Universität Karlsruhe (TH)

Universität Stuttgart

HIGH PERFORMANCE COMPUTING COMPETENCE CENTER BADEN-WÜRTTEMBERG

Scientific Supercomputing Center Karlsruhe

HLRS

Contents

Published and edited by the
High Performance Computing
Competence Center
Baden-Württemberg (hkz-bw)
http://www.hkz-bw.de

The Managing Board

Prof. Dr.-Ing. M. Resch
Universität Stuttgart
HLRS
Allmandring 30
D-70550 Stuttgart

E-mail: resch@hlrs.de

Prof. Dr. W. Juling
Universität Karlsruhe (TH)
Rechenzentrum
Zirkel 2
D-76128 Karlsruhe

E-mail: juling@rz.uni-karlsruhe.de

June 2004

High Performance Computing Competence Center Baden-Württemberg (hkz-bw)

The Ministry of Science, Research and the Arts of the State of Baden-Württemberg, the University of Karlsruhe and the University of Stuttgart have agreed to bundle the personnel and financial resources in the field of high performance computing so that an international competitive offering of high performance computing power can be maintained permanently.

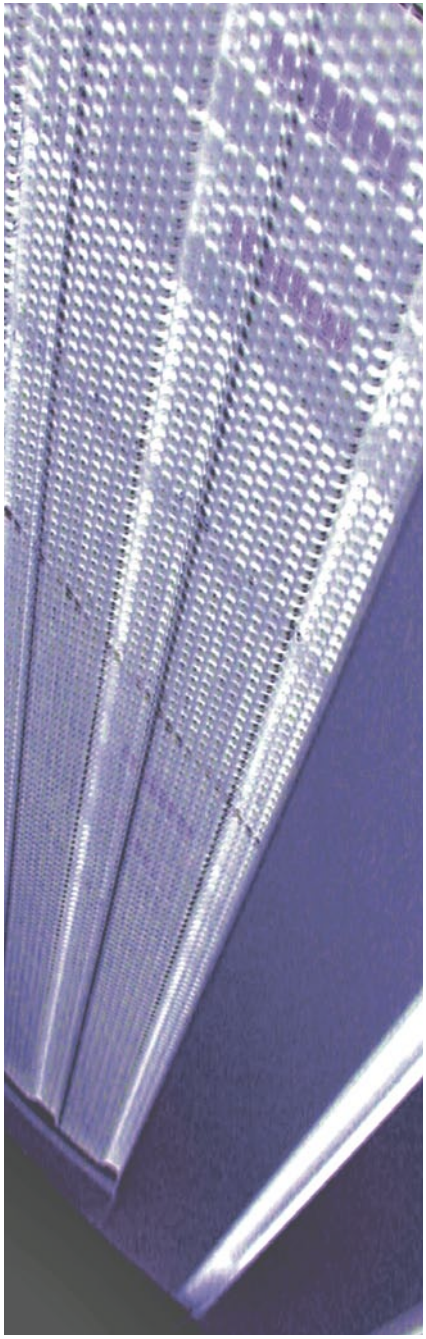
For this reason the Universities of Karlsruhe and Stuttgart have founded the High Performance Computing Competence Center Baden-Württemberg (hkz-bw). Primarily it is established by the High Performance Computing Center of the University of Stuttgart (HLRS) and the Scientific Supercomputing Center of the University of Karlsruhe (SSCK) jointly providing the necessary services.

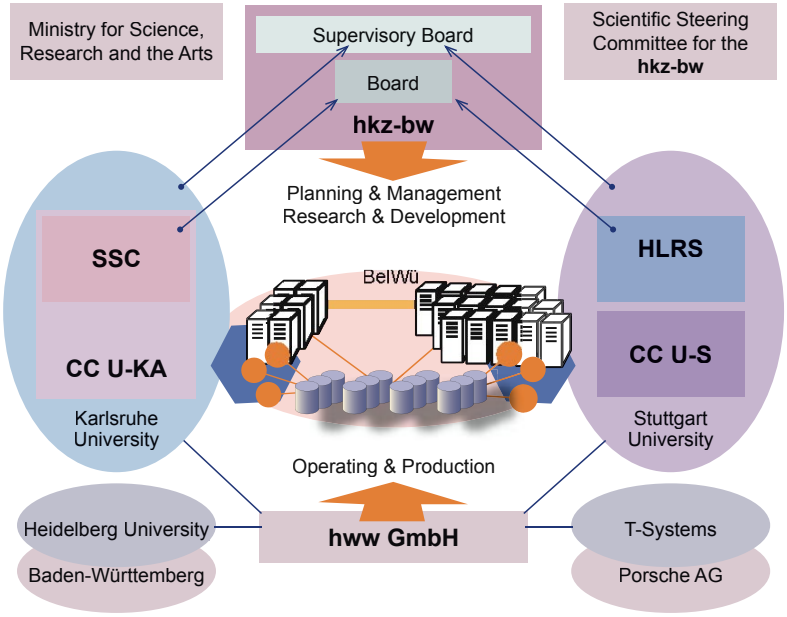
The hkz-bw plans and organizes the utilization of appropriate state resources for science in order to achieve a maximum efficiency by bundling the personnel and financial resources. It performs the tasks of a national scientific high performance computing center by providing scientific users with computing power which cannot be covered locally because of the high system requirements.

