

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

HL-LHC Data Challenge 2024

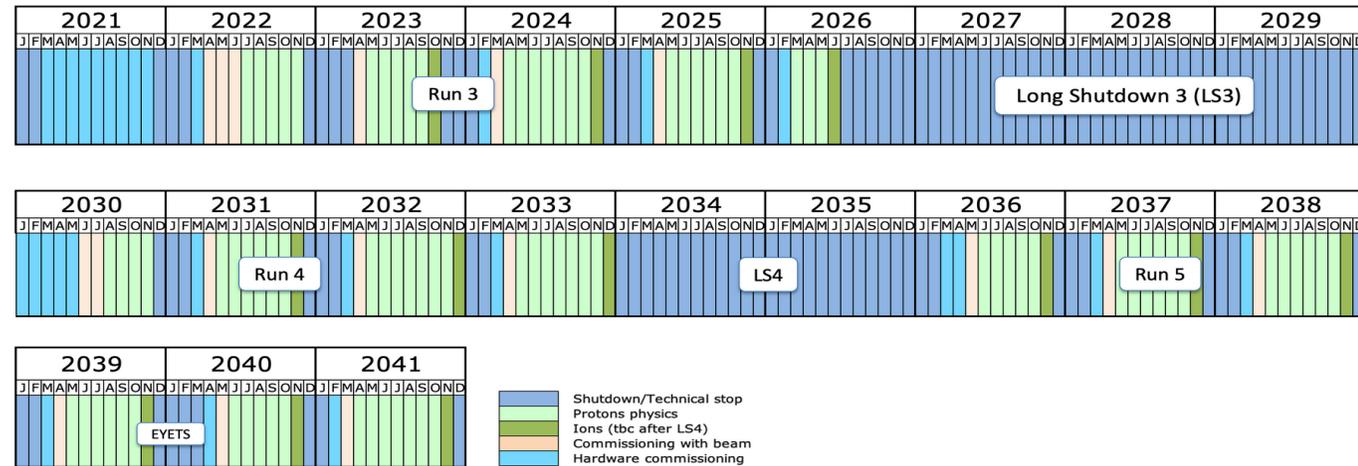
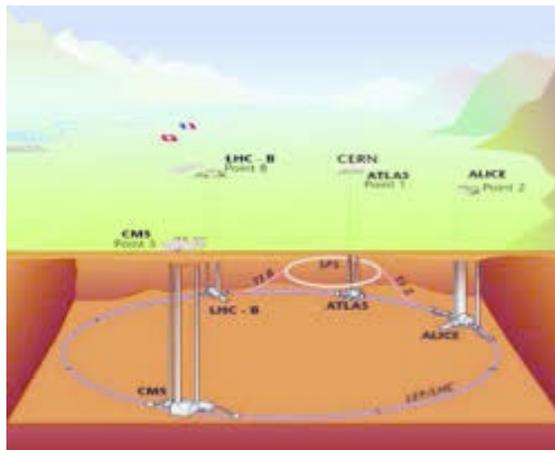
- **Contexte et motivations**
- **De quoi on parle ?**
 - De sites/data center
 - De datas
 - D'infrastructures
 - De services
 - De collaborations
- **Les objectifs du Data Challenge DC24**
- **Les moyens mis en œuvre**
- **Le déroulé et résultats**
- **Conclusion**

• Contexte

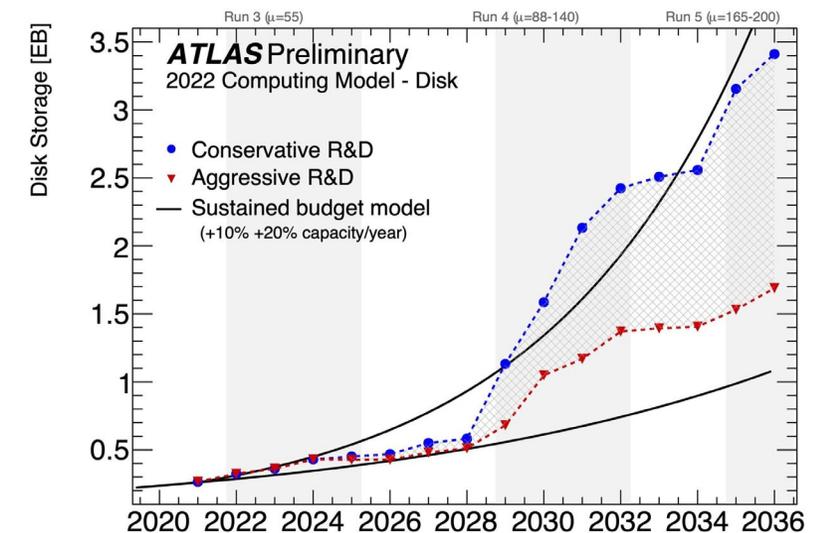
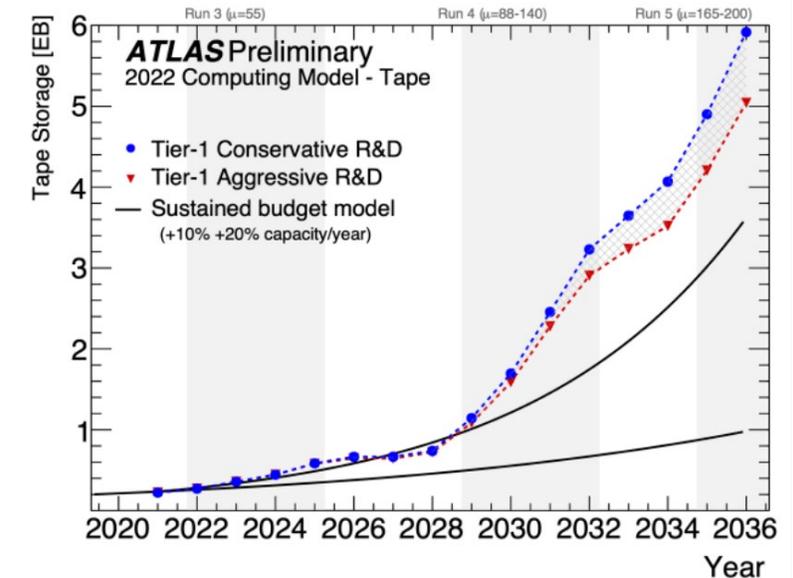
- Préparer les infrastructures informatiques à l'upgrade majeur qui va être fait des expériences de physiques des hautes énergies auprès du collisionneur LHC (Large Hadron Collider :CERN) : High Luminosity LHC
- Pas une révolution mais une évolution de l'infrastructure actuelle distribuée (grille) à l'échelle internationale.

• 4 Collaborations

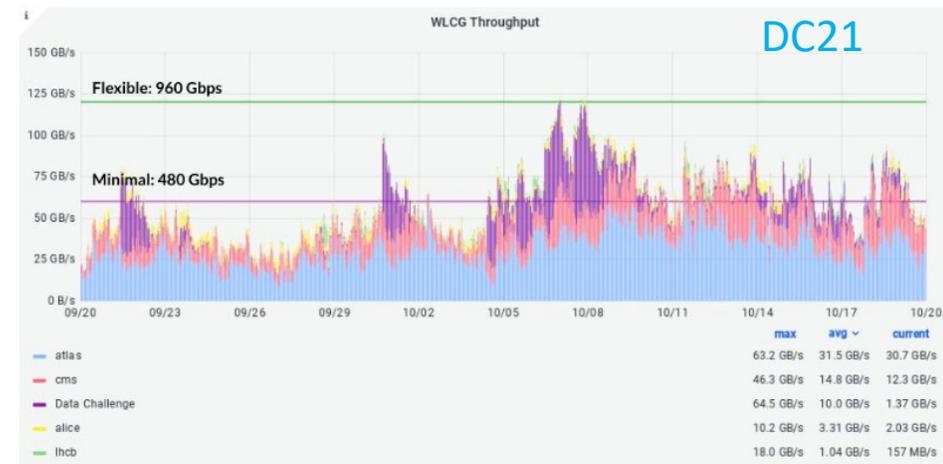
- Alice
- ATLAS
- CMS
- LHCb



- **Meilleure performance de l'accélérateur**
 - Plus de collisions de particules.
 - Plus de superposition de collisions (pileup).
 - Donc plus de data
- **Mise à niveau des expériences du LHC**
 - Quasiment toutes les données sont gardées pour analyses contrairement au LHC ou seulement une partie sortait des détecteurs.
 - Donc plus de data
- **Une problématique de data management à prendre en compte**
- **Donc plus de « calcul » aussi mais hors du scope de ce talk**

