

**Centre de Calcul**  
de l'Institut National de Physique Nucléaire  
et de Physique des Particules

# Data Challenge 2024 : DC24

- **Les motivations**
- **Un peu d'histoire**
- **De quoi on parle ?**
  - Les datas
  - Les infrastructures
  - Les services
  - Les collaborations
- **Les objectifs des DC24**
- **Les moyens mis en œuvre**
- **Le déroulé et résultats**
- **Le point de vue des expériences**
- **Conclusion**

- **Motivation initiale**

- Challenger le fait de pouvoir faire la gestion des données du HL\_LHC (High Luminosity LHC)/RUN4.

- **Motivation initiale bis**

- Challenger le fait de pouvoir faire la gestion des données du HL\_LHC et d'autres expériences ( Belle2, DUNE,...) qui utiliseront les mêmes infrastructures.

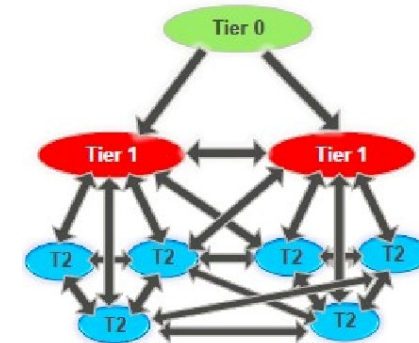
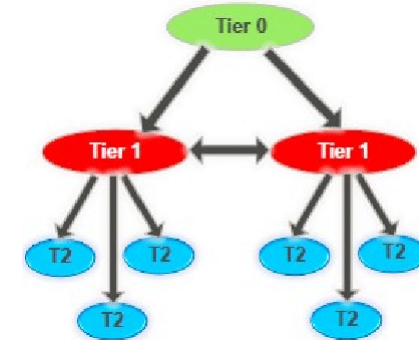


LHC Planning

Last update: April 2023

- **Deux modèles de distribution des données**

- **Modèle hiérarchique ( Monarch )**
  - Basé sur la topologie.
  - Consommateur de stockage.
  
- **Modèle flexible**
  - Modèle assez chaotique.
  - Consommateur de réseau.
  - Nécessite un data management performant.



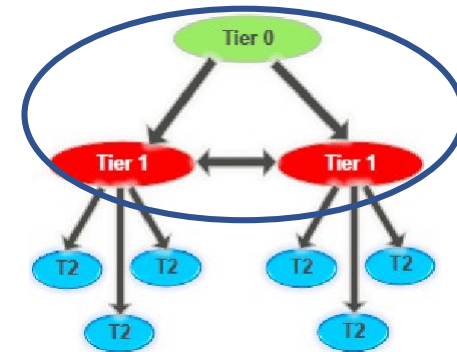
Les data qui sont manipulées par les collaborations peuvent grossièrement être classées en trois catégories.

- **Les données primaires (RAW data)**
  - Sont produites au Tier 0 et vont dans les Tier 1.
  - Sont extrêmement importantes et précieuses mais assez peu utilisées.
- **Les données simulées (Monté Carlo)**
  - Sont produites dans Tier 0, Tier 1 et Tier 2.
  - Ont vocation à être utilisées un peu de partout.
- **Les données d'analyses ( AOD, nano AOD, ESD,... )**
  - Sont produites aux Tier 0 et Tier 1.
  - Sont assez précieuses et ont vocation à être utilisées partout .

## Deux réseaux privés ont été déployés par les NREN nationaux

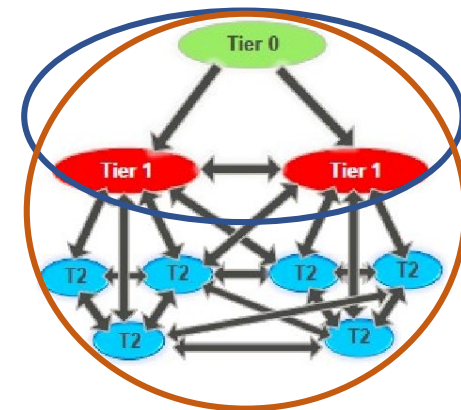
### LHCOPN (Large Hadron Collider Optical Private Network) ●

- Réservé aux trafics WLCG.
- Transferts Tier0-Tier1 et Tier1-Tier1.

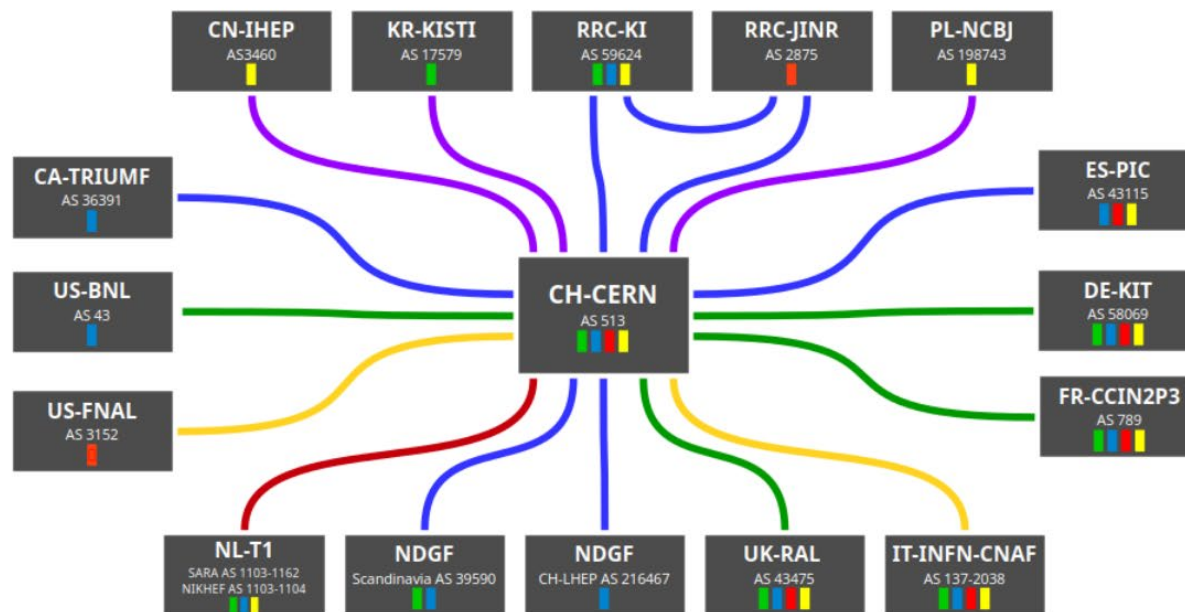


### LHCONE (Large Hadron Collider Open Network Environment) ●

- Ouvert à : LHC, Belle2, Juno, Xenon, Pierre Auger observatory, DUNE.
- Une grande partie (mais pas tous) des Tier2/Tier1 sont connectés dessus.
- Transferts Tier1-Tier1, Tier1-Tier2, Tier2-Tier2.



## LHCOPN



Line speeds:  
20Gbps  
100Gbps  
200Gbps  
400Gbps  
800Gbps

Experiments:  
Alice = Atlas  
CMS = LHCb  
Last update:  
20240209  
edoardo.martelli@cern.ch

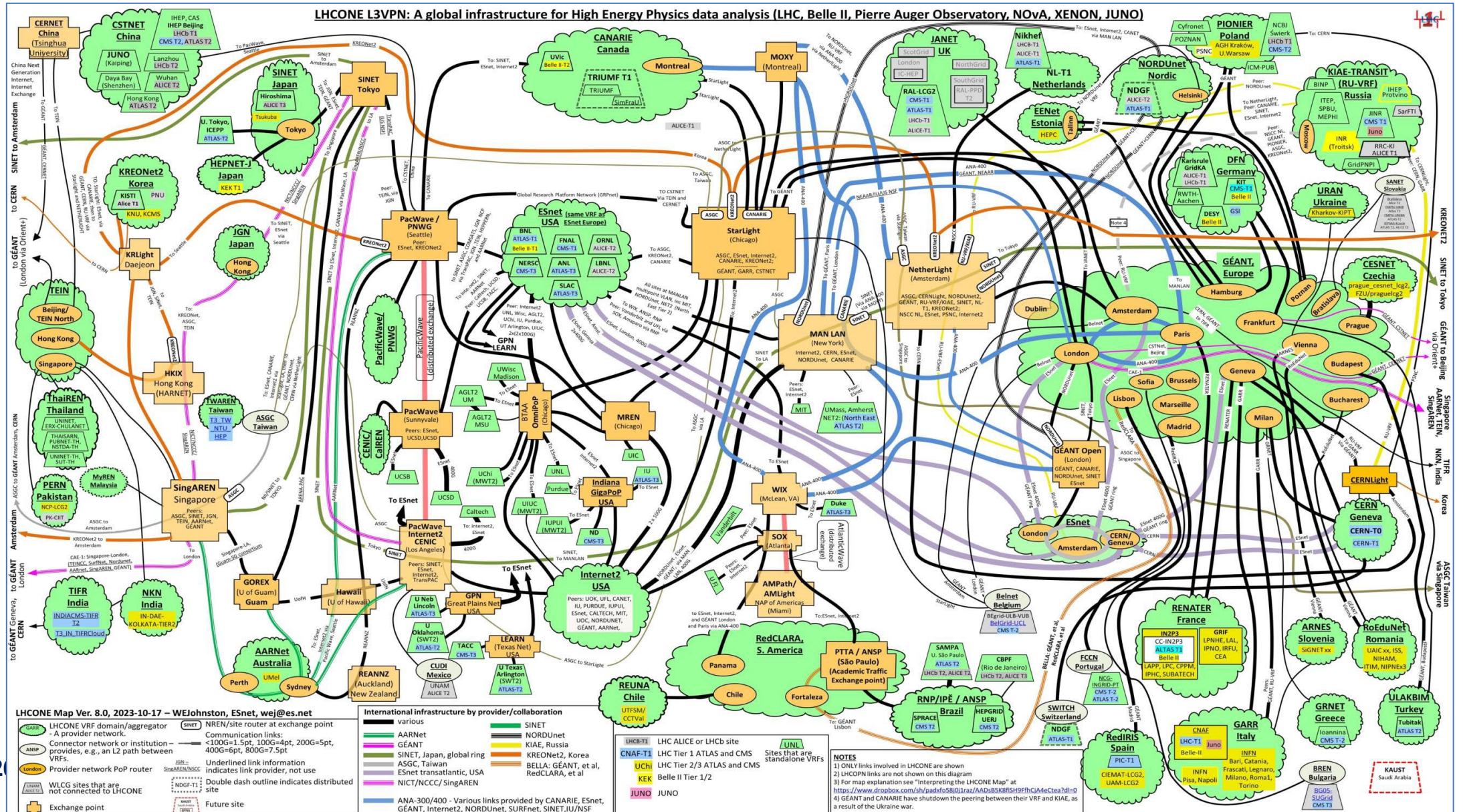
### Numbers

- 17 sites for 15 Tier1s + 1 Tier0
- From 20Gbps to 400Gbps
- 14 countries in 3 continents
- 2.66 Tbps to the Tier0
- CN-IHEP and NDFG-LHEP last connected
- TW-ASGC has left

### Numbers:

- Moved ~619 PB in the last 12 months
- +27% compared to previous year (488PB)
- Peak at ~800Gbps (during DC24)

# De quoi on parle : d'infrastructure réseau





## Stockage sur bande

- Uniquement le Tier 0 et les Tiers 1.
- Des backends divers : HPSS, CTA,...
- Des architectures/configurations diverses.

