

The path to the next WLCG Data Challenge

Christoph Wissing, Mario Lassnig

*with content provided by many colleagues
especially many thanks to Alessandra, Ale&Zach, David&David, Simone, Riccardo, Rizart!*

WLCG Workshop 2022

7 November 2022 - 9 November 2022

<https://indico.cern.ch/event/1162261>



Data Challenges: Set up



WLCG objectives

[Planning document](#)

Export of RAW data from CERN to the T1s

Data processing

Incremental steps until HL-LHC

Accompanying R&D programme

2020 estimation

4.8 Tbps of total network capacity

ATLAS & CMS 400 Gbps flat

ALICE & LHCb 100 Gbps flat

x2 to absorb expected bursts

x2 overprovisioning
for operational flexibility

As of now, this has not changed

	%ATLAS	%CMS	% Alice	% LHCb	ATLAS+CMS	Alice	LHCb	LHC Network Needs	LHC Network Needs
					Network Needs (Gbps)	Network Needs (Gbps)	Network Needs (Gbps)	(Gbps)	(Gbps)
T1					Minimal Scenario in 2027	Minimal Scenario in 2027	Minimal Scenario in 2027	Minimal Scenario in 2027	Flexible Scenario in 2027
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

T1	LHC Network Needs	LHC Network Needs	Data Challenge	Data Challenge	Data Challenge	Data Challenge
	Minimal Scenario in 2027 (Gbps)	Flexible Scenario in 2027 (Gbps)	target 2027 (Gbps)	target 2025 (Gbps)	target 2023 (Gbps)	target 2021 (Gbps)
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 Challenge 2021



Executed as two separate challenges

Disk and network

Separate **Run-3** tape commissioning

Objectives

10% of global HL-LHC traffic at flexible/960 Gbps target

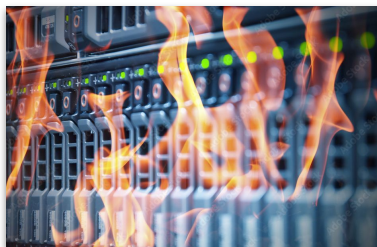
Live system already running in minimum mode

T0-T1 export at minimum/240 Gbps target

Run-3 infrastructure bottleneck discovery

Commissioning of **HTTP-TPC transfer protocol**

Achieve common WLCG **network monitoring**



Implementation

Production infrastructure was used / no separate testbed

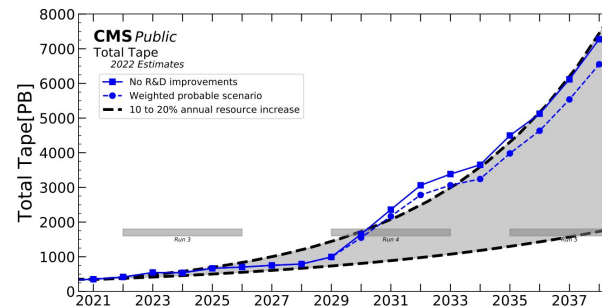
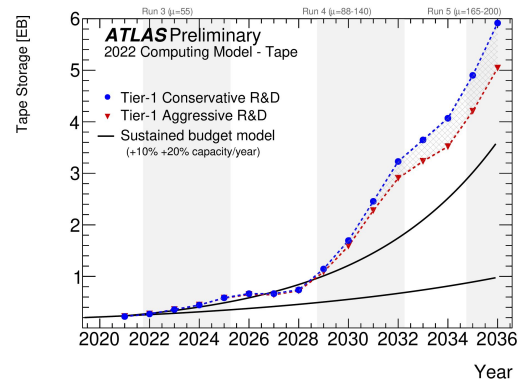
Dedicated system developed to control data challenge rates

Where necessary, additional traffic in **backfill injection** mode for ATLAS and CMS