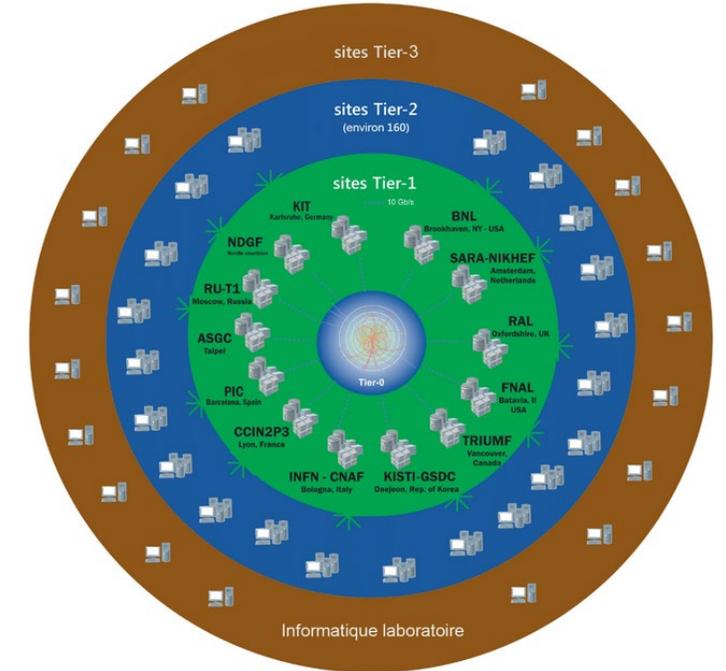


- Depuis ~2020 des initiatives de R&D sont prises pour évaluer les potentielles solutions techniques/organisationnelles/... qui permettrait de satisfaire le besoin
 - Ex DOMA : Présentation faite au JCAD 2019
- En parallèle des ces initiatives un processus de validation a été mis en place pour régulièrement quantifier la montée en capacité vers l'objectif HL-LHC.
- Cela passe par des campagnes (Data Challenge) durant lesquelles l'infrastructure de grille est chargée à son maximum et testé.
 - DC21 : Objectif 10 % de la cible
 - DC24 : Objectif 25 % de la cible
 - Cible HL- LHC :
 - Deux scenarios
 - 9680 Gb/s et 4810 Gb/s



- **Sites de la grille WLCG**

- 3 niveaux de site : Tier 0, Tier 1, Tier 2
 - On a un Tier 0 = CERN
 - On a ~ 14 Tier 1 : CC-IN2P3 pour la France
 - On a ~ 160 Tier 2



- **Ces sites exposent**

- Des services de calcul
- Des services de stockage
- Des services divers de grille (notamment de data management)
- Ont des engagements en ressources, en SLA,...

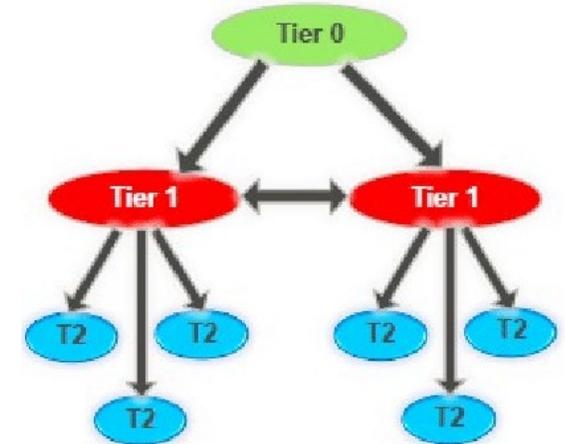
Les data qui sont manipulées par les collaborations peuvent grossièrement être classées en trois catégories.

- **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 modèles de distribution des données**

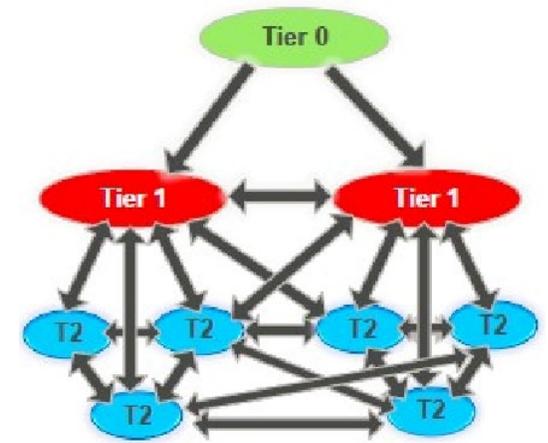
- **Modèle hiérarchique (Monarch)**

- Basé sur la topologie
- Consommateur de stockage



- **Modèle flexible**

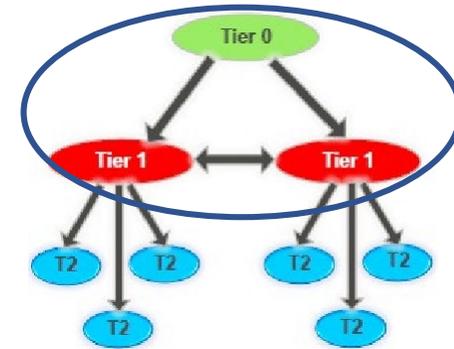
- Modèle assez chaotique
- Consommateur de réseau
- Modèle de référence aujourd'hui



Deux réseaux privés ont été déployés par les NREN nationaux

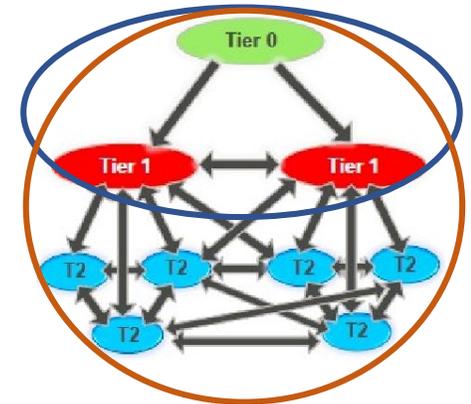
LHCOPN (Large Hadron Collider Optical Private Network) ●

