

## Minute 19/11/2021 – CDG

### Introduzione:

- **CPU**  
20% sotto pledge, manca parte CINECA che non arriva entro fine anno.  
Aggiornamento ad HTCondor 9 -> slide dettagliate Farm -> siamo al 75% dell'upgrade
- **Storage:**  
Siamo pledge 2021. I due lotti della gara 2020 sono stati installati e collaudati. Da inserire in produzione -> slide storage.  
Tape:  
Fornitura cassette nuova libreria. Siamo a pledge 2021. 11 PB liberi
- **Passaggio consegne Luca sui pledge 2022:** da completare. Daniele non ha ancora la tabella. Per il prossimo CDG quadro completo da presentare.
- **Infrastruttura:** nessun intervento importante entro fine anno.
- **Tape challenge -> presentazione storage.**  
Nessun problema significativo. Target LHC raggiunti, ma sottomissioni in momenti diversi.
- **Network challenge.** A. Forti report nel GDB di Novembre. Target da raggiungere 30 e 30 Gb/s -> non raggiunto in ingresso. Forse pochi dati spediti? Organizzata mail -> analisi effettuata da storage e rete. Dal monitoring del Cern, la media in ingresso è 20 Gb/s, interno invece throughput 28 Gb/s. (tra l'8 e il 9 ottobre 40 Gb/s -> no problemi a soddisfare target se c'è sufficiente pressione)  
Richiesti aggiornamenti agli LHC.
- **Prossimo CDG 17 dic (forse sovrapposizione con Alice computing)**

### Farm (sdp):

- L'utilizzo risorse da parte dei gruppi nel mese scorsi. CPU: vedi slide.
- Nuove risorse: Consegna CINECA slitta -> non prima del 2022.
- Al CERN c'è intenzione di usare RHEL in modo massivo. Accordi ancora non definitivi.  
Non tutta la comunità passerà a Red Hat ent.  
Alternative: rocky e almalinux. Alcuni siti indagano (Desy e BNL). In fase di testing.  
Noi attendiamo. Ci attendiamo però che uso RHEL aumenti. Siti affiliati al CERN potranno usufruire dell'accordo con CERN.  
Centos sui wn, sui client, sembra che continuerà a essere usata. Il CERN pensa anche di saltare centos8 e passare direttamente a centos9 (che dovrebbe essere pronta quando la 7 non sarà più supportata).
- Aggiornamenti vari: ritardi consegne hw -> ovirt.
- In arrivo – è in installazione – ui-tier1 su un nuovo server.

- Attivare firewall su una nuova configurazione su tutti gli host del farming – firewalld – che è un wrapper su IPtables. Sui wn era già attivo firewalling con iptables che blocca le connessioni non previste (fuori CNAF). Per i ce rimane aperta solo porta 9619.

#### - **Aggiornamento a Condor9:**

Cambia modalità authZ. Con 8 era host-based. Adesso è security based, basata su Token. I demoni di condor si autenticano tra di loro tramite demoni.

Anche per i CE aggiornamento a Condor5 che utilizza autenticazione con i token.

Colta occasione mese scorso intervento storage per aggiornamento di metà della Farm. Da quel momento due versioni di Condor sono coesistenti. Successivamente il restante 25 %. Ieri pomeriggio ultimo upgrade 25%. Già diversi wn lavorano con Condor9.

Anche gli ultimi ce sono aggiornati a Condor5. Non c'è bisogno di mettere in drain i ce.

Tutti aggiornati tranne ce01 che sarà dismesso (tenuto come di scorta) -> in più però c'è ce07.

Da ieri pomeriggio ce01 non mette in run job nuovi (drain) -> segnato down. Finirà i suoi job running e poi sarà spento.

I ce condor 5 supportano scitokens – per il submitting dei pilot – per il resto serve ancora x509.

Job locali: sn nuovo – sn-02 – per Condor9. AuthZ token-based.

