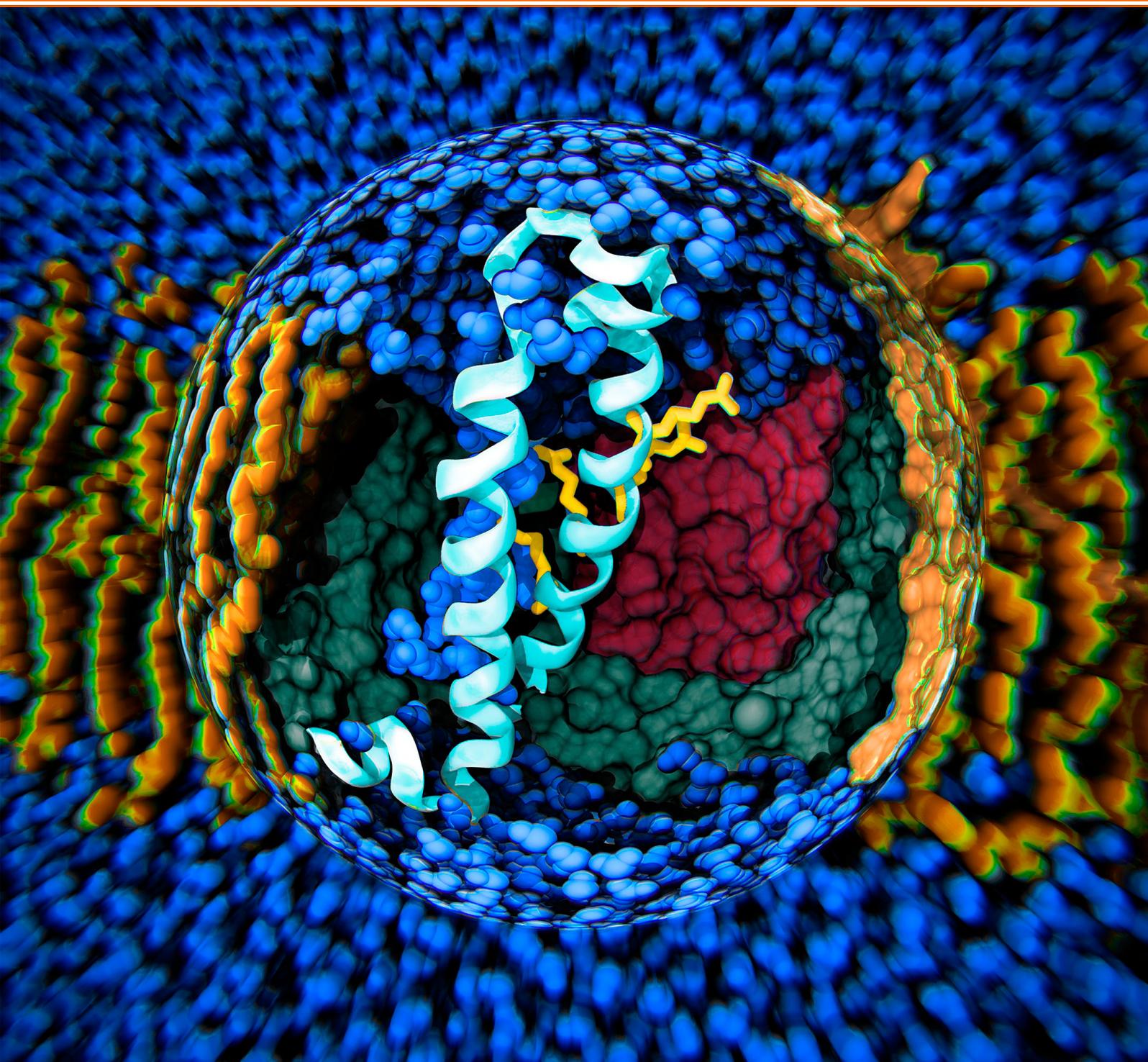


SeRC

Swedish e-Science Research Centre

News

2011:1



The first issue of the SeRC newsletter, presenting the research, goals, visions and management of SeRC