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

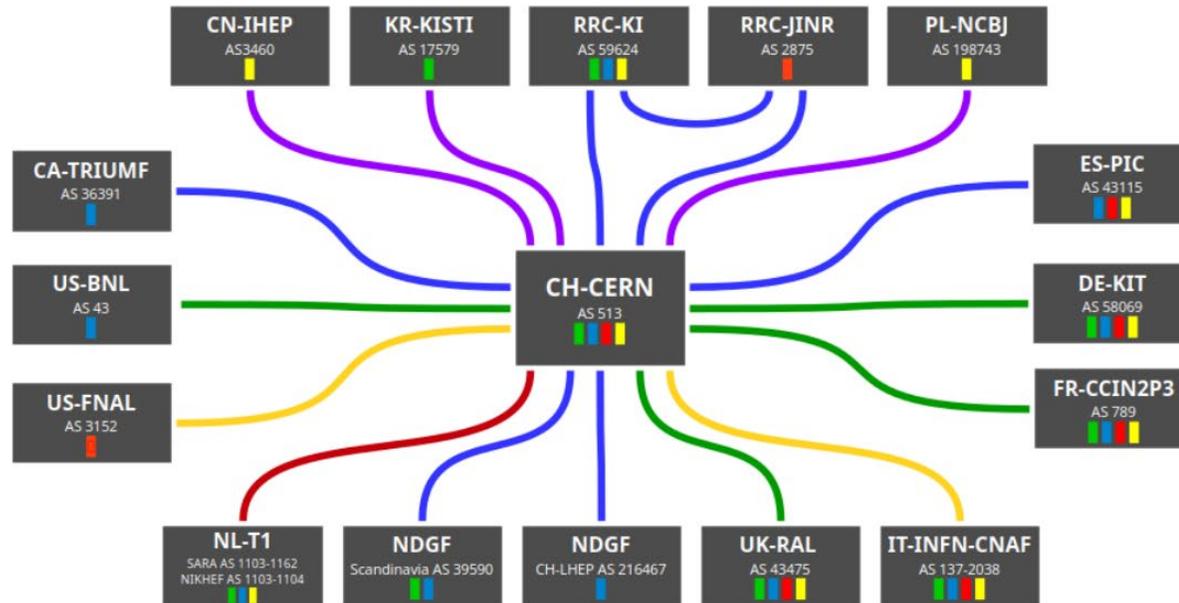


LHCONE (Large Hadron Collider Open Network Environment) ●

- 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 (purple)
 100Gbps (blue)
 200Gbps (green)
 400Gbps (yellow)
 800Gbps (red)

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

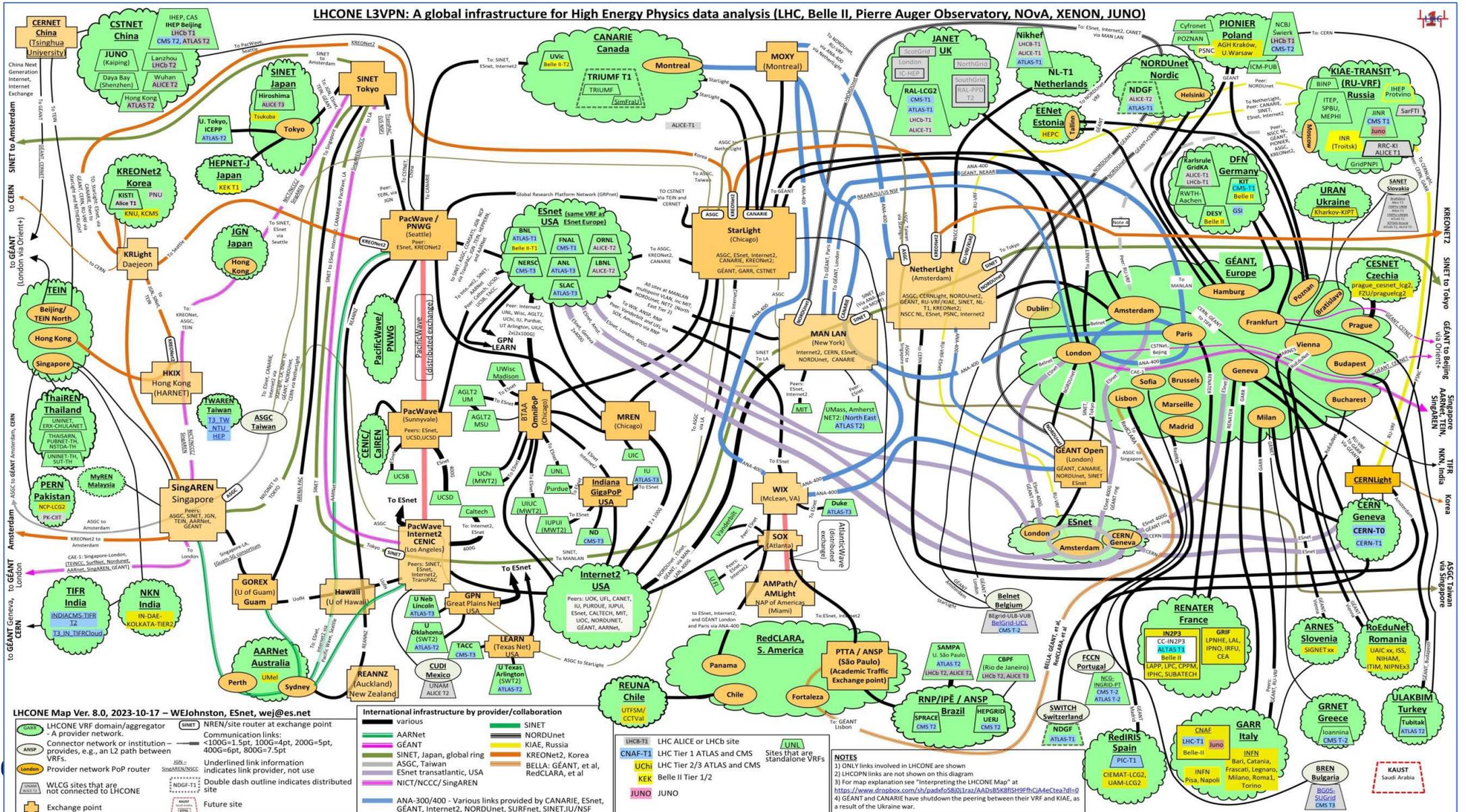
Numbers

- 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 endpoints (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 contrainte de disponibilité, de redondance,.....
- **Services d'authentification**
 - X509 via les certificats/voms
 - Par tokens via Indigo IAM
- **Du monitoring**

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

- **On a les 4 expériences auprès du LHC (CERN)**
 - ALICE, ATLAS, CMS et LHCb
- **Plus d'autres utilisant la même infrastructure de grille**
 - Belle2 (KEK Japon)
 - DUNE (US)

Acteurs du data challenge 2024

Alice

ATLAS

CMS

LHCb

BELLE2

XXXX

Collaborations

AAI

- Authentification
- Autorisation

Monitoring

Gestion des transferts

- FTS

Catalogue

- Rucio

Services

Réseaux

- LHCONE
- LHCOPN

Sites

- Endpoint

Stockage

- Tape
- Disque

LAN

Monitoring

Services

Infrastructure

- **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 ingénieur: Identifier les éléments qui sont limitants**
 - Le réseau ?
 - Le stockage ?
 - Les services ?
- **Pour tous : Optimiser les façons de faire et configurations afin de co construire l'infrastructure « qui va bien et qui fait le job »**

- **3 objectifs 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
Sum	4830	9670	613	1226	2452

Objectifs du data challenge 2024



- Un focus par site et expériences

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

TO															
CERN-PROD source (Write rates)	RRC-KI	ES-PIC	DE-KIT	FR-CCIN2P3	IT-INFN-CNAF	UK-RAL	NDGF (CH-LHEP)	NDGF (Scandinavia)	NL-T1(Nikhef, S	PL-NCBJ	CN-IHEP	RRC-JINR	CA-TRIUMF	US-BNL	US-FNAL
ALICE			5	4	7	1		3	1	0	0	0	0	0	0
ATLAS (injected + prod))		0,8	13	38,4	43,5	27,7	43,5	0	24,4	18,9	0	0	0	28,6	67,4
CMS		0	19	45	45	57	35	0	0	0	0	0	68	0	17
LHCb		0	4,38	23,54	13,14	17,61	34,02	0	0	9,88	8,76	8,93	0	0	0
Total		0,8	36,38	111,94	105,64	109,31	113,52	0	27,4	29,78	8,76	8,93	68	28,6	67,4
Network Capacity[1]	100Gbps	100Gbps	200Gbps (400Gbps)	200Gbps	400Gbps	200Gbps	100Gbps	100Gbps	400Gbps (800Gb	20Gbps	20Gbps	100Gbps	100Gbps	200Gbps	200Gbps
DUNE															
Belle II (from KEK via LHCONE)		0	0	1,9	2,8	3,7	0	0	0	0	0	0	0	0	5,6
CERN-PROD destination (Read rates)															
CERN-PROD destination (Read rates)	RRC-KI	ES-PIC	DE-KIT	FR-CCIN2P3	IT-INFN-CNAF	UK-RAL	NDGF (CH-LHEP)	NDGF (Scandinavia)	NL-T1(Nikhef, S	PL-NCBJ	CN-IHEP	RRC-JINR	CA-TRIUMF	US-BNL	US-FNAL
ALICE			n/a	n/a	n/a	n/a		n/a	n/a						
ATLAS (injected + prod))		0,08	1,64	6,36	6,57	4,19	6,16	0	3,27	3,42	0	0	0	5,94	10,86
CMS		0	15	36	36	45	28	0	0	0	0	0	41	0	25
LHCb		0	3,44	14,26	10,31	13,74	20,62	0	0	6,87	7,65	5,41	0	0	0
Total		0,08	20,08	56,62	52,88	62,93	54,78	0	3,27	10,29	7,65	5,41	41	5,94	10,86
Network Capacity[1]	100Gbps	100Gbps	200Gbps (400Gbps)	200Gbps	400Gbps	200Gbps	100Gbps	100Gbps	400Gbps	20Gbps	20Gbps	100Gbps	100Gbps	200Gbps	200Gbps

