Course Description of DAAD Summer School

Aspects of large scale high speed computing – Building blocks of a cloud

Universities of Freiburg, Göttingen, Cologne
Freiburg
September 2010

This handbook describes the Summer School "Aspects of large scale high speed computing – Building blocks of a cloud" jointly organized by a consortium of German universities and the state university of Paraná (Universidade Federal do Paraná – UFPR).

1 Introduction

The Summer School "Aspects of large scale high speed computing – Building blocks of a cloud" presents a hot topic in both computer science and IT and manufacturing industry. It is organized in cooperation by three German universities with the state university of Paraná (Universidade Federal do Paraná – UFPR) bringing together experienced professors and researcher from various fields in computer science. The DAAD (Deutscher Akademischer Austauschdienst – German Academic Exchange Service) financially supports this Summer School.

2 Lecturers

This Summer School is organized by a group of different lecturers from the German Universities of Freiburg, Göttingen and Cologne and a group of professors of the computer science department of the UFPR.

Dr. Luis C. E. Bona obtained his Ph.D. from Federal University of Technology of Paraná (UTFPR). He is professor at the Informatics Department of the Federal University of Paraná. Bona has been working with research in Computer Science in the area of Distributed Systems since 2000. His research interests include Grid and Cloud Computing, Peer-to-Peer Systems, Digital Preservation and Open Source Software.

Dr. Carlos Carvalho obtained his Ph.D. in Solid State Physics at Paris 6 University in 1990. He is professor at the Physics Department and his current activities are related to management of large storage systems and mirrors, and scientific computing installations.

Dr. Matthew Lewis academic path started at Queen's University in Canada. There, he received his B.Sc. and M.Sc. while studying, and then implementing multiple ciphers in hardware for use in low power mobile devices. He then started his Ph.D. at the University of Freiburg in Germany, where he has focused on accelerating Formal Verification and Test Methods using parallel systems. In this regard, he has developed multiple parallel SAT and QBF solvers, while also playing a large role in the integration and optimization of them for BMC, MC, and ATPG tools. In 2010 he finished his Ph.D., and has since been continuing his work through a postdoctoral position in Freiburg.

Dr. Gerd Rapin Since 2008 Dr. Gerd Rapin works at Volkswagen AG in the field of High Performance Computing with special focus on Computational Fluid Dynamics (CFD). Main tasks are the setup, improvement and evaluation of process chains and cooperations with universities and other research institutions within the framework of pilot projects. Gerd Rapin studied mathematics, computer science and public economy at the Georg-August University Göttingen. He received his Ph.D. in applied mathematics at the Georg-August University Göttingen in 2003. From 1999 until 2008 he was employed as a researcher at the Georg-August University Göttingen, where he gave numerous self independent courses in applied mathematics and mathematical oriented computer science. The time in Göttingen was interrupted by several long term research stays in Turin and Paris. In 2004 Dr. Rapin already contributed to a sommer school about "Unstable Problems" at the Instituto Nacional de Mathematica pura e aplicada (IMPA) in Rio de Janeiro, which was supported by the DAAD. His research deals with mathematical questions in the area of Computational Fluid

Dynamics and High Performance Computing. He has published numerous articles about domain decomposition methods, stabilization techniques or numerical linear Algebra.

Prof. Ch. Schindelhauer Prof. Dr. Christian Schindelhauer graduated 1991 in Computer Science at the University of Darmstadt in Theoretical Computer Science, 1996 Ph.D. in Theoretical Computer Science, Ph.D. thesis "Median und Average Komplexitätsklassen" (Summa cum laude), received 1997 Prof.-Otto-Roth-Preis der Universität Lübeck for an outstanding Ph.D. thesis, 2002 habilitation at University of Paderborn, habilitation thesis "Communication Network Problems", 2002-2006 assistant professor (Hochschuldozent) at the University of Paderborn, received 2004 the research award of the University of Paderborn for "Design of a wireless energy efficient sensor network for super markets". Since 2006 he is a full professor for Computer Networks and Telematics at the Unversity of Freiburg, Germany, where he received 2009 the faculty award for excellent teaching. His research areas are network algorithms, sensor networking, ad hoc networks, parallel and distributed computing, and storage networks. He authored over 50 peer-reviewed conference and journal publications.