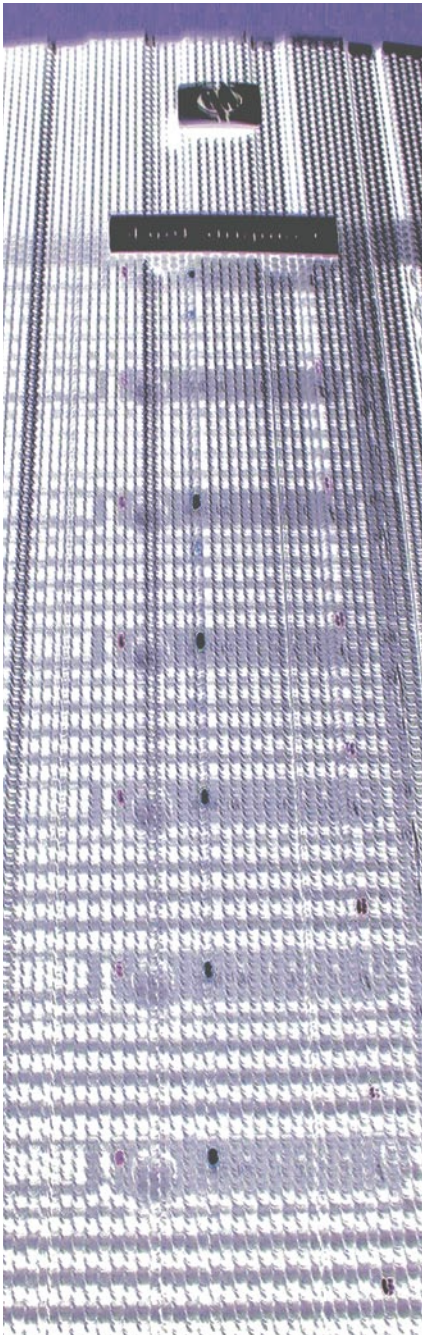
The hkz-bw moreover undertakes the task to further applications for high performance computing and to develop them in close cooperation with the users. In connection with these new developments the center especially pursues the shared utilization of computing resources within the network.

The hkz-bw coordinates the utilization of the resources according to the national and state's scientific steering committee for the allocation of high performance computing capacities. Furthermore the center undertakes the task to coordinate the state's activities with regard to local and national networking.

The hkz-bw is open to further members.







The bodies of the hkz-bw are:

- Managing Board
- Members' Meeting
- Scientific Steering Committee

The managing board is represented by the respective directors of the involved institutions.

The hkz-bw members' meeting consists of two representatives of the concerned universities. It appoints the managing board and decides on all important affairs of the hkz-bw, especially the rules of procedure, the working program, the budget, and the admission of new members.

The Scientific Steering Committee of the National High Performance Computing Center has undertaken additional tasks with respect to the management of the state's high performance computer. Moreover, it advises the managing board of the hkz-bw on the acquisition and utilization of the high performance computers.

The utilization by scientific users is basically effected according to the regulations of the particular university respectively institution. The steering committee decides on the admission.

The systems of the hkz-bw are also available to users from private enterprises against refund of costs via the industrial partners of the hww corporation.

Tasks and Objectives

The hkz-bw undertakes service and development tasks in the field of high performance computing. These are particularly:

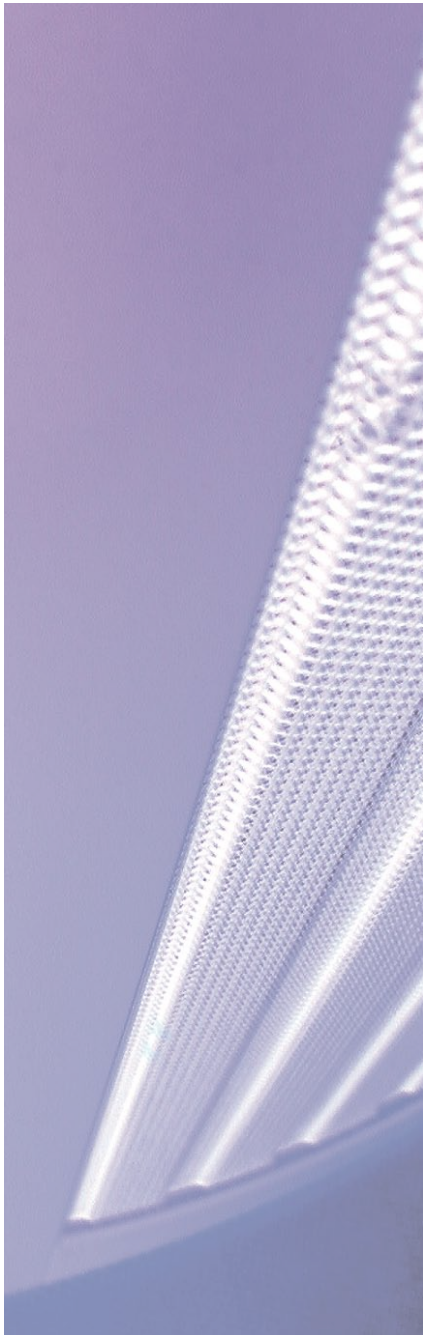
- A continuous and reliable provision of science with high performance computer capacity and the appropriate high performance communication.
- The support of the users for an optimum utilization of the available computer and network infrastructure.
- The coordination and promotion of the utilization of high performance computers and networks.
- The exploration and development of new possibilities of numerical simulation and of new supercomputing applications in cooperation with the users.
- The creation and development of applications, middleware and libraries for high performance computing.
- The cooperation with and support of the research community WiR (Scientific Computing in the State of Baden-Württemberg).
- The promotion of the exchange of experience among the users as well as an active participation.
- The advanced education in the field of high performance computing.