Objectifs du data challenge 2024



• Une planification

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
ALICE	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1
ATLAS	T0 → T1	T0 → T1	T0 → T1 ↔ T1 → T2				
CMS	T0 → T1	T0 → T1	T0 → T1 → T2	T1 → T2	T1 ↔ T2	T1 ↔ T2	T1 ↔ T2
LHCb		T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1	T0 → T1
DUNE	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2	T0 → T1 → T2
Belle II	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)

	Monday 19/02/2024	Tuesday 20/02/2024	Wednesday 21/02/2024	Thursday 22/02/2024	Friday 23/02/2024	
ALICE	T0 → T1	yellow: "reduced minimal" (only T0 export)				
ATLAS	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
CMS	AAA T1 → T2	T0 → T1 ↔ T2	T0 → T1 ↔ T2	T0 → T1 ↔ T2	T0 → T1 ↔ T2	red: flexible scenario
LHCb	T0 → T1	T1 Tape Recall	T1 Tape Recall	T1 Tape Recall	T1 Tape Recall	green: cms minimal T1->T2
DUNE	T0 → T1 → T2					
Belle II	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 IN2P3

- **Réseau**

- Renater a déployé 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**

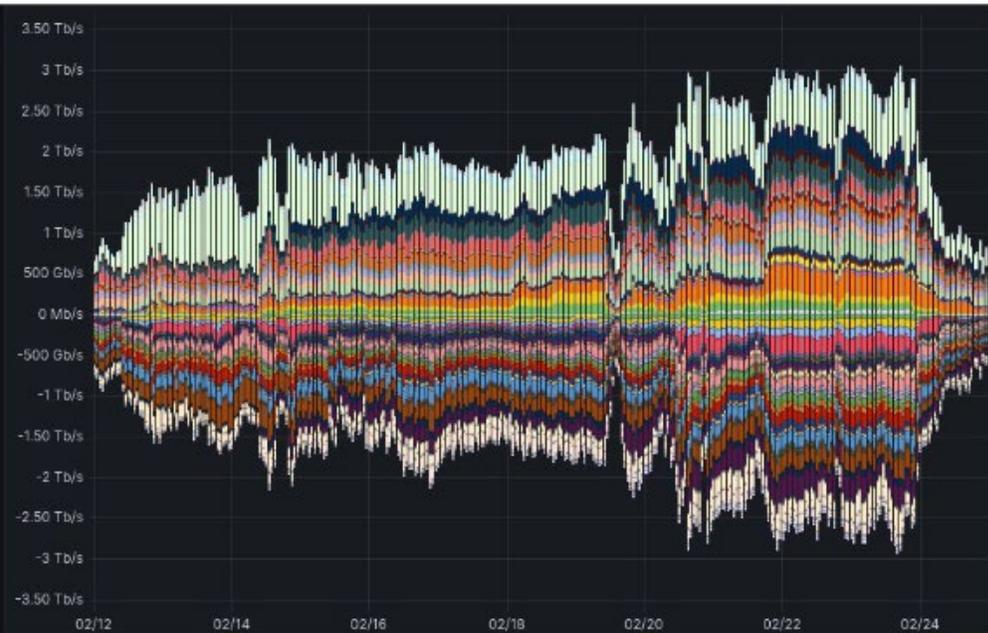
- Les instances de dcache (ATLAS,CMS,LHCb et Belle2) et XrootD (Alice) ont été upgradées afin de mettre en œuvre certaines nouvelles fonctionnalités requises pour le DC24.

- **Monitoring**

- Au fils des besoins certains monitoring ont pu être rajoutés

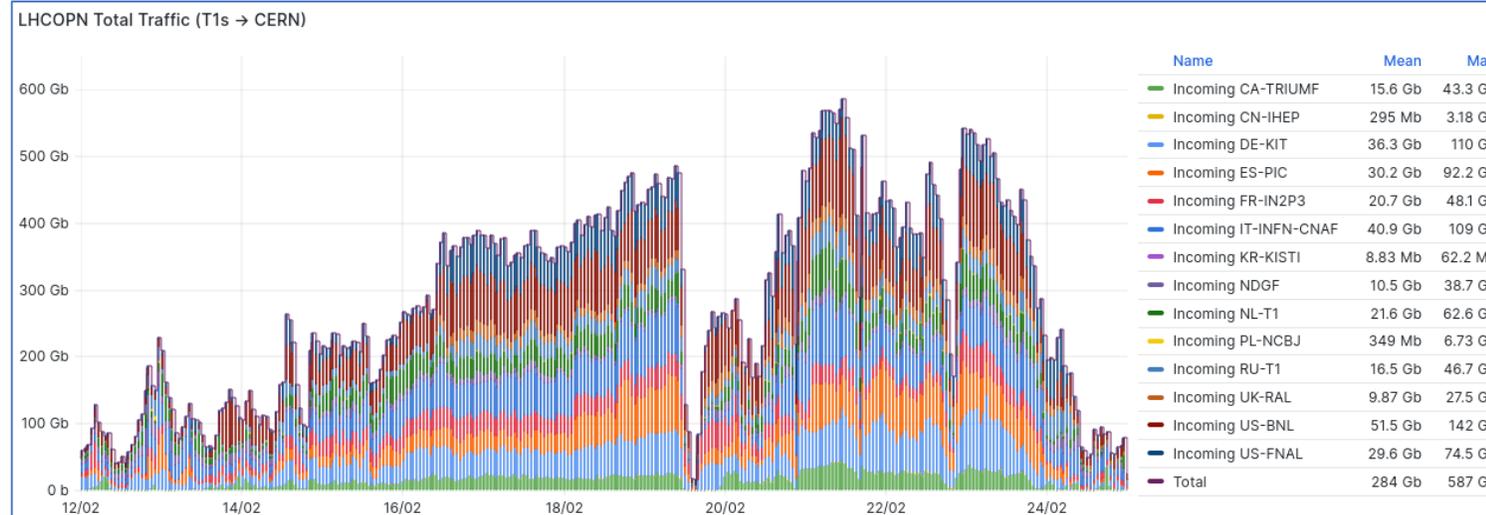
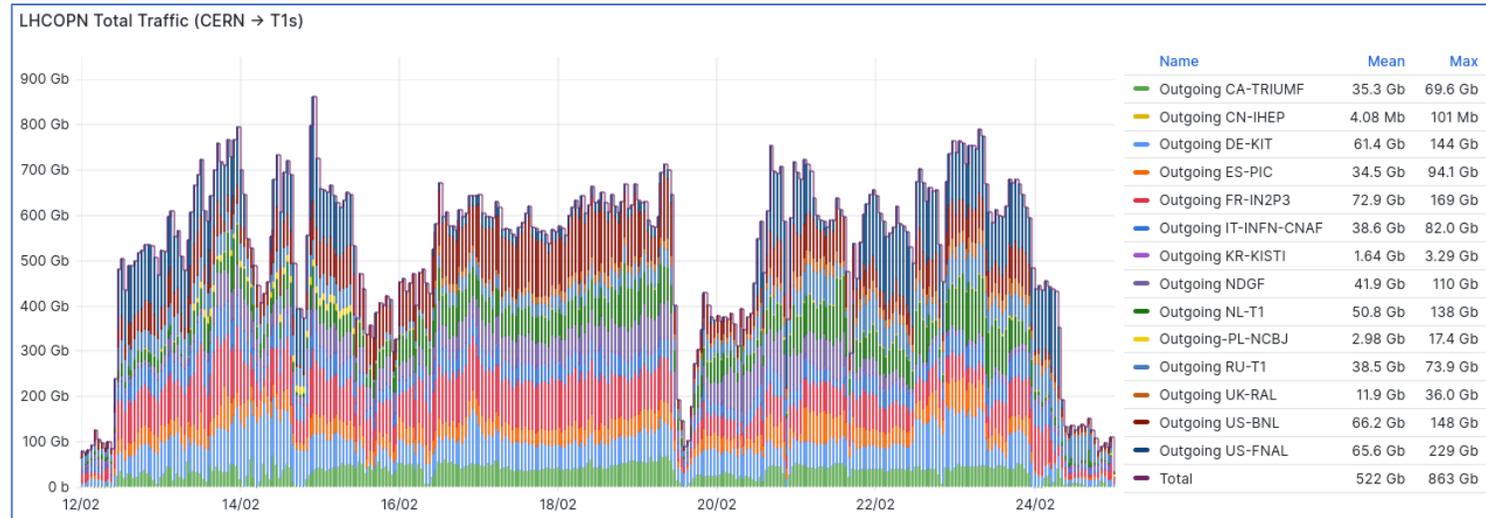
Résultat : Vue globale réseau

