

WLCG network challenges

S. Campana (CERN)



ESNET REQUIREMENTS REVIEW
DOE Office of Science - High Energy Physics Program Office
June 2020 - February 2021

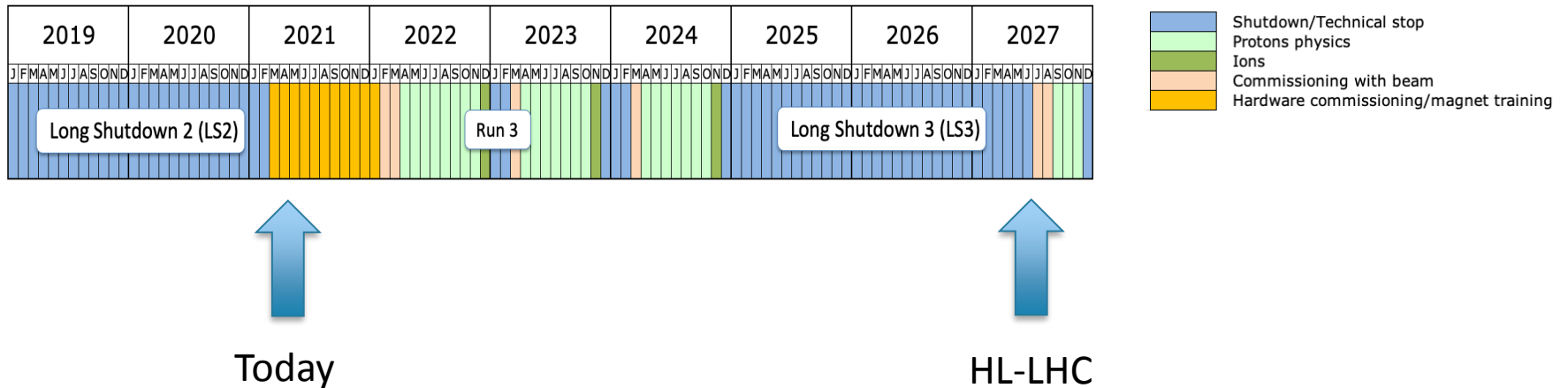
Review Purpose and Process

The primary purpose of an ESnet Requirements Review is to discuss and analyze current and planned science use cases and anticipated data output of a particular program, user facility, or project to inform ESnet's strategic planning, including network operations, capacity upgrades, and other service investments.

HL-LHC Software and Computing Review Committee Report

Network: The availability of a strong and reliable network as a backbone for the data lake model to work is of great importance. Close collaboration with LHC partners and with NRENs is required to ensure that the evolution of the networks is sufficient on the international, trans-atlantic, and national level.

WLCG network needs for HL-LHC



WLCG followed the ESnet review process and extended the findings to the whole of the WLCG infrastructure in preparation for HL-LHC

- Identify the main use cases for HL-LHC
- Define network needs based on those use cases
- Define a plan of data challenges to progressively assess the readiness of the infrastructure

Use Cases

Two use cases will drive the network use at HL-LHC. This is not necessarily the case today

1. Export of RAW data from CERN to T1s
2. Export of RAW data from T1s to T2s for reprocessing

ATLAS and CMS will produce 350 PB of RAW/year/experiment

- To be exported in \sim real time from CERN to the T1s
- To be reprocessed in \sim 3 months largely at the T2s

Network Needs for data export T0 to T1s

350 PB of RAW to be exported in 7M seconds => 400 Gbps/experiment (ATLAS, CMS) + 100Gbps / experiment (ATLAS, CMS) for other data formats + 200 Gbps for Alice and LHCb

Contingency:

- A factor x2 to absorb bursts
- Another factor x2 overprovisioning (avoid running at > 50% network occupancy)

⇒ 4.8 Tbps from CERN to T1s by the time of HL-LHC (1.25 Tbps transatlantic)

Network Needs for Reprocessing

The data at the T1 needs to be staged from tape and exported to the T2s for processing (assuming most of that processing happens outside the T1).

To estimate the network needs, we consider a scenario where 100% of the data collected in the year and stored at a specific T1 is reprocessed in less than three months.

Because 3 months \approx 7M seconds the previous scenario can be used to set the targets for T1 to T2s traffic as well.