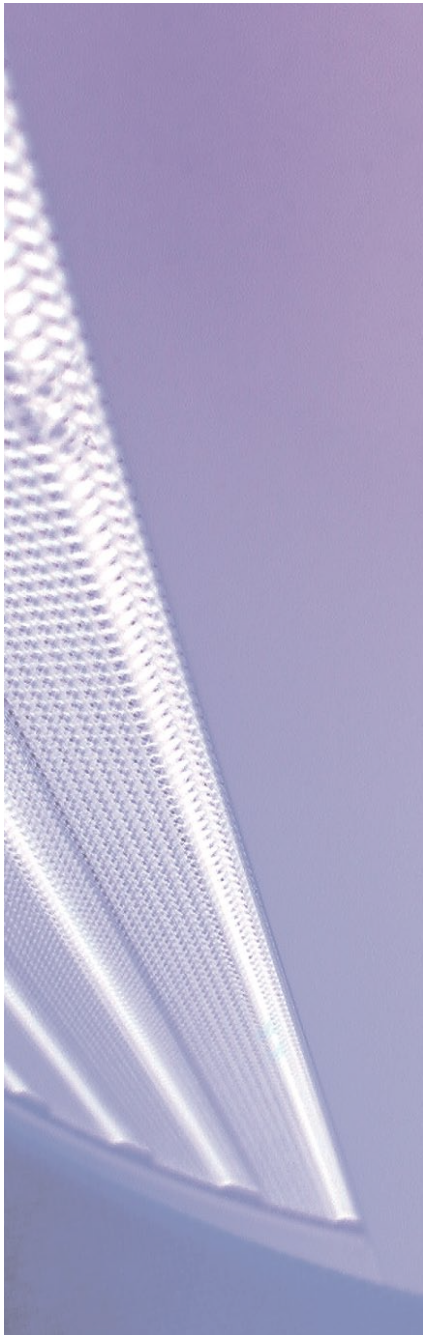
Additional HPC Research Promotion

Within the scope of the hkz-bw additional funds for the development of future technologies in supercomputing are provided by the State of Baden-Württemberg and the involved universities. By these funds further developments in the fields of middleware (especially scheduling and data management), programming environments, tools and application software for grid based supercomputing can be approached as well as new research areas be opened up or supported.

With its program “Modeling and Simulation on High Performance Computers” the Foundation of the State of Baden-Württemberg furthers application-oriented projects, which develop innovative solution methods in their particular specialist area by using the existing HPC infrastructure in the State of Baden-Württemberg. The objective is to establish the methods of HPC beyond the borders of the specialist areas over the medium and long term.

The hkz-bw aims at a close cooperation with these projects in order to develop new application areas for high performance computing.

