

# DOMA/ACCESS data lake tests

Nikolai Hartmann, IDT-UM storage meeting

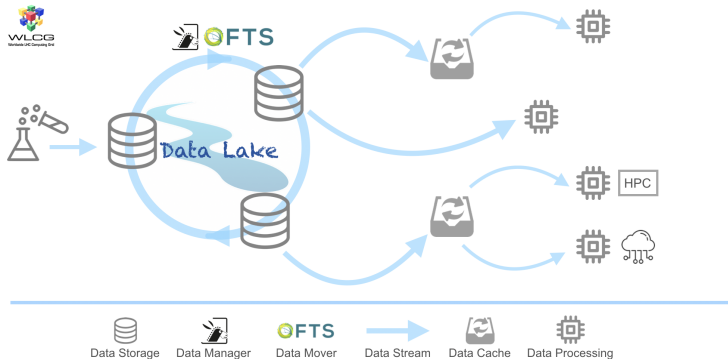
LMU Munich

November 26, 2020



# The WLCG data lake model

Consolidate (managed) storage to fewer large storage centers



Data lakes, latency hiding and caching - Xavier Espinal (CERN)

- Less maintenance effort and cost, trade storage against network I/O  
→ bandwidth expected to increase more than storage
- Allows sites to run computing only, integrate heterogeneous resources
- Caches can help to reduce traffic (for reused data) and hide latency

# Current status/plans for ATLAS

- As a start: one datalake is one RSE (Rucio Storage Element)
- Possibly add on top data lake grouping on Rucio level
  - Mainly for data placement
  - E.g. have one DAOD\_PHYSLITE replica in the “EU data lake”
- XCache (XRootD Proxy File Cache) most studied cache system (not only at ATLAS)
- Test “Virtual placement” (VP)
  - virtually place datasets to cache-only sites to make PanDa send jobs for the same data to same sites (to increase data locality/cache hit rate)
  - yet to figure out how to decide which datasets to assign where (currently random, locality not taken into account)

# Tests planned

Next step (start this year): larger scale remote I/O tests

- Larger scale: ramp up e.g. starting with 100 to several 1k jobs reading remotely
  - first low I/O like simulation
  - next higher I/O like DAOD production, Analysis
  - try running on multiple sites in parallel to test if network saturation effects are limiting
- Compare directIO with copy2scratch where reasonable
  - directIO with TTreeCache enabled
  - compare reading through an XCache
- Possibly use Hammercloud for coordination of test jobs (stress test)
  - currently figuring out technical details
- Try different “Distances”, e.g first within Germany then within Europe
- Observe if/how large efficiency is degraded, where are the bottlenecks

Future (next year):

- Analysis-Facility test
  - maybe an uproot analysis reading a “full set” of DAOD\_PHYSLITE?
- Remote I/O from HPC site

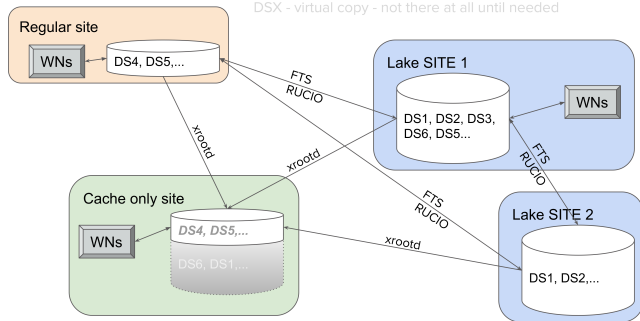
# Backup

# Virtual placement

DSX - primary copy

**DSX** - virtual copy fully or partially cached data

DSX - virtual copy - not there at all until needed



(plot by Ilija)

## Virtually place datasets on cache only sites

- Compromise somewhere between managed storage and no management at all
- Ensures site gets jobs for same data again
- Can increase hit rates