

# Preventivi CMS 2023

Daniele Spiga INFN-PG  
On behalf of CMS-Calcolo Italia

# Agenda



CMS Offline & Computing: breve summary e highlights 2020

Utilizzo delle risorse Tier-1, Tier-2 e Opportunistico (HPC)

- Contributo e prestazioni dei siti italiani ( T1, T2 )
- Attività di integrazione HPC
- R&D sul modello di analisi ( focus sull'integrazione delle risorse di calcolo )

Stato delle Gare

Richieste 2023

- E discussione su “impatto JINR”

# Intro: C&O 2021



La preparazione per Run 3 è stata una priorità assoluta per CMS durante il 2021. Questo include le attività di Software e Computing offline.

Dopo quasi tre anni di spegnimento (LS2), durante il **beam test di ottobre 2021**, CMS ha collezionato con successo circa  $4 \text{ nb}^{-1}$  di collisioni di protoni

- energia di iniezione (senza pileup)

**Il test ha permesso di testare sia il rivelatore che l'infrastruttura di calcolo con dati reali:**

- Trasferimento dati da P5 fino alla scrittura su Tape (T0)
- Prompt calibration loop, prompt processing di collision data.

Eseguiti **scale test Global Pool**: gestione, contemporanea, di mezzo milione di job

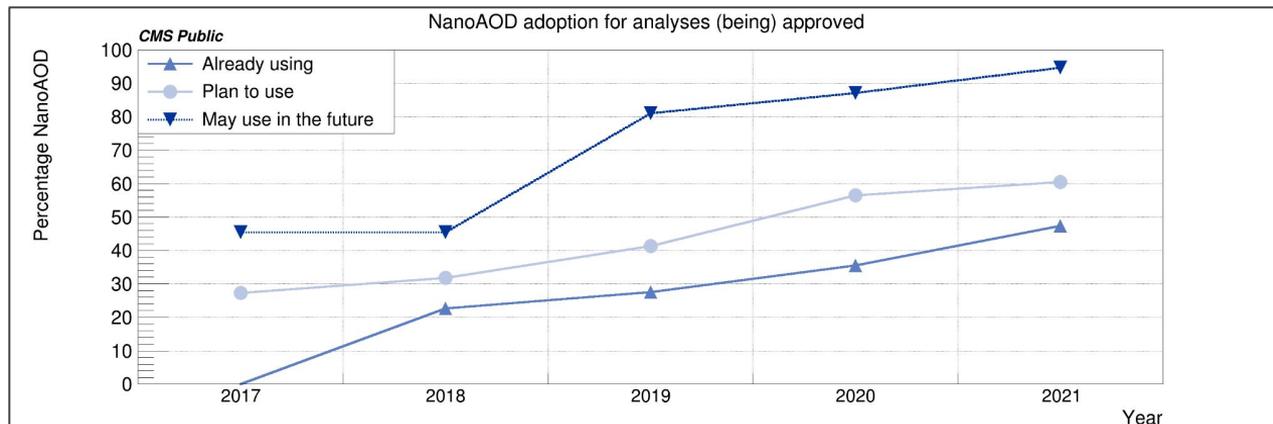
Inoltre il 2021 è stato il **primo anno completo in cui il Data Management è stato gestito con RUCIO**

# NanoAOD: stato adozione e prospettive



L'adozione del formato NanoAOD per l'analisi fisica ha continuato l'incremento anche nel corso del 2021.

- La frazione di analisi che utilizzano NanoAOD ha raggiunto quasi il 40%
- Aspetto chiave con impatto sul computing ( resource request ) future i.e. Fase 2
- Elemento abilitante per l'esplorazione di nuovi approcci all'analisi e quindi nuovi modelli. Da qui l'importanza di fare R&D [vedi dopo]

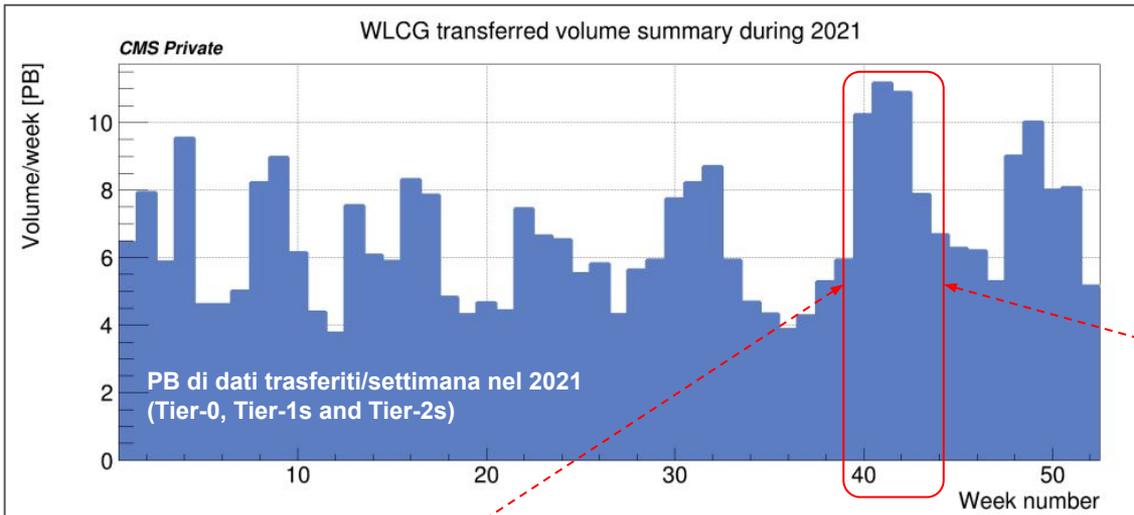


**Nota:** il goal iniziale era è di avere il 50% delle analisi su NanoAOD per Run-3



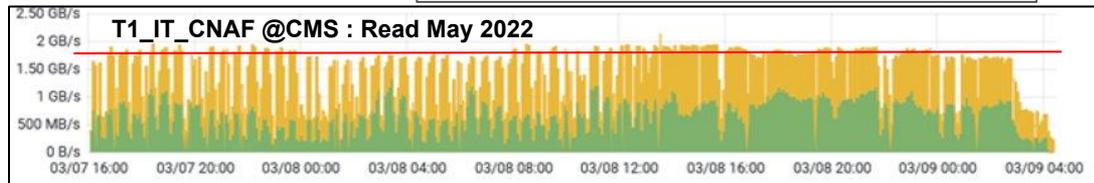
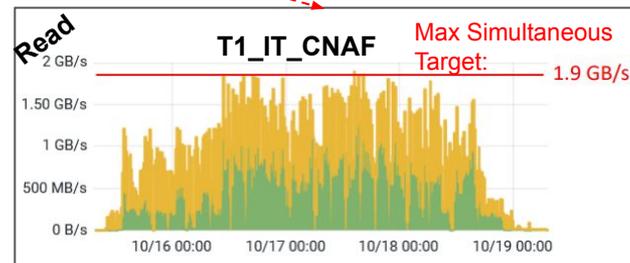
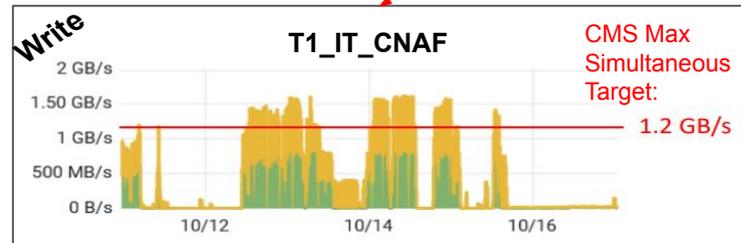
Risultati del survey sull'adozione del formato dati NanoAOD, o che prevede di farlo a breve

# Data Transfer & Tape Challenge(s)



## Challenges su network e Tape ( lettura/scrittura )

- Trasferimenti a velocità artificialmente elevata effettuati simultaneamente con ATLAS, LHCb e ALICE
- **Raggiunti i target al CNAF. Risultati confermati e migliorati nel challenge di primavera '22**



# CMS Software (CMSSW): Supporto a GPU



**Online trigger reco:** Offloading su GPU di 4 componenti principali per HLT

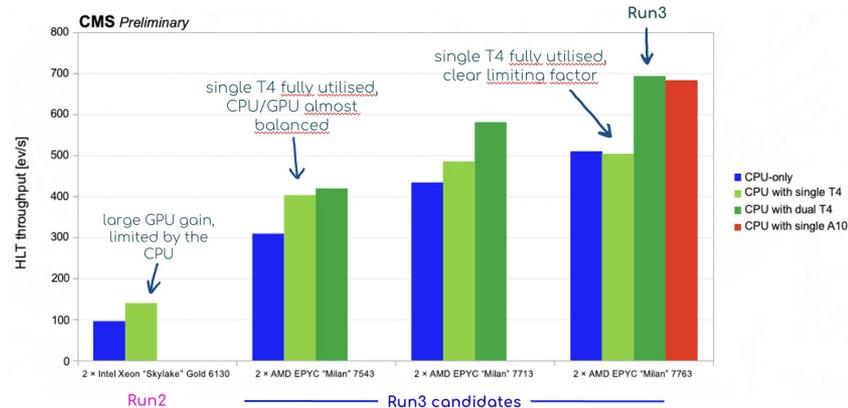
- Ultimi test mostrano +74% rispetto a solo CPU