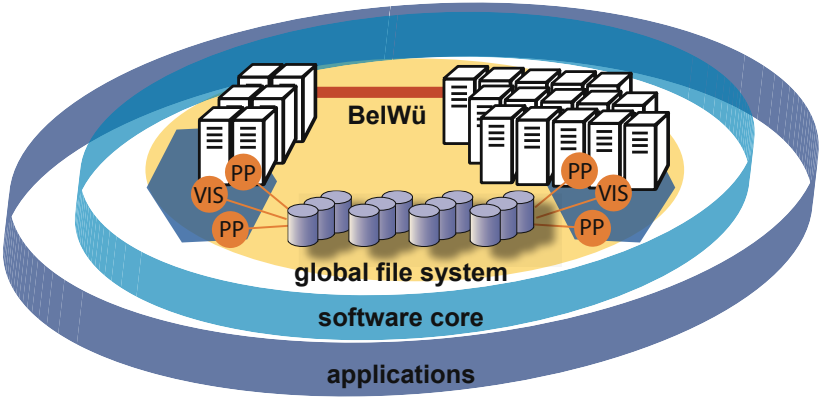




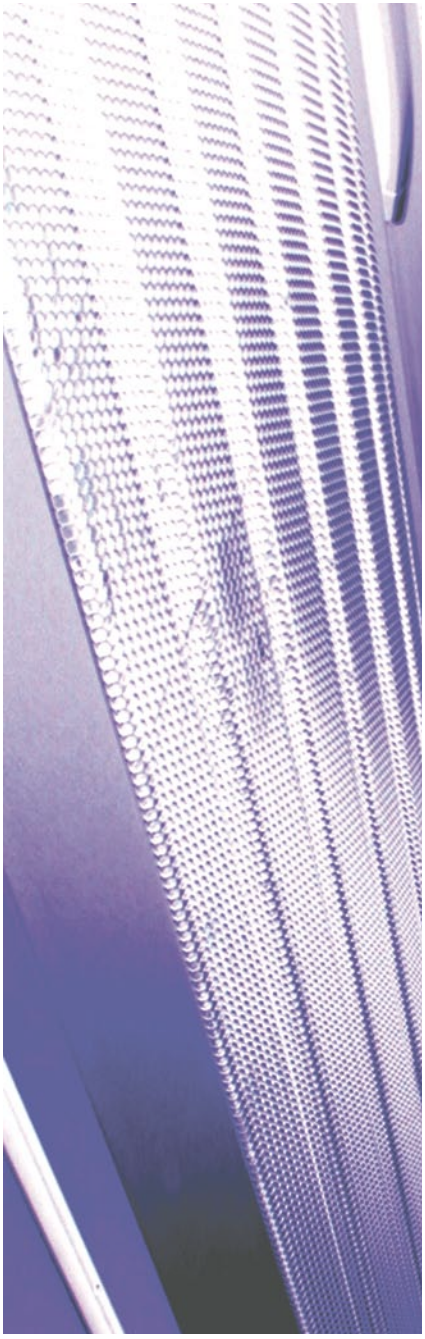
Systems of the hkz-bw

The high performance computer systems of the hkz-bw comprise a supercomputer system of the top category and a high performance computer. Both systems have been chosen in such a way that their different architectures complement each other. So an optimum architecture can be offered for a broad application spectrum. The national high performance computer system, operated at the HLRS, is a vector system, that is characterized by an extremely high storage and communications bandwidth. At the SSCK a high performance computer for the state on the basis of Intel Itanium 2 processors is operated. This system is in particular suited for applications which can profit from the large data cache. Both systems will be operated as a tightly coupled distributed system.

The computers will be directly connected on the basis of the existing network infrastructure in the State of Baden-Württemberg. Thus this installation at the hkz-bw is also to be seen as an example for an ambitious grid computing project lining up with the international initiatives like TeraGrid (USA), NPACI (USA), e-science (UK) or NAREGI (Japan).



The linking of the two systems is physically realized via the fast network connection of the state's research network BelWü with a target bandwidth of 40 Gbit/s. The logical linkage is arranged via a common file system, which from an application's point of view only shows one name space. Within the application the communication between the systems can be performed via the PACX-MPI library developed at the HLRS, which is utilized throughout the world for coupled distributed systems in grid computing. The whole system presents itself to the user as a uniform system with a uniform access. Thereby the hkz-bw infrastructure as part of the D-Grid Initiative integrates "seamlessly" into the future research scenario.



Vector Parallel Supercomputer NEC SX-6X

The High Performance Computing Center Stuttgart (HLRS) and NEC have signed a contract to install a leading edge supercomputing system at Stuttgart in 2004/2005. In an initial phase a medium sized SX-6 system is installed that will later be upgraded to the actual full production system.

Basic Architecture

The supercomputer of the SX-family is based on the well known vector architecture of NEC. The basic building block is a vector processor with the following characteristics:

Clock Rate	**565 MHz**
Parallel Vector Pipes	**2*8**
Processor Speed	**9 GFlop/s**

The theoretical peak performance of the CPU is slightly higher than 9 GFlop/s because of an additional scalar unit. Already this single processor peak performance is outstanding. Eight of these vector processors are integrated into a shared memory system. The main characteristics of such a node are as follows:

Node Speed	**72 GFlop/s**
Total Memory Speed	**288 GB/s**
Memory Speed / Processor	**36 GB/s**
Memory Size	**64 GB**

The outstanding feature is the extremely high bandwidth of the memory. For each operation the system can load or store 4 Bytes. This leads to an extremely high level of sustained performance in the range of 60% for a single processor. The sharing of memory, however, increases costs in a way that a full node operates at a sustained speed of around 50% of the peak speed.
These nodes are interconnected by the well known IXS crossbar switch of NEC. It connects each node to the switch with a bandwidth of 8 GB/s. The MPI communication latency is less than 8 µsec.
The initial cluster of such SMP nodes will consist of 6 nodes so that the initial installation will have the following parameters:

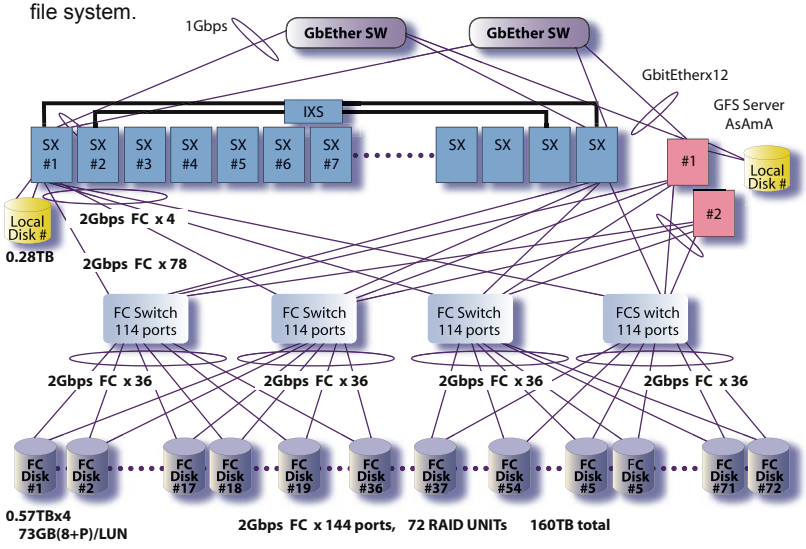
Total Number of Processors	**48**
Total Peak Performance	**434 GFlop/s**
Total Main Memory	**384 GB**
Total Memory Bandwidth	**1.7 TB/s**
Total Communication Speed	**48 GB/s**

This interim system will be used for code porting and software development. Its sustained performance is expected to be about 200-220 GFlop/s which is about a factor of three compared to the old HLRS systems.
The final system will have the same type of architecture. Performance figures for the individual processor will more than double. The number of nodes of the final cluster will be 64 with a total of 512 processors. This will result in a total peak performance of more than 11 TFlop/s and an expected sustained performance of between 3-4 TF/s.

Overall Architecture

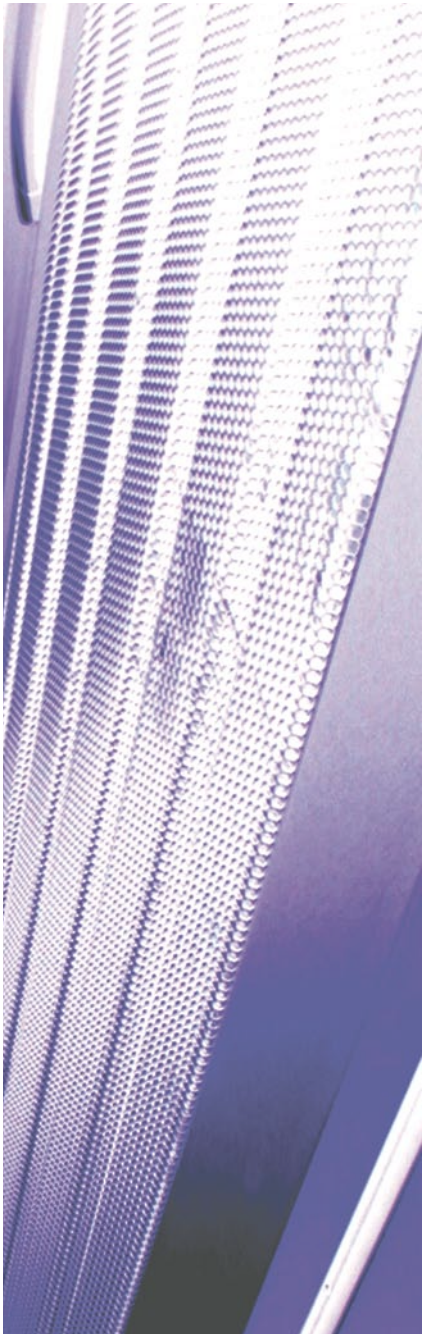
The overall concept consists of the following key parts:

