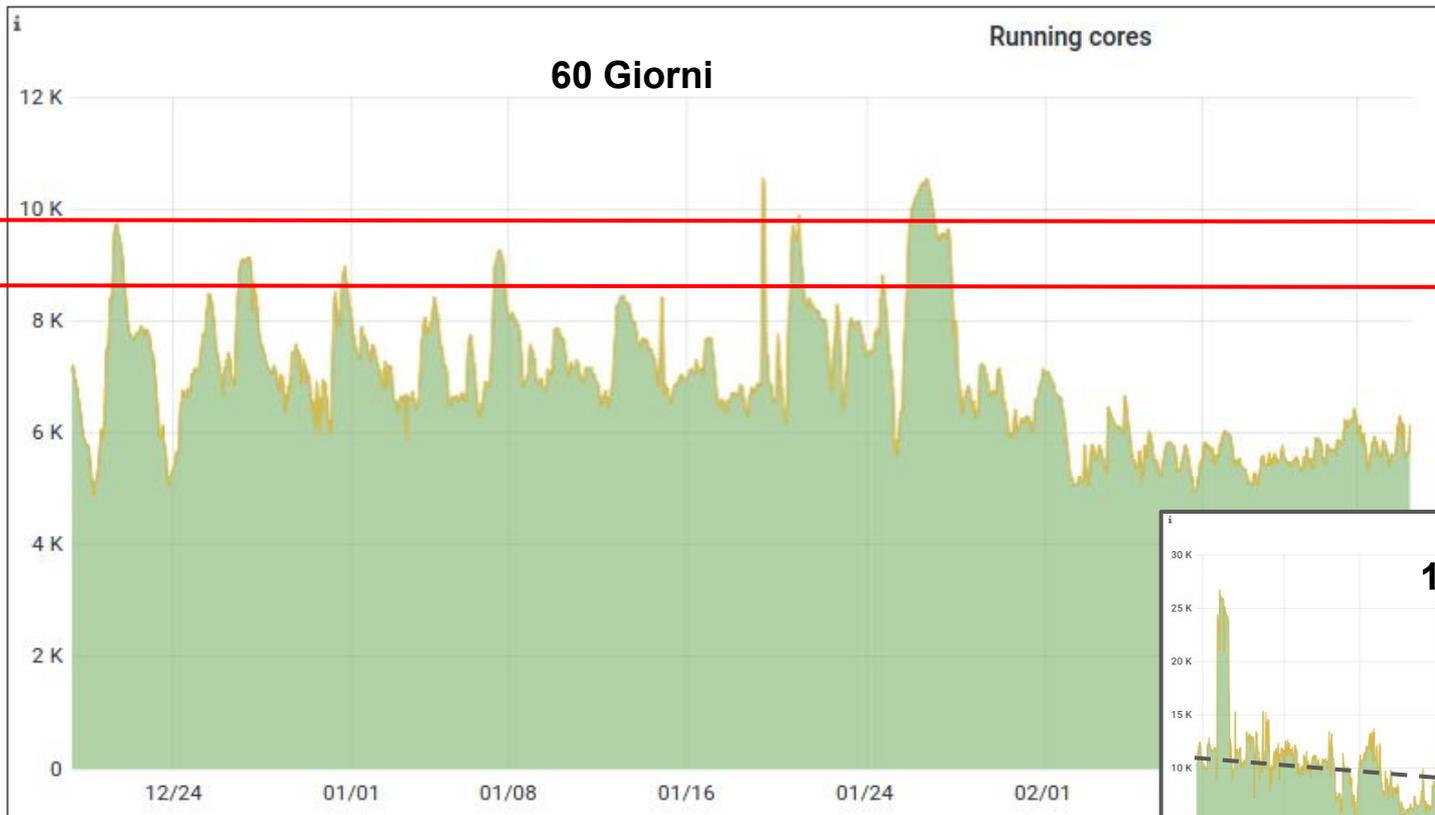


CMS - CdG T1

Daniele Spiga
INFN-PG



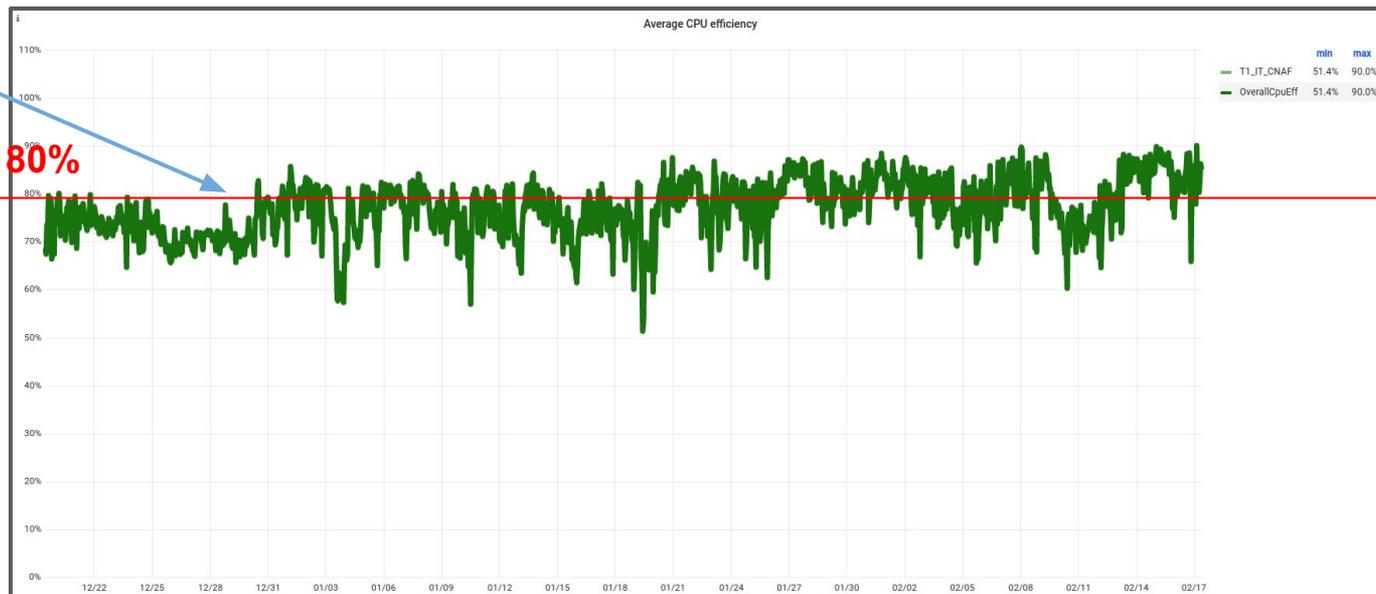
Utilizzo Risorse Tier1





Efficienza di CPU @CNAF

Efficienza
@CNAF





Transizione a Token @CMS: preambolo

Il mese scorso CMS ha iniziato a tracciare sistematicamente i requisiti per la migrazione a token di **tutti i propri servizi**

- Data/Compute, CMSWEB, CRAB etc...

Tuttavia CMS (come tutti gli esperimenti di LHC) è ora sotto pressione esterna (OSG) per migrare presto (ieri :)). OSG dalla versione 3.6 non supporta più l'autenticazione x509

- Implicazioni per HTCondor, trasferimenti di dati. XRootD contiene il proprio stack x509
- Executive: la fine del ciclo di vita della versione 3.5 dell'OSG è stata fissata alla fine di questo mese ma **OSG sta valutando la possibilità di posticipare l'OSG 3.5 EOL al 1° maggio**



Token@CMS: Focus su HTCondor Global Pool

- Sia l'Integration TestBed che il Pool di produzione sono abilitati al supporto di ID e JWT Token

```
SCITOKENS /<TOKEN ISSUER>,<TOKEN SUBJECT>/ <USERNAME>  
Example:  
SCITOKENS  
"https://\cms-auth.web.cern.ch/,490a9a36-0268-4070-8813-65af  
031be5a3" cmspilot  
SCITOKENS  
"https://\cms-auth.web.cern.ch/,bad55f4e-602c-4e8d-a5c5-bd8f  
fb762113" cmspilot
```

