

Stato del Computing in Belle II

Dr. Silvio Pardi

Phone

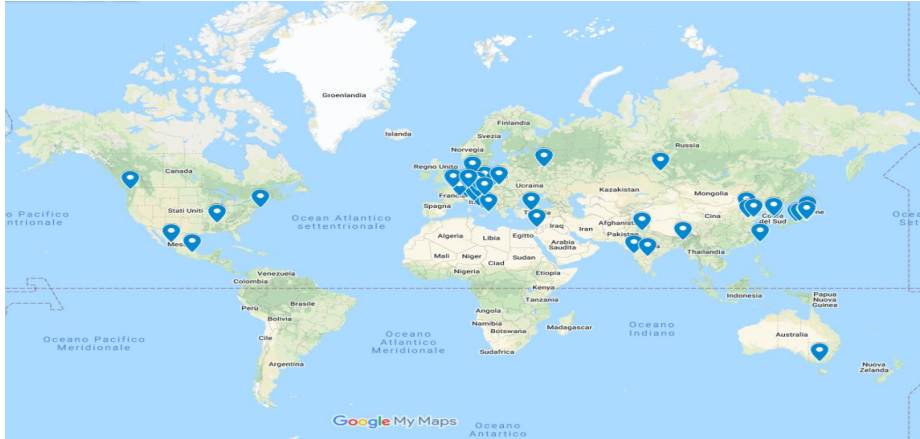
13 Giugno 2022

Siti di Belle II

19 Paesi

55 Siti attivi in DIRAC (5 sono Tier1 LHC)

5 Sistemi di tape



Australia Austria Canada Czech Republic China France Germany India
Israel ITALY Japan Mexico Poland Russia Slovenia South Korea Taiwan
Turkey USA

Dal 2020 stiamo raccogliendo in maniera sistematica le informazioni sulle risorse realmente fornite da ogni sito nel B2GM di ottobre.

Quest'anno abbiamo anticipato la raccolta delle informazioni a Giugno in modo da implementare il prima possibile le variazioni sul sistema DIRAC e migliorare l'utilizzo delle risorse

Attualmente la raccolta delle informazioni è in corso. Report previsto per Venerdì 3 Giugno al B2GM

Site Report 2022

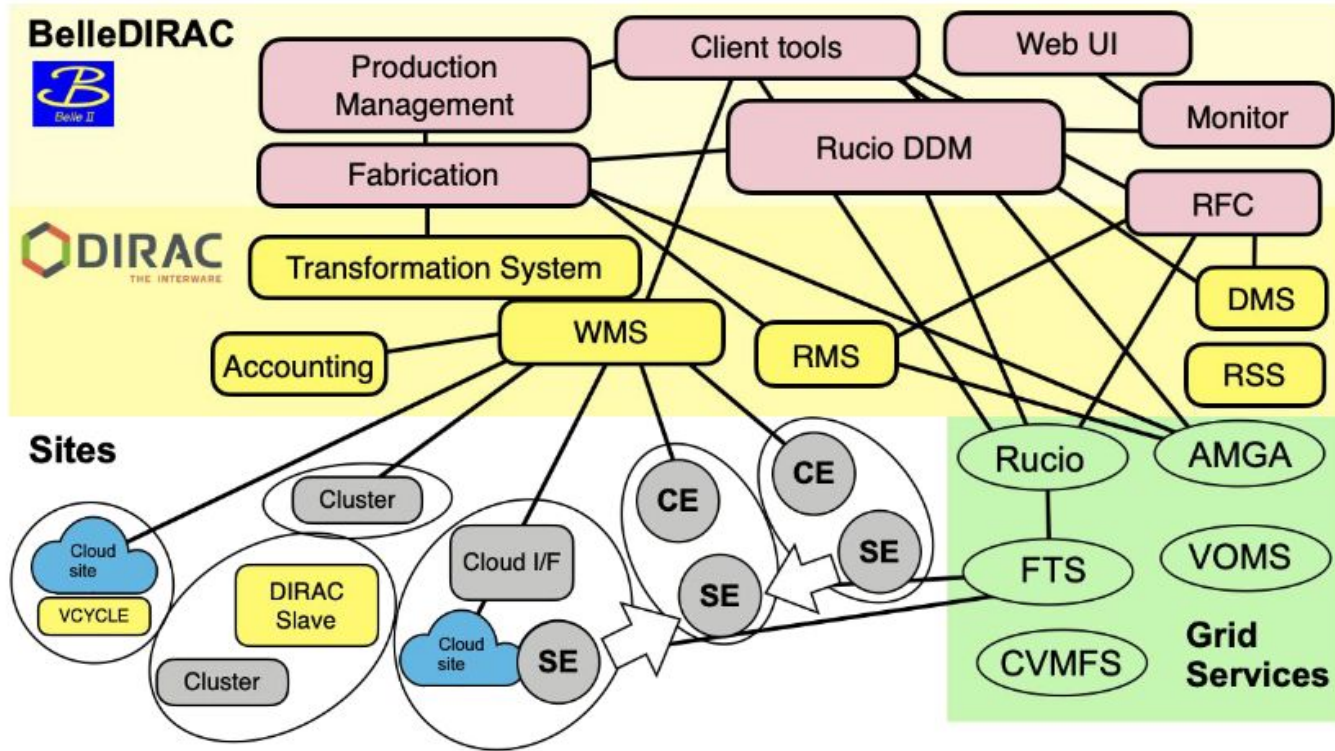
Resources	NOTE	CPU Deployed guaranteed (kHS06)	CPU Deployed guaranteed jobslots	CPU Opportunistic (kHS06)	CPU Opportunistic jobslots	Total CPU (kHS06)	Total Jobslots	Storage DISK	Tape
Production	Total Opportunistic CPU include the BNL core for calibration. CNAF opportunistic are estimated a 10% of declared	466	32k	386	32k	852	64k	15.5PB	12.4PB

