

WLCG DOMA Network Data Challenge 2021

Riccardo DI MARIA

on behalf of the WLCG DOMA Network Data Challenges Working Group

October 2021 - WLCG Operations Coordination Meeting



Disclaimer

- Short report - not comprehensive
- Not conclusive (yet) - full report to follow in DOMA General
- Feedback more than welcome
- [Archive](#) - please comment
- Very-personal feedback (on behalf of anyone)



DC2021

- Goal to demonstrate HL-LHC scale: 10% bandwidth, increasing in subsequent years
 - commission **HTTP Third Party Copy**
- Network target metrics (e.g per site and experiment) available at [HL-LHC document](#)
 - numbers calculated using raw data transfers, i.e. involving tape
- ATLAS and CMS injected data in parallel to production activities
 - usage of “centralised infrastructure” → documentation at [DOMA-TPC Jira](#)
- Alice and LHCb filled the network with normal production activities
 - traffic present for both (centrally monitored for LHCb - but none detected under “Data Challenge” activity label)



Justification for Central Framework

- Common:
 - [transfers machinery](#) and “Data Challenge” activity
 - production/real data and volume
 - monitoring
 - containers to-be-moved centrally created
- ATLAS:
 - production RSEs and nuclei T2s (+1 ex)
 - datasets/files selected by (in collaboration with) DDM
- CMS:
 - “_Test” RSEs and T2s pre-selected by CMS
 - datasets/files selected by transfer team (Data Management/CompOps)



Monitoring

- Leveraging ESCAPE/WLCG existing expertise/dashboards for DC exercises
- FTS-based data source (ElasticSearch) is used for the **common [dashboard](#)**
 - currently including traffic from **ATLAS, CMS & LHCb**
 - special **activity** label for DC activity to filter traffic
 - matrices & plots in place for:
 - files/transfers and throughput
 - data volume
 - transfer error and failures (live links to latest FTS logs)
 - tape related metrics being integrated and filtered accordingly (e.g. read traffic)
- **[Static WLCG Global Throughput](#)** - please do read the NOTE!
- **[Documentation](#)**

Network DC Targets

- [HL-LHC document](#)

	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 (atlantic link)	800	1600	400	240	120	40
	1250	2500	630	380	190	60
Sum	4810	9620	2430	1450	730	240

	%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 (atlantic link)	0	40	0	0	800	0	0	800	1600
					1250	0	0	1250	2500
Sum	100	100	100	100	4000	400	410	4810	9620

- Numbers referring to ingress and egress (disjoint)
- Total ingress+egress
- Minimal: hierarchical model T0-T1-T2 traffic
- Flexible: chaotic model currently more realistic



Current Status of Network Topology for the Link-Tests

ATLAS - 12 T1s and 24 T2s (-1 on Day 1)

- Transfers:
 - from T0 to all T1s
(+dedicated from Day 2)
 - from T1 to all T1s
(full mesh from Day 2)
 - from T0 to all T2s
 - from T1 to all T2s

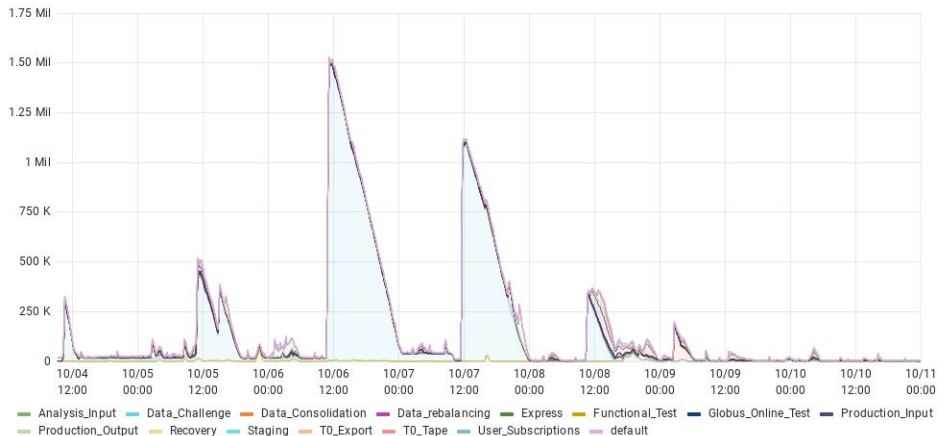
CMS (from Day 3) - 7 T1s and 26 T2s

- Transfers:
 - from T0 to all T1s
 - from T1 to all T1s
 - from T0 to all T2s
 - from T1 to all T2s