Prof. Gerhard Schneider joined the University of Freiburg in 2002. He is the director of the University computing centre and holds a chair for communication systems. From 2003 to 2008 he served as vice-rector of the University and since 2008 he is the University's CIO. Prior to his appointment in Freiburg he was deputy director of the computing centre of the University of Karlsruhe (1992-1997) and of GWDG, the joint computing centre of the University of Göttingen and the Max-Planck-Society (1997-2001). His main research interest is in managing large systems, including the development of new user services, and in long term digital archiving. Latest activities of his research group include security issues of mobile phone systems (GSM). He was/is a partner in various DFG and EU projects.

Dr. Tobias Schubert Since 2000, Dr. Tobias Schubert is a research assistent at the Chair for Computer Architecture, Albert-Ludwigs-University of Freiburg. In 2008 he received his Ph.D. from the same university. From 2006 to the end of 2007, Tobias Schubert was coordinating the Ph.D. Program "Embedded Microsystems" at the Faculty of Engineering. Since 2008, he is heading the "Technology Platform", a subproject within the DFG (German Research Foundation) Transregional Collaborative Research Center "AVACS – Automatic Analysis and Verification of Complex Systems". Besides his research, mainly in the area of developing efficient parallel SAT, QBF, and SMT solvers, he was a teaching assistent in various courses within the last decade. Lastly, Tobias Schubert has presented the results of his research at multiple peer-reviewed conferences and workshops.

Dr. Fabiano Silva obtained his Ph.D. degree in Industrial Informatics from the Federal University of Technology of Paraná. He is a professor at the Informatics Department of the Federal University of Paraná. His current research interests include parallel programming, artificial intelligence and scientific computing.

Dr. Dirk von Suchodoletz is currently holding a position as a lecturer and principal researcher at the chair of "Communication Systems" at the Institute for Computer Science at Freiburg University. Beside the research in several fields he teaches seminars, does the organizing and planning of different courses, prepares lectures and looks after the students scientific work and theses. The courses of the professorship focus on several topics mainly around communication networks and computer systems, starting from the introduction to

computer networking, routing, mobile and telecommunication, location based services and awareness up to covering practical issues like programming special client server applications in Open Source environments. The professorship offers regular lectures for the bachelor and master study programmes of the faculty of engineering on introduction to computer networks and advanced topics since 2003. Practical issues such as programming special client server applications in Open Source environments are part of the supervised student work. Dirk studied mathematics, economics and politics at the Georg August University of Goettingen. His master thesis was on the "Efficient management of large scale computer installations". Dirk received his Ph.D. in computer science at Albert Ludwigs University of Freiburg in 2008.

Dr. Eduardo Todt got his Ph.D. in advanced Automation and Robotics (Universidad Politécnica de Catalunya). He is professor at the Informatics Department of the Federal University of Paraná. Besides academic experience, also worked with industrial automation systems design in industries in Brazil and Germany. Current research interests are embedded systems, robotics, image processing and transportation systems.

Dr. Chantal Weber Before being a research assistant in the Japanese Studies department at the University of Cologne since 2008, Dr. Chantal Weber (responsible for the social programme of the summer school) was regional representative for Asia and conceptual coordinator for international students, scientists and Ph.D. candidates in the International Office of the University of Freiburg (2006–2008). In this function she was contact person for the DAAD and the Alexander von Humboldt Foundation at the university. She has been one of the coordinators of the DAAD projects Profis I and II as well as one of the organisators of the DAAD conference "Ausländerstudium" held in 2008 at the University of Freiburg. Beside being the director of the yearly summer programme for Japanese students with around 120 students participating, Dr. Weber represented the University of Freiburg at several academic fairs and visits to Asian partners as well as coordinated many events for students, scientists and administratives introducing the university. Since 2009 she teaches international students in acadmic writing in Freiburg. Due to her academic carrier in the field of Japanese Studies she has great experience in intercultural communication, theory and practice.

Dr. Daniel Weingaertner He obtained his Ph.D. in Medical Informatics from the Federal University of Paraná in 2007. He is professor at the Informatics Department of the Federal University of Paraná and Head of the Bachelor in Biomedical Informatics course. His research interests include Parallel Image Processing, Hospital Information Systems and Open Source Software.

