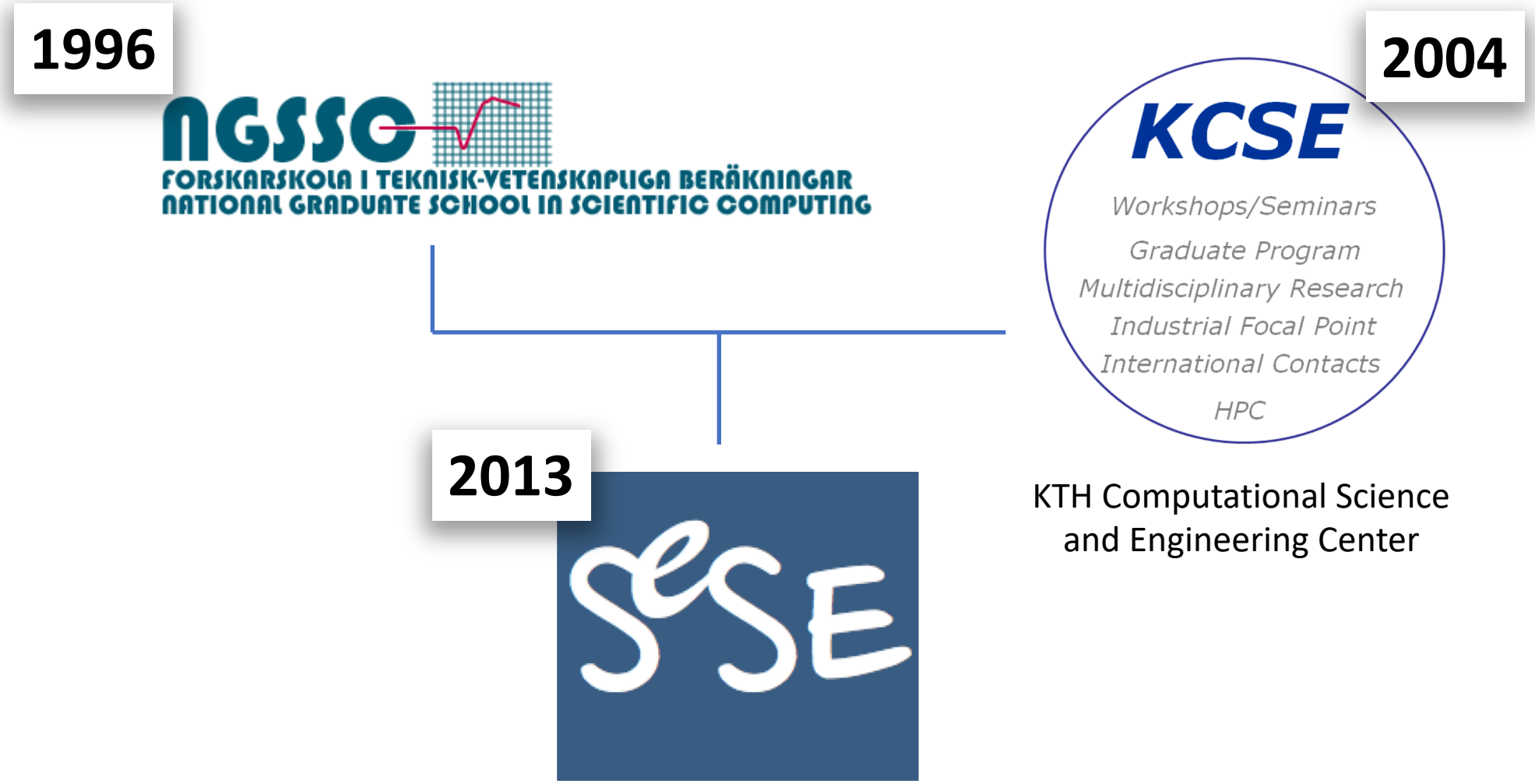
eScience Education Within SeRC

Arvind Kumar
Computational Science and Technology
EECS, KTH Royal Institute of Technology
Stockholm, Sweden
arvkumar@kth.se
https://www.kth.se/profile/arvindku/
http://sese.nu/

SeRC
Swedish e-Science Research Centre

eSCience
THE e-SCIENCE COLLABORATION

History of e-Science education in Sweden



Hast et al. 2015 11th IEEE Conf. on e-Science

Collaboration between two research initiative

Equal Financing from SeRC and eSSENCE

Typical budget: ~150K SEK

Swedish e-Science Education (SeSe) Graduate School

Pavlin Mitev
Uppsala University
Director
Administration

Arvind Kumar
KTH Royal Inst. Of Tech.
Vice Director
Coordination

SeSe Coordination group:
Anna Delin, KTH
Elias Jarlebring, KTH

Gunilla Svensson, Stockholm Uni.

