

Review Calcolo ATLAS Italia

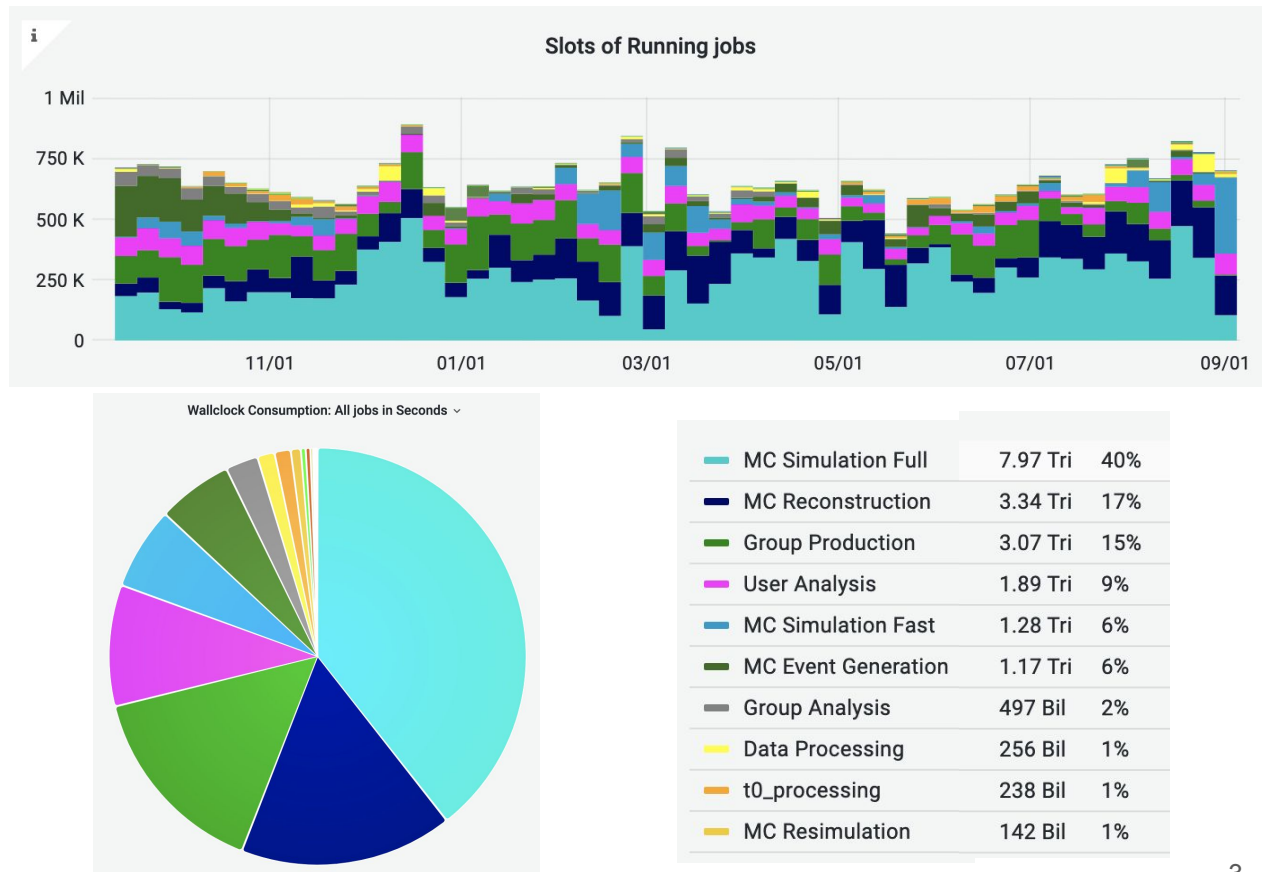
L. Carminati (UniMi), A. Doria (INFN-Napoli), L. Rinaldi (UniBo)
on behalf of the ATLAS-ITALIA computing group

Roma, 7 settembre 2023

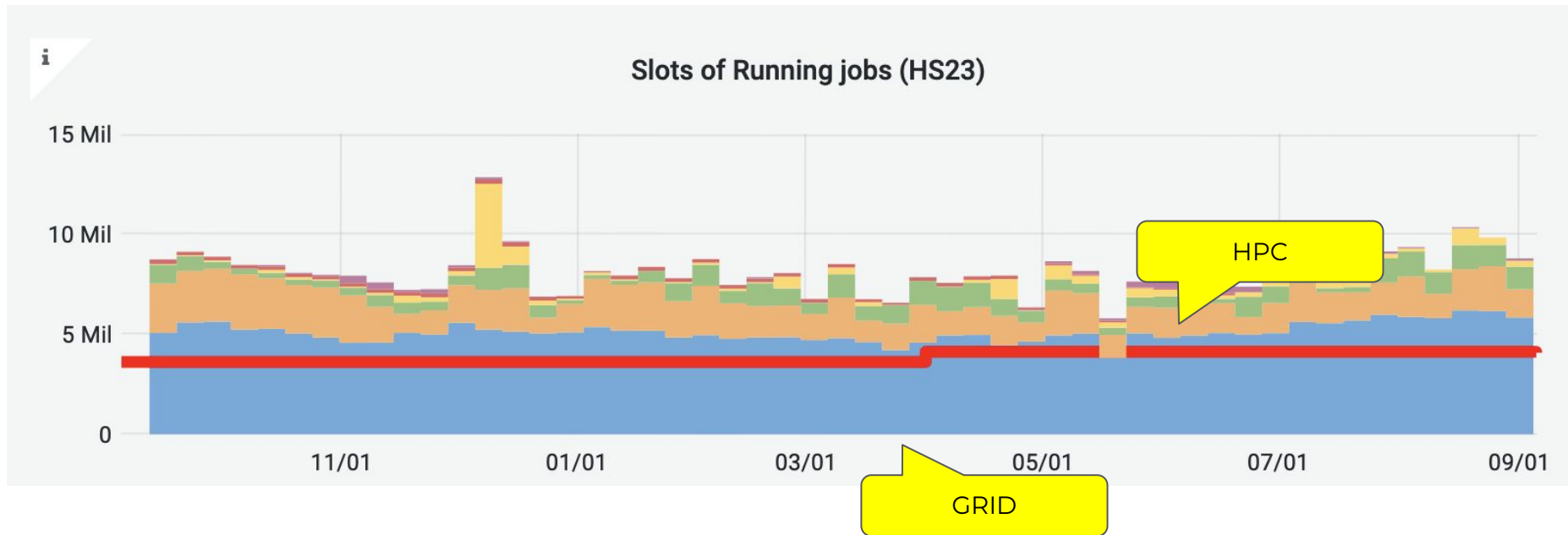
1. Utilizzo risorse calcolo a livello di esperimento
2. Utilizzo risorse calcolo INFN
3. Richieste per il 2024