3 Main Programme

The building blocks of this summer school consist of different sub topics to present a holistic picture of the aspects of large scale high speed computing. The programme follows a bottom-up approach beginning with the generic theoretical brackground of system design, covering system deployment, connectivity, the organizional building blocks and the real life applications using compute clusters e.g. in car manufacturing, particle physics, genomics or simulation.

- Formal verification of combinational and sequential circuits
- VHDL – An introducton to analyzing, modeling, and simulating digital circuits

- Storage Networks – An overview to a major facet of cloud computing
- Managing Clouds – Fast and flexible large scale machine deployment and management, virtualization for better hardware utilization, greener operation
- Cloud Organization – Institutional components, backgrounds, organisational aspects and juridical implications of clouds, security aspects
- Cloud Applications – Modelling and implementation of algorithms in applied mathematics

Each topic is taught in several units which may span 90 minutes for the lecture parts or fill a whole morning or afternoon for more practical oriented sessions like hands-on exercises.

3.1 Formal Verification

The course unit on formal verification teached by Dr. M. Lewis from the Institute of Computer Science in Freiburg focuses on the question of how to prove correctness – in terms of a given property – of a given design. In this course, several methods for hardware verification will be discussed, such as Decision Diagrams, SAT solvers, symbolic methods for checking equivalence of sequential designs, and finally model checking for proving properties. Furthermore, ideas on how to tackle hybrid systems, interacting with a "real" environment (typically having a continuous domain) will be presented:

1. Introduction to formal verification
2. Decision Diagrams
3. Satisfiability Solver
4. Equivalence checking
5. Property checking / Bounded model checking
6. Verification of hybrid systems

The course is supported by Prof. Eduardo Todt from the UFPR.

3.2 VHDL

This course held by Dr. T. Schubert will give an introducton into analyzing, modeling, and simulating digital circuits. The purpose is to provide an introduction to the main aspects of digital circuit design with VHDL. Besides describing the facilities of VHDL, we will make use of a special hardware environment to illustrate them in practice. The course consists of the following units:

1. Short history of VHDL
2. Ingredients of VHDL like entity, architecture and configuration, ports, signals, variables, constants; if- and case-statements; loops.
3. Packages: standard, std_logic.
4. A more complex example, demonstrating the entire work flow when developing hardware with VHDL.

5. Simulation of VHDL designs.

This course will be prepared by Prof. Eduardo Todt from the UFPR. He will introduce rapid system prototyping of embedded systems with VHDL and FPGA-based kits. Additionally the modeling and implementation of combinatorial circuits and sequential circuits are introduced. A part of this lecture will deal with embedded processors.

3.3 Node Interconnection

This lecture is given by Prof. Dr. Christian Schindelhauer, chair of the Computer Networks and Telematics group, University of Freiburg. Storage Networks are a major facet of cloud computing. Nowadays, even private users use storage networks. This development was enabled by the massive price fall of mass storage and the transition of home video from tape recorders to home computers. Another trend is to outsource storage into the Internet where companies like Google, Amazon, and Apple are selling online storage at cheap prices.

1. Introduction to storage systems and technologies.
2. Virtualization of Storage: RAID, SAN, and Internet Storage
3. Networking for large storage systems
4. Data placement in storage systems
5. Data safety and data encodings
6. Peer-to-Peer-Storage Systems

The course will cover the necessary theoretical background as well as the challenges arising during operation.

3.4 Managing Clouds

The topic of ”Managing Clouds” gives an overview on fast and flexible large scale machine deployment and higher level resource management. It is taught by Dr. D. von Suchodoletz of University of Freiburg and he will talk of different strategies like stateless machine operation, diskless booting, optimal operating system roll-out and virtualization. Additionally aspects of better hardware utilization and efficient resource management are introduced to contribute to a ”greener IT” operation. The course is split into the following units:

1. Higher level overview on management of large scale computer setup and deployment, especially on scalability issues, LAN and WAN scenarios.
2. Flexible boot infrastructures with dynamic operating system selection for testing and optimal compute job routing.
3. Distributed filesystems and block devices
4. Virtualization options

This unit is supported by Prof. Carlos Carvalho. He gives an introduction to configuration and update issues for large software mirrors. This will include an overview on different kernel configurations, memory and bandwidth issues and optimizing updates and synchronization