DC24 : LHCOPN CERN –Tier1

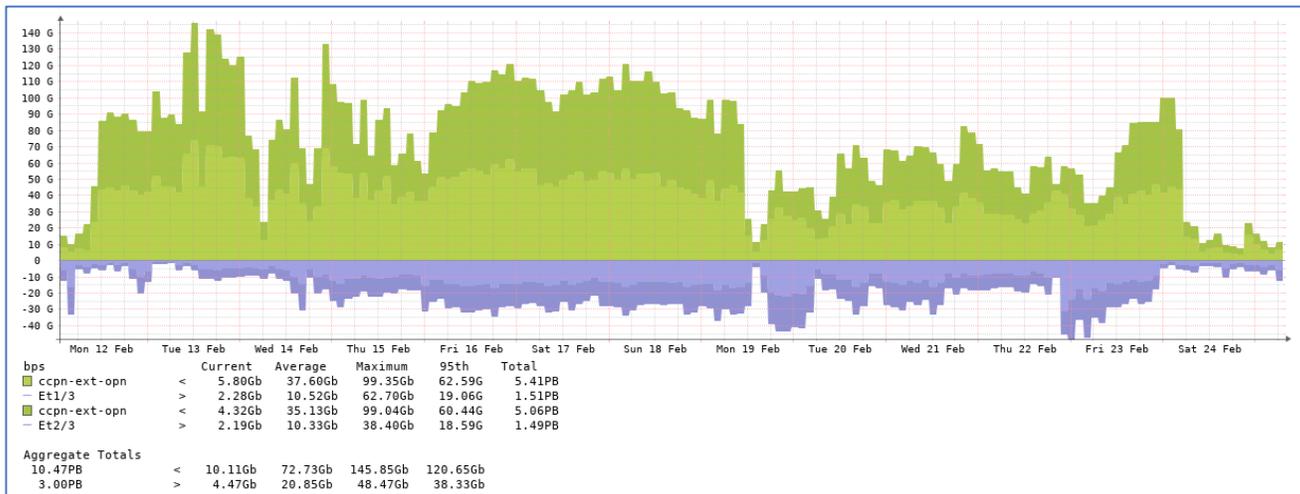
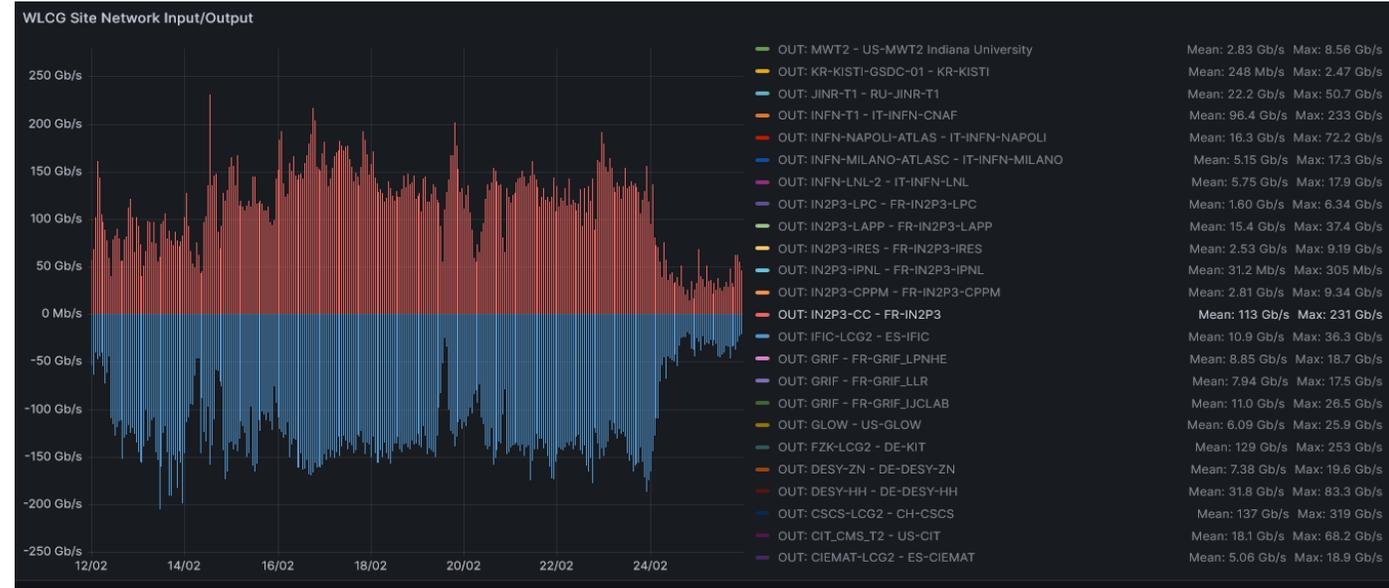


