

Minute CDG 16/12/2022

Introduzione

Slide:

- Pledge 2023 fissati, risorse approvate dal referaggio.
- Crescita netta CSN2 HTC, CSN2 Cloud e CSN2 HPC (tabelle). Crescita anche per CSN3 HTC.
- Stato gare dicembre 2022.
- Stato infrastruttura: interventi manutenzione quadri elettrici nei giorni scorsi ed effort su lavori al Tecnopolo.
- Stato Tecnopolo – immagini sulle slide.

Farm

Slide:

- Utilizzo delle risorse – no particolari discrepanze con i mesi precedenti.
- Nuova gara: ordine inviato lunedì scorso, passati molti mesi e non si sa perché.
- Nuova versione Linux: decisione scaturita la settimana scorsa dal CERN e Fermilab di abbandonare quella che sembrava la soluzione migliore, cioè CentOS stream in favore di Alma Linux. Noi ci stavamo orientando verso Rocky che ci sembra più sostenibile ma CERN e Fermilab sono andati un po' controcorrente. Vi chiederò qualche riscontro su questa decisione.

LA: io come LHCb non sapevo di questa transizione ad Alma Linux e abbandonare CentOS stream. Mi informo e faccio sapere.

DC: Sì, è stato un fulmine a ciel sereno.

AC: da HEPIX avevo già nasato qualcosa. Pensavo più a Rocky, ma va bene perché le versioni sono abbastanza compatibili, però mi interessava sapere qualcosa in più. Questo cosa comporterà? Una rivalutazione del codice? La fisica funziona ancora? Ci vorrà ancora molto tempo per questo? Un anno?

NM: Ho due domande:

1. Come mai da più derivate RH a una sola?
2. Dalla 8 alla 9: noi siamo ancora on hold perché aspettiamo il comportamento dei datacenter. Però abbiamo quasi tutto containerizzato. Noi siamo più favorevoli alla transizione diretta al 9 che dovrebbe portare più benefici?

AC: perché dici "benefici" dalla 9?

N: Ottimizzazione del kernel... Niente di specifico, ma in generale le ottimizzazioni portate dalle nuove versioni.

AC: per la prima domanda: il CERN pensa di utilizzare Alma, ma questo non è preoccupante poiché tutte le versioni sono ottimizzate. Da capire i tempi per i fix delle vulnerabilità e il filo di dialogo con gli sviluppatori.

AD: anche a Napoli pensavamo di passare a Rocky9. Non so a livello di exp se si è discussa questa cosa dopo l'annuncio.

LR: confermo quello che ha detto Alessandra: a livello di exp non sapevamo nulla.

Però la maggior parte dei lavori sono tramite container. A livello di utente finale non mi aspetto grandi variazioni. È giusto?

DC: sì, è giusto. Sono d'accordo.

LR: magari per gli exp più piccoli con installazioni locali del software può essere più sensato avere una scelta del sistema, ma per i grandi exp non cambia molto.

DC: azione su US – girare la domanda di Andrea sui soliti canali di comunicazione e raccogliere i feedback.

CP: è stato fatto un survey utilizzo container. Molti exp non li usano e questo ha un impatto sul lavoro, in particolare sul software, installazioni, ecc...

DC: io la manderei proprio a livello informativo per il passaggio ad Alma che non è stato molto notificato.

DC: Non sono molto preoccupato per la gara CPU, però per la parte storage un po' sì.

AC: Ci tenevo a mettere le cose in chiaro. Non è colpa del CNAF. La burocrazia rallenta sempre tutto.

DC: sì sì, il messaggio era chiaro. Non lo avevo detto io nella mia introduzione. Forse non riusciamo invece a dare lo storage a tutti a gennaio per il nuovo pledge dato che i 14 PB non li avremo a gennaio.

Invece per la parte CPU siamo abbastanza coperti, da marzo almeno. Che ci sarà la rete pronta con il Tecnopolo. Condor al CNAF su Leonardo.

Storage

Slide:

- Traffico ultimo mese aumentato abbastanza.
- Riepilogo sistemi disco in produzione (tabella).
- Acquisti recenti.
- Software in produzione.
- Ticket and problems:
 - Atlas ticket trasferimenti falliti;
 - fallimenti TPC in push mode google cloud;
 - issue aperto agli sviluppatori su saturazione thread XrootD;
 - LHCb: lato FTS è stato risolto bug su come venivano risolti gli host nell'alias degli endpoint StoRM WebDAV. Se volete rifare dei test, dovrebbe andare meglio;
 - No-LHC:
 - Borexino - ancora in corso download via webdav dei dati;
 - in sinergia con User Support stiamo contattando esperimenti per transizione GridFTP -> WebDAV;
- Aggiornamento tape – Traffico lettura e scrittura (tabella).
- Scenari a medio termine per tape. Volontà di dismettere vecchia libreria Oracle spostando i dati sull'altra libreria. Non vogliamo portare al tecnopolo una libreria in dismissione.