3.5 Cloud Organization

The lecture on Cloud Organization is given by Prof. G. Schneider, director of the computer center of Freiburg University. Organisatorial issues which are typical for large institutions like universities are in the focus of the first lecture. It is important to distinguish between theoretical approaches like ITIL and roll based models and their practical realisation. The various processes of how decisions are reached and made will also be explained, since they are tightly linked with the ability to launch new ideas. Then the core of modern distributed computing resources will be explained using existing high speed networks (like GEANT, DFN, Belwue) as examples. Service virtualisation allows for distribution of resources over fast networks and leads to the fundamental concepts of "cloud computing". The advantages and disadvantages of such concepts will be discussed. From this it is possible to analyze the technical consequences of various requests from law enforcement agencies and legislation bodies caused be the easy movability of applications in the cloud. These lead to new security models for cloud services which cannot be enforced by national law. This course is structured in the following way:

1. Organisational issues of large organisations, identity management
2. Frameworks (ITIL) and authentication models (Shibboleth)
3. High speed networks: examples, requirements
4. Virtualisation of services: cloud vs. GRID
5. Cloud services: theory and examples
6. Localization of data, legal consequences

This course also tries to introduce the audience into the requirements of day-to-day operation which are usually not offered for a standard computer science degree.

3.6 Cloud Applications

This course is teached by Dr. G. Rapin from Göttingen University and Volkswagen Corporation (research department). In this lecture an introduction to parallel algorithms in the field of scientific computations is given. After introducing basics like Armdahl's law it will be shown how well-known algorithms for the solution of linear systems like the Jacobi or Gauss-Seidel can be formulated in parallel. To keep the presentation simple we will focus on the efficient parallel solution of linear systems. Modern iterative methods like Krylov methods and their preconditioning are described in detail. Moreover, hands-on lessons will be given, where these algorithms are implemented using C and MPI. Therefore, also a small introduction in MPI is given.

1. Basics of parallel computing
2. Fan-in method, Poisson Problem and Finite Differences, parallel treatment of vectors and matrices
3. Simple Iterative Methods (Jacobi- and Gauss-Seidel method) and their parallelisation
4. Conjugate Gradient (CG) method and preconditioning
5. Introduction to MPI I (Basics, Hands-On exercises)

6. Introduction to MPI II (Point-to-Point communication and collective communication, hands-on exercises)

This lecture will be prepared by a programming course held by the Brazilian partners. The will give an overview on development basics of (taylored) cluster and cloud applications The lecture is supported by the Professors Daniel Weingärtner, Fabiano Silva and Luis C. E de Bona from the UFPR. They will give a preparatory introduction to parallel programming on shared memory systems and will introduce OpenMP, CUDA and the PRAM parallel programming model. Additionally they will talk of the basic building blocks of cloud computing. The introduction to Cloud Computing presents the main concepts and definitions used in the Cloud Computing paradigm. This new paradigm can be defined as a large-scale distributed computing paradigm in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand over the Internet.

3.7 Questions and Answers

Twice a week a special question and answer session will be held by the group of lecturers teaching course units this week. This should give the participants the opportunity to ask questions, catch up with open issues or discuss specifics of the topic which could not be covered in the main lecture units.

4 Social and Intercultural Programm

The DAAD summer school is not just introducing various topics of computer science but intends to give a broad overview on the German university system, academic careers and cultural background. The University of Freiburg itself is an international university which has significant experience in educating and supporting international students. There are many programmes and information centers especially designed for international students and scientists supporting them in their studies but also in their every day life in Freiburg. Based on this experience the social programme during the summer school aims to introduce Germany in general and Freiburg in particular to the students and to prepare them for a study programme in Freiburg.

1. German history and culture: This lesson gives an overview about the recent German history and culture.

2. The University and the City of Freiburg: The University of Freiburg has a history of more than 550 years. Is has always been integrated part of the city and therefore had a big influence. The lessons will give an introduction to the university but also to the city and the region close to France and Switzerland.

3. Intercultural communication and awareness raising: This practical and interactive orientated lessons the cultural awareness will be raised through e.g. role-playing. The students will be given pracitical advises to e.g. avoid cultural shocks.