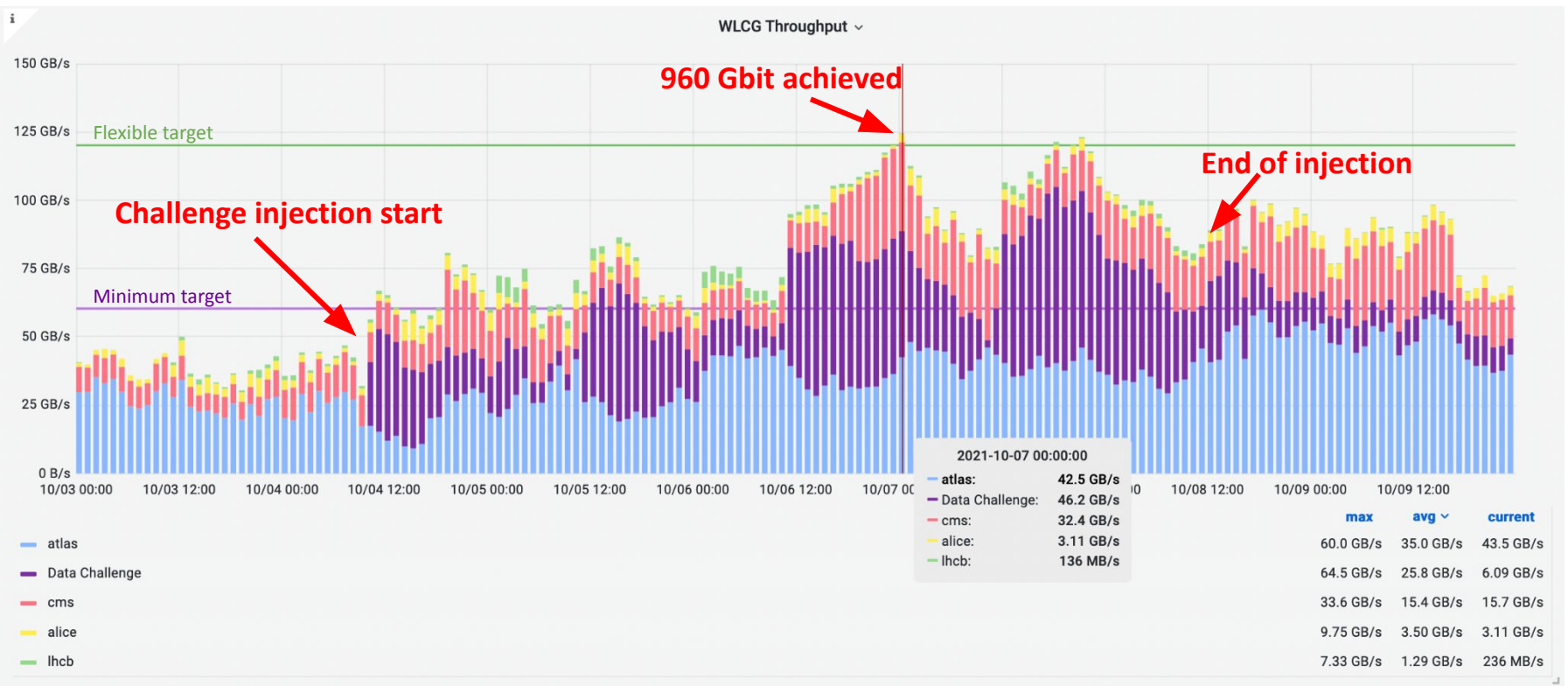
Overlap challenges of all LHC experiments

Use existing **MONIT infrastructure** for monitoring

Ensure inclusion of **XRootD traffic** (ALICE & part of CMS) in evaluation



Data Challenge 2021 measurements



Data Challenge 2021 observations



20 million transfer with burst injection

Short lifetime (1d) for **disk space** reasons

Few configuration changes for storage, Rucio, FTS

Observed transfer **failures unrelated** to challenges

Attempted Transfers	19.8 Mil	Successful Transfers (%)	83.75%	Failed Transfers	3.23 Mil	Failed Transfers (%)	16.25%
Attempted Transfers (vol.)	37.29 PB	Successful Transfers (vol.)	32.12 PB	Failed Transfers (vol.)	5.17 PB	Average Throughput	73.7 GB/s

