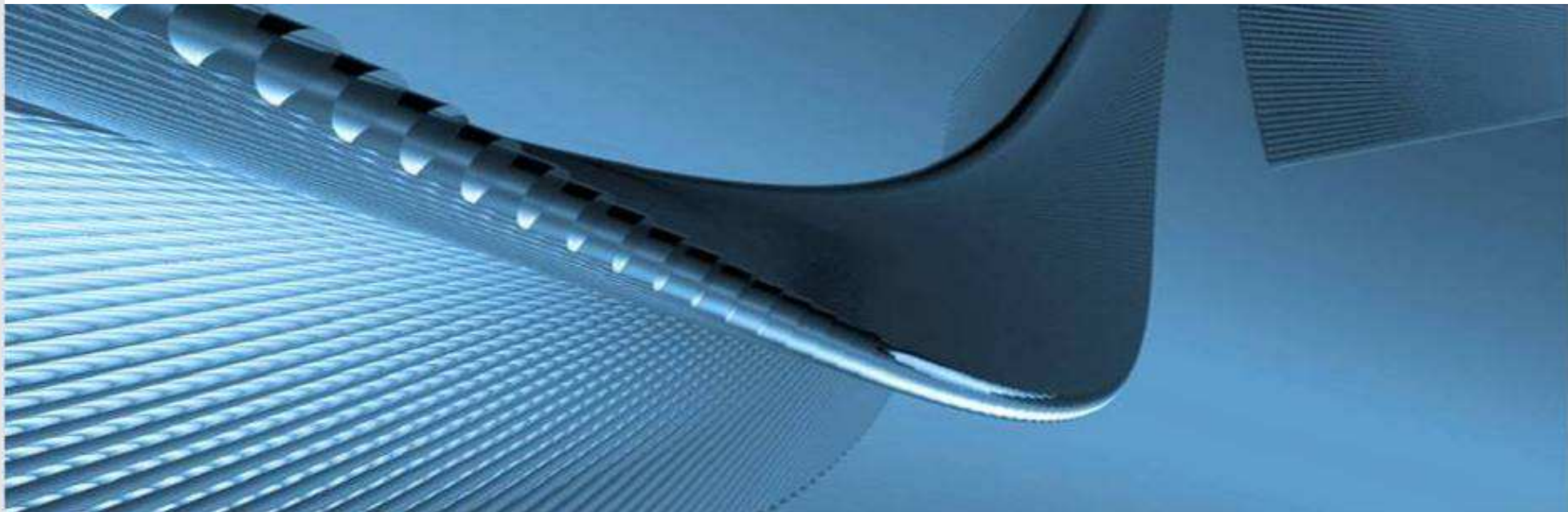


# The usage of Grid Middleware at KIT, C<sup>3</sup> Grid at D-Grid and the Future of Grid Computing

Frank.Schmitz@kit.edu



# Agenda

- n Introduction
  - KIT and SCC, organisation and resources
- n Grid computing and scientists?
- n I/O-Performance and benchmark results
- n Usage and resources
- n European projects
- n D-Grid and C<sup>3</sup>
- n Cloud computing
- n Unicore for HPC
- n Conclusion

# Karlsruhe Institute of Technology (KIT)

## University

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

## KIT...

## Forschungszentrum

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

## infrastructure

## n Karlsruhe Institute of Technology - two strong partners united

- α Universität Karlsruhe (TH)
  - n One of the most research oriented universities in Germany
  - n Nominated to one of three top universities in October 2006
  - n Capitalized with over 100 million € over 5 years
- α Forschungszentrum Karlsruhe
  - n Member of the Helmholtz Association
  - n Internationally acknowledged non-university institution

## n Under the roof of KIT – Steinbuch Centre for Computing

- α Computing Center of the Universität Karlsruhe (TH)
- α Institute of Scientific Computing at the Forschungszentrum Karlsruhe

# Virtual Computing Centre Karlsruhe since 1996



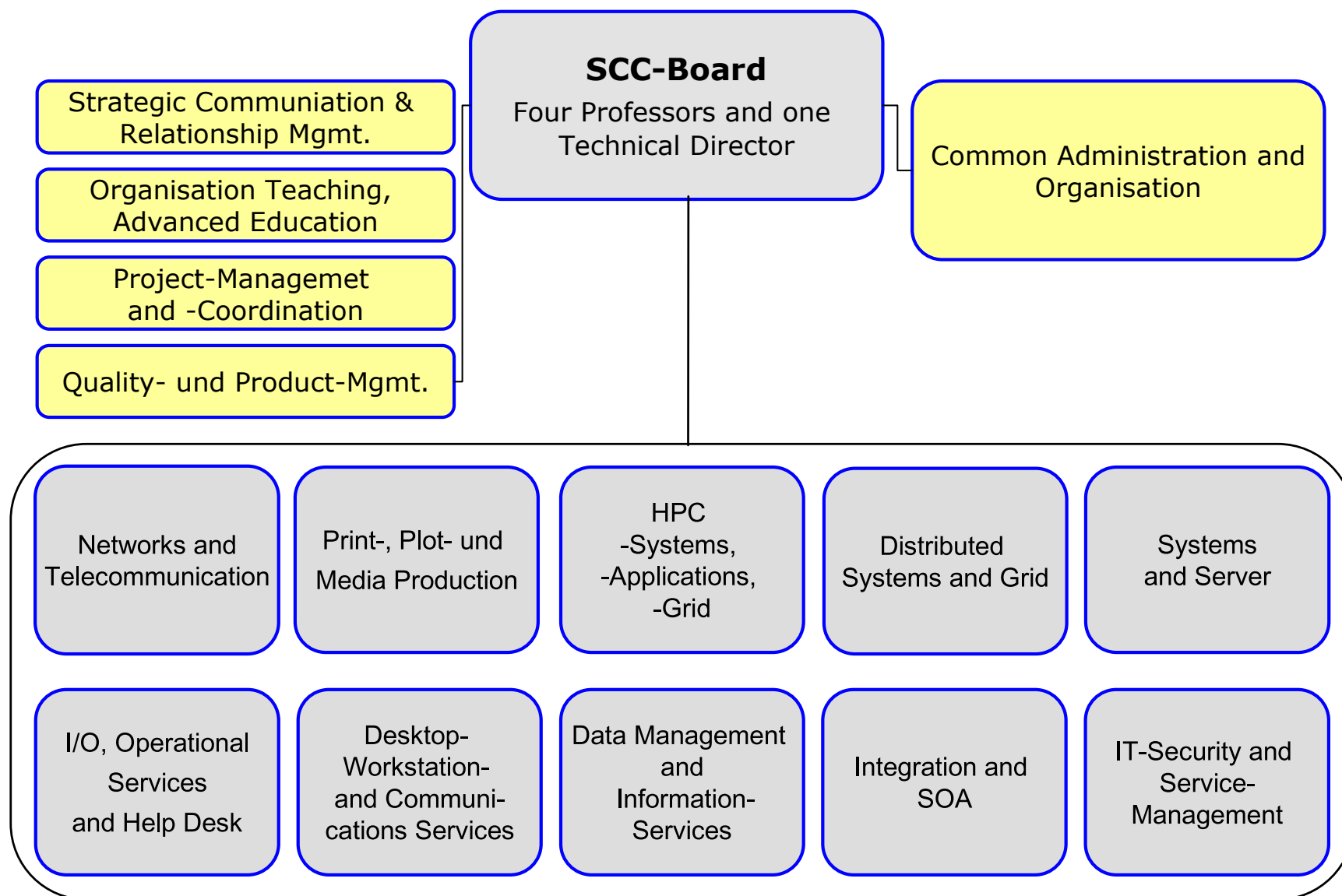
Cooperation in  
High Performance Computing

