DOMA Phase 2

Christoph Wissing, Maria Girone, Mario Lassnig

GDB Meeting

2021-07-14

Some pre-remarks

Similar presentation was given during the DOMA general meeting on June 30th

- Only minor updates since then

Status is present "as is"

- Discussions are still on going

- Feedback is welcome

DOMA overview

Looking at the evolution of the structure of the DOMA R&D for Phase 2

Phase 1   Forum to discuss and foster ideas, define **prototypes** and **explore**

Phase 2   Realizing promising **prototypes into actual services**

Commission the services in the **production infrastructure** during Run-3

New team of co-coordinators

Two-year mandate with some overlap to ensure continuity

Christoph Wissing (DESY CMS), Mario Lassnig (CERN ATLAS), Maria Girone (CERN IT)

Current state

Currently we have three DOMA working groups: Access, TPC, and QoS

- DOMA Access wrapped up main activities with a proposal for next steps
- Many diverse topics have been folded into the TPC working group
- QoS working on several demonstrators

TPC and QoS have bi-weekly interleaved, with DOMA General once monthly

A few extra thoughts and feedback we got

- We see DOMA-related activities in the experiments and facilities, which are not present in DOMA meetings, where we could benefit from bringing them under the same umbrella
- We do not want to change things just for the sake of changing
- We observe that the TPC WG works well with a big list of diverse topics
- Combining and optimising our use of facilities will be key

Reorganisation proposal

This leads us to the point that we need to come up with an idea how Phase 2 could look like

General idea

- Based on the received feedback, **identify the topics** people *should / can / want to* work on

- Set up a **structured approach** towards both Run-3 and Run-4

- Break working group boundaries, be **more flexible** and dynamic

- Bring relevant people together

- Define and follow **measurable objectives**, their key results, and their milestones

DOMA Topics

Wide area networking, transfers and throughput

- Demonstrate that we can move data and utilize networks at the scale required for HL-LHC
- **Natural continuation of present TPC group activities**

    There are ongoing technical activities from phase 1 with milestones by the end of 2021

    Xroot remains central to a number of common things

- Transfer challenges
- Include network developments (SDNs and interfaces to them) and liaise to NET-WG

    Dynamic network provisioning

- Common software activities

    RUCIO   (ATLAS+CMS+beyond)

    FTS   (ATLAS+CMS+LHCb+beyond)

- Complete WebDAV migration (almost done!)
- TAPE interface via HTTP

    gsiftp/SRM decommissioning

DOMA Topics

Commissioning of a data delivery infrastructure

- Leveraging activities of Phase 1 QoS and Access groups, plus experiments & facilities work
- Setup testing infrastructures
- Cache layer deployment
- Assess operations effort and performance
- Liaise to the AAI WLCG WG and token transition
- **Production challenges with prototype setups**
- Storage performance (including archive/tape, disk) incl. QoS classes
- "Analysis facility"-style access incl. liaision with HSF WGs
- Asses files/filesize challenge (10 billion files vs 100GB files)
- Cost modeling and storage accounting

DOMA Topics

Integration with heterogeneous data resources

- **Align existing experiment-specific data solutions towards a common strategy**
- Data management aspects
    - Including HPCs and clouds with experiment DM systems
    - Cache layer deployment for heterogeneous resources integration
    - Related performance assessment
    - Define/execute data challenges
    - Identify testbed sites
- Liaise with ongoing coordination efforts (HPC, Commercial Clouds)
    - Cybersecurity and AAI
    - ASCR/EuroHPC/PRACE/ HPC Collaboration, ...
    - Amazon, Google, Cloudbank, ...
- Define WLCG-wide QoS classes (advertise, costs, accounting, pledges, …)
- "Self-organising" storage (MAS/BNL + Rucio)

Next steps

Continue discussions regarding structure so we can put it in effect during summer

- All feedback is welcome!

Activity survey of all ongoing efforts across the facilities and experiments

- First kick-off mails sent out – some more to come

Identify the most pressing objectives from each topical area

- Assess the experiment's needs, available personnel, and remaining time
- Define the quantifiable results to be reached for Run-3 commissioning
- Based on this, draft the long-term Run-4 plan as far as possible

Remarks

DOMA is no project

- It has no budget
- There is no hierarchy
- Decision and directions have to come by consensus

Success of DOMA depends on engagement and willingness to collaborate

Synergies will payoff

Commonly established infrastructure will be more sustainable