Tutorial 1: Transformation to Cloud Computing
John Rofrano & Hari Ramasamy, IBM T.J. Watson Research Center, USA

Abstract: In this tutorial, attendees will learn about IT infrastructure transformation activities both in the general sense and with a focus on transformation to cloud computing. In contrast to incremental, organic, and routine changes to the infrastructure, transformation activities are planned, highly coordinated, large-scale changes. We will provide a systematic perspective of techniques that can be applied to various aspects of IT transformation, especially those representing traditional pain points the presenters have experienced in real-life transformation engagements. The audience will learn about the drivers for transformation, the economics of migration to cloud, and the steps involved in the transformation process. Starting with techniques for assessing the value of Cloud as a target for transformation, we will describe well-proven techniques for discovering the as-is state of the infrastructure to be transformed, analysis of the discovered information, planning for transformation, automating the migration to cloud, and testing. We will highlight both conceptual and practical challenges in migrating workloads to the cloud, as well as best practices.

About the Speakers:

John Rofrano is a published author and Senior Technical Staff Member in the Services Research department at the IBM T.J. Watson Research Center. John’s current focus is on extreme automation techniques for migrating workloads from traditional data centers into the cloud. During his 29-year career at IBM, John has worked as the Chief Architect for various products such as IBM WebSphere Commerce Edition and IBM Visual Warehouse. John has held various key architectural positions across Software Group, Server & Technology Group, Global Technology Services, and Research.

HariGovind (Hari) Ramasamy is a Research Manager in Services Research at the IBM T.J. Watson Research Center, where he leads programs on cloud resiliency and innovation for technical support services. Hari is the recipient of the IBM Outstanding Innovation Award (2012), IBM Research Division Award (2012), IBM Eminence and Excellence Awards (2012, 2011), C.W. Gear Outstanding Graduate Student Award (2002), and a Best Paper Award from IEEE PRDC (Co-Author, 2002). He obtained his Ph.D. degree in Computer Science from the University of Illinois, Urbana-Champaign.

Tutorial 2: Open Source for Cloud Computing and Practices
Tony Wasserman, Carnegie Mellon University - Silicon Valley, USA

Abstract: Since software licenses are approved in 1998, thousands of high quality software products and components have been developed and released under approved licenses. This constantly growing body of high-quality software has driven the growth and acceptance of free and open source software (FOSS). The rapid growth of FOSS has also transformed the software industry, and has led to new businesses that offer FOSS applications, for which they provide paid support services. As the quality and support for FOSS has grown, some governments and corporations have developed policies that include (or sometimes favor) the use of FOSS in their systems. At the same time, FOSS has continued to grow, playing a major role in such areas as cloud computing and management of large datasets. In the context of Cloud Computing, this tutorial describes the benefits of FOSS software for organizations, the ways it can be most effectively used, and the likely future directions for it.

About the Speaker: Anthony I. (Tony) Wasserman is a Professor of Software Management Practice at Carnegie Mellon Silicon Valley, and the Executive Director of its Center for Open Source Investigation (COSI. Tony is both ACM and IEEE Fellow. Prior to CMU, Tony was a Professor at UC San Francisco, CEO of Interactive Development Environments (IDE), VP of Engineering for a dot-com and later became VP of Bluestone Software, and Director at HP. At IDE, he released Pictures multiuser modeling environment, among the very first commercial products to include open source software. Tony is very active
in the international open source research community, and served as General Chair of the 2009 Int'l. Conf. on Open Source Systems. He is on the Board of Directors of the Open Source Initiative (OSI) and the Board of Advisors of Open Source for America. Within the last year, Tony received the 2013 ACM's SIGSOFT Influential Educator Award, and 2012 Distinguished Educator Award from the IEE's Technical Council on Software Engineering. Tony earned his Ph.D degree in Computer Science from University of Wisconsin – Madison.

**Tutorial 3: Big data Techniques for Process Analysis**

Marcello Leida, Etisalat BT Innovation Center (EBTIC), U.A.E.

**Abstract:** Modern Business process analysis requires an extremely flexible data model and a platform able to minimize response times as much as possible. In order to efficiently analyze a large amount of data, this tutorial illustrates novel technologies that rely on an improved data model supported by a grid infrastructure, allowing storing the data in-memory across many grid nodes and distributing the workload, avoiding the bottleneck represented by constantly querying a traditional database. Both process data and domain knowledge are represented using standard metadata formats: process logs are stored as RDF triples referring to company-specific activity ontologies: RDF triples assign meaning to the process logs, ensuring that they can be universally understood and reusable. The data collected by the process log monitor is translated to a continuous flow of triples that capture the status of the processes. This continuous flow of information can be accessed through the SPARQL query language used to extract and analyze process execution data. Although the query engine has been developed as part of a Business Process Monitoring platform, it is a general purpose engine that can be used in any system that requires scalable analysis of semantic data. The system presented has some unique features such as grid-based infrastructure, extreme scalability, efficient real-time query answering and an on the fly access control layer that will be presented in detail during the tutorial together with future research directions. The tutorial will start from the motivations behind the research, it will introduce the data model used to represent the business processes and finally it will describe in detail the various elements and technologies underpinning the architecture of the developed system.