Virtual Computing Centre since 1996

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

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

# 10 SCC Departments



# HP XC4000 – Baden-Württemberg-State-Supercomputer



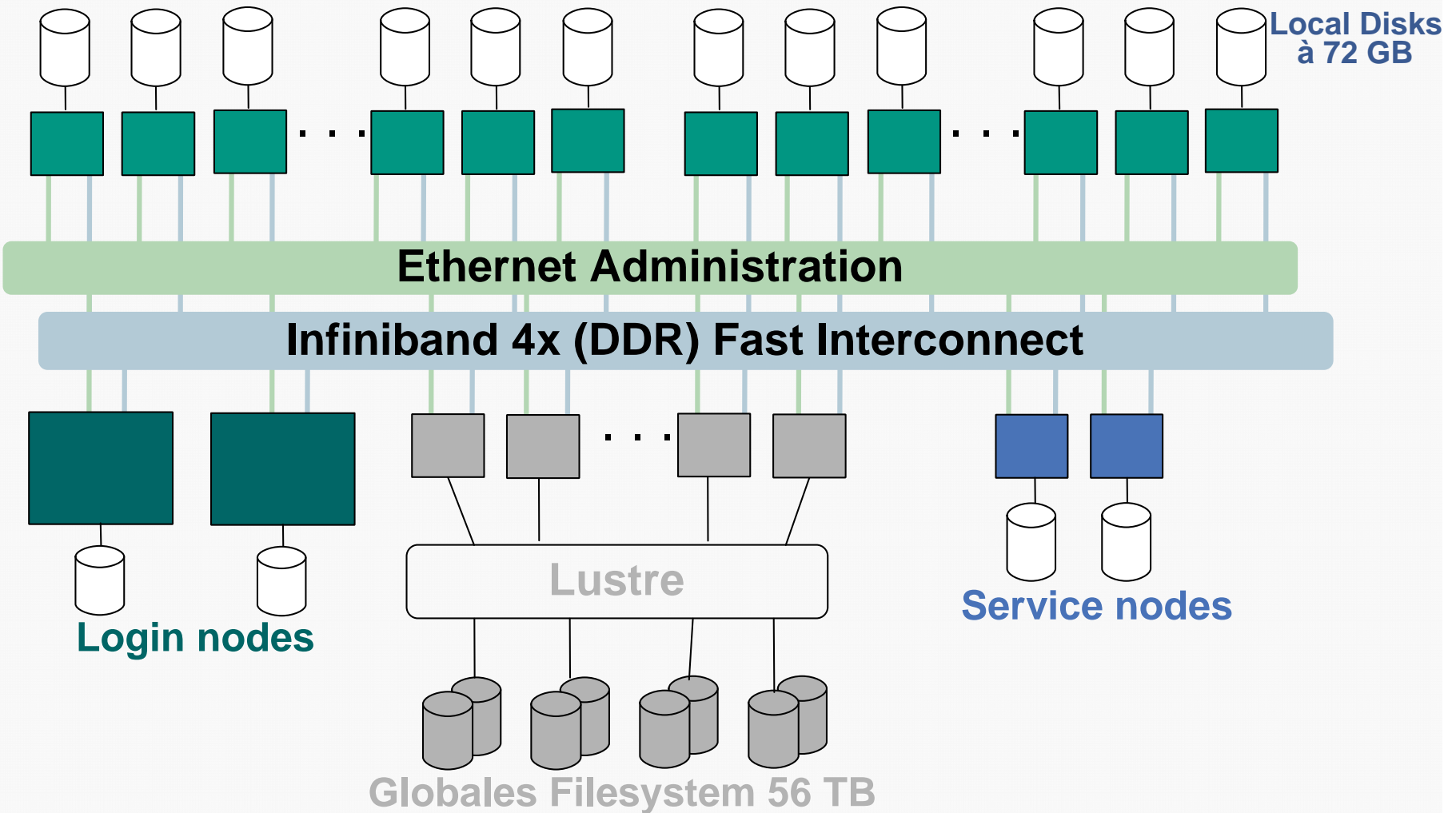
## Phase 2 (Q3 2006)

- » 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

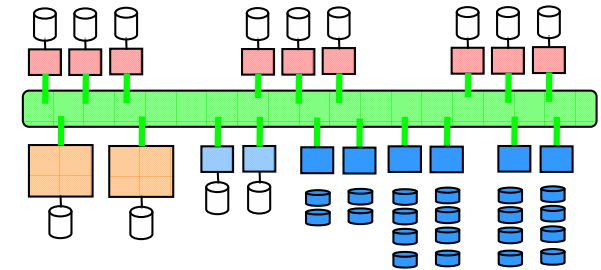
750 Compute nodes





# Parallel File System HP SFS on XC4000

- n Shared storage for all nodes of XC system
- n 10 file server nodes
  - 2 MDS / Admin
  - 2 OSS for \$HOME
  - 6 OSS for \$WORK
- n 56 TB file space
  - 8 TB \$HOME
  - 48 TB \$WORK
- n 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



# Hardware running at KIT

- >11000 Intel and AMD based compute cores mainly for the projects CampusGrid ([www.CampusGrid.de](http://www.CampusGrid.de)), D-Grid ([www.d-grid.de](http://www.d-grid.de)), GridKa ([www.GridKa.de](http://www.GridKa.de)), BW-Grid ([www.bw-grid.de](http://www.bw-grid.de)) and the HPC system for Baden-Württemberg
  - 5 PB online storage, 2PB tape storage
  - 52 Power4 cores
  - 16 vector processors each 32 GB main memory
  - additional more than 17 PByte in the next 3 years
- but floor space is limited and the power consumption is dramatically

# Parts of the CampusGrid Project



a SX-9 node will be added in  
autumn 2008

# Is Grid computing user friendly?

- if users using UNICORE, yes!
- if users want to use GT4 or gLite, no. 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

