



PHP2021 - Physics Highlights Perugia 2020/21

Martedì 26 gennaio 2021

[Virtual Room su MS TEAMS - Ore 15:00](#)

Diego Ciangottini

Integrazione di un modello di distribuzione e analisi dati per HL-LHC basato su sistemi di cache

Abstract: Per l'era HL-LHC è previsto un significativo aumento sia dei requisiti di archiviazione che di elaborazione dei dati raccolti dagli esperimenti. E' inoltre in rapida diffusione un nuovo paradigma di "resource provisioning" che vede in prima linea risorse Cloud ed eterogenee (HPC e simili). Pertanto, gli esperimenti sono stati spinti verso la valutazione di uno scenario dove l'accesso remoto ai dati può avere un ruolo importante nell'evoluzione per l'ottimizzazione della quantità di spazio disco occupato e dell'efficienza di CPU. L'obiettivo di questo contributo è presentare la prima esperienza di integrazione di una federazione INFN di server "cache" istanziati su Tier nazionali e il loro utilizzo in vari scenari di analisi dati, compreso l'utilizzo di risorse di calcolo HPC presso il CINECA con workflow di analisi dell'esperimento CMS. Inoltre, verrà mostrata l'implementazione di una piattaforma di analisi on-demand e multi-utente per utilizzare risorse cloud per l'analisi di dati in un futuro "Data-lake". Saranno presentati la costruzione di prototipi di cosiddetti Data Lakes nel contesto dei progetti IDDLS (progetto INFN-GARR) ed ESCAPE (progetto H2020).

Diego Ciangottini ha conseguito il dottorato in fisica (2015) a Perugia. Sta lavorando come assegnista di ricerca presso l'INFN di Perugia per sviluppare flussi di lavoro innovativi e soluzioni di gestione dei dati su larga scala per LHC.

[LINK ALLA VIRTUAL ROOM](#)