

Minute CdG - 19/07/2024 (<https://agenda.infn.it/event/42709/>)

Partecipanti

- (CMS) Daniele Spiga
- (ATLAS) Lorenzo Rinaldi
- (LHCb) Lucio Anderlini
- (JUNO) Giuseppe Andronico
- (VIRGO) Stefano dal Pra
- (VIRGO) Luca Rei
- (ATLAS) David Rebatto
- (ATLAS) Alessandra Doria
- (ALICE) Mario Sitta
- Lorenzo Silvestri
- Claudio Grandi
- Stefano Bagnasco
- (GAMMA) Alain Goasduff
- (CMS) Tommaso Boccali
- (DUNE/ICARUS) Matteo Tenti
- Alessio Tiberio

T1 overview

slides

Domande

CG: ma le macchine che usiamo a Casalecchio si connettono a Berti o al Tecnopolo?

Zani: no si connettono a Berti

DS: Ho due cose da chiedere: La prima: sui ritardi, CMS aveva dato una deadline di fine luglio riusciamo a rispettarla?

DC: CMS era quello meno in sofferenza e quindi al momento non lo abbiamo esteso, parliamo con Vlad e ti facciamo sapere

DS: La seconda è una curiosità riguardo i problemi di uptime di Leonardo, avete detto che

potreste sopravvivere anche nel caso di downtime. Quello che mi preoccupa soprattutto è il down del filesystem.

DC: dipende perchè sono macchine diskless, l'hypervisor non ha disco

DM: se va giù un sistema operativo, la macchina virtuale dura poco: il job slurm che fa partire la macchina virtuale muore e dobbiamo far ripartire tutto. Dipende come funzionano il filesystem dell'hypervisor

MT: ho saltato la prima parte in cui dicevi che siete pronti per assegnare le quote per gpfs data. Di che tempi parliamo?

DC: lunedì/martedì

MT: per DUNE le richieste 2024 c'era una macchina con GPU che andava a finire tra le hpc bubble, volevo capire se fossero arrivate

DC: come data Q4 2024, quindi tra ottobre e dicembre

Farming

slides

Domande

LR: riguardo alla migrazione a Alma9, se ho capito bene per le ui siamo a buon punto anche con i servizi, per i wn avete un'idea di quando avverrà l'aggiornamento?

DM: dopo l'estate, l'idea era di migrare i WN un pò alla volta, dall'altro lato abbiamo anche un'immagine singularity el7 nel caso in cui proprio non si riuscisse ad usare alma9. Ma siamo disponibili a dare supporto a tutti

DR: sarà una coda per cui i job che finiscono li usano il container o tutti i WN?

AP: usiamo una riga nel submit file in cui si specifica l'immagine, comunque tutti i WN sono in grado di gestire job simili. Inoltre è tutto documentato nella guida

Storage

slides

Domande

LM: Domanda rivolta ad ALICE: questi dati vi servono?

MS: vi diciamo al più presto come trattare questi dati

VS: per lo spazio CMS, c'è un piccolo problema con HSM da risolvere nei prossimi giorni, prevediamo che ci sia qualche PB in meno del pledge.

Uso opportunistico di Leonardo

slides

Domande

SDP: dalle ultime prove fatte prima di andare a chep se si vedevano arrivare richieste per quelle risorse allora si usavano quelle risorse. Eventualmente ti tiro fuori i repository dove era

DC: vale sia per gp che per booster

DS: Si

DC: e l'accounting dove finisce? INFN?

DS: coda a bassa priorità non accountata

CG: a regime, se volessi usare la mia quota su Leonardo c'è modo di accountarla?

DS: non dobbiamo replicare i nodi per n account

CG: si poi quando sarà il caso vedremo come evitare di far usare risorse di altri

DS: l'idea è di usare gli account di servizio.

SDP: riprendo la domanda di cesini sull'account, quelli vengono visti come job regolari da uno schedd di condor

DC: si, mi chiedevo lato cineca

SDP: quello dipende da quale coda slurm fa partire i job

DC: DS hai trovato sempre posto in questa low priority?

DS: si, magari Tommaso può dirci qualcosa in più

TB: comportamento non prevedibile: a volte entra 0 altre volte 100

DC: sulla rete, passa tutto per general IP quando escono i dati?

DS: si

DC: grazie per il lavoro fatto

DS: ringrazio anche DM per il supporto

Atlas

slides

Domande

LR: siamo ancora underpledge

AP: perchè non contate le arm come pledge?

DC: ma non c'è pressione?

