

**International Symposium on  
Grids & Clouds (ISGC) 2023 in  
conjunction with HEPiX  
Spring 2023 Workshop**

**Report of Contributions**

Contribution ID: 1

Type: **Oral Presentation**

## Novel Immersive Sounds Interactivity using GIS Maps Application

*Thursday, March 23, 2023 2:30 PM (30 minutes)*

Among the five central senses we use to perceive the world around us, nothing is more salient than our sense of hearing. Sounds play a very important role in how we understand, behave and interact with the world around us. One can close their eyes, but never their ears. In this research study, we propose design and development of a GIS-based maps application that would allow users to not only navigate, and see pictures of landmark locations in their urban city-environment, but also enable them to hear the immersive spatial sounds of the area. This would add a new interaction paradigm to mobile GIS application design which still has potential to be explored and utilized for better interaction, immersion and usability to the users imparting the final goal of layering additional scopes of realism in GIS map application.

This study will be divided into four phases: initial background research and literature review. Following that, a pilot experiment would be ran with the main motive of understanding what underlying behavior and task-specific factors affect users whilst searching for a landmark tourist destination in city. Once these factors are well understood, user journey maps will be plotted and an application would be developed with service design and human-centered approach. This application would allow users to find selected landmark tourist destinations in the city of Taipei, route and navigate to it, see pictures of the site and most importantly have a feature for listening to the immersive spatial sounds of the place. This adds a novel interaction layer on GIS Maps application with immersive sounds – making the entire process of searching for, and selecting destinations to visit more immersive, enjoyable and authentic.

The final part of the study would focus on evaluating this application's performance, usability and immersion with System Usability Scale (SUS), and Analytic Hierarchy Process (AHP) with input-oriented Data Envelopment Analysis (DEA)

**Primary author:** WANG, Ryan Sheng-Ming (Department of Interaction Design, National Taipei University of Technology)

**Co-author:** PRASAD, Kunal (National Taipei University of Technology (NTUT))

**Presenter:** PRASAD, Kunal (National Taipei University of Technology (NTUT))

**Session Classification:** Data Management & Big Data

Contribution ID: 2

Type: **Oral Presentation**

## A Token based solutions from KIT for SSH with OIDC

*Thursday, March 23, 2023 1:30 PM (30 minutes)*

OIDC (OpenID Connect) is widely used for transforming our digital infrastructures (e-Infrastructures, HPC, Storage, Cloud, ...) into the token based world.

OIDC is an authentication protocol that allows users to be authenticated with an external, trusted identity provider. Although typically meant for web- based applications, there is an increasing need for integrating shell- based services.

This contribution delivers an overview of several tools, each of which provides a solution to a specific aspect of using tokens on the commandline in production services:

- `oidc-agent` is the tool for obtaining `oidc-access` tokens on the commandline. It focuses on security and manages to provide ease of use at the same time. The agent operates on a users workstation or laptop and is well integrated with graphical user interfaces of several operating systems, such as Linux, MacOS, and Windows. Advanced features include agent-forwarding which allows users to securely obtain access tokens from remote machines to which they are logged in.
- `mytoken` is both, a server software and a new token type. Mytokens allow obtaining access tokens for long time spans, of up to multiple years. It introduces the concept of “capabilities” and “restrictions” to limit the power of long living tokens. It is designed to solve difficult use-cases such as computing jobs that are queued for hours before they run for days. Running (and storing the output of) such a job is straightforward, reasonably secure, and fully automisable using `mytoken`.
- `pam-ssh-oidc` is a pam module that allows accepting access tokens in the Unix pluggable authentication system. This allows using access tokens for example in ssh sessions or other unix applications such as `su`. Our pam module allows verification of the access token via OIDC or via 3rd party REST interfaces.
- `motley-cue` is a REST based service that works together with `pam-ssh-oidc` to validate access tokens. Along the validation of access tokens, `motley-cue` may - depending on the enabled features - perform additional useful steps in the “SSH via OIDC” use-case. These include
  - Authorisation (based on VO membership)
  - Authorisation (based on identity assurance)
  - Dynamic user creation
  - One-time-password generation (in case the access token is too long for the SSH-client used)
  - Account provisioning via plugin based system (interfaces with local Unix accounts, LDAP accounts, and external REST interfaces)
  - Account blocking (by authorised administrators in case of a security incident)

- mccli is a client side tool that enables clients to use OIDC access-tokens that normally do not support them. Currently, ssh, sftp and scp are supported protocols.
- oidc-plugin for putty makes use of the new putty plugin interface to use access tokens for authentication, whenever an ssh-server supports it. The plugin interfaces with oidc-agent for windows to obtain tokens.

The combination of the tools presented allows creative new ways of using the new token-based AAs with old and new tools. Given enough time, this contribution will include live-demos for all of the presented tools.

**Primary authors:** Dr GUDU, Diana (KIT); Dr HARDT, Marcus (KIT); Mr ZACHMANN, Gabriel (KIT)

**Presenter:** Dr HARDT, Marcus (KIT)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 3

Type: **Oral Presentation**

## Cagliari 2020: exploiting open data and LHC computing techniques for Smart Cities

*Friday, March 24, 2023 9:00 AM (30 minutes)*

CAGLIARI 2020 is a 25 million euro project funded within the framework of the National Operational Program for Research and Competitiveness – Smart Cities & Communities of the Italian Ministry of Education University and Research.

The project started in 2017 and ended in 2022 developing a pilot system for monitoring traffic and air quality providing innovative and environmentally friendly solutions for urban mobility.

The system exploited data acquisition and computing techniques developed in the context of the LHC experiments.

The partnership includes public and private organisms of the South Sardinia for the development of ICT technologies aimed at optimizing the usage of the “city system” and improving the quality of life for people working and living in the city.

The developed system aims at answering the ever increasing need of innovative tools and technological solutions for the optimization of urban mobility. The approach is based on collecting data traffic flow as well as environmental parameters merging data from different sources and combining them in order to obtain increased security lower travel times and improved air quality. These data are open and available both to operators and managers as well as to citizens. Critical events management is also included. Integration of data from different sources and availability to multiple users are key points in the project.

The pilot system developed consists of a sensors network comprised of:

1. Fixed sensors for the tracking of vehicles. These sensors allow real-time and/or historical analysis, especially helpful in gathering the information required to manage traffic lights systems and sending routing optimization information to interested users;
2. fixed and mobile sensors for the collection of environmental data. Such data will be used to feed decision-making models for the reduction of carbon emissions and the consequent improvement of air quality in the urban area.
3. Mobile devices for the acquisition of the motion habits of people.

The integration of environmental models and smart systems for the management of urban mobility allows optimizing public and private traffic flows as well as to reduce carbon emissions.

CAGLIARI 2020 concept is related to the application of the “netcentric” paradigm by means of a dynamic and pervasive net (the urban information grid) whose nodes can be both fixed and mobile. This feature allows the sensorial integration of the devices distributed in the urban area and turns public transport buses into “mobile platforms” for the urban road system monitoring thanks to the continuous gathering of traffic, carbon emissions and noise pollution data. It is therefore possible to develop models for the analysis of environmental parameters and to provide support tools to policies aimed at curbing traffic flows, energy consumptions, and carbon emissions within urban areas.

Merging of data from multiple sources processing them and making them interoperable and usable to multiple clients is a core element in the Project.

The integration between the aforementioned information and the people’s traveling habits (by means of the anonymous tracking of their mobile phones) allows for the creation of people’s mobility maps.

Cloud services play a key role within the project in supporting the applications dedicated to data

traffic monitoring and analysis. A mixed cloud approach has been adopted with data acquisition services and mediation layer on private cloud and analysis and data fusion on commercial cloud. A micro services approach has been adopted and it is currently operational. The system is scalable and fully interoperable.

**Primary authors:** MASONI, Alberto (INFN National Institute of Nuclear Physics); CARDINI, Alessandro (INFN Sezione di Cagliari - Italy); PUGGIONI, Carlo (INFN Sezione di Cagliari - Italy); Mr MURA, Daniele (INFN Sezione di Cagliari - Italy); Mr ARGIOLAS, Filippo (INFN Sezione di Cagliari - Italy); CASU, Luigi (INFN Sezione di Cagliari - Italy); LA DELFA, Luigi (INFN Sezione di Cagliari - Italy); TACCORI, Mattia (INFN Sezione di Cagliari - Italy); ARESTI, Mauro (INFN Sezione di Cagliari - Italy); FLORE, Walter (INFN Sezione di Cagliari - Italy)

**Presenter:** MASONI, Alberto (INFN National Institute of Nuclear Physics)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 4

Type: **Oral Presentation**

## **Learning Cloud & Grid Computing through implementing innovative solutions to real-world scenarios**

In this paper, we shared on how students and staff from a Singapore tertiary institution acquired and reinforced their learning of Cloud Computing & Grid technologies through participation in applied projects on real-life scenarios in Healthcare, Mobile etc. Centre for IT Innovation (CITI) at School of Information Technology (SIT), Nanyang Polytechnic, is a key platform for staff and students to develop innovative and effective IT Software solutions and services for the industry. A hallmark of SIT is our strong collaboration with industry to provide effective and relevant learning experience to our students.

NYP's practice- and application-oriented teaching and learning philosophy emulates the real-world work environment on campus to provide students with an enriching experience which will prepare them well for their future career. To ensure quality deliverables and adhere to the project schedule, a software engineering process based on Agile/DevOps is being adopted. An important and integral component is industry project work with cost, quality, reliability and deadline constraints, and often requiring multidisciplinary capabilities. The accumulated experience and knowledge from extensive industry project development is systematically captured in a rich repository of application-oriented and established solutions over a wide range of IT and other disciplines.

The paper will share case studies on 3 projects that have been developed by students with staff mentorship in resolving real-life problems and how they applied the knowledge of Cloud and Grid learned in class.

### Case 1: SOSHeart.

One out of three deaths in Singapore is due to heart disease and stroke. Early CPR has been shown to be a vital link in the "Chain of Survival" for out-of-hospital cardiac arrest. One gap is the absence of means for heart attack victims to get the prompt treatment from a CPR-certified citizens. The study aims to develop a mobile app, SOSHeart, to provide location - based services to heart patients to call for help from caregivers as well as providing the personal medical data and other useful information in heart attack emergency. SOSHeart is an Android-based mobile app with cloud support, services and repository

### Case 2 : Presence-Aware Grid Computing

A conventional grid computing allows members to share computing resources in a network of computers in a distributed and coordinated manner. Accessing the grid network is usually done at a user's workstation, either via a command line or web-based interface. The complexity of sharing computing resources in a grid network is hidden from the users; there is no intuitive means for users to assess or monitor the grid network utilization before submitting a job. Thus we are proposing for a new system for implementing grid computing submission, monitoring and scheduling using Instant Messaging in a manner that can minimise if not overcome the limitations of conventional grid computing system

### Case 3: Dementia Patients Locator

With the increase in Singapore's ageing population, comes an increase in the number of persons with dementia. One in 10 seniors aged 60 and above is expected to have dementia. Our solution leverages on cloud, mobile, communication and geospatial technology to build an interactive platform where caregivers can report a lost person with dementia and establish communication channels with the community who may have found his/her loved ones. Finder can provide instant information of the lost person to the caregiver by raising sightings and are informed of the nearest

safe-return point for them to bring the dementia patients to if they have found him/her. Compared with the existing solutions leveraging on social media e.g. Facebook or traditional announcements, e.g. putting up posters, the solution is less dependent on the social network of the caregiver or the manual efforts that they may put in.

**Primary author:** TAN, Nam Beng (Nanyang Polytechnic)

**Presenter:** TAN, Nam Beng (Nanyang Polytechnic)

**Session Classification:** VRE

**Track Classification:** Track 5: Virtual Research Environment (including tools, services, workflows, portals, ... etc.)



Contribution ID: 5

Type: **Oral Presentation**

## **HEPS virtual cloud desktop system based on Openstack, design and implementation (Remote presentation)**

*Wednesday, March 22, 2023 3:00 PM (20 minutes)*

High Energy Photon Source (HEPS) will generate massive experimental data for diversified scientific analysis. The traditional way of data download and analysis by users using local computing environment cannot meet the growing demand for experiments. This paper proposes a virtual cloud desktop system of HEPS based on Openstack, which is used for imaging and crystal scattering experiments in HEPS. Such experiments are characterized by high requirements on image display effect and need high-performance GPU calculation and image rendering, but the experimental site does not have such conditions. In this paper, HEPS virtual cloud desktop system based on Openstack is proposed for imaging and crystal scattering experiments in HEPS, which are characterized by high requirements for image display effect and require high-performance GPU computing and image rendering, but the experimental site does not have this condition. Therefore, how to use the resources of the computing cluster to provide virtual cloud desktops is particularly important. Firstly, we introduce the related research work of this system, including the optimization of Openstack virtual machine, the performance test of light source experiment on virtual machine. Then we introduce the architecture of virtual cloud desktop in detail, PCI pass through and virtual GPU, import of experimental data, storage and export of experimental results, data and network security policies. Finally, we give the actual application of virtual cloud desktop system in the light source experiment, and show the superiority of virtual cloud desktop technology and good application prospects in the field of synchrotron radiation light source.

**Primary authors:** Mr XU, Jiping (IHEP); Mr HU, Qingbao (IHEP)

**Presenter:** Mr XU, Jiping (IHEP)

**Session Classification:** Converging Infrastructure Clouds, Virtualisation & HPC

**Track Classification:** Track 8: Infrastructure Clouds and Virtualizations

Contribution ID: 6

Type: **Oral Presentation**

## A Virtual Research Environment for Integrative Modelling of Biomolecular Complexes with the New Modular Version of HADDOCK.

*Wednesday, March 22, 2023 4:00 PM (30 minutes)*

The prediction of the quaternary structure of biomolecular macromolecules is of paramount importance for fundamental understanding of cellular processes and drug design. In the era of integrative structural biology, one way of increasing the accuracy of modelling methods used to predict the structure of biomolecular complexes is to include as much experimental or predictive information as possible in the process. We have developed for this purpose a versatile information-driven docking approach HADDOCK (<https://www.bonvinlab.org/software>) available as a web service at <https://wenmr.science.uu.nl/haddock2.4>. HADDOCK can integrate information derived from biochemical, biophysical or bioinformatics methods to guide the modelling.

In the context of the BioExcel Center of Excellence for Computational Biomolecular Research (<https://bioexcel.eu>), we have developed HADDOCK3, the new modular version of HADDOCK. It represents a redesign of the HADDOCK2.X series, implementing new ways to interact with the HADDOCK sub-routines and offering more customization. Users can create custom workflows by combining different modules, thus making the workflows tailored to their specific needs. HADDOCK3 has therefore developed to truthfully work like a puzzle of many pieces (simulation modules) that users can combine to more accurately model their systems. The HADDOCK3 workflows are defined in straightforward configuration files, similar to the TOML format (also supported).

In order to facilitate the use of HADDOCK3, in collaboration with the Netherlands eScience Center (<https://www.esciencecenter.nl>) we are developing a customizable, interactive, HTC/Cloud (and HPC)-optimized and reusable Virtual Research Environment for Integrative Modelling of Biomolecular Complexes (<https://github.com/i-VRESSE>), which will consist of three main segments: A workflow builder GUI, an execution layer, and an analysis/storage/sharing workspace. By integrating all steps involved in studying biomolecular interactions, this VRE will lower the steep learning curve for researchers and students from different fields and contribute to reproducible research and FAIR sharing of data.

In this presentation, I will introduce HADDOCK3 and discuss the status of the Virtual Research Environment for Integrative Modelling of Biomolecular Complexes.

**Primary authors:** Dr HONORATO, Rodrigo (Utrecht University); Dr GIULINI, Marco (Utrecht University); Dr TEIXEIRA, João (Utrecht University); Dr VERHOEVEN, Stefan (Netherlands eScience Center); Dr BARBARA, Vrede (Netherlands eScience Center); Dr KALVERA, Peter (Netherlands eScience Center); BONVIN, Alexandre M.J.J. (Utrecht University)

**Presenter:** BONVIN, Alexandre M.J.J. (Utrecht University)

**Session Classification:** VRE

**Track Classification:** Track 5: Virtual Research Environment (including tools, services, workflows, portals, ... etc.)

Contribution ID: 7

Type: **Oral Presentation**

## **Collaborative operational security for Research and Education (Remote presentation)**

*Tuesday, March 21, 2023 5:00 PM (30 minutes)*

We must protect and defend our environment against the cybersecurity threats to the research and education community, which are now acute having grown in recent years. In the face of determined and well-resourced attackers, we must actively collaborate in this effort across HEP and more broadly across Research and Academia (R&E).

Parallel efforts are necessary to appropriately respond to this requirement. We must both share threat intelligence about ongoing cybersecurity incidents with our trusted partners, and deploy the fine-grained security network monitoring necessary to make active use of this intelligence. We must also engage with senior management in our organisations to ensure that we work alongside any broader organisational cybersecurity development programmes.

We report on recent developments in the Security Operations Centres (SOC) Working Group, established by the WLCG but with membership encompassing the R&E sector. The goal of the Working Group is to develop reference designs for SOC deployments and empower R&E organisations to collect, leverage and act upon targeted, contextualised, actionable threat intelligence. This report will include recent experience in deploying SOC capabilities for the first time including network topology considerations, hardware and software provisioning strategies. We also report on experience using containerised SOC training environments in the security thematic CERN School of Computing held in Split in the summer of 2022.

Finally, we discuss ongoing work in the broader community to develop common practices in cybersecurity to support our common response in this area.

**Primary authors:** CROOKS, David (UKRI STFC); VALSAN, Liviu (CERN)

**Presenters:** CROOKS, David (UKRI STFC); VALSAN, Liviu (CERN)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 8

Type: **Oral Presentation**

## Speeding up Science Through Parametric Optimization on HPC Clusters

*Wednesday, March 22, 2023 2:20 PM (20 minutes)*

Science is constantly encountering parametric optimization problems whose computer-aided solutions require enormous resources. At the same time, there is a trend towards developing increasingly powerful computer clusters. Geneva is currently one of the best available frameworks for distributed optimization of large-scale problems with highly nonlinear quality surfaces. It is a great tool to be used in wide-area networks such as Grids and Clouds. However, it is not user-friendly for scheduling on high-performance computing clusters and supercomputers. Another issue is that it only provides a framework for parallelizing workloads on the population level of optimization algorithms, but does not support distributed parallelization of the cost function itself. For this reason, a new software component for network communication – called MPI-Consumer – has been developed.

In Geneva's system architecture, the server node runs the optimization algorithms and distributes candidate solutions to clients. The clients evaluate the candidate solutions based on a user-defined cost function and then send the result back to the server.

When scaling to high-dimensional problems and hundreds or even thousands of nodes, the server's performance is a fundamental challenge, because the speed of answering client requests has a direct impact on the clients' CPU efficiency. We tackle this challenge by making immense use of multithreading on the server. Furthermore, we use asynchronous client requests to hide server response times behind computing times.

Additionally, as the number of compute nodes and the runtime of cluster jobs increase, fault tolerance becomes increasingly important due to the growing probability of errors. Typical MPI programs, however, do not offer fault tolerance, such that client failures or connection issues might result in a crash of the entire system. To address this issue, we have used MPI in a client-server model using asynchronous operations and timeouts to improve fault tolerance.

In some use cases, such as certain hadron physics applications, the cost function itself requires another level of distributed parallelization because its computation requires enormous amounts of CPU time or memory, which go beyond the resources available on single cluster nodes. Using the MPI-Consumer with Geneva, this is no longer a complex or tedious task. The MPI-Consumer provides access to pre-configured subgroups of client nodes, which can be used by domain experts to intuitively parallelize their cost function.

Extensive quantitative evaluation with up to 1000 nodes shows that the MPI-Consumer scales perfectly on HPC clusters and vastly improves Geneva's user experience for high-performance computing. The MPI-Consumer even outperforms some WAN consumers developed earlier for Geneva and, therefore, can be used as a model for the improvement of Geneva as a whole.

The MPI-Consumer has been integrated into the Geneva optimization library and is now available to users [1]. Also, independent of Geneva's parametric optimization functionality, the MPI-Consumer can be used as part of a generic networking library as a scalable implementation of a fault-tolerant client-server model for high-performance computing clusters. Geneva is currently used by scientists at GSI in Darmstadt for fundamental research in hadron physics on the Virgo HPC cluster.

(1) Berlich, R; Gabriel, S; Garcia, Geneva Source Code <https://github.com/gemfony/geneva>.