Network needs for HL-LHC

	%ATLAS	%CMS	% Alice	% LHCb	ATLAS+CMS Network Needs (Gbps) Minimal Scenario in 2027	Alice Network Needs (Gbps) Minimal Scenario in 2027	LHCb Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Flexible Scenario in 2027
T1									
CA-TRIUMF	10	0	0	0	200	0	0	200	400
DE-KIT	12	10	21	17	450	80	70	600	1200
ES-PIC	4	5	0	4	180	0	20	200	400
FR-CCIN2P3	13	10	14	15	450	60	60	570	1140
IT-INFN-CNAF	9	15	26	24	480	110	100	690	1380
KR-KISTI-GSDC	0	0	12	0	0	50	0	50	100
NDGF	6	0	8	0	110	30	0	140	280
NL-T1	7	0	3	8	140	10	30	180	360
NRC-KI-T1	3	0	13	5	50	50	20	120	240
UK-T1-RAL	15	10	3	27	490	10	110	610	1220
RU-JINR-T1	0	10	0	0	200	0	0	200	400
US-T1-BNL	23	0	0	0	450	0	0	450	900
US-FNAL-CMS	0	40	0	0	800	0	0	800	1600
(atlantic link)					1250	0	0	1250	2500
Sum	100	100	100	100	4000	400	410	4810	9620

Minimalistic vs Flexible Scenario

The baseline scenario described so far is “minimalistic” (network is considered and used as a scarce resource).

The experiment computing models evolved between LHC Run-1 and Run-2 to leverage more the network to obtain flexibility in workload scheduling, increase efficiency and turnaround time of processing and analysis activities.

If we want to continue with such a model we estimate an additional x2 network capacity is needed.

Summary of Network Needs

In the flexible scenario, we expect the largest T1s (KIT, IN2P3, CNAF, RAL, BNL, FNAL) to be connected with CERN and the T2s (at least at regional level) through a ~1Tbps network

For the other T1s we expect ~500Gbps.

The numbers from the flexible scenario are coherent with what was presented at the LHCONE/LHCOPN meeting in early 2020.

The minimal scenario is coherent with the numbers presented at the ESNET planning of summer 2020.

Data Challenges

Data transfer targets are derived from the needs. Will be reviewed regularly

	LHC Network Needs (Gbps) Minimal Scenario in 2027	LHC Network Needs (Gbps) Flexible Scenario in 2027	Data Challenge target 2027 (Gbps)	Data Challenge target 2025 (Gbps)	Data Challenge target 2023 (Gbps)	Data Challenge target 2021 (Gbps)
T1						
CA-TRIUMF	200	400	100	60	30	10
DE-KIT	600	1200	300	180	90	30
ES-PIC	200	400	100	60	30	10
FR-CCIN2P3	570	1140	290	170	90	30
IT-INFN-CNAF	690	1380	350	210	100	30
KR-KISTI-GSDC	50	100	30	20	10	0
NDGF	140	280	70	40	20	10
NL-T1	180	360	90	50	30	10
NRC-KI-T1	120	240	60	40	20	10
UK-T1-RAL	610	1220	310	180	90	30
RU-JINR-T1	200	400	100	60	30	10
US-T1-BNL	450	900	230	140	70	20
US-FNAL-CMS	800	1600	400	240	120	40
(atlantic link)	1250	2500	630	380	190	60
Sum	4810	9620	2430	1450	730	240

Data Challenges

Will use the production infrastructure (storage, network, services and tools). Will co-exist and complement the production activity

Will be complemented by network R&D which should enter the challenges when mature (see document for details)

The 2021 challenge provides a first reference of the current capabilities. Will also allow to commission the HTTP 3rd party protocol for Run-3.

The data challenges will be organized in the context of the WLCG DOMA activities

Future work and conclusions

The current planning focuses on the LHC experiments and for 2021 we have an immediate need to focus on the specific aspects in preparation for Run-3. For future challenges we plan to discuss with other sciences e.g. the ones using LHCONE.

A high level view of the asynchronous traffic (Rucio/FTS monitoring) exists. We miss an aggregated low level global view of network utilisation (perfSONAR provides network diagnostic, not network utilisation). An opportunity for LHCONE

The initial planning was discussed with the WLCG sites and regions (in contact with the NRENs). We will need to refine the planning regularly