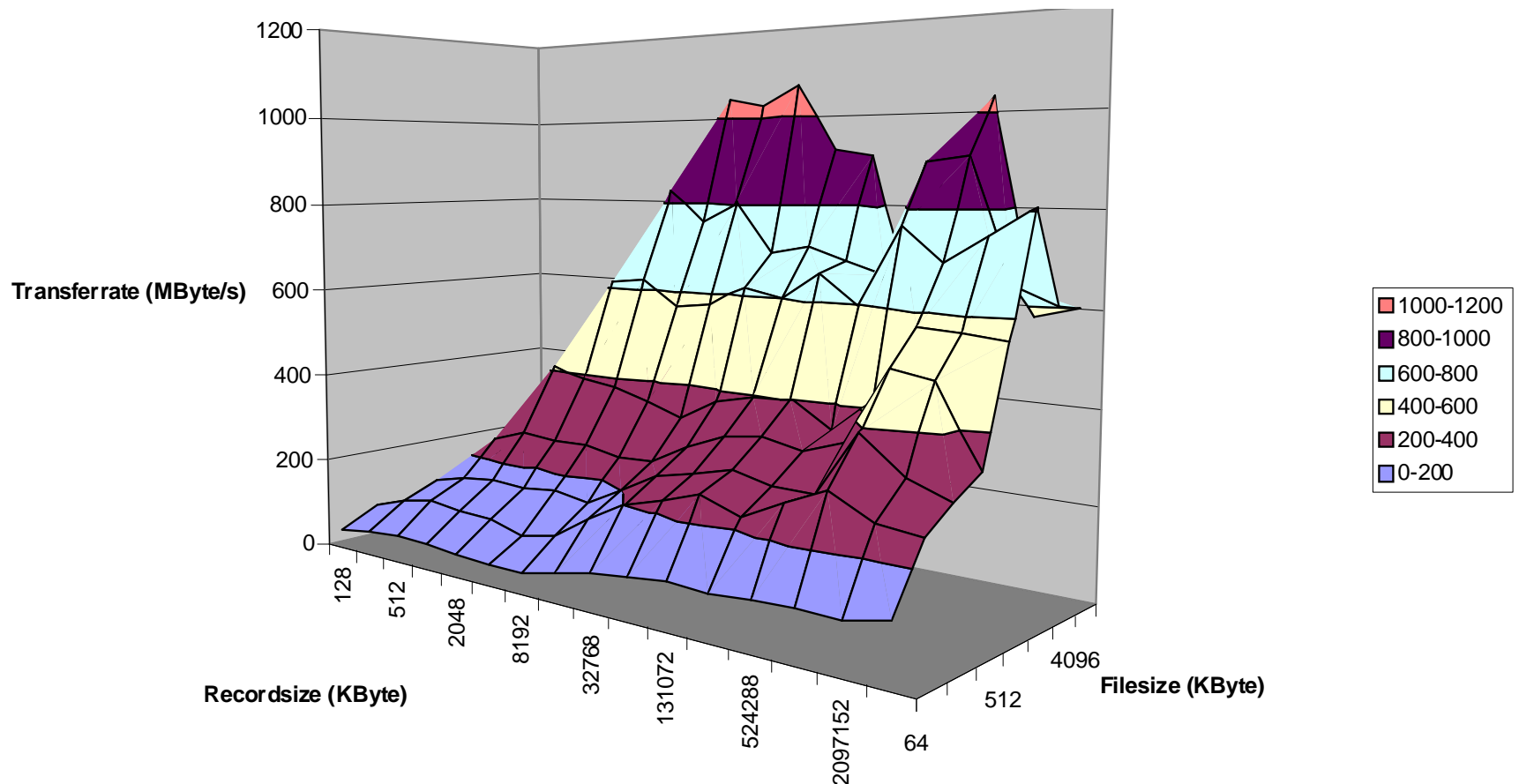
- Why? because CPU is waiting to finish the input or output operation
- But no problem in a only Grid-based environment, CPU is cheap!
- But a problem in the HPC environment, the high-speed network is 30-40% of the node costs depending on the main memory size