**Primary author:** WESSNER, Jonas (GSI Helmholtz Center for Heavy Ion Research)

**Co-authors:** Prof. LUTZ, Matthias F.M. (GSI Helmholtz Center for Heavy Ion Research); Dr SCHWARZ, Kilian (Hochschule Darmstadt University of Applied Sciences); Dr BERLICH, Rüdiger (Gemfony scientific)

**Presenter:** WESSNER, Jonas (GSI Helmholtz Center for Heavy Ion Research)

**Session Classification:** Converging Infrastructure Clouds, Virtualisation & HPC

**Track Classification:** Track 9: Converging High Performance Computing Infrastructures: Supercomputers, clouds, accelerators

Contribution ID: 9

Type: **Oral Presentation**

## Developing Software for Medical Devices (Remote presentation)

*Tuesday, March 21, 2023 2:00 PM (30 minutes)*

Medical devices will almost always be driven by software components. Development for this field of work requires special considerations for patient safety and data privacy and is thus governed by rules alien to other deployment scenarios. The European Union is about to switch to a new regulation framework, the Medical Device Regulations (MDR), replacing the far less comprehensive Medical Device Directives (MDD). The new rules will have a significant impact on software development in the future and, while providing more patient safety, come at the price of significantly increased complexity and thus cost. Other regions in the world either follow or, to the contrary, have relaxed some regulations. The presentation discusses current regulations and difficulties in different areas of the world from the perspective of software development, with the focus and starting point of the new MDR.

**Primary author:** Dr BERLICH, Ruediger (Gemfony scientific)

**Presenter:** Dr BERLICH, Ruediger (Gemfony scientific)

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)

**Track Classification:** Track 2: Health & Life Sciences (including Pandemic Preparedness Applications)

Contribution ID: 10

Type: **Oral Presentation**

## **Case Study on the e-infrastructure built leveraging cloud computing resources to prepare and provide rapid response to COVID-19 at a top-tier Higher Education Institution (HEI) in the USA (Remote presentation)**

*Tuesday, March 21, 2023 4:00 PM (20 minutes)*

From a Public Health perspective, the once-in-a-century Sars-CoV2 (COVID-19) pandemic created unprecedented challenges for Higher Education Institutions (HEIs). HEIs in the USA had to respond rapidly to switch from a mostly in-person mode of instruction to a fully remote mode. These changes had to be immediate and imminent to provide education with less to no impact on the health and safety of the campus community. At one of the top-tier research institutions in the USA, also called a “Public Ivy” school, with about 40,000 enrolled students in an academic year, the University of Maryland’s response to the COVID-19 pandemic highlights the timely establishment of an e-infrastructure that leveraged cloud computing resources with efficient communication strategies to provide real-time updates to the community stakeholders. The strategies were implemented swiftly by adhering to the Public Health guidelines put forth by the Centers for Disease Control (USA), the State of Maryland, the County officials, and the University System of Maryland. Immediately upon understanding the magnitude of the crisis, the University’s pandemic response began with an implementation of a cloud-based medical symptom monitoring web survey which eventually transformed seamlessly into a multi-faceted enterprise-wide web application suite catering to all facets of the University communities such as students, faculty, staff, university health center employees, integrating with public health officials, state and federal COVID-19 testing and vaccination reporting centers. As mandated by several government agencies and the University’s health center, the e-infrastructure project requirements were to be highly available, exhibit elastic scalability, be fault-tolerant, and be accessed worldwide. These were required by the institution to report daily compliance status for thousands of individuals accessing the web application round-the-clock to ascertain their return to campus status based on several factors such as mandatory reporting of testing results (twice a week), the exhibition of symptoms related to COVID-19, submission of vaccination information, and to report close contacts information for contact tracing. The cloud-based web application suite evolved with the development of the pandemic stages by integrating with external vendors and University’s health center database—to import testing results and vaccination records from the State of Maryland’s database hub and the University Health Center’s testing and vaccination data respectively. The application also harnessed cloud-based web reporting tools to provide the general public with an overview of the situation and the internal leadership community with several detailed reports to enable appropriate and timely decision-making. This high-performance e-infrastructure enabled the institution to efficiently identify campus community individuals who are compliant and non-compliant with the COVID-19 guidelines. Some parts of this application suite were also leveraged to be interactive enough for specific campus communities such as supervisors to communicate directly via emails to identify non-compliant individuals. This study performs a unique detailed case study analysis on the strategies, processes, and technologies used to establish an end-to-end e-infrastructure that plays a vital role in managing, maintaining, and controlling the spread of the deadly pandemic within the top-tier HEI campus community. The results of these diligent efforts could be seen with fewer COVID-19 positive test results, fatalities, and eventually becoming one

of the highest vaccinated communities within the state.

**Primary author:** GNANASEKARAN, Rajesh Kumar (the University of Maryland)

**Co-author:** Dr MARCIANO, Richard

**Presenter:** GNANASEKARAN, Rajesh Kumar (the University of Maryland)

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)

**Track Classification:** Track 2: Health & Life Sciences (including Pandemic Preparedness Applications)



Contribution ID: 11

Type: **Oral Presentation**

## IPv4 to IPv6 Worker Node migration in WLCG

*Friday, March 24, 2023 9:00 AM (30 minutes)*

The Worldwide Large Hadron Collider Computing Grid (WLCG) actively pursues the migration from the protocol IPv4 to IPv6. For this purpose, the HEPiX-IPv6 working group was founded during the fall HEPiX Conference in 2010. One of the first goals was to categorize the applications running in the WLCG into different groups: the first group was easy to define, because it comprised of all applications that were not IPv6 ready and would never be. The second was also easy since it covered those applications that were already working successfully with IPv6. Group number 3 consisted of applications that worked under IPv6, but not as smoothly as desired, meaning improvements were required. In 2016 the WLCG management board decided that the storage space of all LHC collaborating Tier-1 centers had to be IPv6 ready until the beginning of 2018 whereas the Tier-2 centers were given time until the end of 2018. However, this was a tight schedule and could therefore not be achieved according to plan. Yet today the storage space of all Tier-1 centers are IPv6 ready and 95% of the Tier-2 centers as well.

After the IPv6 readiness of the storage has been achieved, there are still other services needed to migrate to IPv6. This are for example middleware services like job scheduler, Advanced Resource Connector (arc-ce), HT-Condor and others.

The HEPiX-IPv6 working group pursues and concentrated the next goal on the worker nodes: setting up IPv6 only worker node testbeds. GridKa, however, tackles a bit different route here: we have started migrating IPv4 worker nodes to IPv6 worker nodes. For that purpose, we have set up a highly detailed monitoring system in order to record all inbound and outbound packages. We are using Packetbeat to absorb the packet's header information and transfer it into a Opensearch data base. We analyze the recorded data in the Opensearch Dashboard to identify the packages and that are still running under IPv4. With that information we can identify the applications that are not yet IPv6 ready and actively investigating in the IPv6 migration. Since this applies to all kind of different applications we find different kind of situations or 'pictures' and dealing with them is quite a challenge.

This presentation will show some of the steps that are necessary to implement IPv6 to different applications. The web site (url: : <https://hepixon-ipv6.web.cern.ch>) of the HEPiX-IPv6 working group contains lots of the material brought together by the working group. At the subsection "worker-nodes-migration-ipv6" the major findings for the migration and their solution are maintained.

**Primary author:** HOEFT, Bruno (Karlsruhe Institute of Technology)

**Co-authors:** Dr PETZOLD, Andreas (Karlsruhe Institute of Technology); Dr SCHNEPF, Matthias (Karlsruhe Institute of Technology); Dr FISCHER, Max (Karlsruhe Institute of Technology)

**Presenter:** HOEFT, Bruno (Karlsruhe Institute of Technology)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 13

Type: **Oral Presentation**

## Reducing the Carbon Footprint of Computing

*Friday, March 24, 2023 10:50 AM (30 minutes)*

Over the past year, the soaring cost of electricity in many parts of the world has brought the power-requirements of computing infrastructure sharply into focus, building on the existing environmental concerns around the issue of global warming. We report here on the investigations, and subsequent actions, in the UK to respond to this pressure. The issues we address are both the overall reduction in power-used, but also the quality of the power used, where quality refers to whether the power has larger or small fractions of renewable energy. To reduce power consumption, we have investigated the performance of ARM-based systems compared to AMD processors, using the new HEPsScore test workloads as they have become available. We will show results of our measurements at the University of Glasgow on identically priced ARM and AMD systems, and also compare these with a standard AMD worker-node from the Glasgow Tier-2 sites. We will demonstrate that major power-savings may be possible, and that the work is done in comparable, or faster, times on ARM. The LHC experiments are at different stages of readiness to use ARM architectures and we summarize which workloads currently compile, and which have been physics-validated. Whilst a reduction in the carbon footprint of computing can be achieved with more efficient hardware, it can also be addressed by load-shaping: Reducing the amount of power-used when the fraction of renewable energy in the national power network is low and the fossil-fuel generated power is high. We present an analysis of the UK power-generation data and suggest methods of presenting the reduction in carbon footprint that can be achieved by reducing computing-load at times of peak fossil-fuel use. The ability to load-shape would also allow facilities to respond to calls for reductions in power-use at peak times; a possibility that may be essential in Europe over the winter of 2022-2023. Achieving load-shaping in practice is not easy because LHC workloads run for much longer times than the typical load-shaping timeframe, and hardware does not like being repeatedly power-cycled. We discuss various methods of reducing computing at peak times, and report on various tests and measurements made in the UK designed to investigate the feasibility, such as frequency throttling. Finally, we will summarise the current and future strategy in the UK, how this ties in with the NetZero policies of funding agencies, and the compromises that this will entail.

**Primary authors:** Mr BORBELY, Albert (University of Glasgow); Prof. BRITTON, David (GridPP / University of Glasgow / STFC); Dr SPITERI, Dwayne (University of Glasgow); Dr SIMILI, Emanuele (University of Glasgow); Dr STEWART, Gordon (University of Glasgow); Dr SKIPSEY, Sam (University of Glasgow)

**Presenter:** Prof. BRITTON, David (GridPP / University of Glasgow / STFC)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 14

Type: **Oral Presentation**

## **Transformer-Based Detection Method for DNS Covert Channel (Remote presentation)**

*Friday, March 24, 2023 10:00 AM (30 minutes)*

As network technique continues to flourish, current network attacks against large-scale scientific facilities and science data centers show a more sophisticated trend. In order to evade traditional security detection systems, attackers adopt more stealthy attack methods. The Domain Name System (DNS) protocol is one of the basic protocols used in the network environment of large-scale scientific facilities and science data centers, which usually uses unencrypted data transmission to identify computers reachable through the Internet and is rarely blocked by firewalls under normal conditions. In computer security, a covert channel is a type of attack that creates a capability to transfer information objects between processes that are not supposed to be allowed to communicate by the computer security policy. Attackers exploit the vulnerabilities of DNS protocol to establish covert channels for evading traditional security detection and further launch network attacks by encapsulating hidden information in DNS covert channels, such as remote control and information theft, which seriously affect the network and information security. Therefore, the detection and defense of DNS covert channel are crucial to secure the network of large-scale scientific facilities and science data centers.

At present, many detection methods using machine learning are based on manual features, which usually include complex data preprocessing and feature extraction. Additionally, these methods seriously rely on expert knowledge, and some potential features are hard to discover. Deep learning-based detection methods for DNS covert channel have received increasing attention recently. Deep neural networks can better extract the hidden information, timing relationships, and other deep features of DNS network traffic. Compared with most traditional methods, deep learning-based methods can achieve automatic extraction of data features without manual intervention and implement an end-to-end traffic identification model. However, most deep learning-related detection methods require a large amount of labeled accurate positive and negative sample data. Obtaining huge amounts of labeled accurate DNS network traffic data consumes a lot of labor costs, making these methods difficult to be applied in practical environments. In addition, existing deep learning-based covert channel detection methods still suffer from low recognition rates and long training periods.

In order to solve the above problems, this paper proposes a Transformer-based detection method for DNS covert channel. A Transformer is a deep learning model that adopts the mechanism of self-attention, differentially weighing the significance of each part of the input data. Unlike RNNs, Transformers do not have a recurrent structure, dispensing with convolutions entirely, and the training parallelization allows training on larger datasets. The model is applied in the feature extraction of global dependencies on inputs, fully considering the correlations between the input data and providing parallelized operations, which significantly improves training speed and detection accuracy. Meanwhile, through the Transformer structure's ability to capture long-term dependencies, it can improve the model's ability for long-term prediction, thus improving the accuracy of predicting long-term sequences.

Our method experiments on the DNS network traffic dataset. The results show that the proposed Transformer-based detection method can effectively identify DNS covert channels. This method is also tested in a real network environment and has achieved desired results.

**Primary authors:** SUN, Qian Ran (Chinese Academy Of Sciences); ☒, ☒☒; ☒, ☒☒; ☒, ☒; ☒, ☒☒

**Presenter:** SUN, Qian Ran (Chinese Academy Of Sciences)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 15

Type: **Oral Presentation**

## Progress and Prospects for Online Emotional Feedback Using EEG

*Wednesday, March 22, 2023 4:00 PM (20 minutes)*

In this presentation, we will divide a general EEG study into several steps and introduce the position of this study in EEG research through the issues involved in each step.

After that, we will outline the feature selection method we used in this study that emphasizes the minority class in imbalanced datasets and states the results on imbalanced datasets for the EEG and Emotion dataset and other multi-class classifications.

Finally, we mention our progress and prospects in the other steps of our current study regarding challenges and solution methods.

**Primary authors:** Mr TOKIDA, Yuki (Kanazawa Institute of Technology); NAKAZAWA, Minoru (Kanazawa Institute of Technology)

**Presenter:** Mr TOKIDA, Yuki (Kanazawa Institute of Technology)

**Session Classification:** Artificial Intelligence (AI)

**Track Classification:** Track 10: Artificial Intelligence (AI)

Contribution ID: 16

Type: **Oral Presentation**

## Extension of local dCache instance capacity using national e-infrastructure

*Thursday, March 23, 2023 11:00 AM (30 minutes)*

The Czech WLCG Tier-2 center for LHC experiments ATLAS and ALICE provides computing and storage services also for several other Virtual Organizations from high energy and astroparticle physics. The center deployed Disk Pool Manager (DPM) for almost all (only ALICE VO uses xrootd servers) supported VOs as a solution for storage until recently. The local capacity was extended by a separate instance of dCache server which was operated by CESNET Data Storage unit in a remote location. The exact location has changed during the project, the distance was between 100 to 300 km. This storage extension was based on HSM and was mapped as a separate ATLAS space token where higher latencies were expected. The intended usage was for a non-automatic backup of the LOCALGROUP disk used by ATLAS users from the Czech Republic. Since the usage was relatively low and the system had only one group of users from the ATLAS VO, the effort required for maintenance and frequent updates was not effective.

The DPM project announced the end of support, and we migrated the main Storage Element in CZ Tier-2 to dCache. This brought a possibility of unified solution for an SE. The dCache system at CESNET was stopped and we started to test a new solution with only one endpoint for all users. CESNET Data Unit also changed the underlying technology for data storage - they moved from HSM to CEPH. We mounted one file system as RADOS block device (RBD) on test dCache server and measured properties of the system to compare with storage based on local disk servers. This solution differs from a solution used in the Nordugrid Tier-1 center, where distributed dCache servers use caching on local ARC Computing Elements. Tests included long term stability of network throughput, duration of transfers of files with sizes from 10 MB to 100 GB and variation of transfer time for cases when several simultaneous transfers are executed. The network tests were first executed on an older diskless server and later on a new dedicated test server with surprisingly different results. We used the same tools also to measure differences in transfer performance between local disk servers which are of different age and connected by different speed. Since the results of tests were satisfactory, we will use the external storage first as a dedicated space token for ATLAS and later as a part of a space token located also on local disk servers. We may also use the solution for other Virtual Organizations if the external available space is increased by a sufficient volume.

**Primary author:** CHUDOBA, Jiri (Institute of Physics of the CAS, Prague)

**Co-authors:** MIKULA, Alexandr (FZU); SVATOŠ, Michal (FZU); VOKÁČ, Petr (FZU); CHUDOBA, Michal (Faculty of Mathematics and Physics, Charles University)

**Presenter:** CHUDOBA, Jiri (Institute of Physics of the CAS, Prague)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data

Contribution ID: 17

Type: **Oral Presentation**

# Performance Characterization of Containerized HPC Workloads

*Wednesday, March 22, 2023 2:40 PM (20 minutes)*

Approaching the exascale era, complex challenges arise within the existing high performance computing (HPC) frameworks. Highly optimized, heterogenous hardware systems on the one side, and HPC-unexperienced scientists with a continuously increasing demand for compute and data capacity on the other side. Bringing both together would enable a broad range of scientific domains to enhance the models, simulations and findings while efficiently using the existing and future compute capabilities. Those systems will continue to develop a more and more heterogeneous landscape of compute clusters, varying classical computation and accelerator cores, interconnects or memory and storage protocols and types. Adapting user applications to those changing characteristics is laborious and prevents enhancing the core functions of the applications while focusing on deployment and runtime issues. Consequently, containerization is one key concept to shift the focus back to the actual domain science, removing incompatible dependencies, unsupported sub-programs or compilation challenges. Additionally, an optimized efficiency of the compute systems' usage is reachable, if system owners would be aware of the actual requirements of the containerized applications.

Our proposed work provides a methodology to determine, analyze and evaluate characteristic parameters of containerized HPC applications to fingerprint the overall performance of arbitrary containerized applications. The methodology comprises the performance parameter definition and selection, a comparison of suitable measurement methods to minimize overhead, and a fingerprinting algorithm to enable characteristics comparison and mapping between application and target system. By applying the methodology to benchmark and real-world applications we aim to demonstrate its capability to reproduce expected performance behavior and build prediction models of the application's resource usage within a certain trash-hold. We enable a twofold enhancement of today's HPC workflows, an increase of the system's usage efficiency and a runtime optimization of the application's container. The system's usage efficiency is enabled by container selection and placement optimizations based on the container fingerprint, while the runtime profits from a streamlined, target-cluster-oriented allocation and deployment to optimize the time-to-solution.

The adaption module and the most prospective technology to overcome endless adaption of the application's program code is containerization, which offers portability among heterogeneous clusters and unprecedented adaptability to target cluster specifications. Containers like Singularity, Podman or Docker are well known for cloud usage and micro-service environments. During the last years containers like Apptainer or Charliecloud became also widespread in certain HPC domains, since their capabilities to include high data throughput, intra- and inter-node communication, as well as the overall scalability increased enormously.

We base our approach on the EASEY (Enable exASclae for EverYone) framework, which can automatically deploy optimized container computations with negligible overhead. Today's containers are natively not able to automatically use all given hardware at best, since the encapsulated application varies on computing, memory or communication demands. An added abstraction layer, although enabling many programming models and languages to be executed on very different hardware, is not able to make use of all provided hardware features. An enhanced EASEY framework will support distinct optimization tunings without any human interaction during compilation of the container. Based on the introduced methodology we will demonstrate how these optimizations could impact the performance of the containerized HPC application.

**Primary author:** HÖB, Maximilian (Leibniz-Supercomputing Centre (LRZ))

**Presenter:** HÖB, Maximilian (Leibniz-Supercomputing Centre (LRZ))

**Session Classification:** Converging Infrastructure Clouds, Virtualisation & HPC

**Track Classification:** Track 9: Converging High Performance Computing Infrastructures: Supercomputers, clouds, accelerators



Contribution ID: 18

Type: **Oral Presentation**

## **A distributed framework for security operation center in the application of Institute of High Energy Physics (Remote presentation)**

*Tuesday, March 21, 2023 4:30 PM (30 minutes)*

Security operations center (SOC) frameworks standardize how SOC's approach their defense strategies. It helps manage and minimize cybersecurity risks and continuously improve operations. However, current most of SOC frameworks are designed as the centralized mode which serves for the single organization. These frameworks are hard to satisfy the security operations scenarios that must simultaneously protect several organizations from cyber threats across the wide area network in the synergistic way. In this paper, we propose the distributed security operation center (DSOC) that provides the distributed working mechanism for multiple organizations over the wide area network by combining the security probes. The organizations within the DSOC framework are highly collaborative and mutual trust. The security probes of DSOC are deployed in the different organizations and can parse the network traffic for the organizations by Zeek. Besides, the security probes collect data from these organizations and the collected data is transferred over wide area network to the data analysis center of the DSOC. Especially, the data communication between security probes and data analysis center is encrypted to ensure the data security of every organization. The data analysis center adopts rule-based, AI-based and threat intelligence-based algorithms to detect cyber-attacks. The detection results are input into the automated response module in the DSOC. The automated response module is the client-server structure and the client are installed in the security probe. The server of the automated response sends commands across the wide area network to the target client of the security probe to block the attackers quickly, and meanwhile the communication between client and server in the response processes is encrypted. In addition, the threat intelligence component of DSOC can aggregation intelligence from the organizations and easily share to all organizations based on the distributed security probes. The DSOC also builds the security situation awareness system that visuals the cyber threats of every organization and set the permission to view the security situation by using access control for every organization. The DSOC has been applied to institute of high energy physics (IHEP) and deployed in several collaborative large scientific facilities and scientific data centers since 2021. The excellent security protections are persistently provided to all organizations within the DSOC framework.