pages 2-6

First two years of SeRC

This is the first Newsletter from the Swedish e-Science Research Centre (SeRC) and this issue will give you an overview of what is happening in the centre. SeRC is the leading e-Science centre that came out of the Swedish government's initiative to substantially increase research funding for selected strategic research areas within Sweden, one of them being e-Science. It consists of researchers from four universities, KTH which is the leading partner, Linköping University, Stockholm University and Karolinska Institutet, as well as the supercomputer centres PDC and NSC. I am very happy to report great progress during the first two years of the centre's existence. We have formed ten e-Science communities in which researchers within specified areas collaborate and coordinate their research. In each community several new research projects have been started and faculty has been hired. A total of about 40 projects have been defined and 13 new faculty members have been hired during the first two years. From 2012 the full SRA funding of 30 million SEK per year is available and will be spent on novel e-Science research. SeRC includes about 140 researchers, including PhD-students.

We have gathered all of the SeRC members twice already, first during the Kick-off at Häringe Slott April 2010 and second at the yearly conference in Tammsvik May 2011. These events allowed researchers to meet across disciplines, and together with the community meetings, have been a great inspiration for SeRC researchers to focus on common e-Science aspects in their research. At the Tammsvik meeting we also convened our prestigious advisory board for the first time and received valuable guidance on how to proceed with the development of the centre.

In addition to the new faculty and research projects, there has been much added value already as a result of the formation of SeRC. We have received the honour of organizing the IEEE e-Science conference in Stockholm 5th – 8th of December 2011; we have been successful in the bid to participate in several EU-projects in the e-Science area, such as CRESTA and ScalaLife; we have started a Flagship programme coordinating several of our research projects dealing with management and use of data related to cancer research; we have taken the initiative to create a national portal for coordination of the new e-Science application experts at the computer centres; we signed an MOU between KTH and Linköping University outlining a common strategy for PDC and NSC, including coordination of activities and complementary procurement of computer hardware; and we have through a common funding from the Swedish Research Council and a large contribution from KTH put in place the first Tier 1 computer in Sweden within the European PRACE infrastructure.

You can read much more about this in this newsletter and on our website www.e-science.se.



DAN HENNINGSON
SERC DIRECTOR

e-Science

Traditionally, scientific inquiry has been divided into theory and experiments. Since the introduction of electronic digital computing a new approach based on computation has emerged. This has become a pervasive approach, where data handling and computation can be the core even of applied research, without a pre-existing theory to confirm. Such scientific research, which is carried out by systematically using advanced computer based tools, is collectively known as e-Science. Some of the most high-profile examples of this involve climate research and the human genome project, but simulations are also used to directly generate new knowledge e.g. in fluid dynamics and materials science.

The success of e-Science is based on the rapid development of novel computer technologies including distributed systems, high-performance computing facilities, visualization equipment and fast access to huge datasets, high-throughput communications, as well as new methods and algorithms. Further improvements will require continued effort to build high confidence e-Science tools, adapt tools to architectural changes, build the supporting infrastructure and understand how the tools can be most efficiently utilized to enable scientific discovery. In addition, we believe that there is a need to integrate the new massive-data-centric approach with our traditional strengths in computation and algorithms, harvesting the huge potential for e-Science to expand beyond traditional fields.

Thus, e-Science has brought about a shift in paradigm and the computer based research is playing an increasingly important role. Computer simulations and the handling and processing of enormous data are now established as new ways to acquire knowledge.

Our Vision

Through e-Science enable world leading research within strategic areas

This vision is realized by:

1. The formation of e-Science communities in which we promote the collaboration between e-Science applications, core e-Science and application experts at PDC and NSC.
2. The formation of a strong e-infrastructure node in the European HPC landscape through close collaboration between PDC and NSC.
3. Interfacing with industry and society in order for the research to be of strategic relevance.
4. The promotion of an e-Science curriculum at our universities.

SeRC: The Swedish e-Science Research Centre

SeRC – the Swedish e-Science Research Centre – was formed by the universities in Stockholm and Linköping (KTH, KI, SU and LiU) around the two largest high-performance computing centres in Sweden (NSC at LiU and PDC at KTH). This project has brought together a core of nationally leading IT research teams (tool makers) and leading scientists in selected strategic application areas (tool users).