Resources	NOTE	CPU Deployed guaranteed (kHS06)	CPU Deployed guaranteed jobslots	CPU Opportunistic (kHS06)	CPU Opportunistic jobslots	Total CPU (kHS06)	Total Jobslots	Storage DISK (TB)	Tape (TB)
Calib/Recalibration	DESY and BNL	36,7	3.1k	0	0	36,7	3.1k	500TB	600TB

Resources per Country 2022

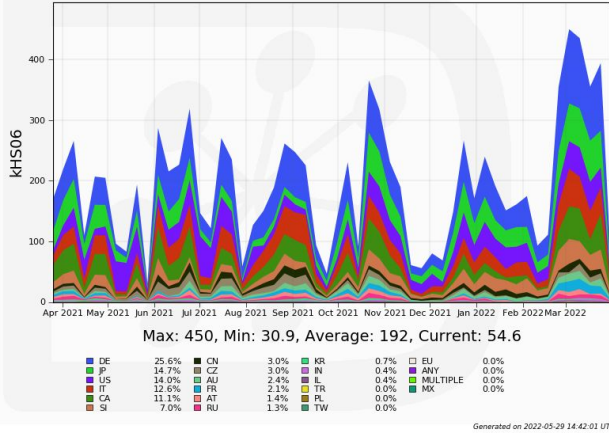
RESOURCES FOR PRODUCTION									
	CPU Deployed guaranteed (kHS06)	CPU Deployed guaranteed jobslots	CPU Deployed Opportunistic (kHS06)	CPU Deployed Opportunistic jobslots	Total Deployed CPU (kHS06)	Total Jobslots	Storage (TB)	TAPE (TB)	Notes
Australia	18	900	10	1000	28	1900	50	0	
Austria	4,8	480	0	0	4,8	480	250	0	
Canada	80	4000	20	1000	100	5000	600	100	N.B. There is not tape but 100TB for RAW Data
China	15	856	0	0	15	856	260	0	
France	11,8	890	2,2	180	14	1070	403	179,22	
Germany	78,02	6424	102,5	8146	180,52	14570	4070	1830	
India	19,58	1100	5,14	161	24,72	1261	0	0	
Israel	2,7	168	0	0	2,7	168	60	0	
Italy	55	5050	95,6	8849	150,6	13899	1772	650	CNAF reported 427 Opportunistic, for this computation considered 10%
Japan	60,3	3256	43,8	2526	104,1	5782	3468	5550	
Korea	0,32	36	1	56	1,32	92	0	0	
Mexico	2,4	144	0	0	2,4	144	0	0	
Poland	2	200	0	0	2	200	10	0	
Russia	13	1156	5	500	18	1656	0	0	
Slovenia	22,5	1800	16	1200	38,5	3000	1210	0	
South Korea	8,576	544	0	0	8,576	544	100	0	
Taiwan	18,33	410	0	0	18,33	410	791,95	0	
The Czech Republic	4,1	400	12,3	1200	16,4	1600	100	0	
Turkey	0,938	128	0	0	0,938	128	130	0	
USA	49,4	4300	73	7000	122,4	11300	2312	4100	Calibration CPU are included also among the opportunistic resources
Production TOT	467	32.242	387	31.818	853	64.060	15.587	12.409	
RESOURCES FOR CALIBRATION									
Germany	3	100			3	100		600	at DESY
USA	33,7	3000			33,7	3000	500		at BNL
Calibration TOT	37	3.100	-	-	37	3.100	500	600	