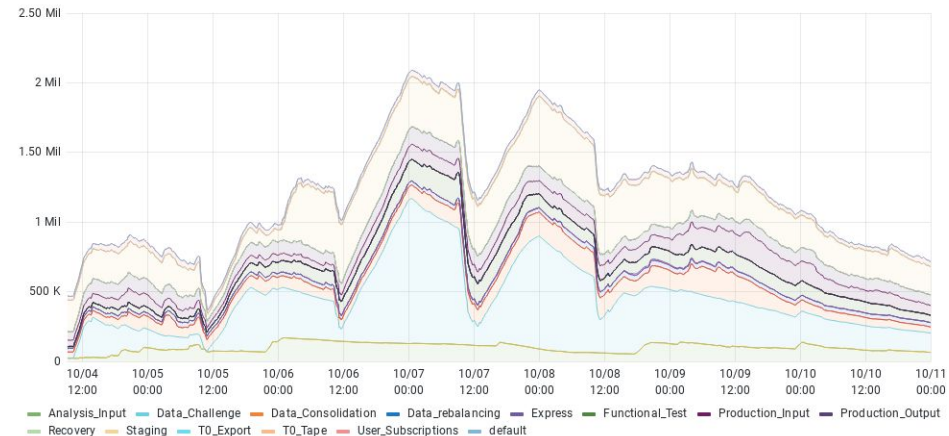
Caveats for Day 1 and 5

ATLAS - DC Week - Conveyor Plots

[Conveyor] Queued Requests by Activity

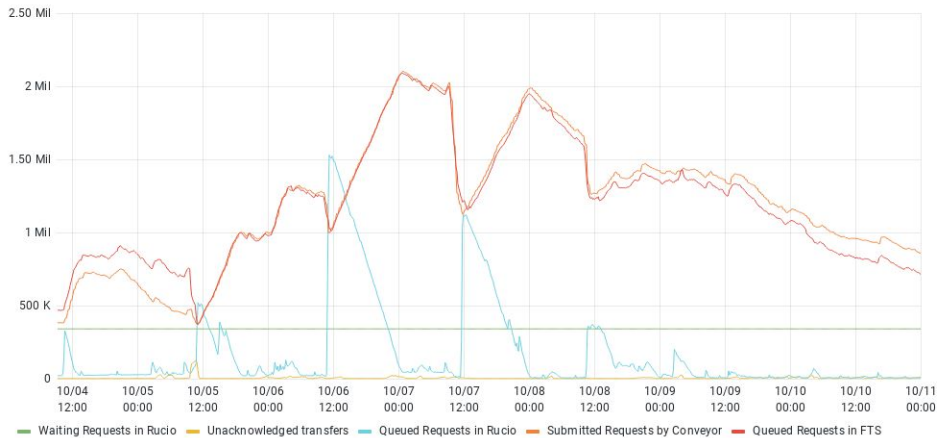


[Conveyor] Submitted Requests by Activity



ATLAS - DC Week - Queues and Deletion Plots

[Conveyor] Queues

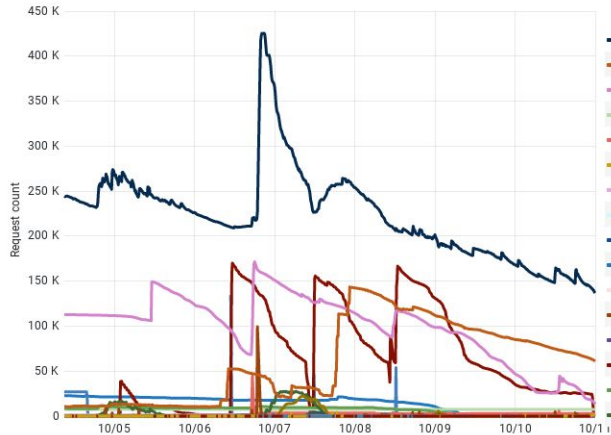


[Reaper] Deletion Rate



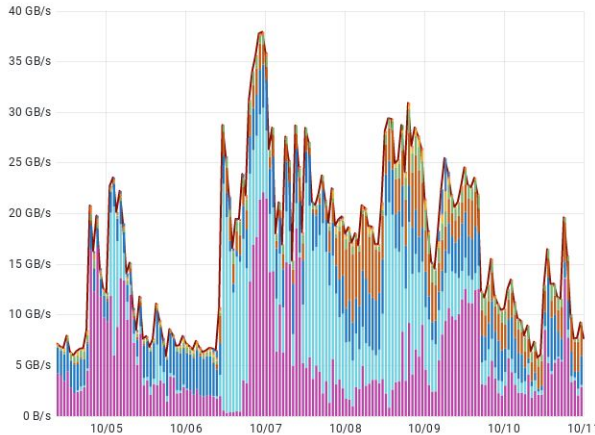
CMS - DC Week - Requests and Transfer Throughput

Requests status



	max	current
Production_Output submitted	425 K	137 K
User_Subscriptions submitted	144 K	61.1 K
Production_Input submitted	171 K	13.1 K
Data_Consolidation queued	7.47 K	7.46 K
Production_Input queued	44.6 K	2.90 K
User_Subscriptions queued	3.99 K	1.45 K
Data_rebalancing queued	1.39 K	1.32 K
Recovery queued	714	605
Functional_Test queued	654	536
Analysis_Input submitted	22.8 K	406
Recovery submitted	468	376
Production_Output queued	99.4 K	289
Functional_Test submitted	1.65 K	185
Data_Challenge submitted	170 K	136
Data_rebalancing submitted	11.9 K	77
Production_Output done	27.5 K	53

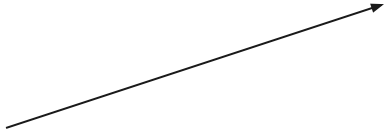
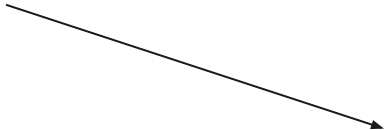
Transfer Throughput



	avg	current	total
Total	16.8 GB/s	7.63 GB/s	2.69 TB/s
Production Output	5.84 GB/s	2.38 GB/s	934 GB/s
Data Challenge	4.52 GB/s	33.9 MB/s	723 GB/s
Production Input	3.23 GB/s	1.21 GB/s	517 GB/s
User Subscriptions	2.04 GB/s	3.81 GB/s	326 GB/s
ASO	771 MB/s	130 MB/s	123 GB/s
Analysis Input	135 MB/s	0 B/s	21.6 GB/s
Data rebalancing	134 MB/s	27.6 MB/s	21.5 GB/s
UNKNOWN	90.4 MB/s	246 kB/s	14.5 GB/s
Functional Test	35.1 MB/s	39.3 MB/s	5.62 GB/s
Recovery	3.54 MB/s	2.57 MB/s	566 MB/s
T0 Export	3.22 MB/s	0 B/s	515 MB/s
T0 Tape	275 kB/s	0 B/s	44.0 MB/s