**About the Speaker: Marcello Leida** received his Ph.D. in Computer Science from the Universita' degli Studi di Milano. His currently works in the Enterprise and Distributed Applications area: leading a research project on Real-time Business Process Analysis. He is responsible for the down-streaming and physical deployment and maintenance in enterprise servers of the research projects developed by the team to production environments and integration of our systems with existing Business Intelligence solutions. His research interests are mainly focused on Semantic Web and Linked Data, Big Data, automatic data visualization techniques, business process analysis and artificial intelligence in general; he published a book, several papers, workshop and journals on these topics. He has been program committee and referee for several conferences and journals and member of IFIP W.G. 2.6 on database semantics.

**Tutorial 4: In Memory Database for Big Data**

Jordan Cao, SAP, USA

**Abstract:** We have entered the era of "big data," where the volume, velocity, and variety of data are exploding at an unprecedented pace. However, the term of Big Data is not clear. It’s actually not primarily about ‘big’, but all about real-time data. The volume and variety are only important because of their impact on velocity. Real-time is the new “real” Big. Now, enterprises are looking for real-time solutions to dramatically accelerate the time from data to decision by minimizing the delay between the transaction and decision. This requires a new approach: in-memory big data solution. This tutorial presents an overview of the in-memory big data industrialization movement, along with real-life working examples and use cases as well as best practices. The tutorial consists of five parts. 1) Fundamental: in-memory big data platform concept, characterization, attributes, value, history, major players, foundational building blocks; 2) Framework: in-memory big data framework, data repository classification, data ingestion, streaming, data compression, scale-out options, and integration; 3) Solutions: techniques, styles, patterns, design, parallel processing, compaction, bulk loading, optimization, real-time, and analytic; 4) Cloud offering: in-memory big data platform on cloud, data loading, IaaS and PaaS; 5) Case study: real-world business scenarios.

**About the speaker: Jordan Cao** is a Sr. product marketing manager for SAP HANA. He has more than 15 years of experience in computing science, including roles as an SAP senior architect and solution manager. Jordan has participated in multiple SAP projects focused on scientific programming, cloud computing, big data, and services oriented related research and works. He has a PhD in cloud computing, service-oriented architecture, and software engineering, as well as an MBA degree both from Arizona State University.
Tutorial 5: Opportunities and Risk of Outcome Based Business
Nianjun (Joe) Zhou, IBM T.J. Watson Research Center, USA

Abstract: Outcome-based business (OBB) is a business practice that links a service provider’s income to the values delivered by the IT and other capabilities to the client. An OBB engagement typically requires establishing a long-term relationship with the client and reduces the risk for the client during the business transformation. A successful OBB engagement requires the service provider owning a deep understanding of the client’s business and the corresponding business value drivers to achieve the desired goal. As a natural evolution of SOA, outcome-based business further requires the IT enabled have to be goal-oriented. In this talk, we will have an overview of outcome-based business, its philosophies and practices. Then we discuss how to use service-value-map to capture the causality relationship to relationship from IT assets and other business initiatives to business goals. We will discuss how to apply this to different industries, education, health care, retails and telecommunication. We will discuss how to use gap analysis to identify gaps in the client’s financial performance, and identify the business capabilities and IT assets that can improve those financial gaps. We category and discuss the risk of OBB from services provider’s perspective. Based on various contract constructions, we discuss the risk in term of fee-at-risk and potential return of investment. Several characteristic variables are defined to capture the risk factor into a quantitative manner. We will also discuss we can evaluate the risk factors through the contract characteristics (like contract term and payback curve) and service-value-map capturing the causality relationship between the measured KPIs and IT assets selected.

About the Speaker: Dr. Nianjun Zhou is a Research Staff Member of IBM T. J. Watson Research Center. He served as the IBM Professional Interest Group Chair for Services Computing from 2010 to 2012. He is selected as a member of 2013 Best of IBM. His current research areas mainly focus on services sciences and service computing to achieve services optimization with information and analytic solutions. Dr. Zhou has many papers and patents in ad hoc and sensor networks, content management, service-oriented architecture, project estimation, and risk assessment. He is serving as Program Committee Co-Chair of IEEE Big Data 2013, and has severed in program communities for IEEE Cloud, ICWS and SCC for several years. Dr. Zhou held a Ph.D degree from RPI, and Master and Bachelor degrees from Peking University.

Tutorial 6: Compute to Data
Masood Mortazavi, Huawei, USA

Abstract: This tutorial includes presentation and discussion of topics related to two technological aspects of cloud computing -- compute and data. By "compute" we mean those facilities in the cloud that enable the running of application logic above an operating system or a container environment. We will discuss the classical, cloud computing environments offered by the leading cloud computing service providers. We will discuss some case studies and look at how these compute technologies have evolved in order to address a wider range of market requirements. By "data" we mean those facilities in the cloud that enable the storage of data for future retrieval, query and manipulation. We will discuss various storage facilities (e.g. file, object and record storage as well as volume services) offered by leading cloud storage providers, and will attempt to discern the trade-offs that each of these facilities have made in order to serve their market.

About the Speaker: Dr. Masood Mortazavi is a distinguished engineer at Huawei Technologies. Earlier, Masood worked in the Cloud Platform Group at Yahoo and led advanced projects related to Yahoo's structured record storage, messaging and application container services. Masood led an international group of engineers at Sun Microsystems, focused on the development of open-source databases such as Apache Derby, PostgreSQL and MySQL. At Huawei's Innovation Center in Santa Clara, Masood works on distributed databases and cloud storage systems.