- Per i siti implica: HTCondor CE >5.x e Batch >9.x... Migrazione Work in progress → **al CNAF Già tutto OK**
 - Per i siti con ARC_CE la storia è diversa (*)
- CMS ha prosposto quindi di partire da US T1+ T2 passare poi agli altri Tier1 EU e poi il resto di HTCondor T2 (poi ARC)
 - Nulla vieta che se un sito è pronto può migrare → i.e. **Al CNAF stiamo esattamente su questo pattern**

(*) Gli ARC CE supportano i token WLCG. Tuttavia, i lavori della griglia HTCondor (la base dei nostri lavori pilota GlideinWMS) non possono ancora impiegarli. L'interazione dei piloti CMS con gli ARC CE tramite token non è quindi ancora disponibile, condizionata all'installazione di una nuova versione di HTCondor nelle nostre fabbriche GlideinWMS.



Token@CMS: alcune NOTE

Evidentemente ci sono alcuni aspetti aperti con impatto più o meno diretto sulla VO di CMS

- **SAM Test (CE):** Vanno abilitati i SAM test il cui stato è : ready to go for HTCondor-CEs “**but waiting on conclusion of subject/capability discussion**”
- **Payload token based:** il token deve vivere per la durata del jobs (meccanismi di “renewal”)
 - Attualmente stiamo parlando di token dei pilots (JWT -scitokens) e IDToken dei demoni HTCondor



Token@CMS Processo di transizione

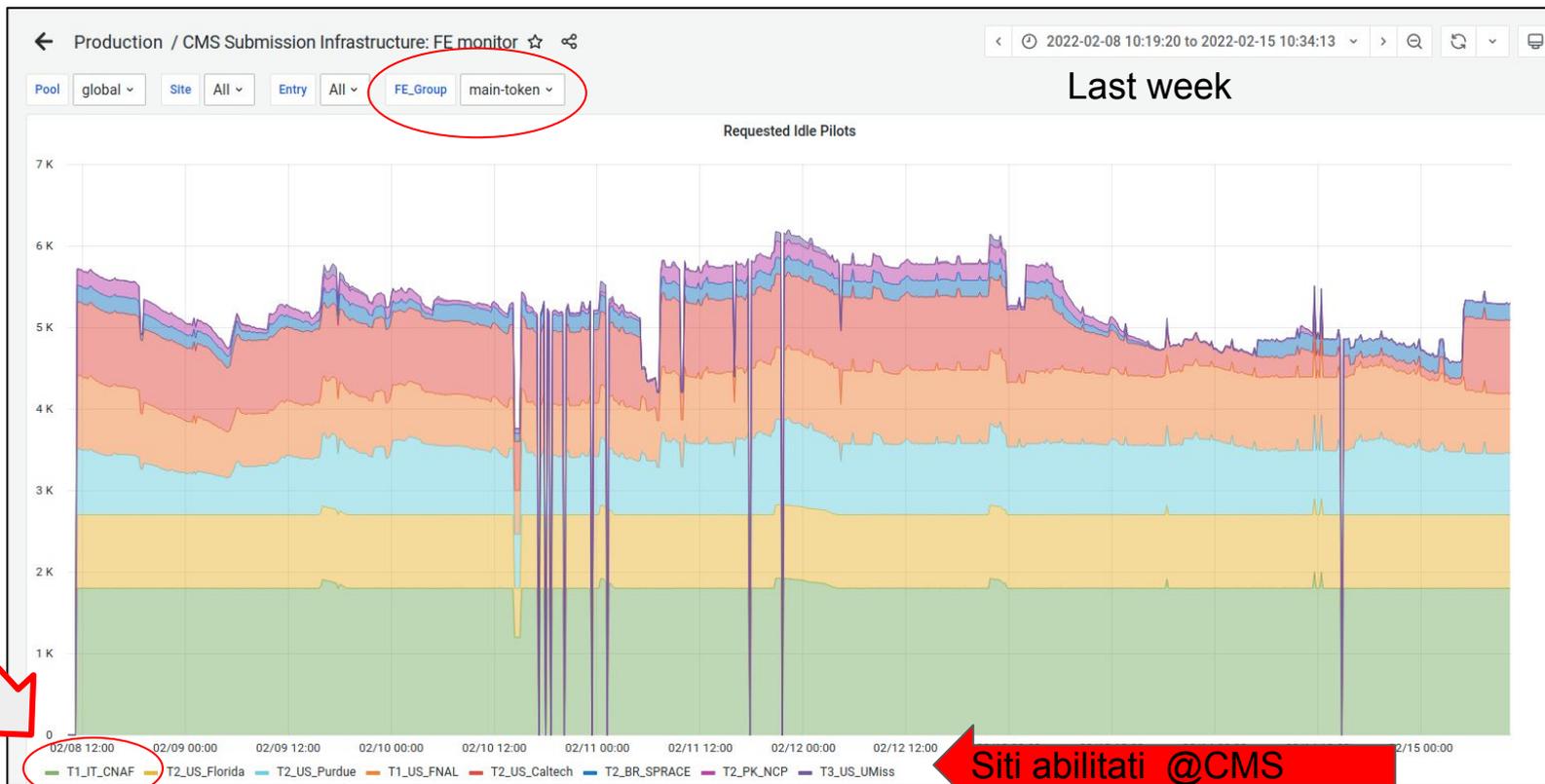
Dobbiamo garantire una **transizione graduale all'utilizzo dei token** per questo CMS ha abilitato il Front-End del Global Pool GlideinWMS per utilizzare contemporaneamente autenticazioni basate su GSI e Scitoken.