DC Week - "Data Challenge" Activity

- Transfers:
 - Attempted: 3.49 M
 - **Successful: 84.09% - Failed: 15.91%**
 - Volume:
 - Attempted: 15.14 PB
 - **Successful: 12.81 PB - Failed: 2.33 PB**
 - Average Throughput: 22.4 GB/s
- 
- 

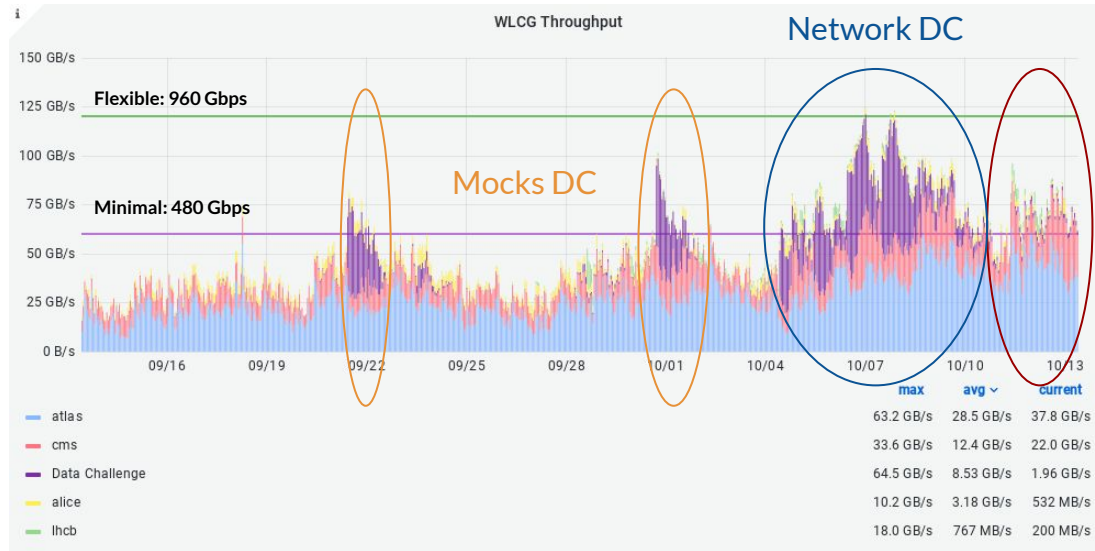
ATLAS

- Transfers:
 - Attempted: 2.88 M
 - **Successful: 86.88% - Failed: 13.12%**
- Volume:
 - Attempted: 11.54 PB
 - **Successful: 10.26 PB - Failed: 1.28 PB**
- Average Throughput: 17.9 GB/s

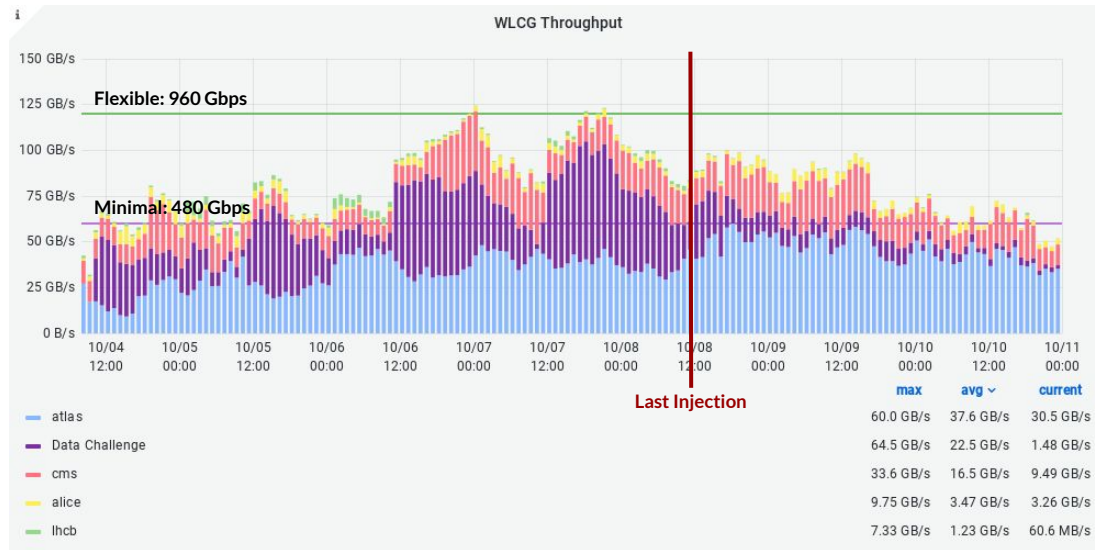
CMS

- Transfers:
 - Attempted: 616 k
 - **Successful: 71.06% - Failed: 28.94%**
- Volume:
 - Attempted: 3.60 PB
 - **Successful: 2.55 PB - Failed: 1.05 PB**
- Average Throughput: 4.95 GB/s