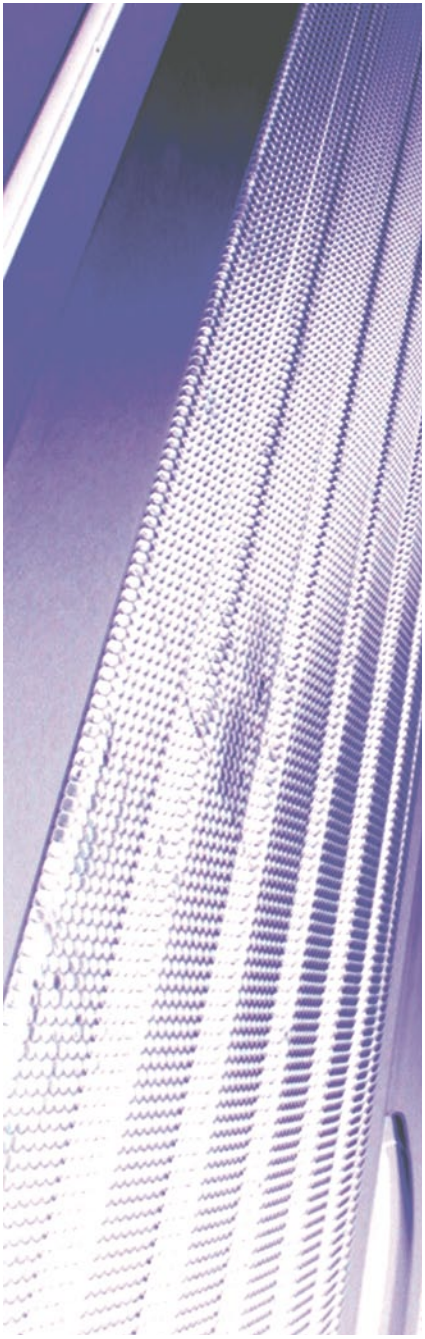
- Compute Cluster: This is the cluster of shared memory vector nodes that is working as the power horse of simulation.
- Pre-Processing Nodes: These are going to be IA64 based 32 processor nodes with a large shared memory in the range of about 0.5 TB.
- Post-Processing Cluster: A cluster for post-processing will be integrated into the concept to transfer the compute load for visualization to an appropriate platform but at the same time keep a close integration of systems.
- Common File System: All three hardware architectures will be integrated through a file system.



Delivery Schedule

The first part of the system was already delivered in March 2004 and is operational since April 2004.
The final system will be installed in two phases starting late in 2004 and is scheduled to be finished by June 2005. The system will then be fully operational by mid of 2005.





Massively Parallel Itanium Supercomputer HP XC6000

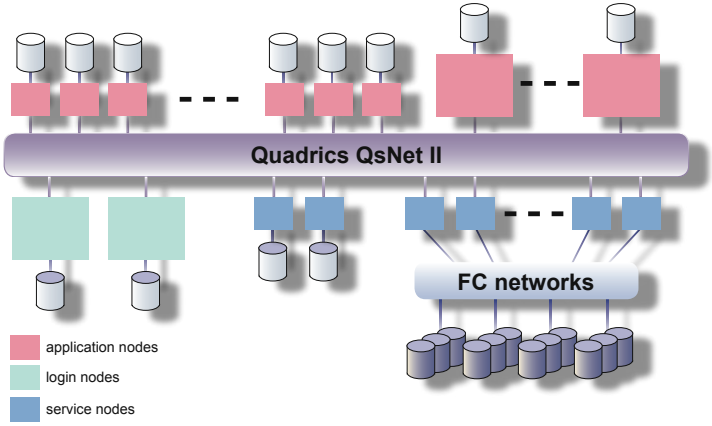
At the SSCK the installation of an HP XC6000 Cluster as high performance computer of the state has started with the beginning of 2004. It will be set up in several stages. The system, operated under Linux, consists of powerful single nodes which contain each two, four or 16 Intel Itanium 2 processors being linked by a fast interconnect (Quadrics QSNet II).

In the final stage of expansion in the beginning of 2006 the redundantly designed, highly available system with 1,200 processors will achieve a computing power of 11 TFlop/s and provide more than 7 TB of main memory.

After the installation of the test cluster with 16 nodes in April 2004 the system expansion of the first phase (end of 2004 / beginning of 2005) will have a peak performance of 2.2 TFlop/s as well as a memory extension of 2.2 TB.

Structure of the HP XC cluster in phase 1:

- 116 two-way nodes with 12 GB main memory each
- 6 sixteen-way nodes with 128 GB main memory each
- Single rail Quadrics QSNet II interconnect
- 10 TB global disk space



In phase 2 (beginning of 2006) this system will be extended by:

- 218 four-way nodes with 24 GB main memory
- Dual rail Quadrics QSNet II interconnect
- 30 TB global disk space

Fast Communications Network and High Scalability

The architecture of the HP XC Cluster is characterized by a clear structure and specialization of the single nodes. Two-way nodes respectively four-way nodes are used for applications being parallelized with MPI. The communications network with a latency of about 3 µs and a bandwidth of about 800 MB/s on MPI level allows a high scalability so that also communication intensive applications with high processor counts can be performed efficiently.

Applications being parallelized according to the principle of shared memory access can be carried out on the sixteen-way nodes and benefit from the shared main memory with 128 GB and the local disk capacity of more than 1 TB per node. The utilization of the sixteen-way nodes is also planned for interactive applications, for pre- and postprocessing as well as for data filtering.

Parallel Cluster File System Lustre

With Lustre the HP XC6000 Cluster possesses a global parallel file system being designed for very large clusters and high I/O bandwidths. By the utilization of several object storage servers (OSS) und meta data servers (MDS) a parallelism of data access and also a redundancy in case of a failure of single servers is achieved.