DIRAC Infrastructure

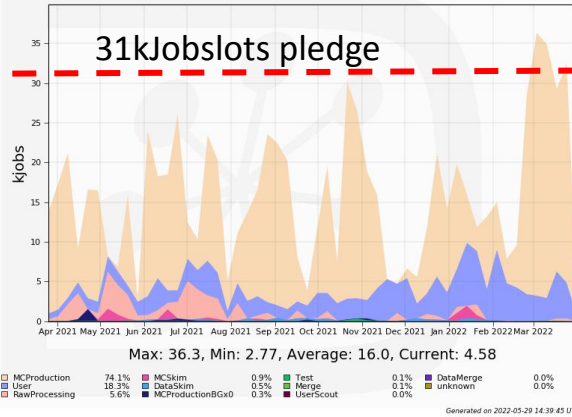


Overall activity in 2021 JFY

Normalized CPU usage by Country
53 Weeks from Week 12 of 2021 to Week 13 of 2022



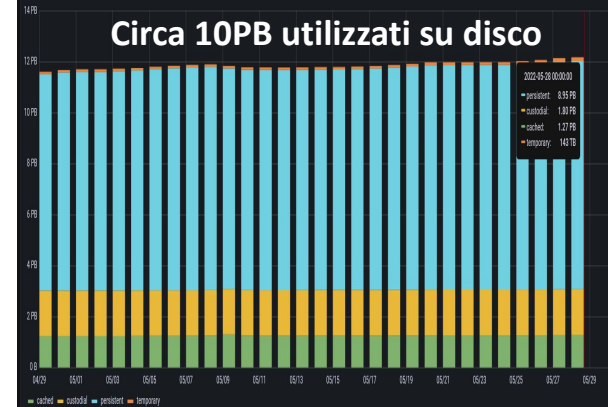
Running jobs by JobType
53 Weeks from Week 12 of 2021 to Week 13 of 2022



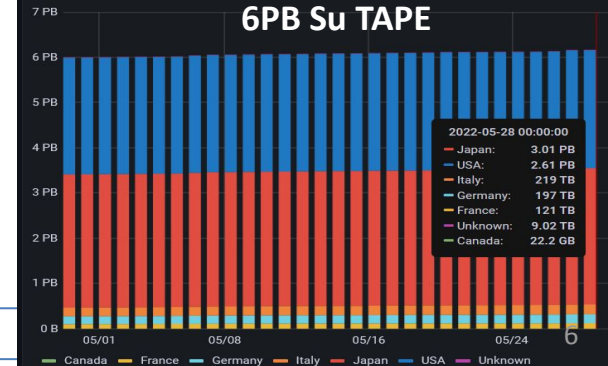
Italian Share 12.6 (Milestone 2021 -11%)

- Attività aumentata rispetto al 2020
- Picchi di ~40 k jobs running
- 31kJobSlots Pledged molte CPU Opportunistiche
- Current User Job 24% (increasing)

Volume per Custodially



Volume per Country

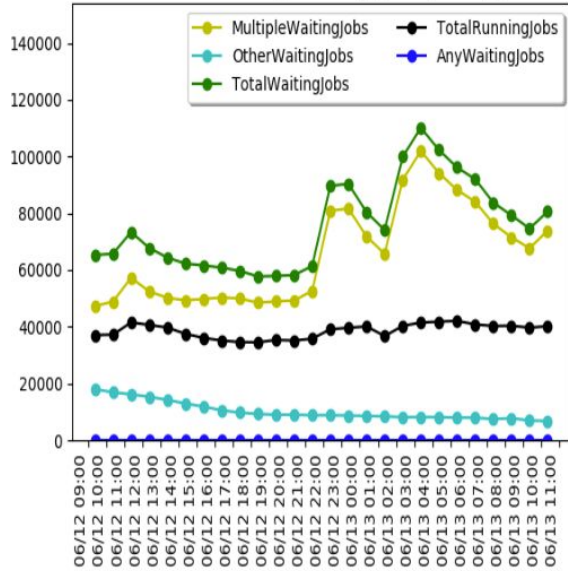


Siti Italiani

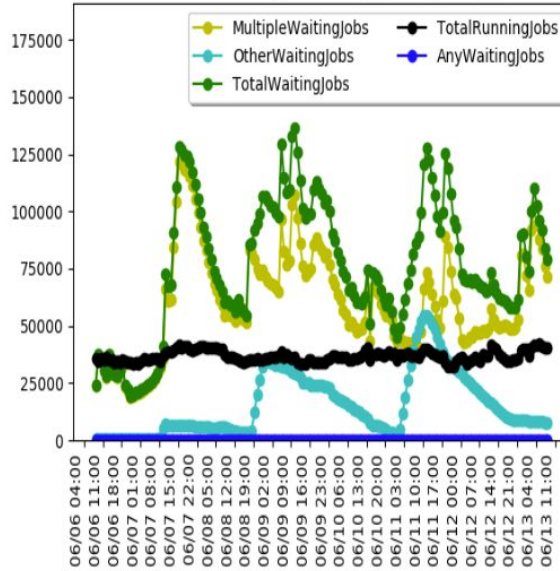
	CPU Pledge (kHS06)	CPU Opport. (kHS06)	Storage (TB)	Tape (TB)
CNAF	27		820	650
Cosenza	1			
Napoli	13	10	390 (+200)	
Pisa	8	10	200	
Torino	6	24	350	
Frascati		0,5	11	
LNL		1		
Roma3		2	2	
TOTALE	55 kHS06	47,5 kHS06	1.973 TB	650 TB

Dalle stime attuali per il 2023 occorre implementare ulteriori 200 TB da mettere al CNAF