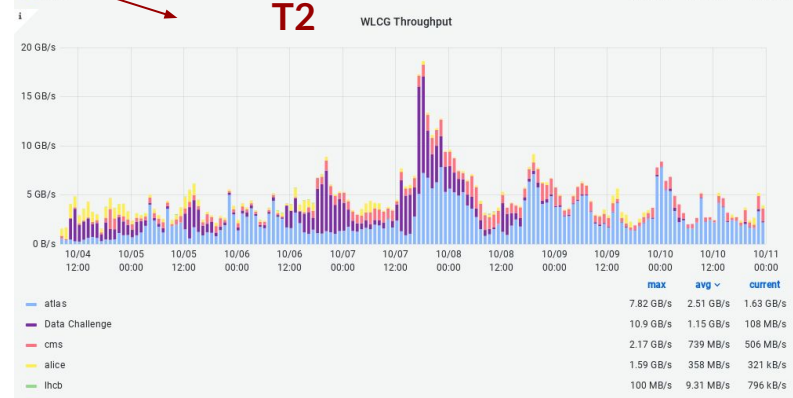
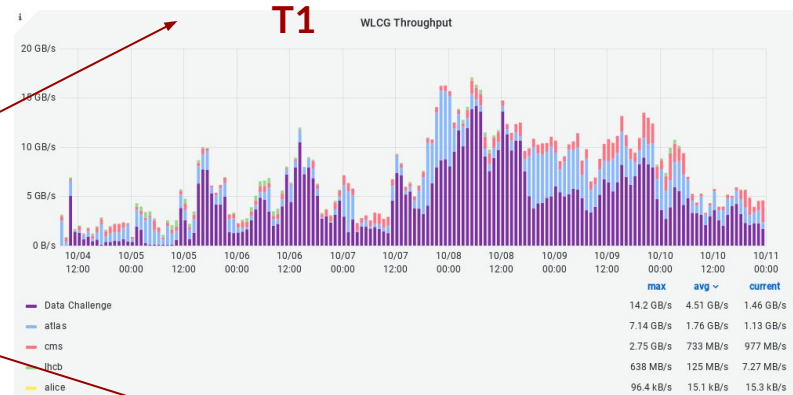
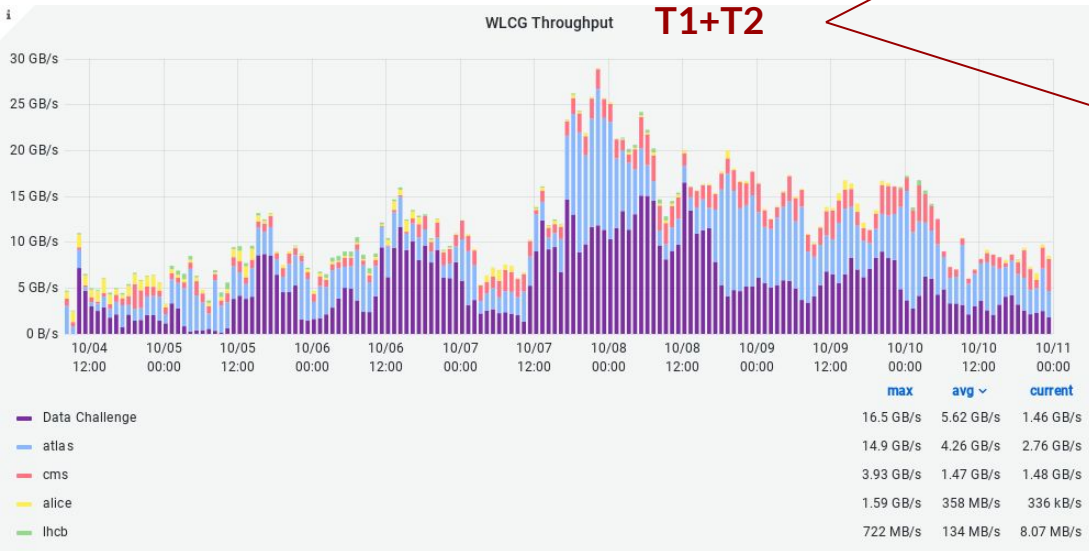
WLCG (FTS+XRootD) Throughput - 30-Day Plot