DC24 : ALL Network, All site IN/out

DC24 : LHCOPN T1 –CERN



DC24 : CC in/out

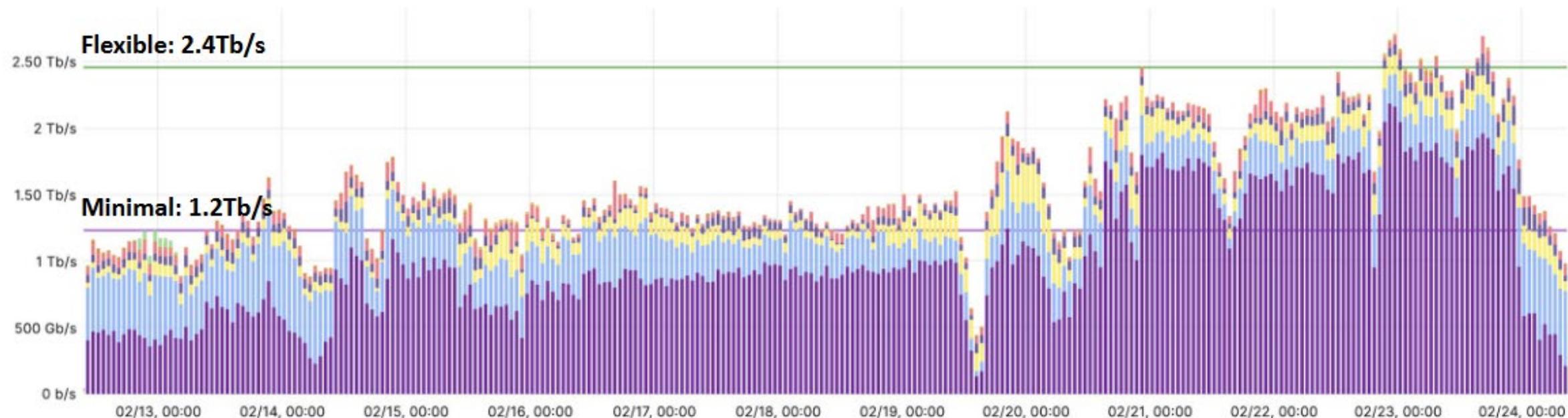


DC24 : CC LHCOPN in/out

Résultats : Vue globale des transferts

WLCG Throughput ⓘ

DC24 : Transferts en une vue



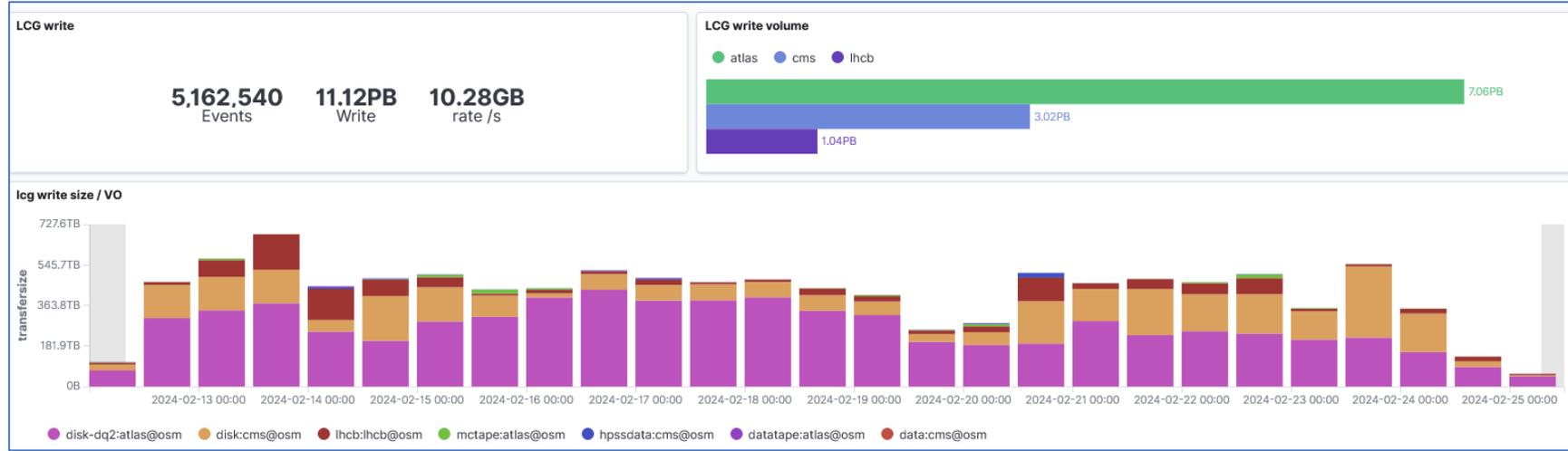
- Data Challenge
- atlas
- alice xrootd
- cms xrootd
- cms
- belle

DC24 met the (main) goals:

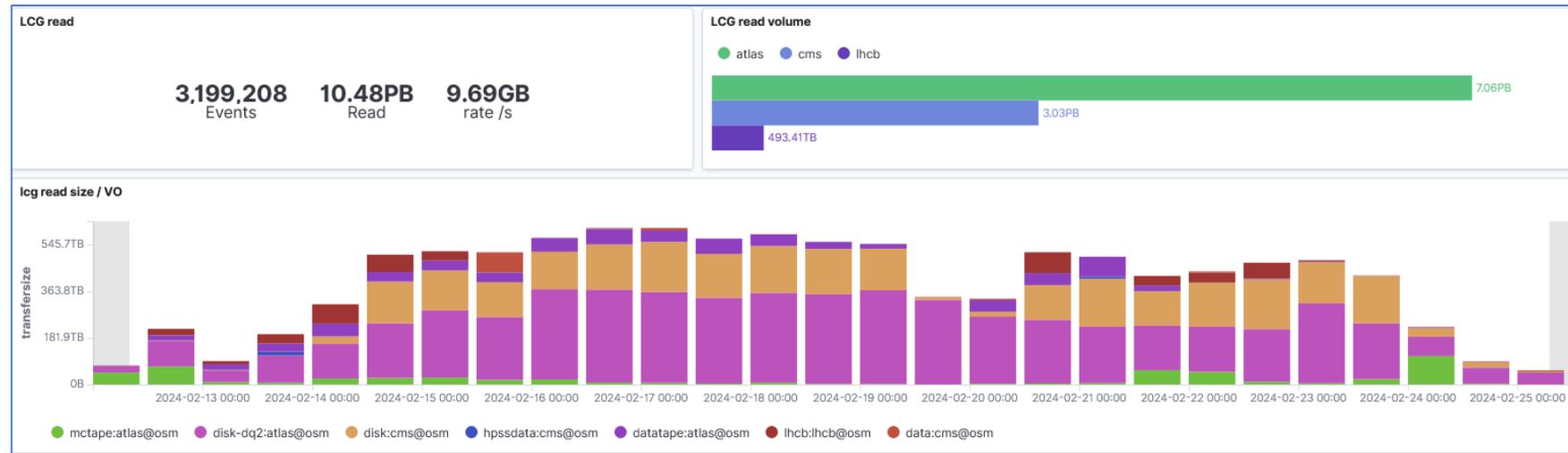
- Achieved full throughput of minimal model (1st week)
- Push for flexible target (2nd week)

	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



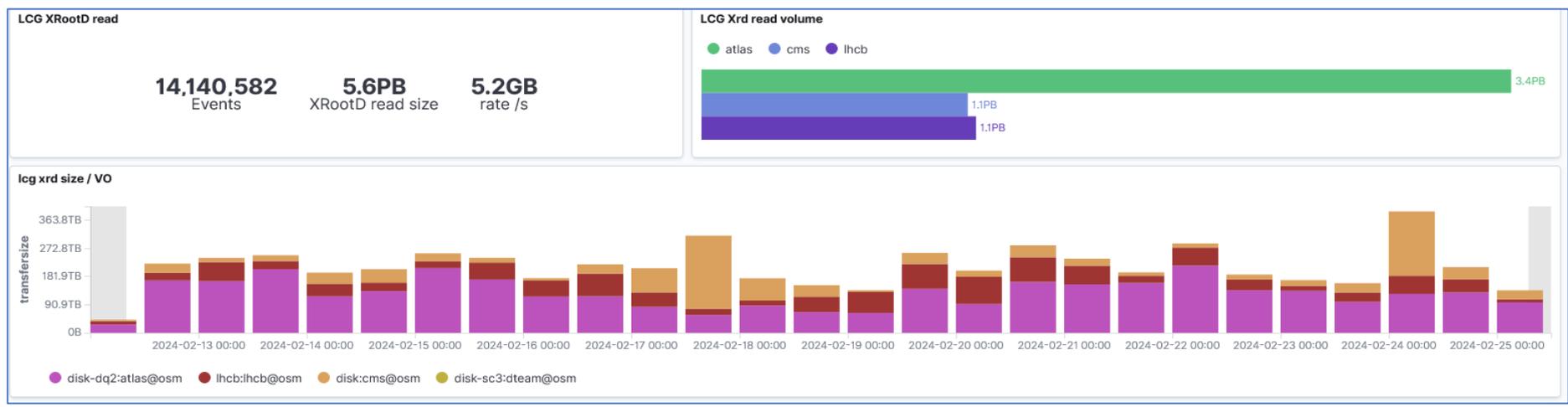
DC24 : Dcache LCG CC Read



Résultats : Vue CC stockage (ATLAS, CMS, LHCb)

