

# **International Symposium on Grids & Clouds 2016 (ISGC 2016)**

**Sunday, March 13, 2016 - Friday, March 18, 2016**

**Academia Sinica**

## **Scientific Program**

Contemporary research is impossible without a strong IT component – researchers rely on the existence of stable and widely available e-infrastructures and their higher level functions and properties. As a result of these expectations, e-Infrastructures are becoming ubiquitous, providing an environment that supports large scale collaborations that deal with global challenges as well as smaller and temporal research communities focusing on particular scientific problems. To support those diversified communities and their needs, the e-Infrastructures themselves are becoming more layered and multifaceted, supporting larger groups of applications. ISGC 2016 continues its aim to bring together users and application developers with those responsible for the development and operation of multi-purpose ubiquitous e-Infrastructures.

The proceedings of the International Symposium on Grids & Clouds (ISGC 2016) will be published in the open access Proceedings of Science (PoS) by SISSA, the International School for Advanced Studies of Trieste. All papers will be free to read and download immediately upon publication. Papers published online will be hosted in PoS electronic journals system and will be fully citable.

## **Keynote Talk**

## **Country Report**

## **Physics (including HEP) and Engineering Applications**

Submissions should report on experience with physics and engineering applications that exploit grid and cloud computing services, applications that are planned or under development, or application tools and methodologies. Topics of interest include: (1) End-user data analysis; (2) Management of distributed data; (3) Applications level monitoring; (4) Performance analysis and system tuning; (5) Workload scheduling; (6) Management of an experimental collaboration as a virtual organization; (7) Comparison between grid and other distributed computing paradigms as enablers of physics data handling and analysis; (8) Expectations for the evolution of computing models drawn from recent experience handling extremely large and geographically diverse datasets.

## **Biomedicine & Life Sciences Applications**

During the last decade, research in Biomedicine and Life Sciences has dramatically changed thanks to the continuous developments in High Performance Computing and highly Distributed Computing Infrastructures such as grids and clouds. This track aims at discussing problems, solutions and application examples related to this area of research, with a particular focus on non-technical end users. Submissions should concentrate on practical applications and solutions in the fields of Biomedicine and Life Sciences, such as (1) Nano-medicine, (2) Drug discovery, (3) Structural biology and / bioinformatics, (4) Medical imaging, (5) Public health applications / infrastructures, (5) High throughput data processing/analysis, (6) Integration of semantically diverse data sets and applications, (7) Cloud-based application examples, (8) Distributed data, computing and services, and (9) Data management issues.

## **Earth & Environmental Sciences & Biodiversity Applications**

Natural and Environmental sciences are placing an increasing emphasis on the understanding of the Earth as a single, highly complex, coupled system with living and dead organisms. It is well accepted, for example, that the feedbacks involving oceanic and atmospheric processes can have major consequences for the long-term development of the climate system, which in turn affects biodiversity, natural hazards and can control the development of the cryosphere and lithosphere. Natural disaster mitigation is one of the most critical regional issues in Asia. Despite the diversity of environmental sciences, many projects share the same significant challenges. These include the collection of data from multiple distributed sensors (potentially in very remote locations), the management of large low-level data sets, the requirement for metadata fully specifying how, when and where the data were collected, and the post-processing of those low-level data into higher-level data products which need to be presented to scientific users in a concise and intuitive form. This session would in particular address how these challenges are being handled with the aids of e-Science paradigm.

## **Humanities, Arts, and Social Sciences (HASS) Applications**

Disciplines across the Humanities, Arts and Social Sciences (HASS) have critically engaged with technological innovations such as grid- and cloud computing, and, most recently, various 'big data' technologies. The increasing availability of 'born digital' data has led to an increasing interest in analysis methods such as natural language processing, social network analysis or machine learning. These developments pose challenges as well as opening up opportunities and members of the HASS community have been at the forefront of discussions about the impact that novel forms of data, novel computational infrastructures and novel analytical methods have for the pursuit of science endeavours and our understanding of what science is and can be. The ISGC 2016 HASS track invites papers and presentations covering applications demonstrating the opportunities of new technologies or critically engaging with their methodological implications in the Humanities, Arts and Social Sciences. We also invite contributions that critically reflect on the impact that ubiquitous and mobile access to information and communication technologies have for society more generally, especially around topics such as smart cities, civic engagement or digital journalism.

## **Virtual Research Environment (including Middleware, tools, services, workflow, ... etc.)**

Virtual Research Environments (VRE) provide an intuitive, easy-to-use and secure access to federated computing resources for solving scientific problems, trying to hide the complexity of the underlying infrastructure, the heterogeneity of the resources, and the interconnecting middleware. Behind the scenes, VREs comprise tools, middleware and portal technologies, workflow automation as well as security solutions for layered and multifaceted applications. Topics of interest include but are not limited to: (1) Real-world experiences building and/or using VREs to gain new scientific knowledge; (2) Middleware technologies, tools, services beyond the state-of-the-art for VREs; (3) Innovative technologies to enable VREs on arbitrary devices, including Internet-of-Things; and (4) One-step-ahead workflow integration and automation in VREs.

## **Data Management**

Data management encompasses the organization, distribution, storage, access, validation, and processing of digital assets. Data management requirements can be characterized by data life stages that evolve from shared project collections, to formally published libraries, to archives of reference collections. Papers are sought that demonstrate the management of data through the multiple phases of the data life cycle, from creation to re-use. Of particular interest are demonstrations of systems that process massive data collections.

