Contribution ID: **120**   Type: **Presentations**

e-Infrastructure for the Multi-Scale Complex Genomics Virtual Research Environment

Thursday, 30 November 2017 14:00 (15 minutes)

3D/4D genomics is one of the next great challenges for biology and biomedicine. While major milestones have been achieved in sequencing, imaging and computation, still understanding of the 3D folding of the chromatin fiber, its role in fundamental cellular processes, and connection with pathology remains a huge challenge. Genomics projects, together with astrophysics, are among the major generators of Big Data, thus being in need for the kind of solutions developed by the MuG VRE. The particularity in managing 3D/4D genomics data lies in the diversity of data formats generated and analysis methods due to the continued advent of new experimental techniques as well as the multi-resolution problem involved in integrated navigation in data that range from sequence to 3D/4D chromatin dynamics. The successful implementation and uptake of MuG VRE solutions is sure to serve as an example for other research communities that may face a high multidisciplinary component and that need to handle very diverse data.

Multiscale Genomics (MuG) Virtual Research Environment (VRE) is developing a cloud-based computational infrastructure to support the deployment of software tools addressing the various levels of analysis of the genome. Integrated tools tackle needs that range from high computationally demanding applications (e.g. molecular dynamics simulations) to the analysis of NGS or Hi-C data, where stress is on data management and high throughput data analysis. The development of such infrastructure includes the building of unified data management procedures, and distributed execution to minimize data transmission and ensure sustainability. The present MuG Infrastructure is based two main cloud systems (Institute for research in Biomedicine, IRB and Barcelona Supercomputing Center BSC), with a satellite installation at EBI's Embassy cloud. The infrastructure is based in openNebula and openStack cloud management systems, and has developed specific interfaces for users and developers. Interoperability of the tools included in the infrastructure is maintained through a rich set of metadata for both tools and data, that allow the system to associated tools and data in a transparent manner. Two alternatives for execution scheduling are provided, a traditional queueing system to handle demand peaks in applications of fixed needs, and an elastic and multi-scale programming model (PyCOMPs, controlled by the PMES scheduler), for complex workflows requiring distributed or multi-scale executions schemes.

The first release of the infrastructure will be presented in November 2017 to the 3D/4D research community.

Topic Area

The EOSC & EDI building blocks

Type of abstract

Presentation (15 minutes)

Primary author: Dr GELPI, Josep Ll. (Barcelona Supercomputing Center (BSC), Barcelona, Spain. Dept. of Biochemistry and Molecular Biomedicine, University of Barcelona, Barcelona, Spain)

Co-authors: Dr HOSPITAL, Adam (Institute for Research in Biomedicine (IRB) Barcelona, Barcelona. Spain,); Dr YATES, Andy (European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK.); Dr LAUGHTON, Charles A (School of Pharmacy and Centre for Biomolecular Sciences, Nottingham, UK); Ms CONEJERO, F. Javier (Barcelona Supercomputing Center (BSC), Barcelona, Spain,); Mr BAYARRI, Genis (Institute for Research in Biomedicine (IRB) Barcelona, Barcelona. Spain,); Dr CAVALLI, Giaccomo (Institute of Human Genetics, UMR9002 CNRS, University of Montpellier, France); Ms CODÓ, Laia (Barcelona Supercomputing Center (BSC), Barcelona, Spain); Dr MARTI-RENOM, Marc (Structural Genomics Group, CNAG-CRG, The Barcelona Institute of Science and Technology (BIST), Spain,); Dr PASI, Marco (School of Pharmacy and Centre for Biomolecular Sciences,

Nottingham, UK); Dr MCDOWALL, Mark (European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, UK.); Prof. OROZCO, Modesto (Institute for Research in Biomedicine (IRB) Barcelona, Barcelona. Spain & Dept. of Biochemistry and Molecular Biomedicine, University of Barcelona, Barcelona, Spain); Dr BADIA, Rosa M (Barcelona Supercomputing Center (BSC), Barcelona, Spain)

Presenter: Dr GELPI, Josep Ll. (Barcelona Supercomputing Center (BSC), Barcelona, Spain. Dept. of Biochemistry and Molecular Biomedicine, University of Barcelona, Barcelona, Spain)

Session Classification: EOSC building block presentations