**Ricostruzione offline:** beneficiare del porting fatto per online ed espanderlo:

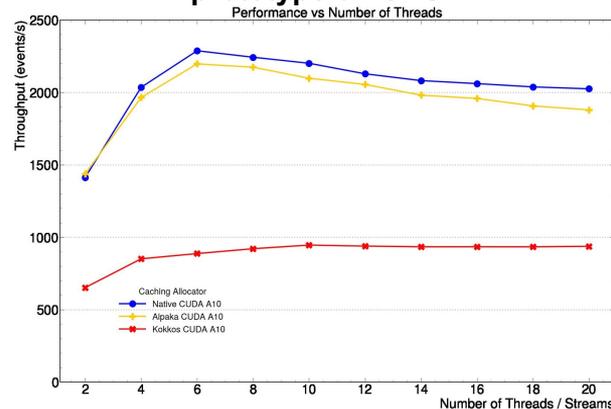
- Sfruttare la disponibilità di GPU opportunistiche su HPC
- Per Phase2 e Run3: puntare a >15% di offload su GPU 2023

**Portabilità:** supportare più acceleratori in modo trasparente:

- Riscrivere lo stesso codice per ciascuna architettura
- Manutenzione più facile
- Obiettivo di sostituire CUDA nativo con Alpaka entro il 2023.



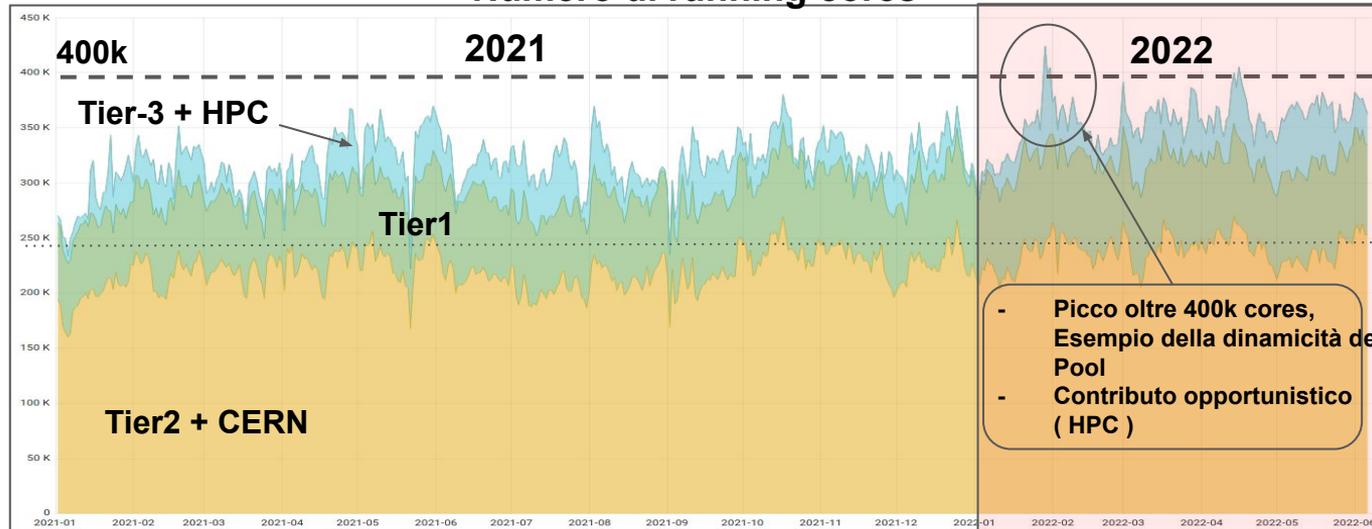
## Throughput results for the patatrack-standalone prototype on GPU



# Risorse di calcolo a CMS: contributi & utilizzo



## Numero di running cores

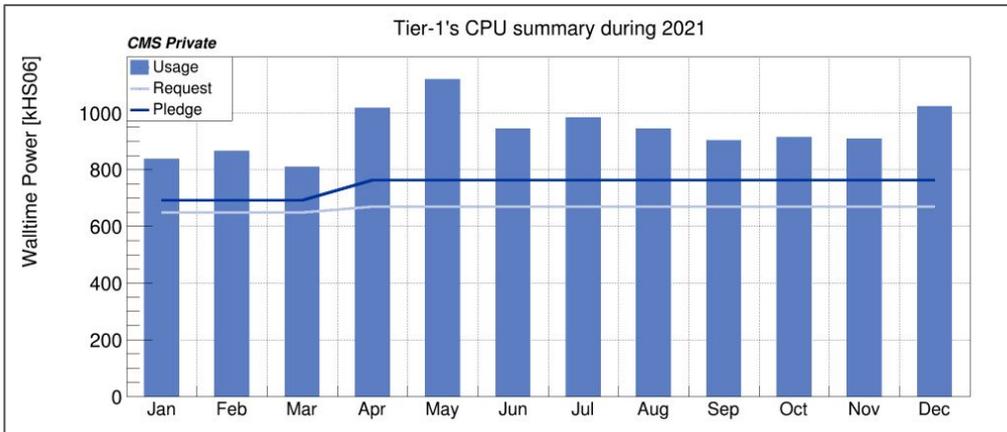


Il 2021 è stato caratterizzato da un ottimo utilizzo delle risorse pledged e opportunistiche sia presso i siti WLCG che HPC [vedi dopo].

Nel 2019 CMS ha avviato le campagne Run 2 Legacy (ricostruzione dell'intero dataset Run 2 e generazione del Monte Carlo associato). Nel 2021 sono stati processati i **campioni per gli studi Run 3 e Run 2 basati su Monte Carlo legacy e non legacy**

- 56,4 miliardi di eventi Monte Carlo unici, simulati principalmente con il toolkit Geant4 (consegnati all'analisi)
- Ricostruzione Legacy dei parked data nel 2018 (circa 10 miliardi eventi)

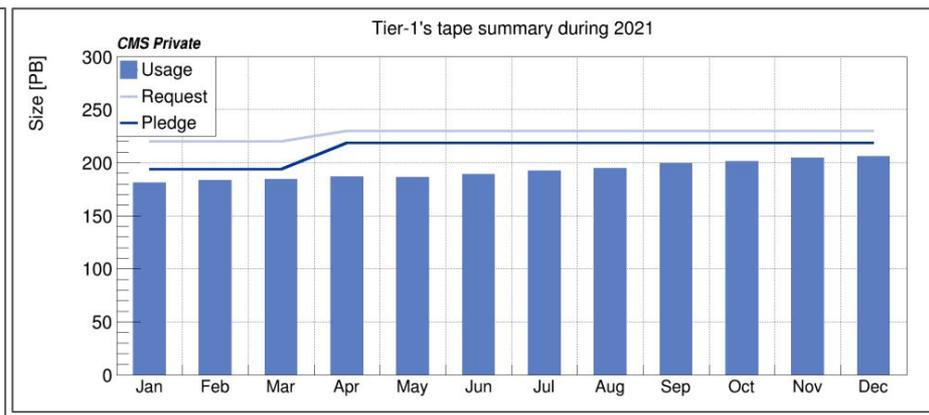
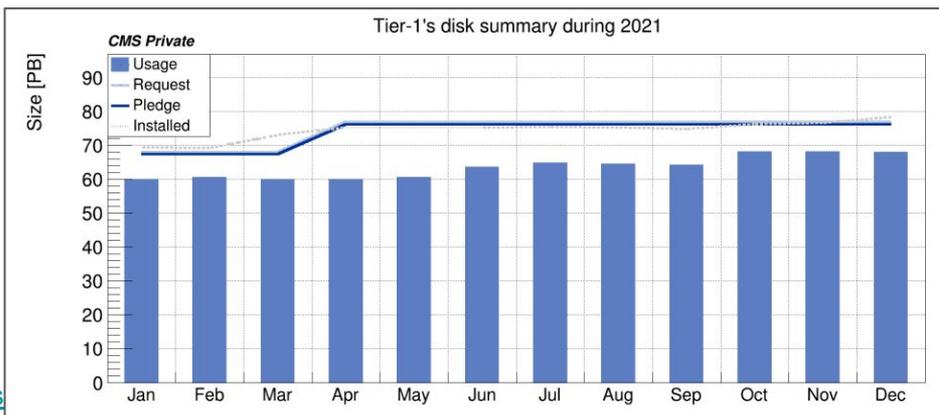
# Utilizzo dei Tier 1 : CPU - Disco e Tape