## 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 EMC<sup>2</sup> disk system)

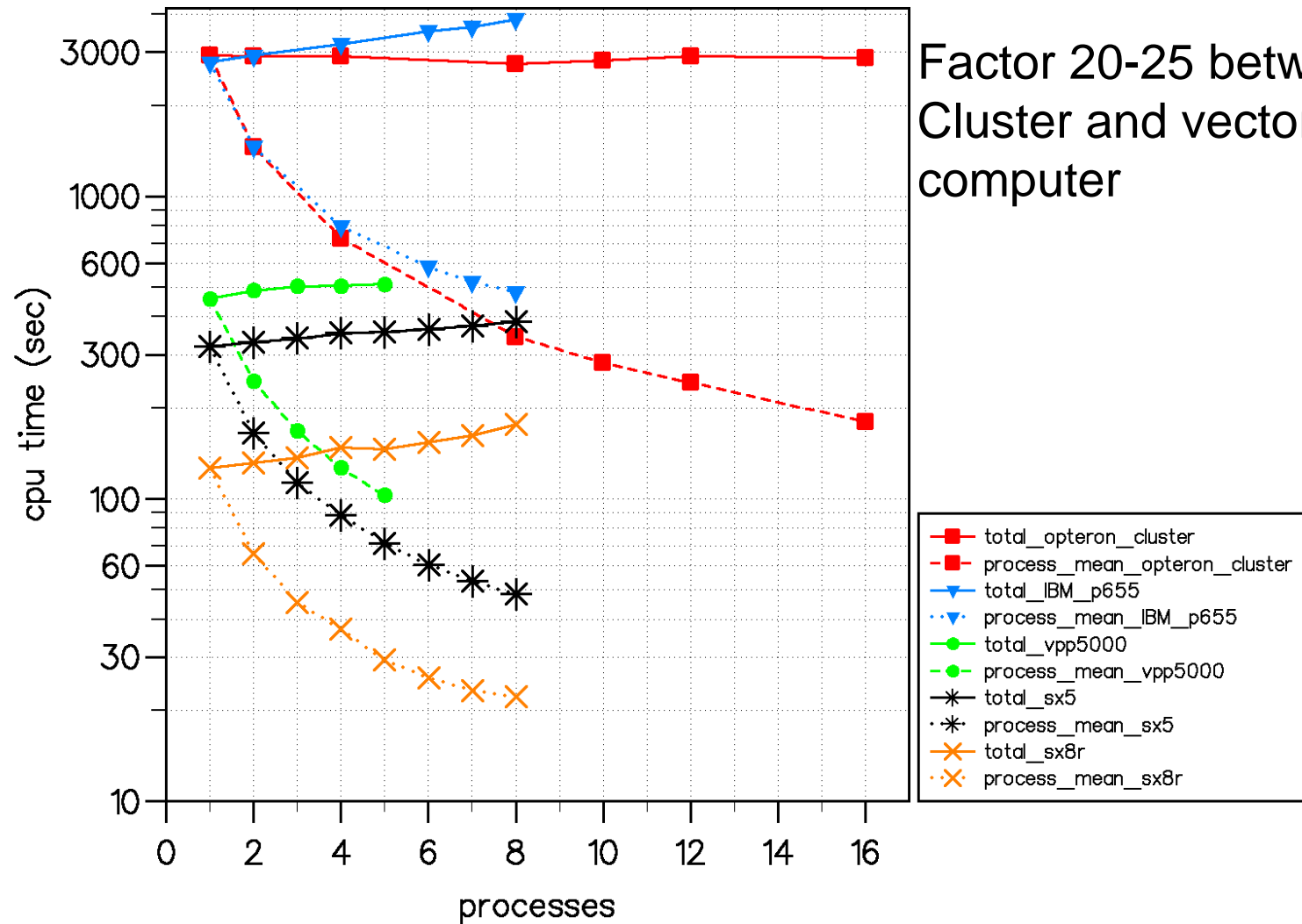


## I/O performance

- 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

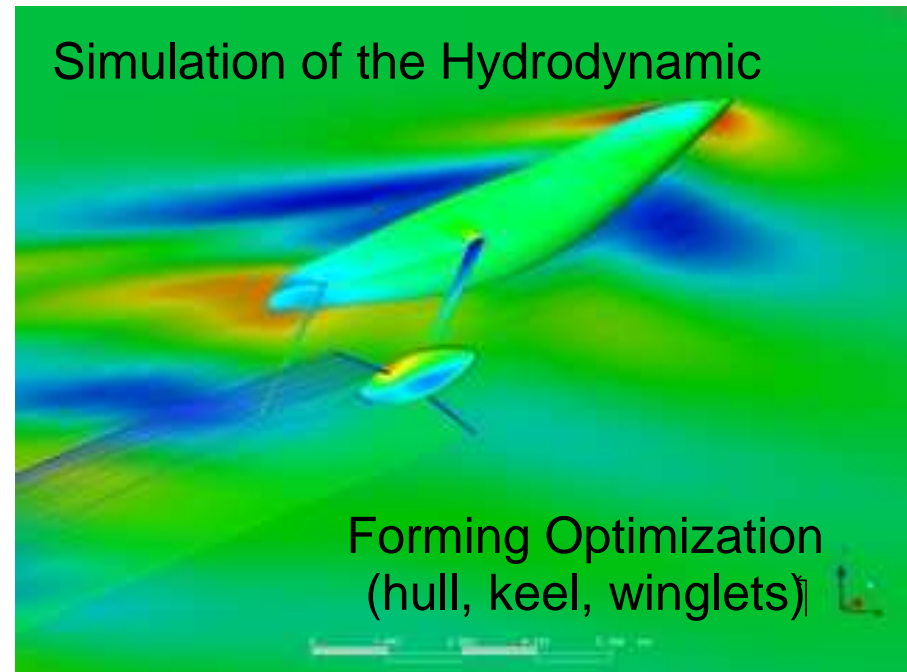
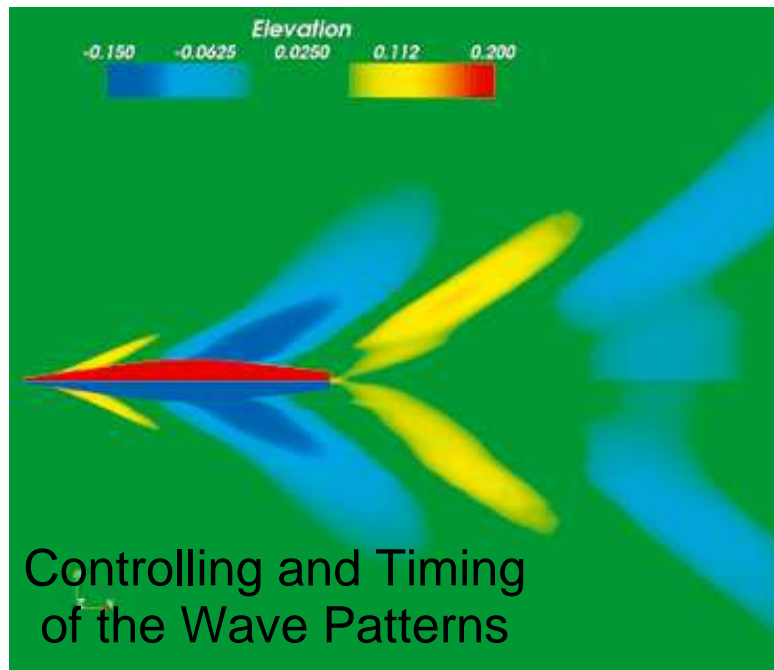


# America's Cup: Shosholoza (1)

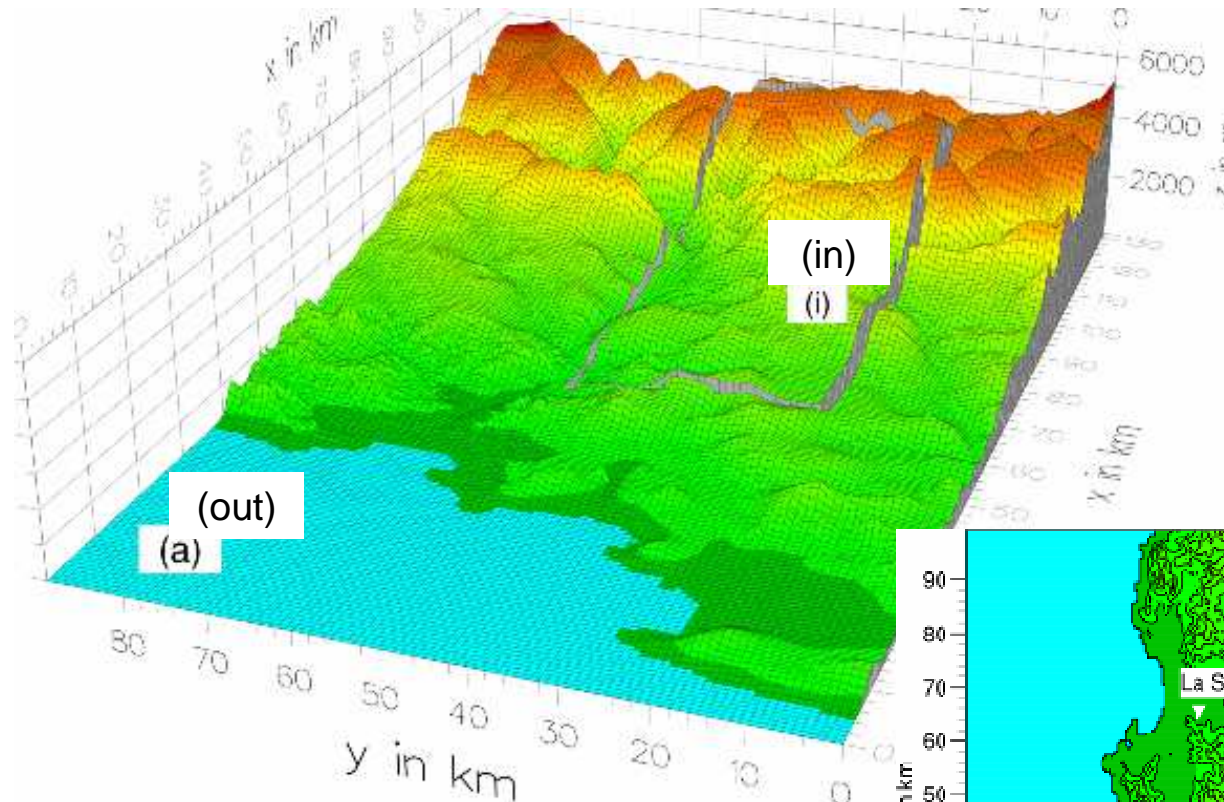
- n Shosholoza Team (South Africa)
- n Main sponsor: T-Systems
- n Head of the Scientific Advisory Board