WLCG (FTS+XRootD) Throughput - DC Week Plot



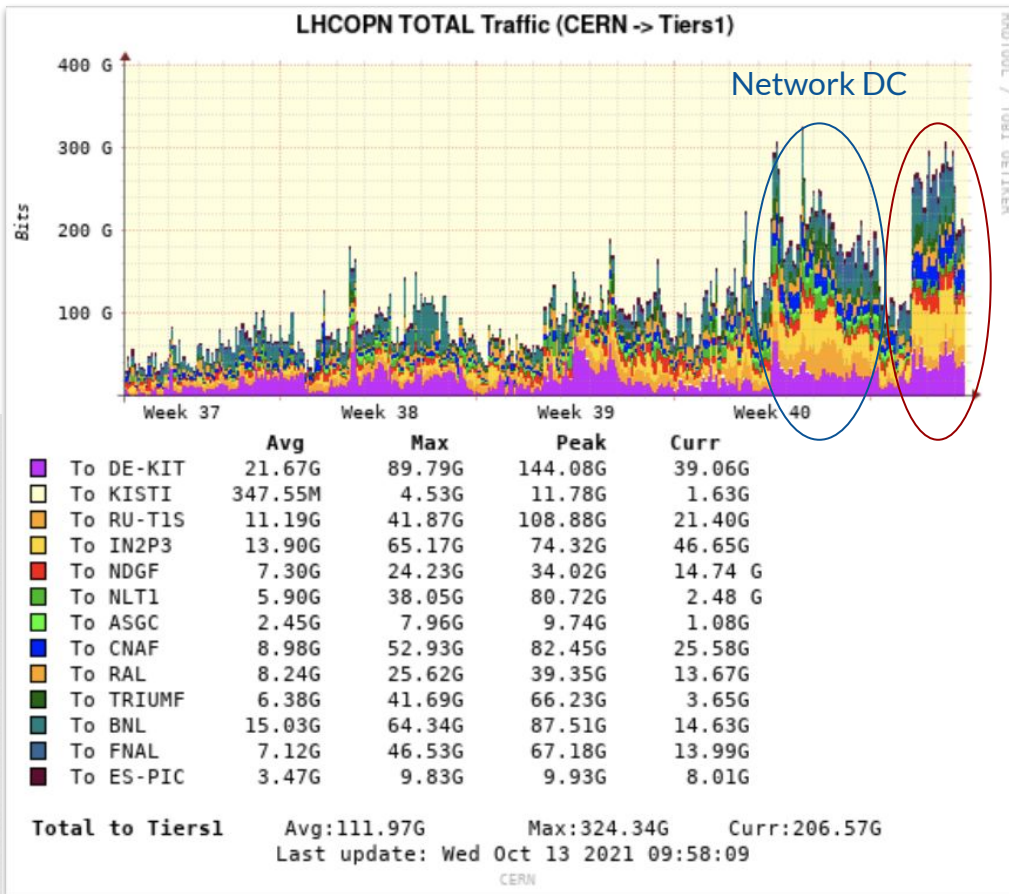
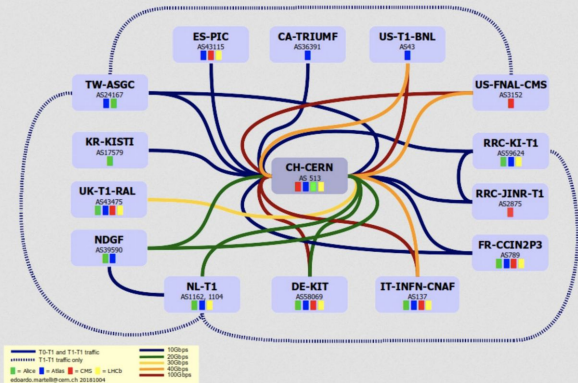
WLCG FTS+XRootD - T0→T1+T2



LHCOPN

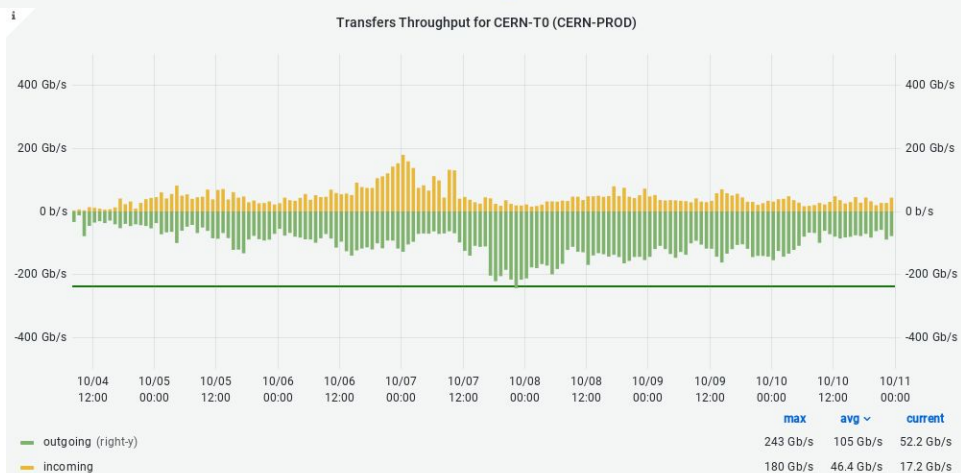
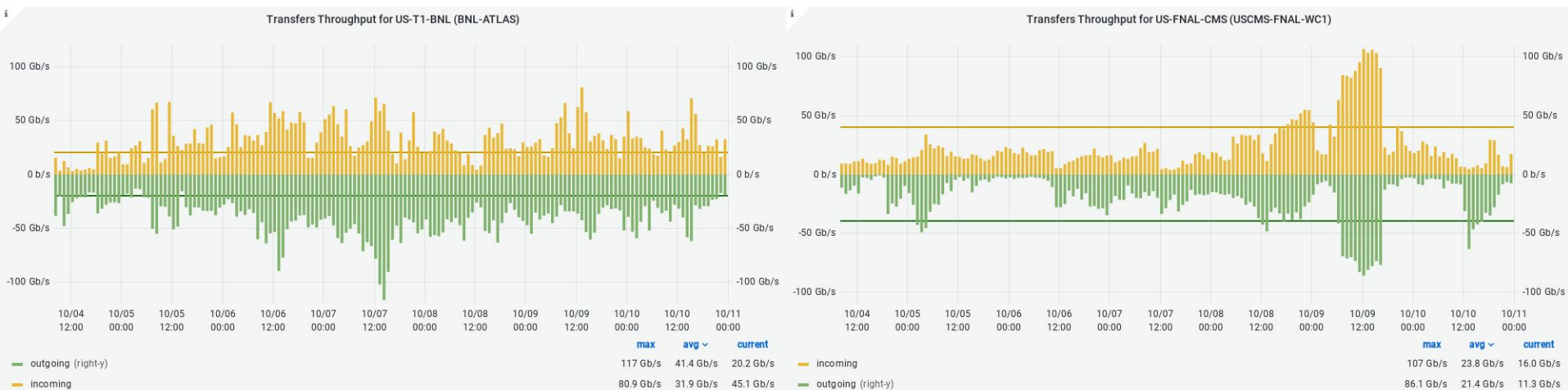
T0-T1 FTS traffic
vs
LHCOPN T1-T0-T1

