INFN-T1 site report

Andrea Chierici

On behalf of

INFN-T1 staff

Outline

- Infrastructure and common services

- Data management & Storage

- Farming

- Projects and activities

Andrea Chierici

Technopole

- Migration to a new, bigger location
- Close to CINECA and ECMWF
- The next gen Italian T1 DC
- Storage resources will be completely hosted
- "Half" of computing power will be borrowed by CINECA (aka Leonardo, aka pre-exascale)
- General setup will be more cloud-oriented

Infrastructure and common services

Atlassian tools

- Adopted atlassian tools for ticketing and documentation
  - ServiceDesk
    - User requests
    - Internal tasks
  - Confluence
    - General documentation and internal how-tos
- INFN-wide lincensing

Collaboration tools

- Heavily using Microsoft 365 tools
  - "MS Teams" has quickly become the most used application
  - Collaborative editing of docs and spreadsheets
- Zoom adoption is increasing (INFN-wide license)

- Thanks to these tools, impact of social distancing was reduced in our activities

Other tools

- OpenDCIM is considered the best candidate for the new asset management tool for INFN-T1

- Open-audit is being tested as audit tool for software inventory

- Refactoring of AAI system
  - Experiences on new tools made in "EPIC Cloud" context
  - Technologies will probably involve FreeIPA, Keycloack and Indigo IAM

Data Management and Storage

Storage resources

- Disk
  - 41 PB installed – 36 PB used
- Last tender: installation ongoing
  - 8.7 PB usable with GPFS back-end (10 TB disks)
  - 5.2 PB raw to be used as CephFS back-end for a pre-production cluster (18 TB disks)
- Tape
  - 96 PB installed – 86 PB used
  - 2 libraries
    - Oracle SL8500, almost full
    - IBM TS4500 in production since Feb 2020, 15 PB installed

Storage services

- Migration from gsiftp to https/XRootD, in view of Globus retirement
  - WLCG DOMA TPC working group initiative
  - Token-based AuthN/Z fully supported on our StoRM WebDAV endpoints
- Deploying several XRootD caches
  - XCache (caching) proxy used also in support of the INFN-T1 extension to a CINECA Marconi KNL partition
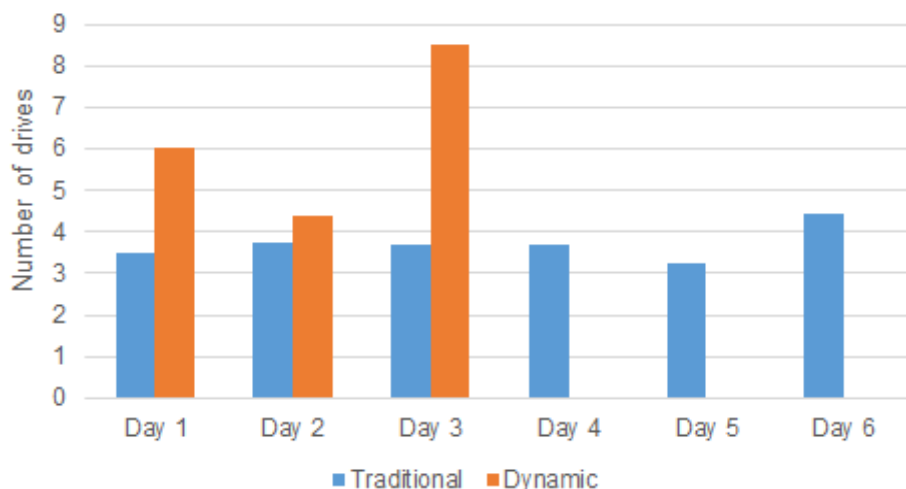 - network connectivity from CINECA to other sites through XrootD at INFN-T1

Dynamic tape resources allocation

- All tape drives shared among all the experiments

- Traditional allocation: static partitioning, manual intervention in case of bulk recalls

- Dynamic allocation (since Jan 2020): drives are automatically allocated based on monitoring information about free drives and pending recalls

