

Richieste di Calcolo per il 2022 da parte degli esperimenti di CSN1 "non LHC"

Andrea Perrotta – INFN Bologna

Riunione gruppo di referaggio del Calcolo INFN

9 settembre 2021

BELLE II

Milestones del Computing di BELLEII per il 2022:

- Conservare su TAPE 20% dei RAW Data presso CNAF (seconda copia)
- Processing e reprocessing dei RAW Data presso CNAF
- Produzione del 12% della produzione Monte Carlo su siti pledge + risorse opportunistiche

N.B. Nel 2022 è previsto uno shutdown per upgrade, da confermare in quali date

Progetto IBISCO

Completate la maggior parte delle GARE.
Per la parte HTC-Grid INFN nella sede di Napoli sono stati installati 8 nuovi Rack ciascuno attrezzato come segue:

1.4 PB di spazio disco per armadio
4-quad-server per un totale di 896 cores
Totale oltre 11PB disco e 71kHS06.

Ulteriore cluster di UNINA con un totale di circa 30kHS06.

Nuovo core switch 100G in fase di installazione.

Prime risorse già disponibili. Nel lungo termine circa 3 PB Potranno essere utilizzati per Belle II e fino a 40kH06.



Resoconto risorse disponibili per Belle II a settembre 2021

SITE	CPU	STORAGE	TAPE
CNAF	27kHS06	650TB	350TB (chiesto estensione a 650TB)
Napoli+Cosenza	13kHS06	390TB	
Pisa	8kHS06 (4 da installare)	200TB da configurare	
Torino	6kHS06	200TB da installare	
TOTALE OGGI	54kHS06	1.440TB (1.040 Utilizzabile)	650TB

Analisi esigenze per l'anno prossimo

Viste le milestone, il long shutdown che interesserà parte del 2022 le esigenze sono le seguenti

	CPU	STORAGE	TAPE
Esigenze 2022 in base alle milestone	47 kHS06	1.810 TB	450 TB
TOTALE Disponibile oggi	54 kHS06	1.440 TB (1.040 Utilizzabile)	650 TB
Da richiedere		+370 TB	

Richieste NETWORKING

In corso Upgrade dell'infrastruttura di rete del sito di Torino. Richiesto contributo a CCR per la connessione delle macchine e storage di Belle II

Richieste Inventariabili 2022

Ma il CNAF chiede di inserirli comunque nelle richieste affinché ne resti traccia

SITO	ACQUISTI	COSTO	Origine fondi
CNAF	300 TB TAPE	5.4 k€	Dalla Giunta. Già forniti in anticipo (stimato 18€xTB)
CNAF	170 TB Disk	23.8 k€	Dalla Giunta. (stimato 140€xTB)
Napoli	200 TB Disk	28 k€	Su fondi di progetto IBISCO (stimato 140€ x TB)
Torino	Contributo Core switch	2.5 k€	Da CCR
VALORE TOTALE		59.7 k€	
TOTALE CSN1		28 k€	Tutto su IBISCO

N.B. I costi a insistere sulla Giunta o su IBISCO verranno comunque riportati nella tabella richieste come da prassi.

Richieste Inventariabili 2022 (updated)

SITO	ACQUISTI	COSTO	Origine fondi
CNAF	300 TB TAPE	5.5 k€	Dalla Giunta. Già forniti in anticipo (18€xTB)
CNAF	170 TB Disk	20.5 k€	Dalla Giunta. (120€xTB)
Napoli	200 TB Disk	24 k€	Su fondi di progetto IBISCO (120€ x TB)
Torino	Contributo Core switch	2.5 k€	Da CCR
TOTALE GIUNTA		26 k€	
TOTALE CSN1		24 k€	Tutto su IBISCO

N.B. I costi a insistere sulla Giunta o su IBISCO verranno comunque riportati nella tabella richieste come da prassi.