# America's Cup: Shosholoza (2)

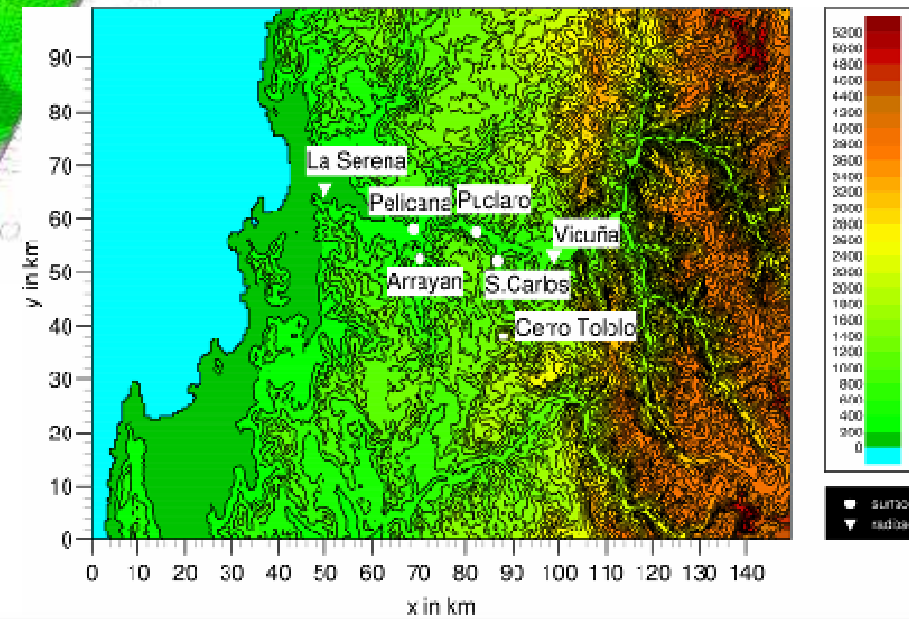


# The influence of a Storage Lake in the Arid Elqui Valley in Chile on the Local Climate



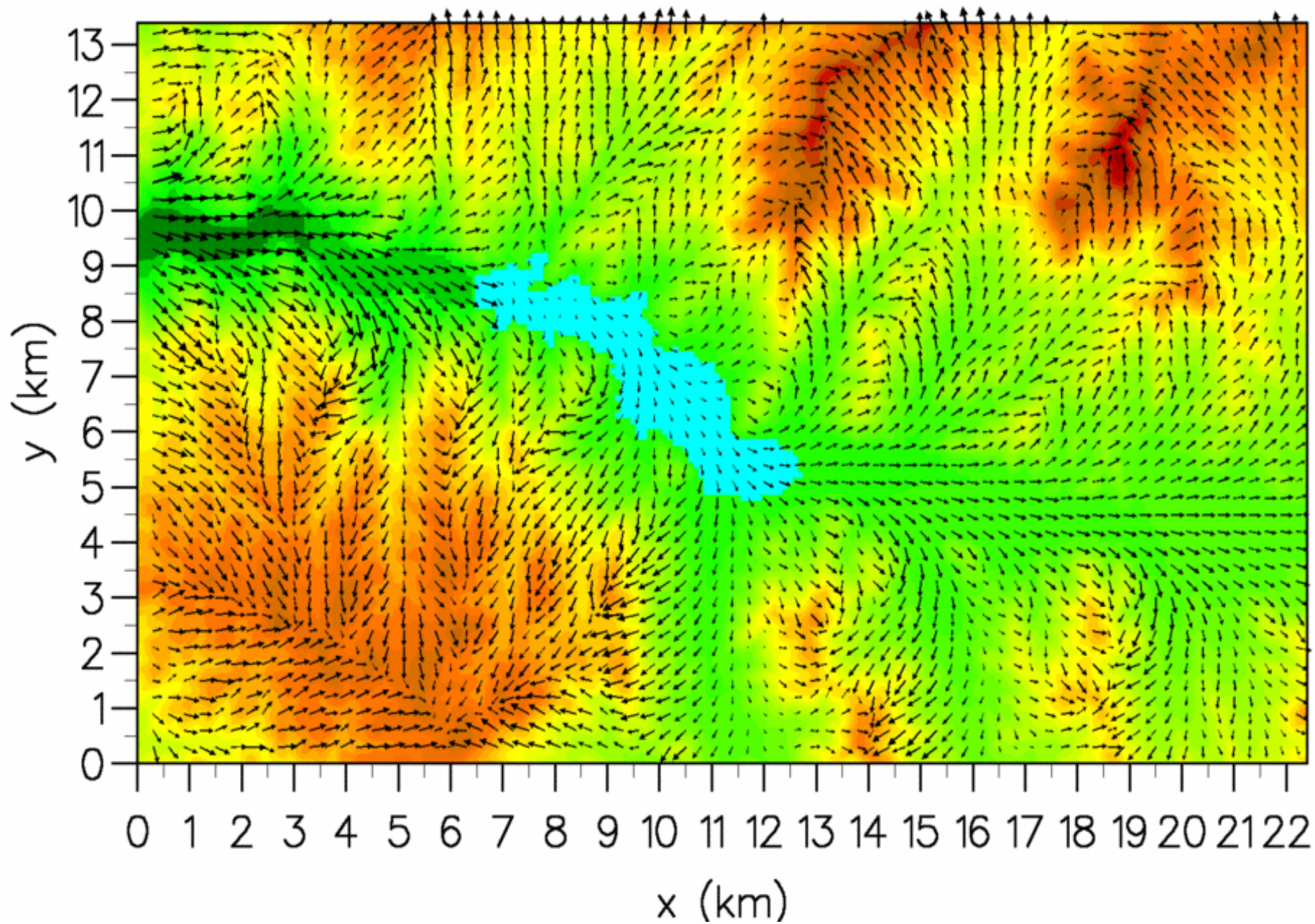
225 x 135 horizontal gridpoints  
100 vertical levels

out:  $\Delta x = 800 \text{ m}$ ,  $\Delta y = 800 \text{ m}$   
 in(1):  $\Delta x = 400 \text{ m}$ ,  $\Delta y = 400 \text{ m}$   
 in(2):  $\Delta x = 200 \text{ m}$ ,  $\Delta y = 200 \text{ m}$   
 in(3):  $\Delta x = 100 \text{ m}$ ,  $\Delta y = 100 \text{ m}$



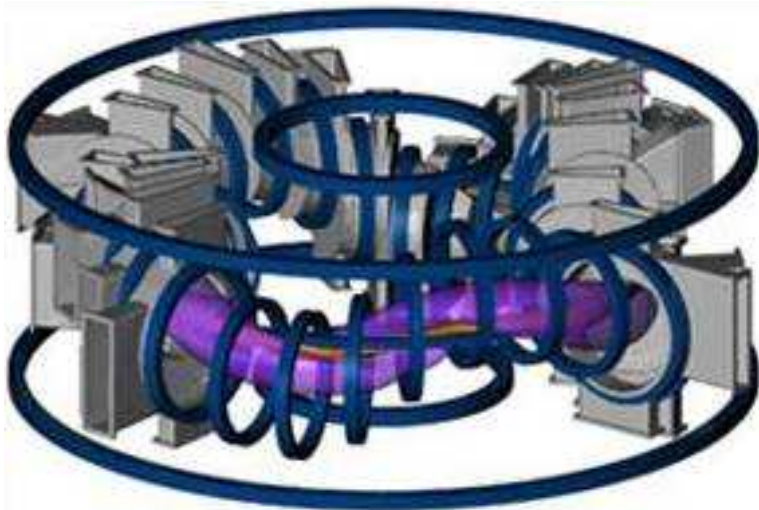
11:02

→ 10 m s<sup>-1</sup>



# HPC Grid-projects in the EU Framework Program

- n 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 .....



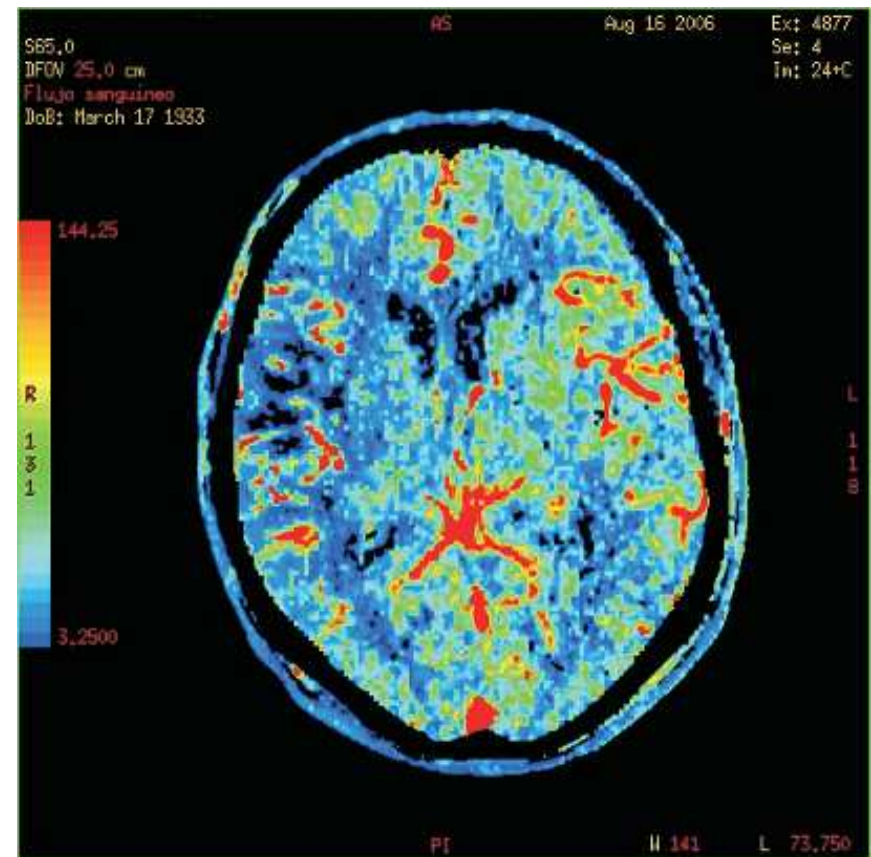
New computing power →  
new powerful results