Anders Hast, Uppsala Uni.
Gunilla Kreiss, Uppsala Uni.
Ingela Nyström, Uppsala Uni.
Ola Spjuth, Uppsala Uni.

Jonas Lindemann, Lund Uni.
…

SeSe Mission

- To provide education in fields where the use of **e-Science** is emerging
- Identify areas where courses within **e-Science** are needed
- A meeting place for graduate students using **e-Science** tools and techniques

e-Science Training/Education

Training

Help users to use the tools & resources efficiently = less user support

Field specific & general courses

Education

Theory of HPC & Neighbouring courses

HPC Software & Hardware

SeSE course format:

- 1 week preparation week
- 1 week on campus
- 1 week project
- Equivalent of 3 to 5 credits

Who can attend SeSE courses?

- SeSE courses is open for PhD students enrolled in a graduate programme at a Swedish university (or Nordic university)
- Travel and accommodation funding available

Course Financing

Course of 10 students minimum		
Developing the course 80 000 SEK	Giving the course 80 000 SEK	Travel grants 60,000 SEK Max 10K/student

SeSe courses

CORE topics

- Scientific computing
- Software tools
- Visualization
- Computational data science

- Focus on general skills for eScience
- The objective is to provide a broad coverage
- Offered more regularly

Advanced/specialized topics

- Climate modelling
- Matrix computations for statistics
- Optimization tools for simulations
- Multiscale modelling
- Fluid dynamics

- Focus on specific applications of eScience
- Advanced topics in a specific area
- For experts
- Usually not offered regularly

Core courses

- **Scientific computing** (programing, accelerators, cloud computing)
  Computational Python – Olav Vahtras, Stockholm
  Applied cloud computing - Andreas Hellander, Uppsala
  PDC Summer School – Erwin Laure, Stockholm
  Introduction to Scientific Computing I & II - Jonas Lindeman, Lund
  Introduction to GPU and accelerator programming for scientific computing &
  Performance Optimisation of Numerical Simulation Codes Michael Schliephake, Stockholm

- **Software tools** (performance and code optimization)
  Maximizing performance in practical HPC applications - Carl Nettelblad, Uppsala
  Scientific software development toolbox - Radovan Bast, Stockholm
  Python for MATLAB users - 2 days workshop – Olivier Verdier, Stockholm

- **Scientific Visualization** – Ingrid Hotz, Linköping, Anders Hast, Uppsala

- **Computational data science** (databases, big data analysis, data management)
  Scientific Data Management - Graham Kemp, Chalmers
  Big Data Analytics (with Hadoop and Spark) - Patrick Lambrix, Linköping

Specialized courses

- **Climate modeling -** Rodrigo Caballero, Stockholm
- **Matrix computations for statistics and applications** - Maya Neytcheva, Uppsala
- **Uncertainty quantification for partial differential equations -** Jan Nordström, Linköping
- **Numerical Solution of Initial Boundary Value Problems -** Jan Nordström, Linköping
- **Semantic Web Technologies -** Olaf Hartig, Linköping
- **Foundations of numerical optimization -** Eddie Wadbro, Umeå
- **Winter School in Multiscale Modeling** - Zilvinas Rinkevicius, KTH
- **Topics in CFD** - Philipp Schlatter, Stockholm
- **Numerical Linear Algebra** - Bärbel Janssen, KTH
- **Tools and Techniques for Simulation and Optimisation** - Martin Berggren, Umeå
- **Potential energy surfaces and dynamics**
- SeRC visualization school - Linköping

Course statistics

Total courses	48
Unique courses	26
Core courses	13
Advanced courses	13

Course selection

- Is the topic ‘hot’ (e.g. AI, Climate modelling)?
- Is the course responsible person an influential scientists or teacher?
- Does the course fit in the 3-week format?
- Does the course covers an advanced topic usually not covered in available programs at the universities?
- Will the course attract at least 10 students?