Utilizzato il 126% della CPU pledged ai Tier-1 - efficienza media del 72%

L'utilizzo del disco è l'86% del totale del disco installato disponibile. ( gestione conservativa per primo anno con Rucio

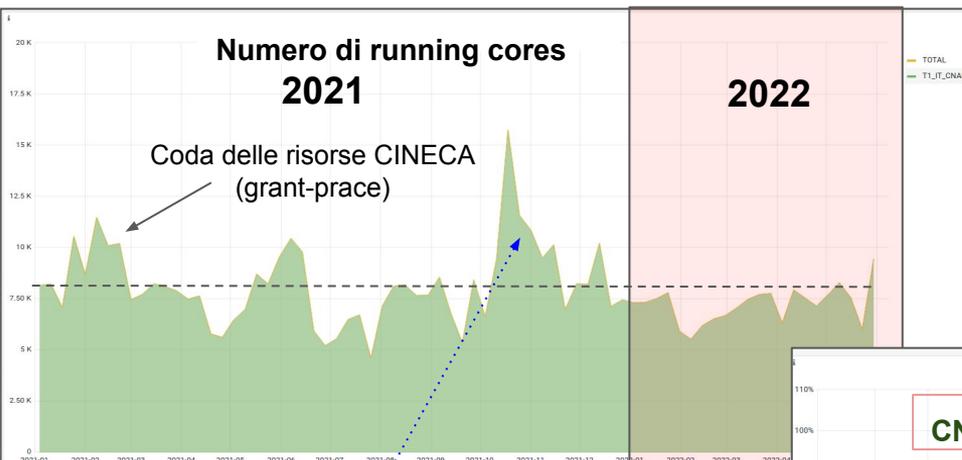
94% del tape in uso (12/2021). Il divario tra richiesto e pledged passa dal 12 al 5%!!



# Il Tier1 @CNAF



Source CERN Monit

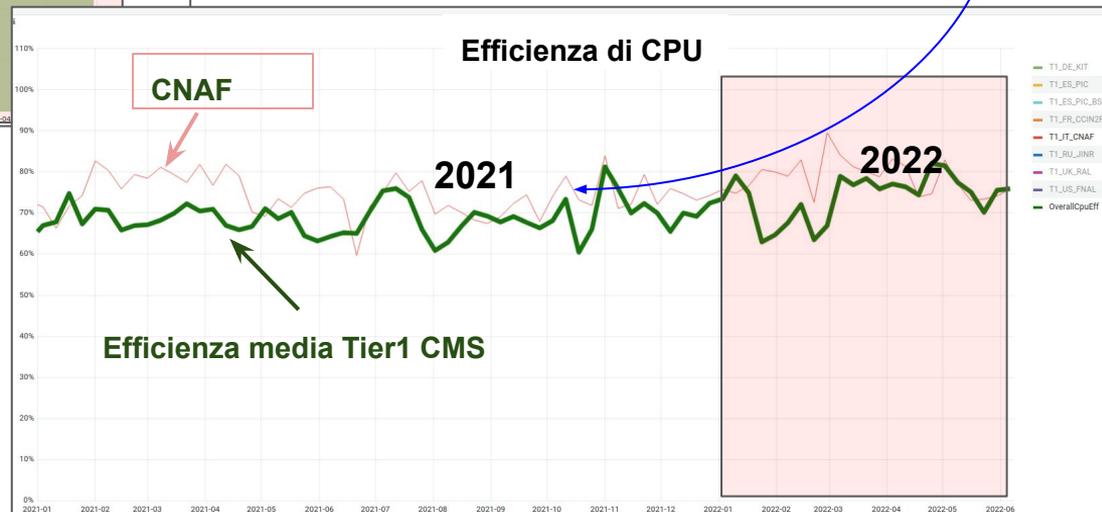


Il CNAF ha avuto ottime prestazioni durante tutto il 2021. Oltre la stabilità, l'efficienza di CPU è stata maggiore della media dei Tier1

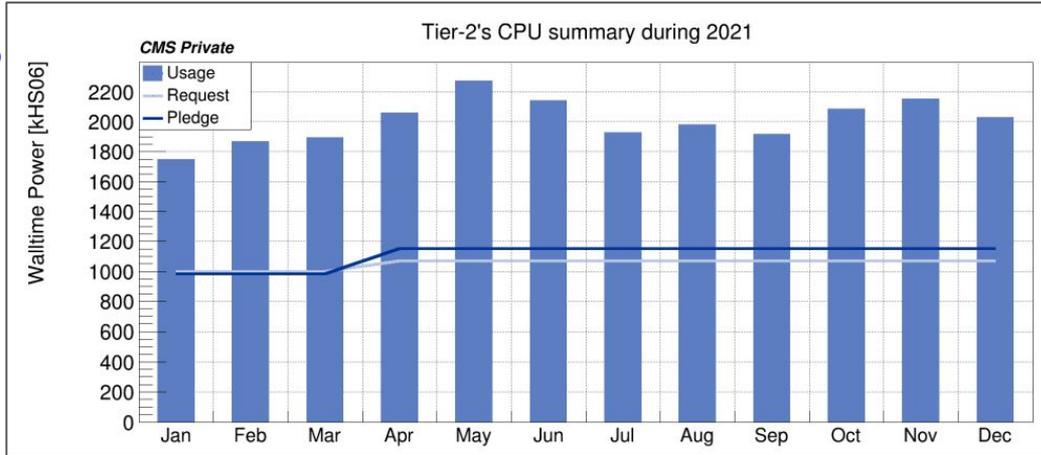
- Contributo di estremo valore per la collaborazione

Combinazione di due effetti

- Attività di fine-tuning dello share del batch al Tier1, questo include l'ottimizzazione della gestione delle slot multi-core vs single-core
- Test di integrazione tecnica i M100



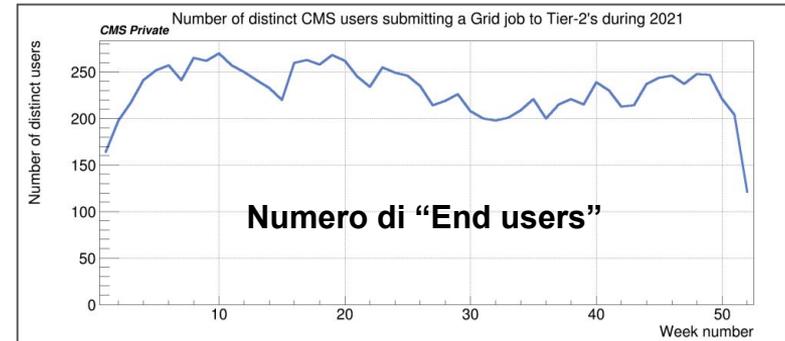
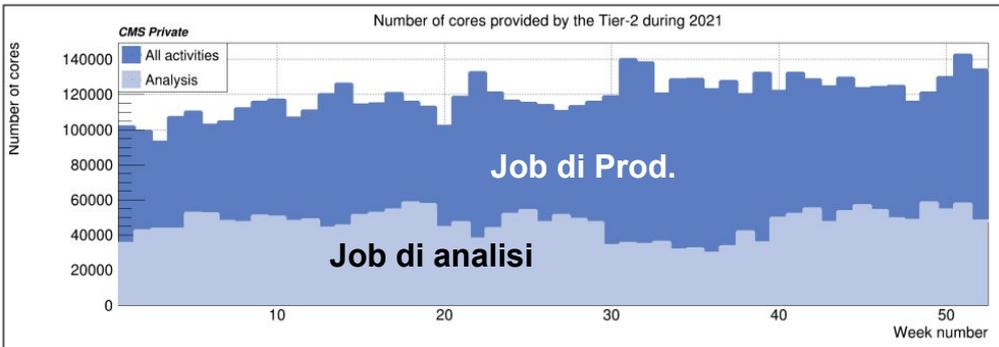
# Utilizzo dei Tier 2



5% circa di overpledge da parte dei Tier-2 rispetto alle richieste di CMS.

