DOMA mini DC

A. Forti, K. Ellis

DOMA

13 November 2024

Why miniDC?

- DC24 considered by everyone as a successful example of an exercise that focuses activities and helps solve problems and optimize infrastructure
  - However it was a very big test trying to test several things at once.
  - DC26 will be twice the rates
- We agreed after GridPP51 to have smaller, more frequent UK-only miniDCs
  - Goal to prepare the UK infrastructure for the next challenge in 2026 or likely 2027
    - No pressure
- miniDCs should be organised in a way that is easily repeatable when a new technology/hardware/network becomes available
  - i.e. either on demand or once a year

RAL tape

- T0 export is supposed to write directly to tape
- DC24 did not explicitly use tape
  - Only LHCb and ALICE tested tape but they wrote and read through echo
- post-DC24 ATLAS tried to write directly to Antares but we found a simple switch from from *disk* ➔ *disk* to *disk* ➔ *tape* wasn't optimal
  - Fast injection/deletion cycle didn't work for tape
- Antares is connected to JANET not LHCOPN yet
  - Gateways are behind firewall
- Do we consider further tape tests useful without LHCOPN?
  - yes, we want to practice a procedure tailored for tape

RAL echo

- DC24 used echo heavily and echo rates improved notably (see previous slide)
- Are there still questions on concurrent experiments traffic?
  - Probably yes ➔ need to test concurrent rates
- Do we want to test SDNs (NOTED/SENSE)?
  - Contingency plan for losing a link
    - wasn't needed during DC24 when it happened
  - More complicated to organise it will require more coordination with CERN and other T1s

UK T2s

- ATLAS tested the T2s for ~20h hours on the last day
  - Manchester, Lancaster and QMUL above production rates
- CMS tested only IC and RALPP
  - Bristol and Brunel were not available
- Changes since DC24
  - Manchester and Brunel have a new storage
  - Bristol has a new backend (HDFS -> cephfs)
  - QMUL is in a new data centre (might have a new front end too!)
- Glasgow, RALPP, Liv, Lancaster, IC, need to be tested more thoroughly even without changes
- Rates would be the same as DC24
  - Full mesh

Rates

Site	Ingress CMS	Egress CMS	Ingress ATLAS	Egress ATLAS	Ingress LHCb	Egress LHCb	Ingress total
T1 disk	4.15	3.55	7	9.5	3.96	2.4	15.11
T1 tape			7.45	2.7	3.96	2.4	11.41
RALPP	0.76	0.22	0.55	0.51			1.31
Manchester			3.85	3.53			3.85
Lancaster			3.5	3.23			3.5
QMUL			3.38	3.03			3.38
Glasgow			1.69	1.53			
IC	3.04	0.9					3.04
Brunel	0.48	0.09					0.48
Bristol	0.41	0.06					0.41

Note ATLAS tape in/egr: T0 export+data consolidation/carousel

- Replicate the DC24 rates from ATLAS, CMS LHCb
  - to start with

ATLAS

- ATLAS agreed countries doing their tests without intervening, i.e. they are not going to do any organisation.
  - Sites still need to agree the period so it can be monitored
  - So far only US and UK are proactively talking about further tests
    - Italy was also interested but have bigger fish to fry right now
- dc_inject will be used
  - Alessandra will do the injections

CMS

- Katy has agreed a mini-DC for UK sites with CMS Data Management
- A small number of non-UK CMS sites have also requested additional testing at some point
  - This could be done simultaneously with UK test, but doesn't have to be
  - Regardless of this, Katy's attention needs to be on UK sites
- Will run the dc_inject tool as used in DC24

LHCb & SKA

- As with DC24, LHCb are:
  - Only planning to test tape endpoints, i.e. Antares
  - No T2 testing
- RAL has some 100Gb gateways to test
  - Originally thought as an SKA test but these will be used also by WLCG
  - The mini-DC is a good chance to test these in different configurations
    - Can be done using CMS (or ATLAS) data.

Dates & proposed tests

- Need to be careful not lose the “mini” along the way doing too many things in one test
  - Agreed to split T2 and T1
- **T2 tests:** Week starting the 9th December 2024
  - Stress-testing T2s
    - Replicate DC24 rates ➔ using only UK sites ➔ for at least 6h
    - Increase rates to max bandwidth at 100Gb/s sites
    - Under discussion enabling jumbo frames
      - Unlikely at this round
- **T1 tests:** February 2025
  - Writes to Echo
    - Cutting LHCOPN to RAL; what rate can be achieved through LHCONE?
    - Test a SDN solution (under discussion)
  - Writes to Antares
    - Direct from CERN (ATLAS, CMS)
    - Multihop via Echo (LHCb)
  - Stress test 100Gb boxes at RAL
    - One writing to Ceph
    - One writing to CephFS

Note on dc_inject

- DC24 dc_inject version used for this miniDC
  - Useful experience from non experts
- ATLAS found manpower to rewrite the tool
  - Xuantong Zhang
- Main goals are
  - Integrate it in rucio
  - Adequate protections for the database access
- Requirements being collected
  - Please check Xuantong slides

People

- ATLAS: Alessandra Forti
- CMS: Katy Ellis
- LHCb: Alexander Rogovskiy
- echo team
- antares team
- T2s: Usual Suspects 😊