- Comparison of two bulk recalls (~300 TB) of CMS (traditional vs dynamic allocation)
  - Increased average throughput by 85%
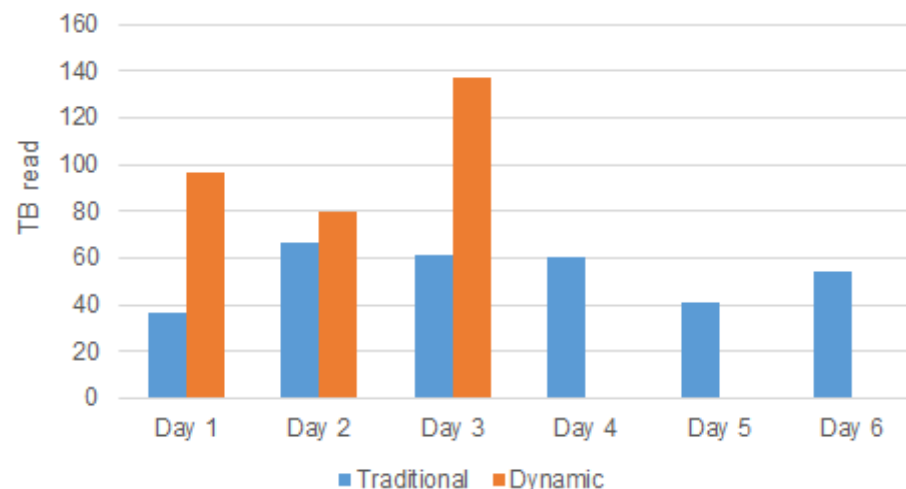
Dynamic tape resources allocation

Sample comparison: CMS real bulk recalls. Similar number of files and TB read.

Max drives used per day:
4.4 traditional vs 8.5 dynamic

Max TB read per day:
66.5 traditional vs 137 dynamic

Farming

Computing resources

- ## Farm power: approx. 440 KHS06
  - 250 kHS06 at CNAF
  - 180 kHS06 at CINECA
  - 10 kHS06 at Bari-ReCaS

- ## Just installed latest tender
  - 42 KHS06 (Epyc 7282)

- ## Within summer we plan to add 240 KHS06
  - New batch of nodes from CINECA (same configuration as current one)
 - We will reach 2021 pledge and will be able to dismiss very old hardware

Andrea Chierici

Batch System

- 6xHTC-CE, 1xSN, 1xCM, ~870WN, 39Kcores
- LSF phase out completed on Jun 2020
- HTC 8.9 testbed to check TokenId auth*n and docker-wn

Andrea Chierici
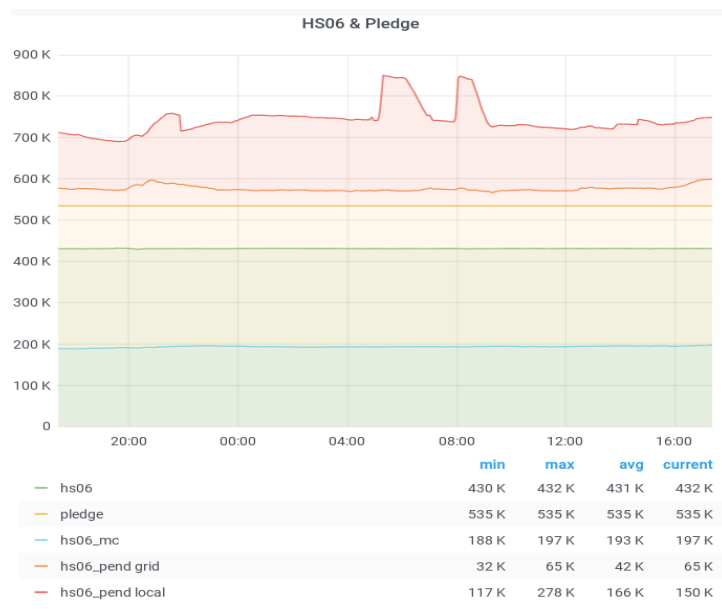
LHC jobs on HPC resources