Tape challenge achieved more continuous flows

NOTED SDN was enabled during challenge

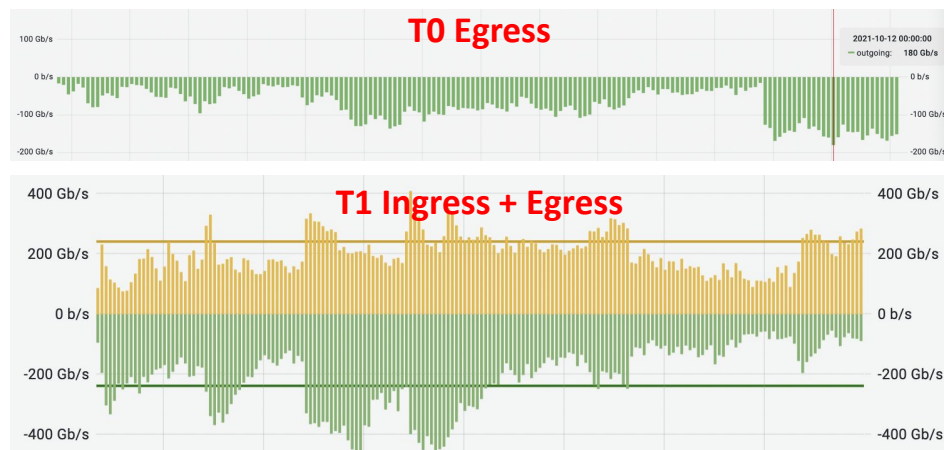
LHCOPN saturation detected

Additional **LHCONE channel** added

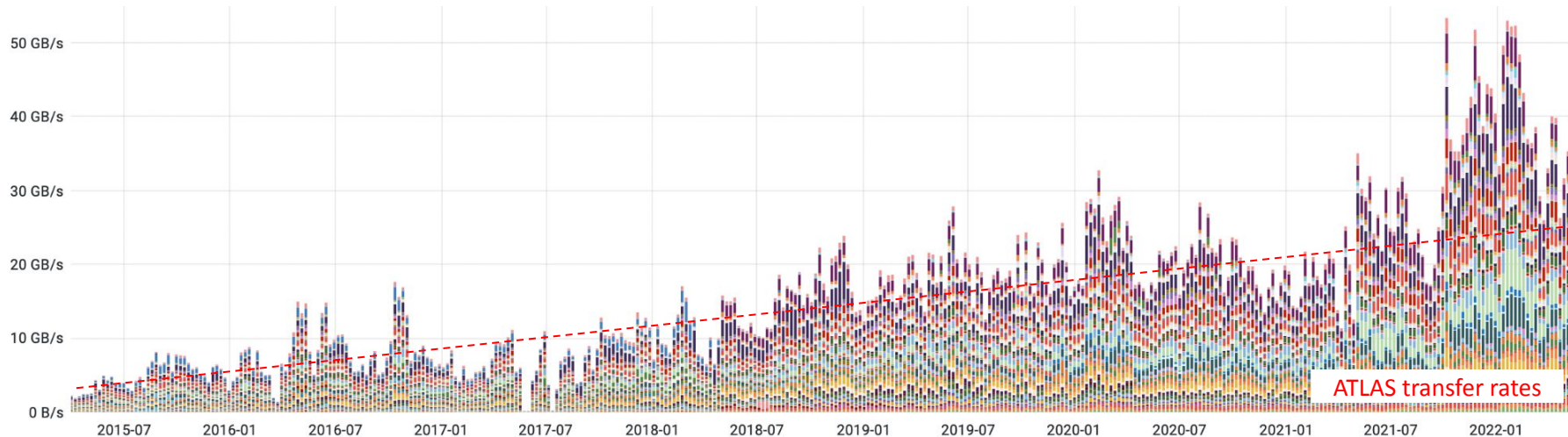
cf. talk by Maria later today

Common monitoring setup very helpful

Detail improvements over the next months



The path to the next challenge



As planned, next challenge would jump from 10% (960 Gbps) to 30% (2880 Gbps) of HL-LHC requirements