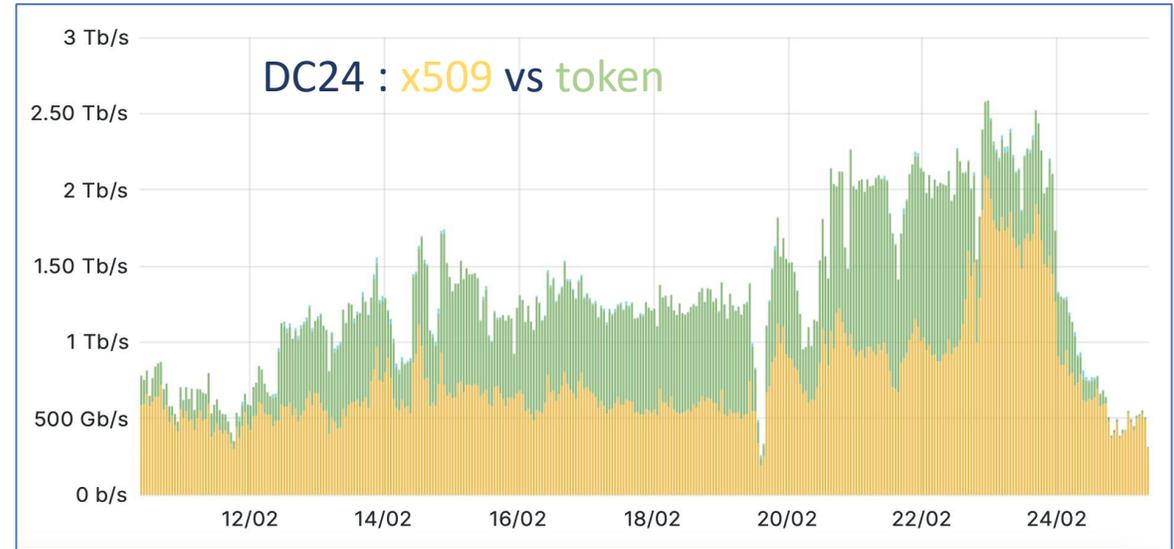


DC24 : Dcache LCG CC Xrood



DC24 : Dcache LCG CC 24heure





Dans le modèle de data management le plus chaotique (seconde semaine) donc que tous les sites Tier0, Tier1 et Tier2 sont sollicités de façon non hiérarchisé

- **Problème de monté en charge au niveau des services de gestion des transferts et des catalogues**
 - Limitations sur les DB sous jacentes.
 - Des « collisions» d'optimisation entre les services de transfert des data et les actions de data management (transferts , effacements, priorités de transfert,...)
- **Le nouveau mécanisme d'identification (via des token vs x509) sous dimensionné et/ou sursollicité.**

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 infrastructures

- 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 service de transfert (FTS) et catalogues (Rucio) est visée.

De nouveaux services qui ont été validés

- Utilisation des tokens dans la gestion des transfert
- Services de gestion des tokens

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

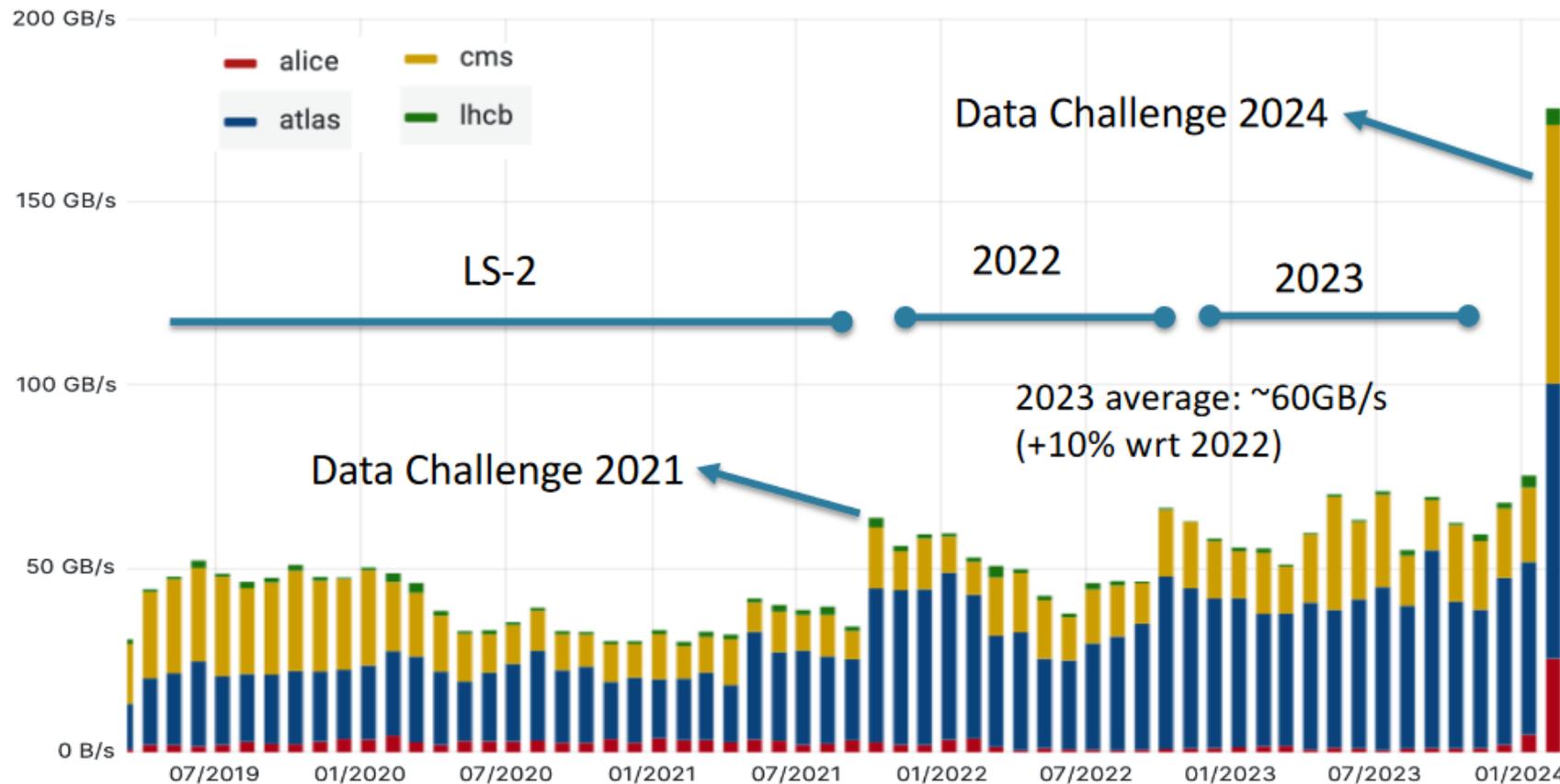
- Des sites
- Des collaborations
- Des NREN
- Des développeurs

Prochain Data Challenge en 2026 avec la possibilité d'initier des mini challenge à l'échelle local