4. Academic writing and research standards in Germany: International students face other difficulties in academic writing than German mother-tongue students. This won't be a German language course, but give an overview on text types used in the academic context. In these lessons the students will get to know the research standards in Germany as well as some aspects of academic writing at a German university.

The social and intercultural programme is split into different parts: lessons on history and culture of Germany in general and Freiburg in particular; lessons on intercultural communication; lessons on the German university system; separat presentation for Master and Ph.D. candidates on studying in Germany; panel dicussions with German and Brasilian scientists. These lessons aim to give Master students and Ph.D. candidates of the Informatics Department at the UFPR specific information about corresponding programmes in Germany and answers to individual questions:

1. Study programme for Master in Germany and Freiburg: requirements, organisation, costs, scholarship possibilities, programmes and graduation, support for international students at the university.

2. Ph.D. in Germany and Freiburg: academic system, requirements, organisation, costs, scholarship possibilities, programmes and graduation, support for international students at the university, International Graduation Academy.

3. German university system: Due to the so called Bologna process and the establishment of the international standard programmes of Bachelor and Master the German university system is undergoing a great change. In this lesson an overview about these systems will be given.

Besides these academic lessons the personal contact in easy-going atmosphere with the German lecturers will bring a peer effect.

5 Background Information on Involved Universities

The University of Freiburg is among the top nine German universities selected in the Initiative for Excellence of the German Federal and State Governments. The technical faculty composed of the Department of Informatics and the Department of Microsystems Engineering.

The Communication Systems Group presented by G. Schneider and D. v. Suchodoletz belongs to the Faculty of Engineering and is attached to the Computing Center of the University of Freiburg. Therefore, its research aspects are influenced by both feasibility and rollout/scalability, and it has access to feedback from a large number of users. The main research fields include large scale cluster and grid computing, rollout of large computer pools, identity management, location-based services, personalization, context-awareness, long term archiving, and impact of IT on law. The computer center runs the Black Forest Grid with over 2500 nodes and operates a part of the Baden-Württemberg state compute cluster. It is deeply involved in the Identity Management of the whole university campus. The chair cooperates with the Bernstein Center for Computational Neuroscience Freiburg and the Radiology Department of the University Hospital, and has partners in the EU, Brazil and the US. Currently they were involved in the EU-funded project PLANETS, working in the fields of e-science, and projects on location-based services and mobile computing.

The Chair of Computer Architecture headed by B. Becker, T. Schubert and M. Lewis belongs to the Faculty of Engineering as well. The research mainly focuses on methods and procedures to ensure the correctness, reliability and robustness of circuits and systems. Our work encompasses initial design stages through manufacturing, and even use in the field. They work closely with many international industrial and university partners – such

as Mentor Graphics, Infineon, LIRMM Montpellier, the University of Iowa and the University of Tokyo - and they play a fundamental role in the continuous technological progress made today on ever increasingly complex systems. This growth in complexity is being fueled by the reducing in costs of producing chips with millions, or even billions of transistors. Additionally, and in contrast to traditional computers and PCs, they are also seeing the rise in embedded systems that consist of the integration of processors, specialized hardware, and software, into onecompact package. Embedded systems are now an indispensable part of everyday life as they appear in cars, trains, aircrafts, medical equipment, industrial automation, and in all the wonderful mobile communication and electronic devices in daily use.

Institute for Numerical and Applied Mathematics of the university of Göttingen is presented by Dr. Gerd Rapin. The institute is involved into a number of research and international cooperation projects like the Research Training School on "Identification in Mathematical Models", large scale research projects "Nanoscale Photonic Imaging", "ARRIVAL", Joint Research Project REIT and INVERS "Deconvolution with sparsity constraints" and a research cluster on "Statistical Regularisation and Qualitative Constraints". The institute organized a summer school on mathematics in Brazil, Rio de Janeiro 2006.

6 Non-Curicular Activities

Beside the curricular activities of the summer school weekly events open to university public will be offered (announced university-wide):

- Lecture on the recent developments of the German university system (Bologna process) and the initiative of excellence.
- Podium discussion with the former vice-rector of the University of Freiburg, Prof. Dr. G. Schneider: The University of Freiburg – recent developments in science and internationalisation
- Panel discussion with Brasilian and German scientists: Internationalisation – Chances and Problems. What can we learn from each other? Where can we support each other?