

CSN2

G. Mazzitelli

- Organizzazione: GdL CSN2-Calcolo, R. Cerulli, F. Di Pierro, M. Duranti, G. Mazzitelli
- Mazzitelli rappresentate in C3SN
- Di Pierro e Duranti referee in C3SN
- monitorare e dare delle linee guida per le attività di calcolo degli esperimenti in CSN2, promuovere scuole, modelli e pratiche nella comunità di astro-particle
- il meeting di novembre della CSN2 e' in buona parte dedicato al calcolo.

The screenshot shows a meeting agenda for the CSN2 event on Tuesday, 14 November 2023. The agenda is organized into sessions:

- 17:30 - 18:00:** KATRIN_TRISTAN: stato e prospettive (25+5). Speaker: Marco Carminati (Istituto Nazionale di Fisica Nucleare). Document: CSN2_KATRIN_TRIS...
- 18:30 - 19:30:** Sessione chiusa del 13 novembre 2023: aggiornamenti della GC.
- TUESDAY, 14 NOVEMBER**
- 09:00 - 13:00:** Sessione aperta: Sessione aperta del 14 Novembre 2023: il calcolo nella DSN2. Aggiornamento esperimenti.
 - 09:00:** Referaggi: proposte di assegnazione pledge 2024. Speaker: Federico Di Pierro (Istituto Nazionale di Fisica Nucleare), Matteo Duranti (Istituto Nazionale di Fisica Nucleare). Documents: Cernuicazioni-GLC..., ReferaggioCabolo?...
 - 09:40:** Assegnazioni Licenze 2024. Speaker: Matteo Duranti (Istituto Nazionale di Fisica Nucleare). Document: Licenze2024.pdf.
 - 10:10:** Aggiornamenti dal GLC.
 - 10:50:** Discussione: calcolo in CSN2.
 - 11:30:** Pausa.
 - 12:00:** DataCloud, gestione ed evoluzione dell'infrastruttura INFN per il calcolo scientifico. Speaker: Gianpaolo Carlino (Istituto Nazionale di Fisica Nucleare). Document: CSN2 1411.pdf.
- 14:30 - 18:30:** Sessione chiusa del 14 Novembre 2023: Stato delle attività.

monitoraggio e line guida

- al fine di monitorare e saper quindi elaborare delle linee guida per gli esperimenti della CSN2 e' stato introdotto all'intento del **PAQ** della CSN2 un paragrafo nel **PR annuale sul calcolo.**
- esempio (CYGNO) https://drive.google.com/file/d/1JznBAVUXPrCIC4wuHGEEqrzkL5_KAd5HO
- questo ci permette anche di capire meglio i modelli di calcolo, esigenze e avere una idea degli sforzi necessari a medio termine.

3.2 Risorse di Calcolo

Referenti del Calcolo dell'esperimento

- Referente: Giovanni Mazzitelli giovanni.mazzitelli@inf.infn.it
- Simulazioni Giulia: D'Imperio giulia.dimperio@roma1.infn.it
- Ricostruzione e Analisi: Emanuele di Marco emanuele.di.marco@cern.ch

Ruolo di INFN nel Calcolo dell'esperimento:

- Benché l'esperimento sia oggi una collaborazione internazionale il computing model e tutta la infrastruttura di storage, simulazione a analisi dati e' italiana e si basa sulla cloud INFN, dove anche i partner internazionali accedono. Non esiste quindi MoU o impegni in materia di calcolo con gli altri partner. L'INFN provvede quindi alle necessita' e al calcolo per tutta la collaborazione.
- L'esperimento ha più volte presentato il proprio computing model, brevemente descritto in reference¹, in varie conferenze internazionali (ISGC22², IEEE22³, CHEP23⁴) che e' comunque in continua evoluzione anche in parallelo agli sviluppi della stessa Cloud INFN. Il modello attualmente si basa sulle risorse in Cloud distribuite fra CNAF e ReCaS, disco S3 su backbone INFN Cloud e TAPE sul Tier1 al CNAF. C'è anche una piccolo R&D in corso su GPU a ReCaS bari. Infine, il computing model e in particolare il data e software management via RUCIO e CVMFS in Cloud, sono parte dello use case spoke2-WP3.4-Pipeline optimization for space and ground based experiments.

Pledge anno in corso:

CPU (HS06)	DISCO (TB)	TAPE (TB)	HPC (descrizione)	Infrastruttura
2200	35+5+5(*)	50	-	Cloud-INFN
1600(**)	30(***)	30(***)		Cloud-INFN
Tot: 3800	75	80		Cloud-INFN

(*) lo storage e' diviso in 3 bucket data, sim e analysis; (**) risorse temporanee - vedi dopo; (***) richiesta di anticipo del 24 sul 23 dato l'alto throughput dei dati rilevato⁵

la distribuzione e l'uso delle pledge assegnate e temporanea su Cloud-INFN al 18/7/23 e' la seguente:

¹ <https://docs.google.com/document/d/1nSVIovyaiVQ35OJX3XT1b2DcEP-waBy1MoAMiFLzpQI/>

² <https://indico4.lwgrid.org/event/20>

³ <https://nssmic.ieee.org/2022/>

⁴ <https://indico.jlab.org/event/459/contributions/11663/>

⁵ https://drive.google.com/drive/folders/13nW_HBaf3DDIT6UYzRr-Mza7KhcVv_0B/

modello richieste

- slide al link <https://agenda.infn.it/event/37772/>
 - con la nascita della C3SN e il nuovo gruppo di referaggio si e' rivisto il **modello di richiesta e proposta di assegnazione delle pladge e licenze**. Gli esperimenti inseriscono nel DB sotto il proprio esperimento le richieste di CPU, disco, tape e licenze di cui i referee scientifici fanno un prima valutazione rispetto alle attività scientifiche previste.
 - a seguire i referee del calcolo, coadiuvati eventualmente da esperti esterni, controllano la **congruità tecnica** e propongono l'assegnazione alla C3SN riversando le richieste nel **DB CALC-Tir1**.
 - questo modello ha funzionato molto bene, ma se come ora il DB **CALC-Tir1** chiude contemporaneamente, e' **possono esserci imprecisioni dell'ultimo momento**.

Licenze

- le licenze sono sempre state supportate dalla CSN2 direttamente, ma comunque i referee del calcolo, dopo la valutazione dei referee scientifici effettuano una **valutazione tecnica soprattutto volta a razionalizzare del prodotto rispetto a quanto offerto dalla CCR e della spesa**
- i fondi per le licenze **non sono più assegnati a gli esperimenti, ma al coordinatore**, in modo da evitare duplicazioni e frammentazioni in ogni sede
- purtroppo il modello di finanziamento delle CSN2 non e' quello della CCR e la logica delle "**collette**" inizia a mostrare grandi problemi.
- in generale **non e' possibile monitorare l'uso** e questo non ci permette di rendere più efficienti richieste e ottimizzare i costi.
- sarebbe necessario spingere maggiormente sui **modelli centralizzati**, non solo per la spesa, il suo controllo e razionalizzazione, ma anche per minimizzare i problemi legali, anche involterai, legati ad uso di licenze in modo non consono.