Cloud-Based Computing Clusters for Data Analytics

The cloud is well known for its elasticity by leveraging abundant resources. Cloud data centres easily host thousands or even millions of multicore servers. Further, these servers are increasingly virtualized for the sake of data centre efficiency. However, the reality is that these resources are often relentlessly exploited particularly to improve applications performance. Although the elasticity facilitates achieving cost efficiency (or the performance to cost ratio), the ultimate efficiency in resource usage (or more broadly data centres) lies in scheduling and resource allocation strategies that explicitly take into account actual resource consumption. The optimization of resource efficiency in clouds is of great practical importance considering its numerous benefits in the economic and environmental sustainability. In this talk, we will discuss resource efficiency in cloud data centres with examples providing alternative approaches to constructing “compute clusters” in clouds to support large-scale distributed processing applications for data analytics applications.

Albert Y. Zomaya is the Chair Professor of High Performance Computing & Networking and Australian Research Council Professorial Fellow in the School of Information Technologies, Sydney University. He is also the Director of the Centre for Distributed and High Performance Computing which was established in late 2009.

Dr. Zomaya published more than 500 scientific papers and articles and is author, co-author or editor of more than 20 books. He served as the Editor in Chief of the IEEE Transactions on Computers (2011-2014) and was elected recently as a Founding Editor in Chief for the newly established IEEE Transactions on Sustainable Computing. Dr. Zomaya also serves as an associate editor for 22 leading journals, such as, the ACM Computing Surveys, ACM Transactions on Internet Technology, IEEE Transactions on Cloud Computing, and IEEE Transactions on Computational Social Systems. Dr. Zomaya is the Founding Editor of several book series, such as, the Wiley Book Series on Parallel and Distributed Computing and Springer Scalable Computing and Communications.

Dr. Zomaya was the Chair the IEEE Technical Committee on Parallel Processing (1999–2003) and currently serves on its executive committee. He is the Vice–Chair, IEEE Task Force on Computational Intelligence for Cloud Computing and serves on the advisory board of the IEEE Technical Committee on Scalable Computing and the steering committee of the IEEE Technical Area in Green Computing. Dr. Zomaya has delivered more than 160 keynote addresses, invited seminars, and media briefings and has been actively involved, in a variety of capacities, in the organization of more than 600 conferences.

Dr. Zomaya is the recipient of the IEEE Technical Committee on Parallel Processing Outstanding Service Award (2011), the IEEE Technical Committee on Scalable Computing Medal for Excellence in Scalable Computing (2011), and the IEEE Computer Society Technical Achievement Award (2014). He is a Chartered Engineer, a Fellow of AAAS, IEEE, and IET. Dr. Zomaya’s research interests are in the areas of parallel and distributed computing and complex systems.
How to Build Reliable Software Systems for Cyber-Physical Applications

Software has seamlessly become a key component of our lives. It is found in almost everything we use or are surrounded with, from simple gadgets and appliances at home, via smartphones, vehicles we drive, to work related computers and systems that automate our industry and economy. Software is increasingly used in systems that interact with the physical world; we call such systems embedded or cyber-physical systems (CPS) and they already make more than 90% of all computer applications. As a man-made and designed engineering product, software is highly prone to bugs and malfunctions. For example, software bugs in vehicles have led to recalls of millions of vehicles by leading manufacturers. The importance of accurate and timely collection of data from multiple sensors, producing control outputs and reacting on any abnormality, is key to relying on software in CPS applications. This talk addresses software dependability and its design by considering software from a new, system’s perspective, where it operates in real-time and interacts with its immediate environment, creating systems with feedback. Software as a System (SaSy) is a new paradigm we propose that adds dimensions of composability and time to what is our traditional perception of software and programs. The answers we seek are related to the fundamental questions of how we compose traditional programs/software into reliable software systems and how we provide software designers with methods and tools that automate this process in a trustworthy way. We present some recent results of research at the University of Auckland that deal with the automated entire design flow of software systems, from specification in system-level design language to implementation on standard or custom-made system-on-chip platforms. As the result, we expect the new software design paradigm to dramatically increase productivity and reliability of software systems for CPS.