LHCOPN

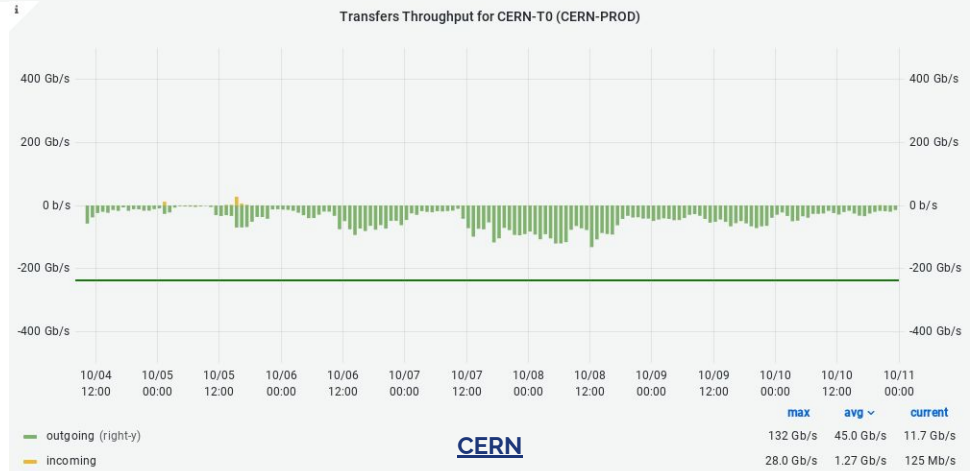
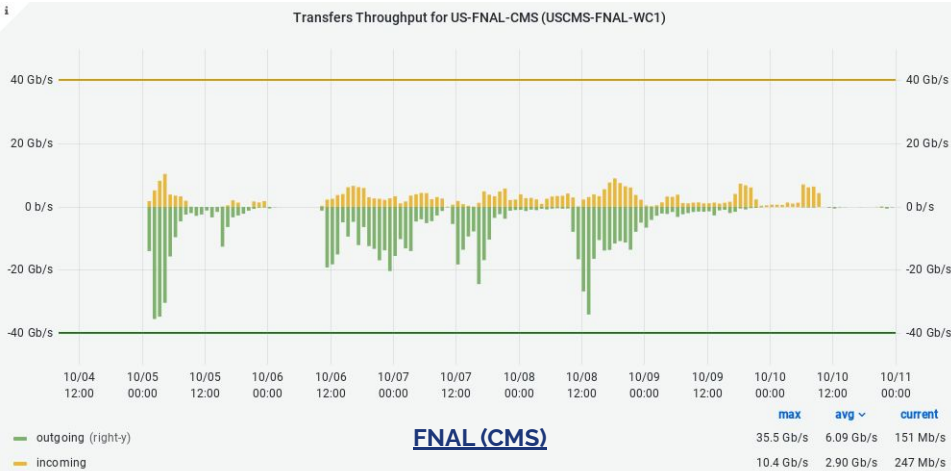
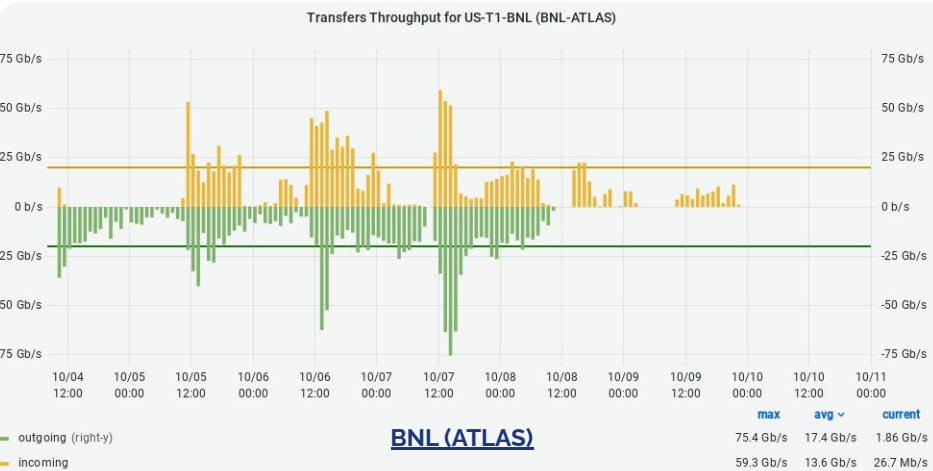