SeRC consists of three parts, the SeRC communities, the supercomputer centres and the application experts. The SeRC communities support collaboration on several levels. Each of the communities comprises computer experts, e-Science method developers and scientists from application areas who jointly run e-Science projects. The projects are characterized by novelty in terms of technology (high end computing, novel architectures, grid, databases, etc.), methodology (new theories, models, methods, algorithms and software) and application (application areas with large potential gains from e-Science tools). SeRC has started ten e-Science communities, four communities focusing on core e-Science methods: distributed and parallel techniques, data management, visualization, and numerical analysis; and six applied e-Science communi-

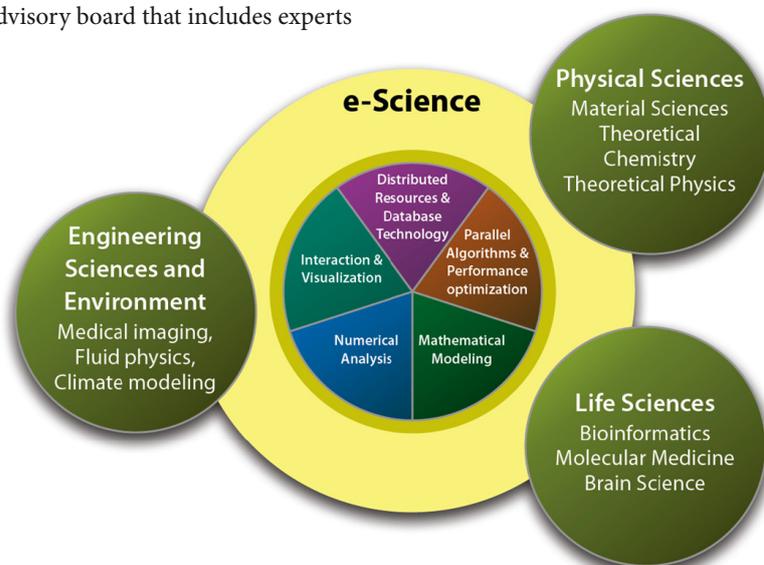
ties: bioinformatics, complex diseases, molecular simulation, climate modelling, FLOW and electronic structure.

SeRC is hosting a majority of the Swedish e-Science infrastructure through the supercomputer centres PDC and NSC. Collaboration between the supercomputer centres and the SeRC communities are strengthened by means of the application experts. These provide e-Science expertise within a certain application area, e.g. electronic structure.

To ensure that commercialization and non-academic use are dealt with at the highest decision making level, SeRC has an advisory board that includes experts

from the private sector alongside international science and e-Science experts. SeRC will also have industry representatives taking an active role in the e-Science communities.

In summary, SeRC, as a Swedish e-Science Research Centre, constitutes a leading visionary e-Science node with a national scope and strong international ties. The four partner universities are together committed to strengthening and shaping the emerging e-Science landscape in Sweden and to give research in this field clear priority in their strategic plans.



Infrastructure

The SeRC infrastructure is based on a collaboration between the National Supercomputer Centre (NSC), LiU, and the PDC Center for High Performance Computing, KTH. Through this alliance and a substantial increase in advanced support staff, the goal is to turn PDC and NSC into comprehensive e-Science enablers. Through PDC and NSC, and together with the Swedish National Infrastructure for Computing (SNIC), SeRC takes a national responsibility for major parts of the Swedish e-infrastructure, by

- Hosting large-scale HPC resources, e.g. the new PRACE computer.
- Financing application software engineers.

A joint NSC and PDC venture is the only Swedish constellation that will have a chance of playing a major role on the European infrastructure arena. The collaboration is initially focusing on areas of importance to the users. One such area is

harmonized user environments, making it easy for the user to simultaneously use computer resources at the two sites. Joint projects and task forces will be set up to achieve common goals and ensure reciprocal knowledge transfer.



Frontpage: A simulation of the voltage-sensor helix of the potassium ion channel. For every heart beat or nerve signal in your body, this small molecule moves to open a channel and transport ions through a cellular membrane. SeRC is internationally leading both in applications and development of widely distributed simulations tools such as the GROMACS toolkit used for this work. Visualization assistance provided by Jyrki Hokkanen, The Finnish IT Center for Scientific Computing (CSC).