Estremamente utile il **beyond pledge** (che è la frazione opportunistica) principalmente fornita da **USA**

- L'utilizzo medio della CPU, integrato per tutto il 2021, è stato quindi di circa il 181% delle risorse pledged



# La federazione dei Tier2 Italiani

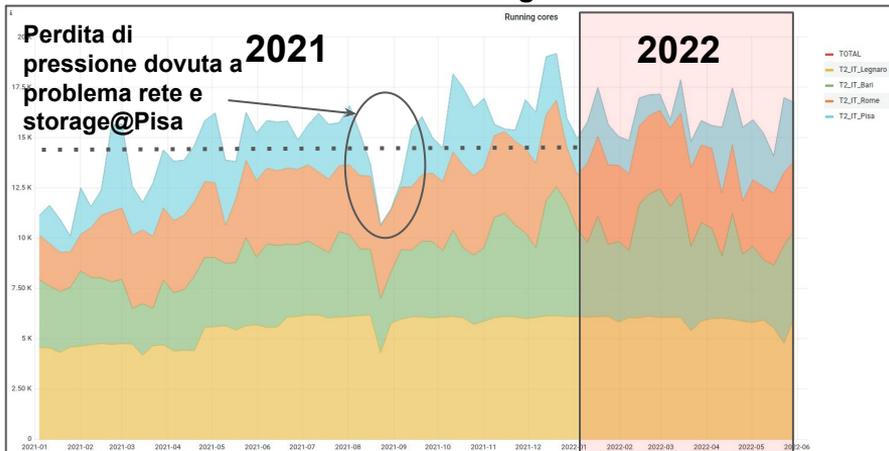


Eccellenti prestazioni dei Tier2 Italiani rispetto alle varie metriche (anche in termini di efficienza di CPU, rispetto ai valori medi)

Roma conferma ampiamente la stabilità che aveva già raggiunto durante il 2020

I Tier 2 italiani sono un contributo fondamentale al computing di CMS, sia in termini di **quantità (13%)** che di **qualità e affidabilità**

### Numero di running cores



### Efficienza di CPU



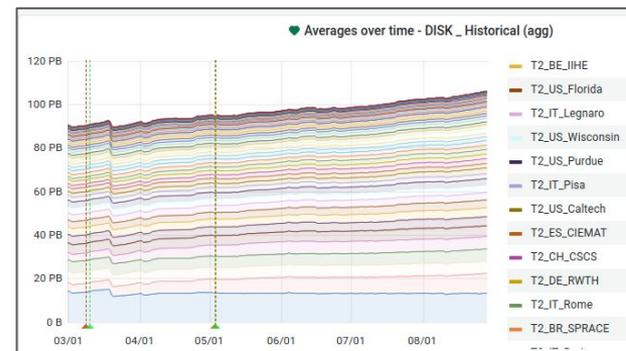
# Utilizzo delle risorse ed efficienze: sintesi



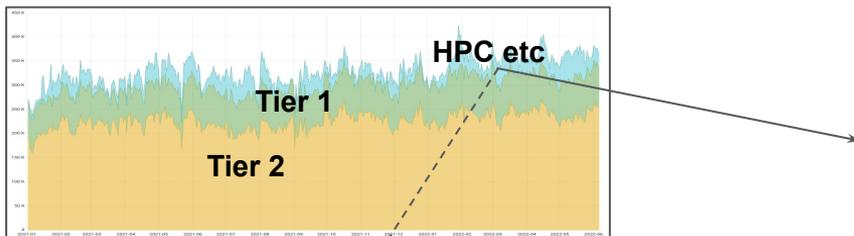
CMS		Pledged	Used	Used/Pledged	Avg. CPU Eff
CPU [kHS06]	Tier-0	500 (423)	582 (589)	135% (137%)	62% (66%)
	Tier-1	764 (693)	974 (837)	127% (121%)	72% (72%)
	Tier-2	1152 (985)	2063 (1837)	179% (186%)	71% (69%)
	<b>Total</b>	<b>2416 (2268)</b>	<b>3546 (3212)</b>	<b>147% (142%)</b>	<b>70% (69%)</b>
Disk [PB]	Tier-0	30 (26)	22 (21)	66% (81%)	
	Tier-1	76 (68)	65 (60)	85% (89%)	
	Tier-2	92 (78)	82 (-)	90% (-)	
	<b>Total</b>	<b>198 (172)</b>	<b>167 (151)</b>	<b>84% (88%)</b>	
Tape [PB]	Tier-0	120 (99)	111 (97)	92% (98%)	
	Tier-1	219 (194)	206 (185)	94% (95%)	
	<b>Total</b>	<b>339 (293)</b>	<b>317 (282)</b>	<b>94% (96%)</b>	

L'uso delle risorse è stato efficiente e in linea con le prestazioni ottenute durante il 2020

Per l'utilizzo del disco il nuovo sistema di monitoring, nell'ambito della transizione di Rucio, è disponibile da Aprile



# Altre risorse : Integrazione e uso di HPC

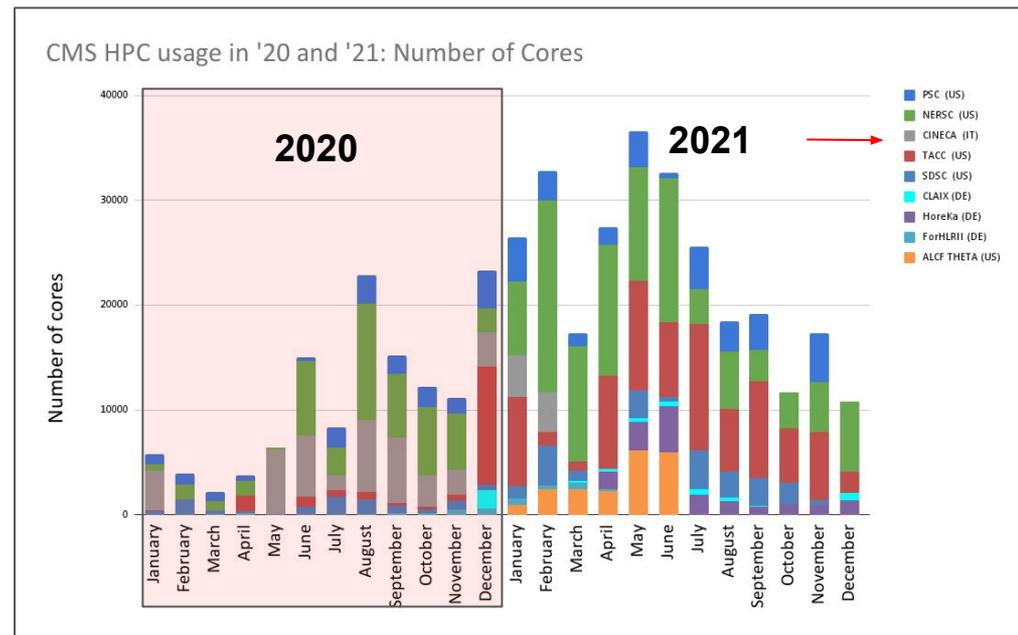


Quantità di risorse di calcolo dei centri  
**HPC usata da CMS nel 2021 è stata tre volte superiore rispetto al 2020**

- (e un ordine di grandezza superiore rispetto al 2019.)

TACC Stampede2 (HEPCloud)	HPCs	18.7	0.9	0.4
TACC Frontera (HEPCloud)		98.7	4.9	2.3
NERSC (HEPCloud)		95.8	4.7	2.2
SDSC (HEPCloud)		44.1	2.2	1.0
PSC (HEPCloud)		41.8	2.1	1.0
BSC		23.7	1.2	0.5
CINECA (CNAF)		1.5	0.1	0.0
ForHLRII (KIT)		2.7	0.1	0.1
HoreKa (KIT)		31.8	1.6	0.7
CLAIX(RWTH Aachen)		2.9	0.1	0.1

**Circa 360 kHS06**



Contributo italiano con risorse **CINECA PRACE** Project Access (30McoreH, PI T.Boccali) per il calcolo LHC

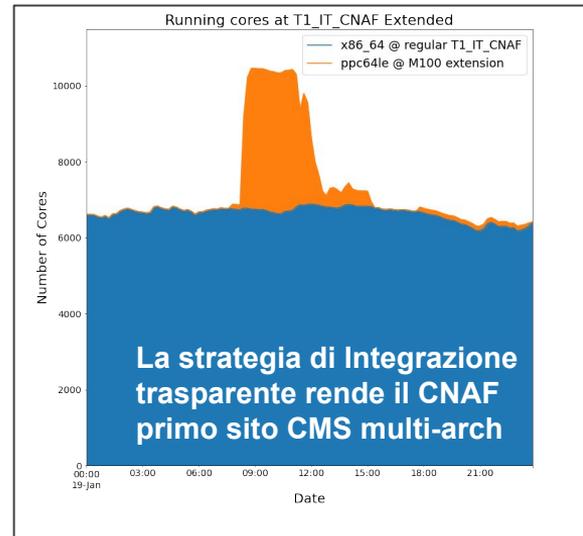
Contributo italiano è relativo ad attività di integrazione e **commissioning di M100** [vedi dopo]