Utilizzo risorse ATLAS (settembre 2022 - settembre 2023) : CPU

- ❑ Circa 650k (average) computing slots ogni giorno per tutto il periodo in esame (pledged e unpledged, tutti i tiers insieme)
- ❑ Intensa attività di produzione MC (~70% includendo generazione, simulazione e ricostruzione) e analisi dati (~30% tra derivations e analisi vera e propria)
- ❑ Efficienza (wallclock consumption kHS23 seconds) ~ 92%

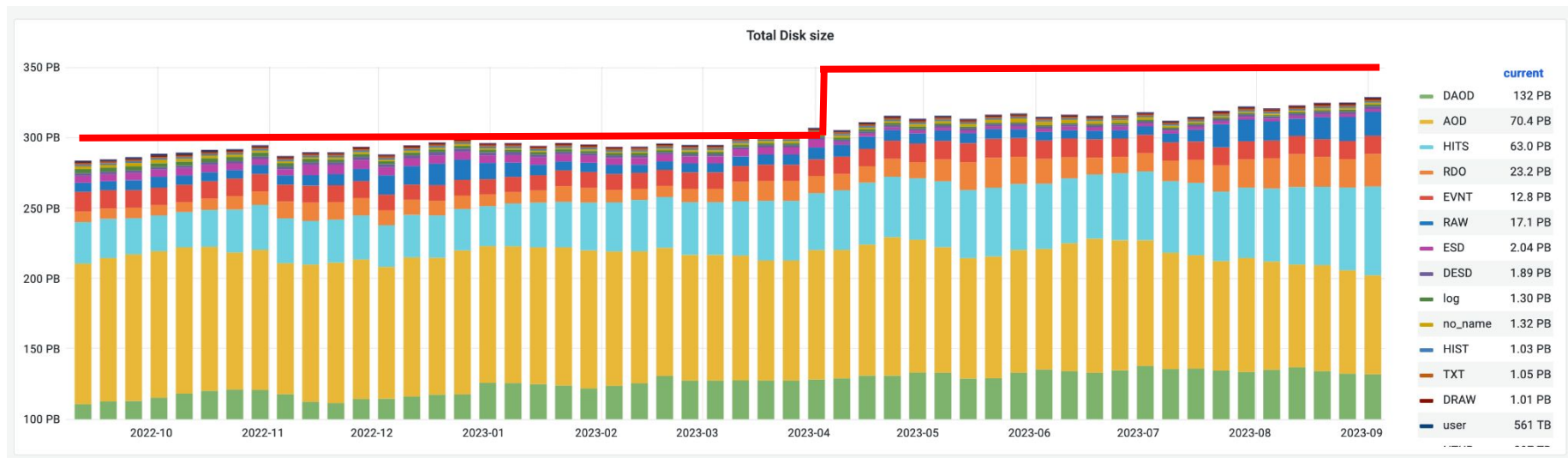


Utilizzo risorse ATLAS (settembre 2022 - settembre 2023) : CPU



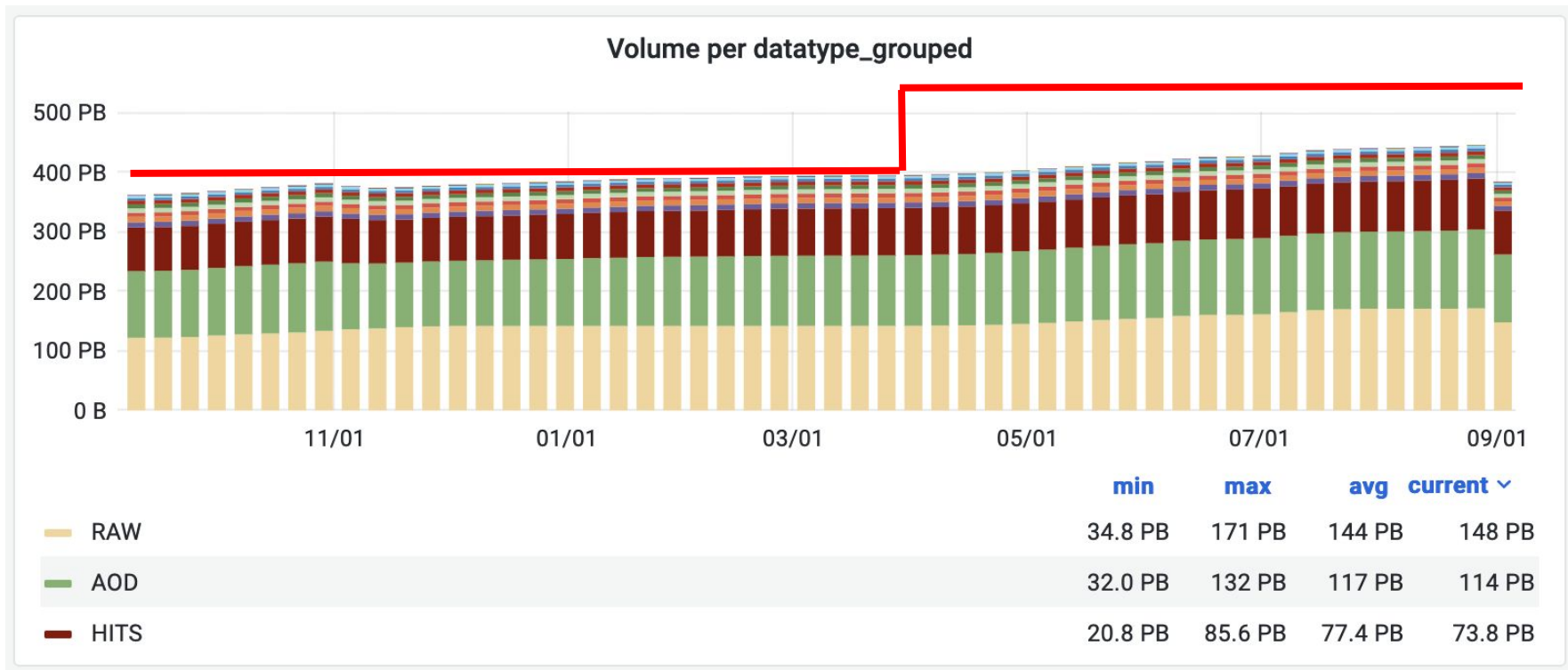
- ❑ Significativo over-pledge da siti grid (~ 30 % del pledge)