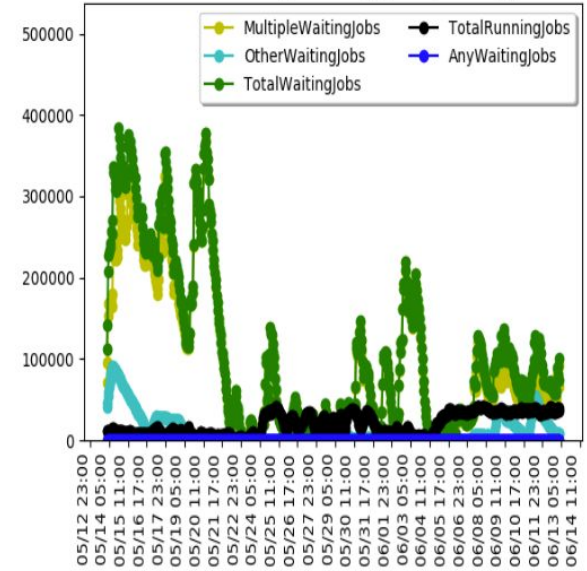
Total Running/Waiting jobs (1 day)



Total Running/Waiting jobs (7 day)



Total Running/Waiting jobs (30 day)



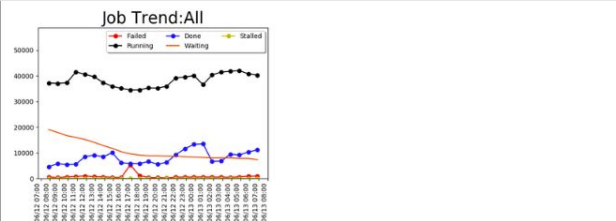
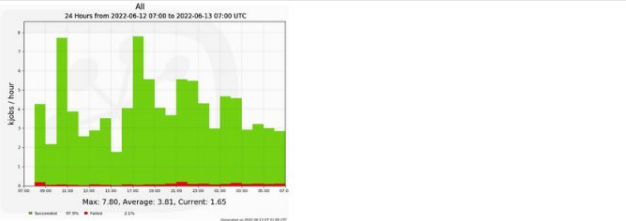
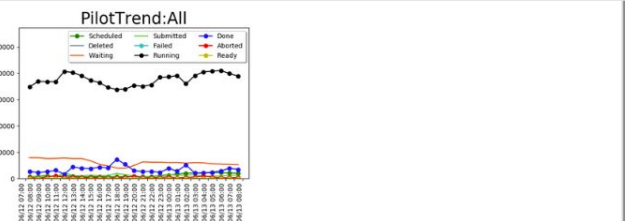
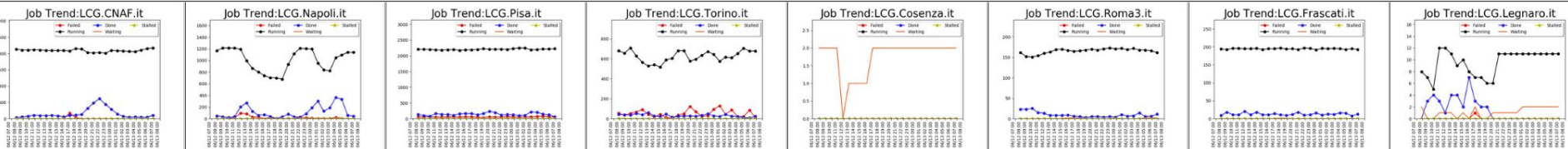
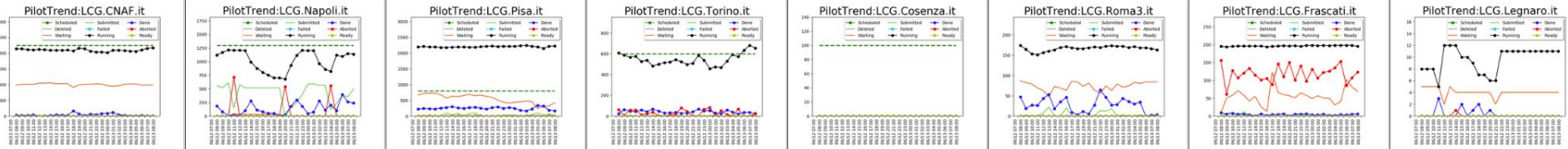
Total Running Job = 40184