- Messo in produzione glideinWMS 3.7.6 il 27 Gennaio

Tecnicamente, **la distribuzione procederà separando i siti in gruppi FE.**

- Un nuovo gruppo FE gestirà l'autenticazione sia con token che proxy, in quest'ordine, con il meccanismo di fallback.
 - **Questo sarà il nuovo gruppo principale, ad es. "Main-token".**
- I siti stanno venendo migrati dall'attuale gruppo "main" (che rimarrà solo proxy) poiché vengono testati per accettare correttamente i progetti pilota basati su token CMS.

Esperienza: La Migrazione è iniziata





NOTA: Token anche su Marconi 100

Su M100 (a differenza di Marconi A2) **non usiamo NESSUN CE dedicato**. La strategia è quella di accendere i pilots di CMS sottoforma di slurm job e teniamo la pressione opportuna da Login Node

- Modello che con Stefano stiamo validando per l'integrazione di Leonardo opportunistic

NON avendo CE **tutto quello detto prima non applica**

- Oggi iniziamo la migrazione alla versione di HTCondor per PPC da 8.X a 9.X
- Questo è funzionale per i test che stiamo facendo con Stefano!

M100: Validazione di fisica stato e progressi

Shc

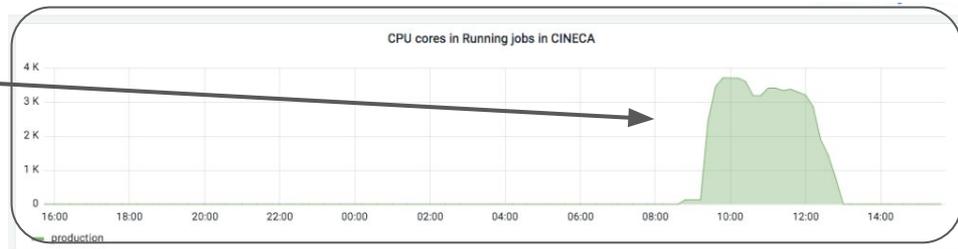
D	A	L	workflow ^	status ^	type ^	priority ^	queue injection ^	job progress ^	event progress ^	lumi progress ^	failure rate ^	Estimated Completion ^	cool off ^
			pdmvserv_RVCMSSW_11_2_0QCD_FlatPt_15_3000HS_14_PPC_220113_171249_3807	running-closed	TaskChain	500000	100.0%	91.7%	6.0%	6.0%	0.0%	10 h 51 m	3
			pdmvserv_RVCMSSW_11_2_0ZEE_14_PPC_220113_171249_3831	running-closed	TaskChain	500000	100.0%	93.0%	8.1%	8.1%	0.0%	9 h 3 m	0
			pdmvserv_RVCMSSW_11_2_0SingleMuPt100_PPC_220113_170901_9586	running-closed	TaskChain	500000	100.0%	97.8%	0.0%	0.0%	0.0%	2 h 39 m	0
			pdmvserv_RVCMSSW_11_2_0DisplacedSUSY_stopToB_M_800_500mm_14_PPC_220113_170857_9026	running-closed	TaskChain	500000	100.0%	97.8%	0.0%	0.0%	0.0%	2 h 40 m	0
			pdmvserv_RVCMSSW_11_2_0QCD_Pt_1800_2400_14_PPC_220113_170857_2448	running-closed	TaskChain	500000	100.0%	0.0%	0.0%	0.0%	0.0%	N/A	0
			pdmvserv_RVCMSSW_11_2_0SingleMuPt10_PPC_220113_170856_169	running-closed	TaskChain	500000	100.0%	97.8%	0.0%	0.0%	0.0%	2 h 39 m	0
			pdmvserv_RVCMSSW_11_2_0ZMM_14_PPC_220113_170855_3697	running-closed	TaskChain	500000	100.0%	97.8%	0.0%	0.0%	0.0%	2 h 39 m	0