- ❑ Significativo contributo da risorse HPC opportunistiche (45 % del pledge, VEGA)

Utilizzo risorse ATLAS (settembre 2022 - settembre 2023) : disco



- Disco occupato ~ 329 PB (~700 M files) su 350 PB (pledge): principalmente dati derivati (DAOD)

Utilizzo risorse ATLAS (settembre 2022 - settembre 2023) : tape

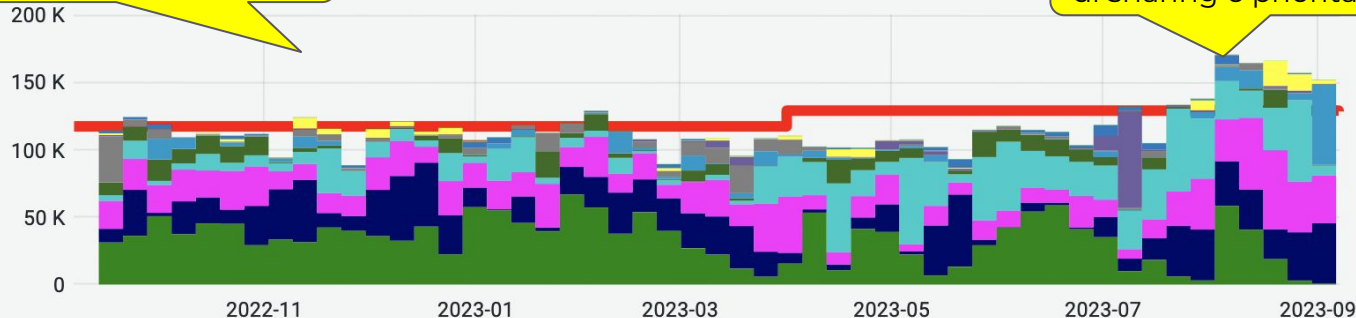


❑ Tape occupato ~ 448 PB (~300 M files) su 534 PB (pledge): mainly RAW data, AOD e HITS

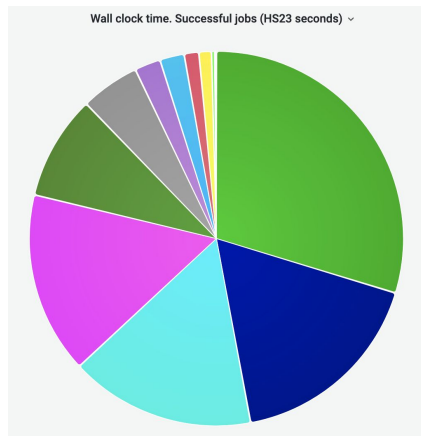
Attività ATLAS al T1 : CPU (settembre 2022-settembre 2023)

- Pledge 121 kHS23
- Average CNAF = 115 kHS23

Slots of Running jobs (HS23)

