

CINECA - update



Dove siamo?

- CE funziona (grazie StefanoDP)
 - Confermato da CMS e da ALICE
 - ALICE: vedere slide di backup, presentata anche nel report ALICE oggi!
- Dress rehearsal fatto da CMS con jobs di analisi (vedere slide CHEP2019)
- Ora il problema e' soprattutto in mano agli esperimenti
 - Come integrare il CE nel loro WMS, e cosa mandarci
- Rete: piu' complicato
 - Diverse iterazioni (Vladimir, DanieleC, StefanoZ)
 - Cerco di fare il riassunto del punto nella prossima slide
- News generali CINECA HPC
- Ma prima di tutto ... slides CHEP2019

SLIDES CHEP

- Le trovate a <https://indico.cern.ch/event/773049/contributions/3474805/>
 - (per la lista dei nomi, ultima slide - lista temporanea, quella che conta sarà quella dei proceedings)
- Talk ben ricevuto, con un sacco di domande
- Mostrato il funzionamento di CMS in full dress-rehearsal
 - Jobs di analisi sottomessi con CRAB via CNAF CE
 - utilizzando la cache XDC al CNAF
 - Report back alla dashboard
 - (ma usando ancora solo GPN)

The system in action

- Grant ends by April 2020; plan is to start large scale productions @ Xmas
- Complete dress rehearsal already done with CMS:
 - Submit via CRAB to the CMS HTCondor Pool, directing to the CE @CINECA (logically belonging to CNAF T1)
 - Slurm starts singularity as a pre-exec (optional, some experiments can run the pilot also in the host system)
 - Payload starts from CVMFS and reads data from the CNAF XrootD proxy cache provided via the XDC EU Project
 - Results are stored on CNAF's storage
 - All the communications including MONIT callbacks in place

```
===== gWMS-CMSRunAnalysis.sh STARTING at Mon Oct 7 18:47:00 CMT 2019 on r105c08103 =====
Local time : Mon Oct 7 18:47:00 UTC 2019
Current system : Linux r105c08103 3.10.0-327.36.3.el7_ppcl 1.5.1.4151.x86_64 #1 SMP Tue Feb 7 22:48:59 UTC 2017 x86_64 x86_64 x86_64 GNU/Linux

MARN $TMPDIR/marconi_scratch/usera07cms/a07cms00/slurm_job.544552/q/ide_U732qe/execute/dir_920) is not a writable directory setting $TMPDIR = $PWD
===== gWMS-CMSRunAnalysis.sh STARTING at Mon Oct 7 18:47:00 CMT 2019 on r105c08103 =====
Local time : Mon Oct 7 18:47:00 UTC 2019
Current system : Linux r105c08103 3.10.0-327.36.3.el7_ppcl 1.5.1.4151.x86_64 #1 SMP Tue Feb 7 22:48:59 UTC 2017 x86_64 x86_64 x86_64 GNU/Linux
```

Read data from XCache:
root://ds-202-06-19.cr.cnaf.infn.it:31094/

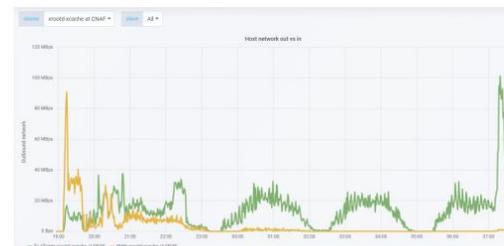
Write output directly @ T1-CNAF:
srm://storm-fe-cms.cr.cnaf.infn.it:8444

== JOB AD: Used Gatekeeper = "r000u1106-fe.marconi.cineca.it r000u1106-fe.marconi.cineca.it:9619"

The proxy cache in action...



- Based on **XDC**(EU project) Xcache implementation
- **Yellow**: starting from cold cache, files are loaded into the cache
- They are served in successive runs (**Green**)



- Given the design where files @ CNAF do not pass through the cache, **1 host @ 10Gbit/s should suffice**
- Cache is needed here mostly in order to **overcome routing limitations**, and to allow accessing the whole federation (and not only CNAF and CERN files)

Rete

- Come al solito: il routing aggiuntivo rispetto alla GPN *_non_* aggiunge funzionalità / maggiori accessi lato CNAF, aggiunge solo banda
- Varie iterazioni, disegno sperato (4 macchine di frontend che facciano routing a 4x10 Gbit/s verso reti CNAF) non possibile
 - Comunque StefanoZ e' a SC con CINECA, sta discutendo con loro se sia solo un prob politico
 - Inoltre, c'era un problema non risolto di come fare routing attraverso 4 macchine
- Disegno suggerito adesso e' leggermente piu' complesso ma completamente trasparente lato CNAF
 - Le 4 macchine (con una eth10gbps verso CNAF, usando Infinera) montano xrootd proxy w/o cache (come nel disegno XDC con TSystems)
 - I WN dovranno lato client fare round robin (previsto dal protocollo Xrootd, con sintassi tipo `root://host1,host2,host3,host4//path/to/file`)
 - Scrivere via GPN su SRM
 - Ok per CMS, per gli altri? (NOTABENE: via GPN continuerà a funzionare tutto)
- E' molto vicino a quello che Vladimir e Daniele avevano suggerito, senza disco

ALICE - vedere report specifico!

- Simulazione MC con software O2
- Deployment ufficiale del framework O2 da CVMFS
- Deployment ufficiale di immagini singularity fornito da CINECA
- Sottomissione su coda HTC via alice VO uihtc-alice (grazie a Stefano Dal Pra)
- Ad ogni sottomissione viene allocato un intero nodo (68 cores x 4 HT)
- Misure di performance in funzione del numero di processi concorrenti/istanze di simulazione per ottimizzare la configurazione di simulazione
- Utile esercizio per l'ottimizzazione del software di simulazione in O2 e per i test di scalabilità
- Risultati mostrati Mercoledì 20/11/2019 in un meeting dedicato:
<https://tinyurl.com/tng7k5z>