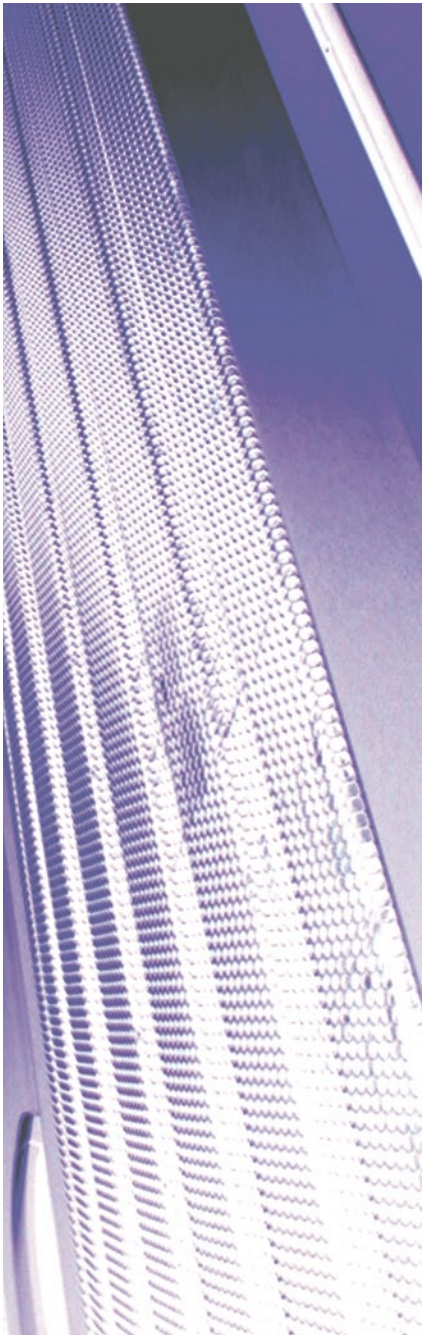
In the first expansion stage 10 TB, in the second stage 40 TB will be available for global file systems. Moreover, every node of the XC Cluster is provided with local disks for temporary files.

High Efficiency by Cache Utilization

The nodes of the HP XC Cluster are based on Intel Itanium 2 processors. These processors are in particular characterized by a high floating-point performance as well as a very large data cache, which is located on the processor chip and therefore can be accessed with a very short latency and extremely high bandwidth. Thus the system is especially suited for application programs being optimized for cache utilization.

Simple Porting of Application Programs

The programming environment and application interfaces of the HP XC6000 system are based on open standards and therefore permit a simple porting of application programs.



High Performance Computing Center Stuttgart (HLRS)

Based on a long tradition in supercomputing at the Universität Stuttgart, the HLRS was founded in 1995 as the first national center for high performance computing in Germany. Since then it has assumed an internationally renowned position in supercomputing. The National Science Foundation Award for High Performance Distributed Computing in 1999 and the HPC Challenge Award 2003 have emphasized this again. Today, HLRS provides services to researchers at universities and research laboratories in Germany and at the same time does a great deal of research in the field of modeling and simulation.

One of the key goals of the HLRS is the provision of the best system for each application. Hence, the decision for any hardware provided is always driven by demand of the users aiming to provide the best solution for a given problem. In consequence HLRS sees its role in evaluating the strengths and weaknesses of existing systems. At least as important is the feedback that HLRS is able to provide to its industrial hardware vendors in intensive partnerships. Thus HLRS has helped to improve HPC-systems by bringing to bear the requirements of simulation in the hardware design process.

On top of this HLRS is as a provider for simulation solutions. This requires taking the lead in bringing together various disciplines working towards solutions in research and engineering. So HLRS has become a breeding place for innovative simulation. Not the least because it has successfully managed to integrate means of virtual reality visualization and supercomputing providing the user with a real simulation workbench that virtually puts the scientist directly into the simulation process.

The NEC-HLRS Teraflop-Workbench Cooperation (www.tflop-workbench.de)

To improve the quality of services and research HLRS works closely with hardware and software vendors. The most important cooperation is the joint program with NEC called the Teraflop-Workbench Partnership.

The aim of this initiative is twofold. On the one hand it is designed to help close the gap between peak performance and sustained performance in supercomputing. On the other hand supercomputers have ceased to be stand alone large systems that inhale huge amounts of data to exhale even larger amounts. As a standard tool current systems are integrated into a workflow of pre-processing, simulation and post-processing. This workflow itself is integrated into the overall workflow of either science or engineering. This puts the supercomputer into a heterogeneous environment of hardware and software. To handle this complexity and heterogeneity is the second part of the cooperation.

High Performance Computing Center Stuttgart (HLRS)
Universität Stuttgart
Allmandring 30
D-70178 Stuttgart

Phone: +49 (0) 711 685 2504
E-mail: resch@hlrs.de
http://www.hlrs.de

Scientific Supercomputing Center Karlsruhe (SSCK)

The Scientific Supercomputing Center Karlsruhe is the service unit of the University's Computing Center that cares for supercomputer customers.
The SSCK will provide support for experts as well as for novices in parallel computing. Our goal is to help our customers in all problems related to scientific supercomputing. Our services are not only confined to advice on how to use supercomputers efficiently but you will also get qualified help if you are looking for appropriate mathematical methods or simply having problems to login.

