets you run on Amazon's proven computing environment. Amazon EC2 provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios.

Conclusion

- KIT is the largest research and development organization in Germany.
- SCC provides a lot of IT resources worldwide.
- Scientists are supported by SCC.
- D-Grid is a success story.
- Applications are not simple to gridify!
- C^3 is the climate project as part of D-Grid.
- The future of Grid computing is …..
- Cloud computing could be a way for standardizing accessing IT infrastructure and to integrate in a quite simple way scientific applications and business applications.

Thank you!

Questions?