BES III

- Richiesto a CCR il contributo di 1 kE per l'upgrade del server del T2 di Torino (come per BelleII)
- La stessa richiesta è stata fatta anche per CNS1: **va rifiutata** (doppio conteggio) se passa quella alla CCR

COMPASS - Torino

- Richiesti **5 kE** per sostituzione CPU obsolete a Torino

Storico finanziamenti
Elastic Farm - TO

	CORE (n)	STORAGE (TB)
2014	30	0
2015	100	20
2016	150	40
2017	150	80
2018	340	80
2019	340 sostituzione cpu obsolete 2014	100
2020	340	120
2021	340 sostituzione cpu obsolete 2015	120

COMPASS - Trieste

Farm di Trieste:

- CPU: 5.9 kHS
- Disco: 55 TB in replica 2 (occupati al 50%) + 400 TB in replica 1 (occupati al 32%)
TOTAL EOS STORAGE (COMPASS+OTHERS) = 860 TB
PS: COMPASS è l'unico esperimento con una partizione /noreplica/
Queste risorse di disco sono tutte dedicate alle analisi dei dati già raccolti e simulazione

- Compass raccoglierà nel 2021-22 20 PB/y nei dataset di loro interesse, fattore di compressione per mDST è circa 20; in più devono storare samples di calibrazione, servizio, etc.(megaDST): totale 120 TB.
- Richiesti **120 TB** di disco a Trieste per la presa dati 2021-22
 - Chiedono 17 kE
 - (considerando invece 120 E/TB → **14.5 kE**)

```
# _____  
# ==> Quota Node: /eos/infnts/compass/  
# _____  
.....  
group   used bytes logi bytes used files aval bytes aval logib aval files filled[%] vol-status ino-status  
project 52.72 TB 26.36 TB 23.25 k- 110.00 TB 55.00 TB 0 - 47.93 ok ignored
```

```
# _____  
# ==> Quota Node: /eos/infnts/compass/noreplica/  
# _____  
# .....  
group   used bytes logi bytes used files aval bytes aval logib aval files filled[%] vol-status ino-status  
project 126.43 TB 126.43 TB 491.11 k- 398.59 TB 398.59 TB 0 - 31.74 ok ignored
```

#

TOTAL EOS STORAGE (COMPASS+OTHERS) = 860 TB

PS: COMPASS is the only experiment with a partition /noreplica/

FCC

- Finanziamenti assegnati per il 2021 erano stati:

Proposte

CNAF

- CPU: 1 kHS = 10 k€
- Disco 100 TB = 14 k€

Si propone il finanziamento come gettone di ingresso nell'infrastruttura INFN.

- Calcolo svolto finora principalmente al CERN e nelle farm locali
- Ritardi nell'imparare le procedure per l'utilizzo delle risorse al CNAF hanno comportato un sottoutilizzo di quanto assegnato al T1
- Puntano a utilizzare le risorse già assegnate: **nessuna nuova richiesta per il 2022**

G-2

Stanno elaborando il modello di calcolo per i prossimi anni nei quali pubblicheranno prima con 4 volte la statistica del Run1 (già acquisita nei Run2 e 3) e poi con ~20 volte per il risultato finale (Run4 acquisito, Run5 da svolgere).

Nel 2019 erano stati assegnati 51 kE per 400 TB di disco a FNAL: acquistati 420 TB, che al netto di RAID e filesystem overhead corrispondono a circa **330 TB utilizzabili per lo storage**. Era stata inizialmente stimata una necessità per il Run5 di 700 TB, 550 TB per i dati (30 TB * 18.4 BNL) e 150 TB per il MC

Riviste le stime quest'anno: **22 TB per BNL per i dati, 50 TB per le simulazioni.**

Necessità 2022: 455 TB – 300 TB disponibili = 125 TB → 125 TB / 80% = **156 TB**

Stima dei costi: 65 E / TB → 10.2 kE. Richiedono **10 kE**

Nessuna richiesta per CPU power addizionale.

KLOE

Richieste per il 2022:

- 26 kE per la manutenzione della tape library e del Cisco switch:
- 5 kE per eventuali sostituzioni di pezzi non in manutenzione

Nel 2021 (e 2020):

- 26 kE per «Manutenzione nuova libreria tape-drive 10 TB, CISCO detector + salaA»: **assegnati**
- 5 keuro per eventuali sostituzioni di pezzi non in manutenzione: **assegnati sub iudice**

LHCf

Produzione (dati e MC, grezzi e ricostruzione) **centralizzata al CNAF**

Farm locali utilizzate per le analisi.

