Karlsruhe Institute of Technology

SX-9, a good Choice ?

Steinbuch Centre for Computing (SCC)

Forschungszentrum Karlsruhe
in der Helmholtz-Gemeinschaft

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

Introduction

- Karlsruhe Institut of Technology (KIT)
- Steinbuch Centre for Computing (SCC) at KIT
- The SCC environment
- Why SX-9
- SX-9 problems!
- Benchmarks
- Do we have a software or hardware problem?
- Further investments to be done by NEC
- Conclusion

Karlsruhe Institut of Technology

Universität Karlsruhe (TH)

+

Forschungszentrum Karlsruhe

=

Karlsruhe Institute of Technology

CC of the University

+

IWR of the Research Centre

=

SCC

After 1.7.2009, the Karlsruhe Institut of Technology will be a legal entity.

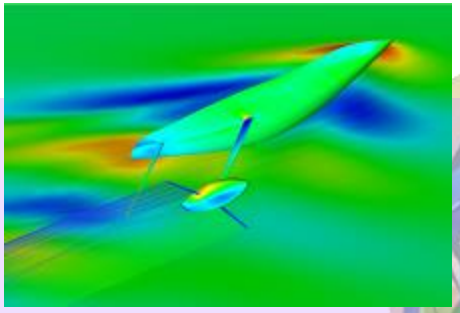
Steinbuch Centre for Computing (SCC)

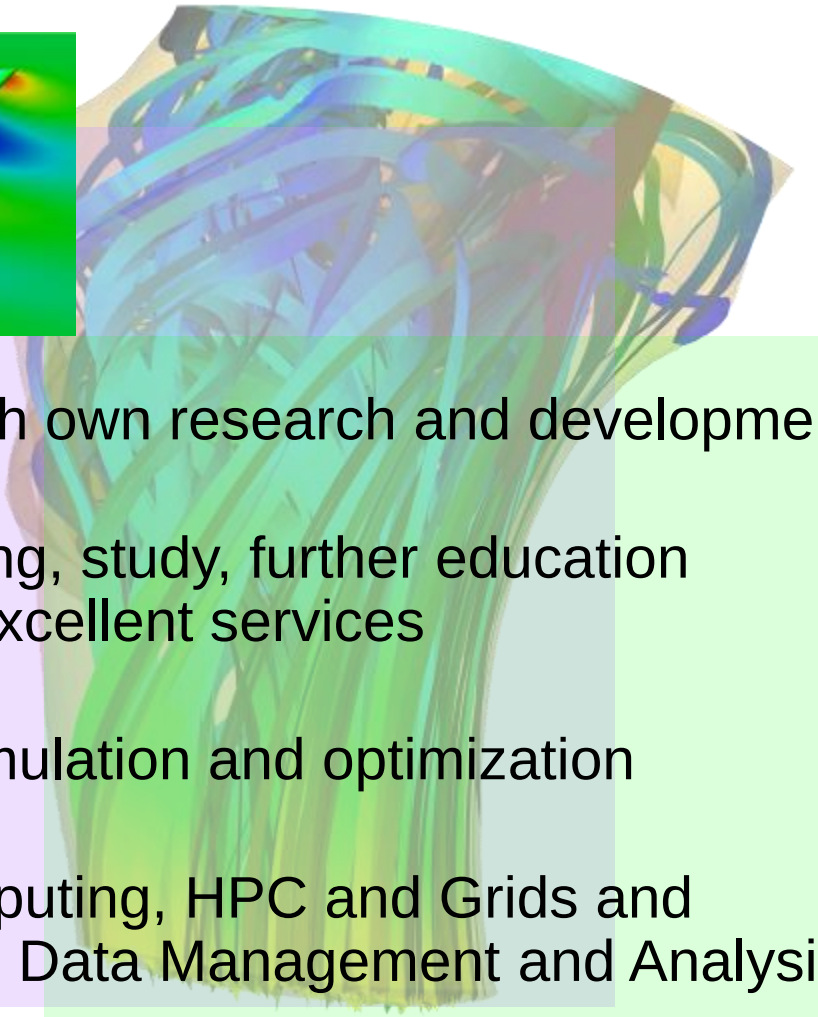
SCC at
Karlsruhe University

SCC at
Research Center

- Founded on January 1st, 2008
- Information Technology Center of KIT
- Merger of the Computing Center of Karlsruhe University and the Institut of Scientific Computing of the Research Center Karlsruhe
- One of the largest scientific computing centers in Europe