e-Science courses at the universities

SG2212 Computational fluid dynamics
DD2360 Applied GPU Programming
MAI0122 Numerical Solution of Initial Boundary Value Problems
TNM067 Scientific Visualization
DD2420 Parallel Computer Systems
Deep Learning in Data Science
DD2421 Machine Learning
TDDC78 Programming parallel computers: Methods and Tools DH2320
BB2490 Analysis of Data from High-Throughput Molecular
Biology Experiments
KB7004 Bioinforamtics
(TFYA75) Computational Physics, Graduate Project
93FY51 Fysik fördjupning
SF2520 Applied Numerical Methods
SF2522 Computational Methods for Stochastic Differential Eq.
SF2524 Matrix computations for large-scale system
SG2221 Wave motions and stability
DD2370 Computational Methods for Electromagnetics
MAI0129: Stochastic Galerkin Methods for Partial Differential Eq.
TDDD56 Multicore and GPU Programming
DD2257 Visualization
732A54/TDDE31 Big data analytics
SF2521 Numerical Solutions of Differential Eqs
Numerical algorithms for data-intensive science

SG2224 Applied Computational Fluid Dynamics
DD2356 Methods in High Performance Computing
DD1354 Models and Simulation
TDDD43 Advanced Data Models and Databases
Deep Learning in Data Science
SG3114 Computational Fluid Dynamics
DD3360 Applied GPU Programming
Deep Learning in Data Science
DD3354 Probabilistic Graphical Models
Deep Learning, Advanced
DD3431 Machine Learning
Computational Physics, Graduate Project
SF3565 Program Construction in C++ for Scientific Computing
SF3580 Numerical linear algebra
SG3112 Turbulence
DD3258 Introduction to High Performance Computing
DF21500 Multicore Computing
Advanced Data Models and Databases
2992 Biostatistics III: Survival analysis for epidemiologists
SF3561 The Finite Element Method
SF3584 Preconditioning for Linear Systems
Programming Frameworks for Deep Learning

- At least 30 second cycle and 12 third cycle ‘e-science’ courses exist at KTH and LU
- There is an overlap between available II/III cycle courses and SESE courses
- In general there is a lack of course on computational biology, both at SESE and in KTH/LU

Data Collected by Elias Jarlebring

SeRC
e-Science Education Coordination group

Arvind Kumar
EESC
Computational
Brain Science

Elias Jarelbring
Applied Maths
Scientific
Computing

Stefano Markidis
EECS
High Performance
Computing

Lucie Delemote
Biophysics
Computational
Biophysics

Towards a strategy for future SeRC courses

1. The SeSE advanced courses should reflect the current research focus in SeRC.
   We propose that each MCP should offer at least one advanced course every year

2. Introduce new courses on Computational Neuroscience (BrainIT), Medical Image Analysis (e-Science for Cancer Prevention and Control) & Machine Learning

3. SeRC board should persuade people in different MCPs to give courses

4. The self funded courses can continue as it is for now and only when we need to pay for these course we can decide whether to continue or not

5. SeSe should collaborate with other large-scale research collaboration networks and link with their course e.g. WASP courses, computational fluid dynamics courses

6. SeSE provide guidelines to improve pedagogical aspects of the PhD level courses
   This is needed as SeSe courses require a lot of online interaction.

Challenges of eScience education within SeSe

- Format of the courses (1+1+1)
  - Many students do not finish the course assignments in time
  - Requires extra effort to modify a regular course into a 1 week course module
  - Students come from diverse background
  - Usually there is no (*objective*) criterion for student selection
- Identification of advanced courses that are relevant to more students
- Top-down decisions on offered courses – students usually don’t get to suggest courses
- Relatively limited set of courses

Challenges of eScience education

- ‘Context’ in which e–Science can be introduced in more courses (even at undergrad level)
  - to reach students at very early stages to foster the eScience culture
- Balance between ‘analytical approach’ and ‘numerical simulation’ based understanding
- ‘Interactive’ course material
  - lecture slides/videos are usually not very effective
  - Interactive environments such as JUPYTER NOTEBOOKs (Neuroscience: AK), WIKI (Numerical Analysis: EJ) provide more effective learning experience
- Balance between learning ‘scientific contents’ and ‘software tools’
- Student evaluation
- Maintenance of course contents up to date - software tools/platforms keep evolving
- Availability of cyberinfrastructure
- eScience ≠ e-Learning

SeSe Courses in 2019

Spring Semester

- Maximizing performance in practical HPC applications
- Introduction to climate modelling
- Advanced Molecular Dynamics

To sign up for the courses go to http://ww.sese.nu

Fall Semester

- Computational python
- Semantic web technologies
- Introduction to high performance computing
- Numerical solutions of initial boundary value problems
- Matrix computations in statistics with applications
- SeRC visualization school