## **Big Data**

The continuous innovation of the analysis model, resource federation, distributed infrastructure, dynamic data and computing, workflow management, and smart applications are all vital to the storage, management and analysis of Big Data. Characterized by Volume, Variety and Velocity, Big data is creating a new generation of scientific process and the discovery mechanism. In order to maximise the data values and to speed up novel discoveries, multidisciplinary approach, dynamic provisioning of storage resources and long-term data preservation are primary focuses for the moment. This track aims to attract novel research, application and technology on Big Data analysis. Submissions should focus on both the conceptual modelling as well as the techniques related to Big Data analytics.

## **Networking, Security, Infrastructure & Operations**

Networking and the connected e-Infrastructures are becoming ubiquitous. Ensuring the smooth operation and integrity of the services for research communities in a rapidly changing environment are key challenges. This track focusses on the current state of the art and recent advances in these areas: networking, infrastructure, operations, and security. The scope of this track includes advances in high-performance networking (software defined networks, community private networks, the IPv4 to IPv6 transition, cross-domain provisioning), the connected data and compute infrastructures (storage and compute systems architectures, improving service and site reliability, interoperability between infrastructures, data centre models), monitoring tools and metrics, service management (ITIL and SLAs), and infrastructure/systems operations and management. Also included here are issues related to the integrity, reliability, and security of services and data: developments in security middleware, operational security, security policy, federated identity management, and community management. Submissions should address solutions in at least one of these areas.

## **Infrastructure Clouds and Virtualisation**

This track will focus on the use of cloud computing, mainly but not exclusively Infrastructure-as-a-Service (IaaS) and virtualization technologies in large-scale distributed computing environments in science and technology. We solicit papers describing underlying virtualization and "cloud" technology, scientific applications and case studies related to using such technology in large scale infrastructure as well as solutions overcoming challenges and leveraging opportunities in this setting. Of particular interest are results exploring usability of virtualization and infrastructure clouds from the perspective of scientific applications, the performance, reliability and fault-tolerance of solutions used, data management issues. Papers dealing with the cost, price, and cloud markets, with security and privacy, as well as portability and standards, are also most welcome.

## **Interoperability**

Interoperability of federated ICT-infrastructures (with Grid, Cloud, HPC, HTC, HPTC, data or network resources) are mandatory to address the needs of modern e-science. Researchers from all scientific disciplines make intensive use of multiple resource types in various e-infrastructures on the regional, national and international level. To enable an easy and intuitive access, interoperability is of major importance. Interoperability includes a spectrum from rapidly-prototyped solutions to open-standards based interoperability of components interfering via standardized interfaces. Topics of interest include but are not limited to: (1) Real-world production use cases of scientific applications using standards-based, interoperable ICT-infrastructures; (2) Methods, design principles and solutions for standards-based interoperability; and (3) Operational security solutions enabling secure interoperability.

## **Business Models & Sustainability**

Understanding how a particular e-Infrastructure component can be created and sustained requires answering two pairs of questions: What resources are needed to create it, and how can those resources be assembled and applied? And What resources are needed to sustain it, and how can those resources be assembled and applied? This track seeks contributions around business models and sustainability relating to e-Infrastructure components including: (1) Business models around e-Infrastructures and their components, (2) Sustainability of e-Infrastructure components, (3) Initiatives to understand the cost of delivering e-Infrastructures components, (4) Planning strategies and methodologies around e-Infrastructure components, and (5) Planning strategies and methodologies around e-Infrastructure components.

## **Massively Distributed Computing and Citizen Sciences**

This track welcomes contributions dealing with technologies, concepts and applications for very large distributed systems, desktop grids and volunteer computing. Additional focus will be on support of citizen science, making the massively distributed computing systems available to individual researchers through e.g. Science gateways and other kinds of portals, specific interfaces to connect and use the systems, but also new ways how to contribute and to combine volunteered and institutional computing resources. The topics will cover new technologies of the related software frameworks, recent application developments, as well as infrastructure operation and user support techniques. Special focus will be on the (1) Interoperability with other and integration in other e-infrastructures (2) Data management and (3) Quality of service in such environments (4) Novel uses of volunteer computing and Desktop Grid (5) Best practices and (social) impacts.

## **High Throughput & Supercomputing Systems and their Integration**

With the growing availability of powerful computing resources through public grids (e.g., EGI and OSG) and public/private clouds (e.g., Amazon EC2), as well as through coordinated access to supercomputing resources (e.g. PRACE), it has become possible to develop and deploy applications that exploit many and very different powerful computing resources as possible. These resources include High Performance and High Throughput Computing (HPTC) infrastructures, comprised of supercomputers and clusters, using general purpose, accelerators, GPUs and many-core processors. However, despite many developments, it is still quite challenging to effectively access, aggregate and manage the variety of available resources, which under control by different resource providers. This session will solicit recent research and development achievements and best practices in exploiting these computing resources available around the

world. The topics of interest include, but are not limited to the followings: (1) Experiences, use cases and best practices on the development and operation of large-scale HPTC applications (2) Delivery of and access to HPTC resources through grid and cloud computing (as a Service) models (3) Integration and interoperability to support coordinated federated use of different HPTC e-infrastructures (4) Use of virtualization techniques to support portability across different HPTC systems (5) Robustness and reliability of HPTC applications and systems over a long-time scale.

## **Plenary**

### **EGI Federated Cloud for Open Science Tutorial**