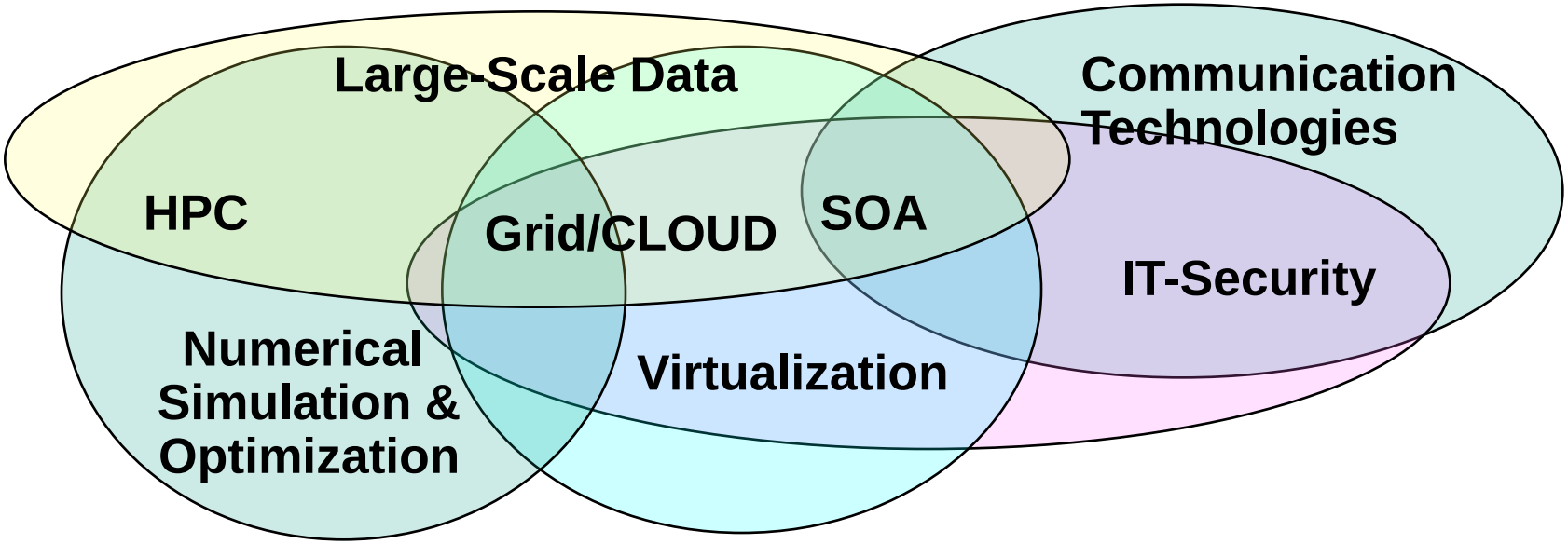
Mission of the SCC





- IT-Services under one roof with own research and development
- Promotion of research, teaching, study, further education and administration at KIT by excellent services
- Major center for modelling, simulation and optimization
- Leading role in Scientific Computing, HPC and Grids and Clouds as well as Large Scale Data Management and Analysis