Viene creato token utente appena si apre la propria ui, per cui basta cambiare solo il nome del sn. Granularità più fine da parte dell'utente che può chiedere.

#### **Storage (VS):**

- Ultimi mesi più traffico verso esterno – tape data challenge.

- Traffico interno sorpassato un 1Tb/s legato al permissioning di Ceph. Modificando funzionalità di FS ceph.

- Aggiunto huawei ai sistemi di prod. 9 GB spazio netto -> il piano è di assegnarli ad Atlas.

- Il collaudo concluso con successo con numeri di prestazioni più alti rispetto al target (capitolato) vd slides. Manca slitte e cavi da montare.

- **Piano di spostamenti:** sistema pronto per la prod ma per ottimizzare è necessario effettuare alcuni spostamenti. Allocare ad Atlas intero sistema. Così il pledge 2021 di ATLAS sarà su un intero sistema.

Questo permette di liberare 4 PB su sistema di DDN. E anche 2 PB su OS.

#### **Dettagli slides -> numeri esatti e spostamenti.**

Una parte andrà CMS e altra parte ad Alice.

Per questi spostamenti ci vorrà almeno una settimana.

Dopo questi movimenti ci sarà spazio libero per altri exp – darkside e ams.

Così riusciamo a soddisfare le richieste di tutti exp.

LR: gli spostamenti sono trasparenti per gli utenti? Sì.

#### **Installazione CEPH:**

Cluster installato e collaudato. Fase di test di pre-prod. Accordo con Nofero per testare.

Storage PB-N – assegniamo 500TB per cominciare con xrootd.

Condivisione dell'esperienza con RAL.

FN: RAL contattati da me.

- SW in prod: gpfs aggiornato ovunque - versione stabile e sicure. Vedi slides per dettagli versioni.

Problemi:

- CMS: debug problemi xrootd – vediamo necessità configurazione limite che possono essere aperti dal processo xrootd, alzando limite con carico molto alto si riscontra cattivo funzionamento.

Per risolvere: aggiornato xrootd e rimossi limiti **senza successo**. Debug ongoing.

Ticket aperto da CMS: aggiornamento redirector – da discutere prima con Daniele.

Ticket rimozione files – azione lato exp.

Accesso alla Storage Area – test in corso tape http trasferimenti.

- Atlas:

Ticket trasferimenti durante down -> chiuso.

Ticket errori staging – problema configurazione -> risolto da Enrico.

- LHCb:

Ticket trasferimenti falliti -> problemi hw -> risolti.

Server non raggiungibile ipv6 -> problema lato nostro risolto.

Files rimossi appena finito il job -> exp deve rivedere il workflow.

**Tape:**

- Throughput ultimi due mesi un po' più alto. Quasi 300TB al giorno in lettura.

- Arrivati 5.2 PB di nastri; installati questa settimana. Siamo a pledge.

- Problemi vecchi per mancanza di tape su libreria nuova adesso sono ok - scritture erano dirottate sulla vecchia. Adesso sono tornate sulla nuova libreria. Info in slides.

- Tape challenge 11-15/10 in tabella rate run3. Verifica dei dati durante il challenge.

Nel totale scrittura 5.14 GB/s e lettura 5.72.

Throughput r/w oltre 7 GB/s.

Dettagli slides.

- I test – rate complessivi nei plot – sono finiti il 15/10 ma in realtà un po' dopo per letture finite dopo.  $w > 4$  GB/s e  $r = 2.3$  GB/s. Però non state effettuate contemporaneamente.

- Scritture LHC in slide. Riportato throughput da buffer a tape. Rate superiori ai target richiesti dagli exp.

**Importante capire per quanto tempo l'exp ha bisogno che il massimo possa essere soddisfatto al CNAF.** Non è detto che rate resistano su lunghi periodi i.e. 2 settimane.

- Letture in momenti diversi – la somma non è rilevante. CMS target non toccato sempre.