Total Waiting Job = 80564



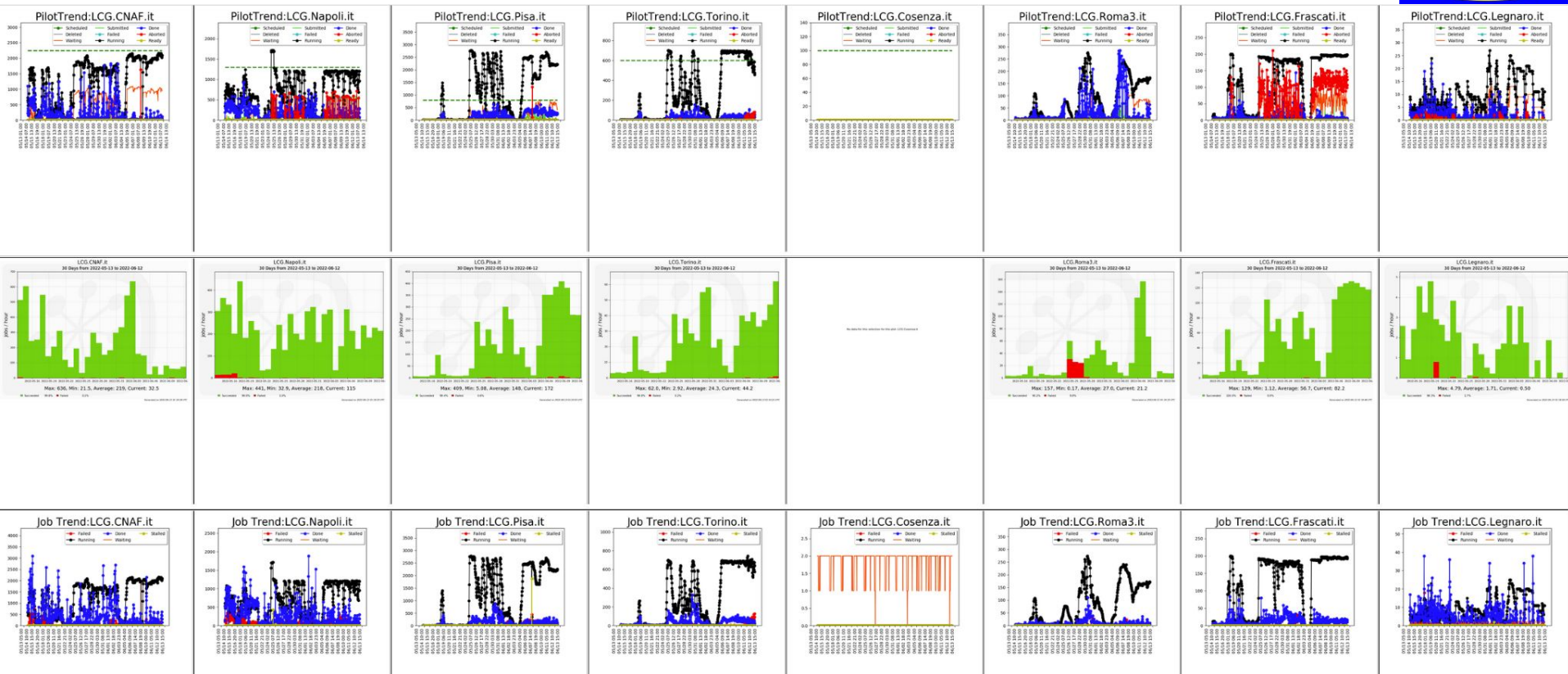
Risorse di Calcolo siti Italiani Oggi 13 Giugno 2020