- Goal: Run typical HEP job workloads on HPC resources, using ordinary Grid Middleware

- Use case: **30M** hours on Marconi A2 HPC cluster (CINECA):
    - INTEL KNL, 68 core x 4 thread = 272 slot
    - 96GB RAM, (10 reserved to the OS)
    - Batch System: Slurm
    - Outgoing connectivity only toward CERN, CNAF

- "whole node" jobs only
    - **A singlecore job is accounted 68 times its WCTime.**

- The initial grant was extended (expired on 28 Feb)

- Total usage: **93M** hours

Core h by VO
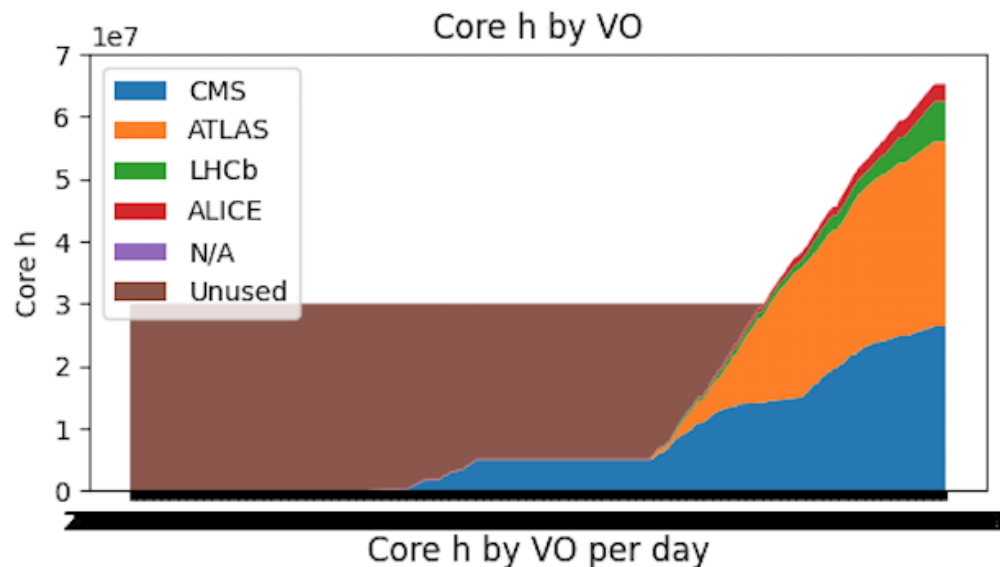
- From a CMS report:
- 92% of successful jobs is compatible with what seen at more standard sites
- CNAF seen as a processing site able to serve > 17k threads (~7k at CNAF and 10k at CINECA).

Developments

- Bare-metal management via Ironic
  - Gain flexibility in handling computing resources
- K8s on OpenStack
- HTC and HPC farms integration
  - Resources provided via Jupyter notebooks
- Dynamic farm expansions
  - HTC, K8S, OpenStack

Other projects and activities

EPIC Cloud

- **E**nhanced **Pr**I**vacy and **C**ompliance
- Our ISO27001 certified zone
  - During these days we are re-certifying (change of scope)
- We provide computing resources to some "high security demanding" medical research collaborations
- We had to learn new work methodologies
  - Very useful in Technopole transition

Cloud@CNAF

- Increased available resources
  - Added GPU machines
  - New storage based on Ceph under deployment in one of our cloud region.
- Some infrastructure upgrades ongoing (hardware and software)
- Cloud@CNAF is now federated with INFNCloud.
  - CNAF is one of the two backbone sites of INFNCloud (Openstack geographically distributed)
  - Cloud@CNAF is now federated in order to provide scarce resources like GPU
- Openstack upgrade to Train/CentOS 7 scheduled
  - Testbed already running
  - Waiting for CentOS 8 situation to "clarify" before Ussuri/Victoria upgrades
- See https://www.cloud.infn.it

Questions?