# Integrazione e commissioning M100



Abbiamo integrato M100 nell'infrastruttura di calcolo di CMS come estensione del T1\_CNAF. **E' la prima facility (large-scale) NON x86 @CMS (M100 è Power9 + nvidia v100 )**

- Co-design e commissioning del workload management di CMS (WMAgent) con supporto di arch. PowerPC
  - Appena validato anche il **supporto multi-architettura** che è **fondamentale per l'uso ottimale di CNAF** (pledge + opportunistico vista della **migrazione al Tecnolo**)
- Eseguita **la validazione tecnica di Power9**
  - Ora stiamo finalizzando la validazione delle fisica. Fa da **apripista per l'integrazione di SUMMIT ( US ) e permetterà l'uso in produzione**
- **Validata l'integrazione delle risorse GPU ( V100 )**
  - Usando Workflow di test estratti da HLT



-**PowerAtCMS:** “PRACE 24th Call for Proposals for Project Access proposal No. 2021250016 “ **per 8MCoreH ( 1M GPU hours )** **Approved**

- **T.Boccali (+M.Pierini/D.Spiga)**

-**Sinergie e ricadute su attività HTC-HPC al Tecnolo ( CNAF Reloaded )**

-**Progetti EU - interTwin**

[RelMon](#) Global Report: DataReport [main...](#)

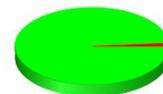
## Release Validation workflows on Data && Montecarlo

### Summary

107442 COMPARISONS:

- SUCCESS: **98.8%** (106102)
- NULL: **0.0%** (11)
- FAIL: **1.2%** (1329)

[To the DQM GUI...](#)



### Releases:

- CMSW\_12\_3\_0
- CMSW\_12\_3\_0

### Statistical Test (Pvalue threshold):

- chi2 (1E-05)

# R&D su Analysis Model



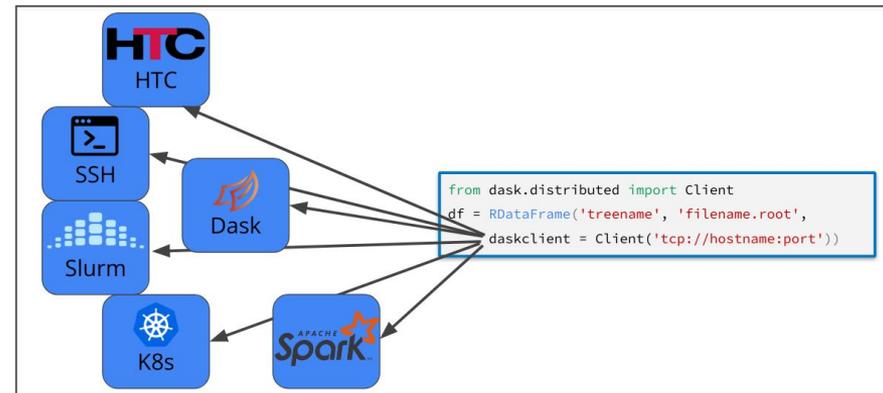
I nuovi framework di alto livello ( ROOTDataFrame/Coffea) consentono di valutare nuovi approcci all'analisi “From event looping to declarative analysis”

**Motivazione:** contribuire allo sviluppo di soluzioni tecniche per supportare lo sviluppo di nuovi **modelli di analisi, principalmente interattivi:**

- “from batch to interactive mode” e contribuire alla definizione del modello per l'uso delle risorse

## Highlights:

- Sistema che sia analysis **framework agnostic**
- Una integrazione che consenta di aumentare il **throughput (evts/s)**, che ottimizza l'uso delle risorse
- Stimolare l'adozione di **NanoAOD**
- **Integrazione GRID/Cloud/HPC**



**Domanda:** Possiamo utilizzare l'infrastruttura WLCG in questo tipo di approccio?

- Modello ibrido con offloading su risorse eterogenee e distribuite

# Multi site Analysis Facility model @INFN



Implementato un testbed con poche risorse “offerte” da INFN-Cloud e qualche di WN fuori manutenzione ai T2. Stiamo proponendo alla comunità un sistema modulare:

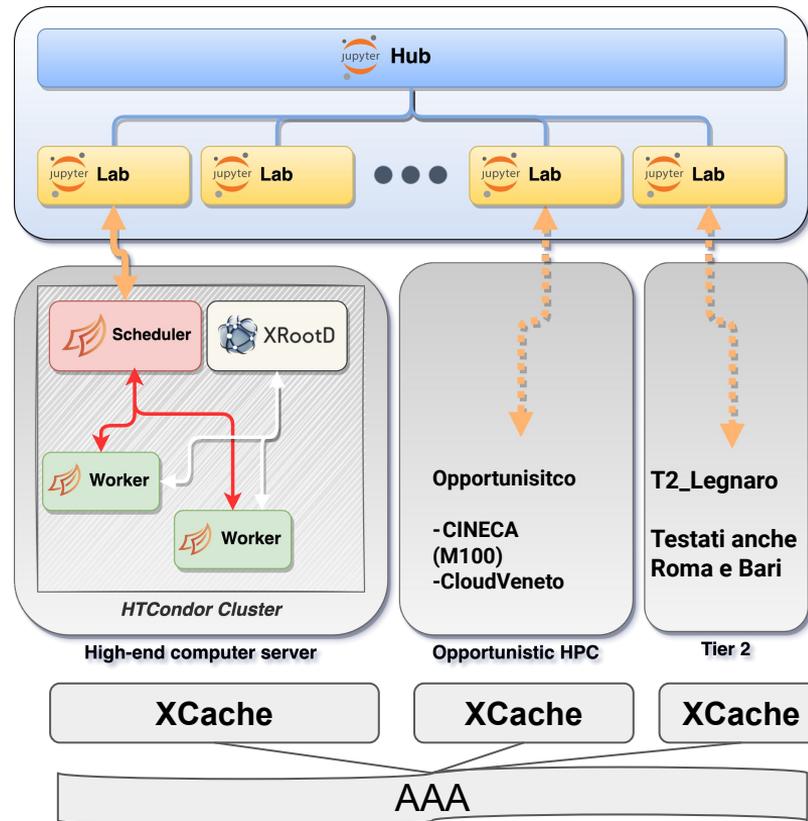
- **Per fare benchmarking:** Quantificare i guadagni ( throughput, event-loop rate, network, overall time...) e l’impatto sul computing ( Data Management, resource provisioning ). **In test con analisis VBS (PhD)**
- **Experiment agnostic :** Architettura e tecnologie sostanzialmente CMS free
- **Riuso di soluzioni e risorse e competenze:** Non reinventare la ruota e sfruttare l’esperienza maturata ( grid, cloud, applicazioni, frameworks etc )

-[HSF -AF Forum](#) co-convenership

- Sinergie e ricadute su INFN: INFN-Cloud, integrazione CINECA e Progetti esterni ( egi-ace, interTwin )

- Primi contatti con LHCb

- Presenti nell’ Analysis Tools Task Force di CMS e organizzato WS CMS Italia dedicato a Feb.22



# Stato delle gare 2022



## Legnaro:

- **CPU:** ordine Consip, hardware Installato e on-line
- **DISCO:** pochi TB, l'acquisto è stato messo come 6/5 della gara 2021
  - HW della gara 2021 è stato appena consegnato

## Pisa:

- **CPU:** Acquisto è Work in Progress.
- **Storage:** Acquistato e consegnato. Sta per essere messo on-line

## Roma:

- **CPU:** Installato ma non è ancora on-line
- **Storage:** installato ed è già on-line

## Bari:

- **Tutto il pledge 2022 è installato e on-line**
  - **Tutta la CPU è su PON IBISCO mentre per il disco 1TB è ancora in manutenzione ( dimissione 2024 )**
  - NOTA che tutta la CPU andrà in dismissione in bulk nel 2026; analogamente 2PB di disco andranno in dismissione nel 2027.



# Richieste CMS risorse di calcolo 2023



Approvato RRB 28 Aprile 2022

Table 6: CMS resource request for 2023. The percentage changes with respect to the 2022 approved request are shown.