Application Experts

SeRC has hired application experts working closely with researchers in the e-Science communities and with programmers and computational scientists at PDC and NSC. The goal is to provide support and enable better use of the e-Science resources, e.g. advanced code optimization, parallelization, choice of architectures. The application expert support is mainly in specific key areas where there is both a significant need for technical assistance among current SeRC researchers, direct impact on strategic research areas and clear synergy effects

from corresponding co-investments in the research part by SeRC.

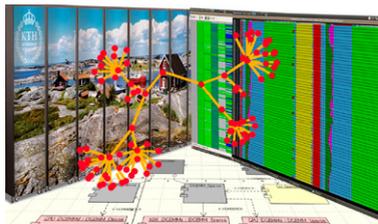
To realize a national application expert infrastructure we believe it is inevitable to achieve a critical mass of application experts that work together in complementing areas, rather than diluting the effort and have scattered experts working with individual groups. In particular, although several groups in SeRC develop internationally widely used codes that could qualify for specific support (GROMACS, DALTON, Simson, PENCIL, EMTO etc.), we believe

it is better to integrate all experts in a joint support infrastructure, even if they spend significant time on some of these codes.

The application experts work in close contact with relevant Swedish research groups, independent of their affiliation, spending time both at NSC/PDC and together with the researchers. With a larger infrastructure we will also achieve flexibility where we can support new projects on relatively short notice, as well as efficient sharing of knowledge between adjacent areas. The SeRC application experts are financed by SRA funds channeled through SNIC.

Core e-Science communities

Distributed and Parallel Techniques

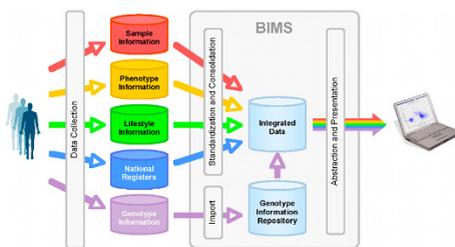


The overall research goals of the DPT community are twofold: first, to ease the development and porting of parallel applications for future large-scale systems and second, to provide easy to use, efficient, and secure frameworks for distributed, data-centric applications.

For parallel applications we focus on the use of task-oriented programming models, high-level equation-based object-oriented textual/graphical programming models and their efficient compilation, the exploitation of multi-level parallelism and associated performance monitoring and analysis.

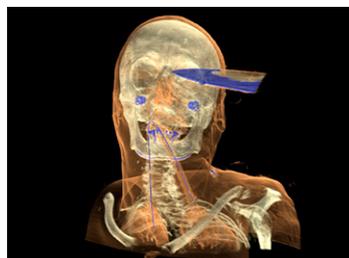
For distributed applications we focus on efficient and secure frameworks for (federated) data centric applications. Specifically, we are developing distributed storage services for cloud computing environments that support strong data consistency, authorization and auditing of data accesses, and efficient replication algorithms.

Data Management



The aim of the research is to develop standards, methods, and tools for data-intensive research. It will deliver unique values for researchers through the integration of server, storage, networking and software for management and sharing of data and knowledge in a secure environment.

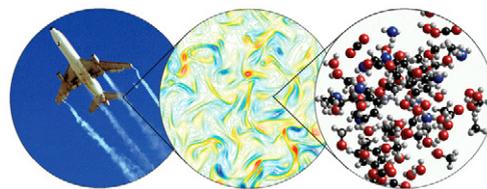
Visualization



The human visual sense is superior to today's computers in terms of perceiving and interpreting content in images. Using vision, high bandwidth can be created between digital data representations and the user. It is this human capability that visualization builds upon by generating images (moving – or not, interactive – or

not) representing the content of large and complex data sets. In image science the goal is to translate complex spatio-temporal patterns into forms that can be understood by humans, through a series of processing steps to extract descriptions of the objects embedded in the data. To achieve this goal there is a need to develop practices for interacting with the data sets, visualized in a large variety of ways, for sense making to the user in her context. This will turn the visualization and image processing system into an increasingly autonomous companion to the e-Scientist.