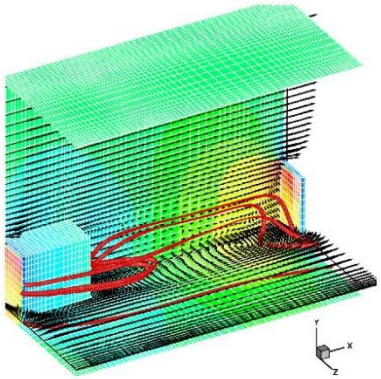
SCC – Competencies and Synergies

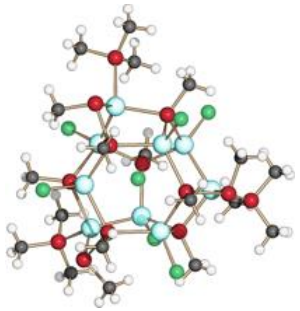


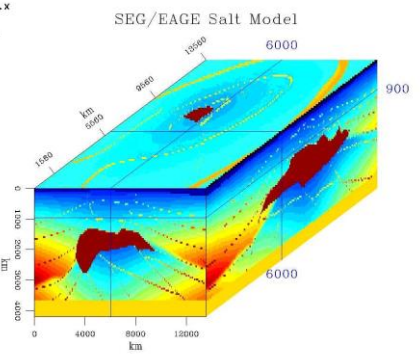
- Synergies between R&D and IT-Services:
  - IT-Management & Process Integration
  - Distributed Systems, Grid & Virtualization
  - IT Security & Service Mgmt., innovative Network Technologies
  - HPC & Cluster Computing, Visualization

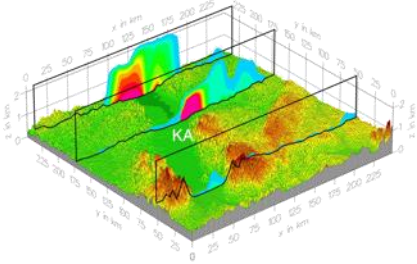
SCC – Main Topics of Research

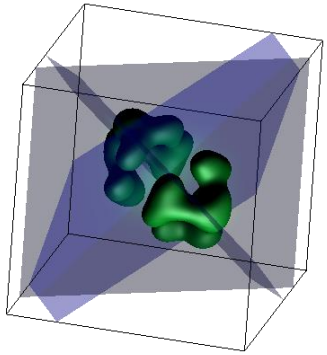
- High Performance Computing,
  Cluster Computing,
  HW-aware Computing
- Grid und Cloud Computing,
  Virtualized Service Environments,
  Service-Oriented Architecture – SOA
- Large-Scale Data Management & Analysis,
  Large-Scale Data Facility
  (e.g. WLCG, Genom Research Data)
- Energy-efficient large System Environments
  (Green IT)
- Process Integration and
  Service Management





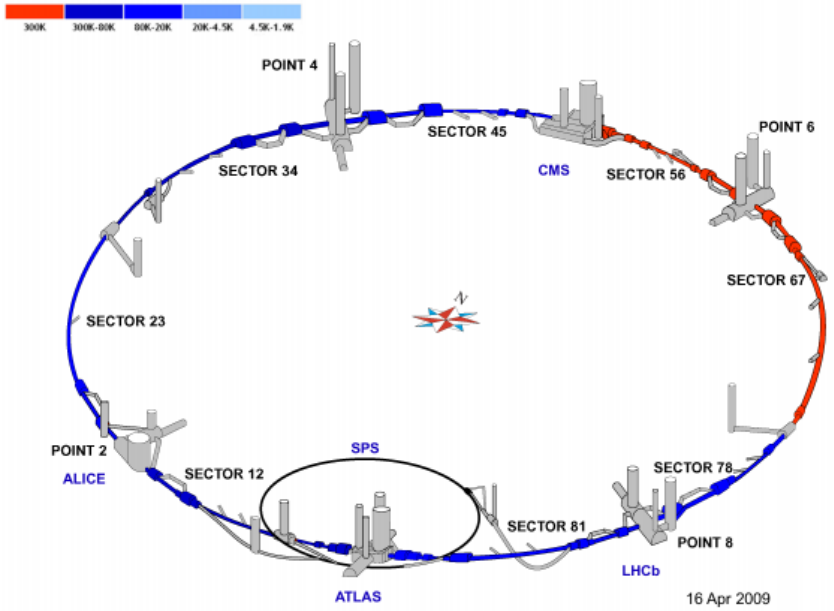






Environment at SCC

- HPC-Clusters based on Intel or AMD processors with InfinBand or Quadrics interconnects (XC1, XC2, IC1, OPUS^{IB}, bwGRiD, ~ 9000 cores)
- GridKa cluster for the LARGE HADRON COLLIDER (LHC) project at CERN (~ 8500 cores, PBytes of disk storage and tape, no InfiniBand interconnect)