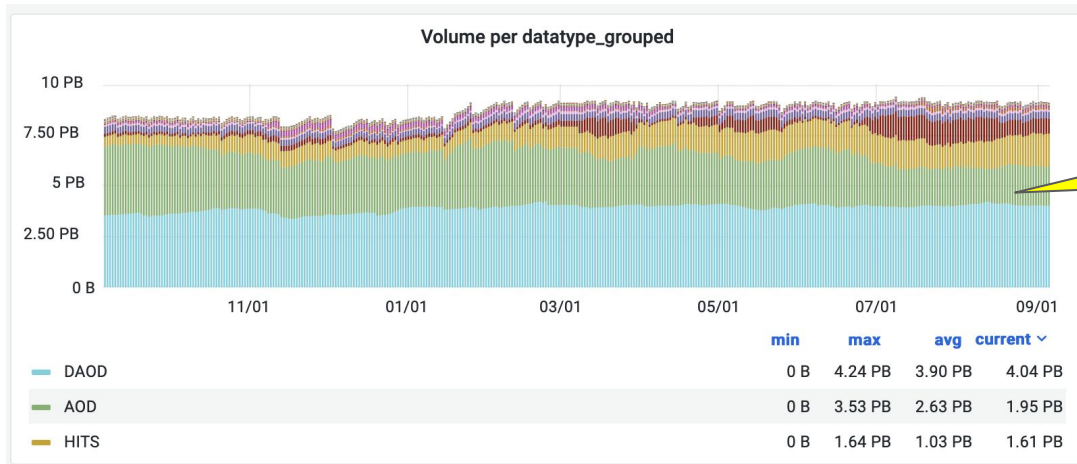


- ❑ Breakdown attività
 - ❑ produzione di MC (generazione, simulazione e ricostruzione)
 - ❑ preparazione dati secondari per l'analisi (group production),
 - ❑ analisi (user analysis)



	Value	Percent
Group Production	1.02 Tri	30%
MC Reconstruction	595 Bil	17%
MC Simulation Full	548 Bil	16%
User Analysis	537 Bil	16%
MC Event Generation	307 Bil	9%
Group Analysis	173 Bil	5%
MC Merge	74.7 Bil	2%
MC Simulation Fast	69.9 Bil	2%
MC Resimulation	40.8 Bil	1%

Attività al T1 : occupazione disco e tape

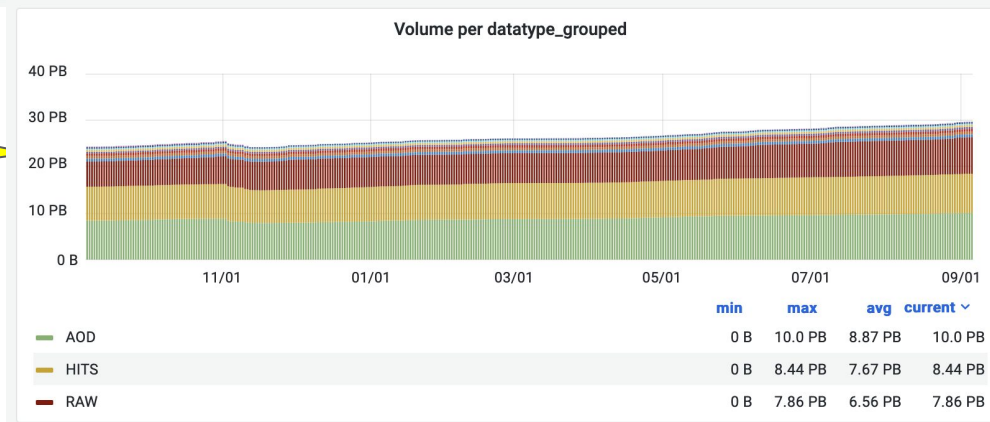


Disk

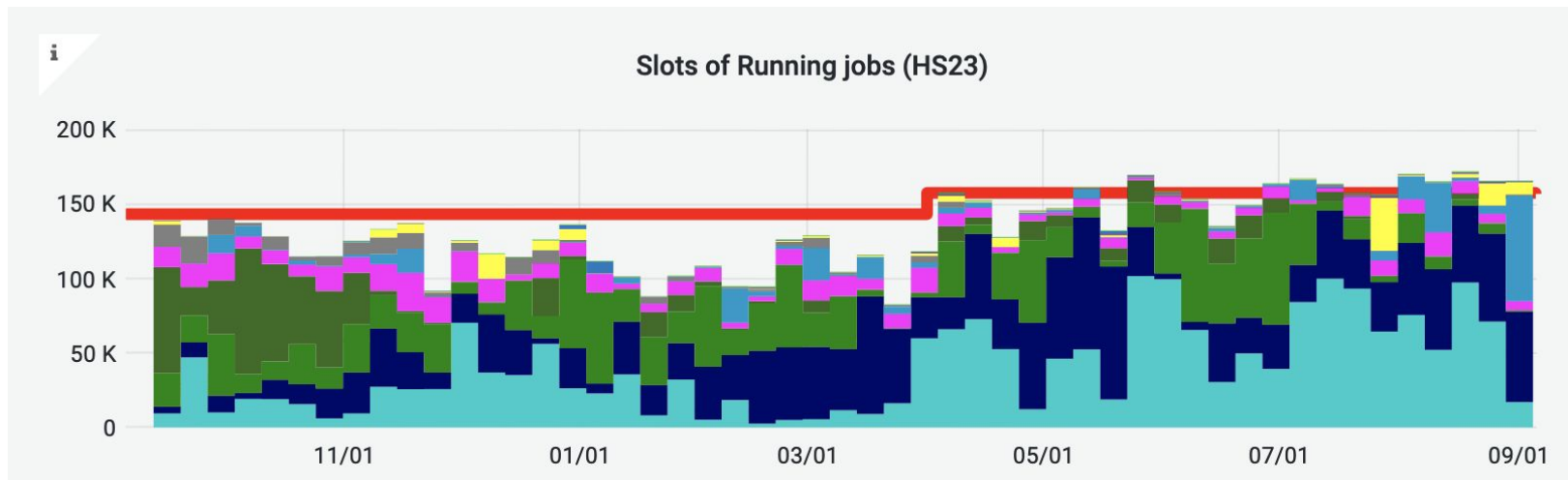
9.50 PB occupato (10.21 PB disponibile)

28.13 PB occupato (31.77 PB disponibile)

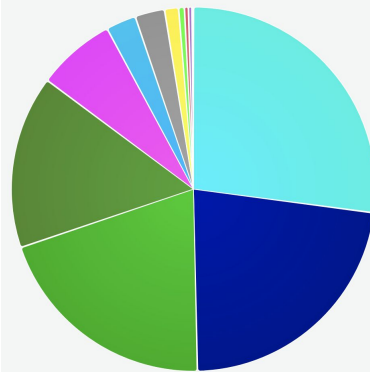
Tape



Attività ATLAS ai T2 : CPU (settembre 2022-settembre 2023)