			pdmvserv_RVCMSSW_11_2_0zPTT_1500_14_PPC_220113_170855_6288	running-closed	TaskChain	500000	100.0%	91.1%	13.3%	13.3%	0.0%	11 h 40 m	0
			pdmvserv_RVCMSSW_11_2_0MinBias_14TeV_PPC_220113_170854_69	running-closed	TaskChain	500000	100.0%	78.6%	0.0%	0.0%	0.0%	32 h 40 m	0
			pdmvserv_RVCMSSW_11_2_0SingleMuPt1000_PPC_220113_170854_4460	running-closed	TaskChain	500000	100.0%	0.0%	0.0%	0.0%	0.0%	N/A	0
			pdmvserv_RVCMSSW_11_2_0TTbar_14TeV_PPC_220113_170854_8822	running-closed	TaskChain	500000	100.0%	0.0%	0.0%	0.0%	0.0%	N/A	0
			pdmvserv_RVCMSSW_11_2_0Upsilon1StoMuMu_14_PPC_220113_170854_6465	running-closed	TaskChain	500000	100.0%	98.4%	0.0%	0.0%	0.0%	1 h 59 m	0
			pdmvserv_RVCMSSW_11_2_0BuMixing_14_PPC_220113_170852_1082	running-closed	TaskChain	500000	100.0%	95.7%	0.0%	0.0%	0.0%	5 h 19 m	0
			pdmvserv_RVCMSSW_11_2_0H125GGluonfusion_14_PPC_220113_170852_6119	running-closed	TaskChain	500000	100.0%	94.6%	9.6%	9.6%	0.0%	6 h 50 m	0
			pdmvserv_RVCMSSW_11_2_0NuGun_PPC_220113_170852_9914	running-closed	TaskChain	500000	100.0%	98.9%	0.0%	0.0%	0.0%	1 h 19 m	0
						500000	100.0%	0.0%	0.0%	0.0%	0.0%	N/A	0
						500000	100.0%	97.2%	6.7%	6.7%	0.0%	3 h 29 m	0

Abbiamo effettuato la migrazione al **setup di produzione**, il Global Pool (eravamo su integration testbed)

- Eseguito il primo workflow ufficiale di Release Validation (Montecarlo)

Abbiamo aggiunto proxy xrootd (**Lucia/Storage**) per accesso alla Federazione AAA

- Abbiamo avviato il primo workflow di validazione (RelVal) sui dati



8.02.2021



M100: debugging GPU integration

Abbiamo validato il setup del pilot per l'esecuzione di CPU (only) e CPU+GPU jobs e funziona!

```
condor_status slot1@glidein_38410_1102677510@r221n16.m100.cineca.it -af SlotType cpus gpus memory
```

```
Partitionable 128 4 322140
```

Il nostro obiettivo però è anche di saper usare le slot solo CPU (worse case)

Questo non sta funzionando per due motivi

- Problema di fine tune di match making (stiamo migrando ad uno scenario dove la main expression, ovvero quella “hard-coded” sia minimale, e le estensioni siano gestite localmente (local match expression)
- Problema più articolato che stiamo debuggando con CMS+HTcondor devel ovvero per i job CPU-only il matchmaking avviene ma poi lo schedd non riesce a fare il claim :(..