## Stockage sur disque

- Divers middleware de stockage de grilles sont utilisés par les sites.
- Dcache, EOS, Storm, XRootD, ECHO,...
- Des backends divers : serveur DAS, CEPH, GPFS,...

**Sont utilisés principalement les protocoles : webdav, gridftp, xrootd, srm qui sont accessibles via des endpoint (des urls)**

- <https://ccdavcms.in2p3.fr:2880>
- root://ccxrdatlas.in2p3.fr:1094
- srm://cclhcbtape.in2p3.fr:8443

## Si on veut faire du data management il faut un certains nombre de services

- **Services de transfert**
  - File transfert service (FTS)
    - Copie des données d'un endpoint à un autre en utilisant une ensemble de mécanismes d'optimisation.
    - Ordonnance également ces transferts en fonction de la charge et de la bande passante.
- **Services de meta catalogue applicatif**
  - Rucio ou autre
    - C'est lui qui sait où se trouve les données et il connait leurs caractéristiques.
    - C'est lui qui initie les transferts, les effacements,... en fonction des demandes, des capacités de chacun des sites, des priorités, des contraintes de disponibilité, de redondance,.....
- **Services d'authentification**
  - X509 via les certificats/voms.
  - Par tokens via Indigo IAM.
- **Du monitoring**

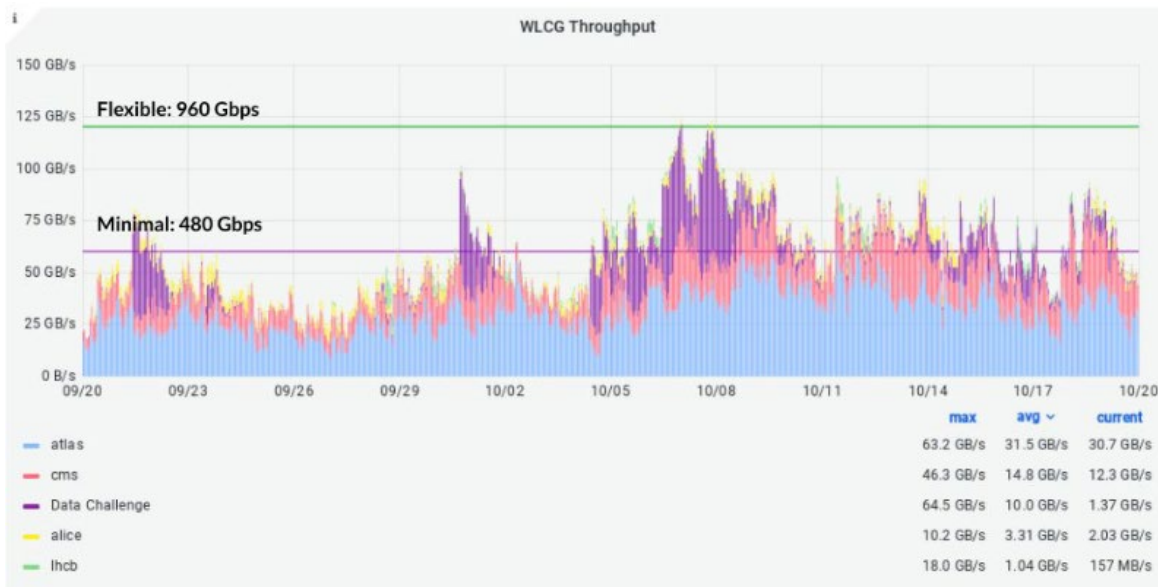
- **Chaque expérience a son data management.**
- **Chaque expérience a son usage de ses données.**
- **Chaque expérience a ses ressources.**
- **Chaque expérience a sa façon de faire.**
- **Chaque expérience est maître d'œuvre dans la gestion de ses données.**

# Donc un data challenge c'est quoi

- **S'assurer (ou plutôt que les expériences s'assurent) qu'il est possible de faire interagir données, infrastructures, services de gestion afin d'atteindre un certain niveau de performance et au passage tester quelques nouvelles fonctionnalités.**
- **Pour les ingénieurs: Identifier les éléments qui sont limitants**
  - Le réseau ?
  - Le stockage ?
  - Les services ?
- **Pour tous : Optimiser les façons de faire et configurations**

# Objectifs du data challenge 2024

- Le DC24 est l'étape 2 dans la série de Data Challenge qui a pour objectifs de valider un objectifs correspondant à 25 % de la cible HL-LHC
  - 1,2 Tb/s et 2,4 Tb/s
- En 2021 on a eu le DC21



DC 21

	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 challenge initial plan

- **3 objectifs principaux pour 2024**
  - Mesurer les capacités de transfert site à site en visant une capacité de 25 % des besoins HL-LHC.
  - Valider de nouveaux services/ technologies.
  - Tester de nouvelles fonctionnalités.

- **Les stockages sur bande ne seront sollicités que de façon marginale.**
  - On teste donc les réseaux, les services et les systèmes de stockage sur disque.
- **Le challenge consiste à rajouter de l'activité sur la production quotidienne habituelle.**
- **Une grille d'objectifs pour chaque site est définie et déclinée par expérience, pour deux scénarios et sur 2 semaines.**
  - 12 Février – 25 Février
- **On en profite pour tester l'authentification par token et quelques fonctionnalités réseau et monitoring.**

T1	HL-LHC Network Needs (Gbps) Minimal Scenario	HL-LHC Network Needs (Gbps) Flexible Scenario	Data Challenge minimal thresholds 2024 in&out	Data Challenge flexible thresholds 2024 in&out	Data Challenge flexible total rates 2024 in+out
CA-TRIUMF	200	400	25	50	100
DE-KIT	600	1200	90	180	360
ES-PIC	200	400	25	50	100
FR-CCIN2P3	570	1140	85	170	340
IT-INFN-CNAF	690	1380	90	180	360
KR-KISTI-GSDC	50	100	10	20	40
NDGF	140	280	20	40	80
NL-T1	180	360	25	50	100
NRC-KI-T1	120	240	0	0	0
PL-NCBJ	20	20	8	10	10
UK-T1-RAL	610	1220	85	170	340
RU-JINR-T1	200	400	0	0	0
US-T1-BNL	450	900	60	120	240
US-FNAL-CMS	800	1600	90	180	360
(atlantic link - only US, CA not included)	1250	2500	150	300	600
<b>Sum</b>	<b>4830</b>	<b>9670</b>	<b>613</b>	<b>1226</b>	<b>2452</b>

# Objectifs du data challenge 2024



T0

CERN-PROD source (Write rates)	RRC-KI	ES-PIC	DE-KIT	FR-CCIN2P3	IT-INFN-CNAF	UK-RAL
ALICE				5	4	7
ATLAS (injected + prod))		0.8	13	38.4	43.5	27.7
CMS		0	19	45	45	57
LHCb		0	4.38	23.54	13.14	17.61
<b>Total</b>		<b>0.8</b>	<b>36.38</b>	<b>111.94</b>	<b>105.64</b>	<b>109.31</b>

Network Capacity[1]	100Gbps	100Gbps	200Gbps (400Gbps)	200Gbps	400Gbps	200Gbps
---------------------	---------	---------	----------------------	---------	---------	---------

DUNE

Belle II (from KEK via LHCONE)		0	0	1.9	2.8	3.7	0
--------------------------------	--	---	---	-----	-----	-----	---

CERN-PROD destination (Read rates)	RRC-KI	ES-PIC	DE-KIT	FR-CCIN2P3	IT-INFN-CNAF	UK-RAL
ALICE			n/a	n/a	n/a	n/a
ATLAS (injected + prod))		0.08	1.64	6.36	6.57	4.19
CMS		0	15	36	36	45
LHCb		0	3.44	14.26	10.31	13.74
<b>Total</b>		<b>0.08</b>	<b>20.08</b>	<b>56.62</b>	<b>52.88</b>	<b>62.93</b>

Network Capacity[1]	100Gbps	100Gbps	200Gbps (400Gbps)	200Gbps	400Gbps	200Gbps
---------------------	---------	---------	----------------------	---------	---------	---------

Ex : Objectifs globaux ( In/out) par expériences pour le scenario flexible dans le cadre des transfert Tier0-Tier1



# Objectifs du data challenge 2024



## Ex : Calendrier des flux testés

	12/02/2024	13/02/2024	14/02/2024	15/02/2024	16/02/2024	17/02/2024	18/02/2024
<b>ALICE</b>	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1
<b>ATLAS</b>	T0 → T1	T0 → T1	T0 → T1 ↔ T1 → T2	T0 → T1 ↔ T1 → T2	T0 → T1 ↔ T1 → T2	T0 → T1 ↔ T1 → T2	T0 → T1 ↔ T1 → T2
<b>CMS</b>	T0 → T1	T0 → T1	T0 → T1 → T2	T1 → T2	T1 ↔ T2	T1 ↔ T2	T1 ↔ T2
<b>LHCb</b>		T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1
<b>DUNE</b>	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2
<b>Belle II</b>	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1

### SUMMARY

T0 exports minimal rates

(ALICE+ATLAS+LHCb+CMS)

529.7 Gbps

650.3 Gbps

650.3 Gbps

650.3 Gbps

650.3 Gbps

650.3 Gbps

650.3 Gbps

T0 exports (DUNE + Belle II)