Zoran Salcic is a professor of computer systems engineering at the University of Auckland, New Zealand. He leads research in the domains of embedded systems and industrial informatics at the Department of Electrical and Computer Engineering. He has the BE (1972), ME (1974) and PhD (1976) degrees in electrical engineering from the University of Sarajevo. He did most of his PhD research at the City College New York (CCNY). His main research interests include embedded and real-time systems, cyber-physical systems, custom-computing machines, multi-core systems on chip, design automation tools for embedded systems, hardware-software co-design, models of computation and languages for cyber-physical systems and wireless sensor and actuator networks. He has published more than 300 peer-reviewed journal and conference articles, books and book chapters and numerous technical reports. He is a Fellow of the Royal Society (Academy of Science) New Zealand and recipient of Alexander von Humboldt Research Award in 2010.
Keynote Speaker III (Day 2)

Mazin Yousif
Global Chief Technology Officer
Vice President of Architecture for the
Royal Dutch Shell Global Account
T-Systems, International

Cyber-Physical Systems – Evolution and Challenges

Starting from the 1950’s, science fiction novels and movies have depicted interactions and collaborations between machines and humans beyond the typical Human-Computer Interaction (HCI) as we know it. The movie “Colossus: the Forbin Project” from 1970 is a great example. It may take us, really, very long time to build a cyber-physical system similar to Colossus, but we can show we are on some trajectory to a goal where cyber-physical systems will constitute an integral part of our social fabric. When will that happen? No one can predict. However, we are seeing them in islands of industry verticals such as manufacturing and healthcare and they will likely evolve to be deployed in more industry verticals as the various facets of technology mature. The reason cyber-physical is becoming to some extent a reality is we have been able to digitize almost everything and whatever we digitize, we connect, analyze and make it part of our ecosystem. At the same time, computing paradigms have evolved considerably enabling all types of social innovations under trends such as smart mobility, cloud computing, big data and the Internet of Things.

This talk will look at CPS Systems from various points of view including evolution, architecture, challenges and whether they will be able to advance them to their full potential or only to some limited potential.

Mazin Yousif is the Global Chief Technology Officer and Vice President of Architecture for the Royal Dutch Shell Global account at T-Systems, International. Before joining T-Systems, he spent some time with IBM Canada GTS, Avirtec and Intel. He was also a Professor at Louisiana Tech University.

Dr. Yousif serves as a Technology Advisor, Non-Executive Director and Board Member for two start-up companies: Simudyne and Stafford & Associates. He chairs the Advisory Board of the European Research Consortium for Informatics and Mathematics (ERCIM). He is also the Editor-In-Chief of the IEEE Cloud Computing Magazine. He founded the NSF Industry/University Cooperative research Center for Autonomic Computing and delegated to three Professors to complete the needed paperwork. Dr. Yousif was an adjunct professor in several universities including Duke, NCSU, OGI and Arizona. He has served as the General Chair or Program Chair for many conferences and serves in the editorial board of many journals. He is a frequent speaker in academic and industry conferences on topics related to cloud computing, big data and autonomic computing. He has also published extensively and was an IEEE Distinguished Visitors Program speaker from 2008 – 2013.
**Keynote Speaker IV (Day 2)**

Hans W. Guesgen  
School of Engineering and Advanced Technology  
Massey University, New Zealand

---

**Activity Recognition in Smart Environments: When and Where makes a Difference**

In recent years, researchers have developed various approaches to recognising human activities based on sensor data, ranging from logic-based approaches to probabilistic machine learning approaches. Although the reported successes are promising, determining the correct activity from sensor data alone is often impossible, since sensors can only provide very limited information and human activities are inherently complex. This talk explores how spatial and temporal information can be used to improve the process of recognising human activities in smart environments such as smart homes.