CMS		'22 Approved Request - Spring '21	'23 Preliminary Request - Fall '21	'23 Final Request - Spring '22	Increase wrt '22	
					Abs.	Perc.
CPU [kHS06]	Tier-0	540	720	720	180	33%
	Tier-1	730	800	800	70	10%
	Tier-2	1,200	1,350	1,350	150	13%
	<b>Total</b>	<b>2,470</b>	<b>2,870</b>	<b>2,870</b>	<b>400</b>	<b>16%</b>
Disk [PB]	Tier-0	35	45	45	10	29%
	Tier-1	83	98	98	16	18%
	Tier-2	98	117	117	19	19%
	<b>Total</b>	<b>216</b>	<b>260</b>	<b>260</b>	<b>45</b>	<b>20%</b>
Tape [PB]	Tier-0	155	228	228	73	47%
	Tier-1	260	316	316	56	22%
	<b>Total</b>	<b>415</b>	<b>544</b>	<b>544</b>	<b>129</b>	<b>31%</b>

L'italia è (ancora) il  
**13%**, sta crescendo  
ma quest'anno la  
consideriamo ancora il  
13%

# Dismissioni Tier-2 per il 2023



Lo scenario delle dismissioni del '23 gode di una **“risonanza costruttiva”** dovuta all’effetto gara unica 2018 ( conclusa effettivamente nel 2019 )

	Disco (TBN)	CPU (kHS06)	\$Disco	\$CPU
Bari	0	0	0	0
Pisa	0	4.14	0	62100
Legnaro	0	11	0	165000
Roma1	0	6.7	0	100500
<b>TOT</b>	<b>0</b>	<b>21.84</b>	<b>0</b>	<b>327600</b>

Tuttavia questo ha un effetto poi sul 2024 [vedi dopo]



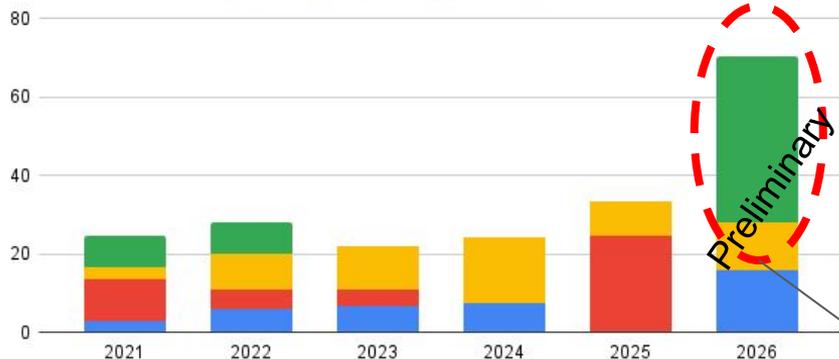


# Dismissioni ( CPU e DISCO ): Proiezioni future

**NOTA Bene:** Il prossimo ciclo di bilancio ( 2024 ) vedrà la dismissione di una ingente quantità di storage distribuita su quattro siti

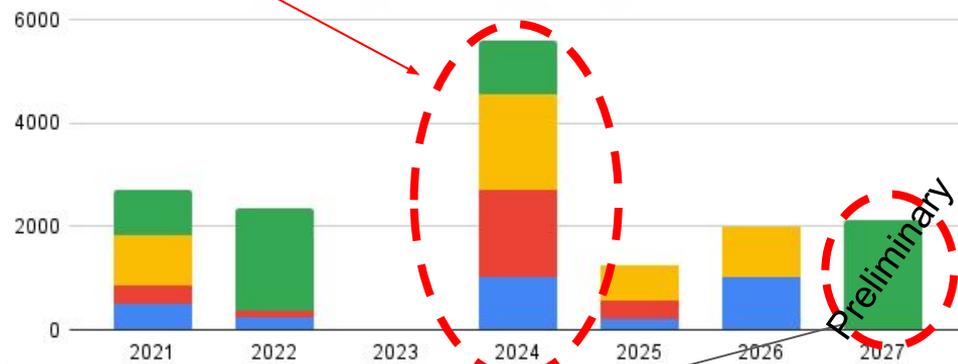
CPU - kHS06

■ Bari ■ Legnaro ■ Pisa ■ Roma1



DISCO - TBN

■ Bari ■ Legnaro ■ Pisa ■ Roma1



Effetto bulk injection a Bari PON IBISCO

# Nuove richieste 2023 ( Delta di crescita )



	Disco (TBN)	CPU (kHS06)	\$Disco	\$CPU
Bari	617.5	4.87	86450	73050
Pisa	617.5	3.87	86450	58050
Legnaro	617.5	1.87	86450	28050
Roma1	617.5	2.57	86450	38550
<b>TOT</b>	<b>2470.0</b>	<b>13.18</b>	<b>345801</b>	<b>197700</b>

0.00
1.00
3.00
2.30
6.30

Gli acquisti CPU 2022, come per il 2021, **sono stati fatti ad 8 ad HS06** ( invece di 10 )  
 Il surplus di CPU lo scaliamo sul delta del 2023.

Questo porta un risparmio di 63Ke ( 32% sulla crescita di CPU )

	Disco (TBN)	CPU (kHS06)	\$Disco	\$CPU
Bari	617.5	4.87	86450	73050
Pisa	617.5	3.87	86450	73050
Legnaro	617.5	4.87	86450	73050
Roma1	617.5	4.87	86450	73050
<b>TOT</b>	<b>2470.0</b>	<b>19.48</b>	<b>345801</b>	<b>292200</b>

Scenario senza risparmio acquisti CPU 2022

# Complessivamente per i 4 Tier2

Risorsa	Euro
1 HS06	15
1 TBN	140
1 TB Tape	10

**INFN**  
 La rete è  
 scorporata e  
 inserita in  
**DATA CLOUD**

	TOT (euro)	disco TBN	disco Eur	CPU kHS06	CPU Eur	Rete Eur	Server Eur
Bari	179370.89	617.5	86450	4.9	73050	8705.5175	11165.0245
Pisa	232593.89	617.5	86450	8.0	120150	11531.5175	14462.0245
Legnaro	314970.89	617.5	86450	12.9	193050	15905.5175	19565.0245
Roma1	253950.89	617.5	86450	9.3	139050	12665.5175	15785.0245
<b>TOT</b>	<b>980886.57</b>	<b>2470.01</b>	<b>345801.40</b>	<b>35.0</b>	<b>525300.00</b>	<b>48808.07</b>	<b>60977.10</b>

**NOTA:** Il surplus acquisti CPU 2022 incide per un **risparmio di circa 9%** rispetto al totale iniziale

	TOT (euro)	disco TBN	disco Eur	CPU kHS06	CPU Eur	Rete Eur	Server Eur
Bari	179370.89	617.5	86450	4.9	73050	8705.5175	11165.0245
Pisa	249543.89	617.5	86450	9.0	135150	12431.5175	15512.0245
Legnaro	365820.89	617.5	86450	15.9	238050	18605.5175	22715.0245
Roma1	292935.89	617.5	86450	11.6	173550	14735.5175	18200.0245
<b>TOT</b>	<b>1087671.57</b>	<b>2470.01</b>	<b>345801.40</b>	<b>41.3</b>	<b>619800.00</b>	<b>54478.07</b>	<b>67592.10</b>

# Effetto JINR

Tolto JINR dal pledge 2022  
E non dalle richieste 2023



Tier-1 Resource	'22 Pledge	'23 Request	'23 Pledge increase	
			Abs.	Perc.
Tape [PB]	220	316	96	44%
Disk [PB]	70	98	28	40%
CPU [kHS06]	612	800	188	31%

JINR, è il secondo T1 più grande di CMS. In termini di pledges 2022: **35 PB tape, 13 PB disk and 240 kHS06 CPU**

- 13%, 16% and 33% delle richieste totali dei Tier-1

Lo scenario in cui scompare un sito non ha effetto sulle richieste dell'esperimento. Le richieste non diminuiscono a meno di effettuare tagli alla fisica

- Il problema è come aiutare per soddisfarle

Table 6: CMS resource request for 2023. The percentage changes with respect to the 2022 approved request are shown.

CMS		'22 Approved Request - Spring '21	'23 Preliminary Request - Fall '21	'23 Final Request - Spring '22	Increase wrt '22	
					Abs.	Perc.
CPU [kHS06]	Tier-0	540	720	720	180	33%
	Tier-1	730	800	800	70	10%
	Tier-2	1,200	1,350	1,350	150	13%
	<b>Total</b>	<b>2,470</b>	<b>2,870</b>	<b>2,870</b>	<b>400</b>	<b>16%</b>
Disk [PB]	Tier-0	35	45	45	10	29%
	Tier-1	83	98	98	16	18%
	Tier-2	98	117	117	19	19%
	<b>Total</b>	<b>216</b>	<b>260</b>	<b>260</b>	<b>45</b>	<b>20%</b>
Tape [PB]	Tier-0	155	228	228	73	47%
	Tier-1	260	316	316	56	22%
	<b>Total</b>	<b>415</b>	<b>544</b>	<b>544</b>	<b>129</b>	<b>31%</b>



# Mitigazioni considerate da CMS

Stiamo pianificando di mettere in atto diverse misure di mitigazione che riducono i requisiti per il 2023 (e 2022):

- Migliore compressione RAW con l'algoritmo LZMA: nell'ottimizzazione finale
  - guadagno di spazio fino al 10%
-  Campagna di cancellazione del nastro in corso, già prevista nella richiesta del 2022, ma l'abbiamo resa più aggressiva:
  - Eliminazione di 69 PB nei Tier 1 e 12 PB nel Tier 0
- Mitigazione dovuta a fattori esterni, in particolare LHC darà meno dati di quanto previsto in fase di richieste 2022 (Circa Fattore 2 )