18.5 Gbps (belleII)

18.5 Gbps (belleII)

18.5 Gbps (belleII)

18.5 Gbps (belleII)

18.5 Gbps (belleII)

18.5 Gbps (belleII)

18.5 Gbps (belleII)

	Monday 19/02/2024	Tuesday 20/02/2024	Wednesday 21/02/2024	Thursday 22/02/2024	Friday 23/02/2024	
<b>ALICE</b>	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	yellow: "reduced minimal" (only T0 export)
<b>ATLAS</b>	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	blue: minimal scenario
<b>CMS</b>	AAA T1 → T2	T0 → T1 ↔ T2	T0 → T1 ↔ T2	T0 → T1 ↔ T2	T0 → T1 ↔ T2	red: flexible scenario
<b>LHCb</b>	T0 → T1	T1 Tape Recall	T1 Tape Recall	T1 Tape Recall	T1 Tape Recall	green: cms minimal T1->T2
<b>DUNE</b>	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	
<b>Belle II</b>	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 == SURF , T1 == FNAL, T2 == Storage sites

- **Un outil d'injection qui demande au Meta Catalogue (rucio) de rajouter de l'activité de transfert selon plusieurs paramètres : Sites cibles, taille des données, charge max sur les liens,.....**
- **Un monitoring global qui donne une vue d'ensemble et par expériences du niveau de performance dans les transferts.**
- **Des personnes sur site ( réseau et du stockage) qui surveillent/monitorent et éventuellement agissent.**
- **Des personnes des expériences qui injectent les data et surveillent/monitorent .**
- **Des experts dans les services ( RUCIO,FTS,IAM,...) qui surveillent/monitorent et éventuellement agissent.**
- **Un canal d'échange pour tout ce beau monde.**
- **Un réunion quotidienne pour débriefer et organiser le jour d'après.**

## Au niveau du CC

- **Réseau**

- Renater a déployé ( in extremis) un upgrade de nos liens LHCOPN et LHCONE qui sont donc passés de 100Gb/s à 200 Gb/s chacun.
- Les objectifs assignés au CC nécessitaient cet upgrade.

- **Stockage**

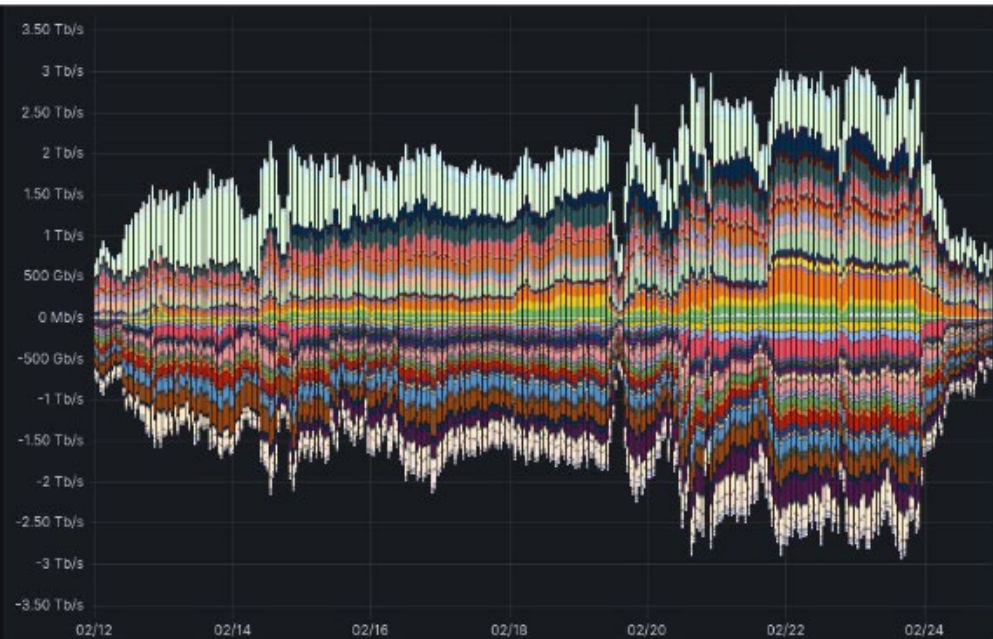
- Les instances de dcache (ATLAS,CMS,LHCb et Belle2) et XrootD (Alice) ont été upgradées afin de mettre en œuvre certaines nouvelles fonctionnalités requises pour le DC24.

- **Monitoring**

- Au fils des besoins certains monitoring ont pu être rajoutés.

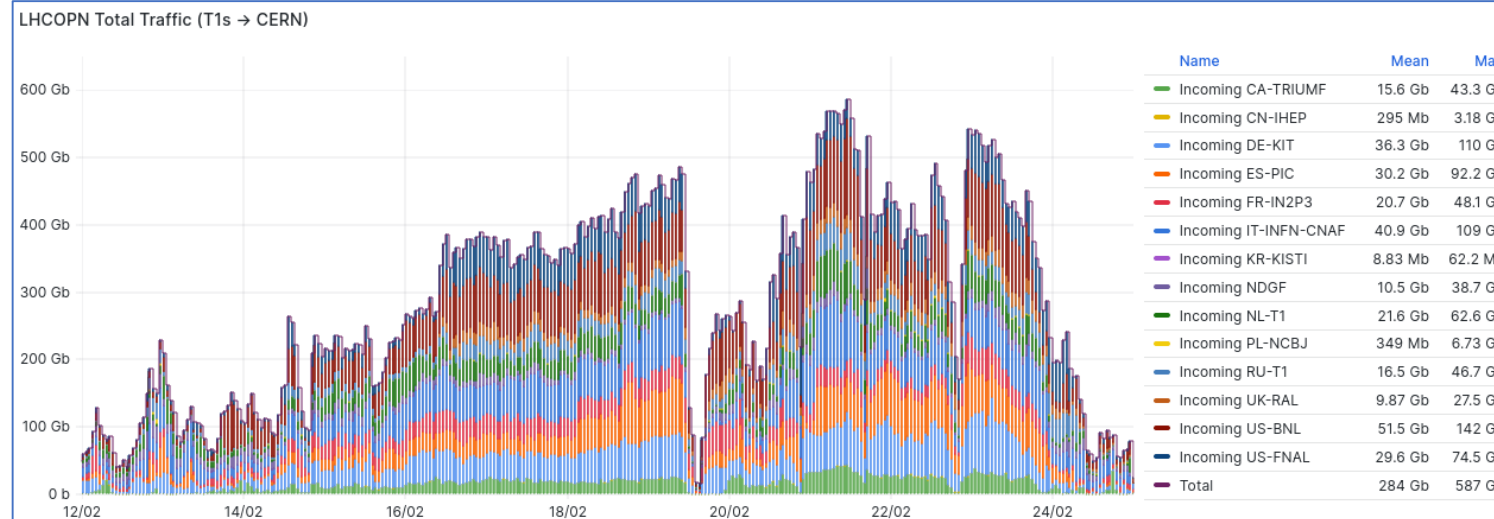
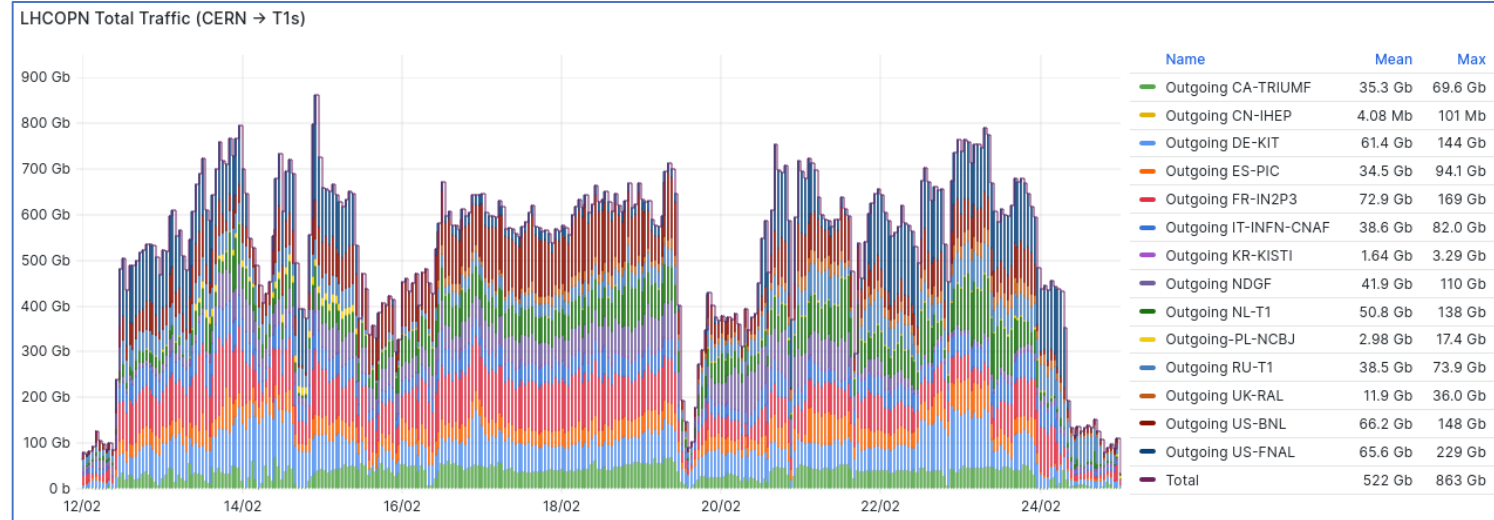
# Résultat : Vue globale réseau