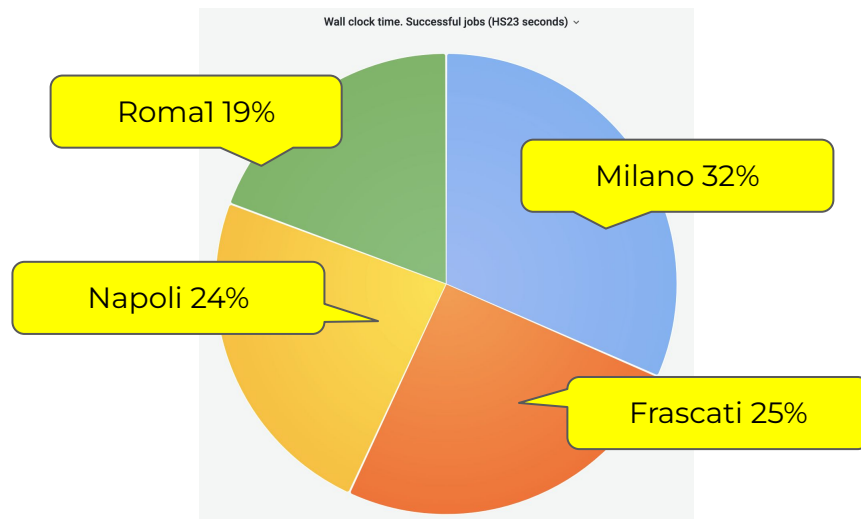
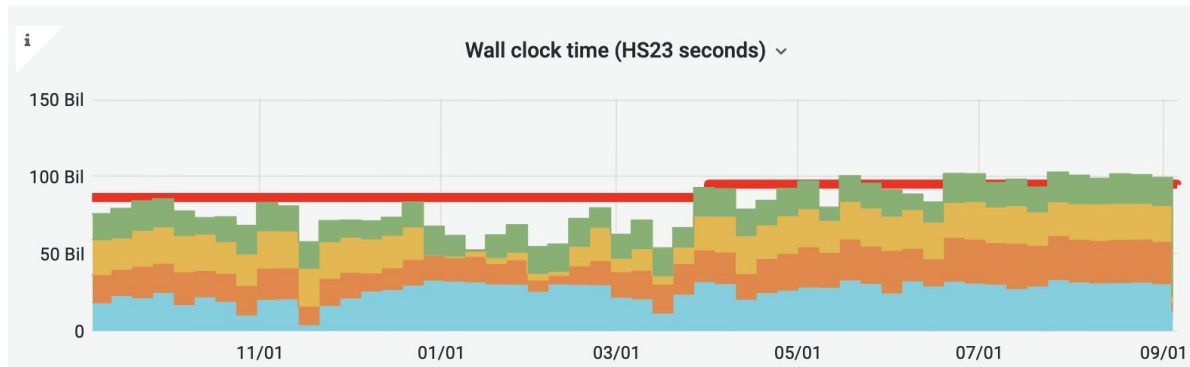
- ❑ Tipico share T2 :
 - ❑ MC (generation, simulation e reconstruction)
 - ❑ produzione di dati secondari (group production) e analisi utenti



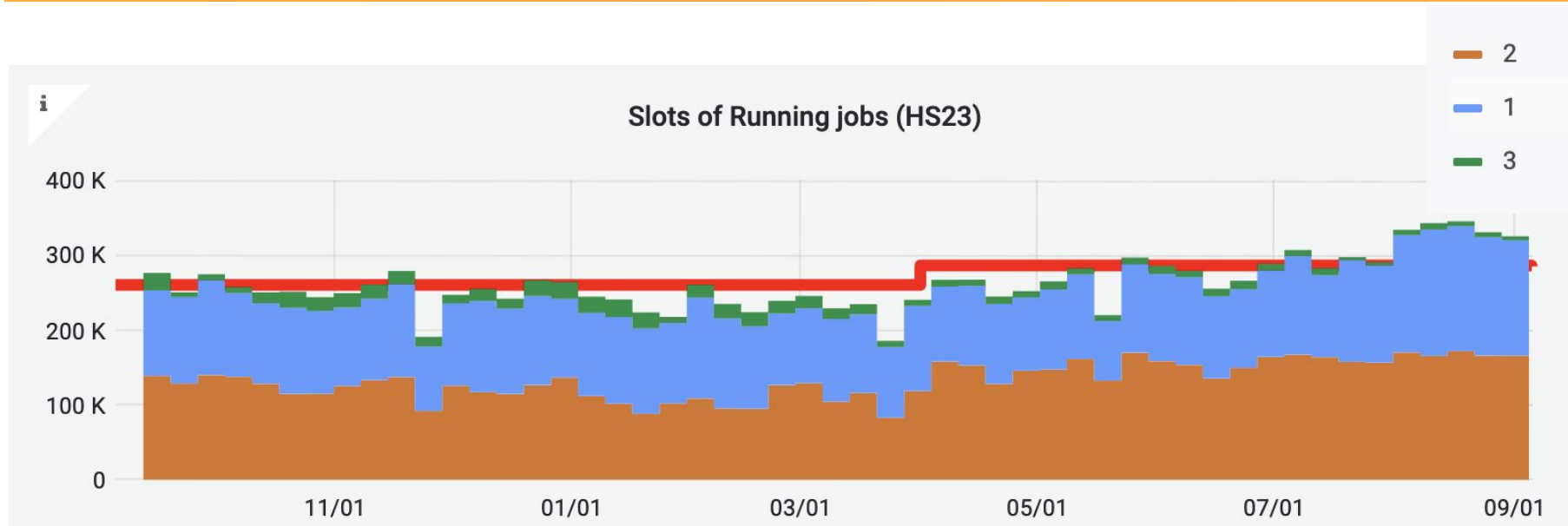
	Value	Percent
MC Simulation Full	1.06 Tri	27%
MC Reconstruction	883 Bil	23%
Group Production	787 Bil	20%
MC Event Generation	600 Bil	15%
User Analysis	270 Bil	7%
MC Simulation Fast	102 Bil	3%
Group Analysis	100 Bil	3%
Data Processing	45.6 Bil	1%