<http://people.na.infn.it/~spardi/mon.html>

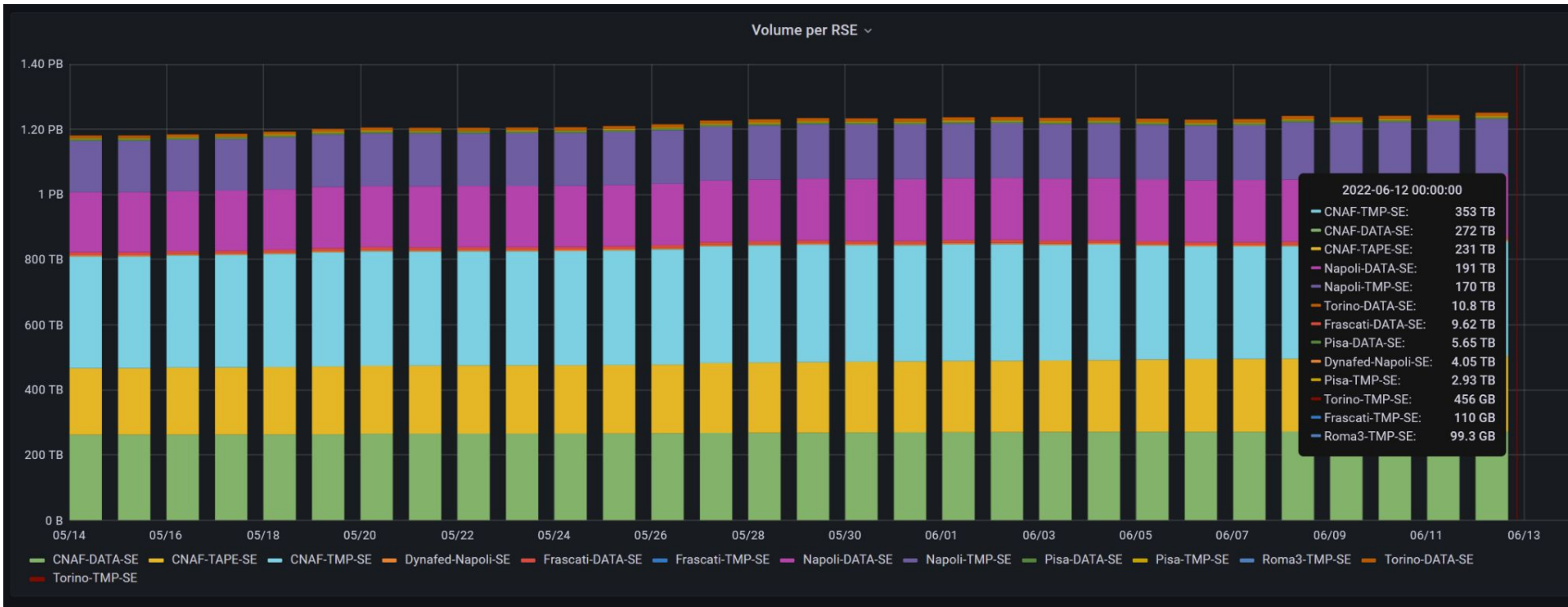


<http://people.na.infn.it/~spardi/mon.html>

Risorse di Calcolo siti Italiani Ultimo Mese



Utilizzo Storage Siti Italiani



<https://monitoring.sdcc.bnl.gov/pub/grafana/d/ddm-site-accounting/ddm-site-accounting?orgId=1&var-groupby=RSE&var-Country=Italy&var-Site=All&var-RSE=All&var-SiteType=A&var-Account=All&var-Campaign=All&var-Release=All&var-DataLevel=All&var-Scope=All&var-Experiment=All&var-Custodiality=All&var-BeamEnergy=All&viewPanel=3452>

Commenti sull'utilizzo delle risorse

- Utilizzo CPU dei siti italiani è sempre sulla linea del pledge e oltre nei momenti di utilizzo intenso dell'infrastruttura.
- Utilizzo dello Storage: Non ancora bilanciato. Verificato che dopo il cambio di deputy computing coordinator (da Kato-san a Michel Villanueva) non è stato ancora aggiornato lo share sugli storage. Verrà fatto sperabilmente nelle prossime settimane.

Belle II Global Sites Meeting

A febbraio abbiamo organizzato un evento di coordinamento di tutti i siti di Belle II sui seguenti topic:

- Phase-out of SRM/GridFTP
- Enabling Token Based Authentication
- WebDAV deployment
- HTTP-Third-Party-Copy (TPC) support
- Storage Accounting via JSON File

Grande partecipazione, stabiliti due goals:

- Implementare webdavs su tutti gli storage per Giugno
- Testare la token based authentication sui siti entro Giugno

Attivazione davs in scrittura

E' iniziato il processo di migrazione definitiva vs davs con l'abilitazione in scrittura come primo protocollo.

Per ora sulla base dei risultati dei test di TPC sono stati configurati i seguenti storage pilota

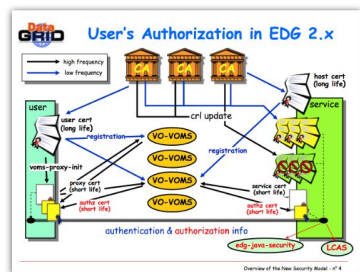
CNAF-TMP-SE, KEK-DISK-TMP-SE e IN2P3CC-TMP-SE

<https://agira.desy.de/browse/BIIDCO-4521>

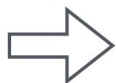
Tutti gli storage Italiani sono pronti per utilizzare davs in scrittura.

Token Based Authentication in WLCG

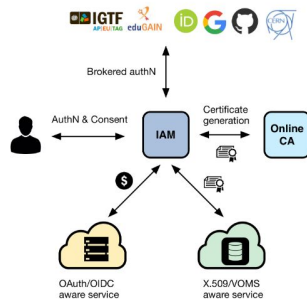
Current, X.509 based AAI



Move beyond X.509



Future, token-based AAI



WLCG Roadmap

<https://twiki.cern.ch/twiki/bin/view/LCG/WLCGAuthorizationWG>

https://docs.google.com/document/d/11fcZU8fEsfjDiSkjh95nVr4tNXLPCA_xwr2SwriBpiw/edit#

Utilizza protocolli standard OAuth/OpenID Connect. L'italia in particolare il CANF è leader nell'implementazione di questa tecnologia in abito WLCG.

HTCondor-CE e i siti americani in generale (OSG) stanno accelerando il processo di dismissione dell'autenticazione via certificato








M.7	Mar 2022	All storage services provide support for tokens	M.1	WLCG Ops, Storage providers
	?	All VOs shut off VOMS-Admin		
	Sept 2022	End of HTCondor support for GSI Auth (link)		
M.8	Oct 2022	Rucio transfers performed with token auth in production	M.7	Rucio, Experiments
M.9	Mar 2023	Experiments stageout & data reads performed via tokens.	M.7	Experiments
M.10	Mar 2024	X.509 client auth becomes optional.	M.9, M.8, M.4	Experiments