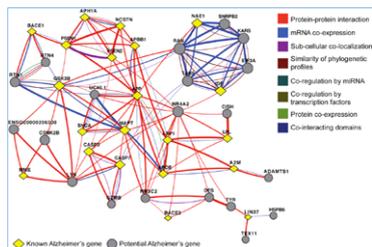
Numerical Analysis



Development of numerical algorithms is critical in successful e-Science based research, often matching or outperforming improvements of hardware in terms of speed gains. The numerical analysis group works together with several other e-Science communities in their projects and also lead its own projects. The work is focusing on algorithm development and analysis for multiscale, multiphysics and stochastic problems related to application fields like complex and multiphase flow, high frequency wave propagation and molecular dynamics.

Application e-science communities

Bioinformatics



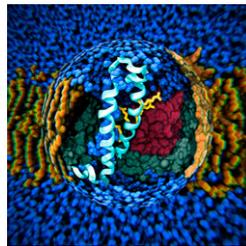
Bioinformatics is a strong strategic field in Sweden, with world-leading research on membrane protein structure prediction, maintenance of high-profile sequence databases (Pfam, InParanoid), and the Stockholm-Linköping groups have been driving the formation of the Swedish BILS node of the ELIXIR infrastructure for bioinformatics in Europe. This community builds on the constellation of Stockholm Bioinformatic Center, SBC, (SU + KTH), Center of Biomembrane Research, CBR, (SU), the new high-throughput genomics core facility in SciLifeLab (KTH + SU + KI), and additional participants from LiU.

Complex Diseases



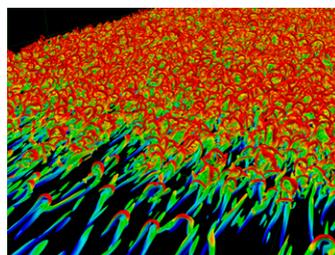
For the first time in medical history, the technical prerequisites exist for merging cascades of molecular data from biological samples, with images of the brain or other body parts, with longitudinal information on lifestyle and health, and with clinical data from hospital databases and from disease registers. High-throughput molecular technology allows mining a large number of genes, transcripts, metabolites or proteins in order to detect association with indicators for health and disease. Adding the interplay with environmental and lifestyle factors allows new insight into the complex multi-factorial mechanisms underlying human health. The Complex Diseases community has recently started a SeRC flagship project: e-Science for Cancer Prevention and Cure – eCPC, in collaboration with the communities Data Management and Distributed and Parallel techniques.

Molecular Simulation



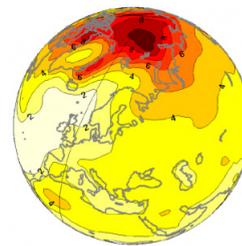
Particle-based modeling such as molecular dynamics or Monte Carlo simulation is a very powerful tool to study time-dependent processes using methods based on statistical mechanics. Some of the currently most important areas include modeling of biological macromolecules, materials, condensed matter and nanophysics. A typical example is the development of new drug molecules to interact with specific proteins such as the ion channel on the cover of this report, or to understand complex dynamics on the molecular level. SeRC is internationally leading both in applications and methodological research, and several widely used tools such as the GROMACS toolkit are developed by SeRC groups.

FLOW



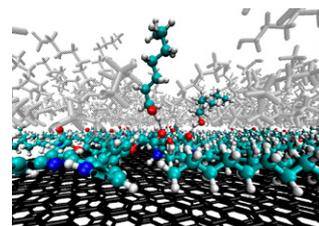
The FLOW community includes application groups from fluid mechanics, geophysical flow, aeronautics, and astrophysics, some of which already collaborate in the Linné FLOW and Bolin centres. There is also a strong connection to the strategic areas of transport (laminar wings) and energy (wind energy). The development of efficient and accurate methods for the simulation of turbulent flow is one of the core activities. Major breakthroughs are expected via numerical simulations of various types of turbulence; such situations are at present often only attainable in experiments. Members of the FLOW community are also active within European supercomputer activities such as PRACE and DEISA/DECI.