**Primary authors:** WANG, Jiarong (Institute of High Energy Physics); LIU, Junyi; SUN, Qianran; YAN, Tian; AN, Dehai

**Presenter:** WANG, Jiarong (Institute of High Energy Physics)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 19

Type: **Poster Presentation**

## A Study of Heavy-Ion Beam Simulation using Geant4 on KISTI-5 Supercomputer

We introduce the result of heavy ion beam simulation using Geant4 version 11.0.2. There has been relatively little focus on accelerator-based study of secondary particles in the nuclear physics, biophysics, and dark matter research. In this study, we have compared our simulation result with experimental data on collisions between the liquid hydrogen target and uranium beams which take the energy of 1 GeV per atomic mass. Since the Geant4 is more accurate than any other simulation packages for particle-matter interaction, we have used it for this study. We examine the validity of the Geant4 model with the comparison between simulation and experiment result. Next, we examined the various physics models built in Geant4 to determine the optimized model. We checked how each physics model describes the expected physical phenomena. We checked the cost of CPU time to find what the best cost-effective physics model is. As a result, we have found the optimized physics model for our study. Finally, we simulated the interaction between particle and matter to study the secondary ions created by beam-target collision. The uranium plate target with thickness 1~10mm is used for simulations. The proton beam with energy range 0.1~1 GeV per atomic mass is used for simulations. We execute Geant4 simulations to determine the characteristics of secondary heavy-ion beams.

We have used the KISTI-5 supercomputer [Nurion Knight Landing (KNL)] to run the simulation. Nurion KNL is equipped with 8,305 nodes and 68 cores per node. To find the cost-effective physics model on uranium beam to liquid hydrogen target, the CPU time of calculation is measured by using 64 nodes. Same condition is used to study the secondary beam from proton beam to uranium target. These results will help the experiments for secondary particle at RAON, which is an upcoming facility for heavy-ion collision experiments in Korea.

This research was supported by the National Research Foundation of Korea (NRF) grant fund by the Korean government (MSIT) (No. 2021R1F1A1064008). This research is also supported by the National Supercomputing Center with supercomputing resources including technical support (KSC-2022-CHA-0003).

**Primary authors:** Dr KIM, Kyungho (KISTI); Prof. CHO, Kihyeon (KISTI/UST)

**Presenters:** Dr KIM, Kyungho (KISTI); Prof. CHO, Kihyeon (KISTI/UST)

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 20

Type: **Poster Presentation**

## A study of double dark photon visible decay mode at $e^+e^-$ colliders using HPC

Dark sector, which includes dark matter and dark photon, is one of the most important and challenging research subject in particle physics. Dark matter is known to play a crucial role in the formation of cosmic structure including stars, galaxies, and clusters of galaxies. However, we do not know the identity of dark matter to date. In order to find identity of dark matter, various searches including direct search, indirect search, and particle collider search are actively on-going. However, dark matter study heavily depends on computing simulation due to small cross-section and unknown parameters. The unknown parameters are dark photon mass and coupling constant. We studied double dark photon visible decay mode at electron-positron colliders at center-of-mass energy  $\sqrt{s} = 10.58\sim 1000$  GeV using KISTI-5 supercomputer.

According to a theory, dark photon can decay into muon pair. From this theory, we studied dark photon ( $A'$ ) in the decay mode of  $e^+e^- \rightarrow A' A'$  and  $A' A' \gamma$  where each dark photon decays into muon pair. We used a simplified model for the generation of the dark matter. At first, we scanned  $\sqrt{s}$  and dark photon mass in order to know dependency of cross-section with them. Because lots of event generation are required for this scan, we conducted our study efficiently using KISTI-5 supercomputer. Based on the scan, we found the dark photon mass for the maximum cross-section for each collider. For given dark photon masses, we generated one million events of signal modes with MadGraph5 at  $\sqrt{s}$  of each collider using KISTI-5 supercomputer. To study optimization of event generation using KISTI-5 supercomputer, we compared CPU time and memory which are used at  $\sqrt{s}$  of each collider. The result will help to optimize the CPU time and memory for event generation using high performance computing.

Then, we performed detector simulation using Delphes applying those dark photon masses. We reconstructed dark photon from muon pairs. Finally, we found the expected number of event and detector efficiency for each collider. The results of this study will help to search for double dark photon visible decay mode at electron-positron colliders.

**Primary authors:** Mr PARK, Kihong (Korea Institute of Science and Technology Information (KISTI)); Dr KIM, Kyungho (KISTI); Prof. CHO, Kihyeon (KISTI/UST)

**Presenters:** Mr PARK, Kihong (Korea Institute of Science and Technology Information (KISTI)); Dr KIM, Kyungho (KISTI); Prof. CHO, Kihyeon (KISTI/UST)

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 22

Type: **Oral Presentation**

## On Definitions in the Humanities and Social Sciences: What is a good definition of Culture?

*Wednesday, March 22, 2023 2:00 PM (30 minutes)*

In the Humanities and Social Sciences, Big Data, and technologies of the Digital Humanities have helped to substantiate academic work. For the portability of data however, attributes and values require definitions to preserve their intended meaning through time and space. The standard practice of researchers to rely on ontologies, even for the study of cultures and societies, however, is colonial: It assumes a single truth, relying on individually seen variables. Many notions, from initiation rite to burial, from food to tool, or from disease to health are conceptualized differently, depending on the community, village, region, language, religion, or discipline of the researcher. This is because knowledge of natural, psychological and social phenomena is complex and includes interlocking cycles of multiple factors. In such complex views, classifying a snake as prey or predator, wind as harmful or beneficial, milk as healthy or unhealthy depends on the specific complex of conceptualized interactions. And it is this variety of conceptualizations that is one of the nuclei of studies in the Humanities and Social Sciences.

In this light, we investigate the potential and actual role of definitions for research in the SSAH in a) binding together a set of attributes and values in their mutual dependence as specified in the definition, and b) as a way to summarize or label different or even contradicting historical and contemporaneous conceptions. We propose to include these conceptions as essential component of data-sets, as definitional space of the attributes and values and for the empirical falsification.

Our endeavor starts by classifying and analyzing types of definitions as used in relation to research in the SSAH, using examples from the century-old struggle to define “culture” or Culture. In other words, we set out to create a dataset which for one or several definitions of culture collects and describes tabular data. These data have been collected directly from cultural sites, such as housings, temples, graveyards etc. all over Asia for 15 years, documenting through photos cultural practices and the way they change through migration, time, and space. This data set also recollects artifacts documented in old good paper books published 40 years ago. With such a span of data addressing cultural anthropology, the question of how to design a good definition of Culture becomes unavoidable.

Carl Popper, among others, distinguishes *nominal definitions* of a word, e.g. “culture”, from *essential definitions* of a notion, e.g. Culture. The first type of definition answers the question “when to use a word”, the second the question “what is it what we talk about”.

In Social Sciences, the essence of a notion, e.g. of Culture, is not fixed by a genome or a formula. It remains the construct of a theory. The use of essential definitions thus has been criticized as being unnecessary in addition to the obligatory formulation of a theory: In most academic publications, essential definitions don’t match the theory and are therefore not evaluated in the same way as the theory. In this case we speak of *weak essential definitions*. In this paper, we want to counter this criticism of essential definitions. We argue that essential definitions that match a theory, a type we will call a *strong essential definition* or *heuristic definition*, if carefully crafted, projects smoothly into the design of attributes, values and their interrelations for operationalization, data collection and data analysis. Heuristic definitions help to focus and conceptualize, to plan the empiric evaluation and eventually can be falsified and discarded. We claim that they should be the beginning of each academic enterprise and the roots for the compilation of portable data-sets.

**Primary authors:** STREITER, Oliver (National University of Kaohsiung); Mrs ZHAN, Yaqing (University of Hamburg); Dr GOUDIN, Yoann (Research Center for Asia-Pacific Languages and Cultures at the National University of Kaohsiung)

**Presenter:** STREITER, Oliver (National University of Kaohsiung)

**Session Classification:** Humanities, Arts & Social Sciences

Contribution ID: 23

Type: **Oral Presentation**

## High performance Geant4 simulations of electromagnetic processes in oriented crystals

*Friday, March 24, 2023 9:30 AM (30 minutes)*

Electromagnetic processes of charged particles interaction with oriented crystals provide a wide variety of innovative applications in high-energy frontier physics, accelerator physics, detector physics, nuclear physics and radiation therapy. A small piece of crystal material could be used as -an intense source of X- and gamma-ray radiation for nuclear physics and cancer treatment, -a positron source for future collider projects, i.e. both linear and circular e+e- colliders (ILC, FCC-ee) as well as for muon colliders, -a beam manipulation instrument for particle detector R&D at tens of existing electron synchrotrons as well as for ultra-high energy fixed-target experiments at existing and future collider projects (LHC, FCC-ee) to measure, CP-violation processes and physics beyond Standard Model -a compact crystalline electromagnetic calorimeter and -a compact plasma wakefield accelerator as well.

The design of every of these applications require simulations. Simulations of these processes imply a very detailed charged particle trajectory calculation and a lot of computational power as well. Geant4 simulation toolkit [1] is perfect for the development of these applications, since it allows one to carry out detailed simulations of a complete experimental setup. It is a Monte Carlo code written on C++, simply parallelizable, including intrinsic multithreading parallelization and MPI parallelization as well. It also includes a rich collection of physical models of particles interaction with matter, wide capabilities to implement a complicated geometry of materials and a number of scoring methods as well.

We present a new simulation model of electromagnetic processes in oriented crystals “ChannelingFastSimModel” implemented into Geant4 using so called Fast Simulation Interface. It allows one to create a model by means of the inheritance G4VFastSimulationModel Geant4 class, and in particular the following functions:

- IsApplicable, a condition if a particle corresponds to the list of particles applicable for this model (electrons and positrons for ChannelingFastSimModel)
- ModelTrigger, a specific condition to launch the model (the particle energy and the angle to be within ChannelingFastSimModel limits)
- DoIt, the main function of the model executing if the IsApplicable and ModelTrigger conditions are fulfilled.

The key advantage of Fast Simulation Interface is automatic switching off of all the standard Geant4 processes on the step when ChannelingFastSimModel is executing. This allows one to avoid its conflict with Geant4 physical processes as well as makes the ChannelingFastSimModel independent of the standard Geant4 physics lists. This model is simply addable to the Geant4 examples already existing, which makes it very simple to use for different applications. In addition, it supports standard Geant4 parallelization methods.

We perform simulations with ChannelingFastSimModel on supercomputers of KISTI and CINECA supercomputing centers. We validate the model with the experimental data. We compare the simulation performance on different architectures and optimize the simulation code.

[1] J.Allison et al., NIM A 835, 186-225 (2016).

A. Sytov is supported by the European Commission (TRILLION, GA. 101032975). We acknowledge partial support of the INFN through the MC-INFN project. We acknowledge the CINECA award under the ISCRA initiative, for the availability of high performance computing resources and support. This work is also supported by the National Supercomputing Center with supercomputing

resources including technical support (KSC-2022-CHA-0003).

**Primary authors:** Dr SYTOV, Alexei (INFN; KISTI); Prof. CHO, Kihyeon (KISTI/UST); Dr BANDIERA, Laura (INFN)

**Presenters:** Dr SYTOV, Alexei (INFN; KISTI); Prof. CHO, Kihyeon (KISTI/UST)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 24

Type: **Oral Presentation**

## **Influencers Matchmaking for Startups B2B-Branding interactivity: Case Studies of LinkedIn Data Mining**

*Wednesday, March 22, 2023 2:30 PM (30 minutes)*

Influencers Matchmaking for Startups B2B-Branding interactivity: Case Studies of LinkedIn Data Mining

Marketing in the fourth industrial revolution offers a global opportunity to the rise of micro-influencers. Influencer marketing allows any social media user with particular influence within the social media platform to help amplify and circulate the campaign of the brand to the influencers' audience. Especially with more business decision makers on social media, social media provides an advertisement opportunity for B2B brands. LinkedIn as a social platform that is primarily used for professional networking offers a suitable environment to promote B2B brands mainly by its functions that are built around helping firms create brands, build relationships and connect with existing and potential customers. However, not only that influencer-marketing for B2B brands is not widely practiced on LinkedIn, it can be challenging to perform with the tremendous amount of data available and lack of effective approach to run a successful micro-influencer campaign.

This research proposed a data analysis approach in clustering and classifying the available data provided on LinkedIn to provide marketers actionable insights on connecting with the right micro-influencers to conduct influencer marketing campaigns. This analysis will leverage the data provided by LinkedIn by using a web scraping method. The data acquired from the user's activity will then be classified and analyzed to determine potential matches between micro-influencers and brands. The purpose of this study aims to break through the new approach of influencer marketing, particularly for B2B brands on LinkedIn.

To test the hypothesis of the proposed research question, Job to Be Done (JTBD) analysis will be implemented to help understand stakeholders behavior of improving this research. The research identifies and gathers the target users based on the persona, which then used to design the most suitable content marketing and respective incentives to match the influencer marketing concept. To enable the analysis, data from LinkedIn API will be retrieved and analyzed to provide actionable insights for the user to connect with potential brands and influencers. In the end, brand interactivity will be measured through the LinkedIn post performance analytics. The target participants will be B2B brands and all active professional users with various degrees of social media presence on LinkedIn.

The proposal discusses the usage of web scraping and data mining to create a pointer that allows influencer marketing approaches to become more targeted and data oriented on LinkedIn. We conclude with design consideration on creating an influencer management platform for B2B brands on LinkedIn that offers effective advertisement opportunities to help brands grow and increase the brand credibility. The proposed analysis will help brands to make decisions through real-time visual reports of the campaign and available pool of micro-influencers.

Keywords: Data mining, influencer marketing, LinkedIn, B2B brands, brand interactivity.

**Primary authors:** PRAHASTIWI, Agnes Mutiarawati (National Taipei University of Technology); Prof. SHEN-MING, Wang (National Taipei University of Technology)

**Presenters:** PRAHASTIWI, Agnes Mutiarawati (National Taipei University of Technology); Prof. SHEN-MING, Wang (National Taipei University of Technology)

**Session Classification:** Humanities, Arts & Social Sciences



**Track Classification:** Track 4: Social Sciences, Arts and Humanities (SSAH) Applications

Contribution ID: 26

Type: **Oral Presentation**

## **Leveraging TOSCA orchestration to enable fully automated cloud-based research environments on federated heterogeneous e-infrastructures (Remote Presentation)**

*Wednesday, March 22, 2023 4:30 PM (30 minutes)*

In the last years cloud computing has opened up interesting opportunities in many fields of scientific research. Cloud technologies allow to scale applications and adapt quickly, ease the adoption of new software development methods (e.g. DevOps), accelerating time to value.

However, the lack of integration of the existing infrastructures and the consequent fragmentation of the resources are still a barrier to a broader adoption of these technologies.

Starting from the times of the INDIGO-Datacloud project (2015-2017) we have been developing a set of solutions for implementing a seamless and transparent access to geographically distributed compute and storage resources, mainly the INDIGO IAM, a modern authentication and authorization system, and the INDIGO PaaS, a suite of microservices that allow to federate multiple providers and orchestrate cloud deployments via TOSCA.

At the beginning of 2021, INFN inaugurated a national multi-site cloud infrastructure (INFN Cloud), that is currently exploiting and extending the INDIGO solutions to provide an extensible portfolio of services tailored to multi-disciplinary scientific communities, spanning from traditional IaaS to more elaborate PaaS and SaaS solutions. Some examples are: data analytics and visualisation environments based on Elasticsearch and Kibana, file sync & share solution based on OwnCloud with replicated backend storage, web-based multi-user interactive development environment for notebooks, code and data built on JupyterLab, kubernetes clusters, HTCondor on-demand clusters, Spark clusters integrated with Jupyter, Cloud storage solutions, etc. Moreover, the INFN Cloud service catalogue includes integration with the Kubernetes ecosystem and customizations for specific use-cases, e.g. the exploitation of GPUs for machine learning projects or pre-installed experiment software for data analysis.

The topology of each service is described through a TOSCA template, whereas the provisioning of the cloud resources is orchestrated through the INDIGO PaaS, that is able to schedule the request on the best provider of the federation; finally, the configuration of the resources is fully automated through ansible roles. All these technical details are hidden to the final users that can request the instantiation of the services through a user-friendly web portal.

Security is another key aspect that is carefully addressed in our platform. First of all, we adopt consistent authentication and authorization rules defined at the different IaaS, PaaS and SaaS levels, providing user/group isolation. The recipes used for automating the installation are then developed by IT experts that take care also of implementing secure configurations and updating the recipes as soon as a vulnerability is discovered. Finally, the INDIGO PaaS system allows also to perform deployments on private networks (through a bastion properly configured by the provider) where the deployed services can be reached through dedicated VPNs.

In this contribution we will provide details about both the platform architecture, the high-level service implementation strategy and the expected lines of further development.

**Primary authors:** ANTONACCI, Marica (INFN); SALOMONI, Davide (INFN)

**Co-authors:** CIASCHINI, Vincenzo; COSTANTINI, Alessandro (INFN-CNAF); DONVITO, Giacinto (INFN); DUMA, Doina Cristina (INFN - CNAF); FANZAGO, Federica; GIORGIO, Emidio; ITALIANO, Alessandro; SGARAVATTO, Massimo; SPIGA, daniele (INFN-PG); SPINOSO, Vincenzo; STALIO, Stefano; VERLATO, Marco; TRACOLLI, Mirko

**Presenters:** ANTONACCI, Marica (INFN); SALOMONI, Davide (INFN)

**Session Classification:** VRE

**Track Classification:** Track 5: Virtual Research Environment (including tools, services, workflows, portals, ... etc.)

Contribution ID: 27

Type: **Poster Presentation**

## Certification Mechanism to Assure Software Reliability with Digital Signature

Internet of Things (IoT) platforms are widely deployed both in scientific and social e-infrastructures. Security is one of crucial issues to assure secure data collections and analyses on the platforms. For example, the vulnerability of software running on an IoT device may give opportunity attackers to insert malicious processes in the software (or code) such as data leakage and falsification. Code developers, system administrators and users of an IoT platform need to confirm that codes, including applications, middleware and operating systems, used in the platform is reliable to avoid security incidents.

Code signing is commonly used to prove the reliability of codes. The code developer signs the code and publishes it with his/her digital signature, so that the user of the code can verify the integrity of the code and the code developer. However, the digital signature here does not assure that the code does not include vulnerability. Vulnerability scanning tools are also used to detect vulnerability in a code. The tools verify if the scanned code includes known vulnerability. For example, Docker Scan enables the code developer or the user to verify vulnerabilities included in container images. While these tools assure the reliability of the code at the moment of scanning, they do not verify the reliability of the code after scanning. If a new vulnerability is found after the user run the scanned code, the tools do not have means to stop the code.

In this presentation, we propose a software certification mechanism to provide an assurance of software reliability with digital signature. We assume that the code is reliable, if the integrity and the developer of the code are proved, and if the code has passed vulnerability scanning. The proposed architecture is organized by a code developer, a code user, a code validator and a certificate authority. The code developer signs the code and submit it to the code validator. The code validator is a third-party organization that provides a code verification service. The service scans the code submitted by the code developer and sign the code if it confirms that the code does not include known vulnerability. The code user can confirm the reliability of the code by verifying the digital signature of the code, and the user runs the reliable code on IoT devices or servers. The certificate authority issues digital certificates for code signing and clients (a code developer and a code validator). A new vulnerability may be found after scanning, or at the moment of the code validator signed the code. Once the new vulnerability is found, the certificate authority revokes the digital certificate used for signing the vulnerable code and notifies the code user, so that the user can immediately make decision, e.g., stopping the running code or applying the workaround.

We present the detailed architecture of the proposed mechanism and discuss its operational issues: Which entity (an application source codes, library codes, OS packages) should the code validator sign? How does the certificate authority immediately notify certificate revocations to users?

This work is supported by JST, CREST Grant Number JPMJCR21M3.

**Primary authors:** SHIMIZU, Sayako (National Institute of Informatics); SAKANE, Eisaku (National Institute of Informatics); NISHIMURA, Takeshi (National Institute of Informatics); AIDA, Kento (National Institute of Informatics); Mrs TAKEFUSA, Atsuko (National Institute of Informatics)

**Presenter:** SHIMIZU, Sayako (National Institute of Informatics)

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 28

Type: **Oral Presentation**

## The implementation of Data Management and Data Service for HEPS (Remote presentation)

China's High Energy Photon Source (HEPS), the first national high-energy synchrotron radiation light source and soon one of the world's brightest fourth-generation synchrotron radiation facilities, is being under intense construction in Beijing's Huairou District, and will be completed in 2025.

To make sure that the huge amount of data collected at HEPS is accurate, available and accessible, we developed an effective data management system that is aimed at automating the organization, transfer, storage, distribution and sharing of the data produced from HEPS experiments. First, the general situation of HEPS and the construction progress of the whole project are introduced. Second, the architecture and data flow of the HEPS DMS are described. Third, key techniques and new function modules implemented in this system are introduced. For example, the process of automatic data tracking when using a hierarchical storage policy is illustrated, and how the DMS deals with the metadata collection when an emergency occurs such as beamline network interruption. Finally, the progress and the effect of the data management and data service system deployed at testbed beamlines of BSRF are given.