Somma di tutti gli effetti di cui sopra:

- Riduce il fabbisogno di Tape di 6,4 PB nei Tier 1, 9,9 PB nel Tier 0

**Applichiamo il totale della mitigazione di 16 PB ai Tier 1.**

- Il modello di calcolo CMS è abbastanza flessibile da permetterlo

CMS ha già iniziato “la ricerca” di risorse per l'integrazione di un nuovo Tier 1

- **operazione con effetto a medio/lungo termine**



# Richieste senza JINR + Mitigazioni

Tier-1 Resource	'22 Pledge	'23 Request	'23 Pledge increase		'23 Needs Mitigated	'23 Mitigated need increase	
			Abs.	Perc.		Abs.	Perc.
Tape [PB]	220	316	96	44%	300	80	36%
Disk [PB]	70	98	28	40%	93	23	33%
CPU [kHS06]	612	800	188	31%	750	138	23%

Rinormalizzando lo share delle FA che ospitano i Tier1 si ottiene che l'Italia passa dal 13% al 15.5% (**solo su Tier 1 NON si applica ai Tier 2**)

Site	Share [%]	
	taken from CRIC	renormalized w/o JINR
FNAL	40	47.7
CNAF	13	15.5
IN2P3	10.3	12.3
KIT	10	11.9
RAL	6.5	7.8
PIC	4	4.8
<b>Total</b>	<b>83.8</b>	<b>100</b>

# Crescita Tier1 CNAF 2023

RICHIESTE 2023				
T1	pledge 2022	pledge 2023	Incremento	Eur Delta
CPU (kHS06)	94.9	104	9.1	136500
DISK (TBN)	10790	12740	1950	273000
TAPE (TB)	33800	41080	7280	72800
				482300

## Richieste Finali

RICHIESTE 2023 no JINR CON Mitigazione				
T1	pledge 2022	pledge 2023	Incremento	Eur Delta
CPU (kHS06)	94.9	116.25	21.35	320250
DISK (TBN)	10790	14415	3625	507500
TAPE (TB)	33800	46500	12700	127000
				954750

Impatto globale con mitigazione è questo

## Come sarebbe senza mitigazioni applicate da CMS

RICHIESTE 2023 no JINR no Mitigazione				
T1	pledge 2022	pledge 2023	Incremento	Eur Delta
CPU (kHS06)	94.9	124	29.1	436500
DISK (TBN)	10790	15190	4400	616000
TAPE (TB)	33800	48980	15180	151800
				1204300

- Nessuna richiesta per il 2022: il run time ridotto e le mitigazioni messe in atto da CMS consentono di operare con i siti rimanenti

- **Per il 2023 ci sono mancanze significative:** le priorità di CMS in ordine di importanza sono: **prima il Tape**, poi il disco Tier-1 e infine la CPU Tier-1.
- **NOTA:** Questa injection è da considerare **una-tantum** e non da reiterare ( Poi si torna al flat-budget )



# Organigramma Offline & Computing



## Coordinators

D. Piparo, J. Letts

### Core Software

S. Muzaffar, M. Kortelainen

### Simulation

V. Ivantchenko, S. Bein

### Monitoring & Analytics

B. K. Jashal, F. Legger

### Analysis Infra. & Support

S. Belforte, K. Ellis

### Generators \*

S. Bhattacharya, G. Chahal

### Computing Resources Board

J. Hernandez, K. Bloom

### Computing Operations

C. Paus, A. Wightman

### Workload/Data Mgt. Devel.

K. Lannon, D. Ciangottini

### Offline Release Planning

A. Perrotta, Q. Li

### Upgrade R&D and TDR

D. Elvira, TBD

### Machine Learning \*

J.R. Vlimant, G. Kasieczka

### Dyn. Res. Provisioning

D. Spiga, C. Wissing

### Reconstruction

J. Pata, C. Caputo

### Upgrade Software

P. Srimanobhas, A. Di Florio

### Web Services & Security

A. Pfeiffer, P. Paparrigopoulos

### L1 Software \*\*

C. Callol, E. Palencia Cortezon

### Facility Services

G. Bagliesi, S. Lammel

### Resource Management

J. Flix, D. Lange

### Submission Infra.

A. Perez-Calero, M. Mascheroni

### DPOA \*\*\*

K. Lassila-Perini, C. Lange (Dep.)

\* Joint with Physics / \*\* Joint with L1 DPG / \*\*\* Joint with CB  
Mandates can be found [here](#).



E responsabilità  
italiane

# Tabella responsabilità



Giuseppe	Bagliesi	L2	Pisa	Facilities Services, Resp. Nazionale Computing
Stefano	Belforte	L2	Trieste	Analysis Infrastructure and support
Diego	Ciangottini	L2	Perugia	Data Management & Workload Management
Adriano	di Florio	L2	Bari	Upgrade Software
Federica	Legger	L2	Torino	Monitoring & Analytics
Andrea	Perrotta	L2	Bologna	Offline Release Planning
Daniele	Spiga	L2	Perugia	Dynamic Resource Provisioning
Massimo	Biasotto	L3	LNL	Resp. tier2 LNL
Giacinto	Donvito	L3	Bari	Resp. tier2 Bari
Enrico	Mazzoni	L3	Pisa	Resp. tier2 Pisa
Shahram	Rahatlou	L3	Roma1	Resp. tier2 Roma

# Backup



# Il 13% dell'Italia

## ANNEX 1

PhD Scientists per Funding Agency Based on the Annually Revised Annex 13 of the M&O MoU

The List of Names is Available at

[https://cms-docdb.cern.ch/cgi-bin/DocDB/RetrieveFile?docid=398&filename=CMS\\_PhD\\_List\\_2021.pdf&version=16](https://cms-docdb.cern.ch/cgi-bin/DocDB/RetrieveFile?docid=398&filename=CMS_PhD_List_2021.pdf&version=16)

(Count closed on 17 September, 2020)

Institute FA	PhD #	PhD %
Austria	11	0.78%
Belgium-FNRS	24	1.71%
Belgium-FWO	14	1.00%
Brazil-FAPESP	7	0.50%
Brazil-RENAFAE	21	1.49%
Bulgaria	9	0.64%
CERN	70	4.98%
China	32	2.28%
Colombia	4	0.28%
Croatia	10	0.71%
Cyprus	8	0.57%
Ecuador	2	0.14%
Egypt	3	0.21%
Estonia	7	0.50%
Finland	12	0.85%
France-CEA	15	1.07%
France-IN2P3	40	2.85%
Germany-BMBF	54	3.84%
Germany-Helmholtz	37	2.63%
Hungary	12	0.85%
India	29	2.06%
Iran	5	0.36%
Ireland	1	0.07%
Italy	185	13.17%
Korea	34	2.42%
Kuwait	1	0.07%
Latvia	3	0.21%
Lithuania	4	0.28%
Malaysia	2	0.14%
Mexico	14	1.00%
Montenegro	1	0.07%
New Zealand	2	0.14%
Pakistan	4	0.28%
Poland	10	0.71%
Portugal	8	0.57%
RDMS-DMS	26	1.85%
RDMS-Russia*	70	4.98%
Serbia	4	0.28%
Spain	41	2.92%
Sri Lanka	3	0.21%
Switzerland ETHZ	18	1.28%
Switzerland PSI	7	0.50%
Switzerland University	10	0.71%
Taipei	14	1.00%
Thailand	3	0.21%
Turkey	17	1.21%
United Kingdom	56	3.99%
USA-DOE	304	21.64%
USA-DOE-NP	30	2.14%
USA-NSF	84	6.12%
USA-OTHER	4	0.28%
<b>Grand Total</b>	<b>1405</b>	<b>100%</b>

\*The figures include PhDs financed from sources external to the Funding Agency

2022

Hungary	12	0.85%
India	29	2.06%
Iran	5	0.36%
Ireland	1	0.07%
Italy	185	13.17%
Korea	34	2.42%
Kuwait	1	0.07%

2023

Iran	6	0.42%
Ireland	1	0.07%
Italy	197	13.64%

# Data Transfer V2



T0 -> T1s BUFFER (FTS monitoring)

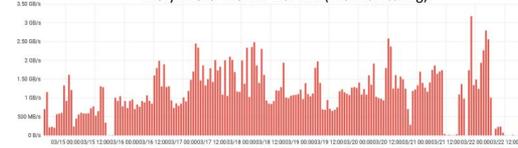


- Target avg rate: **0.37 GB/s**

T0 -> CNAF buffer

- Avg rate: 0.32 GB/s
- Test data amount: 234 TB
- Periods with no transfers due to the stress on T0 disk