Climate Modelling



The climate modelling community engages scientists from the Bert Bolin Centre for Climate Research, SU, SMHI, and FLOW. The research concerns fundamental processes in the climate system and its natural evolution and variability as well as the effect of changes imposed by man's ever-increasing impact on the system through emission of greenhouse gases and aerosols, and changes in land-use, vegetation and hydrology. We have a leading role in a European consortium of institutes from nine countries on the development and application of a new Earth System Model named EC-Earth. e-Science components in the climate modelling community include development of efficient coupling techniques between model components and numerical techniques. Other areas of interests are, scalability, techniques on handling the large amount of data generated by the models as well as e-science tools to facilitate model development and evaluation.

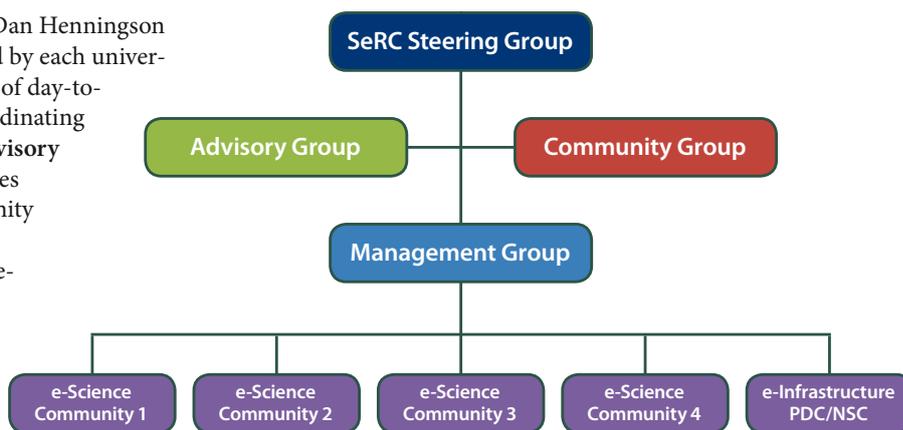
Electronic Structure



This community focuses on quantum-mechanical calculations of materials properties from first principles. The computational methods used are based on density functional theory and Hartree-Fock methods, but also on multiscale simulation techniques. The scientists involved in the community are theoretical physicists and chemists, as well as materials scientists. More powerful computational methods in this research area are expected to enable a realistic description of dynamics, complex transport problems (for example optics and thermoelectrics) and strongly correlated systems.

Organization of SeRC-leadership

SeRC is led by a formal **Steering Group** with Dan Henningson as Director and with representatives appointed by each university. A smaller **Management Group** takes care of day-to-day business and a **Community Group** is coordinating the communities. There is also an external **Advisory Group** with national and foreign representatives from industry as well as academia. A **Community Coordinator** is leading each community and reports to the Community Group. The Management Group consists of people from the Steering and Community Groups where the members may vary depending on assignments. The Steering group and the Community group often have their meetings together.



Steering and Community group



Dan Henningson, KTH,
Director,
Steering group



Bengt Persson, LiU, KI,
Steering group,
NSC Director,
Bioinformatics
Coordinator



Olivia Eriksson, KTH, SU,
Steering group,
SeRC Coordinator



Philipp Schlatter, KTH,
FLOW Coordinator



Anders Ynnerman, LiU,
Co-director,
Steering group,
Visualization Coordinator



Olof Runborg, KTH,
Steering group,
Numerical Analysis
Coordinator



Ola Spjuth, KI,
eCPC Flagship
Coordinator



Juni Palmgren, KI, SU,
Steering group,
Complex Diseases
Coordinator



Anna Delin, KTH,
Steering group,
Electronic Structure
Coordinator



Gunilla Svensson, SU,
Steering group,
Climate Coordinator



Jan-Eric Litton, KI,
Data Management
Coordinator



Berk Hess, KTH,
Molecular Simulation
Coordinator



Erik Lindahl, KTH,
Steering group



Erwin Laure, KTH,
Steering group,
PDC Director,
Distributed and Parallel
Techniques Coordinator

Annual meeting
The third annual meeting of SeRC
will take place the 24th – 25th of April 2012.
Location to be announced.