- 2 x SX-8R each running 8 processors, 256 GByte main m
  a SX-9 system, 16 processors, 1 TB main memory
  for all three systems 80 TByte GFS disk storage
- A Power6 Blade system running AIX with 112 cores and IB (JS22)

Why SX-9

- A large memory node 1 TB and 16 processors, but 16 processors is not a must, better 8, but also 4 could be enough!
- Application CPU-performance should be a factor 2-3 compared with SX-8R, should be realistic because SX-8R has a memory bottleneck! (100 GFlop/s peak for SX-9 compared with 35 GFlop/s of SX-8R)
- Scalar unit should be a factor 2 faster on SX-9.
- Power consumption should be less
- Floor space should be less
- Seamless integration into the existing SX-8R world including the front-end environment, technology upgrade of the SX-8R
- sustainable development for the KIT-users up to 2011
- Pre-stage for users to go to larger NEC vector systems

SX-9 problems!

- Five processors of the SX-9 node fails in the first quarter of 2009.
- For changing the failed processor you have to
  - shut-down the system
  - repair and add some cooling liquid
- Scalar CPU-performance is worse than assumed
- Vector performance achived is bad compared with SX-8R
- SUPER-UX 17.1 on SX-8R and SUPER-UX 18.1 on SX-9
- For the users we have at the moment two different front-end systems, two different environments, different schedulers.
- Air cooling problems (35 kW), but we have enough cool water!
- Cooling problem with one memory hub (last Sunday!)

Benchmarks

- Triad (d=a+b*c)

Vector length	SX-8R (vector)	SX-9 (vector)
100	0.07s	0.08s
1000	0.31s	0.30s
10000	2.97s	3.20s
100000	30.05s	31.90s
1000000	302.30s	319.00s