# HPC Grid-projects

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



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



# C<sup>3</sup>-Grid: Collaborative Climate Community Grid in the D-Grid project ([www.C3Grid.de](http://www.C3Grid.de))



## n 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

## n Goal

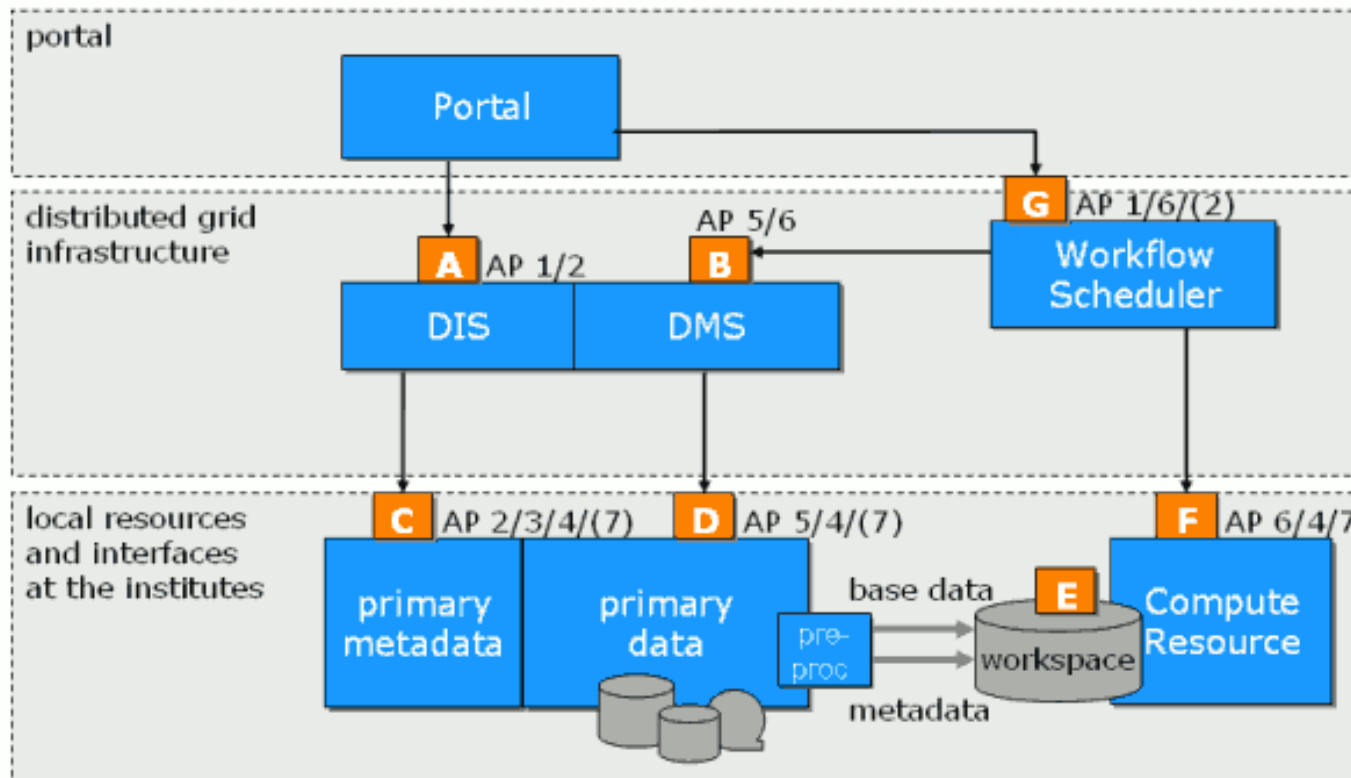
- Connection of distributed climate data archives
- Uniform access and central metadata catalogue
- it is not a prior aim to participate in compute Grid resources!

## n 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.



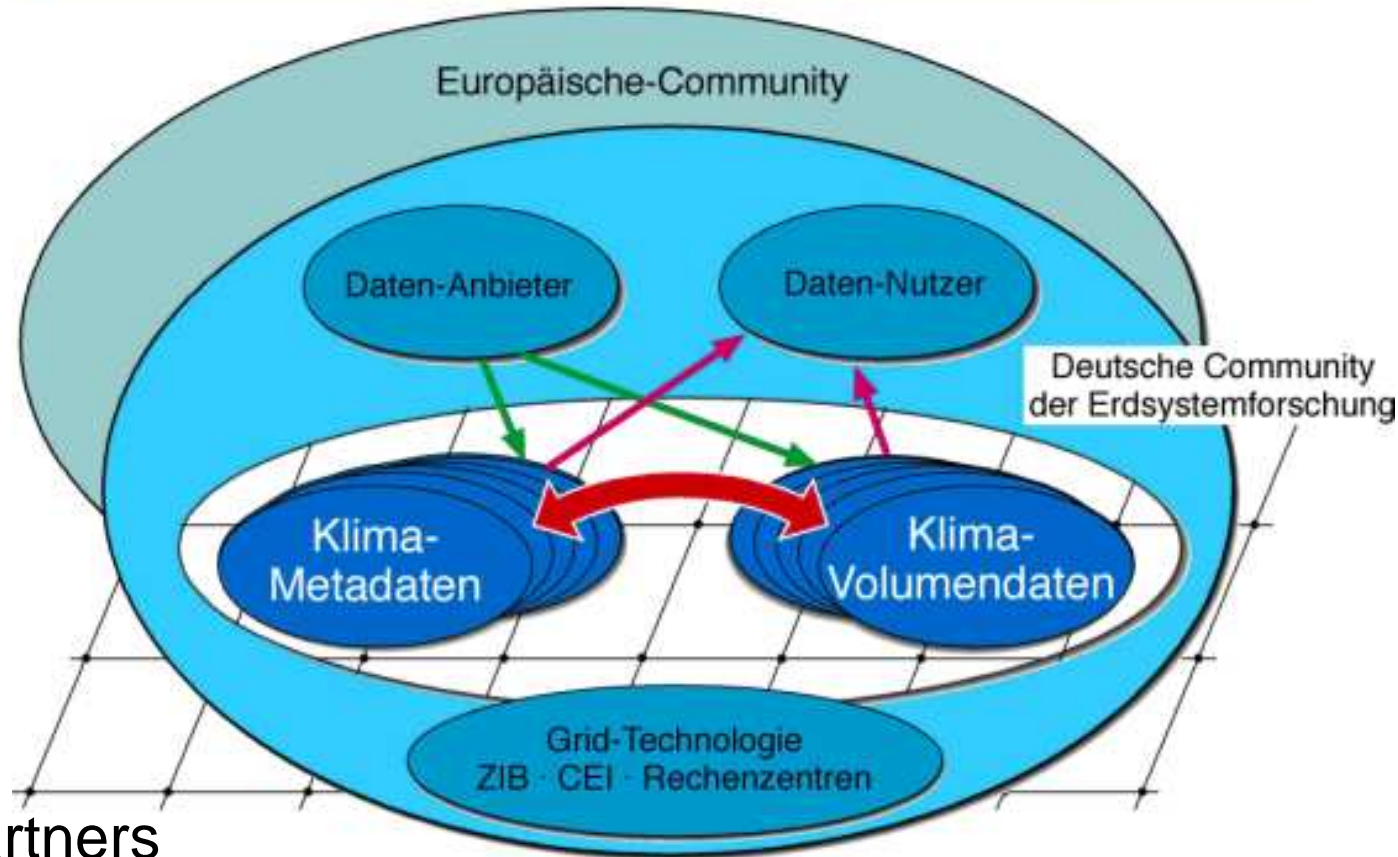
## n Solution



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.