Hans Guesgen is the Chair of Computer Science in the School of Engineering and Advanced Technology at Massey University in Palmerston North, New Zealand. He holds a diploma in Computer Science and Mathematics of the University of Bonn, a doctorate in Computer Science of the University of Kaiserslautern, and a higher doctorate (Habilitation) in Computer Science of the University of Hamburg, Germany. Before joining Massey University, he worked as a research scientist at the German National Research Center of Computer Science (GMD) at Sankt Augustin, as a post-doctoral fellow at the International Computer Science Institute in Berkeley, California, and as a lecturer, senior lecturer and associate professor at the Computer Science Department of the University of Auckland, New Zealand. Hans has published more than 100 refereed papers in various areas, including ambient intelligence, smart environment, and spatio-temporal reasoning. He is one of the nine inaugural senior members of the Association for the Advancement of Artificial Intelligence (AAAI) and has served on the editorial board of several international journals, including Applied Intelligence, Universal Computer Science, and the Journal of Ambient Intelligence and Smart Environments.
Securing Critical Cyber Infrastructures with Predictive Capability

Cyber infrastructures for critical applications, such as in the areas of homeland security, transportation, finance, telecommunications, energy, and health care, attract frequent attacks from sophisticated and well supported attackers with abundant human and computer resources. Hence, it is important to develop cost-effective approaches to securing critical cyber infrastructures. In order to achieve such a goal, security approaches with predictive capability would be needed. In this address, the challenges, current state of art, and future research directions for developing cost-effective approaches with predictive capability for securing critical cyber infrastructures will be discussed. These approaches also need to incorporate human behaviors and be capable of operating in partially observable environments.

Stephen S. Yau is Professor of Computer Science and Engineering and the founding director of Information Assurance Center at Arizona State University (ASU), Tempe, Arizona, USA. He served as the chair of the Department of Computer Science and Engineering at ASU in 1994-2001. Previously, he was on the faculties of Northwestern University, Evanston, Illinois, and University of Florida, Gainesville.

He served as the president of the Computer Society of the Institute of Electrical and Electronics Engineers (IEEE) and was on the IEEE Board of Directors, and the Board of Directors of Computing Research Association. He served as the editor-in-chief of IEEE COMPUTER magazine. He organized many major conferences, including the 1989 World Computer Congress sponsored by the International Federation for Information Processing (IFIP), and the IEEE Annual International Computer Software and Applications Conference (COMPSAC) sponsored by IEEE Computer Society. He has served as a honorary co-chair and general co-chair of the IEEE World Congress on Services and co-located conferences on Service Computing, Cloud Computing, Web Services and Mobile Services.

His current research includes services and cloud computing systems, cyber security, trustworthy computing, software engineering, internet of things, and ubiquitous computing. He has received many awards and recognitions, including the Tsutomu Kanai Award and Richard E. Merwin Award of the IEEE Computer Society, the IEEE Centennial and Third Millennium Medals, and the Outstanding Contributions Award of the Chinese Computer Federation. He is a Life Fellow of the IEEE and a Fellow of the American Association for the Advancement of Science. He received the B.S. degree from National Taiwan University, and the M.S. and Ph.D. degrees from the University of Illinois, Urbana, all in electrical engineering.
Keynote Speaker VI (Day 3)

Jinjun Chen
Director, Lab for Data Systems and Visual Analytics
Global Big Data Technologies Centre
Faculty of Engineering and IT
University of Technology Sydney, Australia

Searchable Symmetric Encryption: Potential Attacks, Practical Constructions and Extensions

Data outsourcing has become one of the most successful applications of cloud computing, as it significantly reduces data owners' costs on data storage and management. To prevent privacy disclosure, sensitive data has to be encrypted before outsourcing. Traditional encryption tools such as AES, however, destroy the data searchability because keyword searches cannot be performed over encrypted data. Though the above issue has been addressed by an advanced cryptographic primitive, called searchable symmetric encryption (SSE), we observe that existing SSE schemes still suffer security, efficiency or functionality flaws. In this research, we further study SSE on three aspects. Firstly, we address the search pattern leakage issue. We demonstrate that potential attacks exist as long as an adversary with some background knowledge learns users' search pattern. We then develop a general countermeasure to transform an existing SSE scheme to a new scheme where the search pattern is hidden. Secondly, motivated by the practical phenomenon in data outsourcing scenarios that user data is distributed over multiple data sources, we propose efficient SSE constructions which allow each data source to build a local index individually and enable the storage provider to merge all local indexes into a global one. Thirdly, we extend SSE into graph encryption with support for specific graph queries. E.g., we investigate how to perform shortest distance queries on an encrypted graph.

