Data Challenges & Monitoring WS

A. Forti

WLCG Ops Coordination

1 April 2021

Data Challenges

- As part of the planning for HL-LHC [Data Challenges are foreseen](#) from this year every 2 years until the start of HL-LHC.
- Increasing transfer rates until we reach the required level for HL-LHC
  - 10% of the target 2021, 30% in 2023, 60% in 2025 and 100% in 2027
- The overall aim is to prepare the WLCG infrastructure and demonstrate that we can use the bandwidth effectively
  - "The DC should follow and drive the expansion of the network capacity"

DC: targets

	Data Challenge target 2027 (Gbps)	Data Challenge target 2025 (Gbps)	Data Challenge target 2023 (Gbps)	Data Challenge target 2021 (Gbps)
T1				
CA-TRIUMF	98	59	29	10
DE-KIT	312	187	94	31
ES-PIC	93	56	28	9
FR-CCIN2P3	281	169	84	28
IT-INFN-CNAF	336	202	101	34
KR-KISTI-GSDC	25	15	7	2
NDGF	71	43	21	7
NL-T1	94	56	28	9
NRC-KI-T1	62	37	19	6
UK-T1-RAL	296	177	89	30
RU-JINR-T1	52	31	15	5
US-T1-BNL	227	136	68	23
US-FNAL-CMS	454	273	136	45
(atlantic link)	681	408	204	68
Sum	2400	1440	720	240

- Transfer rates out of the T1s based on the experiments needs.
  - Rates may be revised over time

DC: infrastructure

- All 4 experiments in parallel and all sites
- Use experiments production infrastructure
  - Storage, network, services, tools,....
- In parallel with experiments activities
- DC not just demonstrators
- DC different from but overlapping with
  - Exp workflows (local and remote stage-in/out, streaming)
  - Data Carousel (tape stage-in/out & data distr T1 - T2)
  - Networking R&D
  - HPC challenges
- All use networking and eventually we will have to integrate everything
- But approach should be step by step
  - Develop the tests methodology
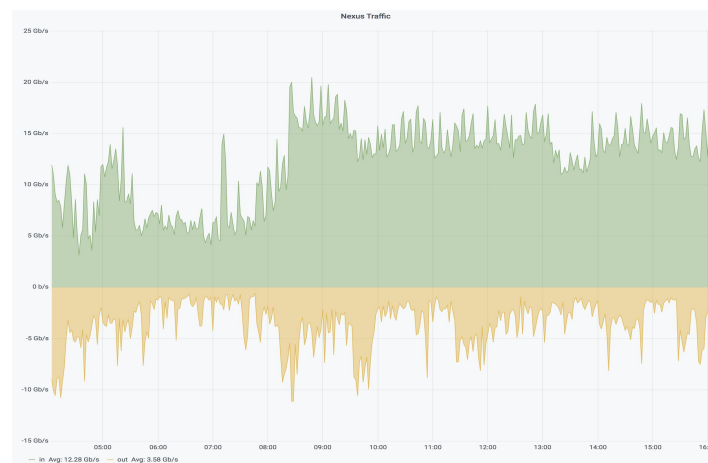 - Work in parallel if we can

DC2021: HTTP-TPC

- DC2021 will also be used to commission HTTP-TPC
- Currently being enabled by ATLAS, CMS & LHCb as production transfer protocol replacing gsiftp
- TAPE transfers also with HTTP in the future
  - Data carousel in ATLAS ~11% of traffic

DC2021: Monitoring

- Most important part to get right for DC2021
  - No monitoring blind transfers
- Monitoring has to be useful also in the future
  - Part of common infrastructure
- Do we have all the monitoring we need? 🧐
  - Do we have all the metrics we want to evaluate
  - Is the aggregate WLCG transfer monitoring enough?
  - Can we integrate more information?
  - Other DOMA activities may require integration
    - tape activities, network tests
  - Sites have network bandwidth can we collect that
    - perfsonar not the right tool for this
    - We need real time traffic at sites

PIC monitoring for NOTED

Monitoring WS

- Other DOMA activities like network tests, tape stage-in challenges may require dashboards
- Monitoring keeps on being mentioned at all these meeting as needing to be improved
- Monitoring workshop idea was proposed at the last DOMA general meeting to have a coherent picture of what we might require.
  - For DC2021 we already said monitoring is the part we need to have in common

Common effort

- Data Challenges are by definition a common effort that requires the involvement of experiments, network people, monitoring and sites

Communication

- Email wlcg-doma-tpc@cern.ch
- Or chat Mattermost channel
  - Probably both
- Propose to kickstart activity with topical meeting at the next TPC meeting 3rd March
  - Meeting is now on 1st and 3rd Wednesdays

Links

- [HL-LHC network needs and data transfer challenges](#)