International Symposium on Grids & Clouds 2017 (ISGC 2017)

Sunday, 5 March 2017 - Friday, 10 March 2017

BHSS, Academia Sinica

Scientific Programme

Global Challenges: From Open Data to Open Science

The unprecedented progress in ICT has transformed the way education is conducted and re-search is carried out. The emerging global e-Infrastructure, championed by global science communities such as High Energy Physics, Astronomy, and Bio-medicine, must permeate into other sciences. Many areas, such as climate change, disaster mitigation, and human sustaina-bility and well-being, represent global challenges where collaboration over e-Infrastructure will presumably help resolve the common problems of the people who are impacted. Access to global e-Infrastructure helps also the less globally organized, long-tail sciences, with their own collaboration challenges.

Open data are not only a political phenomenon serving government transparency; they also create an opportunity to eliminate access barriers to all scientific data, specifically data from global sciences and regional data that concern natural phenomena and people. In this regard, the purpose of open data is to improve sciences, accelerating specifically those that may bene-fit people. Nevertheless, to eliminate barriers to open data is itself a daunting task and the bar-riers to individuals, institutions and big collaborations are manifold.

Open science is a step beyond open data, where the tools and understanding of scientific data must be made available to whoever is interested to participate in such scientific research. The promotion of open science may change the academic tradition practiced over the past few hundred years. This change of dynamics may contribute to the resolution of common chal-lenges of human sustainability where the current pace of scientific progress is not sufficiently fast.

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, but also in big-data solutions to deal with the explosion in genomic data. 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 Drug discovery, Structural biology, Bioinformatics, Medical imaging, Public health applications / infrastructures, High throughput (grid and cloud-based) data processing/analysis, Distributed data computing and services, and Big data management issues. Submissions should ideally highlight how the availability and use of Big Data has enabled new processes for or dramatically evolved the scope of their research.

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 data analytic 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, machine learning and text mining. 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 2017 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 following subjects: (1) 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, and digital journalism; (2) philosophical and methodological reflections on the development of the techniques and the approaches by which data scientists use to pursue knowledge.

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.

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 a 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 & Big Data

The rapid growth of the data available to scientists and scholars – in terms of Velocity and Variety as well as sheer Volume – is transforming research across disciplines. Increasingly these data sets are generated not just through experiments, but as a byproduct of our day-to-day digital lives. This track explores the consequences of this growth, and encourages submissions relating to two aspects in particular: firstly, the conceptual models and analytical techniques required to process data at scale; secondly, approaches and tools for managing and creating these digital assets throughout their lifecycle.

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.

Business Models, Policy & Long-term 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, how can those resources be assembled and sustained by the different stakeholders? This track seeks contributions around business models, policy and long-term 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, (5) Economic engagement of user communities and policy makers and (6) how to go beyond local, national or community-specific financial boundaries and think global.

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.

Supercomputing, High Throughput, Accelerator Technologies and Integrations

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 becomes 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

Keynote