Risorse attualmente dedicate al CNAF:

- **100 TB disco**, in questo momento liberi 33,6 TB (dopo «pulizia» iniziale)
- **11 kHS06**
- (No tapes)

LHCf

Per la presa dati 2015 si occupano circa 25 TB: 1.5 TB di dati e 23.5 TB di simulazione.

Nel run del 2022 si punta a aumentare la statistica (dei dati) di un fattore 10, mentre la stessa quantità di simulazione prodotta per la precedente campagna sarà sufficiente.

- Dati $\rightarrow 1.5 \text{ TB} * 10 = 15 \text{ TB}$
- MC $\rightarrow 23.5 \text{ TB} \approx 25 \text{ TB}$

Richiedono **10 TB addizionali al CNAF**. ($38.5 - 33.6 = 5 \text{ TB}$, e dicono che potrebbero recuperare qualcosa «utilizzando degli algoritmi di generazione ottimizzati e recuperando un po' di spazio disco extra»... ma un po' di contingenza gliela lascerei)

Nel 2020 LHCf poteva usare 11000 HS06. Essendo la generazione degli eventi e la simulazione delle interazioni processi abbastanza time-consuming, se non fosse un problema aumentare un po' (10%?) la disponibilità di CPU rispetto alla situazione attuale, sarebbe comunque molto utile e apprezzato, anche se **non indispensabile**.

MUCOL



MuCol

Risorse importanti erano state assegnate alla riunione di bilancio del 2021 (dalla slide di Gianpaolo alla CSN1 di Perugia, settembre '21)

Non è chiara l'esigenza di una statistica così elevata, si propone quindi di fornire risorse pledged per 1/3 della richiesta e usare opportunisticamente le risorse (ampiamente) disponibili al CNAF e in IBISCO per il restante 2/3. Il disco non è utilizzabile opportunisticamente

Proposte di finanziamento

- PD – Tier2: **10 kE** (possibile anticipo)
 - Non si finanziano risorse fuori dei siti ufficiali dell'infrastruttura INFN
- CNAF (disponibili da inizio anno)
 - CPU 3 kHS = **30 kE**
 - Storage: 150 TB di disco = **21 kE** e 150 TB tape **3 kE**

IBISCO BA

- CPU: 7 kHS = **70 kE** e Disco: 300 TB = **42 kE**

IBISCO NA

- CPU: 3 kHS = **30 kE** e Disco: 300 TB = **42 kE**

PON IBISCO: risorse disponibili tra primavera e estate 2021. Possibile utilizzo di altre risorse necessarie in maniera opportunistica

Incontrate inizialmente diverse difficoltà al CNAF, per i seguenti motivi:

- il limite di RAM dei job su grid, poiché hanno molti job che utilizzano dai 10 ai 30 GB di RAM
- il proxy renewal (48 ore durata massima del proxy), la cui durata non è sufficiente per le applicazioni previste

Dopo diversi tentativi e miglioramenti nel codice, sono riusciti a girare solamente un tipo di job al CNAF: la **simulazione del beam-induced background**.

Adottato solo recentemente il seguente **modello di calcolo**:

- Simulazione del beam-induced background al CNAF
- Simulazione segnale e ricostruzione segnale+beam-induced background a Padova (+Trieste)

Le risorse su Padova hanno permesso di realizzare la maggior parte degli studi fino ad oggi, e sono state finanziate quasi interamente dall'Università (+ un contributo di 10k euro dall'INFN come anticipo 2020). Le risorse di Trieste fanno parte di una cloud già esistente.

Per quanto riguarda i job al CNAF, a causa dei problemi incontrati, arrivati a regime solo circa un mese fa.

Inviati grafici sull'utilizzo delle CPU al CNAF (see next slide), che parte da giugno dal momento che prima facevano essenzialmente solo test:

- Media giornaliera di circa 72 HS06, che dovrebbero mantenere nel prossimo periodo.
- Per quanto riguarda lo spazio disco, circa 20 TB su 150 TB occupati.

