

20 September 2024

Report LHCb

Lucio Anderlini
INFN Firenze



Matteo Barbetti
CNAF



Long-summer short

- Mid-July, LHCb started a **reconstruction campaign** involving all Tier-1 sites
- End-July, the **transfer of data to the new OceanDisk** in Tecnopolo, caused a temporary downtime (1 week)
- Early August, **LHCb tried to recover the backlog** sending and processing data to CNAF at the maximum agreed bandwidth (1.7 GB/s)
- Late August, the OceanDisk designed for thin-provisioning Cloud storage had degraded performance due to large occupancy, this made transfer to tape slower, until the file-system was completely full and **CNAF declared downtime** (snowball).
- Early September, with CNAF still in downtime, solutions to resist to the intense IO during reconstruction campaigns are being studied (*e.g. tape buffer on separate storage*)

What we have learnt so far

- While mitigations exist, OceanDisk may become **unpredictably slower** when the underlying storage is full. *Full filesystem full \neq underlying storage.*
- It is possible to identify **hot regions** of the LHCb filesystem and use GPFS placement policies to use dedicated, smaller hardware. Successful tests with **tape buffer** (possibility of migrating other hot directories under investigation).
- The number of **reconstruction jobs must be limited**. CNAF will never cope with the whole farm dedicated to reconstruction jobs (*and that is ok!*). However the current limit (1200/20000 jobs) is too tight. Tuning this limit is challenging because it depends on FTS activity.
- **Storm-webdav** manages all the write operations (FTS-TPC and PUT). Upon heavy load of GPFS, it saturates the available threads and starts failing. Unfortunately, errors are reported only partially in the *log and the monitoring is a bit obscure*, which does not help operations.

This week

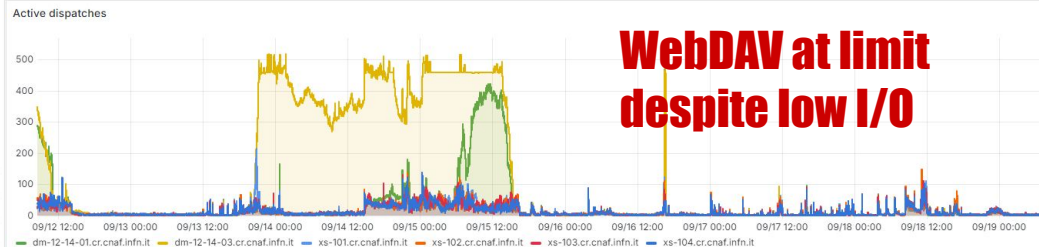
Good news

LCG.CNAF.it (ES) Back-log recovered

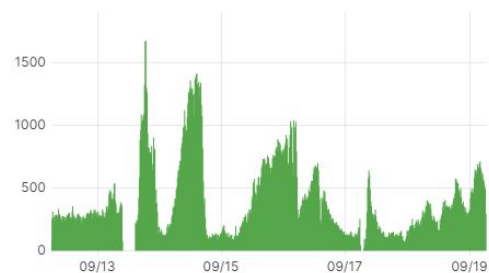


Bad news

WebDAV at limit despite low I/O

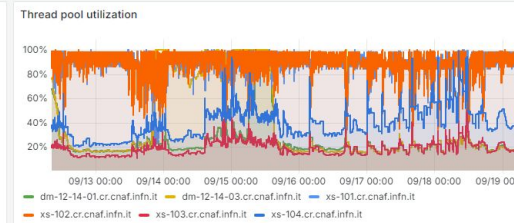
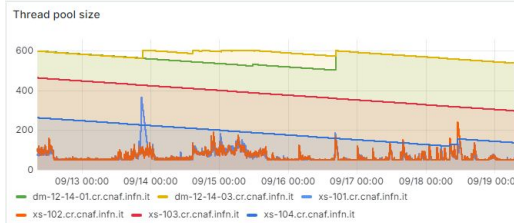


Successful upload from LCG.CNAF.it



Failed upload from LCG.CNAF.it

Low error rate



GPFS speed Migration at 1.7 GB/s, buffer empty



What next?

