



Contribution ID: 78

Type: **Presentazione orale**

IRCSS Sant'Orsola Computational Genomic platform

Tuesday, 21 May 2024 17:15 (15 minutes)

Il primo sequenziamento del genoma umano ha richiesto 32 anni, ma grazie alle moderne tecnologie di Next-Generation Sequencing di DNA e RNA siamo in grado di sequenziare 40 genomi in un paio di giorni, producendo TB di dati testuali, più di 100 GB per singolo genoma. Ciò pone una grossa sfida per le infrastrutture di calcolo, le quali devono essere in grado di gestire queste grandi quantità di dati e processarli attraverso pipeline efficienti e che utilizzino in modo efficace risorse di calcolo eterogenee (come CPU, GPU, cluster HPC) incluso lo storage.

Al contempo diventa mandatorio adottare interfacce utente intuitive e facilmente accessibili da parte di differenti figure specializzate come biologi, fisici, ingegneri, medici ecc. In ultimo luogo, il formato dei dati e le API di accesso devono rispettare standard internazionali di interoperabilità per massimizzarne la portabilità e l'esecuzione delle pipeline in ambienti cloud.

Con il presente contributo vogliamo descrivere lo stato dell'arte relativo allo sviluppo della piattaforma Genomica Computazionale nel contesto di collaborazione tra INFN e l'IRCSS AOU Sant'Orsola, principale ospedale di ricerca universitario di Bologna.

La piattaforma attualmente è realizzata su due differenti tenant nello stack IaaS OpenStack di EPIC (Enhanced Privacy and Compliance) Cloud, il servizio di INFN Cloud atto a soddisfare i requisiti dei progetti ed esperimenti che operano su dati clinici, biomedici e genomici. Al fine di garantire il massimo rispetto delle leggi, dei regolamenti e delle linee guida in vigore (come il GDPR, le autorizzazioni del Garante per la Protezione dei Dati personali e le linee guida dello European Data Protection Board), EPIC Cloud è certificata secondo gli standard per la gestione sicura delle informazioni riconosciuti a livello globale, come ISO/IEC 27001:2013, ISO/IEC 27017:2015 e ISO/IEC 27018:2017.

In questo lavoro verranno mostrate le tecnologie adottate e in particolare i miglioramenti implementati per la migrazione da una infrastruttura tradizionale basata su code verso una infrastruttura cloud, utilizzando un approccio a microsistemi con una particolare attenzione alla sicurezza dei dati.

Primary authors: COSTANTINI, Alessandro (Istituto Nazionale di Fisica Nucleare); MARTELLI, Barbara (Istituto Nazionale di Fisica Nucleare); SINISI, Francesco (Istituto Nazionale di Fisica Nucleare); SERGI, Giusy (Istituto Nazionale di Fisica Nucleare); GASPARETTO, Jacopo (Istituto Nazionale di Fisica Nucleare); MAGENTA, Letizia (Istituto Nazionale di Fisica Nucleare); GIANGREGORIO, Tania (AOU Sant'Orsola); PIPPUCCI, Tommaso (AOU Sant'Orsola)

Presenter: GASPARETTO, Jacopo (Istituto Nazionale di Fisica Nucleare)

Session Classification: Sessione "Infrastrutture ICT e calcolo distribuito"

Track Classification: Infrastrutture ICT e Calcolo Distribuito