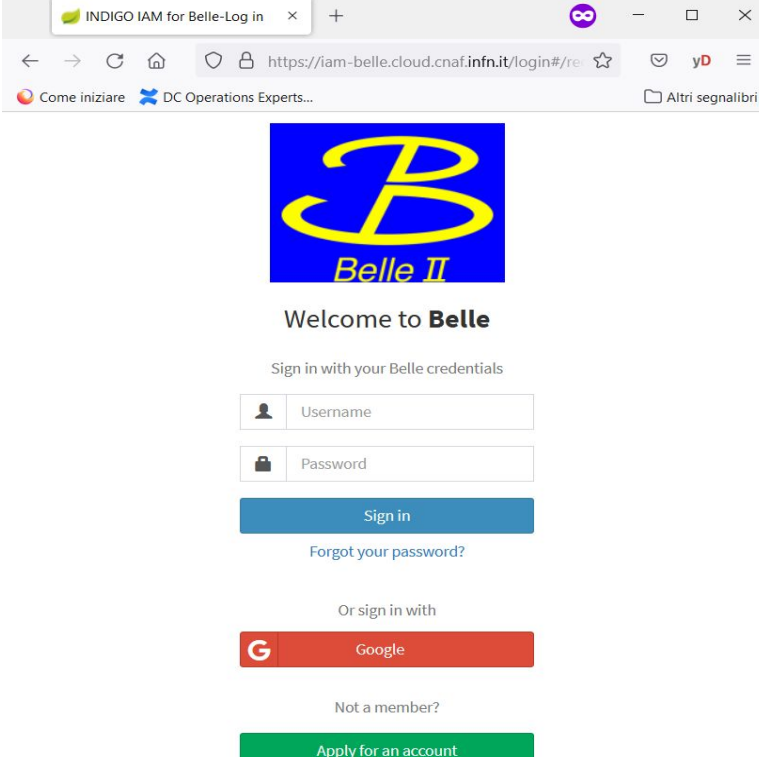
Testing IAM Implementation

Abbiamo iniziato a lavorare sull'implementazione della token based authentication. Il CNAF ha messo a disposizione un server Indigo-IAM i test. Il servizio di produzione sarà ospitato da KEK.

<https://iam-belle.cloud.cnaf.infn.it/>

Abbiamo un set di siti pilota. N.B. Attualmente DIRAC Non supporta completamente i token.


Sub-Tasks				+	...
1. IAM Configuration		IN PROGRESS	Unassigned	...	
2. CNAF: HTCondor-CE and STORM		IN PROGRESS	Silvio Pardi		
3. NAPOLI: HTCondor-CE		IN PROGRESS	Silvio Pardi		
4. DESY: HTCondor-CE and dCache		IN PROGRESS	Thomas Hartmann		
5. BNL: HTCondor-CE and dCache		IN PROGRESS	Hironori Ito		
6. IN2P3CC: HTCondor-CE and dCache		IN PROGRESS	Aresh Vedae		
7. KIT: HTCondor and dCache		IN PROGRESS	Matthias Schnepf		



INDIGO IAM for Belle-Log in

https://iam-belle.cloud.cnaf.infn.it/login#/re

Come iniziare DC Operations Experts... Altri segnalibri



Welcome to **Belle**

Sign in with your Belle credentials

Username

Password

Sign in

Forgot your password?

Or sign in with

Google

Not a member?

Apply for an account

Test HTCondor-CE con token based authentication

I token in prima battuta verranno utilizzati solo per la sottomissione dei pilotjobs.

Tutti i siti italiani utilizzano HTCondor-CE.

Dovremmo coordinarci per avere tutti i siti pronti.

Andrebbe bene fine 2022 come milestone?

HEP Benchmarks project

Three main components

– HEP Workloads

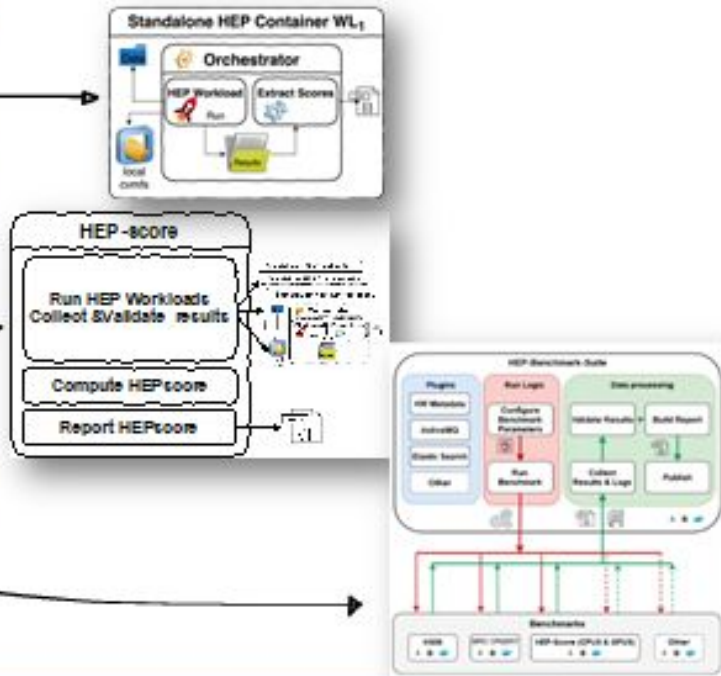
- **Standalone reference** HEP applications
 - No network needed at runtime to access sw and data
- Common build infrastructure

– HEP Score (the benchmark)

- Orchestrate the run of a series of HEP workloads
- Compute the **HEPscore** value
- Report whole set of WL results

– HEP Benchmark Suite

- Framework to run likewise multiple benchmarks
 - HEPscore, HS05, SPEC CPU2017...