Per ATLAS superato abbondantemente.

- Test utili. Organizzazione migliorabile.

- Tempistiche non chiare. Non grande coordinamento tra gli exp.

DA chiarire scritture LHCb. Aspettiamo altro test.

Nofero: rate per 3 mesi. Trasferire 20 PB in tre mesi.

DC: forse anche per agli exp, target non raggiunto per carico basso.

EF: Alice problemi mancanza cassette.

Comunque non chiaro. CMS atlas oltre target. Alice più basso parte iniziale.

SPer: 2.24 valore di picco in presa dati mentre 1.2GB è la media in un anno. Il tempo è difficile da calcolare ma si oscilla tra questi valori.

In termine di lettura circa 1 GB/s per un paio di mesi – reprocessing di fine anno. Numeri nel 19/2 cdg LHCb. Torna con i numeri del tape challenge.

Sui problemi lati exp – Chris – EOS dati non spediti. EF: Pochi gateway schierati. (mancanza hw lato exp)

### Esperimenti:

- **ALICE**

- Utilizzo risorse – ultimo mese bene – girato a pledge, anche se risorse ridotte al T1.
- Iniziamo a usare multicore. Disco siamo sotto-pledge.
- Test beam a fine ottobre: nessun impatto.
- Stiamo configurando servizi xrootd per CEPH. Sarà creata un'istanza dedicata, da mettere in prod).
- WShop a Frascati a dicembre.

- **CMS**

- Utilizzo risorse: plot slide. Storico 3 mesi. Tutto ok.
- Efficienza: 80% media in tutti i Tier1 di CMS, ma CNAF sopra la media.
- Condor e Token: lato CMS ok. CNAF primo sito europeo. Si passa dai test manuali a sistematici. Setup integration testbed con un gruppo in cui tutto è pronto.
- Debug XrootD: Slide + spiegazione Lucia parte storage.
- Attività di R&D su infra per l'analisi. Challenge: data flow. Capire la cosa più efficiente ed economica.
- Marconi100: sottomessa la call per il proposal CMS. Sottomessa, da validare.
- Review tecnico-scientifica.
  1. Utilizzare parte CPU replicando ciò che fatto per grant precedente.
  2. Dataset per training ricostruzione end to end per il calorimetro elettrico di CMS - basata su machine learning.
- Novità: parte GPU. Che per ora non serve per validazione della fisica. In slide gli obiettivi.
- GPU per validazione grant.
- Due problemi: la piattaforma x86 (fase autodiscovering non funziona per mancate librerie, Tommaso dovrebbe aver risolto) e Match making – colloquio con condor CMS risolutivo → nelle prossime settimane tutto pronto da testare.
- Tape: da capire i numeri per tempi a max rate – poi comunicati ad Enri. Però sembra tutto ok.

SDP: configurati condor ce per accettare job con GPU – wn con due k40. Utilizzati da Atlas in passato. Test sarebbero utili per noi.

DS: necessità è che i piloti che arrivano al CNAF devono appartenere al gruppo che usa GPU. Inoltre CMS match making dà priorità a GPU. (non sono ancora risorse pledged)  
Comunque va bene – sentirsi offline.

- **LHCb**

- Operazioni as usual. Solito montecarlo e job utenti.
- Campagne 15/16 reprocessing fermate per problemi di validazione. Ma impiegheranno poco tempo.
- Efficienza elevata.
- Situazione CPU: under-pledge in linea con quello generale della farm.
- La migrazione di condor ha avuto un impatto quasi irrilevante.

- **NO-LHC**

**Slide preparate da reparto US.**

- NEWCHIM: Stefano Zani: capito e risolto comportamento anomalo router di frontiera CNAF verso LHC-OP1. È stato molto utile, anche se complesso.
  - Virgo nessun problema – Luca Rei: grid è ok.
- Parte online: grossi avanzamenti, ma nessun problema. Tutto funziona correttamente.