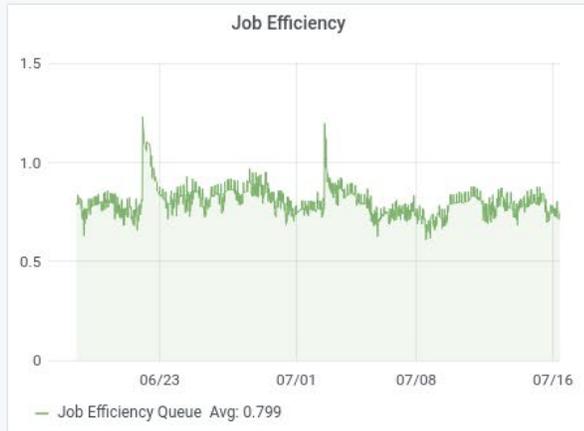
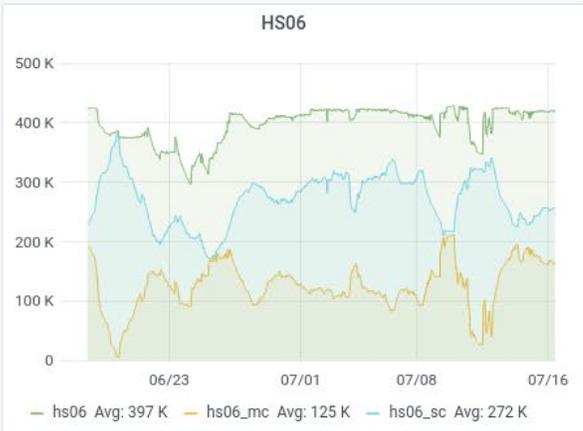
Un altro aspetto è il rinvio della strategia di Snowmass all'anno prossimo: avevano previsto un maggiore utilizzo delle risorse al CNAF anche da parte di utenti esterni (ad esempio dai colleghi USA) ma a causa dello spostamento di circa un anno della deadline questo non è avvenuto.

«Con una estrapolazione, utilizzando le risorse del CNAF allo stesso ritmo dell'ultimo mese, dovremmo arrivare ad occupare circa 70/150 TB alla fine del 2021.»

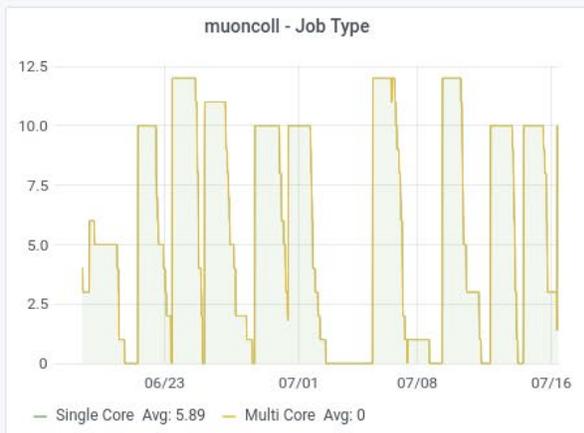
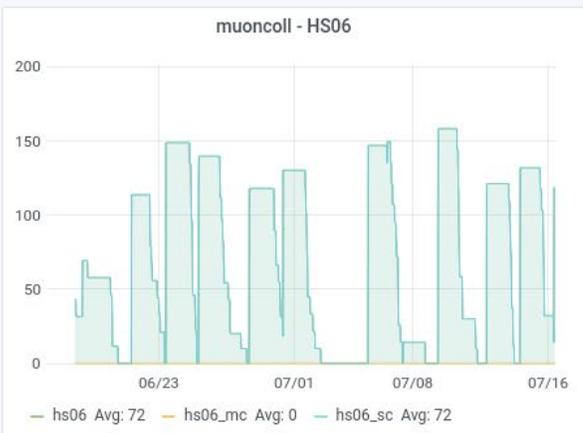
«Per questi motivi riteniamo che siano sufficienti anche per tutto il 2022, e non chiediamo nessun incremento.»



Total



muoncoll



MUCOLL: e le risorse del PON?