LHCb has paused CNAF to work off the large backlog, we are now ready to restart.

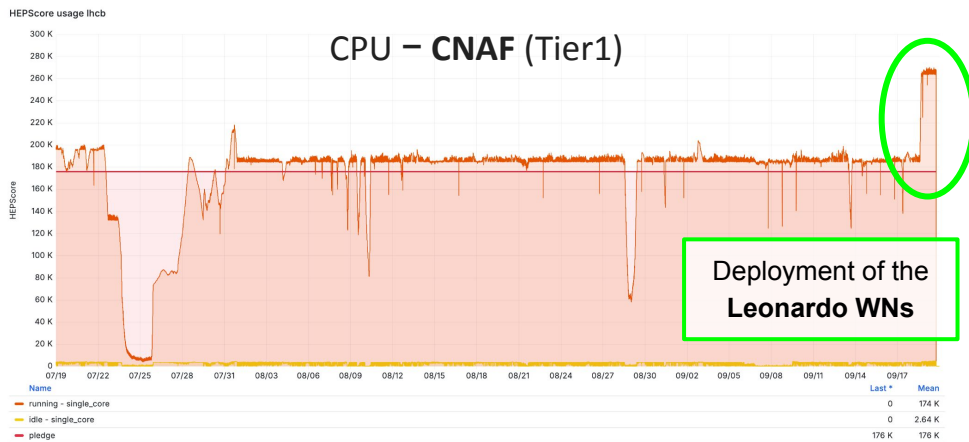
LHCb may restart this afternoon its “normal flow” at CNAF.

To be understood: the weight of “Working group production” jobs in terms of IO. Most probably, not negligible → some tuning will be needed

This will allow for full test of the mitigation strategy put in place and will allow to evaluate the need for additional effort (e.g. hot storage area for /gpfs_lhcb/disk/lhcb/buffer).

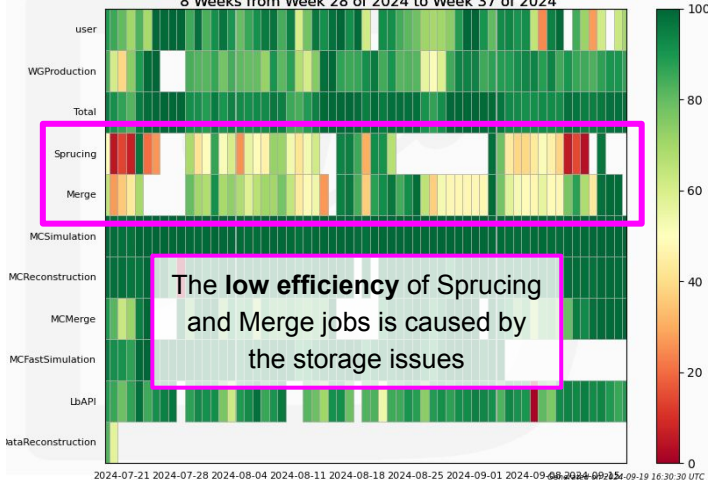
A word on CPU

Despite the storage problems of August, the use of CPU resources has remained at **nominal values** even if it has required to restrict the number of Sprucing and Merge jobs.



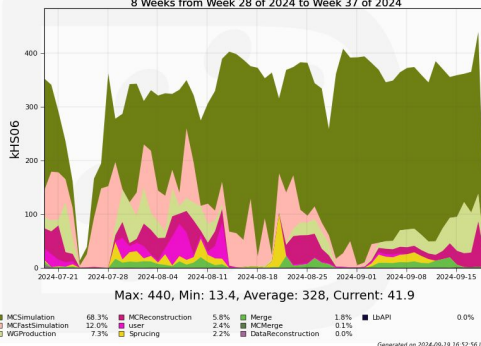
Job CPU efficiency by JobType (CNAF)