**Primary authors:** HU, Hao (Institute of High Energy Physics); Mr WANG, Haofan (IHEP); Mr LUO, Qi (IHEP)

**Presenter:** HU, Hao (Institute of High Energy Physics)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data

Contribution ID: 29

Type: **Oral Presentation**

# Efficient Deep Reinforcement Learning with Probability Mask in Online 3D Bin Packing Problem

*Wednesday, March 22, 2023 4:20 PM (20 minutes)*

3D Bin Packing is the problem of finding the best way to pack several cargos into a container in order to maximize the container density. Moreover, some problems have constraints such as weight, stack-ability, fragility, and orientation of cargo pieces. Since the 3D Bin Packing problem is known to be NP-hard, and an exact solution is hard to be obtained in a reasonable time. Therefore, various approximate solution methods have been proposed. We focused on methods using deep reinforcement learning (DRL) to overcome its weakness: its inapplicability to large-scale problems.

In this study, we propose a method to incorporate heuristic computation into the solution of the bin-packing problem using deep reinforcement learning, we propose a method that applies ideas such as the Bottom-Left and Best-Fit methods without searching all the space in the container. The proposed method presents candidate solutions in advance by applying ideas such as the Bottom-Left and Best-Fit methods. Then, a MASK is created with a certainty or binary value that indicating whether the cargo can be placed or not. The MASK is used to narrow the action space by multiplying it by action probabilities produced by DRL; thereby it leads to improve efficiency of training. This method significantly reduces the search space while maintaining solution accuracy, and is shown to be effective for efficient learning and reduced computational cost.

Through these efforts, we demonstrated the usefulness of using probability distributions with MASK to present candidate solutions using heuristics, and showed the possibility of applying deep reinforcement learning to more complex problems. The proposed learning method improves learning efficiency and achieves performance comparable to that of conventional methods.

In the future, we plan to conduct experiments on the problem of packing cargo of various shapes and materials.

**Primary authors:** Mr NAKAJIMA, Takumi (Osaka University); Prof. LEE, Chonho (Okayama University of Science); Prof. MASHITA, Tomohiro (Osaka University)

**Presenter:** Mr NAKAJIMA, Takumi (Osaka University)

**Session Classification:** Artificial Intelligence (AI)

**Track Classification:** Track 10: Artificial Intelligence (AI)

Contribution ID: 30

Type: **Oral Presentation**

# Building International Research Software Collaborations in Physics

*Thursday, March 23, 2023 9:40 AM (40 minutes)*

Building successful multi-national collaborations is challenging. The scientific communities in a range of physical sciences have been learning how to build collaborations that build upon regional capabilities and interests over decades, iteratively with each new generation of large scientific facilities required to advance their scientific knowledge. Much of this effort has naturally focused on collaborations for the construction of hardware and instrumentation. Software has however also become a critical element to design and maximize the physics discovery potential of large data intensive science projects. To fully realize their discovery potential a new generation of software algorithms and approaches is required. Building these research software collaborations is challenging and inherently international, matching the international nature of the experimental undertakings themselves. Initiatives such as the HEP Software Foundation have been instrumental in establishing international research software collaborations in high-energy physics, in particular between European and North American researchers.

This talk is about a new initiative, HSF-India, aiming to implement new and impactful research software collaborations between India, Europe and the U.S. The experimental scope of this project is relatively broad, aiming to bring together researchers across facilities with common problems in research. The research and development scope is on three primary topics: analysis software and integrated facilities for analysis; simulation techniques including generators and Artificial Intelligence based approaches; and enabling open science. By exploiting national capabilities and strengths, an immediate mutual benefit of the international collaboration will be a training network that enables early-career researchers to pursue impactful research software initiatives in ways that advance their careers in experimental data-intensive science. In this presentation, we will describe the scope of this initiative, its mechanisms for fostering new collaborations, and ways for interested research groups to get involved. We will also discuss thoughts towards broadening our initiative to foster more general collaborations in research software projects between Asian researchers and European/North American researchers who are already jointly pursuing “team-science” endeavors in research software for high-energy, nuclear and astro-particle physics.

**Primary authors:** LANGE, David (Princeton University); Dr ELMER, Peter (Princeton University)

**Presenter:** LANGE, David (Princeton University)

**Session Classification:** Keynote Speech III

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 31

Type: **Oral Presentation**

## Enabling Communities - Building trust for research and collaboration

*Thursday, March 23, 2023 11:00 AM (30 minutes)*

### Enabling Communities - Building trust for research and collaboration

When exploring the world of Federated Identity, research communities can reap considerable benefit from using common best practices and adopting interoperable ways of working. EnCo, the Enabling Communities task of the GÉANT 4-3 and GÉANT 5-1 Trust and Identity Work Package, provides the link between those seeking to deploy Federated Identity Management and the significant body of knowledge accumulated within the wider community. Individuals from EnCo aim to ensure that outputs from projects (e.g. AARC) and groups (e.g. WISE, FIM4R, IGTF, REFEDS) are well known, available and kept up to date as technology changes. Since many of these groups are non-funded, it's vital for their survival that projects such as the GÉANT project sponsor individuals to drive progress and maintain momentum. The ultimate aim is to enhance trust between identity providers and research communities/infrastructure, to enable researchers' safe and secure access to resources.

As we commence the work programme for the GEANT-5 phase, which starts in 2023, it is a good moment to review the impact on the trust and identity world achieved through the previous programme, and how global engagement can be promoted as the community gets ever more interconnected. The next GEANT programme will build on the same open structures of WISE, FIM4R, IGTF and REFEDS, so that shared knowledge is maintained and updated in the future - something essential for interoperability, trust and security.

The Federated Identity Management for Research (FIM4R) community is a forum where Research Communities meet to establish common requirements, combining their voices to send a strong message to FIM stakeholders. For example, in 2020 people from EnCo were among those who led efforts to produce a position paper on the EOSC identity management strategy from the perspective of research communities as well as the rebooting of the FIM4R activities post-pandemic

The WISE community promotes best practice in information security for IT infrastructures for research. EnCo has been and is leading several activities within WISE. This includes the Security for Collaborating Infrastructures working group, which has produced a guidance document to encourage self-assessment against the SCI Trust Framework and is working towards updating the AARC Policy Development Kit (PDK). Also, since information security processes need periodic exercise, the community organises challenges for communications response and mitigation of incidents affecting collaborative communities, and at times even deep forensics - all to make sure communities are prepared, and the various tests complement each other.

REFEDS is the voice that articulates the mutual needs of research and education identity federations worldwide. EnCo has been leading and participating in several activities on both assurance (the REFEDS Assurance Suite) and security to increase the level of trust in federations (SIRTFI). Trust in community for AARC proxy services is further promoted with the IGTF guidance on secure attribute authority operations and exchanging assurance for the infrastructures.

Our target audience are the communities and the infrastructures providing their services.

Aims of the presentation:

- The audience will learn about essential trust, policies and guidance
- Raise awareness of the availability of common resources, including those owned by WISE, FIM4R, REFEDS, IGTF



- Promote participation in these bodies and groups
- Share news of progress, e.g. Updates on the PDK, Sirtfi, Assurance
- Inform about future activities, e.g. trust for proxies and moving towards tokens
- Get input on our new activities

**Primary authors:** GROEP, David (Nikhef); KREMERS, Maarten (SURF); KELSEY, David (STFC-RAL)

**Presenter:** KREMERS, Maarten (SURF)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 32

Type: **Oral Presentation**

## Recent activities of the WISE Security for Collaborating Infrastructures (SCI) working group

*Tuesday, March 21, 2023 4:00 PM (30 minutes)*

The mission of the “WISE” community is to enhance best practice in information security for IT infrastructures for research. WISE fosters a collaborative community of security experts and builds trust between those IT infrastructures. Through membership of working groups and attendance at workshops these experts participate in the joint development of policy frameworks, guidelines, and templates.

The WISE Security for Collaboration among Infrastructures Trust Framework was first presented to the ISGC Conference in 2013. Since then the trust framework was updated to version 2 in 2017 and WISE activities were last presented to ISGC in 2021.

Since ISGC2021, the WISE Security for Collaborating Infrastructures (SCI) working group, in collaboration with Trust and Security activities in the GEANT GN4-3 Enabling Communities task and other projects and Infrastructures, has completed work on the guidance for self assessment of an Infrastructure’s maturity against the SCI Trust Framework.

Work on updating the Policy Development Kit, an output of the EU Horizon 2020 projects Authentication and Authorisation for Research Collaborations (AARC/AARC2) continues. An updated template policy for the Service Operations Security Policy has been completed. The working group has also worked on updating the AARC guidance on the handling of Data Privacy (and GDPR) issues in the operational access data collected by Infrastructures. Work on updating the AARC PDK policies relating to the Communities of Infrastructure Users is also underway.

This talk will report on the WISE SCI Working Group activities since ISGC2021 together with details of the new policy templates and guidelines produced.

**Primary authors:** GROEP, David (Nikhef); KELSEY, David (STFC-RAL); KREMERS, Maarten (SURF)

**Presenter:** KREMERS, Maarten (SURF)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 33

Type: **Oral Presentation**

## Multi-domain anycasted high availability for stateful services in RCauth.eu, now made simple

*Friday, March 24, 2023 11:50 AM (30 minutes)*

Use of ‘anycasting’ internet addresses (‘IP anycast’) in load balancing and high availability, and for traffic engineering reasons, is a widely deployed technique for content delivery networks to lower latency for access to frequently accessed content such as web pages and video. Using the properties of the Border Gateway Protocol (BGP) as a variable-length path-vector protocol for routing internet packets, distinct hosts in multiple places in the internet announce the same network address to serve the same content. This provides redundancy of service provisioning, and at the same time offers the possibilities for traffic engineering by varying the perceived path length in the ‘default-free zone’ of the global routing table.

The most common deployment of anycast is a single organization managing all underlying hosts, and then announcing their hosts either using their own autonomous system, or from a range of autonomous systems all under single administrative control. The provisioning hosts themselves are also usually ‘stateless’ – they either service static content or obtain any state required from upstream sources that are not publicly exposed.

The RCauth.eu federated token translator is a service that issues end-user ‘PKIX’ certificates with globally unique, persistent, and non-reassigned identifiers based on eduGAIN-federated authentication. However, the uniqueness and non-reassignment must be guaranteed by the service itself, and hence it maintains state in a back-end database that is consulted and updated on issuance of each certificate.

The initial deployment of RCauth.eu consisted of simple hardware security module and security controls at a single site, Nikhef in Amsterdam, which could sustain a very low issuance volume. For deployment in more communities and infrastructures, and in the European Open Science Cloud, a most robust solution was required. A collaboration of Nikhef (Amsterdam, The Netherlands), GRNET (Athens, Greece), and STFC (Didcot, Oxford, UK) therefore initiated a more robust setup using a distributed RCauth.eu service, where each site hosts a fully replicated instance. Since the user experience must be consistent (a persistent, unique, and mostly unchanging credential based the user’s federated identity), the service has to be supported by a distributed database that retains near-synchronous state across all instances. However, since the expected total issuance volume for RCauth.eu is unlikely to exceed the capacity of one instance, the primary purpose of the distributed setup is to provide redundancy and rapid fail-over, rather than load balancing.

In building the distributed RCauth.eu, we reviewed several distributed high availability techniques that aim to remove single points of failure, and work without operator intervention. Since the transaction flow in token translation can take several minutes (due to the user authentication interaction with the home organisation), failures occurring during that period must be absorbed, and be independent of the settings on the client devices. It should also work across administrative domains and across countries and regions. Based on those requirements, we selected BGP Anycast as the most appropriate technology, but engineered the system in such a way that it minimally affects existing systems and network operations. We demonstrate that we can build a stateful anycasted service across three countries and two autonomous systems, achieve rapid (seconds-scale) failover, can synchronise databases over transport-protected L4 virtual circuits while maintaining a consistent database state. And by considering an integrated approach of service and host management, internal routing, and eBGP engineering, we show how to build a highly available multi-domain and multi-national service without requiring additional autonomous system resources.

**Primary author:** GROEP, David (Nikhef and Maastricht University)

**Co-authors:** SALLÉ, Mischa (Nikhef); GKINIS, Kyriakos (GRNET); LIAMPOTIS, Nicolas (GRNET); FURNELL, Will (UKRI-STFC); JENSEN, Jens (UKRI-STFC)

**Presenter:** GROEP, David (Nikhef and Maastricht University)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 34

Type: **Oral Presentation**

## Overcoming obstacles to IPv6 on WLCG

*Friday, March 24, 2023 9:30 AM (30 minutes)*

The transition of WLCG storage services to dual-stack IPv4/IPv6 is nearing completion after more than 5 years, thus enabling the use of IPv6-only CPU resources as agreed by the WLCG Management Board and presented by us at earlier ISGC conferences. Much of the data is transferred by the LHC experiments over IPv6. All Tier-1 storage and over 90% of Tier-2 storage is now IPv6-enabled, yet we still see IPv4 transfers happening when both endpoints have IPv6 available or when remote data is accessed over the network from worker nodes.

The monitoring and tracking of all data transfers is essential, together with the ability to understand the relative use of IPv6 and IPv4. This paper presents the status of monitoring IPv6 data flows within WLCG and the plans to improve the ability to distinguish between IPv6 and IPv4. Furthermore, the Research Networking Technical Working Group has identified marking the IPv6 packet header as one approach for understanding complex large data flows. This provides another driver for full transition to the use of IPv6 in WLCG data transfers.

The agreed endpoint of the WLCG transition to IPv6 remains the deployment of IPv6-only services, thereby removing the complexity and security concerns of operating dual stacks. The working group is identifying where IPv4 can be removed and investigating the obstacles to the use of IPv6 in WLCG. Why do transfers between two dual-stack endpoints still use IPv4? This work is presented together with the obstacles defeated, those remaining, and those outside of our control.

**Primary authors:** HOEFT, Bruno (Karlsruhe Institute of Technology); KELSEY, David (STFC-RAL); MARTELLI, Edoardo (CERN); CHUDOBA, Jiri (Institute of Physics of the Czech Academy of Sciences)

**Presenter:** CHUDOBA, Jiri (Institute of Physics of the Czech Academy of Sciences)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 35

Type: **Oral Presentation**

## CNAF experience in support of the JUNO distributed computing model

*Thursday, March 23, 2023 11:30 AM (30 minutes)*

The Italian WLCG Tier-1 located in Bologna and managed by INFN CNAF provides computing and storage resources to several research communities in the fields of High-Energy Physics, Astroparticle Physics, Gravitational Waves, Nuclear Physics and others. Among them, the Jiangmen Underground Neutrino Observatory (JUNO), devoted to the construction and operation of a neutrino detector located underground in Kaiping, Jiangmen in Southern China, will employ a computing infrastructure geographically distributed in Chinese, Russian, French and Italian datacenters. The detector data rate is expected to be of the order of 2 PB per year, continuously transferred from the detector site to the INFN Tier-1 in Italy. To guarantee the optimal operations among all the aforementioned sites, a series of periodic network and data management challenges have been performed.

In this talk, the technologies involved to set up the cross-continent data transfer (e.g. StoRM Web-DAV, EOS, dCache, XrootD, FTS, Rucio) are presented, together their performance.

**Primary author:** RENDINA, Andrea (INFN - CNAF)

**Co-authors:** Dr MORGANTI, Lucia (INFN - CNAF); Dr ZANI, Stefano (INFN - CNAF); Dr DE GIROLAMO, Donato (INFN - CNAF); Dr CHIARELLI, Lorenzo (GARR); Dr FORNARI, Federico (INFN - CNAF)

**Presenter:** RENDINA, Andrea (INFN - CNAF)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data

Contribution ID: 36

Type: **Oral Presentation**

## Data management for the InterTwin project

*Thursday, March 23, 2023 1:30 PM (30 minutes)*

InterTwin is an EU-funded project that started on the 1st of September 2022.

The project will work with domain experts from different scientific domains in building a technology to support the emerging field of digital twins.

Digital twins are modelled for predicting the behaviour and evolution of real-world systems and applications.

InterTwin will focus on employing machine-learning techniques to create and train models that are able to quickly and accurately reflect their physical counterparts in a broad range of scientific domains.

The project will develop, deploy and “road harden” a blueprint for supporting digital twins on federated resources.

For that purpose, it will support a diverse set of science use-cases, in the domains of radio telescopes (Meerkat), particle physics (CERN/LHC and Lattice-QCD), gravitational waves (Einstein telescope), as well as climate research and environment monitoring (e.g. prediction of flooding and other extreme weather due to climate change).

The ultimate goal is to provide a generic infrastructure that can be useful in many additional scientific fields.

In the talk, we will present an overview of the interTwin project along with the corresponding architecture.

Its focus will be on the federated data management layer that is designed to support both the training and exploitation of digital twins within the different scientific domains.

The challenges faced when designing the architecture will be described, along with the solutions being developed to address them.

**Primary authors:** MANZI, Andrea (EGI); FUHRMANN, Patrick (DESY/dCache.org); MILLAR, Paul (DESY); WETZEL, Tim (Deutsches Elektronen-Synchrotron DESY); SPIGA, daniele (INFN-PG)

**Presenter:** WETZEL, Tim (Deutsches Elektronen-Synchrotron DESY)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data

Contribution ID: 37

Type: **Oral Presentation**

## Experiments support at INFN-T1

*Friday, March 24, 2023 11:20 AM (30 minutes)*

The Italian WLCG Tier-1 located in Bologna and managed by INFN, provides computing resources to several research communities in the fields of High-Energy Physics, Astroparticle Physics, Gravitational Waves, Nuclear Physics and others. The facility is hosted at CNAF. Although the LHC experiments at CERN represent the main users of the Tier-1 resources, an increasing number of communities and experiments are also being supported in all of their computing activities. Due to this demanding user base, an efficient support system is needed in order to assure a smooth and appropriate exploitation of the computing infrastructure.

In this framework, such role is played by the Tier-1 User Support group, which acts as a first level support and represents the entry point for services and support requests, and problem reports. The group makes use of multiple systems to meet the different needs and specificities of the supported experiments. Moreover, the group continuously maintains detailed knowledge based in the form of an online users guide.

The communication channels are represented by ticketing systems and also by mailing lists used for a more direct communication with users, allowing to notify maintenance interventions, downtimes and more in general all the new features and services provided by the datacenter.

In this talk, the ticketing systems, tools, platforms and services that User Support offers, and the internal organization of the department will be described. Future workflow plans in view of the DATA CLOUD project, which will require an increasing effort, will also be presented.

**Primary author:** Dr LATTANZIO, Daniele (INFN - CNAF)

**Co-authors:** Dr RENDINA, Andrea (INFN-CNAF); Dr PASCOLINI, Alessandro (INFN-CNAF); Dr MORGANTI, Lucia (INFN-CNAF); Dr FORNARI, Federico (INFN-CNAF); Dr PELLEGRINO, Carmelo (INFN-CNAF); Dr CESINI, Daniele (INFN-CNAF)

**Presenter:** Dr LATTANZIO, Daniele (INFN - CNAF)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations



Contribution ID: 38

Type: **Oral Presentation**

## The design of the unified identity authentication system for HEPS (Remote Presentation)

*Thursday, March 23, 2023 2:30 PM (30 minutes)*

The Institute of High Energy Physics of the Chinese Academy of Sciences is a comprehensive research base in China engaged in high-energy physical research, advanced accelerator physics and technology research and development and utilization, and advanced ray technology and application.

The Single Sign On (SSO) system of the High Energy Institute has more than 22,000 users, the calculation cluster (AFS) users 3.2K, Web applications of more than 150, and more than 10 client applications. With the development of the high energy institute, international cooperation has become more and more frequent, so the SSO system of the high energy institute has been generated. The SSO system of the High Energy Institute integrates all personnel systems and AFS user accounts. And realize the Chinese certification federation CARSI and international federation EduGain access, not only realize the unified account management of the within the place, but also gradually realize domestic universities and international organization certification.

**Primary authors:** HU, Hao (Institute of High Energy Physics); LUO, qi (The Institute of High Energy Physics of the Chinese Academy of Sciences)

**Presenter:** LUO, qi (The Institute of High Energy Physics of the Chinese Academy of Sciences)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 39

Type: **Oral Presentation**

## An FPGA implementation of Variational Autoencoders for anomaly detection in HEP (Remote presentation)

*Thursday, March 23, 2023 11:30 AM (30 minutes)*

If new physics does exist at the scales investigated by the Large Hadron Collider (LHC) at CERN, it is more elusive than expected.

Finding interesting results may be challenging using conventional methods, usually based on model-dependent hypothesis testing, without substantially increasing the number of analyses. Thus, standard signal-driven search strategies could fail in reaching new results, and unsupervised machine learning techniques could fill this critical gap.

Such applications, running in the trigger system of the LHC experiments, could spot anomalous events that would otherwise go unnoticed, enhancing the LHC's scientific capabilities.

The most basic unsupervised machine learning technique is an autoencoder (AE) with a bottleneck. It is constructed using a network that translates a high-dimensional data representation onto itself to create an average or typical entity.

Standard autoencoders are recommended for unsupervised jet classification, but they are known to have problems in more general applications.

The AE learns to compress and rebuild the training data very well, but when new, untrained data is run through the trained AE, it will produce a considerable loss or reconstruction error.

By using the AE, it is possible to search for data that differs significantly from training data or even training data that is a small subclass of anomalous instances.

In addition, the AE fails if the anomalous data is structurally simpler than the dominant class because the AE can encode simpler data with fewer features more efficiently.