Hanno fatto richieste per Bari, senza tenere conto che quelle risorse erano già state richieste l'anno scorso col PON:

- CPU 7 kHSE : richiesti 70 kE (su PON)
- Disco 120 TB : richiesti 17 kE → 14.5 kE (su PON)

Da verificare i tempi di consegna delle risorse finanziate col PON

MUONE

- Avranno in prospettiva bisogno di risorse di Calcolo centrali per simulazioni e storage addizionale dei loro dati
- Per il 2022 **non ritengono di fare richieste di calcolo all'INFN**

NA62

- L'esperimento usa le risorse del T1 per produzioni Montecarlo e non per analisi. Tutta l'analisi viene fatta al CERN e non sono mai stati fatti grossi sforzi per usare la grid.
- I file prodotti al CNAF vengono poi copiati in UK, "assemblati" e spostati al Cern. La quantità di cpu utilizzata non è regolare ma varia a seconda della tornata di produzione MC.
- Limitato utilizzo anche per la continuazione di NA62, KLEVER
- Le stime di utilizzo per la cpu per il prossimo anno sono uguali a quelle utilizzate nel 2021 e già disponibili (o almeno il gruppo che si occupa del MC non ha avvisato di aver bisogno di incrementare le risorse).

Nessuna richiesta

PADME

- La collaborazione è al momento impegnata nell'analisi dei dati del Run 2, settembre-dicembre 2020, per la quale i mezzi di calcolo e di storage erano stati allocati negli anni scorsi.

Per il 2022 PADME sta chiedendo al Comitato Scientifico LNF l'assegnazione di tre mesi di fascio del Linac di Frascati per effettuare misure specifiche per la ricerca della particella X17. La richiesta verrà discussa a novembre. Se verrà approvata, **si dovrà prevedere lo storage per i nuovi dati.**

Da inizio novembre 2020 attivata la procedura di **zero-suppression on-line**, cosa che ha portato la dimensione degli eventi raw da circa 800kB a circa 300kB. Assumendo quindi una rate di trigger di 55Hz (50Hz di trigger di fascio + 5Hz di trigger cosmici e off-beam), 3 mesi di presa dati al **100% di efficienza** richiedono circa 150TB di storage. *[circa 130 TB, se non ho sbagliato i conti (AP)]*

Al CNAF sono già stati assegnati abbastanza nastri per ospitare la nuova statistica e non sono quindi necessarie risorse aggiuntive. Vorrebbero invece **ampliare lo spazio disco assegnato a PADME sul Tier2 di Frascati:** al momento sono disponibili 180TB, interamente occupati dai dati del Run 2, sui quali la collaborazione è impegnata con l'analisi. Per poter proseguire con le attività di analisi anche durante il periodo di presa dati del 2022 sono quindi **necessari 150TB aggiuntivi.**

Dei 180TB attualmente disponibili, circa **80TB** sono stati acquistati nel 2017 e andranno fuori manutenzione nel corso del 2022: bisogna quindi programmarne la sostituzione.

- Riassumendo, per il 2022 PADME avrà bisogno di un totale di 230TB di disco a LNF (al netto del RAID6): 150 TB per ospitare i nuovi dati e 80TB per rimpiazzare i dischi che vanno fuori manutenzione
- Su Frascati c'è un totale di 10k€ di fondi residui del 2021: hanno già chiesto di inserirli nella gara per lo storage di Atlas e **contano di acquistare con questi circa 70 TB di disco.**

Visto il prezzo ottenuto per il disco nella gara per ATLAS direi che possono acquistare **almeno** tutti i rimpiazzi

Le richieste per la CSN1 sono quindi per i rimanenti 160 TB, che andranno ovviamente **sub judice all'approvazione del periodo di presa dati da parte del Comitato Scientifico LNF a novembre.**

Lato **CPU**, ritengono che gli HEPSpec attualmente disponibili al CNAF e al Tier2 di Frascati (finora in realtà sottoutilizzati) siano **sufficienti** a garantire le ricostruzioni e le produzioni MC previste per i prossimi mesi. Una parte di questi HS andrà anch'essa fuori manutenzione nel 2022: è tuttavia ragionevole rimandare la richiesta per la sostituzione al prossimo anno, quando saranno più chiare le reali necessità.