DC24 : LHCOPN CERN –Tier1

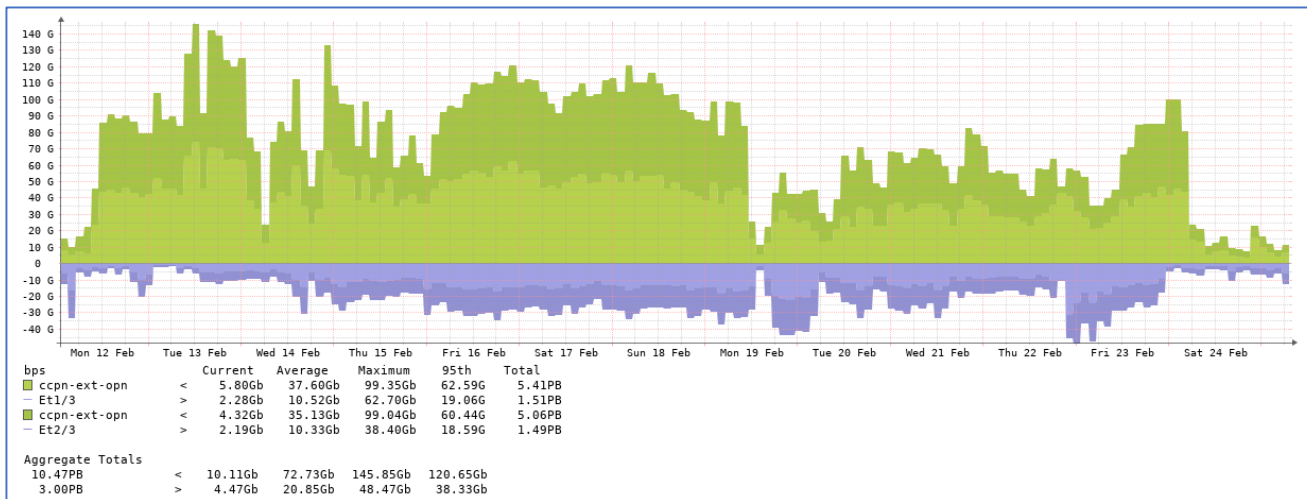
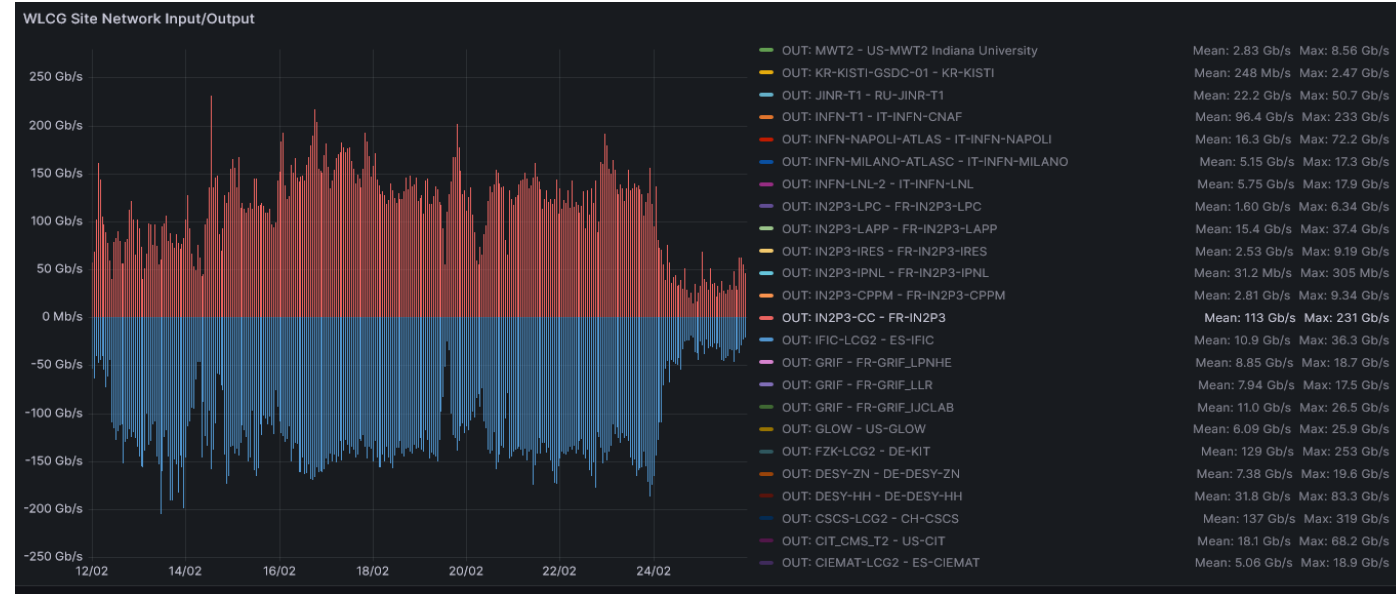


DC24 : ALL Network, All site IN/out

DC24 : LHCOPN T1 –CERN



## DC24 : CC in/out

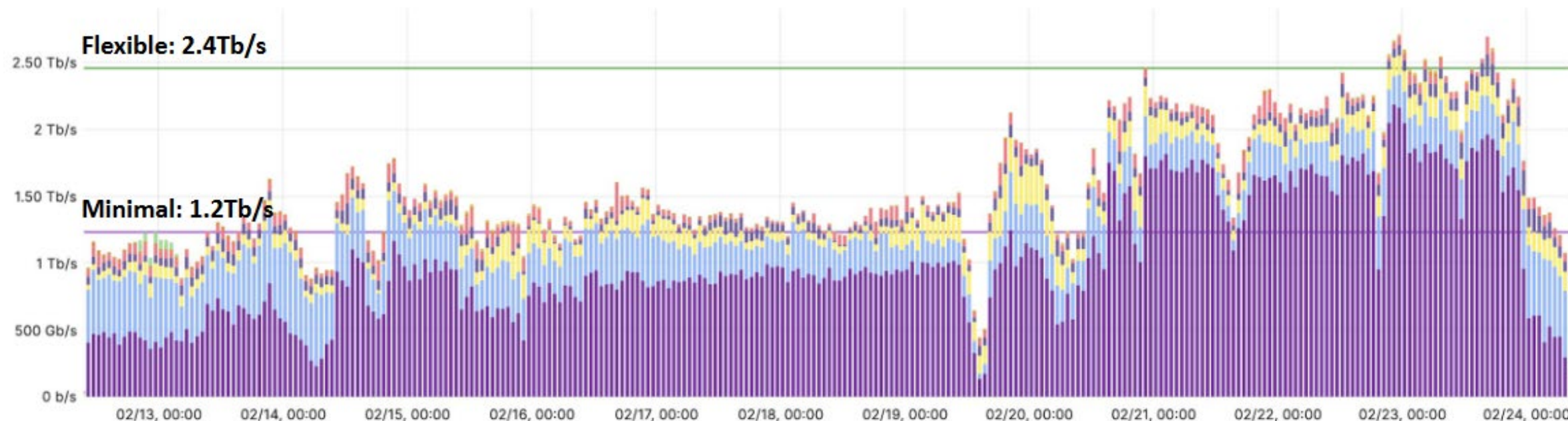


## DC24 : CC LHCOPN in/out

# Résultats : Vue globale des transferts

WLCG Throughput ⓘ

DC24 : Transferts en une vue



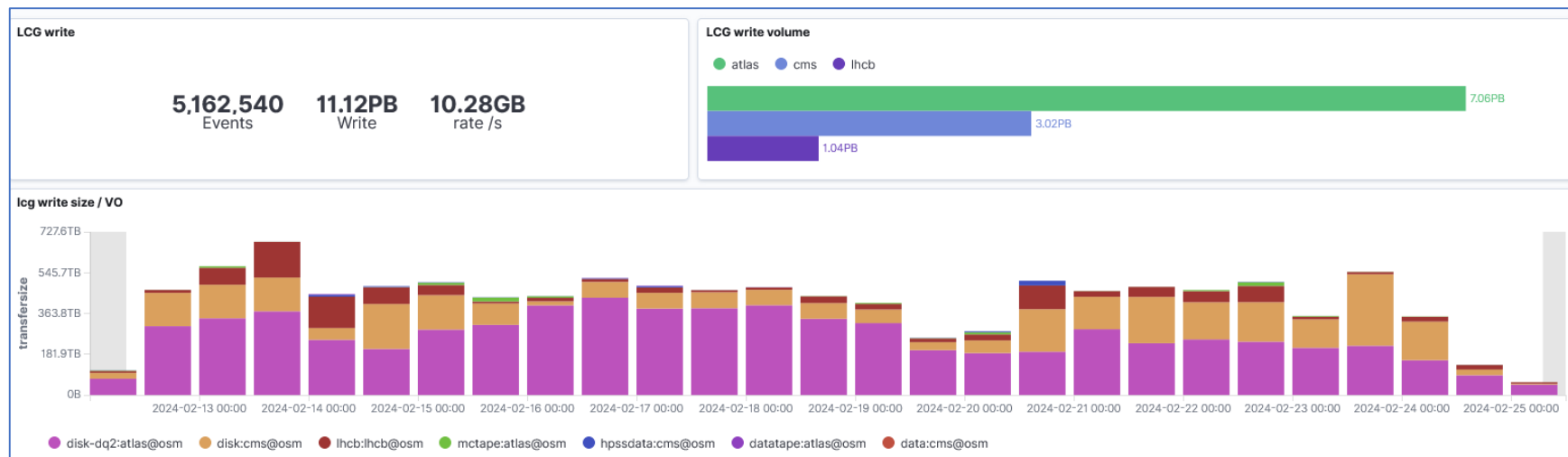
- Data Challenge
- atlas
- alice xrootd
- cms xrootd
- cms
- belle

DC24 met the (main) goals:

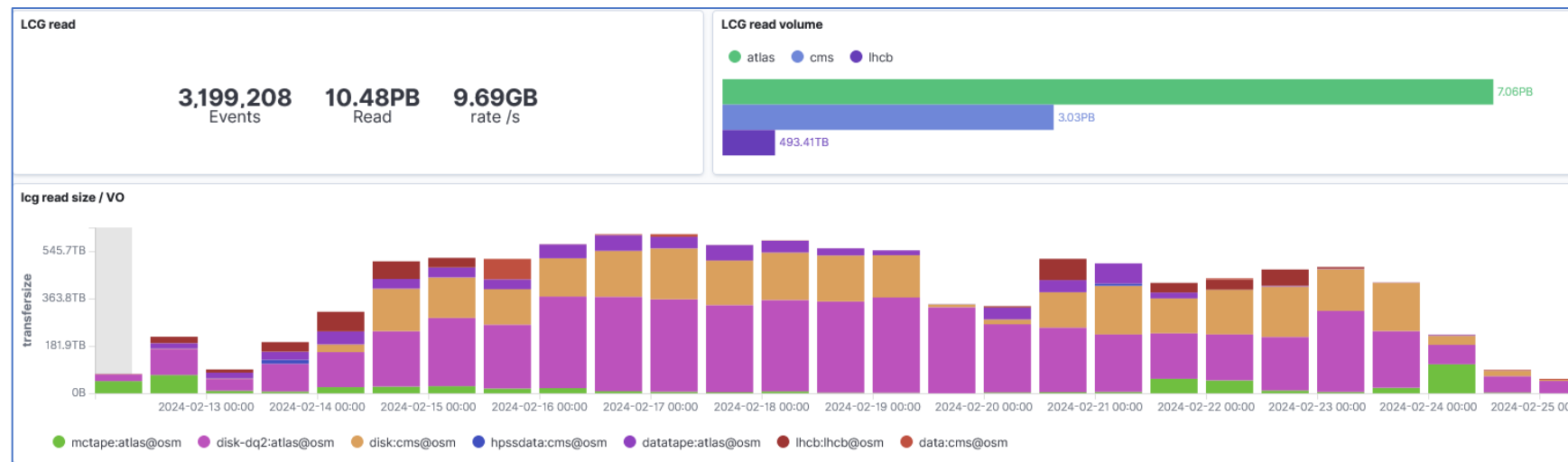
- Achieved full throughput of minimal model (1st week)
- Push for flexible target (2nd week)

	max	avg	current
	2.19 Tb/s	1.02 Tb/s	211 Gb/s
	625 Gb/s	304 Gb/s	567 Gb/s
	349 Gb/s	115 Gb/s	71.4 Gb/s
	191 Gb/s	67.4 Gb/s	42.7 Gb/s
	271 Gb/s	57.2 Gb/s	75.0 Gb/s
	38.9 Gb/s	9.45 Gb/s	17.1 Gb/s