8 Weeks from Week 28 of 2024 to Week 37 of 2024



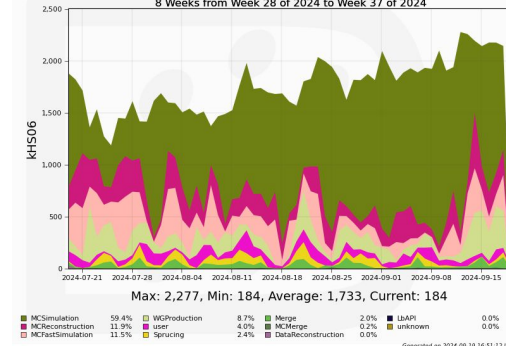
from DIRAC: CNAF (Tier1 + Tier2)

Normalized CPU usage by JobType (CNAF)
8 Weeks from Week 28 of 2024 to Week 37 of 2024



from DIRAC: all Tier1 sites

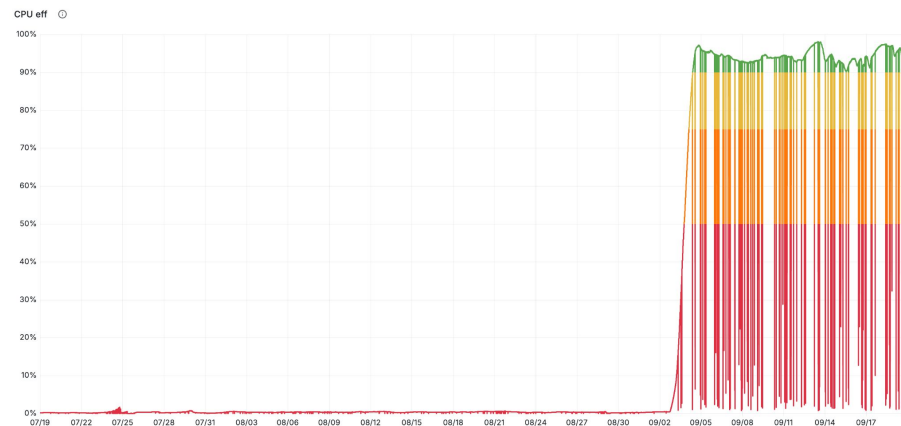
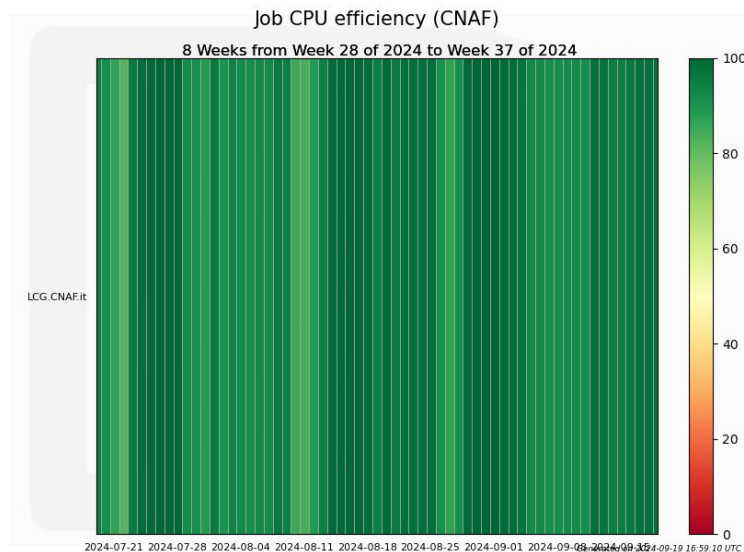
Normalized CPU usage by JobType (T1 sites)
8 Weeks from Week 28 of 2024 to Week 37 of 2024



[SOLVED] Monitored efficiency inconsistency

With the migration to HTCondor 23, the monitoring of the efficiency as measured by DIRAC and Tier1@CNAF was **inconsistent**.

The inconsistency resulted from an **interaction problem** between Singularity and `cgroups`, that prevented HTC from tracing the Singularity process.



Conclusion

Summer is over...

Thanks to the CNAF team for the great effort trying to cope with such a new workflow on a new hardware.