Dr Jinjun Chen is a Professor from Faculty of Engineering and IT, University of Technology Sydney (UTS), Australia. He is the Director of Lab for Data Systems and Visual Analytics in the Global Big Data Technologies Centre at UTS. He holds a PhD in Information Technology from Swinburne University of Technology, Australia. His research interests include scalability, big data, data science, data intensive systems, cloud computing, workflow management, privacy and security, and related various research topics. His research results have been published in more than 130 papers in international journals and conferences, including ACM Transactions on Software Engineering and Methodology, IEEE TSE, IEEE TPDS, IEEE TCC, IEEE TC, IEEE TSC, and IEEE TKDE.

He received UTS Vice-Chancellor's Awards for Research Excellence Highly Commended (2014), UTS Vice-Chancellor's Awards for Research Excellence Finalist (2013), Swinburne Vice-Chancellor's Research Award (ECR) (2008), IEEE Computer Society (CS) Outstanding Leadership Award (2008-2009) and (2010-2011), IEEE CS Service Award (2007), Swinburne Faculty of ICT Research Thesis Excellence Award (2007). He is an Associate Editor for ACM Computing Surveys, IEEE Transactions on Big Data, IEEE TKDE, IEEE TCC, and other journals such as Journal of Computer and System Sciences, JNCA. He is the Chair of IEEE CS’s Technical Committee on Scalable Computing (TCSC), Vice Chair of Steering Committee of Australasian Symposium on Parallel and Distributed Computing, Founder and Coordinator of IEEE TCSC Technical Area on Big Data and MapReduce, Founder and Steering Committee Co-Chair of IEEE International Conference on Big Data and Cloud Computing, IEEE International Conference on Big Data Science and Engineering, and IEEE International Conference on Data Science and Systems.
Big data analytics has attracted close attention from both industry and academic because of its great benefits in cost reduction and better decision making. As the fast growth of various global services, there is an increasing need for big data analytics across multiple data centers (DCs) located in different countries or regions. It asks for the support of a cross-DC data processing platform optimized for the geo-distributed computing environment. Although some recent efforts have been made for geo-distributed big data analytics, they cannot guarantee predictable job completion time, and would incur excessive traffic over the inter-DC network that is a scarce resource shared by many applications. To fill in this gap, we study on minimizing the inter-DC traffic generated by Map-Reduce jobs targeting on geo-distributed big data, while providing predicted job completion time. To achieve this goal, we formulate an optimization problem by jointly considering input data movement and task placement. Furthermore, we guarantee predictable job completion time by applying the chance-constrained optimization technique, such that the Map-Reduce job can finish within a predefined job completion time with high probability. To evaluate the performance of our proposal, we conduct extensive simulations using real traces generated by a set of queries on Hive. The results show that our proposal can reduce 55% inter-DC traffic compared with centralized processing by aggregating all data to a single data center.

**Song Guo** is currently a full professor at the University of Aizu, Japan. His research interests are mainly in the areas of wireless network, cloud computing, big data, and cyber-physical systems. He has published over 300 papers in international major journals and conferences in these areas. He serves/served in editorial boards of IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Emerging Topics in Computing, IEEE Communications Magazine, Wireless Networks, Wireless Communications and Mobile Computing, and many other major journals. He has been the general/program chair of numerous international conferences. Dr. Guo is a senior member of IEEE, a senior member of ACM, and an IEEE Communications Society Distinguished Lecturer.