It is possible to overcome the disadvantages of AE by substituting a different classification measure for the reconstruction error.

A possible alternative approach to the reconstruction error in the case of Variational Autoencoders (VAEs) is to derive a metric from the latent space embedding.

The work will consist of implementing a VAE model targeted at FPGA (Field Programmable Gate Array) hardware architecture in order to determine the best latency and resource consumption without sacrificing model accuracy.

Models will be optimized for classification between anomalous jets and QCD jets images, in an unsupervised setting by training solely on the QCD background.

A comparison will be made between the reconstruction error and a latent space metric to determine the best anomaly detection score that enhances the separation of the two classes.

The goal of the model is to reconstruct the input data information as accurately as possible.

Additionally, because of the design of the VAEs architecture, the high-dimensional data representation is transformed into a compressed lower-dimensional latent distribution during the encoding stage.

Subsequently, the decoder learns stochastic modelling and aims to generate input-like data by sampling from the latent distribution.

The information about each dataset instance hidden in the high-dimensional input representation should be present in the latent space after training and the model can return the shape parameters describing the probability density function of each input quantity given a point in the compressed space.

With this application at the LHC, it could ideally be possible to classify the jets and even find anomalies using this latent space representation.

A companion tool based on High-Level Synthesis (HLS) named HLS4ML will implement deep learning models in FPGAs.

Furthermore, a compression and quantization optimization of neural networks will reduce the model size, latency, and energy consumption.

We expect VAE will find many uses in science, outperforming classical or standard deep learning baselines and even being able to solve challenges for physics beyond the Standard Model that were previously unsolvable.

Ideally applicable in a wide spectrum of signal background discrimination through anomaly detection, this application is expected to produce excellent results in a variety of fields.

**Primary author:** Mr VALENTE, Lorenzo (University of Bologna)

**Co-authors:** LORUSSO, Marco (Alma Mater Studiorum - University of Bologna); ANZALONE, Luca (University of Bologna); BONACORSI, Daniele (University of Bologna)

**Presenter:** Mr VALENTE, Lorenzo (University of Bologna)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 40

Type: **Oral Presentation**

## AAI Prototyping for SKA (Remote presentation)

*Thursday, March 23, 2023 2:00 PM (30 minutes)*

The Square Kilometre Array (SKA) telescope's computing platform is being developed through an agile process, with teams from across the SKA Regional Centres (SRCs) developing the SRCNetwork (SRCNet) infrastructure the SKA will need.

One such area of development is the SRCNet's Authentication and Authorisation Infrastructure (AAI), which is currently led by an agile team, Purple Team, with broad international membership. This team's goal is to prototype an interoperable and scalable AAI solution for the SKA SRCNet, considering the requirements and existing infrastructures at different countries SRCs. The team will aim to use lessons learned so far from both the SRCs and other science communities, such as CERNs Worldwide LHC Computing Grid (WLCG) and UK IRIS (eInfrastructure for Research and Innovation for STFC). Part of this was facilitated by producing an AAI Landscape Report which documents tools and the technologies, policies, and procedures which underpin AAI, and national and international implementations.

This talk will cover the development progress made so far, detailing design decisions, the research work underpinning them, and the forward plans for the SRCNet AAI development.

**Primary author:** DACK, Thomas (STFC - UKRI)

**Presenter:** DACK, Thomas (STFC - UKRI)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 42

Type: **Poster Presentation**

## COMBINATION OF LOW LEVEL AND HIGH LEVEL FEATURES TO IMPROVE THE EFFICIENCY OF CONTENT BASED IMAGE RETRIEVAL BASED ON THE EMR MANIFOLD RANKING ALGORITHM

### ABSTRACT

Effective Manifold Ranking (EMR) is a technique used widely in Content-based Image Retrieval (CBIR) to rank the images in a database via measuring and ranking the similarity between each image with a given query image where the images are represented by different features.

In this paper, a combined of low level and Deep features is proposed to create an EMR graph is proposed. In the details, Deep features are extracted by using pre-train CNNs which these networks have shown strong performance in the aspect of image discrimination.

Experimental results show that the combination method of low-level and high-level features significantly increased the accuracy of the EMR manifold ranking algorithm.

**Keywords:** CBIR, Effective Manifold Ranking, Convolution Neural Networks , Fuzzy C-means, Low level features, High level features.

**Primary authors:** Dr DAO VAN, TUYET (Vietnam National Space Center, Vietnam Academy of Science and Technolgt); Dr HOANG VAN, QUY (Hong Duc University, Ministry of Education and Training)

**Presenter:** Dr DAO VAN, TUYET (Vietnam National Space Center, Vietnam Academy of Science and Technolgt)

**Track Classification:** Track 10: Artificial Intelligence (AI)

Contribution ID: 43

Type: **Oral Presentation**

## Were the 2022 Tonga Eruption Meteotsunamis Caused by the Proudman Resonant?

*Friday, March 24, 2023 9:45 AM (45 minutes)*

This study uses numerical methods to explore the causes of meteotsunamis in the Atlantic Ocean, Caribbean Sea, and the Mediterranean Sea after the eruption of the Tonga Volcano on January 15, 2022. Topics are focused on the role of the Proudman resonant effect on the tsunami induced by the Lamb waves. The linear and nonlinear shallow water equation, fully-nonlinear and weakly-dispersive Boussinesq equations, weakly-nonlinear and weakly-dispersion Boussinesq equations, and the three-dimensional Navier Stokes equations are solved for simulating and discussing the phenomena of the Tonga Tsunami. In terms of boundary conditions, the atmospheric pressure data from the Central Weather Bureau of Taiwan and the pressure data from the ten-meter meteorological tower of the Central University are used. In terms of model verification, the free-surface elevation data of the tide station recorded by the Central Weather Bureau of Taiwan and the pressure data of the undersea cable are used. A wide range of Froude numbers of the moving pressure is introduced for understanding the effect of the Proudman resonant. The result shows that the free-surface elevation is positive right under the moving pressure if  $Fr > 1.0$ , a negative free-surface elevation is observed if  $Fr \leq 1.0$ . In the end, the moving-solid algorithm is introduced in the Navier-Stokes model for studying the stern wave generated by the moving pressure. The result shows that the wave amplification factor is about 5 in the case of Proudman resonant of the bow wave, and it is about 1.2 in the case of Tonga Tsunami. However, the wave amplification factor of the stern waves reaches 20 in the case of the Tonga Tsunami. Detailed results and discussion will be presented at the conference.

**Primary author:** WU, Tso-Ren

**Co-authors:** Mr HUANG, Yi-Xuan (Graduate Institute of Hydrological and Oceanic Sciences, National Central University); Dr LIN, Chun-Wei (Graduate Institute of Hydrological and Oceanic Sciences, National Central University); Ms TSENG, Hsin-Yi (Graduate Institute of Hydrological and Oceanic Sciences, National Central University)

**Presenter:** WU, Tso-Ren

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

**Track Classification:** Track 3: Earth & Environmental Sciences & Biodiversity Applications

Contribution ID: 44

Type: **Oral Presentation**

## Enhancing Spatial Reasoning Capability Using Virtual Reality Immersive Experience

*Wednesday, March 22, 2023 3:00 PM (30 minutes)*

Spatial reasoning is the ability to think about objects in two and three dimensions. Spatial reasoning skills are critical in science, art, and math and can be improved with practice. This research's main objective is to explore how virtual reality (VR) immersive experiences can enhance spatial reasoning capability. Past research revealed the vast differences between traditional user experiences and immersive experiences. VR uses tools to create artificial digital worlds that simulate physical ones. With head-mounted displays, users can detach their senses of sound, sight, and space from their surroundings to fully 'immerse' in simulated, computer-generated realities. Immersive experiences are everyday in the consumer space, particularly in the world of video gaming, but have become rapidly adopted by learning and education.

In this study, we propose a two-stage comparison experiment to explore the spatial reasoning skills of participants, who are selected from diversified backgrounds with reasonable spatial abilities and experience in virtual reality. In the first stage, participants will experience traditional hand drawing techniques in the physical environment before moving forward to drawing in a VR environment with Gravity Sketch application, and in the second stage, they are invited to conduct the traditional drawing session once more. Each session is designed as the following flow: begins with task introduction, then object and drawing process brainstorming, followed by object drawing, and ends with object related questions. During the experiment, Galvanic Skin Response (GSR) device is attached to participants in order to collect users' reflection pattern, followed by interviews of which the results are analyzed using text analysis technique to obtain more insights into participants' thoughts, and lastly, a survey is conducted to measure learning performance and immersive tendency of the participants. The preliminary results support the hypotheses and reveal that user immersion has a significant impact on user efficiency, user effectiveness, and user satisfaction, which are related to learning outcomes, and hence, to user spatial reasoning capability. These results have inspired us to plan further investigation regarding how the differences in backgrounds affect the process for a more comprehensive understanding about this topic. This research provides more insights into the applications of VR in learning spatial reasoning, which could be utilized and developed in educational settings, especially in STEAM (science, technology, engineering, art, mathematics), and other aspects as well.

**Primary authors:** Prof. WANG, Ryan Sheng-Ming (Department of Interaction Design, National Taipei University of Technology); Ms PHAN, Minh Nhan (Department of Interaction Design, National Taipei University of Technology); Ms DANG, My Linh (Department of Interaction Design, National Taipei University of Technology); Ms DAI, Yu-Xuan (Department of Interaction Design, National Taipei University of Technology); Mr LIN, Shao Ying (Department of Interaction Design, National Taipei University of Technology); Mr CHEN, Chun Chieh (Department of Interaction Design, National Taipei University of Technology)

**Presenters:** Ms PHAN, Minh Nhan (Department of Interaction Design, National Taipei University of Technology); Ms DANG, My Linh (Department of Interaction Design, National Taipei University of Technology); Ms DAI, Yu-Xuan (Department of Interaction Design, National Taipei University of Technology); Mr LIN, Shao Ying (Department of Interaction Design, National Taipei University of Technology)

nology); Mr CHEN, Chun Chieh (Department of Interaction Design, National Taipei University of Technology)

**Session Classification:** Humanities, Arts & Social Sciences

**Track Classification:** Track 4: Social Sciences, Arts and Humanities (SSAH) Applications



Contribution ID: 45

Type: **Oral Presentation**

## Physics analysis workflows and pipelines for the HL-LHC

*Thursday, March 23, 2023 12:00 PM (30 minutes)*

High Energy Physics analysis workflows commonly used at LHC experiments do not scale to the data volumes expected from the HL-LHC. A rich program of research and development is ongoing to address this challenge, proposing new tools and techniques for user-friendly analysis. The IRIS-HEP Analysis Grand Challenge (AGC) provides an environment for prototyping, studying and improving workflows in the context of a realistic physics analysis. The AGC defines an analysis task, based on publicly available Open Data, which captures the relevant technical requirements and features that physicists need. It furthermore provides a reference implementation that addresses this task.

The IRIS-HEP AGC reference implementation makes use of novel pieces of cyberinfrastructure from the HEP Python ecosystem and is executed on modern analysis facilities (e.g. coffea-casa and others). A coffea-casa analysis facility prototype is used as a testbed for the AGC, offering the possibility for end-users to execute analysis at HL-LHC-scale data volumes. This facility adopts an approach that allows transforming existing facilities (e.g. LHC Tier-2 and Tier-3 sites) into modular systems, using Kubernetes as the enabling technology. Multiple data delivery mechanisms and caching strategies are available for fast and efficient data reduction.

This contribution provides an overview of ongoing R&D work, describes the status of the AGC project, and showcases the envisioned analysis workflows and pipelines on analysis facilities.

**Primary authors:** HELD, Alexander (University of Wisconsin–Madison (US)); SHADURA, Oksana

**Presenters:** HELD, Alexander (University of Wisconsin–Madison (US)); SHADURA, Oksana

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 46

Type: **Oral Presentation**

## Building a PaN analysis platform using EOSC

*Friday, March 24, 2023 10:00 AM (30 minutes)*

The European Open Science Cloud (EOSC) is a key framework through which the EC is fostering the collaboration and interoperability of scientific research across domains to make services and data easily accessible to a broader audience and benefit from synergies.

The EC is establishing the EOSC by funding a series of projects, supporting either individual science domains to get up to speed with distributed cloud mechanisms and by encouraging multiple domains to collaborate and build their analysis stack on core EOSC services. EOSC-Future is one of the latter.

The activity within EOSC-Future is divided between work that applies generally and individual projects that each target a specific scientific domain. One such domain-specific project is building a Photon and Neutron (PaN) analysis platform. This project is building a service through which researchers may take advantage of EOSC services and resources when analysing data from various PaN facilities in Europe.

In this paper, we provide an overview of EOSC and EOSC-Future before describing the EOSC PaN analysis platform. The design and underlying architecture is presented together with the initial analysis use case. Some examples are given that demonstrate, concluding with a road-map for the services future.

**Primary authors:** Dr SANTONI, Gianluca (ESRF); FUHRMANN, Patrick (DESY/dCache.org); MILLAR, Paul (DESY)

**Presenter:** FUHRMANN, Patrick (DESY/dCache.org)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 47

Type: **Oral Presentation**

## **Distributed Data Management System at IHEP (Remote presentation)**

*Thursday, March 23, 2023 2:00 PM (30 minutes)*

A Distributed Grid-Data Management System serving BES, JUNO, CEPC experiments has been built at IHEP since 2014 based on DIRAC file catalog system. Meanwhile, more experiments, such as HERD or JUNO experiment with different data scales and complicated data management demands in data production, enforce us to make attempts on developing more flexible and experiments scenarios-oriented grid-data management system.

Considering its reliability, scalability and automation, the Rucio system became an option to IHEP grid data management infrastructures. In order to better understand Rucio system ability and performance, we designed its roles in current data production workflow and customized data policies corresponding to different experiments requirements.

This talk will introduce the design of JUNO/HERD experiment data flow and rucio's integration into present distributed data management system. The progress of Third-party-copy protocol adaption and token-based authentication system (IAM) deployment at IHEP, which work as the grid data base services and infrastructures, will also be included in this presentation.

**Primary authors:** ZHANG, Xuantong (IHEP, CAS); ZHANG, Xiaomei (IHEP, CAS)

**Presenter:** ZHANG, Xuantong (IHEP, CAS)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data

Contribution ID: 48

Type: **Oral Presentation**

## **K2I - Automated anomaly detection in the chemical footprint of surface water using machine learning**

*Wednesday, March 22, 2023 4:40 PM (20 minutes)*

River and lake water is a major resource for drinking water, food production and various industrial and agricultural purposes and it hosts or feeds many sensitive ecosystems. Therefore assuring the absence of potentially harmful chemicals is a vital issue for environmental and economic sustainability. Tens of thousands of different chemicals are present in fluctuating amounts in surface water bodies, many naturally occurring, others emitted human activities. Laboratories tasked with monitoring the water quality by using targeted analysis can only detect a small subset of the chemical substances present in the water. Therefore so called Non-Target-Screening (NTS) is increasingly used by labs to perform more comprehensive monitoring. This procedure typically relies on liquid chromatography in combination with high-resolution mass spectrometry (LC- HRMS). These produce a large number of signals which are difficult to evaluate for human researchers, especially when many samples are involved. We employ data processing techniques and pattern recognition methods like Autoencoders to structure the data and perform anomaly detection. A high degree of variance in measurements and processing workflows results in low comparability of data from different laboratories, which we alleviate with data alignment processes. The K2I project aims at fostering collaboration between laboratories and research institutions working towards the goal of advanced automated water quality monitoring. A joint platform for uploading and processing raw LC-HRMS data including a cloud based Datalake and processing pipeline is being developed. A standardized processing workflow is being established which is enhanced by anomaly detection to speed up the discovery of unusual changes in water bodies.

The measurements consist of signal peaks that correspond to a specific retention time (RT) in the chromatograph and a certain mass to charge ratio (m/z) which is determined by the mass spectrometer after the substances have been ionized. These peaks can be scanned for unknown combinations of RT and  $m/z$ , indicating the presence of so far unregistered chemicals and for strong or recurring signals that have not been attributed to a known cause yet. Neural networks like Autoencoders can be trained on historical data, to recognize common components and then spot deviations from these normal patterns. It can be advantageous to narrow down the source of emission to compare measurements taken at different locations, both in the same water body (different sites at the same river or lake) and in separate waters. Thus, the combination of LC-HRMS data from different sampling locations and laboratories, which can be enhanced with spatial and temporal coordinates and additional information such as known environmental influences could be used to more effectively notice and track micro pollutants in surface water across larger regions.

**Primary authors:** Mrs PAUW, Viktoria (Leibniz Rechenzentrum); Mr HAYEK, Mohamad (LRZ); Ms SHOJAEI, Elham (LRZ); Mr BADER, Tobias (Landeswasserversorgung (Ulm)); Mr HACHINGER, Stephan (LRZ)

**Presenter:** Mrs PAUW, Viktoria (Leibniz Rechenzentrum)

**Session Classification:** Artificial Intelligence (AI)

**Track Classification:** Track 10: Artificial Intelligence (AI)

Contribution ID: 49

Type: **Oral Presentation**

## **Document Management System and Application of Institute of High Energy Physics, CAS (Remote presentation)**

*Thursday, March 23, 2023 2:00 PM (30 minutes)*

The Institute of High Energy Physics, Chinese Academy of Sciences is a comprehensive research base engaged in high energy physics, advanced accelerator physics and technology, advanced ray technology and its application, and has built a series of large-scale scientific facilities in China, such as Beijing Electron Positron Collider (BEPC), China Spallation Neutron Source (CSNS), High Energy Photon Source (HEPS), etc.

The IHEP Document Management System (IHEP Docs) is designed and built in order to solve the management and utilization problems of a large number of high-value unstructured data such as research documents, technical documents and management documents generated in the process of research activities, research management and construction of large-scale facilities. The system is adopted the concept of unstructured data middle center and the architecture of unstructured content data bus which connects the two key platforms, the unified management services and the integrated application services, and deployed with lightweight cloud services, in order to ensure the efficient connection of different services and the efficient flow of unstructured data.

The management platform provides front-end services such as organization structure, document management, authentication and authority management, etc. The application platform provided an integrated application services such as collaborative editing of office documents, preview of CAD documents, OCR identification of picture documents, smart collaborative sheet, customizable document workflows, powerful search engine with AI, etc. Meanwhile, Anti-virus system and data loss prevention system are also integrated to protect the security of the IHEP Docs system and user data all the time.

IHEP Docs is expected to eliminate the data silos among different systems and different departments, realize an effective life-cycle management and a fully exploited value of the documents.

**Primary authors:** Mr QI, Mengyao (Institute of High Energy Physics, Chinese Academy of Sciences); Ms WANG, Li (Institute of High Energy Physics, Chinese Academy of Sciences); Mr LUO, Qi (Institute of High Energy Physics, Chinese Academy of Sciences); Mr SUN, Zhihui (Institute of High Energy Physics, Chinese Academy of Sciences); Mr QI, Fazhi (Institute of High Energy Physics, Chinese Academy of Sciences); HOU, Fengyao (Institute of High Energy Physics, CAS)

**Presenter:** HOU, Fengyao (Institute of High Energy Physics, CAS)

**Session Classification:** VRE

**Track Classification:** Track 5: Virtual Research Environment (including tools, services, workflows, portals, ... etc.)

Contribution ID: 50

Type: **Oral Presentation**

## A Study of Authentication Proxy Service for Various Research Communities

*Thursday, March 23, 2023 11:30 AM (30 minutes)*

GakuNin, an identity and access management federation in Japan, has provided a stable trust framework to academia in Japan so far. For common services that all constituent members of university or institution use such as e-journal service the framework has worked well. There are many research communities: data science, material science, high energy physics, and research project using high performance computing resources. However, unfortunately those communities do not always rely on identity providers joining GakuNin because all identity providers in GakuNin do not always satisfy the requirement from the communities. As a result, a trust framework has been forced to be formed in each research community. Many of users in the research communities are also members of IdPs that join GakuNin. It is natural for users to demand to use home organization account for services in the research communities. In other words, users should not want to manage several accounts in their academic activities. In order to solve the situation, GakuNin have launched a new working group. The goal of the working group is to build a new trust framework focused on identification and authentication. The new trust framework will be useful for research communities in Japan, namely, it must be enabling collaboration with business sector, promoting international collaboration, and also ensuring world-wide interoperability.

In order to make the new trust framework actually effective, we need a system that realizes the concept of the new trust framework. Namely it must be able to mediate between identity providers and services provided by various research communities and to bridge the gap between requirements from the services and credentials issued from the identity providers. Based on the basic idea mentioned the above, we have developed a new authentication proxy service, called "Orthros". Orthros supports the new GakuNin trust framework, bridge between identity providers and service providers, and enable identity assurance and authenticator assurance levels (IAL/AAL) management and also attribute assurance. In general, the requirements from service providers can be organized from IAL's or AAL's point of view. In order to satisfy the requirement of IAL, Orthros must be able to cooperate not only home organization identity providers operated by universities or institutions but also existing identity providers operated by governmental agencies, IT service vendors, social networking services or nonprofit organizations, because it is possible for home organization identity provider by itself not to be able meet the requirement from service providers. In this paper, we describe the details of the new authentication proxy service, Orthros. We explain the design and implementation of Orthros and the features in details. The future development plan of Orthros is also mentioned.