Since the first supercomputer was installed in Karlsruhe in 1983 the Computing Center of the University has been continuously engaged in supercomputing. Besides the operating of the machines Karlsruhe has always been establishing expert groups for scientific supercomputing. In the 1980s and 1990s experts of the Computing Center in Karlsruhe tuned numerical libraries for vector computers as well as microprocessor based parallel supercomputers. From this time on close cooperation with other institutes of the university and industrial companies has been initiated.

At the Computing Center solvers for arbitrary systems of nonlinear partial differential equations and iterative linear solvers have been developed. Thus, the experts at the SSCK know about the needs and problems of their customers from their own experience.

High Performance Technical Computing Competence Center (HPTC³)

Together with Hewlett Packard and Intel, the SSCK has established the High Performance Technical Computing Competence Center (HPTC³). The main goals of this partnership are the further development of HP XC systems to make it even more usable and reliable for the customers of hkz-bw. This includes development, test and early deployment of advanced tools supporting the application development process. In close cooperation with end users and independent software vendors the HPTC³ will extend the portfolio of application packages that is available on HP XC6000 clusters. The tuning and optimization of applications to reach highest possible performance is another challenging goal of HPTC³.

Scientific Supercomputing Center Karlsruhe (SSCK)
Universität Karlsruhe (TH)
Rechenzentrum
Zirkel 2
D-76128 Karlsruhe

Phone: +49 (0) 721 608 8011
E-mail: ssc@rz.uni-karlsruhe.de
http://www.rz.uni-karlsruhe.de/ssc/

High Performance Computers for Science and Industry Corporation (hww)

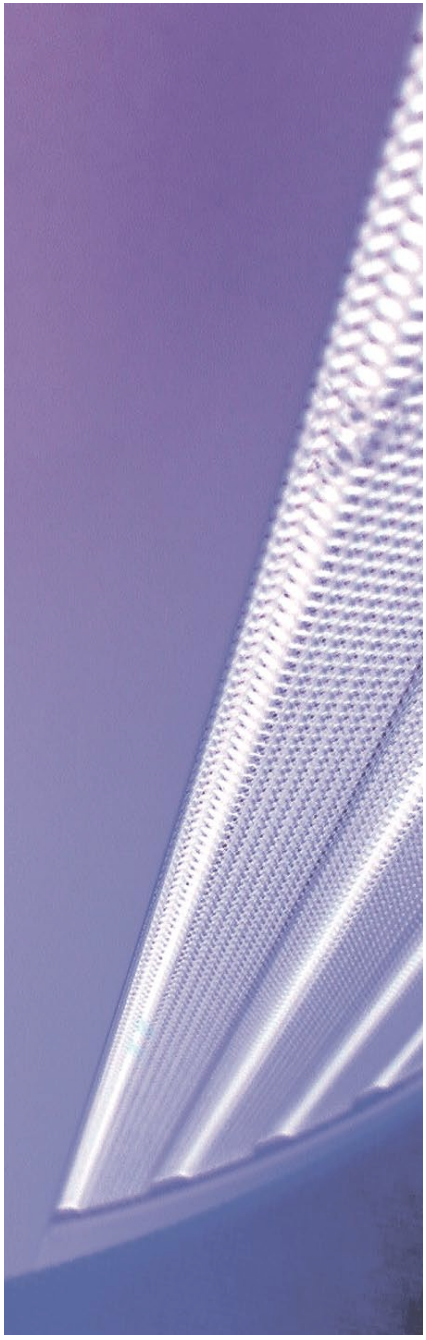
The operation of the hkz-bw systems as well as the utilization by industry is organized by an institutional cooperation of industrial and academic partners. The hww corporation (High Performance Computers for Science and Industry Corporation) was founded in 1995 by debis Systemhaus (dSH, now T-Systems), the Porsche AG, the University of Stuttgart and the State of Baden-Württemberg. Immediately after that also the University of Karlsruhe was affiliated to the corporation. The University of Heidelberg as well as the joint venture of T-Systems and of the German Aerospace Center DLR, dSH-Solutions for Research (SfR) joined the company later. The industry and the public sector are represented with equal shares.

Problem Specific Computer Platforms

The aim of the hww, which operates the common university and industry systems by order of the hkz-bw, is the achievement of synergy effects concerning the acquisition, utilization and operation of high performance computers. So a powerful system of the same architecture instead of several medium-sized systems permits all partners the solution of complex problems and thereby an improved competitiveness. Moreover, a broader spectrum of computer architectures can be made available to each partner. In this way the most powerful and cost efficient computer platforms can be offered. Furthermore, a reduction of cost is achieved by economies of scale and a more efficient assignment of personnel.

Another advantage of this cooperation is the know-how transfer from research to industry being realized by a series of common projects.

The access of industry customers is effected by the organizational structures of T-Systems, which acts at the market as application service provider (ASP) via the web portal http://www.hpcportal.de/.



hww GmbH
Fasanenweg 9
D-70771 Leinfelden-Echterdingen

Phone: +49 (0) 711 972 42 894
E-Mail: hpc@t-systems.com