### n Partners

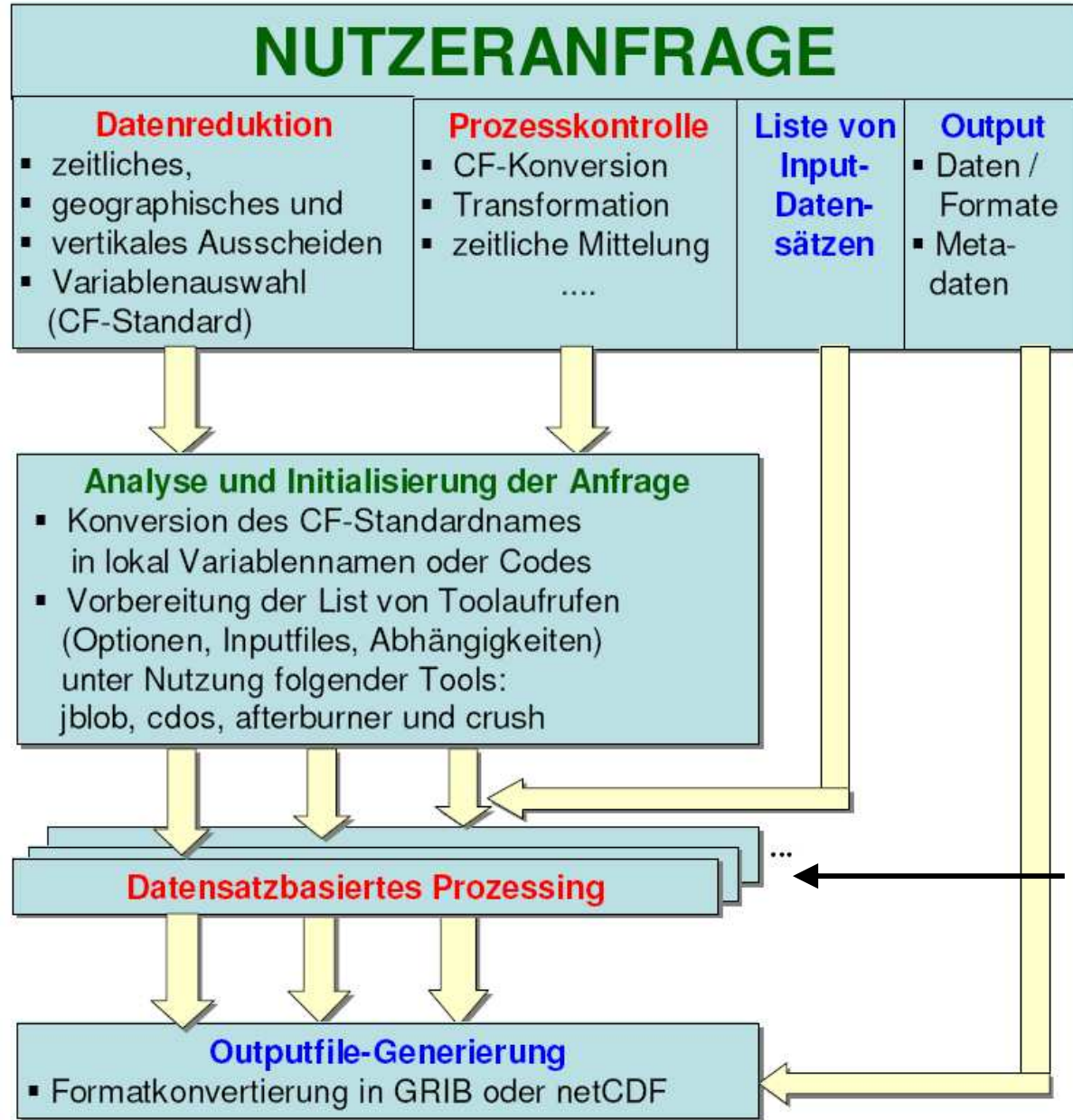
- Official Partners
- KIT with the Institute of Meteorology and Climate Research and the SCC are associated to the project

- n 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 C<sup>3</sup>

n Example

Workflow



Virtualized data center with guaranteed data lifetime and compute resources

## n 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!

n 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.

- ⊠ Can be used as a so called “submit engine”
- ⊠ Is a Grid middleware
- ⊠ used in D-Grid, at KIT in the project CampusGrid on vector computer, InfiniBand cluster and power4 systems from IBM
- ⊠ It's very practical in use!
- ⊠ <http://sourceforge.net/projects/unicore>
- ⊠ It's a must to have a certificate!
- ⊠ What to do à [www.d-grid.de/index.php?id=216&L=1](http://www.d-grid.de/index.php?id=216&L=1)