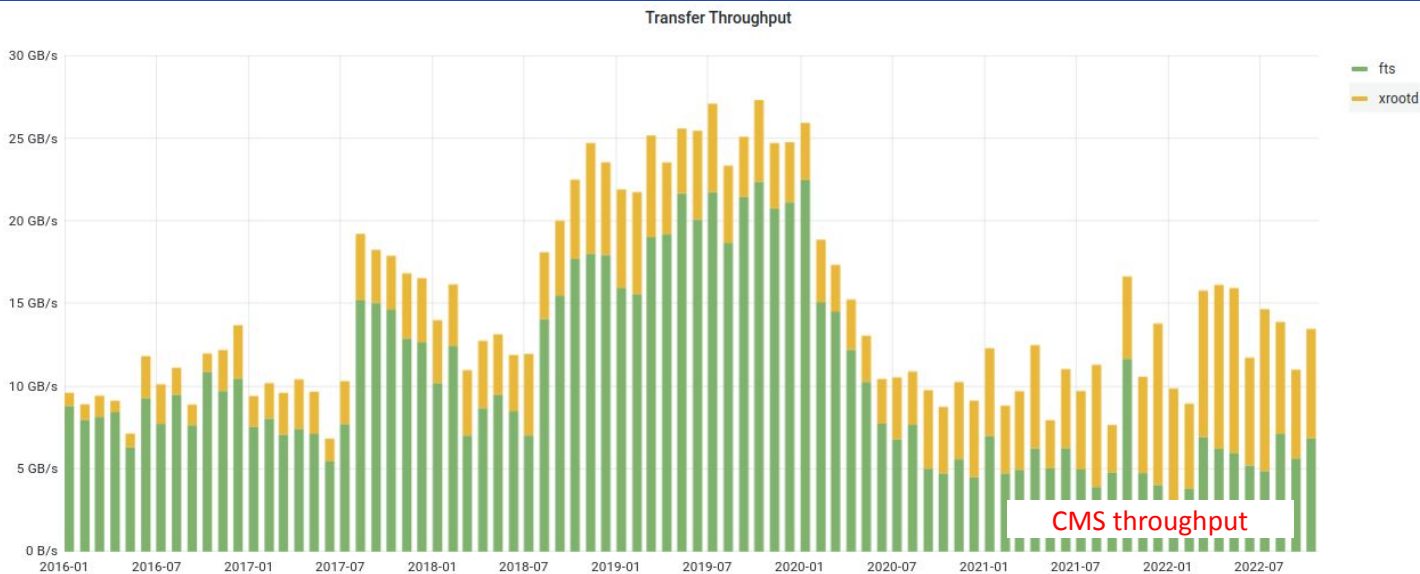
Largest single step increase in the decade-long plan

We have to realistically target the flexible model

Minimum model relies on purely **hierarchical flows**, which works only for the T0-T1 Export case

Experiment dataflows are already **complex**, and new **interactive analysis models** are coming

The path to the next challenge



Various optimisations established at the end of LHC Run 2

Unscheduled XRootD traffic is important for CMS

Monitoring known to be incomplete – improvements are ongoing

Should be also become part of the next challenge

Updated schedule



HL-LHC timeline has been revised

Next "cycle" key WLCG/DOMA topics

- Token-based authn/z
- Software-defined networking
- Flow & packet marking
- TAPE HTTP interaction



Last updated: January 2022

Data Challenge 2021 outcome, recommendations, major points to address

<https://zenodo.org/record/5767913> (detailed 23 pages)

XRootD monitoring integration of standard storage and XCaches

Extend the length of the challenge from few days to 2+ weeks

- Demonstrate sustained traffic capability instead of peak traffic
- Time needed to test different SDN scenarios

Is filling the links even possible?

- Available data centre IO vs. available bandwidth vs. /dev/null ?
- Dedicated Tier-2 test possible?

LHC Experiment Questionnaire



When?

After the processing of the last 2023 **Heavy Ion run** has finished

Before 2024 **pp run** starts

⇒ early-ish 2024

Not during ISGC week (19-24 March)

What?

Specific focus on the test of **SE tokens** for storage, and **migration to IAM**

Monitoring with **IPv6 flow labels**

Demonstrate **SDNs** (*SENSE, AutoGOLE, NOTED, ALTO/TCN*) on selected production sites

Tape challenges were part of the Run-3 export commissioning, necessary to repeat?

Test **peering** with commercial clouds if possible

How?

Start with a series of distributed, constrained, and isolated **ramp-up challenges**

Independently organised, and report via WLCG/DOMA

Hardware purchasing greatly affects data challenge scope, influence on sites non-negligible

Early integration of Tier-1s and Tier-2s in the planning

Instead of a Data Challenge, possibility of **stress tests** instead?

Revisit the original requirements, reduce to 20% or 25% challenge?

Kindle discussion with **non-LHC experiments** for possible future combined Data Challenges

Data Challenge Side Meeting



Discussion with network experts after LHCOPN/LHCONE side meeting

Minutes: <https://docs.google.com/document/d/1zWZMR5U6-nhX1Zo8u-TR0Nt9L3sUiTyQAo9c9Qsits/edit>

Major points

- Clear description of expected bandwidth (average and peak) and access patterns

- Expectation management for Tier-2s (w.r.t. deployed capacity)

- Assess complexity of challenge with more formal metrics and milestones

- Start more formal cooperation with non-LHC experiments

 - At least DUNE and Belle 2 have already expressed that they want to participate

 - Engage with SKA, and potentially ITER, for long-term data challenge alignment (science data ~2028)

- Generalise the idea of isolated testing of new features (esp. SDNs) on production infrastructure

- Official revisit of the bandwidth requirements

 - New HL-LHC schedule: Run4 in 2029 and first PU200 with Run5

 - e.g., <25% challenge would lead to most sites not having to purchase new network hardware