## DC24 : Dcache LCG CC Write



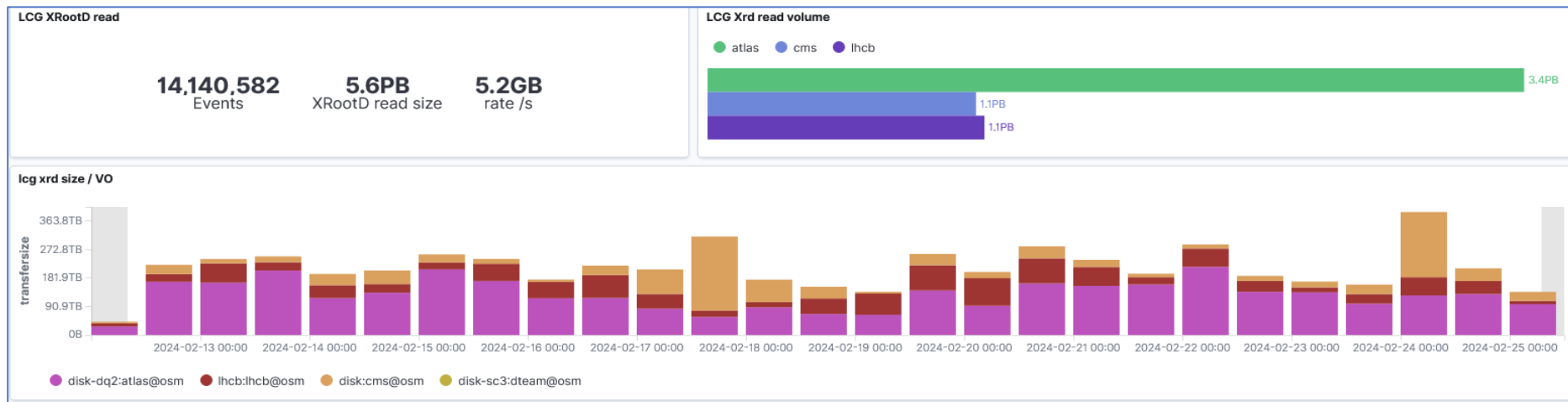
## DC24 : Dcache LCG CC Read



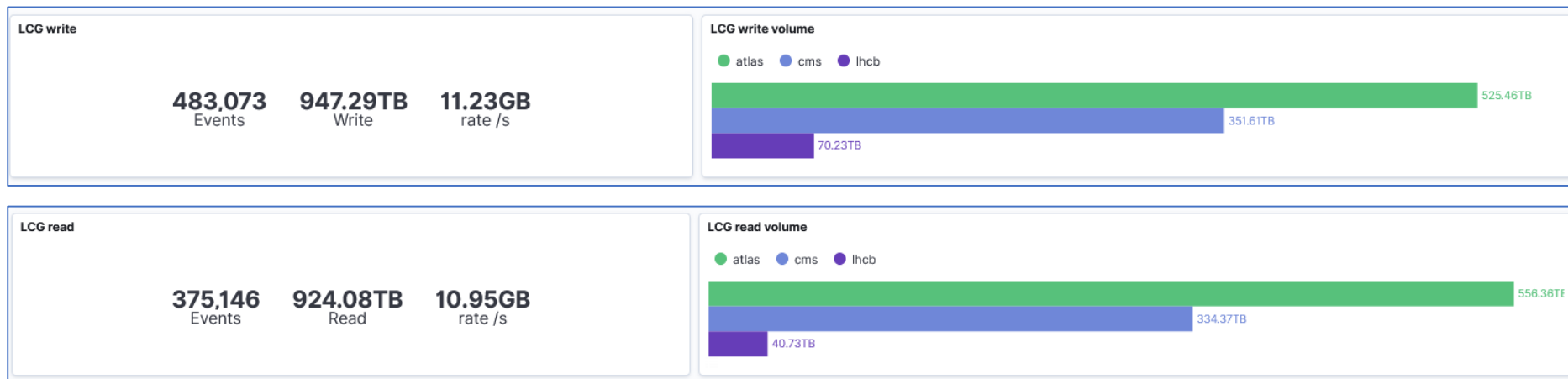
# Résultats : Vue CC d'archive ( ATLAS, CMS, LHCb )



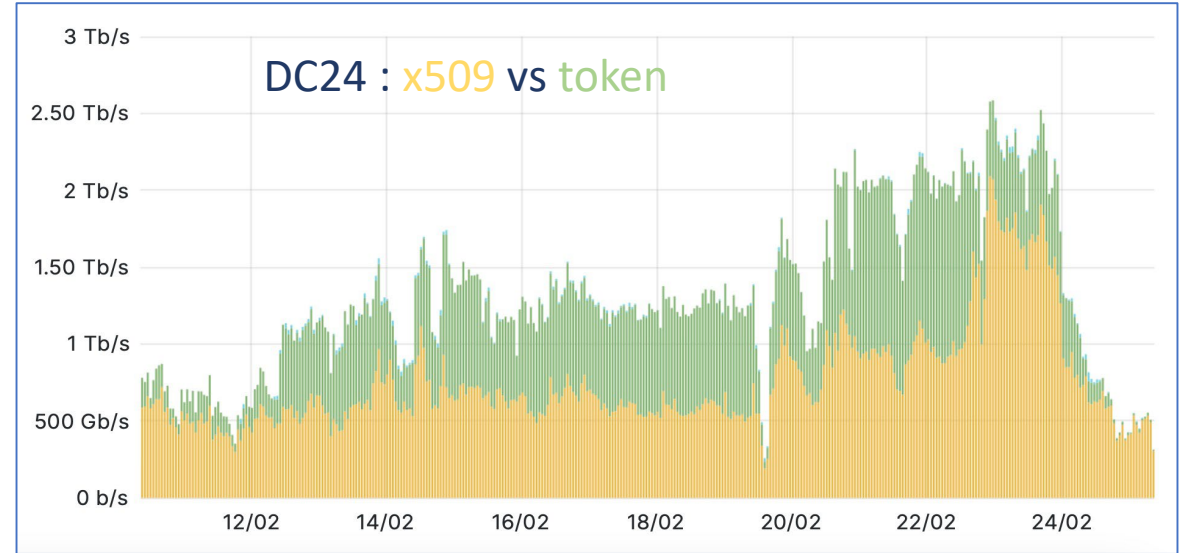
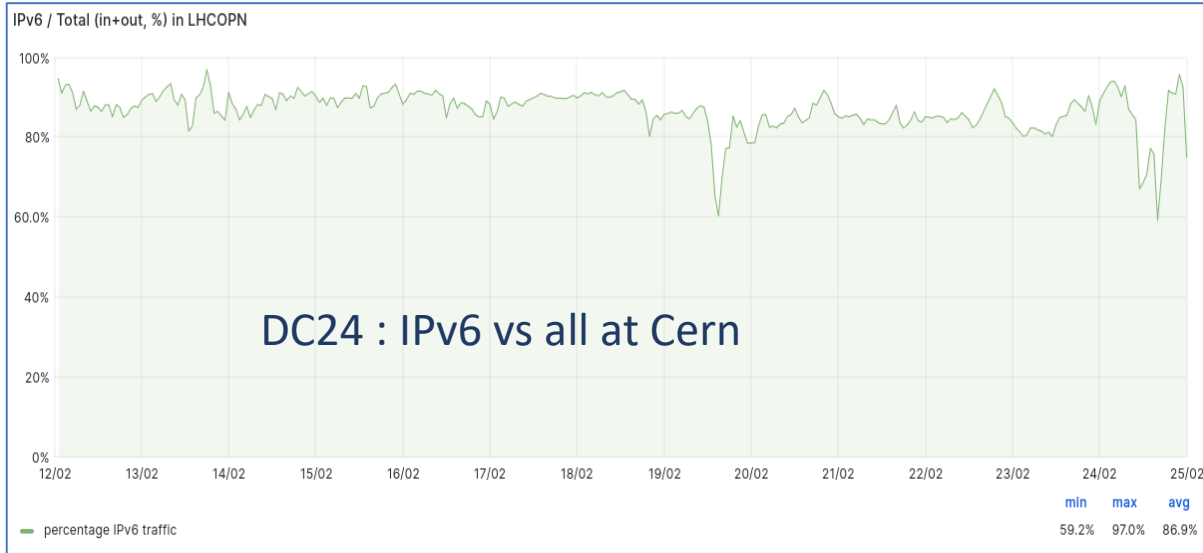
DC24 : Dcache LCG CC Xrootd



DC24 : Dcache LCG CC 24heure

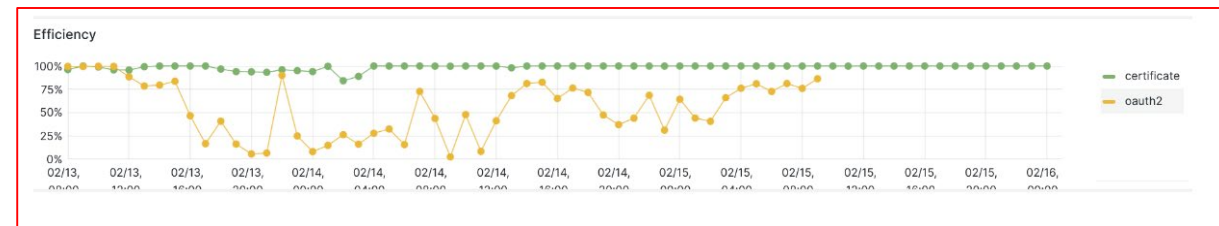






Experiment	Issued tokens	Max. number of tokens in DB	Peak token request rate	Typical token request rates
ATLAS	2.6 M	1.03 M	5 Hz	3 Hz (12 days)
CMS	2.7 M	0.97 M	200 Hz	60 Hz (6h), 20 Hz (10h), 1 Hz (11 days)
LHCb	3.4 M	1.65 M	120 Hz	25 Hz (2 days), 1 Hz (10 days)