# Globus Toolkit 4 (GT4)

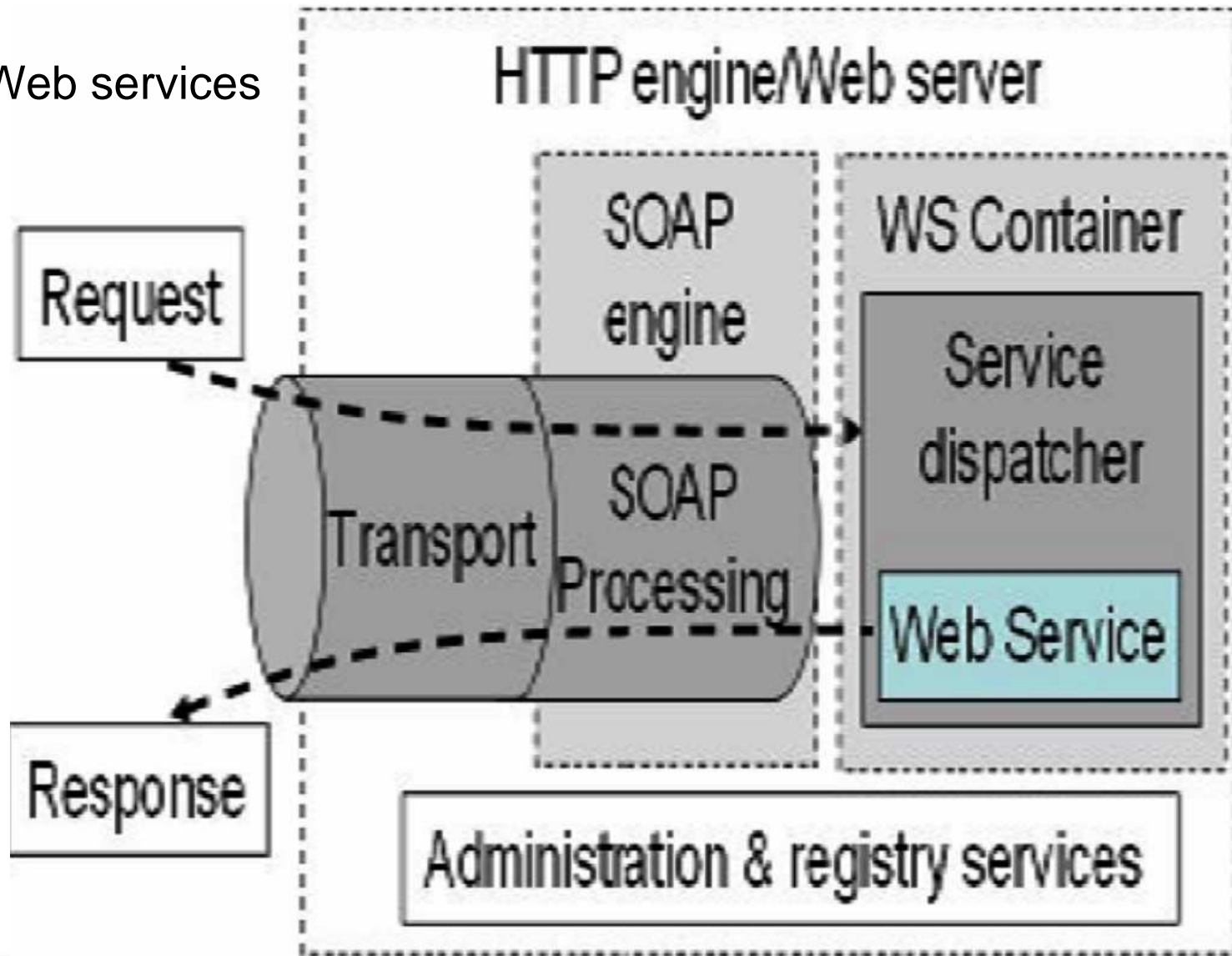
- n "The Globus®Toolkit is an open source software toolkit used for building grids." [[www.globus.org/toolkit](http://www.globus.org/toolkit)]
- n GT4 is a WSRF compliant application development framework.
- n GT4 is collection of ready-to-use of high-level Grid services.
- n Much more a Grid based solution than Unicore
- n It's a must to have a certificate!
- n What to do à  
[www.fzk.de/fzk/idcplg?IdcService=FZK&node=2450&document=ID\\_057021](http://www.fzk.de/fzk/idcplg?IdcService=FZK&node=2450&document=ID_057021)  
[http://gks06.fzk.de/slides/gt4-intro\\_Olaf\\_Schneider.pdf](http://gks06.fzk.de/slides/gt4-intro_Olaf_Schneider.pdf)
- n You have to use a special machine for login
- n It's based on Grid services
- n Will be available soon for the vector system SX8-R

What GT4 provides?

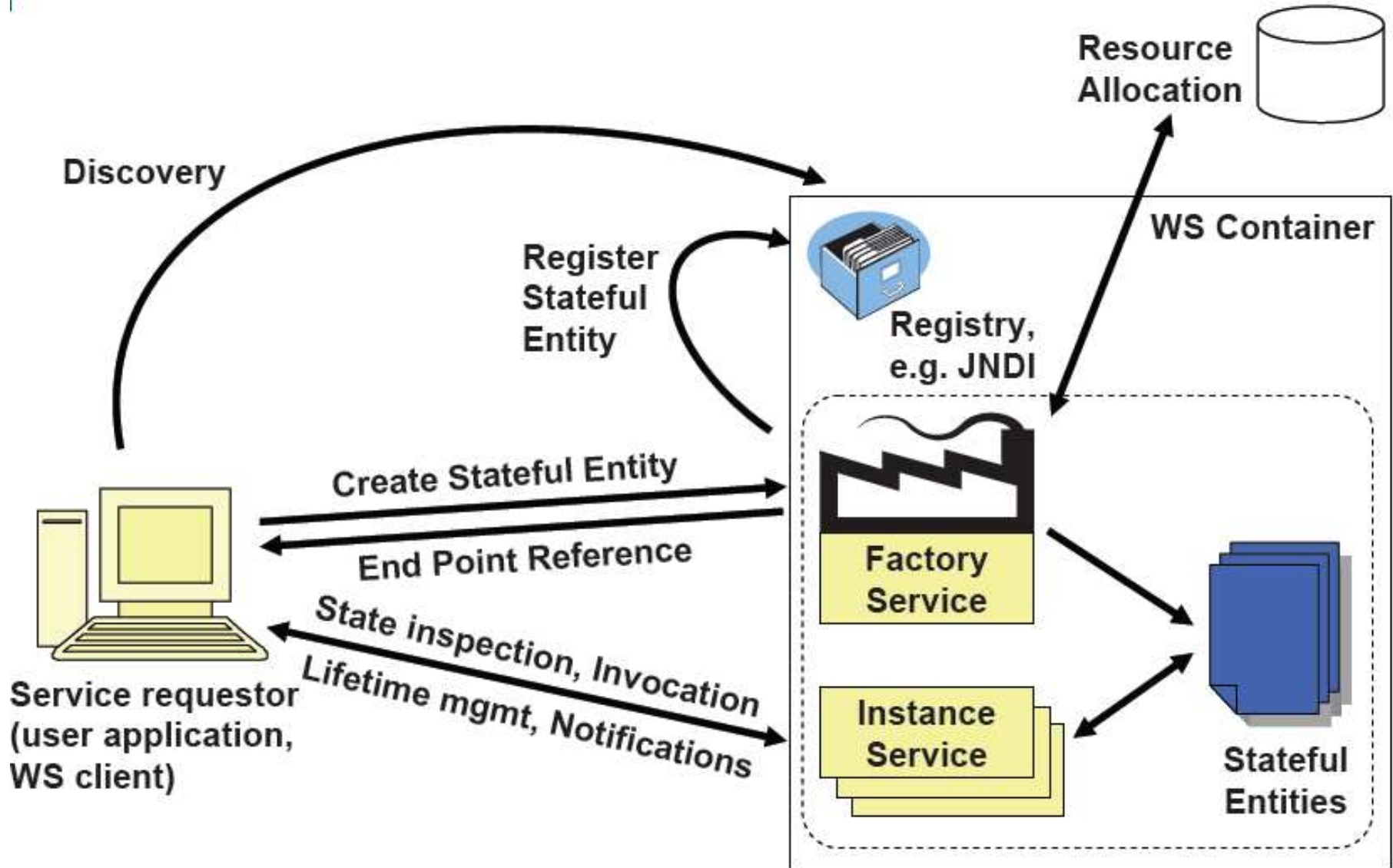
- n **collection of solutions**, which frequently comes up when trying to build collaborative distributed applications
- n implementation of **standards**:
  - use of existing standards (IETF, W3C, OASIS, GGF)
  - provide reference implementations of new/proposed standards
- n a set of **services** focused on infrastructure management
  - common interfaces to heterogeneous resources
- n standard-based security infrastructure
- n client APIs and command line programs for accessing services
- n a **toolkit** for building
  - new Web Services
  - service-oriented architectures (SOA)



- Web services



# Globus Toolkit 4 (GT4)



# Conclusion

- n KIT is the largest research and development organization in Germany.
- n SCC provides a lot of IT resources worldwide.
- n D-Grid is a success story.
- n Applications are not simple to gridify!
- n C<sup>3</sup> is the climate project as part of D-Grid.
- n The future of Grid computing is .....
- n Cloud computing could be a way for standardizing accessing IT infrastructure and to integrate in a quite simple way scientific applications and business applications.
- n Unicore
- n GT4

# Thank you!

# Questions?