Attività ATLAS ai T2 : CPU

- ❑ In generale buon comportamento dei T2 italiani per l'ultimo anno di attività
 - ❑ reliability/availability ~ 95 %
 - ❑ Efficienza ~ 95%
 - ❑ Valori in linea con i migliori siti ATLAS
- ❑ Leggero under-pledge (~12%) sulla CPU fornita (average pledge 149 kHS23, computing power fornita 132 kHS23)
 - ❑ problemi di condizionamento in alcuni siti
- ❑ Tutti i siti sono ora up&running :
 - ❑ Ottima performance del sito di Milano che sta lavorando su tutti i workflows (ammodernamento apparati di rete)



CPU performance summary (all sites)



Globalmente leggera sofferenza dei siti italiani

- ❑ 271 kHS23 pledge, 260 kHS23 delivered: T2 e T1, contributo dei T3 (12 kHS23, ~ 5%) leggermente sotto la norma.
- ❑ Overall efficiency ~ 95% (in linea con l'efficienza della media dei siti ATLAS)

Situazione e disco e TAPE

	Disco occupato (PB) (dati aprile 2023)	Disco allocato (PB) (dati aprile 2023)	Teorici
Frascati	1.76	2.02	3.15
Milano	1.66	1.98	2.3
Napoli	4.14	4.59	4.15
Roma	1.59	1.75	2.45
tot.	9.15	10.34	12.05

Disco gara 2022 sta
apparendo online

- Pledge 2022 : 9.94 PB
- Pledge 2023 : 11.76 PB

	Disco occupato (PB) (dati giugno 2023)	Disco allocato (PB) (dati giugno 2023)	Tape (PB) (dati giugno 2023)
CNAF	9.54	10.21	31.77

- Pledge 2022 (disco): 10.44 PB
- Pledge 2023 (disco): 12.24 PB
- Pledge 2022 (tape): 24.48 PB
- Pledge 2023 (tape): 31.7 PB

Contributo italiano (settembre 2022 - settembre 2023)

Contributo italiano alle attività di computing di ATLAS :

- ❑ CPU : overall ~ 6% wallclock time totale in HS23-sec (T1+T2+T3, solo risorse grid, T0 escluso)
 - ❑ leggermente sotto il contributo nominale per via del consistente overpledge di alcune cloud + contributo delle risorse HPC
- ❑ Disco occupato (datadisk, calibdisk, groupdisk e localgroupdisk) : overall ~6% (escludendo T0)
 - ❑ leggermente sotto share alcune risorse ancora da mettere in linea
- ❑ Tape (T1 only) : ~9%
 - ❑ perfettamente in linea con lo share (9%)

ATLAS		2022			2023		2024		
		C-RSG recomm.	Pledged	Used	C-RSG recomm.	Pledged	Request 2024 req. /2023 C-RSG	C-RSG recomm.	
CPU	Tier-0	740	544	512	740	740	936	126%	936
	Tier-1	1300	1349	1545	1430	1520	1516	106%	1516
	Tier-2	1588	1616	3151	1747	1841	1852	106%	1852
	HLT	n/a	n/a	311	n/a	n/a	n/a	n/a	n/a
	Total	3628	3509	5519	3917	4101	4304	110%	4304
	<i>Others</i>			2161					
Disk	Tier-0	32.0	32.0	33.0	40.0	40.0	49.0	123%	49.0
	Tier-1	116.0	130.0	139.0	136.0	150.5	163.0	120%	163.0
	Tier-2	142.0	142.0	141.0	168.0	160.0	200.0	119%	200.0
	Total	290.0	304.0	313.0	344.0	350.5	412.0	120%	412.0
Tape	Tier-0	120.0	120.0	102.0	174.0	174.0	207.0	119%	207.0
	Tier-1	272.0	280.0	281.0	353.0	360.3	452.0	128%	452.0
	Total	392.0	400.0	383.0	527.0	534.3	659.0	125%	659.0

ATLAS C-RSC recommended 2024

	ATLAS 2024 pledge	Italy share 2024	ATLASIT 2024 pledge	Richieste 2024	Richieste 2024 (k€)	ATLAS 2023 pledge	Italy share 2023	ATLASIT 2023 pledge
CPU T1 (HS06)	1516000	0,09	136440	7740	77400	1430000	0,09	128700
Disco T1 (TB)	163000	0,09	14670	2430	291600	136000	0,09	12240
Tape T1 (TB)	452000	0,09	40680	8910	89100	353000	0,09	31770
Totale (k€)	458100							

- ❑ Richieste calcolate rispetto al pledge del T1 (non rispetto alle risorse effettive attualmente disponibili)
- ❑ Mantenuto lo stesso share per ATLAS-IT (9%)
- ❑ Obsolescenza non inclusa
- ❑ Fattori di conversione : 120 €/TBn (disco), 10 €/HS23 (cpu), 10 €/TBn (tape)