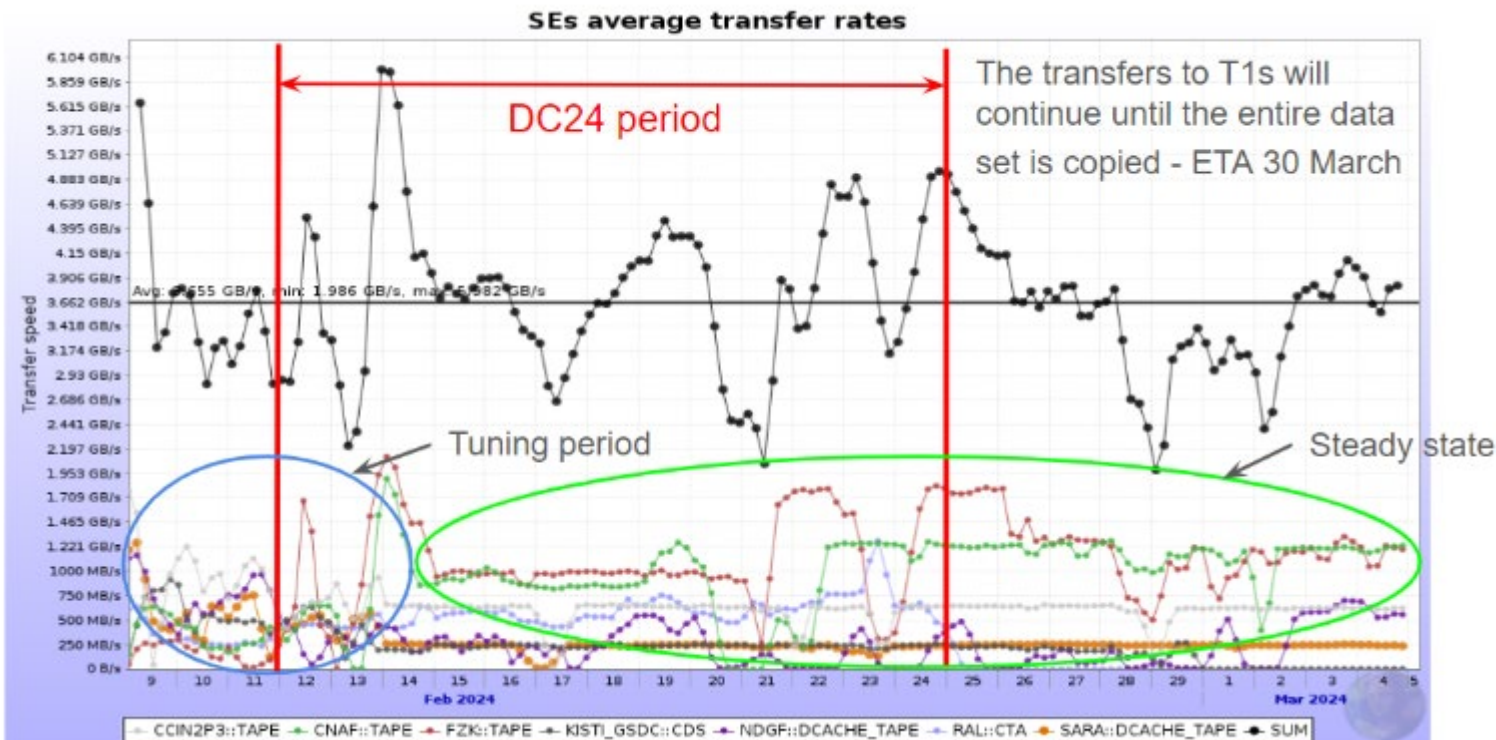
Table 4.4a: Token counts across experiments



Transfert Eff : token vs certificat

# ALICE

Pour ALICE les DC24 ont coïncidés avec la période de transfert vers les Tier1 des data des collisions plomb plomb collectés en 2023. Par conséquent les tapes ont été sollicitées



Centre	Target rate GB/s	Average achieved GB/s
CNAF	0.8	0.98 (+20%)
IN2P3	0.4	0.6 (+40%)
KISTI	0.2	0.25 (+22%)
GridKA	0.6	1.12 (+90%)
NDGF	0.3	0.35 (+15%)
NL-T1	0.1	0.25 (+150%)
RAL	0.1	0.58 (+500%)
CERN	10	14.2 (+40%)

Débits attendus et réalisés ( moyenne)

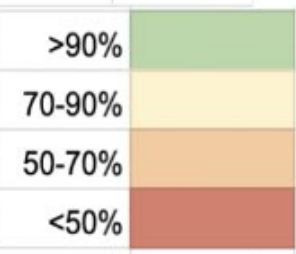
## ATLAS

Table: DC24 (src)	Site WAN (Gb/s)	Common to all scenarios	DC24 minimal scenario			DC24 flexible scenario			FTS active inbound / outbound
	Usable by ATLAS	T0 Export	Total Gb/s & bandwidth		Space [TB/24h] (deletions/hour)	Total Gb/s & bandwidth		Space [TB/24h] (deletions/hour)	
Site			$\sum$ Ingress	$\sum$ egress		$\sum$ Ingress	$\sum$ egress		
CERN-PROD	891	257.0	23.4	282.5	246 (3505)	88.9	392.8	937 (13330)	454 / 2037
BNL-ATLAS	400	60.0	84.5	67.1	892 (12681)	119.8	124.9	1263 (17964)	719 / 851
FZK-LCG2	144	32.0	55.9	35.5	590 (8386)	92.9	65.5	980 (13939)	473 / 410
IN2P3-CC	177	38.0	59.8	43.0	631 (8976)	93.5	77.7	987 (14032)	543 / 429
INFN-T1	62	23.0	36.3	26.0	383 (5447)	61.2	46.1	645 (9177)	230 / 209
NDGF-T1	149	15.0	44.6	23.3	471 (6692)	95.6	33.7	1009 (14345)	593 / 106
SARA-MATRIX	238	15.0	31.0	16.4	327 (4650)	60.1	30.2	634 (9020)	164 / 139
pic	85	11.0	17.1	12.5	181 (2570)	29.0	20.9	306 (4355)	141 / 150
RAL-LCG2	177	38.0	64.7	40.3	683 (9709)	92.8	81.0	978 (13915)	1595 / 663
TRIUMF-LCG2	100	25.0	38.2	27.8	402 (5723)	60.0	50.9	632 (8996)	322 / 434

Objectifs DC24 Tier1 ATLAS

Day	Scenario	BNL-ATLAS		FZK-LCG2		IN2P3-CC		INFN-T1		NDGF-T1		pic	
		dst	src	dst	src	dst	src	dst	src	dst	src	dst	src
1	T0 → T1	25.68	N/A	29.76	N/A	35.6	N/A	21.84	N/A	12.56	N/A	10.48	N/A
2	T0 → T1	35.1	N/A	13	N/A	41	N/A	23.52	N/A	9.79	N/A	14.5	N/A
3	T0 → T1 ↔ T1 → T2	61.6	67.1	47.4	42.2	43.8	39.3	32.1	28	7.72	26.5	18.4	10.8
4	T0 → T1 ↔ T1 → T2	65.3	79.7	61.8	58.5	64.6	47.2	31.8	50.1	4.92	22.7	30.3	15.2
5	T0 → T1 ↔ T1 → T2	63	116	81.3	78.4	75.6	56.6	37.8	52.3	7.59	18.1	32.7	13.1
6	T0 → T1 ↔ T1 → T2	73.7	98.9	85	77.9	71.1	51	39.1	60	4.8	20.2	29.5	21.8
7	T0 → T1 ↔ T1 → T2	65.7	94	79.6	102	63.6	44.8	33.7	69.5	2.2	11.2	33.6	43.8
8	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	52.8	77.3	59.5	56.5	38.9	50.8	33.7	20	2.99	33.1	24.5	19.1
9	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	87.9	80.7	51.6	63.6	40.1	34.8	46.1	48.6	2.41	33	39.3	28.8
10	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	90	95.9	43.7	97.5	39.6	36.8	47.6	50.5	21.9	32.4	54	43.4
11	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	110	96.8	58.8	82.1	42.1	44.6	55.9	53.4	16.3	44.8	50.7	38.3
12	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	89.8	84.2	52.4	51.8	34	38.7	64.6	56.4	27.2	67.2	48	38.3

Day	Scenario	RAL-LCG2		SARA-MATRIX		TRIUMF-LCG2		T2 summary		T0 summary	
		dst	src	dst	src	dst	src	dst	src	dst	src
1	T0 → T1	12.16	N/A	12.64	N/A	19.92	N/A	N/A	N/A	N/A	188
2	T0 → T1	12.5	N/A	18.9	N/A	24.2	N/A	N/A	N/A	N/A	201
3	T0 → T1 ↔ T1 → T2	16.7	40.2	34.3	65.3	33.3	27.6	299	141	19.8	141
4	T0 → T1 ↔ T1 → T2	25.2	44.7	35.8	92.2	35.5	28.3	346	124	19.6	173
5	T0 → T1 ↔ T1 → T2	23.1	52.2	36.3	89.2	49.2	46.3	387	134	25.9	197
6	T0 → T1 ↔ T1 → T2	27.4	23.6	30.6	95.5	40.9	41.1	337	104	20.3	201
7	T0 → T1 ↔ T1 → T2	27.6	20.4	47.2	86.5	53.7	43.4	341	91.7	17.1	190
8	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	29.4	47.1	37.7	29.1	37.3	19.9	400	311	54	100
9	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	32.3	39.1	59.4	84	51.7	42.7	447	330	89.8	139
10	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	43.9	43	92.9	72.3	62.8	52.5	435	337	94.4	97
11	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	51.9	56	111	73.8	66.8	42.1	445	406	127	138
12	T0 ↔ T1 ↔ T1 ↔ T2 ↔ T2 ↔ T0	72.7	58.8	115	70.8	72.9	31.5	418	407	158	174