Tape Challenge

FTS Global Traffic during DC Week

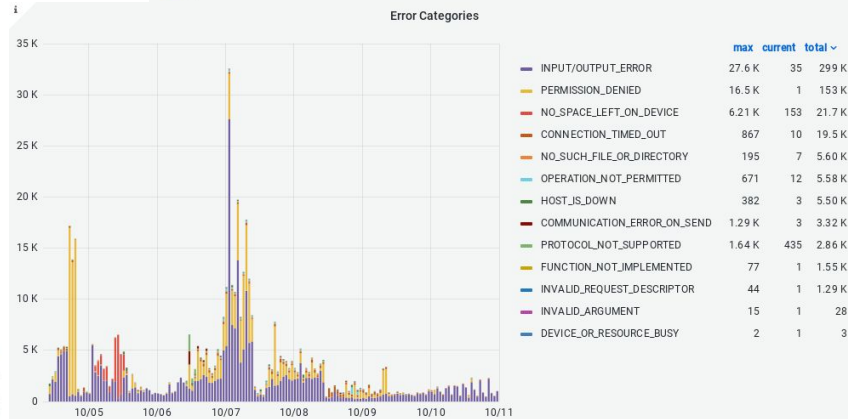
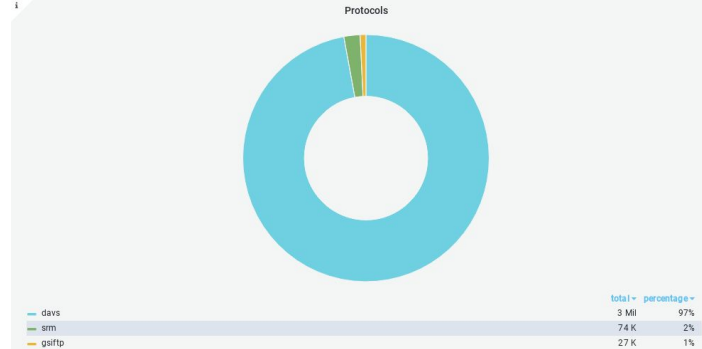
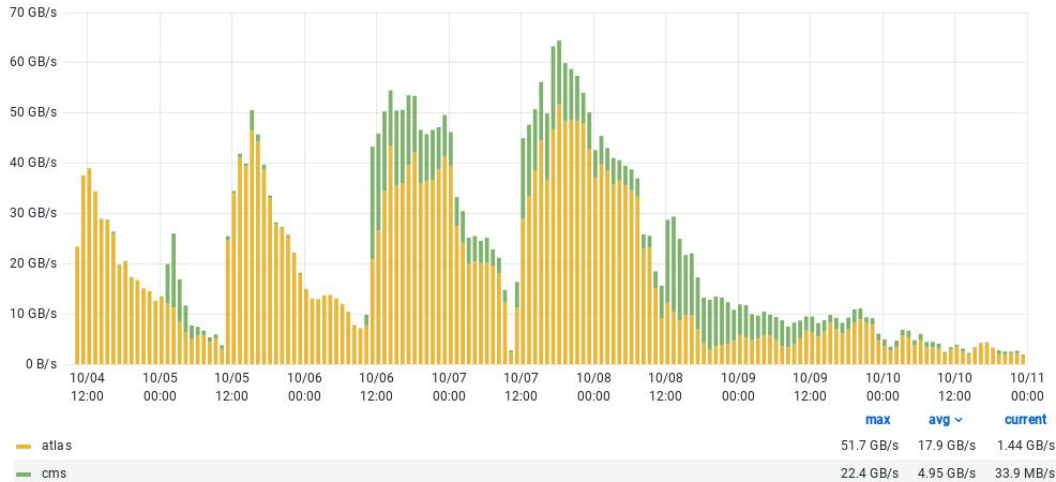


FTS DC-Activity-only Traffic during DC Week



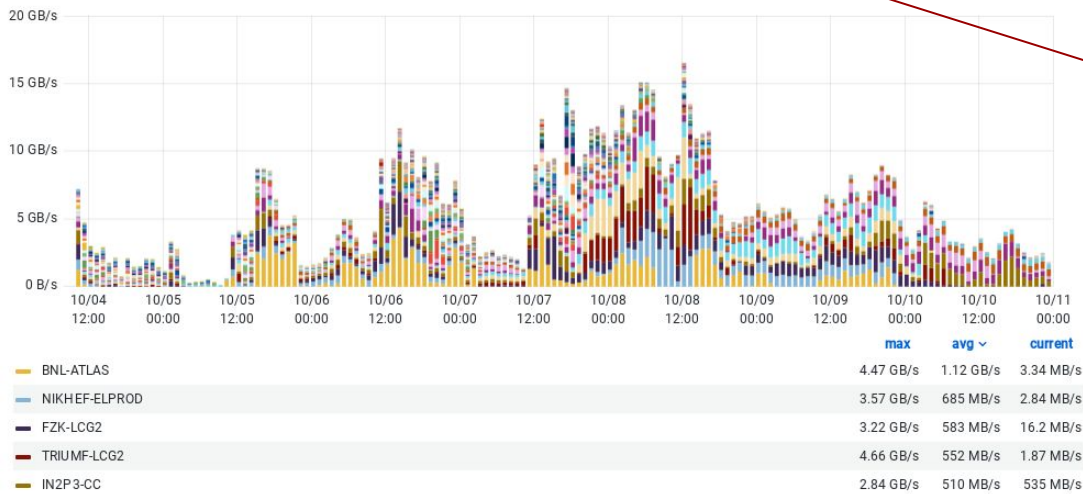
Transfers Throughput DC-Activity-only - DC Week Plot