ATLAS C-RSG recommended 2024

	ATLAS 2024 pledge	Italy share 2024	ATLASIT 2024 pledge	ATLAS 2023 pledge	Italy share 2023	ATLASIT 2023 pledge
CPU T2 (HS06)	1852000	0,09	166680	1747000	0,09	157230
Disco T2 (TB)	200000	0,07	14000	168000	0,07	11760

- ❑ Fattori di conversione : 120 €/TBn (disco), 10 €/HS23 (cpu), 10 €/TBn (tape)
- ❑ Mantenuto lo stesso share ATLAS-IT (9% CPU e 7% disco)

Richieste dettagliate T2 : CPU

CPU (HS06)	Frascati	Milano	Napoli	Roma1	Totale	Pledged
Totale CPU nei siti per il 2023 (HS06)	38312	39708	41886	40028	159934	157230
Obsolescenza 2024 (HS06)	10004 (0/10004)	5878 (5878/0)	0	0	15882 (5878/10004)	
Delta 2024 (HS06)	2500	2500	2500	2500	10000	
Totale richieste 2024 (HS06)	12504	8378	2500	2500	25882	
Totale CPU nei siti per il 2024 (HS06)	40812	42208	44386	42528	169934	166680

Previsione spesa (€)	125040	83780	25000	25000	258820
----------------------	--------	-------	-------	-------	--------

Richieste dettagliate T2 : disco

DISCO (TB)	Frascati	Milano	Napoli	Roma1	Totale	Pledged
Totale disco nei siti per il 2023 (TB)	3148	2309	4147	2444	12048	11760
Obsolescenza 2024 (TB)	0	789	0	864	1653	
Delta 2024 (TB)	500	500	500	500	2000	
Totale richiesta 2024 (TB)	500	1289	500	1364	3653	
Totale disco nei siti per il 2024 (TB)	3648	2809	4647	2944	14048	14000

Previsione spesa (€)	60000	154680	60000	163680	438360
----------------------	-------	--------	-------	--------	--------

Richieste finanziarie T2

	Frascati	Milano	Napoli	Roma1	Totale
Spesa CPU (€)	125040	83780	25000	25000	258820
Spesa DISCO (€)	60000	154680	60000	163680	438360
Totale (€)	185040	238460	85000	188680	697180
Overhead server (CSN1)	12952	16692	5950	13207	48802
Overhead rete (DATA CLOUD)	10502	12760	4500	9684	37447
Totale (overheads inclusi) (€)	208494	267912	95450	211571	783429

- ❑ Overheads :
 - ❑ quota server (su CSN1): 7% del totale CPU + disco
 - ❑ quota rete (su Datacloud): 6% CPU + 5% disco
- ❑ Fattori di conversione : 120 €/TBn (disco), 10 €/HS06 (cpu), 10 €/TBn (tape)

Backup

Richieste dettagliate T2 : CPU

CPU (HS06)	Frascati	Milano	Napoli	Roma1	Totale	Pledged
Totale CPU nei siti per il 2023 (HS06)	38312	39708	41886	40028	159934	157230
Obsolescenza 2024 (HS06)	10004 (0/10004)	5878 (5878/0)	0	0	15882 (5878/10004)	
Delta 2024 (HS06)	2500	2500	2500	2500	10000	
Totale richieste 2024 (HS06)	12504	8378	2500	2500	25882	
Totale CPU nei siti per il 2024 (HS06)	40812	42208	44386	42528	169934	166680

Previsione spesa (€)	125040	83780	25000	25000	258820
----------------------	--------	-------	-------	-------	--------

Totale risorse PNRR da acquisire nel 2023 (HS06)	51194	51194	121624	25605	249617	
Pledge 2023 promesso ad ATLAS	19866	11426	3000	17521	51813	
Risorse PNRR 2023 non ATLAS (HS06)	31328	39768	118624	8084		
Avanzo dopo copertura pledge 2024 ATLAS	18824	31390	116124	5584		

Richieste dettagliate T2 : disco

DISCO (TB)	Frascati	Milano	Napoli	Roma1	Totale	Pledged
Totale disco nei siti per il 2023 (TB)	3148	2309	4147	2444	12048	11760
Obsolescenza 2024 (TB)	0	789	0	864	1653	
Delta 2024 (TB)	500	500	500	500	2000	
Totale richiesta 2024 (TB)	500	1289	500	1364	3653	
Totale disco nei siti per il 2024 (TB)	3648	2809	4647	2944	14048	14000

Previsione spesa (€)	60000	154680	60000	163680	438360
----------------------	-------	--------	-------	--------	--------

Lordi : -30% netti

Totale risorse da acquisire PNRR 2023 (TB)	2379	2230	5655	1200	11464	
Pledge 2023 promesso ad ATLAS	0	0	0	0	0	
Risorse PNRR 2023 non ATLAS (TB)	2379	2230	5655	1200	11464	
Avanzo dopo copertura pledge 2024 ATLAS	1879	941	5155	-164		

Stato risorse al 2023

- ❑ Storage:
 - ❑ Gara comune disco 2022 : disco arrivato nei siti, online nelle prossime settimane

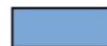
- ❑ CPU :
 - ❑ Risorse 2023 (PNRR) arrivate nei siti

Piano di lavoro

Data sample	Activity	2023		2024				2025			
		Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Run 2 Data	DAOD Production and User Analysis										
	DAOD_PHYS Production										
Run 2 MC	New Production for Ongoing Analyses										
	DAOD_PHYS Production										
Run 3 Data	Tier-0 Reconstruction										
	Partial Reprocessing										
	Full Reprocessing										
	Delayed stream Reconstruction										
	DAOD Production and User Analysis										
Run 3 MC	Generation										
	Simulation										
	Reconstruction										
	Re-Reconstruction										
	DAOD Production and User Analysis										
Upgrade MC	Production and Analysis										
Heavy Ions	First pass Reconstruction										