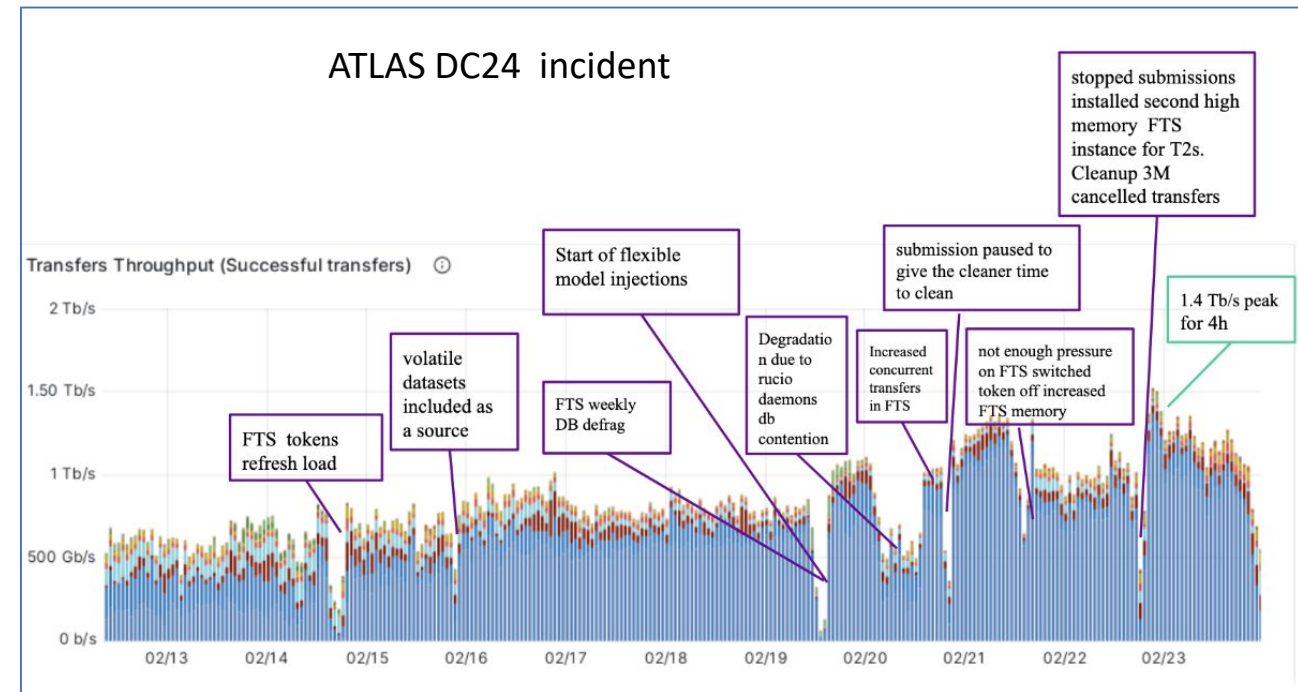


## De façon globale :

- Les tokens ont été utilisés avec tous les Tier1 (inclus CC-IN2P3)
- Problèmes de surcharge des services intermédiaires (Rucio, FTS, IAM)
- Problèmes de configuration ou de hardware des systèmes de stockage

## Au CC

- Problème de surcharge des connexions de stockage qui a déclenché un arrêt de la production d'ATLAS. Donc ATLAS a dû réduire la charge des transferts du DC et donc les débits.



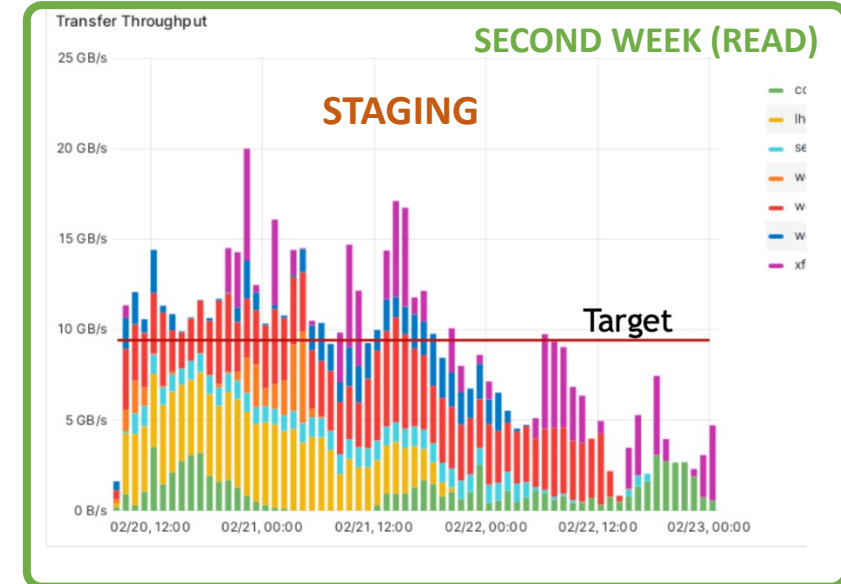
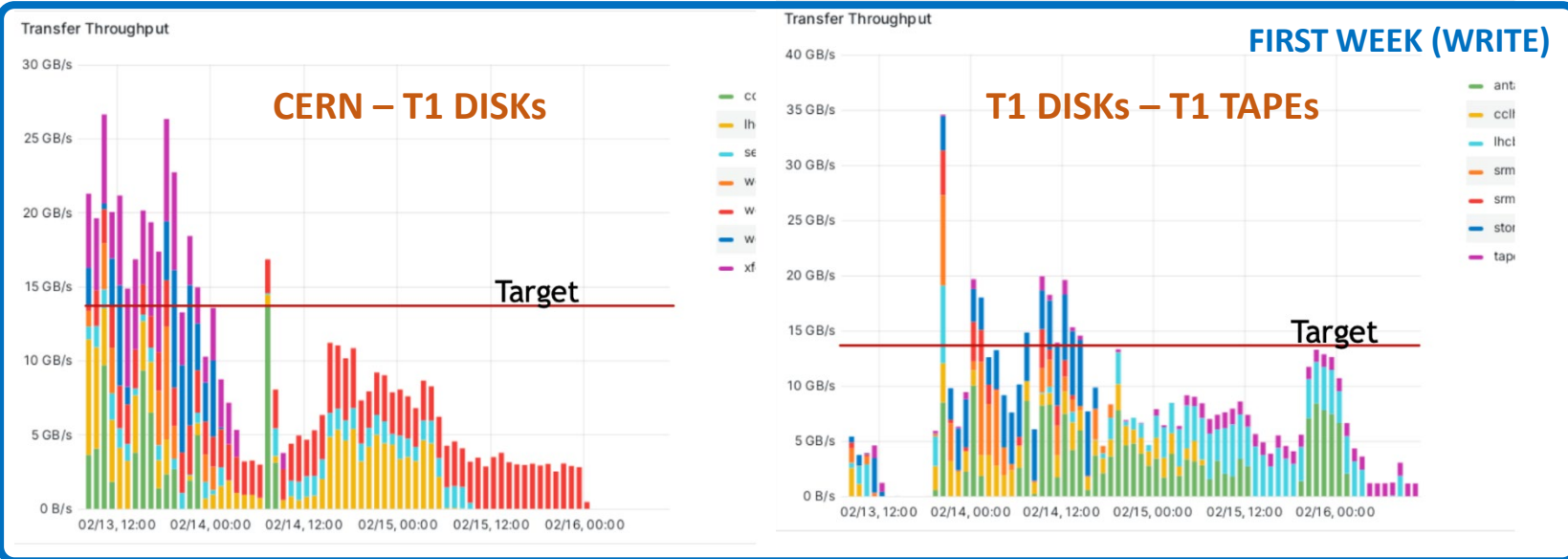
## ATLAS a souhaité rejouer après le DC24 certains tests

Site	T0 Export	DC24 best rates on day 1,2	% of expected rates	T0-T1 one T1 at the time	% of expected rates	T0-T1 collective	%
<u>BNL-ATLAS</u>	60	<u>31.5</u>	53%	<u>61.3</u>	102%	5.36	9%
<u>FZK-LCG2</u>	32	<u>26.4</u>	83%	<u>42.2</u>	132%	7.84	25%
<u>IN2P3-CC</u>	38	<u>43</u>	113%	<u>50.9</u>	134%	57	150%
<u>INFN-T1</u>	23	<u>19.3</u>	84%	<u>33.5</u>	146%	36.08	157%
<u>NDGF-T1</u>	15	<u>13.8</u>	92%	<u>28.2</u>	188%	30.88	206%
<u>SARA-MATRIX</u>	15	<u>12.2</u>	81%	<u>274.1</u>	1827%	24.96	166%
<u>pic</u>	11	<u>12.3</u>	112%	<u>18.1</u>	165%	17.28	157%
<u>RAL-LCG2</u>	38	<u>15</u>	39%	<u>27.2</u>	72%	4.24	11%
<u>TRIUMF-LCG2</u>	25	<u>23.9</u>	96%	<u>27.2</u>	109%	35.2	141%

POST DC24 T0 export



## LHCb



- **Débits cible atteints:**
  - CERN – T1 DISKS (14GB/s) pour écriture des données
  - T1 DISKS – T1 TAPES (14GB/s) pour écriture des données
  - T1 TAPES – T1 DISKS (9,58 GB/s) pour lecture des données
- **Remarques:**
  - Usage partiel de tokens que pour l'écriture des données et que pour certains sites (inclus CC-IN2P3)
  - Problèmes de surcharge du service IAM pour les tokens et FTS
  - Problèmes liés aux systèmes de stockage de plusieurs sites