Transfers Throughput

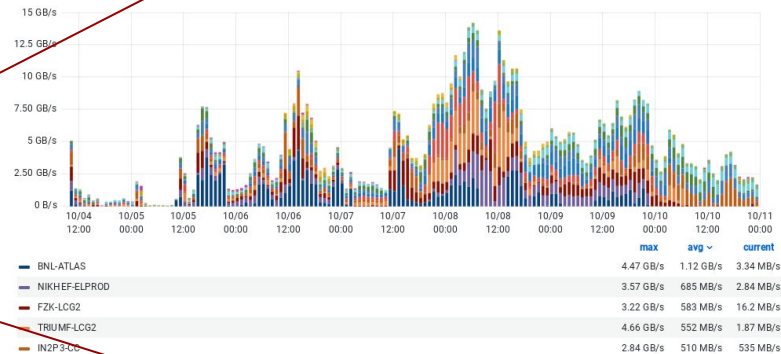


DC-Activity-only - T0→T1+T2

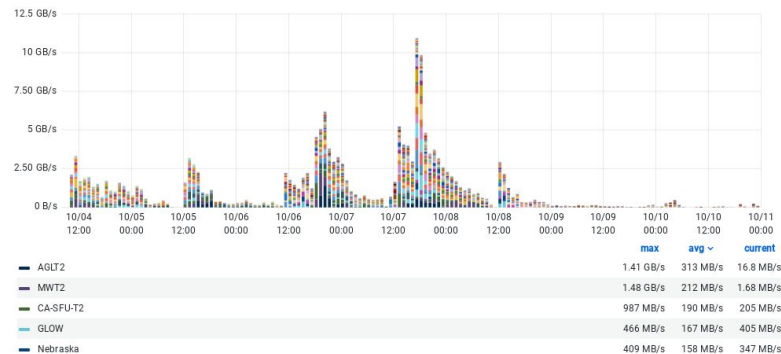
Transfers Throughput **T1+T2**



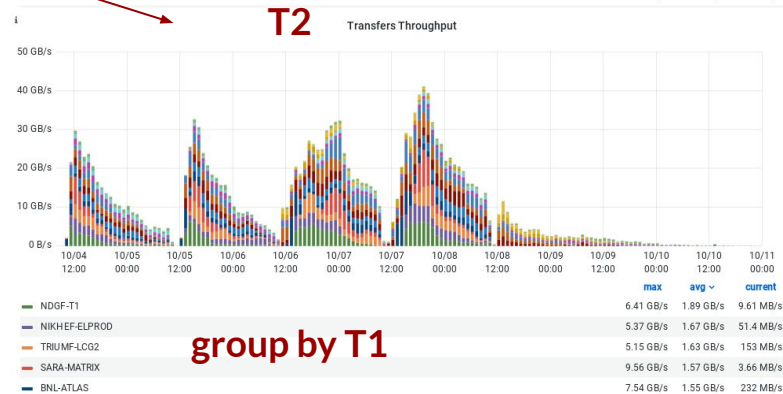
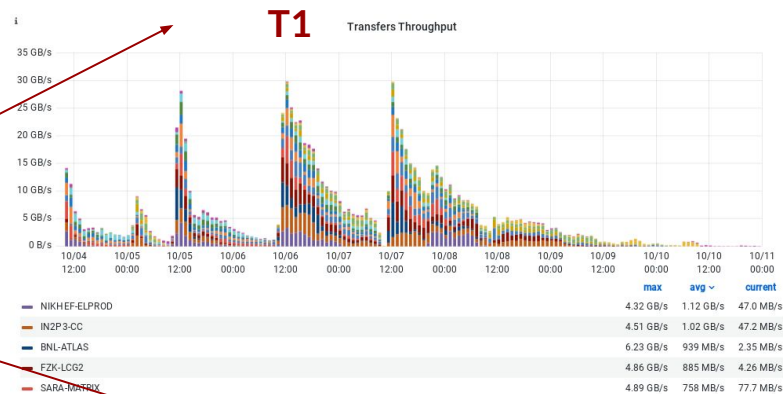
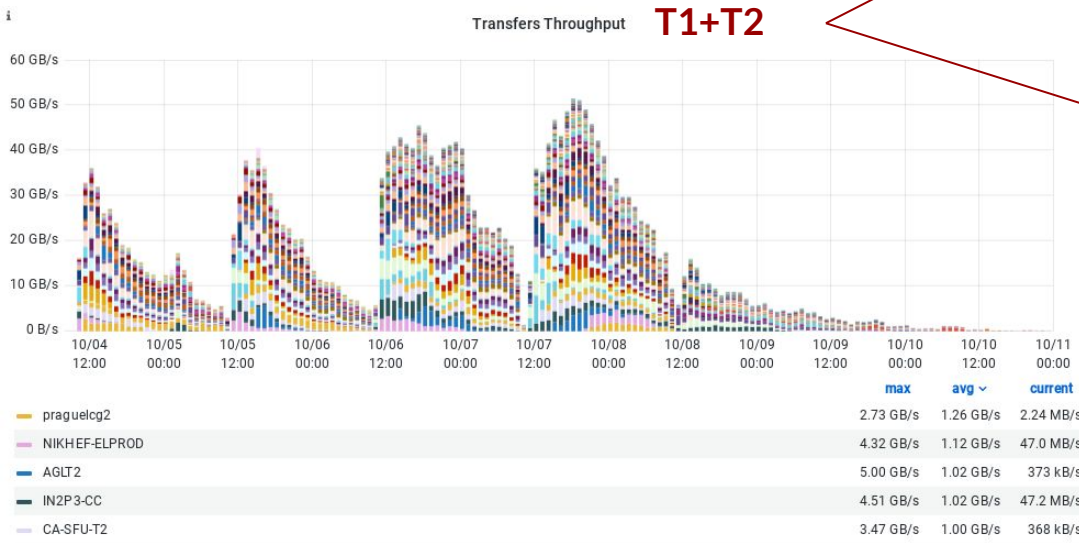
T1 Transfers Throughput



T2 Transfers Throughput



DC-Activity-only - T1->T1+T2





Operational Aspects

- Using real data for these challenges will become trickier towards 2027
 - Load generator or cleanup campaign should help
- Selection of data requires knowledge of Rucio and DB
 - Centralised - dedicated personpower
- Test infrastructures should be avoided at all cost
- Volume is strictly linked to number of tests - T0/T1s affected by a-priori decisions
- Still a non negligible amount of differences wrt Rucio implementations
 - Rucio upstream should be the must
- Monitoring = gazillion of dashboards, datasources, code, people = no-go
 - Common/centralised + XRootD fixed once and for all → DC proved feasibility!



Conclusions

- Data Challenge 2021 was a success overall
 - Further studies should be carried out
 - Dedicated tests might be needed as follow-up
 - Discussions are a must
-
- Contact: doma-data-challenges-development@cern.ch