Activity on the grid by users

Accounting since last B2GM

- User job 24% dell'attività sulla grid (vs 23% rispetto all'ultimo B2GM)

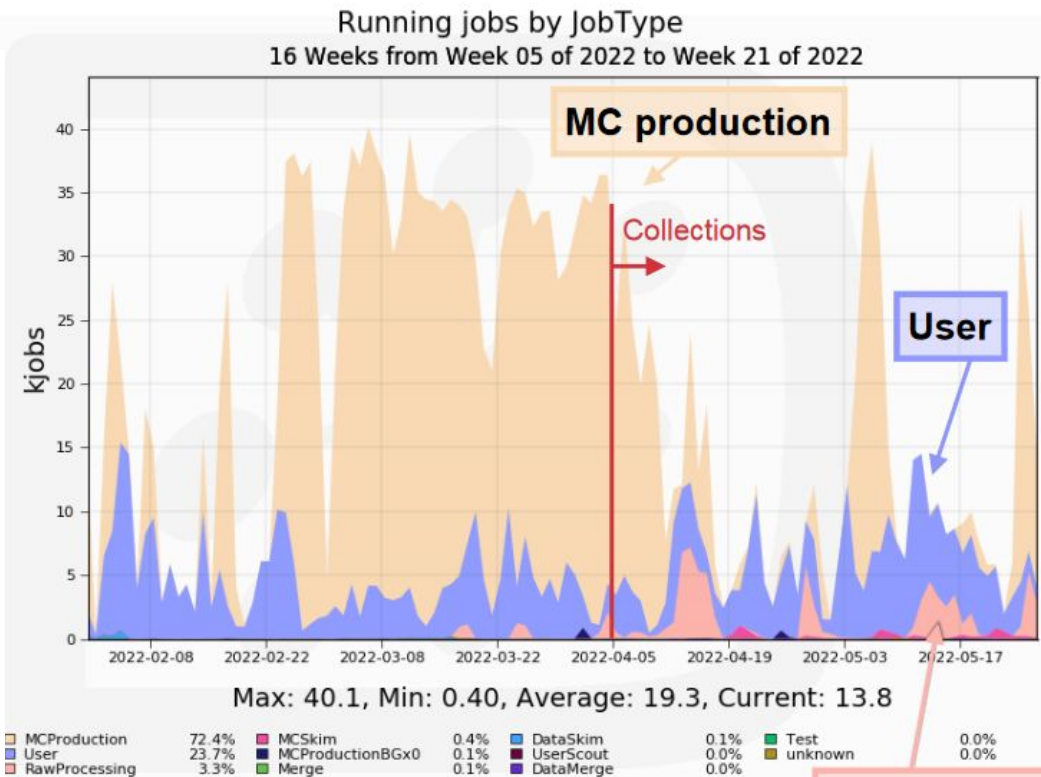
Quando non ci sono attività di produzione gli user job non utilizzano tutte le risorse disponibili:

Problemi conosciuti:

- Limiti nell'accesso all'input sand box
- CPU Time over-estimated

Problemi occasionali:

- Job Optimized in stuck
- Job Submission issue



Raw processing

Problemi conosciuti

- **Limitazioni per l'accesso alla Input SandBox:** Problema anche in LHCb (utente DIRAC) In studio la possibilità di utilizzare gli Storage Element per l'upload dei file di input invece del server DIRAC. Alcuni test svolti da DESY
- **CPU Time-overestimation:** Per default il CPU time viene stimato utilizzando i parametri statici presenti in DIRAC. Nella maggior parte dei casi è sovrastimato. In conseguenza di ciò solo pochi siti/code risultano in grado di eseguire il job. Ci sono delle idee per mitigare ma richiedono ancora del lavoro (e del tempo). Nel frattempo quello che si chiede agli utenti è di stimare il CPUtime:
<https://confluence.desy.de/display/BI/GBasf2+FAQ#GBasf2FAQ-CPUTimevalueetospecify>

Code Multicore

E' da poco iniziata un'attività per l'attivazione di code multicore. Per ora UVic e KIT si sono proposti per i test mettendo degli effort specifici.

Ancora da verificare tutto il workflow dal pilot all'applicazione.

Idea di cominciare con una specifica classe di job poi di capire come estendere l'utilizzo.

Seconda copia output job utente

Un problema comune per gli utenti è la difficoltà a scaricare l'output dei file per problemi contingenti allo storage specifico che contiene tali dati.

Proposta in corso è di creare una seconda copia temporanea degli output dei job utente, con una lifetime breve (pochi mesi o settimane) per fornire un backup e mitigare questa problematica.

- Approccio seguito da altri esperimenti
- Può essere gestito da rucio
- Richiede maggiore spazio disco (da tenere in conto per le stime)