Ricontattato Emanuele Leonardi per discutere le richieste

Sommario richieste per il 2022 (in risorse)

	CPU T1	CPU T2	CPU PON	Disco T1	Disco T2	Disco PON	Tape T1	Tape T2	Tape PON	Altro (kE)
BELLEII	0	0	0	170	0	200	300	0	0	2,5
BESIII	0	0	0	0	0	0	0	0	0	1
COMPASS	0	0,5	0	0	120	0	0	0	0	0
FCC	0	0	0	0	0	0	0	0	0	0
G-2	0	0	0	0	83,33337	0	0	0	0	0
KLOE	0	0	0	0	0	0	0	0	0	31
LHCF	1	0	0	10	0	0	0	0	0	0
MUCOL	0	0	7	0	0	120	0	0	0	0
MUONE	0	0	0	0	0	0	0	0	0	0
NA62	0	0	0	0	0	0	0	0	0	0
PADME	0	0	0	0	160	0	0	0	0	0

Sommario richieste per il 2022 (in Euro)

	CPU T1	CPU T2	CPU PON	Disco T1	Disco T2	Disco PON	Tape T1	Tape T2	Tape PON	Altro		TOT	GE	CSN1	PON
BELLEII	0	0	0	20400	0	24000	5400	0	0	2500		52,3	25,8	2,5	24
BESIII	0	0	0	0	0	0	0	0	0	1000		1	0	1	0
COMPASS	0	5000	0	0	14400	0	0	0	0	0		19,4	0	19,4	0
FCC	0	0	0	0	0	0	0	0	0	0		0	0	0	0
G-2	0	0	0	0	10000	0	0	0	0	0		10	0	10	0
KLOE	0	0	0	0	0	0	0	0	0	31000		31	0	31	0
LHCF	10000	0	0	1200	0	0	0	0	0	0		11,2	11,2	0	0
MUCOL	0	0	70000	0	0	14400	0	0	0	0		84,4	0	0	84,4
MUONE	0	0	0	0	0	0	0	0	0	0		0	0	0	0
NA62	0	0	0	0	0	0	0	0	0	0		0	0	0	0
PADME	0	0	0	0	19200	0	0	0	0	0		19,2	0	19,2	0
												228,5	37	83,1	108,4

Per i costi unitari, assumo quelli utilizzati per le assegnazioni del 2021 per CPU e Tape, e 120 E/TBn per il disco:

- CPU → 10 E /kHS
- DISCO → 120 E / TBn
- TAPE → 18 E / TB

	2016	2017	2018	2019	2020	2021
• CPU T1:	14	10	8	10	10	10
• CPU T2:	12	10	8	10	10	10
• Disco T1:	240	200	180	160	140	140
• Disco T2:	220	200	180	170	140	140
• Tape:	25	25	25	25	18	18

Rimosse le richieste "inaccettabili":

- Finanziamento dei Core-switch per la farm di Torino (BELLE, BES), finanziate separatamente dalla CCR
- Possibile richiesta di CPU aggiuntiva per LHCf (non verrà probabilmente neanche presentata)
- Richiesta "duplicata" di disco e CPU per il sito di Bari per MUCOL, già presentata col PON 2021

Ridotta di 20 TB la richiesta di disco (SJ) per possibile extra-run di Padme

Aggiustamento richieste per il 2022 (in Euro)

	CPU T1	CPU T2	CPU PON	Disco T1	Disco T2	Disco PON	Tape T1	Tape T2	Tape PON	Altro		TOT	GE	CSN1	PON
BELLEII	0	0	0	20400	0	24000	5400	0	0	0		49,8	25,8	0	24
BESIII	0	0	0	0	0	0	0	0	0	0		0	0	0	0
COMPASS	0	5000	0	0	14400	0	0	0	0	0		19,4	0	19,4	0
FCC	0	0	0	0	0	0	0	0	0	0		0	0	0	0
G-2	0	0	0	0	10000	0	0	0	0	0		10	0	10	0
KLOE	0	0	0	0	0	0	0	0	0	31000		31	0	31	0
LHCF	0	0	0	1200	0	0	0	0	0	0		1,2	1,2	0	0
MUCOL	0	0	0	0	0	0	0	0	0	0		0	0	0	0
MUONE	0	0	0	0	0	0	0	0	0	0		0	0	0	0
NA62	0	0	0	0	0	0	0	0	0	0		0	0	0	0
PADME	0	0	0	0	15600	0	0	0	0	0		15,6	0	15,6	0

127,0 27 76,0 24

Di cui 20.6 kE SJ