Everywhere -> CNAF BUFFER (FTS monitoring)



- Avg rate: **1.09 GB/s**
- Overall data amount: 775 TB
- Production data amount: 541 TB
- Peak: >2.5 GB/s

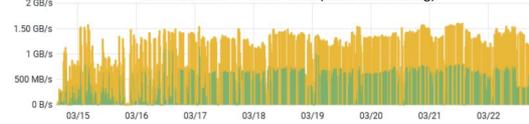
Everywhere -> T1s BUFFER (FTS monitoring)



Everywhere -> CNAF buffer

- Avg rate: **1.09 GB/s**
- Overall data amount: 775 TB
- Production data amount: 541 TB
- Peak: >2.5 GB/s

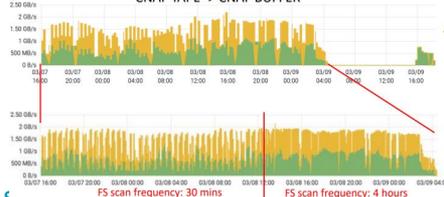
CNAF BUFFER -> CNAF TAPE (CNAF monitoring)



- Avg rate: **1.09 GB/s**
- Overall data amount: 775 TB
- Production data amount: 541 TB
- Peak: >1.5 GB/s (7 hours)

- Target rate: just put stress on the system and verify results
- Overall CMS RUN3 target avg rate: **1.9 GB/s**
- A certain inefficiency due to server load caused by frequent FS scan until 8 Mar 12h
- All data from Oracle library (data older than April 2020).
  - The rate should be better with newer data (recall from IBM library).
  - We are available to repeat the test with newer or mixed data.

CNAF TAPE -> CNAF BUFFER



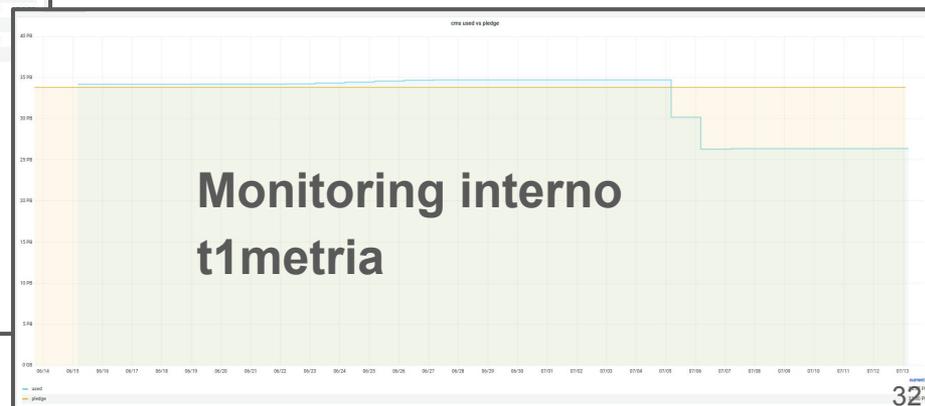
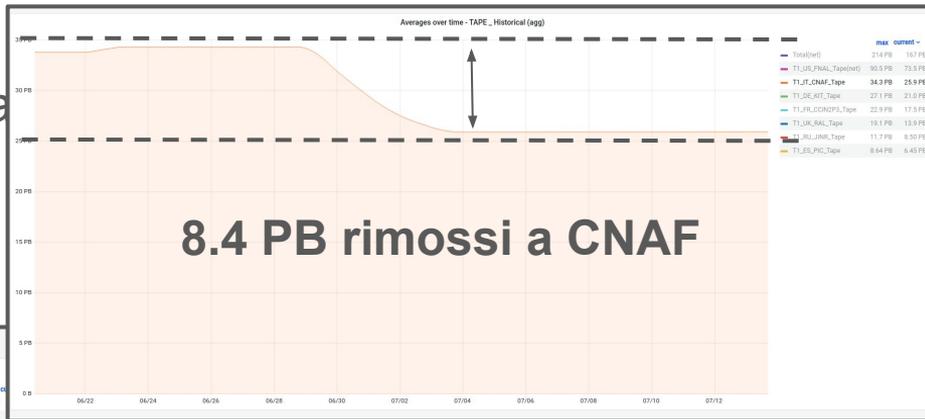
- Overall rate (many hours without requests): **1.03 GB/s**
- Recall rate with continuous requests: **1.46 GB/s**
- Best recall rate (9 hours): **1.8 GB/s**

# Tape Deletion:



**Quanto:** Target  $\sim 80$  PB  $\rightarrow$  1/4 di tutto il Ta

- Total 81.6 PB (**69PB T1**, 12PB T0)

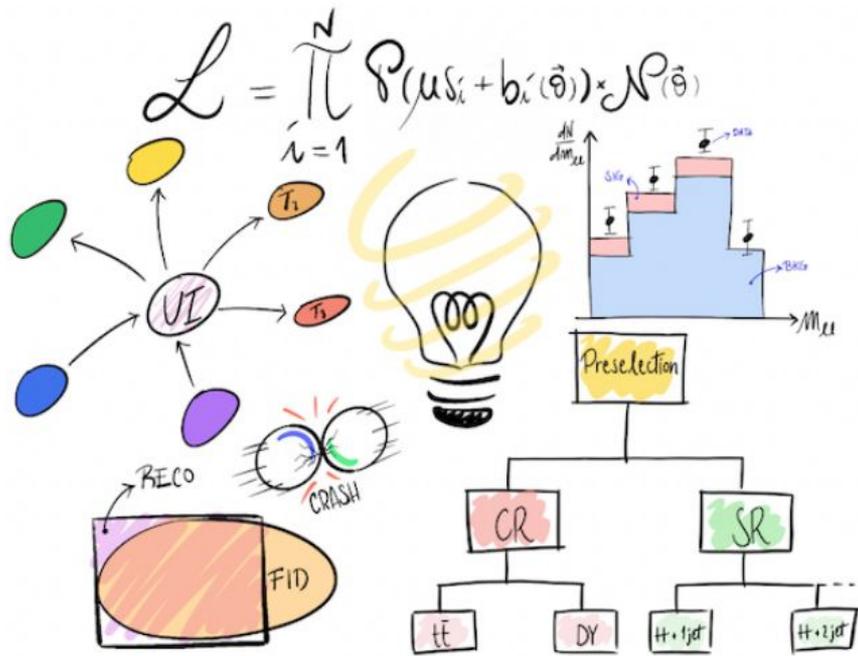


# PowerAtCMS



processed. The PowerAtCMS project has a twofold ambition:

- use Marconi 100 for the CMS Experiment, by replicating the infrastructure we used for the PRACE Project Access Grant #2018194658 on Marconi A2 [lhc@hpc];
- implement a mechanism to logically partition single nodes with “mostly CPU bound” workflows with “mostly GPU bound” ones; the two workflows collaborate towards a common goal, fully exploiting the node capabilities;
- an End-to-End Deep Learning reconstruction workflow is trained to substitute the most time consuming classical algorithms in CMS: the reconstruction of particle showers in the Endcap Hadron Calorimeter Reconstruction (HGCal), their accurate simulation when generating synthetic datasets for physics studies, and the linking of these clusters to those from other detector components, in the global event reconstruction of CMS events, based on the concept of particle flow.



# analisi dati @CMS Italia

**RESCHEDULED**  
NEW DATE ANNOUNCED

~~24-25 Febbraio 2022~~  
**10-11 Marzo 2022**  
**Firenze**



# Preliminare



## Summary

Preselection		
	Legacy	RDF
Overall time	7h	0.5h
Overall rate	602 Hz	8099 Hz
Event-loop rate	1239 Hz	8725 Hz
Overall network read	540 GB	34 GB
Average RSS per-node	Ca. 15 GB	Ca. 15 GB

Postselection - 1st scenario		
	Legacy	RDF
Overall time	1h*	0.083h
Overall rate	509 Hz	5961 Hz
Event-loop rate	XXX Hz	XXX Hz
Overall network read	12 GB	10 GB
Average RSS per-node	Ca. 1 GB	Ca. 15 GB

Postselection - 2nd scenario		
	Legacy	RDF
Overall time	1.5h	0.25h
Overall rate	45.6 Hz	404.3 Hz
Event-loop rate	XXX	XXX
Overall network read	84 GB	12 GB
Average RSS per-node	Ca. 5 GB	Ca. 15 GB

The screenshot shows the homepage of the Analysis Facilities Forum. The navigation bar includes links for Home (HSF), Working Groups, Activities, Meetings, Communication, Projects & Support, and About. The main heading is "Analysis Facilities Forum" followed by an "Overview" section. The overview text describes the forum's purpose in supporting HEP data analysis through shared resources and collaboration. A "What is an Analysis Facility?" section is also visible at the bottom.