Full steam



Partial



Minimal

Summary



- Evgen will be a non-negligible fraction of our processing in HL-LHC
- Investing in improvements is certainly valuable
 - Significant gains for CPU-based running are in the pipeline
- Several groups are developing GPU-based implementations
 - Mostly focused on LO generation ([NLO in time for HL-LHC?](#))
 - Still very early days
- We (ATLAS) are involved at several levels in these developments
 - Expecting to be beta testers / early adopters for some
- We might expect to be able to make a serious statement in one or the other direction in time for the Software and Computing HL-LHC TDR in 2024/5

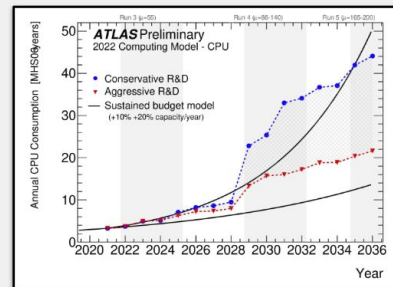
Thanks for help and input: Enrico Bothmann, James Catmore, Taylor Childers, Ale Di Girolamo, Matthew Gignac, Chris Gutsche, Stefan Hoeche, Walter Hopkins, Charles Leggett, Josh McFayden, Emanuele Re, Stefan Roiser, Nick Styles



Why Use Accelerator Resources?



- Because they're here / there
 - Large fraction of computing power comes from GPUs in most HPCs and increasing number of HTCs
 - Ixplus-gpu, grid nodes, swan instances, cloud, under your desk and more
- Because they're energy efficient
 - Computing sites are increasingly limited by the amount of power they consume
 - GPUs are more power efficient than CPUs per FLOP
- Because we always need more computing power
 - But is it cheaper or more efficient to get there with GPUs or FPGAs instead of traditional CPUs?
- Because everyone else is doing it
 - ALICE, LHCb, CMS, ChatGPT, cryptobros





Summary



ATLAS has many different avenues exploring the use of accelerators for both the online EF Tracking for offline reconstruction and analysis.

We have not yet decided whether we **NEED** GPUs/FPGAs for TDAQ / EF.

- technology decision will be in 2025
- many results of this investigation will be leveraged by offline software to take advantage of existing GPU resources (grid, cloud, HPC, laptop, etc)
 - we **WILL** use GPUs in offline reco and analysis

Usage of Machine Learning techniques in both online and offline continues to increase.

- can experience tremendous benefits from accelerators, especially in training of models
- working on finding efficient ways of integrating ML as a service into our workflows

Impatto nuovo analysis model

- We were also asked to comment on the uptake of our new analysis model and its effect on future grid resources
- The good news: the monitoring shows people are moving to use DAOD_PHYS!

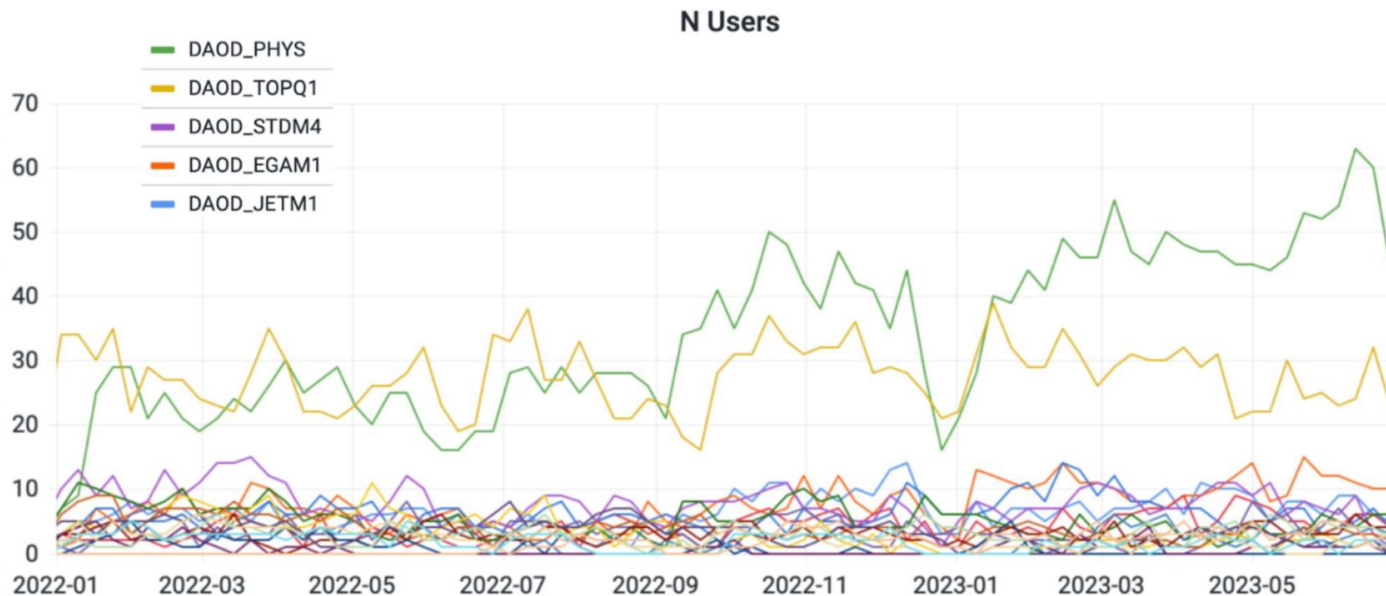


Figure 3: Number of unique users in a given week accessing all the various derivation formats.