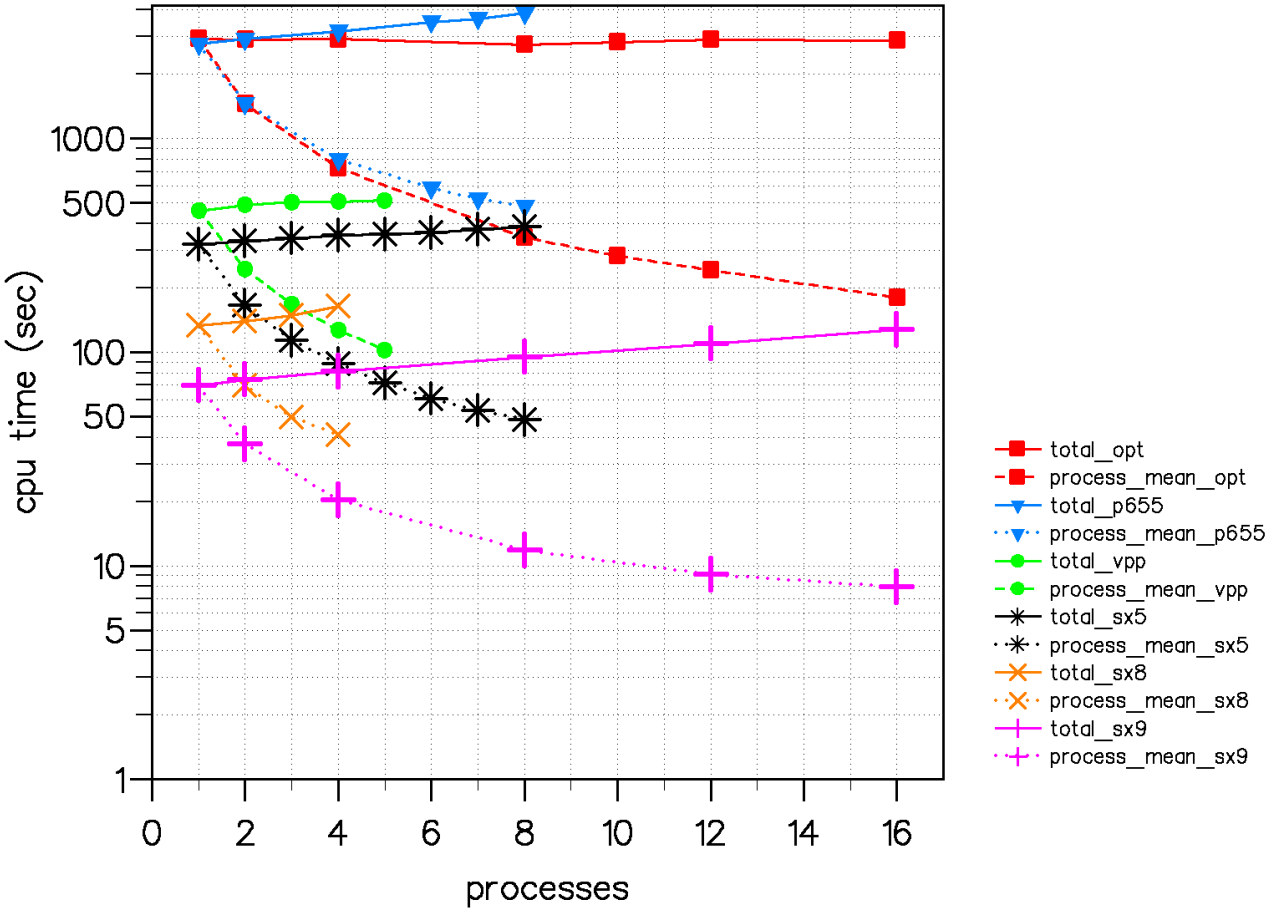
- Solution of a band-matrix

	SX-8R	SX-9
vector	15.8s	11.3s (5.7s)*
scalar	716.9s	925.6s

* you can't use the code on SX-8R, without ADB!

Benchmarks

- COSMO/LM weather code on different systems:



Do we have a software or hardware problem?

- We have a slow scalar unit for SX-9. 👎
- Compiler is not using the ADB automatically, directives are needed. We don't know the information, the compiler is writing into ADB → tuning guide 👎
- It seems, the libraries are not optimized for SX-9 yet! 👎
- Now it looks like a stable hardware. ☺
- The newest Super-UX is not supported on SX-8R 👎
- KIT users have jobs with a vector rate of 94% up to 99% with average vector size of more than 1000. But in all cases it isn't good enough to reach a speed-up > 2 against the SX-8R 👎
- Memory usage near the whole shared main memory and heavily communication results in growing elapsed time for all jobs on the node. 👎 → bank conflicts

Do we have a software or hardware problem?

- DGEMM should run with more than 90 GFlop/s on a single CPU (Schoenemeyer, November 2008), but only 84 archived! 👎
- Tuning-guide for SX-9 needed?

Further investments to be done by NEC

- Speed-up the existing scalar units at costumer sides by a factor of two.
- Further investigation in the used compiler technology:
  - ADB techniques
  - tuning handbook
- Harmonize the SX-8R and SX-9 environment, same compiler versions, libraries, operating system, scheduler, ….
- Supporting customers by the NEC benchmark team.
- Cool the system using the onside cold water installation.
- Better explanation of hard- and software and the interference.

Conclusion

- Why customers should use NEC vector systems in the future?
  - Because of reliability?
  - …… GFlops/Watts?
  - …… GFlops/m^2 floorspace
  - …… serial application performance?
  - …… general purpose use?
  - …… main memory?
  - …… price?

  - …… TCO of the solution for the main applications?

Thank You for the Attention