VF: per AMS. Dei due ddn, uno è per lo storage AMS. Gli interventi impattano?

EF: no, è trasparente. Per questa migrazione dei dati non sono previsti down. In caso di interventi abbiamo delle parti spare, quindi non ci dovrebbero essere problemi.

LA: ricevuto il messaggio. Ho parlato con Chris. Vorremmo fare un'altra mini-challenge a gennaio anche con RAL. Grazie.

LM: Perfetto.

LHCb

Slide:

- Sotto-utilizzo delle risorse perché stiamo pagando i ritardi dovuti alla pandemia. Da un mese siamo under-pledge.
- Cineca M100: use-case completo, sviluppato un set up di addestramento di reti neurali che passa attraverso slurm. Per sottomettere questi job servono le credenziali cineca. Prossimo step: riuscire a sottomettere senza essere utenti cineca ma essere nella VO lhcb. Per questo sarebbe l'ideale fare

sottomissione a un Condor CE. Step successivo: sottomissione via Dirac. L'applicazione al momento non è containerizzata.

- Report uso cloud: stiamo facendo un po' di progressi.
- Tape challenge: a gennaio test di scrittura massiccio su tape.
- Migrazione Alma Linux: a fine gennaio vorremmo avere le prime build con alma linux 9.

DC: su M100, a parte la macchinaria, la fisica è validata per power9?

LA: in realtà no. Sono user job e la parte nostra è esclusivamente in python. Gira out of the box. Non hanno osservato differenze tra M100 e x86. Quello che succede a livello computazionale è dentro la GPU e quindi non mi aspetto grosse differenze.

DC: in futuro per Leonardo dovremmo avere anche la stessa configurazione per l'accesso alla parte booster con GPU. Potremo riutilizzare questo lavoro. Facciamo una cernita e proviamo a scrivere agli exp.

CP: qui è stato abbastanza semplice. Ci sono spesso cartelle utente condivise sul FS e non va bene. Con i "du" ad hoc scopriamo quanti file sono allocati solo temporaneamente.

ATLAS

LR: si possono spengere i GridFTP. La SA atlashotdisk si può cancellare tutta.

Le directory SAM invece bisogna tenerle. Anche quelle locali vanno tenute. Di test invece si può cancellare tutto.

Le altre configurazioni vanno tutte bene.

ALICE

Slide:

- Calcolo (plot slide) – giriamo un po' underpledge.
- Storage: 2 PB liberi. A gennaio inizierà campagna di scrittura massiccia su tutti i Tier.
- Problema non critico al momento: parte gpfs del disco con test IPV6, da investigare.

DC: Francesco era stato erroneamente rimosso dalla lista CdG.

FN: nessun problema.

DC: la scrittura in modo massiccio è su disco?

FN: Sì.

EF: la cancellazione prevede solo il disco?

FN: sì.

DC: per i problemi ipv6?

FN: dobbiamo ancora guardarci.

DC: calo di pressione.

SDP: periodi job pending, ma Alice rimaneva sotto. C'è un fatto sistematico. È legato al meccanismo di share fatto da me un paio di anni fa. Per come è fatto, ci sono errori sistematici. Altri interventi potrebbero però essere la causa. Negli ultimi giorni, Alice sta sopra.

FN: non è importante, volevo solo segnalarlo.

DC: whole-node potrebbe avere un impatto non banale. Martin non voleva coinvolgere i tier-1 per questi test, però se siete disponibili. Parliamone.

FN: nel caso in cui ci fossero state possibilità di test.

SDP: era adatto M100 ma ci vogliono payload.

FN: non siamo pronti.

DC: o lenoardo quando arriverà.

VS: un po' meno utilizzo partizione CEPH. Ci sono feedback? Meno utilizzo? Problemi?

FN: su Ceph nessun problema. Lo storage e i test non danno problemi, ma abbiamo scritto poco in previsione di gennaio.

VS: se ti offriamo più spazio su ceph invece di GPFS?

FN: chiedo, ma non penso sia un problema.

VS: principalmente dobbiamo dimezzare.

FN: è un impatto grosso, quindi prima chiedo e poi ti faccio sapere.

NO – LHC

NM: Ci sono avvisi prima di bannare un utente?

SDP: dipende dalle circostanze.

CP: dipende dall'occorrenza. Prima il ban, poi vengono killati i processi, poi viene notificato l'utente e il referente informatico. Generalmente senza preavviso, in situazioni molto gravi.

Per esempio invece con "rsync", prima mandiamo una mail all'utente. Se però questo crea problemi gravi si deve intervenire. Le attività sono indicate nella nostra AUP. Siamo in generale abbastanza gentili.

DC: L'attività per DarkSide a fronte dello spazio esaurito è stata molto utile. Potrebbe essere utile farlo anche con altri esperimenti che hanno lo spazio quasi esaurito, in un'ottica collaborativa.

AOB

DC: potremmo fare un Focus su qualche esperimento specifico nei prossimi mesi. Da qui a gennaio possiamo ragionarci.