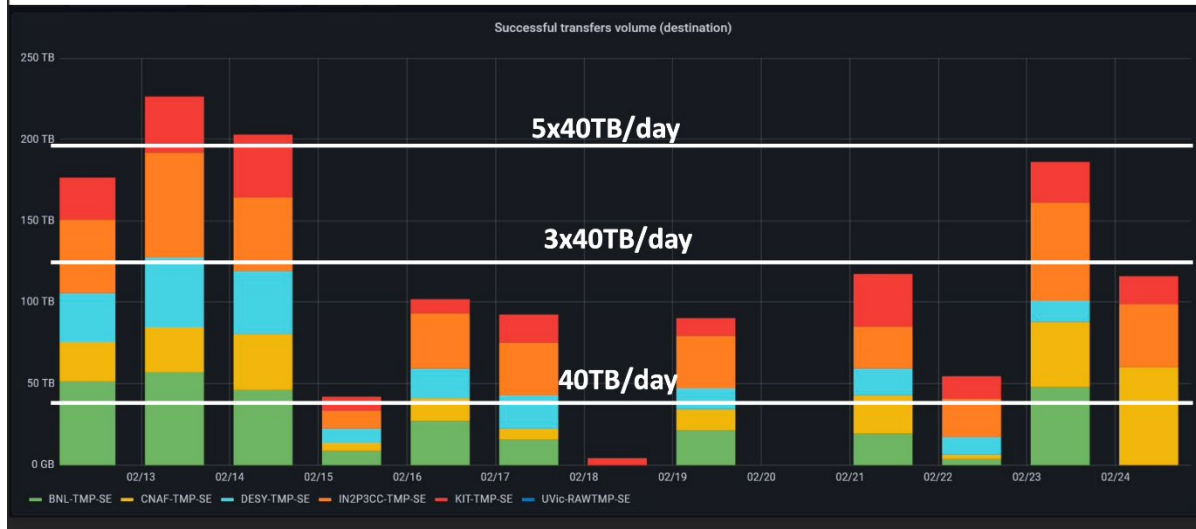
- **Débits cible CERN – CC-IN2P3 partiellement atteints:**
  - 1.53GB/s pour écriture
  - 1.42GB/s pour lecture
- **Remarques:**
  - Problème de performance non lié au CC-IN2P3 (première semaine)
  - Problème de saturation d'espace de stockage temporaire (deuxième semaine)

Targets, GB/s			Achieved, GB/s			Ratio (achieved/target)		
Site	Write	Stage	EOS-Disk	Disk-Tape	Tape-Disk	EOS-Disk	Disk-Tape	Tape-Disk
CNAF	2.05	1.60	3.45	2.74	1.41	1.68	1.34	0.88
GRIDKA	2.74	1.66	2.50	1.65	3.35	0.91	0.60	2.01
IN2P3	1.53	1.20	2.56	1.42	1.05	1.67	0.93	0.88
NCBJ	1.02	0.89	0.953	0.602	0.798	0.93	0.59	0.90
PIC	0.51	0.40	1.21	0.553	1.05	2.37	1.08	2.63
RAL	3.96	2.40	2.68	2.64	3.28	0.68	0.67	1.37
SARA	1.15	0.80	2.77	1.39	1.17	2.40	1.20	1.46

## BELLE2

## Traffic per Day vs Goals

Belle II



- **Débits cible KEK – T1 DISKS atteints:**
  - 120TB/d (11.1 Gbit/s) pour le scénario minimal
  - 200TB/d (18.5 Gbit/s) pour le scénario maximal
- **Remarques:**
  - Pas de tokens
  - Débit moyen de 26 Gb/s atteint dans les fenêtres de trafic maximal générées par les expériences LHC.

@CC-IN2P3



- **Débits cibles KEK – CC-IN2P3 atteints:**
  - Minimal: 18To/jour (1.7 Gb/s)
  - Maximal: 30To/jour (2.8 Gb/s)

## CMS

## CMS schedule

- Different activities on different days helped us 'ramp' up and spot issues

Date	12 Feb	13 Feb	14 Feb	15 Feb	16 Feb	17 Feb	18 Feb	19 Feb	20 Feb	21 Feb	22Feb	23 Feb
	T0 export	T0 export	T0 export	T1 export	T1 export	T1 export	T1 export	AAA	T0 export	T0 export	T0 export	T0 export
			T1 export		Prod. output	Prod. output	Prod. output		T1 export	T1 export	T1 export	T1 export
									Prod. output	Prod. output	Prod. output	Prod. output
									AAA	AAA	AAA	AAA
Scenario(s)	1	1	1,2	2	2,3	2,3	2,3	4	1,2,3,4	1,2,3,4	1,2,3,4	1,2,3,4
Rate (GB/s)	31	31	62	31	62	62	62	31	125	125	125	125
Rate (Gb/s)	250	250	500	250	500	500	500	250	1000	1000	1000	1000

- T1 export: Day 3, T1s exported to 'nearby' T2s. Different source for Day 4
- The final 4 days involved pushing all activities to achieve our target of a sustained rate of 125GB/s (1000Gbit/s)

7

## T1\_FR\_CCIN2P3

Coped very well in the first week with writes at the lower rates. This is consistent with the very good pre-DC24 test results which tested at the T0Export (days 1-3) scale.

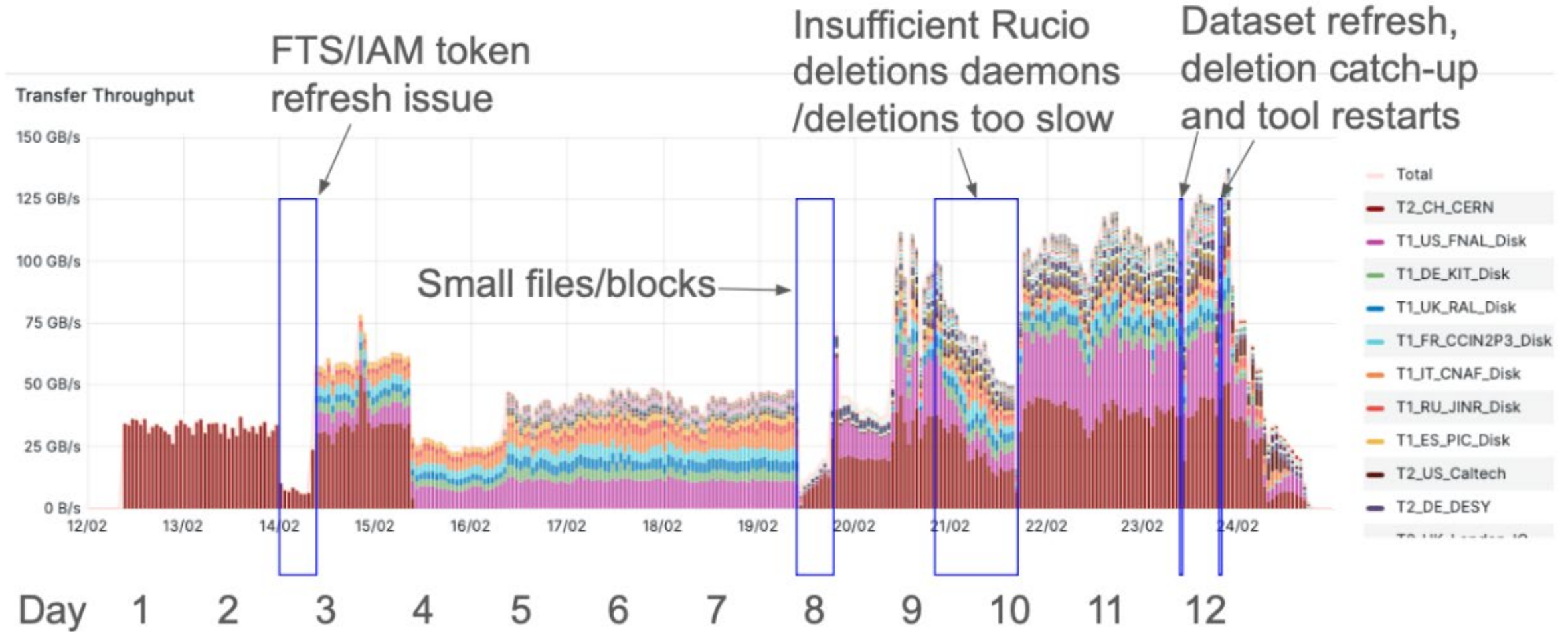
Our records say we changed the FTS config for IN2P3 twice - clearly the first time it didn't stick. Second attempt was on final day

IN2P3

Expected	Observed	Ratio	Expected	Observed	Ratio
as DEST			as SRC		
3.35	<u>3.66</u>	1.09	0	0	N/A
3.35	<u>3.68</u>	1.10	0	0	N/A
3.35	<u>4.47</u>	1.33	3.88	<u>3.98</u>	1.03
0	0	N/A	3.04	<u>3.41</u>	1.12
1.4	<u>1.75</u>	1.25	5.08	<u>4.52</u>	0.89
1.4	<u>2.06</u>	1.47	5.08	<u>4.48</u>	0.88
1.4	<u>1.71</u>	1.22	5.08	<u>4.43</u>	0.87
0.69	<u>0.96</u>	1.39	0	0	N/A
7.66	<u>4.39</u>	0.57	5.02	<u>2.84</u>	0.57
8.34	<u>4.87</u>	0.58	5.02	<u>3.25</u>	0.65
8.34	<u>3.56</u>	0.43	5.02	<u>3.83</u>	0.76
8.34	<u>7.42</u>	0.89	5.02	<u>5</u>	1.00

Performance here limited by FTS configuration?





## FTS

- **Une surcharge de la base de donnée**
  - Une des cause principale était le rafraichissement des tokens
- **Un mismatch entre les optimisation que FTS voulait faire et les objectifs du DC**
  - FTS maximise les nombre de transfert dans les liens
  - Les DC maximisent ( ou plutôt cherche maximiser) le débit

## Ruci

- **Une difficulté dans la mise en oeuvre des règles (rules) de pilotage du data management**
  - Liée à la gestion de la charge
  - Une conséquence significative a été une difficulté à supprimer les data.

## Une série de tests techniques réseau a également fait partie du DC24

- **Packet marking**
  - Afin d'améliorer le monitoring sur les réseau LHCONE et LHCOPN les flux seront marqués pour les identifier.
    - Deux approches différentes entre IPV4 et IPV6.
- **Routage des flux**
  - Modification dynamique des routes réseau en fonction des besoins et de la charge.
    - NOTED (Network Optimized Transfer of Experimental Data) : via SDN
- **Optimisation de la couche réseau**
  - Jumbo frame.
  - Gestion de la congestion des paquets au niveau kernel (Packet pacing).
    - Algorithme BBR versus BIC.

## Au niveau des infrastructure

- Pas de soucis au niveau des réseaux.
- Pas de soucis ( quelques configs à optimiser) au niveau des stockages.

## Des limitations identifiées au niveau des services de gestion des transferts

- Un refactoring des service FTS est en cours.
- Une plus forte intrication entre FTS et Rucio est en cours.

## De nouveaux services qui ont été validés

- Utilisation des tokens dans la gestion des transferts.
- Services de gestion des tokens ( stabilité et scalabilité).

## Une organisation qui a fonctionnée malgré un nombre important de parties prenantes

- Des sites.
- Des collaborations.
- Des NREN.
- Des développeurs.

**Rapport final :** <https://zenodo.org/doi/10.5281/zenodo.11401877>

## Data Transfers

Monthly data transfer throughput between WLCG sites (GB/s) – 5 years