**Primary authors:** Dr SAKANE, Eisaku (National Institute of Informatics); Prof. NAKAMURA, Motonori (Kyoto University); Dr KOMIYAMA, Yusuke (National Institute of Informatics); Dr KOMURA, Takaaki (Kyoto University); NISHIMURA, Takeshi (National Institute of Informatics); SHIMIZU, Sayako (National Institute of Informatics)

**Presenter:** Dr SAKANE, Eisaku (National Institute of Informatics)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations



Contribution ID: 51

Type: **Oral Presentation**

## Analyze VIIRS night data in Indonesia during 2018 - 2022 (Remote presentation)

*Friday, March 24, 2023 10:50 AM (20 minutes)*

Forest fires that occurred in Indonesia started in 1998 and reached its peak in 2015 where in that year almost half of the world was affected by forest fires. One of the technologies used by NOAA is VIIRS (The Visible Infrared Imaging Radiometer Suite) which began operating in 2011. Most forest fires in Indonesia are mostly carried out by humans whose interests are to expand land, especially oil palm lands, which cover an area of tens to hundreds of thousands of hectares once land clearing is carried out. The peak of the forest fires occurred in 2015 where massive forest fires occurred and caused environmental impacts that not only affected the Indonesian people but also Indonesia's neighboring countries and even had their impacts felt in almost all Asian countries. In 2018, VIIRS released a product, namely VIIRS night, where data was released in one area, namely Indonesia. This data is issued daily for areas in Java, Sumatra, Kalimantan, Sulawesi and Indonesia in general. This VIIRS night data shows the possibility of burning at night. With these data, we began to identify certain areas that tend to burn forests at night, especially areas where large-scale forest fires frequently occur. With the existing data released per day per area, ANN (Artificial Neural Network) is used. The processes used start from Data Cleaning and Processing, Neural Net Creation, Train the ANN, Test the ANN. By using the ANN network, clustering per region of Indonesia, especially Sumatra, Kalimantan and Sulawesi where these areas are areas where forest fires are common, will be able to predict the possibility of forest burning, especially at night.

Keyword : VIIRS , Forest Fire , Artificial Neural Network

**Primary authors:** Dr HERMANTO, Beni Rio (ITB); SUHARDIMAN, basuk (itb)

**Presenter:** SUHARDIMAN, basuk (itb)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

**Track Classification:** Track 3: Earth & Environmental Sciences & Biodiversity Applications

Contribution ID: 52

Type: **Oral Presentation**

# Anomaly Detection in Data Center IT Infrastructure using Natural language Processing and Time Series Solutions

*Wednesday, March 22, 2023 5:00 PM (20 minutes)*

Data centers house IT and physical infrastructures to support researchers in transmitting, processing and exchanging data and provide resources and services with a high level of reliability. Through the usage of infrastructure monitoring platforms, it is possible to access data that provide data center status, e.g. related to services that run on the machines or to the hardware itself, to predict events of interest. Detecting unexpected anomalies is of great significance to prevent service degradation, hardware failures, data losses, and complaints from users. In the context of the data center of the Italian Institute for Nuclear Physics, which serves more than 40 international scientific collaborations in multiple scientific domains, including high-energy physics experiments running at the Large Hadron Collider in Geneva, we have performed a set of studies based on service log files and machine metrics.

Starting from our initial study aimed at combining a subset of log files and monitoring data information to detect anomaly patterns [1] involving heterogeneous unstructured data, natural language processing solutions have been applied to log files to identify words and sequences of terms as anomalies. Good results have been obtained, revealing thousands of anomalies verified by exploiting log-service messages. By defining an ad hoc clustering algorithm, various types of anomalies at the service level have been identified and grouped together. Furthermore, the adoption of a multivariate time series anomaly detection technique, called JumpStarter [2], enabled us to compute anomaly scores on monitoring data to identify the timeframe where we could overlap services and monitoring data anomalies to perform predictive maintenance analysis.

In the present work, we aim at validating the above mentioned model by considering critical scenarios and extending the range and type of monitoring data. By using error reconstruction algorithms based on, but not limited to, principal component analysis, clustering techniques, and statistical anomaly detection solutions, we plan to achieve a faster, real-time, detection of anomalies taking into consideration also the collection of past events. Furthermore, the relationship between the identified anomalies and the threshold-risk values will be assessed and shown as a dynamic level of risks to be used for predictive maintenance management. The defined pipeline can be exported to other data centers because of the usage of open source code for its implementation. It has to be considered that training and related inference may vary depending on the amount of data provided by the data center.

## References

- [1] Viola, L; Ronchieri, E; Cavallaro, C. Combining log files and monitoring data to detect anomaly patterns in a data center. *Computers*, 11(8):117, 2022. doi: <https://doi.org/10.3390/computers11080117>
- [2] Ma, M.; Zhang, S.; Chen, J.; Xu, J.; Li, H.; Lin, Y.; Nie, X.; Zhou, B.; Wang, Y.; Pei, D. Jump-starting multivariate time series anomaly detection for online service systems. In *Proceedings of the 2021 USENIX Annual Technical Conference (USENIX ATC 21)*, Virtual, 14–16 July 2021; pp. 413–426.

**Primary authors:** Dr COSTANTINI, Alessandro (INFN CNAF); Dr SALOMONI, Davide (INFN CNAF); Dr DOINA, Duma Cristina (INFN CNAF); Dr RONCHIERI, Elisabetta (INFN CNAF); Dr GIOMMI, Luca (INFN CNAF)

**Presenters:** Dr COSTANTINI, Alessandro (INFN CNAF); Dr SALOMONI, Davide (INFN CNAF); Dr DOINA, Duma Cristina (INFN CNAF); Dr RONCHIERI, Elisabetta (INFN CNAF); Dr GIOMMI, Luca (INFN CNAF)

**Session Classification:** Artificial Intelligence (AI)

**Track Classification:** Track 10: Artificial Intelligence (AI)

Contribution ID: 53

Type: **Oral Presentation**

## **DIRAC: OIDC/OAuth2 based security framework (Remote Presentation)**

*Thursday, March 23, 2023 1:30 PM (30 minutes)*

The DIRAC Interware is the framework for building distributed computing systems which allows to integrate various kinds of computing and storage resources in a transparent way from the user's perspective. Up until recently, the client communications with DIRAC were based on a custom protocol using X.509 PKI certificates. Following the recent move towards OIDC/OAuth2 based security infrastructure, the DIRAC client/server protocol was enhanced to support both proxy certificates and tokens. The new framework has components for user authentication and authorization with respect to the DIRAC services. It also has a Token Manager service for maintaining long-living tokens needed to support asynchronous operations on the user's behalf. The tokens now can be used to access computing resources such as HTCondorCE and ARC Computing Elements as well as cloud sites. Enabling access to the storage resources is also in the development. In this contribution we will describe the architecture of the DIRAC security framework and details of its implementation and usage in dedicated or multi-community DIRAC services.

**Primary authors:** TSAREGORODTSEV, Andrei (CPPM-IN2P3-CNRS); Mr LYTOVCHENKO, Andrii (CPPM-IN2P3-CNRS)

**Presenter:** TSAREGORODTSEV, Andrei (CPPM-IN2P3-CNRS)

**Session Classification:** VRE

**Track Classification:** Track 5: Virtual Research Environment (including tools, services, workflows, portals, ... etc.)

Contribution ID: 54

Type: **Poster Presentation**

## IMPROVING THE EFFICIENCY OF IMAGE RETRIEVAL BASED ON LEARNING THE DISTANCE MEASURE WITH RELEVANCE FEEDBACK

### ABSTRACT

The paper presents an effective method to learn the distance measure between images based on two processing phases: filter high-distinct images with query image and combine the feature sets of images. Our proposed algorithm calculates the non-linear distance measure of images, using the Fuzzy integral and based on the relevance feedback. The experiments have demonstrated that the performance of content-based image retrieval increases significantly when using the proposed measure. In the next study, we will continue the research direction to train and determine the measure on any image.

**Keywords:** Content-based image retrieval, relevance feedback, fuzzy measure.

**Primary authors:** Dr HOANG XUAN, TRUNG (University of Business and Technology, Ministry of Industry and Trade); Dr DAO VAN, TUYET (Vietnam National Space Center, Vietnam Academy of Science and Technology)

**Presenters:** Dr HOANG XUAN, TRUNG (University of Business and Technology, Ministry of Industry and Trade); Dr DAO VAN, TUYET (Vietnam National Space Center, Vietnam Academy of Science and Technology)

**Track Classification:** Track 10: Artificial Intelligence (AI)

Contribution ID: 55

Type: **Oral Presentation**

## **Secure deployments of Galaxy Servers for analyzing personal and Health Data leveraging the Laniakea service (Remote presentation)**

*Tuesday, March 21, 2023 4:40 PM (20 minutes)*

Data security issues and legal and ethical requirements on the storage, handling and analysing genetic and medical data are becoming increasingly stringent. Some regulatory obstacles may represent a gap to data sharing and the application of Open Science and Open Access principles. In this perspective, Task 6.6 of the EOSC-Pillar (<https://www.eosc-pillar.eu/>) project aimed to analyze the regulatory compliance of the integrated and interoperable PaaS level data analysis service (Laniakea) for ELIXIR and the Life Science community in general, the result of an interaction between Galaxy services and data repositories. Starting from the research activity carried out in the context of Task 6.6, the work aims to define the ethical and legal requirements that must be respected in order to guarantee an adequate balance between data and privacy protection and effective application of the FAIR, OS and OA principles.

From a technological point of view, we have implemented the necessary measures to improve the security of the entire service. In particular, the goal is to guarantee the creation of isolated and secure environments to carry out data analyses. To do this, we have focused on two critical aspects: data access and storing and network access control to the service.

User data isolation is accomplished by encrypting the entire storage volume associated with the virtual machine, using the Linux kernel encryption module. The level of disk encryption is completely transparent to software applications, in this case Galaxy. The procedure has been completely automated through the web Dashboard of the PaaS orchestration service (<https://github.com/indigo-dc/orchestrator>), taking advantage of Hashicorp Vault for storing user passphrases. After authenticating on the Dashboard, the user enables data encryption when setting up a new instance. The Dashboard contacts Hashicorp Vault to get a token that can only be used once. The token is passed to the encryption script on the virtual machine, a random passphrase is generated, the volume is encrypted, unlocked and formatted. Finally, the encryption script accesses Vault using the one-time token and stores the passphrase which will be accessible, at any time, only to the user via the Dashboard. This strategy makes it possible to create secure encryption keys and at the same time prevent user credentials or the encryption passphrase from being transmitted unencrypted to the virtual infrastructure, compromising its security.

The INDIGO Cloud orchestration system (PaaS layer) that allows the automatic deployment of the data analysis service has been extended to be able to manage the creation of virtual environments on a private network, taking advantage of the isolation of L2 virtual networks at the tenant level guaranteed by the cloud provider and automatically configuring appropriate security groups that monitor network traffic. In this way, access to the various analysis environments is blocked from the external network and also the traffic between virtual machines instantiated on the same private network, but belonging to different deployments, is filtered. The access to the service provided to users leverage VPN server at the tenant level. To improve the user experience, VPN authentication has been integrated with the authentication and authorization system, INDIGO IAM (<https://github.com/indigo-iam/iam>), used by the entire PaaS/IaaS stack and based on OpenID Connect. This way, users don't have to create additional accounts/credentials, but can use federated authentication. In particular, the solution implemented for the VPN is based on the OpenVPN open source software and on a PAM module developed ad-hoc to allow authentication

via IAM.

The solutions described have been tested and validated on the ReCaS-Bari (<https://www.recas-bari.it/index.php/en/>) cloud and on the INFN-Cloud (<https://www.cloud.infn.it/>) multi-site distributed infrastructure.

**Primary authors:** DONVITO, Giacinto (INFN); Dr TANGARO, Marco Antonio (CNR and INFN, Italy); ANTONACCI, Marica (INFN); Dr FOGGETTI, Nadina (INFN, Italy)

**Presenter:** Dr TANGARO, Marco Antonio (CNR and INFN, Italy)

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)

**Track Classification:** Track 2: Health & Life Sciences (including Pandemic Preparedness Applications)

Contribution ID: 56

Type: **Oral Presentation**

## **An open source Block chain as a service solution for health science applications (Remote presentation)**

*Tuesday, March 21, 2023 2:30 PM (30 minutes)*

The use of big data in the field of omics and biomedical studies is the enabling factor for finding new insights with sufficient statistical confidence. When dealing with such data, several issues have to be addressed, related to the personal identifiable information (PII) often included in datasets and subject to the European General Data Protection Regulation (GDPR), which imposes particular organizational and technical measures, aimed to protect patients' privacy. In this contribution we describe an open-source Blockchain-as-a-service solution, developed and deployed on the INFN Cloud infrastructure, and a health-related decentralized application (DApp) using this service. The main target of this DApp is to manage personal information while preserving patients' rights, thanks to the trusted, tamper-proof, traceable and accountable distributed digital ledger provided by the blockchain.

**Primary authors:** Dr MARTELLI, Barbara (INFN - CNAF); Dr SALOMONI, Davide (INFN CNAF )

**Presenter:** Dr MARTELLI, Barbara (INFN - CNAF)

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)

**Track Classification:** Track 2: Health & Life Sciences (including Pandemic Preparedness Applications)



Contribution ID: 57

Type: **Oral Presentation**

## **KIG: a tool for carbon footprint monitoring in physics research (Remote presentation)**

*Friday, March 24, 2023 10:50 AM (30 minutes)*

Greenhouse gas (GHG) emissions have been recognized as accelerators of the Global climate change phenomenon and several human activities take part in it. In particular, the contribution of the computing sector is significant and deemed to grow. While on one side unprecedented discoveries have been obtained thanks to the increasing computational power available, on the other the heavy reliance on power-eager resources might lead scientific research to become energetically unsustainable as a result of overlooking the burst of energy-intensive operations, resulting, as an example, from the spread of AI in most research fields, including Physics.

In order to guarantee the sustainability of research, all the stakeholders, namely users and data centers, should be able to keep track, analyze and report the carbon footprint and energy-intensiveness associated to their operations, in addition to the currently adopted performance metrics. By doing so, the stakeholders can reach a deeper understanding of the burden related to their operations and take informed decisions. For instance, users might plan energy optimizations of the workflow while data centers might adopt different management policies to abate the footprint of the facility. In this work, we introduce an open tool, written in C++, that allows users and data centers to easily keep track, analyze and report the energy requirements and carbon footprint in gCO<sub>2</sub>e of their computing tasks. Such tool should help shedding some light on the often not-so-trivial trade-off between performance and environmental Footprint. By gathering detailed data, such tool should also trigger meta-analyses on the behaviour of algorithms as well as computing infrastructures with a view to better leveraging said resources. In the following, sample Physics research-related use-cases are discussed to present the tool.

**Primary author:** MINARINI, Francesco (Alma Mater - Università di Bologna)

**Presenter:** MINARINI, Francesco (Alma Mater - Università di Bologna)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 58

Type: **Oral Presentation**

## **Resilience of the VO membership vetting process**

*Thursday, March 23, 2023 12:00 PM (30 minutes)*

This presentation reports on a series of exercises that checked the steps of the vetting process to gain VO membership for Check-in users. EGI Check-in accepts a range of identity providers on different trust levels, ranging from social media accounts where the identity provider can only guarantee that someone was in control of a mobile phone number or an email address.

**Primary author:** GRENIER, Baptiste (EGI Foundation)

**Co-author:** GABRIEL, Sven (Nikhef/EGI)

**Presenter:** GABRIEL, Sven (Nikhef/EGI)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 59

Type: **Oral Presentation**

## Scalable training on scalable infrastructures for programmable hardware (Remote presentation)

*Wednesday, March 22, 2023 2:00 PM (20 minutes)*

The increasingly pervasive and dominant role of machine learning (ML) and deep learning (DL) techniques in High Energy Physics is posing challenging requirements to effective computing infrastructures on which AI workflows are executed, as well as demanding requests in terms of training and upskilling new users and/or future developers of such technologies.

In particular, a growth in the request for training opportunities to become proficient in exploiting programmable hardware capable of delivering low latencies and low energy consumption, like FPGAs, is observed. While training opportunities on generic ML/DL concepts is rich and quite wide in the coverage of sub-topics, a gap is observed in the delivery of hands-on tutorials on ML/DL on FPGAs that can scale to a relatively large number of attendants and that can give access to a relatively diverse set of ad-hoc hardware with different hardware specs.

A pilot course on ML/DL on FPGAs - born from the collaboration of INFN-Bologna, the University of Bologna and INFN-CNAF - has been successful in paving the way for the creation of a line of work dedicated to maintaining and expanding an ad-hoc scalable toolkit for similar courses in the future. The practical sessions are based on virtual machines (for code development, no FPGAs), in-house cloud platforms (INFN-cloud infrastructure equipped with AMD/Xilinx Alveo FPGA), Amazon AWS instances for project deployment on FPGAs - all complemented by docker containers with the full environments for the DL frameworks used, as well as Jupyter Notebooks for interactive exercises. The current results and plans of work along the consolidation of such a toolkit will be presented and discussed.

Finally, a software ecosystem called Bond Machine, capable of dynamically generate computer architectures that can be synthesised in FPGA, is being considered as a suitable alternative to teach FPGA programming without entering into the low-level details, thanks to the hardware abstraction it offers which can simplify the interaction with FPGAs.

**Primary authors:** COSTANTINI, Alessandro (INFN-CNAF); Prof. BONACORSI, Daniele (University of Bologna); SPIGA, Daniele (INFN-PG); SALOMONI, Davide (INFN); MICHELOTTO, Diego (INFN-CNAF); DUMA, Doina Cristina (INFN - CNAF); LORUSSO, Marco (Alma Mater Studiorum - University of Bologna); MARIOTTI, Mirko (Department of Physics and Geology, University of Perugia); VERONESI, Paolo (INFN); Dr TRAVAGLINI, Riccardo (INFN)

**Presenter:** LORUSSO, Marco (Alma Mater Studiorum - University of Bologna)

**Session Classification:** Converging Infrastructure Clouds, Virtualisation & HPC

**Track Classification:** Track 8: Infrastructure Clouds and Virtualizations

Contribution ID: 60

Type: **Oral Presentation**

## Recent developments in the data analysis integrated software system of HEPS (Remote presentation)

Recent advances of X-ray beamline technologies, including the high brilliance beamlines at next-generation synchrotron sources and advanced detector instrumentation, have led to an exponential increase in the speed of data collection. As a consequence, there is an increasing need for a data analysis platform that can refine and optimize data collection strategies online and effectively analyze large volume data after the data collection.

Data analysis integrated software system (Daisy) has been designed to address the requirements of the next-generation advanced Synchrotron radiation source, such as the High Energy Photon Source (HEPS). Daisy aims to support on-site data analysis services with fast feedback and interaction, and offline analysis of large-scale data.

We provide an update on recent developments in the Daisy framework as well as on the custom application for specific scientific areas developed based on Daisy. This involves a general-purpose graphical user interface for visualization of raw or processed data, a web-based interactive computing platform for remote access, a web-based application for micro and nano computed tomography, the data-processing pipeline for biological macromolecular crystallography beamlines based on artificial intelligence, PDF data-processing pipeline for X-Ray Diffraction beamlines, spectral analysis application for X-ray absorption fine structure beamlines. The future developments of Daisy are also discussed.

**Primary author:** HU, Yu (IHEP)

**Co-authors:** Dr LIU, Rui (IHEP, CAS); Dr SUN, Haokai (IHEP, CAS); Dr LIU, Jianli (IHEP, CAS); Dr FU, Shiyuan (IHEP, CAS); WANG, Lei (IHEP, CAS); Mr HU, Qingbao (IHEP, CAS); QI, Fazhi (Institute of High Energy Physics, CAS); Dr ZHAO, Haifeng (IHEP, CAS)

**Presenter:** HU, Yu (IHEP)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 61

Type: **Oral Presentation**

## **cryo Electron Microscopy in the European Open Science Cloud**

*Tuesday, March 21, 2023 3:00 PM (30 minutes)*

In this pandemic times our group has coordinated large National and International Consortia to understand, through cryo Electron Microscopy (cryo-EM), both key issues on SARS Cov 2 spike dynamic (Melero et al., IUCrJ. 2020) as well as specific properties of mutations that were prevalent in Europe at certain periods (Ginex et al., PLoS Pathog. 2022), as part of our work in the European Research Infrastructure for Integrative Structural Biology, Instruct-ERIC. These works have been complemented with others in Bioinformatics as developers of one of the few ELIXIR Recommended Interoperability Resources (Macias et al., Bioinformatics. 2021).

Along this hard work we have learnt many lessons and realized many new needs that are now guiding our efforts in cloud computing and in data management. In the following we will briefly review some of these new developments.