LR: i buchi non mi preoccupano perchè sono dei transienti

DC: è la pancia dopo il buco, nella prima settimana di luglio che mi preoccupa

AP: la pressione la vedo su arm

LR: questa cosa non si può sistemare? tutto quello che non è arm è extra pledge

AP: io vedo una media di 120 kHS06

DC: ci guardiamo

LR: vorrei fare la richiesta di escludere dal conteggio il contributo del lavoro su arm

CP: magari ha senso mettere un'infrastruttura dedicata? per noi è difficile considerare queste risorse come extra. Come ci regoliamo con le SLA e le operations?

LR: semplicemente se mettete i nodi in drain non li usiamo

DC: ci ragioniamo

CMS

slides

Domande

LHCb

slides

Domande

VS: come vi avevamo già riportato via mail, il monitoring sullo spazio non è affidabile in questo periodo perché aggiungiamo nuovo storage al TP e migriamo i dati da quello vecchio al TP. Per questo motivo, il monitoring potrebbe riportare spazio libero in più, quindi abbiamo richiesto di limitare le nuove scritture sul disco nelle prossime settimane a 3PB.

LA: va benissimo se la limitazione è 3PTB, però rispetto ai 7-8 PTB usati al momento, possiamo scrivere di più?

VS: sì, la quota sarebbe disponibile già adesso (uguale al pledge)

LA: quindi posso ignorare il 7.13 che vedo su t1metria e andare oltre?

VS: sì, aggiorniamo adesso la quota

DC: pensavamo di dimensionare il buffer ad 800TB e ridurre il pin a 12 ore invece che tre giorni.

VS: 800TB sono troppi, la permanenza sul disco di questi dati è abbastanza limitata (qualche ora) quindi possiamo diminuire il pin a meno di 3 giorni (che è il default), come suggerito da Lucio. Alcuni tool permettono di settare il pin file per file durante la scrittura. Forse Andrea può commentare meglio.

AR: forse con client SRM, ma dovremmo guardarci.

LA: la mia preoccupazione è che se riduciamo troppo lo spazio, potremmo vedere degli errori sul trasferimento di file fts (i file vengono migrati prima di essere trasferiti dal buffer al disco). La mia proposta era quella di prendere qualche misura prima della presa dati.

VS: Da fts si può sapere quando il file è stato chiamato e quando è stato trasferito.

DC: riguardo la slide sulle tape: quand'è che arrivi al pledge? novembre?

LA: si attorno a novembre, ci aspettiamo che per quelle non ci sia tape nel pledge

DC: ti aspetti di usare l'overpledge? o l'overpledge l'avevamo pledged l'anno scorso?

LA: mi aspetto di usare un tesoretto di 9 peta

no LHC - us

slides

Domande

DC: g-2 li stiamo sentendo?

AP: non si sono fatti risentire dopo l'onboarding. Magari li ripinghiamo settimana prossima

DC: okok

CP: confermo quello che dice AP abbiamo inviato 2-3 email con minute e guide e domande a cui non abbiamo ricevuto risposte. Visto che lo storage è arrivato possiamo ripingarli.

Giro di tavolo

JUNO

GA: stiamo pensando di passare allo IAM voms ma ci manca un ultimo passaggio a quel punto la decisione è quella di spegnere voms admin in cina e passare ad IAM voms. Volevo aggiungere anche qualcosa riguardo lo storage, abbiamo sofferto parecchio ma mi pare che la situazione sia risolta, giusto?

DC: si assegneremo le quote 2024 verso lunedì-martedì

JLAB

nulla da riportare

Virgo

SDP: con virgo stanno cominciando a funzionare dei sistemi di checkpointing: un job che

ha esaurito il tempo utile può ricominciare da dove era rimasto e portare a termine il suo lavoro con una sequenza di vite successive. Nei siti WLCG per richiesta degli LHC questo non è possibile, al CNAF è inibito anche per la sottomissione dei job locali, a me piacerebbe attivare il checkpointing almeno per virgo e le sottomissioni locali.

AP: ne discutiamo con farming e vediamo come agire

SDP: a seconda degli utenti potrebbe tornare utile anche nei contesti di Leonardo in cui le risorse tendono ad essere più ballerine o per i job slurm

DC: va bene

CP: domanda per stefano: com'è implementato il checkpointing? il processo salva lo stato della memoria sul disco quindi se viene killato può ripartire da quel punto lì. Quali sono le richieste di storage?

SDP: l'intera sandbox viene preservata e rilanciata nello stesso posto in cui si trova, lo implementi lato condor specificando un exit code, condor non è appesantito, semplicemente ricicla la sandbox. Poi il checkpointing è realizzato dall'utente per esempio con un file di append. Non è una panacea per tutti i mali ma per file di pochi MB si può usare e non è pesante per lo storage

AOB

nulla da riportare

DC: il prossimo cdg sarà il 20/09.