The first one is termed “ScipionCloud”, and it is a service registered in the EOSC Marketplace where users from the Instruct Research Infrastructure can deploy a cluster in the cloud to process the data acquired at an Electron Microscopy facility. This cluster has all cryoEM packages and software needed to obtain a 3D structure and is powered by EOSC agreed-on computing resources on the back-end. This means that scientists with minimal computational background (or compute resources of their own) can access the latest tools as well as powerful computational resources to obtain a refined 3D structure to be published and shared with the community. The service quality is assessed through the SQAaaS utility, developed in the same project, that allows to check different quality metrics in software development projects. In addition, the tool permits to evaluate the FAIRness of service outputs that are stored in public repositories. Finally, minimal modifications of the service are needed to deploy a similar cluster in the AWS cloud. Documentation of this service can be found in the Scipion website (<https://scipion.i2pc.es/>) and in arXiv:2211.07738

The second development address the issue of lack of standardization, specially in the area of information (Image) processing, trying to improve the FAIRness of cryoEM workflows. In this way we have developed the tools to export the image processing workflow in Common Workflow Language, using a CryoEM ontology and depositing workflows in WorkflowHub. On this front we have published a cryoEM ontology in the following catalogues: Ontology Lookup Service (<https://www.ebi.ac.uk/ols/ontologies/cryoem>), BioPortal (<https://bioportal.bioontology.org/ontologies/CRYOEM>) and FAIRsharing (10.25504/FAIRsharing.q47I0t). Additionally, we have developed a Scipion plugin to publish processing workflow template in WorkflowHub in the form of a RO-Crate object with Scipion JSON and CWL workflow (enriched with previous ontology) + diagram + metadata.

On a further effort to first map/understand the current situation on raw data deposition in public databases, we first performed an analysis of the pre pandemic situation and the current one in terms of deposition of cryoEM data, and we were surprised to see that, essentially, nothing had been learnt through the pandemic, and that there was a very widespread lack of raw data structural data, with also substantial differences among “Regions” (Europe deposits around 10% of the acquired data, the US 2%, and virtually none from Asia). To act on this situation, an ambitious push towards data sharing is being performed by Instruct-ERIC, coordinated from this laboratory. Still at the pilot stage, we are developing a strategy across the entire infrastructure, starting at the individual facilities which are to archive quality annotated data and share the data in a Federated manner using either Onedata or IRODS (both solutions are under testing). The whole interplay

among Facilities will be orchestrated from Instruct Hub through new extensions of its project management system ARIA.

In short, almost three intense years of work in the context of the current SARS Cov2 pandemic, marked by the need to increase efficiency and be better prepared for the likely events of future pandemics.

**Primary authors:** Prof. SORZANO SANCHEZ, Carlos Oscar (CNB-CSIC, Instruct-ERIC); Ms SANCHEZ, Irene (CNB-CSIC, Instruct-ERIC); Ms DEL CANO, Laura (CNB-CSIC, Instruct-ERIC); CARAZO GARCIA, Jose Maria (National Center for Biotechnology - CNB - CSIC)

**Presenter:** CARAZO GARCIA, Jose Maria (National Center for Biotechnology - CNB - CSIC)

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)

**Track Classification:** Track 2: Health & Life Sciences (including Pandemic Preparedness Applications)

Contribution ID: 62

Type: **Poster Presentation**

## Design and implementation of security policy for HEPS computing platform (Remote presentation)

Based on k8s cluster, HEPS computing platform creates a container computing environment to provide analysis services for users. The computing platform provides a container data analysis environment based on jupyterlib with the jupyterhub web page as the entry point. The platform uses CVMFS to store the software library, and the container environment accesses the CVMFS by CSI. The Lustre is used to store user experiment data, map storage volumes to the container virtualization environment in localhost mode, and provide read/write data access services for users.

HEPS platform uses k8s tool to manage LAN computing resources and create container environment for users to use. WAN users are authenticated by Oauth2.0 to access the LAN container environment for data analysis. The container environment that provides interactive functions for users needs to meet both communication requirements of accessing scale data of WAN and experimental data of LAN. In this service mode, how to effectively limit the activity range of hackers after they invade the container environment and how to quickly locate the container and login users after the SOC system detects the attack behavior, requires a series of security policies to protect the security of HEPS computing platform. In view of the above security problems, this paper introduces the design scheme and application effect of the security policy of HEPS container computing platform from the aspects of k8s network security policy, login behavior audit, network information association analysis, and so on, so as to realize configurable management of the activity scope of user analysis environment and traceability of abnormal container environment, so as to ensure the security of HEPS computing platform.

**Primary authors:** HU, Qingbao (IHEP); Mr XU, Jiping (IHEP); ZENG, SHAN (IHEP); Mr CUI, TAO (IHEP); YAN, Tian (IHEP)

**Presenter:** HU, Qingbao (IHEP)

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations

Contribution ID: 63

Type: **Oral Presentation**

## LHCb Run3 computing model

*Thursday, March 23, 2023 12:00 PM (30 minutes)*

LHCb is one of the four high energy physics experiments at the Large Hadron Collider at CERN, Switzerland. During the second long technical break of the LHC (LS2) that took place from 2018 to 2022, LHCb underwent major upgrades. These upgrades concern not only the actual detector, but also the computing model driving the physics analysis.

The big challenge for the new computing model of Run 3 will be the increased throughput from the upgraded detector by a factor 30, without corresponding jump in the offline computing resources. Full software trigger and selective persistency allow to mitigate this factor, nevertheless we have to scale from 0.65 GB/s (Run2) to 10GB/s (Run3).

We reviewed our data management strategies by favoring LAN transfers over WAN copies, and tailoring our workflows for a faster distribution and eviction from the experiment site. Large cross experiment scale tests were performed during LS2 in order to validate our new approach, and ensure that both our software and the infrastructure can sustain the load.

The way we process data to extract relevant physics figures also had to be considered. This impacts the centralized productions as well as the way physicists perform their analysis on a day to day basis. We make extensive use of the Turbo model, already implemented in Run 2, to reduce the computing and storage needs. The offline reconstruction has been replaced by an online reconstruction, saving a lot of CPU time spent on the grid. Finally, Analysis Productions were introduced in order to leverage the full power of the DIRAC - the WMS and DMS grid middleware used by LHCb - transformation system for the user analysis.

The Monte Carlo simulations largely dominate our CPU needs, and represent about 95% of the total CPU work on the grid. Improvements to the simulation software, as well as the introduction of fast heavily filtered simulations lead to a significantly decreased CPU work per event.

This paper presents the challenges of the upgraded LHCb computing model, the solutions we have implemented to address them, the outcome of our large scale tests, as well as the experience we draw from the 2022 commissioning year.

**Primary author:** HAEN, Christophe (CERN)

**Presenter:** HAEN, Christophe (CERN)

**Session Classification:** Data Management & Big Data

**Track Classification:** Track 6: Data Management & Big Data



Contribution ID: 64

Type: **Oral Presentation**

## HTCondor as batch system in India-CMS Tier-2 Grid

After the retirement of Globus Toolkit, *HTCondor* has been widely adapted in WLCG sites. The reason of the popularity of HTCondor as a batch system could be the maturity of software functionality and strong development and support team, which is getting better day by day. It is open source and fulfils most of the requirements of a user or groups. Integration with other batch systems (i.e. PBS/Torque, LSF, SLURM etc.) and deployment in grid or cloud are couple of features of this batch system.

Tata Institute of Fundamental Research (TIFR), Mumbai, India is hosting one of the biggest Tier-II (T2) grid center in World Wide LHC Computing Grid (WLCG) for Compact Muon Solenoid (CMS) experiment. There are two registered sites for TIFR, 1) Tier-II (T2\_IN\_TIFR) and 2) Tier-III (T3) with the registered name T3\_IN\_TIFRCloud. Where, T3 is able to scaleup the CPU as opportunistic resources including clouds. In this talk, I would present and discuss the overall experience of HTCondor batch system, which has been used as job submission in both grid and cloud, since 2017 at TIFR.

**Primary author:** PATEL, Puneet Kumar (Tata Institute of Fundamental Research, Mumbai, India)

**Presenter:** PATEL, Puneet Kumar (Tata Institute of Fundamental Research, Mumbai, India)

**Session Classification:** Physics and Engineering Applications

**Track Classification:** Track 1: Physics and Engineering Applications

Contribution ID: 65

Type: **Oral Presentation**

## Towards exascale science on Fugaku

*Wednesday, March 22, 2023 9:00 AM (40 minutes)*

Fugaku, one of the first 'exascale' supercomputers of the world, since the beginning production, has been one of the most important R&D infrastructure for Japan, especially in producing groundbreaking results to realize Japan's Society 5.0. Such results have been obtained by the immense power and versatility of Fugaku, allowing complex modern workloads involving not only classical physics simulations, but tight coupling with AI methods and its ability to handle Big Data, based on the standard software ecosystem. Efforts are underway for the successor of Fugaku, FugakuNEXT, to be deployed around 2029, as well as formulation of a hybrid infrastructure encompassing hybrid quantum-HPC computing.

**Presenter:** Prof. MATSUOKA, Satoshi (RIKEN)

**Session Classification:** Keynote Speech II

Contribution ID: 66

Type: **Oral Presentation**

## **Accelerating Science and Learning: The Race for the Digital Restoration of Damaged Historical Material**

*Friday, March 24, 2023 12:20 PM (40 minutes)*

**Presenter:** Prof. SEALES, William B (University of Kentucky)

**Session Classification:** Closing Keynote & Ceremony

Contribution ID: 67

Type: **Oral Presentation**

## **Open Scholarly Communication for Social Sciences and Humanities: supporting the science and maximising its impact for society**

*Tuesday, March 21, 2023 9:10 AM (40 minutes)*

OPERAS is the Research Infrastructure for open scholarly communication in the field of Social Sciences and Humanities. This keynote will focus mainly on introducing 3 different acceleration mechanisms: 1. Discovery Acceleration by presenting the GoTriple platform, 2. Dissemination Acceleration of publications by presenting the DIAMOND action, 3. Quality Acceleration by presenting the OAeBu project facilitating trusted data exchange between publishers. These three types of services are contributing the global strategy of OPERAS Research Infrastructure in order to improve science and support innovation.

**Presenter:** LEGRE, Yannick (OPERAS)

**Session Classification:** Opening Ceremony & Keynote Speech I

Contribution ID: 68

Type: **Oral Presentation**

## The History of the Accept and Rise of Geant4

*Thursday, March 23, 2023 9:00 AM (40 minutes)*

Geant4 is an open-source software toolkit that has been in development since 1994 and is used to simulate the interactions between particles and matter. It is widely used in a variety of fields, including high energy physics, nuclear physics, accelerator physics, medical physics, and space science. The first paper on Geant4, published in Nuclear Instruments and Methods in Physics Research A in 2003, has received more than 16,000 citations according to the SCOPUS database and is the second most frequently cited paper in fundamental physics.

Geant4 is the first large-scale software designed and developed using object-oriented technology in particle physics. Its design is based on an earlier effort in Japan called ProDigi. The Japanese team introduced object-oriented analysis, design, and development to Geant4, which was a new technology at the time. Without these technologies, likely, Geant4 would not have been accepted as a detector simulation toolkit in particle physics or achieved success in other fields. In this presentation, the speaker will discuss the unknown history of Geant4 and how its adoption of object-oriented technology contributed to its success.

In addition to discussing the technical aspects of Geant4's development, the speaker will also delve into the cultural aspects of the collaboration. Geant4 has more than 100 collaborators from different parts of the world, which has greatly enriched the project and led to various challenges. The speaker will discuss how these cultural differences have affected the development process and how they have been addressed.

**Presenter:** Dr SASAKI , Takashi (KEK)

**Session Classification:** Keynote Speech III

Contribution ID: 69

Type: **Oral Presentation**

## **eScience Activities in Singapore (Remote presentation)**

*Tuesday, March 21, 2023 11:20 AM (20 minutes)*

**Presenter:** Prof. TAN, Tin Wee (NSCC)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 70

Type: **Oral Presentation**

## **eScience Activities in Australia (Remote presentation)**

*Wednesday, March 22, 2023 11:40 AM (20 minutes)*

**Presenter:** Dr WALSH, Carmel (ARDC)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 71

Type: **Oral Presentation**

## **eScience Activities in Thailand (Remote Presentation)**

*Wednesday, March 22, 2023 12:00 PM (20 minutes)*

**Presenter:** Dr VORAKULPIPAT, Chalee (NECTEC)

**Session Classification:** e-Science Activities in Asia Pacific



Contribution ID: 72

Type: **Oral Presentation**

## **eScience Activities in Japan**

*Wednesday, March 22, 2023 11:00 AM (20 minutes)*

**Presenter:** AIDA, Kento (National Institute of Informatics)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 73

Type: **Oral Presentation**

## **eScience Activities in China (Remote presentation)**

**Presenter:** CHEN, Gang (Institute Of High Energy Physics)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 74

Type: **Oral Presentation**

## **eScience Activities in Taiwan**

*Wednesday, March 22, 2023 11:20 AM (20 minutes)*

**Presenters:** YEN, Eric (ASGC); Mr LEE, Felix (ASGC); Ms YOU, Jingya (ASGC)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 75

Type: **Oral Presentation**

## **eScience Activities in Korea**

*Tuesday, March 21, 2023 11:00 AM (20 minutes)*

**Presenter:** AHN, Sang Un (Korea Institute of Science and Technology Information)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 76

Type: **Oral Presentation**

## **eScience Activities in Indonesia (Remote Presentation)**

*Tuesday, March 21, 2023 12:00 PM (20 minutes)*

**Presenter:** Dr SUHARDIMAN, Basuki (ITB)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 77

Type: **Oral Presentation**

## Different Facets of Distance Geometry

*Monday, March 20, 2023 4:00 PM (30 minutes)*

The Distance Geometry Problem (DGP) asks whether a simple weighted undirected graph  $G=(V,E,d)$  can be realized in the  $K$ -dimensional Euclidean space so that the distance constraints implied by the weights on the graph edges are satisfied. This problem was proven to be NP-hard in the context of graph embeddability, and has several applications. In this talk, we will focus on various currently ongoing works in this very rich research context: (1) We will talk about a particular class of DGP instances where, under particular assumptions, it is possible to represent the search space as a binary tree, and where in ideal situations, vertex positions can be assigned to each of its nodes. In real-life applications, however, the distances are generally provided with low precision, and they are actually likely to carry measurement errors, so that local continuous search spaces are actually assigned to each tree node. This gives rise to special combinatorial problems which are locally continuous. (2) We will review the main applications of the DGP, ranging from structural biology, passing through sensor network localization and adaptive maps, until the dynamical component is included in DGP instances for computer graphics applications. (3) Finally, we present an alternative computing approach for the solution of the DGPs in dimension 1, where an analog optical processor is employed for the computations, which is based on some properties of laser light beams.

**Presenter:** Prof. MUCHERINO, Antonio (Institut de Recherche en Informatique et Systèmes Aléatoires, University of Rennes 1, France)

**Session Classification:** Workshop on Novel Computational Methods for Structural Biology

Contribution ID: 78

Type: **Oral Presentation**

## A Curvilinear-Path Umbrella Sampling Approach to Characterizing Thermodynamics and Mechanisms of Biomolecular Interactions

*Monday, March 20, 2023 4:30 PM (30 minutes)*

Protein-protein and protein-ligand interactions are central in biological mechanisms. These interactions can be classified into thermodynamics and mechanistic pathways. Estimating accurate and reliable interaction energetics along the thermodynamic pathway is one of the ongoing challenges in computational biophysics. Umbrella sampling simulation-based potential of mean force calculations is one of the methods to estimate the interaction energetics. Previously this method was implemented by first choosing a predefined path of dissociation, which is often chosen as a straight-line/vectorial path. However, there are several unresolved issues such as choices of predefined direction, corrections of potential of mean force to standard free energy of binding, etc. To unleash these limitations, we developed a curvilinear-path umbrella sampling molecular dynamics (MD) simulation approach to address some of the issues. We have applied the new method for evaluating the standard free energy of binding for the barnase-barstar protein-protein system and then on a protein-ligand system, where the interaction energetics of FKBP12-rapamycin protein-ligand system is estimated. The computed energetics for both systems are in good agreement with the experimental values. The revealed mechanistic insight for the protein-protein complex matches very-well with the computationally expensive adaptive biasing MD based brute-force methods. Further, we also conducted the simulations of dissociation reactions of ternary complex FKBP12-rapalog-FRB, which indeed demonstrated a tug-of-war between FRBP12 and FRB to bind with the rapamycin, and revealed that the rapamycin prefers to bind with FKBP12 more than FRB. Thus, the glue-like molecule rapamycin and other rapalogs seem to follow a step-wise path of forming FKBP12-rapalog complex first and then the ternary complex with FRB. Thus, the developed curvilinear-path approach offers accurate and reliable binding energetic, is sensitive enough to distinguish the change in interaction energetics upon mutations, and can reliably reveal mechanistic details towards the fulfillment of the characterization.

**Presenter:** Dr JOSHI, Dhananjay (Research Center for Applied Science, Academia Sinica)

**Session Classification:** Workshop on Novel Computational Methods for Structural Biology

Contribution ID: 79

Type: **Oral Presentation**

## Quantum Machine Learning for Structure-Based Virtual Screening of the Entire Medicinal Chemical Space

*Monday, March 20, 2023 5:00 PM (30 minutes)*

It has been estimated based on the graph theory that there are at least 1060 organic molecules that are relevant for small-molecule drug discovery. Using machine learning to estimate the binding free energies for screening of large chemical libraries to search for the tightly binding inhibitors would take a considerable amount of computational resources, yet it is not possible to explore the entire biologically relevant chemical space. Quantum computing provides a unique opportunity to accomplish such a computational task in the near future. Here, we demonstrate how to use 512 occupancies to describe the structures of protein-ligand complexes, how to convert the classical occupancies to the quantum states using nine qubits, and to estimate the binding free energies ( $\Delta G_{\text{bind}}$ ) of the complexes using quantum machine learning. We showed that it is possible to use only 450 parameters to prepare the quantum states for describing the structure of one protein-ligand complex. In this work the entire 2020 PDBbind dataset was adopted as the training set, and we used 45 parameters as the first attempt to construct the model for predicting the binding free energies ( $\Delta G_{\text{bind}}$ ). The Pearson correlation coefficient (PCC) between the estimated binding free energies and the corresponding experimental values are 0.49. By slightly increasing to 1,440 parameters for constructing the neural network model for the prediction of the  $\Delta G_{\text{bind}}$ , the PCC is improved to be 0.78, which is even slightly better than to the results achieved by recent classical convolutional neural network models using more than millions of parameters. In this work, for the first time, we demonstrated the feasibility of using quantum computers to explore the entire medicinal chemical space with a concrete, implementable approach.

**Presenter:** Prof. LIN, Jung-Hsin (Research Center for Applied Science, Academia Sinica)

**Session Classification:** Workshop on Novel Computational Methods for Structural Biology



Contribution ID: 80

Type: **Oral Presentation**

## Coffea-casa analysis facility

*Thursday, March 23, 2023 2:30 PM (30 minutes)*

The Coffea-casa analysis facility prototype provides physicists with alternative mechanisms to access computing resources and explore new programming paradigms. Instead of the traditional command-line interface and asynchronous batch access, a notebook-based web interface and interactive large-scale computing is provided. The facility commissions an environment for end-users enabling execution of increasingly-complex analyses, as demonstrated by the Analysis Grand Challenge (AGC) and other examples.

In this contribution, we describe the Coffea-casa facility design and the key features focusing on its modularity and portability. This facility is designed to be Kubernetes-native, allowing to adopt an approach to transform existing facilities (e.g., LHC Tier-2 sites) into highly composable systems. Targeting more generic configurations makes the facility modular and easily expandable and re-deployable on other sites.

**Presenter:** SHADURA, Oksana (University Nebraska-Lincoln (US))

**Session Classification:** VRE

Contribution ID: 82

Type: **Oral Presentation**

## **eScience Activities in Mongolia**

*Tuesday, March 21, 2023 11:40 AM (20 minutes)*

**Presenter:** Dr OTGONSUVD , Badrakh (Mongolian Academy of Sciences)

**Session Classification:** e-Science Activities in Asia Pacific

Contribution ID: 83

Type: **Oral Presentation**

## Goal Setting: Opening Remarks

*Tuesday, March 21, 2023 2:00 PM (15 minutes)*

Tosh discusses the educational paradigm for New Education Normal. Topics such as “What are the future skills?” and “Ambiance for Authentic Learning: How to implement such skills in authentic learning?” are elaborated.

**Presenter:** YAMAMOTO, Tosh (Kansai University)

**Session Classification:** Education Informatics Workshop

Contribution ID: 84

Type: **Oral Presentation**

## Showcases at different Educational Tiers

*Tuesday, March 21, 2023 2:15 PM (15 minutes)*

This workshop aims at identifying the essential issues involved in the fundamental components of authentic education, especially in the realm of authentic learning, in the New Education Normal and then demonstrates some experimental showcases at various tiers.

In this part of the session, innovative new educational practices are showcased. We intend to share some successful educational experiences with participants and offer triggers to have them devise new and authentic learning for the future generation.

From various tiers of education, renounced educators share their experiences in teaching and learning to make us rethink education in the New Education Normal era.

**Presenter:** YAMAMOTO, Tosh (Kansai University)

**Session Classification:** Education Informatics Workshop

Contribution ID: 85

Type: **Oral Presentation**

## **City Auncel—Analyzing Learners' Multiple Representations Literacy in the Socio-scientific Issue Inquiry Game Based on GIS Information**

*Tuesday, March 21, 2023 2:30 PM (20 minutes)*

The Scenario-Issue-Resolution (SIR) instructional model is introduced to nurture students' abilities to tackle complex problems that are grounded on scenario-based issues. It teaches to intrigue students' foresight. SIR is oriented from issue-based inquiry and grounded on the socioscientific issues (SSI) that lie in resolving open and ill-structured world problems that are controversial with conflicts between groups with different perspectives.

**Presenter:** Dr SHIH, Juling (National Central University)

**Session Classification:** Education Informatics Workshop

Contribution ID: 86

Type: **Oral Presentation**

## **Incorporating Regional Social Aspects and Gamification in STEAM education.**

*Tuesday, March 21, 2023 2:50 PM (20 minutes)*

Dr. Takemata's team developed a STEAM curriculum to nurture K-12 students' computational thinking skills in the surrounding living environment. With the concepts of SDGs, the purpose of the learning is to think seriously about the future of the society where they will live as full-fledged members of the society. The proposed hands-on and heads-on workshop enhances active learning in PBL to motivate students' curiosity in their surrounding social environment.

**Presenter:** Dr TAKEMATA, Kazuya (Kanazawa Inst. of Tech.)

**Session Classification:** Education Informatics Workshop

Contribution ID: 87

Type: **Oral Presentation**

## **Academic Writing COIL, Tourism, Essay Writing, Press Release Writing**

*Tuesday, March 21, 2023 3:10 PM (20 minutes)*

English Writing classes were conducted between Kansai University and Chihlee University for several semesters in the virtual classroom, where all students worked in teams to interact and acquire writing skills in English. Based on their interests and curiosities, the students explore their societies and compare their cultural values and lifestyles. Through writing activities, they acquired how to organize their thoughts in the form of infographics and mind mapping and to write in paragraphs. They also acquired skills to express themselves in rich media such as blog writing, infographic presentation, and pitch videos. Throughout the course, all students as well as all teams were on the same page regarding learning in the virtual classroom.

**Presenters:** Prof. CHEN , Ru-Shan; Prof. WANG, Yi-Chien

**Session Classification:** Education Informatics Workshop

Contribution ID: 88

Type: **Oral Presentation**

## **Social Entrepreneurship with Global Collaborative Learning**

*Tuesday, March 21, 2023 4:00 PM (20 minutes)*

Collaborative Online International Learning (COIL) courses have been collaboratively conducted between the Department of Business and Management at Nanyang Polytechnic in Singapore and Kansai University/Kansai University of International Studies in Japan. Students learning and interaction are all conducted in the virtual classroom asynchronously. The realm of learning is to acquire business skills to build a startup company with the mindset of SDGs and the future of society and the world. The students build teams of the same interests and learn to be ready for their startup companies through simulation learning.

**Presenters:** Prof. ONG, Benson (Nanyang Polytechnic); Prof. PANG, Chris (Nanyang Polytechnic)

**Session Classification:** Education Informatics Workshop



Contribution ID: 89

Type: **Oral Presentation**

## **Integration of E-portfolio into General Education Classroom and Automate Classification Model for E-portfolio**

*Tuesday, March 21, 2023 4:20 PM (20 minutes)*

With the application of the model of design thinking, Dr. Nakazawa incorporates features of e-portfolios into reflective learning for authentic assessment. The self-awareness of the values from learning and the process of meta-cognitive learning activities are the keys to authentic assessment in New Education Normal.

**Presenter:** NAKAZAWA, Minoru (Kanazawa Institute of Technology)

**Session Classification:** Education Informatics Workshop

Contribution ID: 90

Type: **Oral Presentation**

## Trust-Building and Negotiation Practicum

*Tuesday, March 21, 2023 4:40 PM (20 minutes)*

Before the Pandemic, university students and corporate staff of the HR departments from IBM, ANA, and Fuji Xerox, among others, gathered in Tokyo once a year to have a communication and negotiation workshop to enhance their negotiation skills. The purpose is to develop and enrich their human resource skills through the interaction of multi-tiered age groups. During the Pandemic, such activities were interrupted. As we approach the end of the Pandemic, organizers have started an innovative workshop to maintain the quality on a larger scale than before.

**Presenters:** Prof. TAGAMI , Masanori; YAMAMOTO, Tosh (Kansai University)

**Session Classification:** Education Informatics Workshop

Contribution ID: 91

Type: **Oral Presentation**

## **Further Discussion and Conclusion**

*Tuesday, March 21, 2023 5:00 PM (30 minutes)*

**Presenter:** YAMAMOTO, Tosh (Kansai University)

**Session Classification:** Education Informatics Workshop

Contribution ID: 92

Type: **Oral Presentation**

## **General introduction to docking, integrative modelling and HADDOCK**

*Monday, March 20, 2023 9:00 AM (1h 30m)*

**Presenter:** BONVIN, Alexandre M.J.J. (Utrecht University)

**Session Classification:** Integrative modelling with HADDOCK - Introducing the new modular HADDOCK3 version

Contribution ID: 93

Type: **Oral Presentation**

## **Computer practical: Antibody-antigen docking with the HADDOCK server**

*Monday, March 20, 2023 11:00 AM (1h 30m)*

HADDOCK2.4 antibody-antigen docking tutorial: This tutorial demonstrates the use of HADDOCK2.4 for predicting the structure of an antibody-antigen complex using information about the hypervariable loops of the antibody and either the entire surface of the antigen or a loose definition of the epitope. This tutorial does not require any Linux expertise and only makes use of our web servers and PyMol for visualisation/analysis.

<https://www.bonvinlab.org/education/HADDOCK24/HADDOCK24-antibody-antigen/>

**Presenter:** BONVIN, Alexandre M.J.J. (Utrecht University)

**Session Classification:** Integrative modelling with HADDOCK - Introducing the new modular HADDOCK3 version

Contribution ID: 94

Type: **Oral Presentation**

## Lecture & practical: Introduction to HADDOCK3 and practical

*Monday, March 20, 2023 2:00 PM (1h 30m)*

HADDOCK3 antibody-antigen docking: This tutorial demonstrates the use of HADDOCK3 for predicting the structure of an antibody-antigen complex using information about the hypervariable loops of the antibody and either the entire surface of the antigen or a loose definition of the epitope. It illustrates the modularity of HADDOCK3 by introducing a new workflow not possible under the current HADDOCK2.X versions. As HADDOCK3 only exists as a command line version, this tutorial does require some basic Linux expertise.

<https://www.bonvinlab.org/education/HADDOCK3/HADDOCK3-antibody-antigen/>

**Presenter:** BONVIN, Alexandre M.J.J. (Utrecht University)

**Session Classification:** Integrative modelling with HADDOCK - Introducing the new modular HADDOCK3 version

Contribution ID: 95

Type: **Oral Presentation**

## **Intro about the EGI and the EGI infrastructure**

*Tuesday, March 21, 2023 2:00 PM (20 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific

Contribution ID: **96**

Type: **Oral Presentation**

## **EGI VO for AP**

*Tuesday, March 21, 2023 2:20 PM (30 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific



Contribution ID: 97

Type: **Oral Presentation**

## **DEMO - How to use the EGI Cloud**

*Tuesday, March 21, 2023 2:50 PM (20 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific

Contribution ID: 98

Type: **Oral Presentation**

## **DEMO - Getting started with the EGI Notebooks**

*Tuesday, March 21, 2023 3:10 PM (20 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific

Contribution ID: 99

Type: **Oral Presentation**

## **Hands-on with the EGI Notebooks and Replay services**

*Tuesday, March 21, 2023 4:00 PM (30 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific

Contribution ID: **100**

Type: **Oral Presentation**

## **Approach to reproducible data science with EGI and EOSC**

*Tuesday, March 21, 2023 4:30 PM (15 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific

Contribution ID: **101**

Type: **Oral Presentation**

## **Overview of the hands-on - Full data lifecycle**

*Tuesday, March 21, 2023 4:45 PM (45 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** EGI Tutorial: Regional infrastructure for reproducible open science in Asia Pacific

Contribution ID: **102**

Type: **Oral Presentation**

## **Robotics and Artificial Intelligence for Pollen Monitoring (Remote presentation)**

*Wednesday, March 22, 2023 2:00 PM (30 minutes)*

**Presenter:** Dr BUTERS , Jeroen (ZAUM - HMGU/TUM, Munich)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 103

Type: **Oral Presentation**

## **Is an NWP-Based Nowcasting System Suitable for Aviation Operations? (Remote presentation)**

*Wednesday, March 22, 2023 2:30 PM (40 minutes)*

**Presenter:** Dr PARODI , Antonio (CIMA, Savona)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **104**

Type: **Oral Presentation**

## **Large-scale flow patterns and their relation to summer lightning in Europe (Remote presentation)**

*Wednesday, March 22, 2023 3:10 PM (20 minutes)*

**Presenter:** Dr GHASEMIFARD , Homa (ESSL, Berlin)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop



Contribution ID: **105**

Type: **Oral Presentation**

## **Wine in the Cloud – Smart Viticulture (Remote presentation)**

*Wednesday, March 22, 2023 4:00 PM (30 minutes)*

**Presenter:** Dr HARSH , Piyush (TERRAVIEW, Winterthur)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **106**

Type: **Oral Presentation**

## **tbd – slots for energy efficiency**

**Presenter:** TBD

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **107**

Type: **Oral Presentation**

## **Environmental Computing and Energy Efficiency at Leibniz Supercomputing Centre**

*Thursday, March 23, 2023 1:30 PM (45 minutes)*

**Presenter:** KRANZLMULLER, Dieter (LMU Munich)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **108**

Type: **Oral Presentation**

## **tbd – slots for energy efficiency**

**Presenter:** TBD

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **109**

Type: **Oral Presentation**

## **Risk Management**

*Monday, March 20, 2023 9:00 AM (1h 30m)*

Risk Management is the tool to organise your limited resources to provide efficient operational security services to your organisation.

In this session we will first give an introduction to Risk Management, and discuss different methods which could be applied. After discussing the concept of Risk Management we will then dive into its components, define risks for an example organisation and do the risk assessment for some of the identified risks. In the hands-on section we will look into the usage of MITRE ATT&CK for threat modeling.

**Presenter:** GABRIEL, Sven (Nikhef/EGI)

**Session Classification:** Security Workshop

Contribution ID: **110**Type: **Oral Presentation**

## **SSC-2023-03**

*Monday, March 20, 2023 11:00 AM (1h 30m)*

The Security Service Challenge SSC-2023-03 is looking into the CMS's to use the distributed compute infrastructure. In particular we look into the joining elements of the multiple security teams involved. This particular exercise goes well beyond EGIs logical borders, which over time got more fuzzy with the advent of new technologies, and additional resources available to our user community. In this session we will demonstrate the SSC framework developed by us and show first results.

**Presenters:** ROORDA, Jouke (Nikhef); GABRIEL, Sven (Nikhef/EGI)

**Session Classification:** Security Workshop

Contribution ID: 111

Type: **Oral Presentation**

## **Threat Intelligence and Security Operations Centres (Remote presentation)**

*Monday, March 20, 2023 4:00 PM (1h 30m)*

In the current research and education environment, the threat from cybersecurity attack is acute having grown in recent years. We must collaborate as a community to defend and protect ourselves. This requires both the use of detailed, timely and accurate threat intelligence alongside fine-grained monitoring.

In this session we explore aspects both of sharing appropriate intelligence and the conceptual design of a security operations centre.

**Presenter:** CROOKS, David (UKRI STFC)

**Session Classification:** Security Workshop

Contribution ID: **112**

Type: **not specified**

## **Discussion**

**Session Classification:** Security Workshop



Contribution ID: 113

Type: **Oral Presentation**

## **Effects of transport on a biomass burning plume from Indochina during EMeRGe-Asia identified by WRF-Chem**

*Friday, March 24, 2023 9:00 AM (45 minutes)*

**Presenter:** Dr LIN , Chuan-Yao (RCEC, Academia Sinica)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 114

Type: **Oral Presentation**

## **Upper ASEAN Wildland Fire Special Research Unit (WFSRU) Activities Updates**

*Thursday, March 23, 2023 11:45 AM (25 minutes)*

**Presenter:** Dr TANPIPAT, Veerachai (HII)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 115

Type: **Oral Presentation**

## **Introduction to Sentinel Asia and its Data Systems for Disaster Response**

*Thursday, March 23, 2023 12:10 PM (20 minutes)*

**Presenter:** Dr TAKEI, Goro (JAXA)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 116

Type: **Oral Presentation**

## **Increasing Heavy Rainfall and Flood Events in Malaysia**

*Thursday, March 23, 2023 11:20 AM (25 minutes)*

**Presenter:** Prof. LIEW , Ju Neng (UKM)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 117

Type: **Oral Presentation**

## Asia Regional Collaborations

*Thursday, March 23, 2023 11:00 AM (20 minutes)*

**Presenter:** YEN, Eric (ASGC)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 118

Type: **Oral Presentation**

## **Impact of earth observation satellite data on local DRR activities (Remote presentation)**

*Friday, March 24, 2023 11:10 AM (25 minutes)*

**Presenter:** Ms TETANGCO , Jelina Tanya H. (ASTI)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **119**

Type: **Oral Presentation**

## **EGI Notebook and Replay Services**

*Friday, March 24, 2023 11:35 AM (35 minutes)*

**Presenter:** LA ROCCA, Giuseppe (EGI Foundation)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: **120**

Type: **Oral Presentation**

## Wrap Up

*Friday, March 24, 2023 12:10 PM (10 minutes)*

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop



Contribution ID: **121**

Type: **Oral Presentation**

## Opening Remarks

*Tuesday, March 21, 2023 9:00 AM (10 minutes)*

**Presenters:** MATYSKA, Ludek (CESNET); CHANG, Yuan-Hann (Institute of Physics, Academia Sinica)

**Session Classification:** Opening Ceremony & Keynote Speech I

Contribution ID: 122

Type: **Oral Presentation**

## **Generative AI: How it changes our lives?**

*Wednesday, March 22, 2023 9:40 AM (40 minutes)*

**Presenter:** Dr WANG, Yu-Chiang Frank (NVIDIA)

**Session Classification:** Keynote Speech II

Contribution ID: 123

Type: **Oral Presentation**

## Resource Balancing in AI/HPC Accelerated Solutions

*Tuesday, March 21, 2023 9:50 AM (40 minutes)*

Accelerating time to market, time to service and time to science through computing, the challenge we encountered today may not be in the speed of data processing. To streamline huge amount of data without blocking and queuing in an AI/HPC infrastructure design is crucial.

In this speech, a case study about resource balancing in an AI driven inspection and decision systems from a world top notch smart manufacturer will be presented and will be showcased the design of the infrastructure with scale up/out upgrade path as the demands grow. The similar concept can also be employed in the problem solving for the other fields of AI/HPC applications like high-energy physics, nuclear and astrophysics.

To strengthen the weakest link of the chain in this particular case, the team examined and analyzed the problem from the manufacturing processes, the system with proper balancing infrastructure design between computing capacity and the performance of tier0 storage has been delivered with plug and play readiness in rack scale.

Last not least, a clear and present challenge and solutions in AI/HPC systems will also be quickly addressed at the end part of the session.

**Presenter:** Dr LYNN, Andrew (Super Micro)

**Session Classification:** Opening Ceremony & Keynote Speech I

Contribution ID: 124

Type: **Oral Presentation**

## Introduction

*Tuesday, March 21, 2023 11:00 AM (5 minutes)*

**Presenter:** YEN, Eric (ASGC)

**Session Classification:** APGridPMA Meeting

Contribution ID: 125

Type: **Oral Presentation**

## **Update of TAGPMA and Token Transition (Remote presentation)**

*Tuesday, March 21, 2023 11:05 AM (30 minutes)*

**Presenter:** SIMMEL, Derek (Pittsburgh Supercomputing Center)

**Session Classification:** APGridPMA Meeting

Contribution ID: 126

Type: **Oral Presentation**

## **IGTF and AAI for Academics**

*Tuesday, March 21, 2023 11:35 AM (30 minutes)*

**Presenter:** GROEP, David (Nikhef and Maastricht University)

**Session Classification:** APGridPMA Meeting

Contribution ID: 127

Type: **Oral Presentation**

## **AAI Development in HPCI and Gakunin**

*Tuesday, March 21, 2023 2:00 PM (25 minutes)*

**Presenter:** SAKANE, Eisaku (National Institute of Informatics)

**Session Classification:** APGridPMA Meeting

Contribution ID: 128

Type: **Oral Presentation**

## **Self-Audit and Status Report of All CAs**

*Tuesday, March 21, 2023 2:25 PM (35 minutes)*

**Session Classification:** APGridPMA Meeting



Contribution ID: 129

Type: **Oral Presentation**

## **Future Meetings & Discussion**

*Tuesday, March 21, 2023 3:00 PM (30 minutes)*

**Session Classification:** APGridPMA Meeting

Contribution ID: **130**

Type: **Oral Presentation**

**TBC**

*Tuesday, March 21, 2023 5:00 PM (20 minutes)*

**Presenter:** HONORATO, Rodrigo (Utrecht University)

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)

Contribution ID: 131

Type: **Oral Presentation**

## **IT in Citizen Science: A Case Study in Germany (Remote presentation)**

*Wednesday, March 22, 2023 4:30 PM (30 minutes)*

**Presenter:** Dr BATSAIKHAN , Anudari (LRZ)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 132

Type: **Oral Presentation**

## Wrap Up

*Wednesday, March 22, 2023 5:00 PM (30 minutes)*

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 133

Type: **Oral Presentation**

## **Simulation of Geothermal Heating/Cooling Potential within a City**

*Thursday, March 23, 2023 2:15 PM (30 minutes)*

**Presenter:** PAUW, Viktoria (Leibniz Rechenzentrum)

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 134

Type: **Oral Presentation**

## Wrap Up

*Thursday, March 23, 2023 2:45 PM (15 minutes)*

**Session Classification:** Joint DMCC, UMD & Environmental Computing Workshop

Contribution ID: 135

Type: **Oral Presentation**

## **Introduction**

*Wednesday, March 22, 2023 2:00 PM (10 minutes)*

**Presenter:** WADENSTEIN, Mattias (NeIC)

**Session Classification:** GDB Meeting

Contribution ID: 136

Type: **Oral Presentation**

## **ASGC site update**

*Wednesday, March 22, 2023 2:10 PM (20 minutes)*

**Presenter:** YEN, Eric (ASGC)

**Session Classification:** GDB Meeting



Contribution ID: 137

Type: **Oral Presentation**

## **Experiences from ICEPP migration from DPM to dCache (Remote presentation)**

*Wednesday, March 22, 2023 2:30 PM (20 minutes)*

**Presenter:** SAITO, Masahiko (ICEPP, The University of Tokyo)

**Session Classification:** GDB Meeting

Contribution ID: 138

Type: **Oral Presentation**

## **IHEP migration from DPM to EOS, status update (Remote presentation)**

*Wednesday, March 22, 2023 2:50 PM (20 minutes)*

**Presenter:** YAN, Xiaofei (IHEP)

**Session Classification:** GDB Meeting

Contribution ID: 139

Type: **Oral Presentation**

## **LHC networking in Asia**

*Wednesday, March 22, 2023 3:10 PM (20 minutes)*

**Presenter:** AHN, Sang Un (Korea Institute of Science and Technology Information)

**Session Classification:** GDB Meeting

Contribution ID: 140

Type: **Oral Presentation**

## **Belle II computing update (Remote presentation)**

*Wednesday, March 22, 2023 4:00 PM (30 minutes)*

**Presenter:** VILANUEVA, Michel Hernandez

**Session Classification:** GDB Meeting

Contribution ID: 141

Type: **Oral Presentation**

## **IAM Token Hackathon summary (Remote presentation)**

*Wednesday, March 22, 2023 4:30 PM (30 minutes)*

**Presenter:** DACK, Thomas (STFC - UKRI)

**Session Classification:** GDB Meeting

Contribution ID: 142

Type: **Oral Presentation**

## **WLCG token migration update (Remote presentation)**

*Wednesday, March 22, 2023 5:00 PM (30 minutes)*

**Presenter:** LITMAATH, Maarten

**Session Classification:** GDB Meeting

Contribution ID: 143

Type: **Oral Presentation**

## **Identity Management and Federation in Asia (Remote presentation)**

*Tuesday, March 21, 2023 12:05 PM (25 minutes)*

**Presenter:** SMITH, Terry (AAF)

**Session Classification:** APGridPMA Meeting

Contribution ID: 144

Type: **Oral Presentation**

## **Windows Forensics Essentials**

*Monday, March 20, 2023 2:00 PM (1h 30m)*

**Session Classification:** Security Workshop



Contribution ID: 145

Type: **Oral Presentation**

## **Legal issues concerning genomic and genetic data in Laniakea**

*Tuesday, March 21, 2023 4:20 PM (20 minutes)*

**Presenter:** FOGGETTI, Nadina

**Session Classification:** Health & Life Sciences (including Pandemic Preparedness Applications)