Rapport final : <https://zenodo.org/records/11401878>

Data Transfers

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



BACKUP

Questions



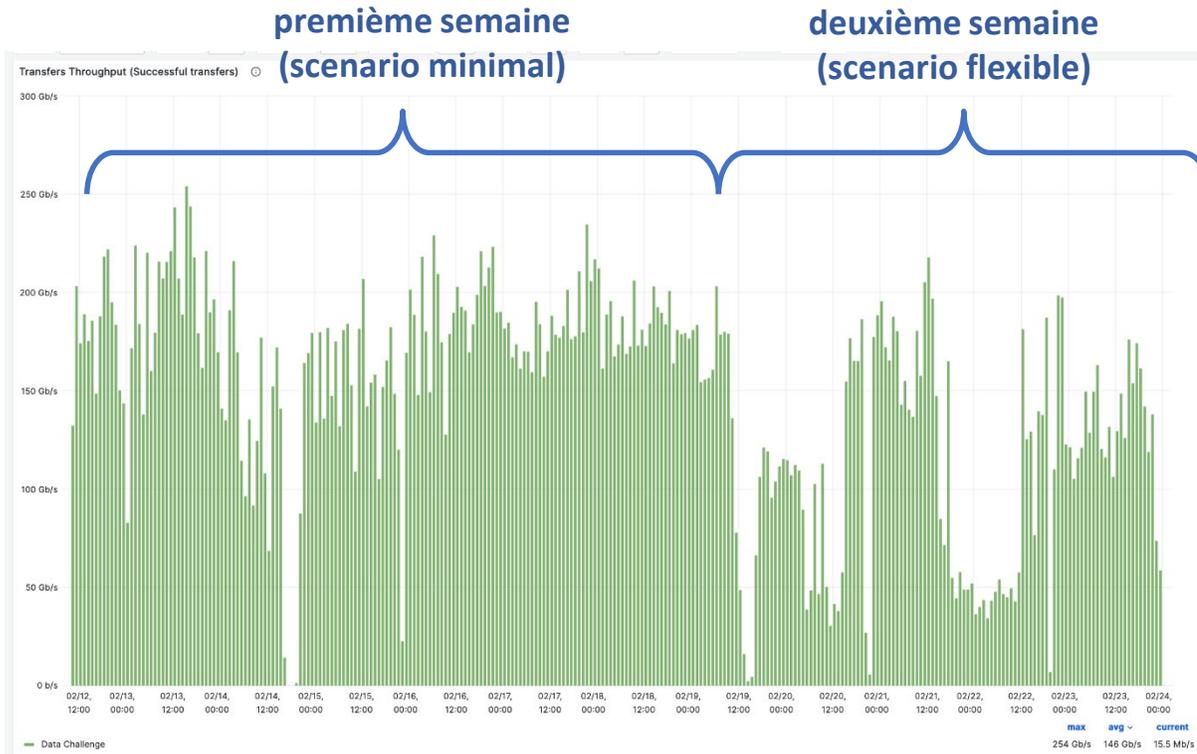
ATLAS

• **Débits cible CERN – T1 DISKs pas atteints:**

- 242Gb/s pour la première semaine
- 286Gb/s pour la deuxième semaine

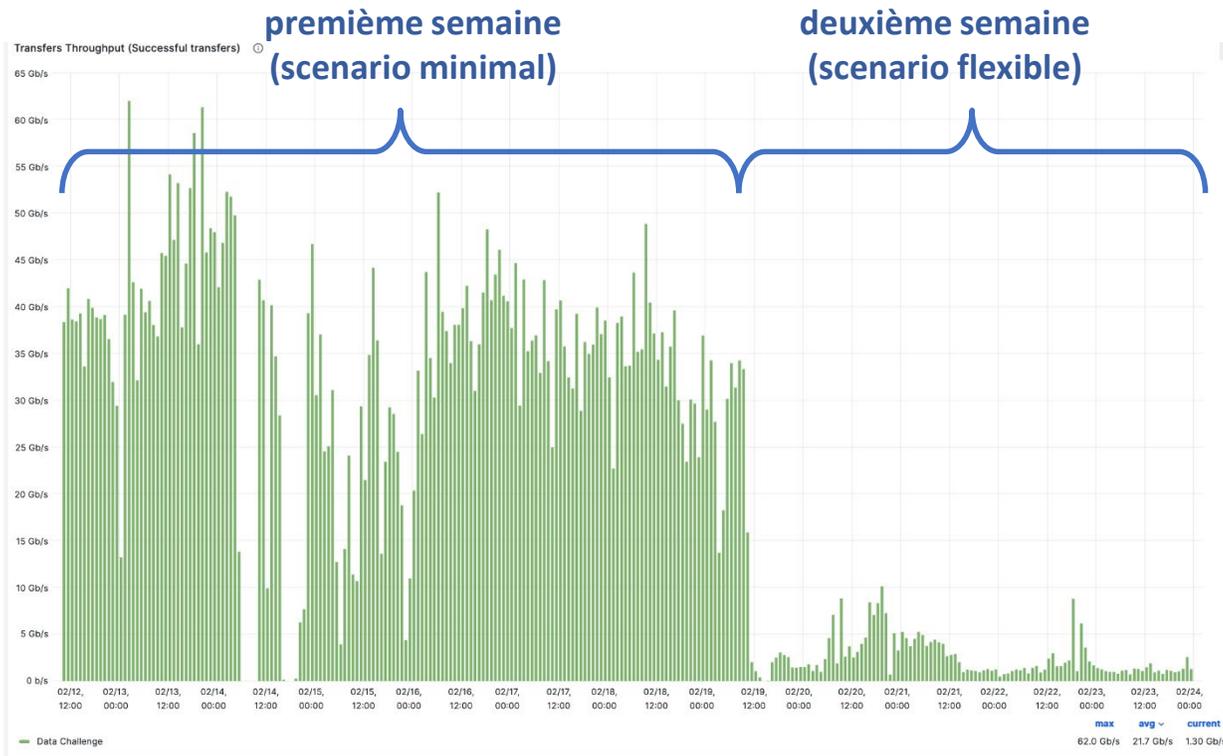
• **Remarques:**

- 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



T1 Site	Minimal (T0→T1)			Flexible (T0→T1)			Flexible (T0+T1→T1)		
	model	reality	[%]	model	reality	[%]	model	reality	[%]
BNL-ATLAS	60.0	25.9	43	68.4	21.2	31	82.1	57.1	70
FZK-LCG2	32.0	34.1	107	39.0	13.2	34	59.4	43.2	73
IN2P3-CC	38.0	36.4	96	44.2	1.4	3	59.1	21.4	36
INFN-T1	23.0	22.0	96	28.3	8.9	31	39.4	47.6	121
NDGF-T1	15.0	0.7	5	24.4	0.0	0	52.2	0.0	0
SARA-MATRIX	15.0	17.9	119	19.3	32.8	170	36.2	84.6	234
pic	11.0	13.8	126	13.3	4.2	32	18.1	35.7	198
RAL-LCG2	38.0	12.5	33	44.4	29.7	67	56.9	48.4	85
TRIUMF-LCG2	25.0	26.0	104	29.3	12.5	43	38.6	54.0	140
Σ (no NDGF)	242.0	188.6	78	286.3	123.9	43	389.8	392.0	101

- **Débits cible CERN – CC-IN2P3 pas atteints:**
 - 38Gb/s pour la première semaine
 - 44.2Gb/s pour la deuxième semaine

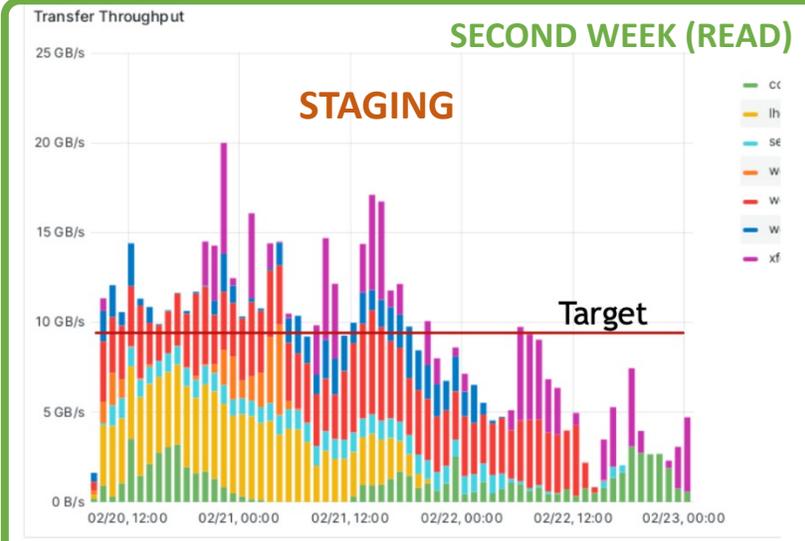
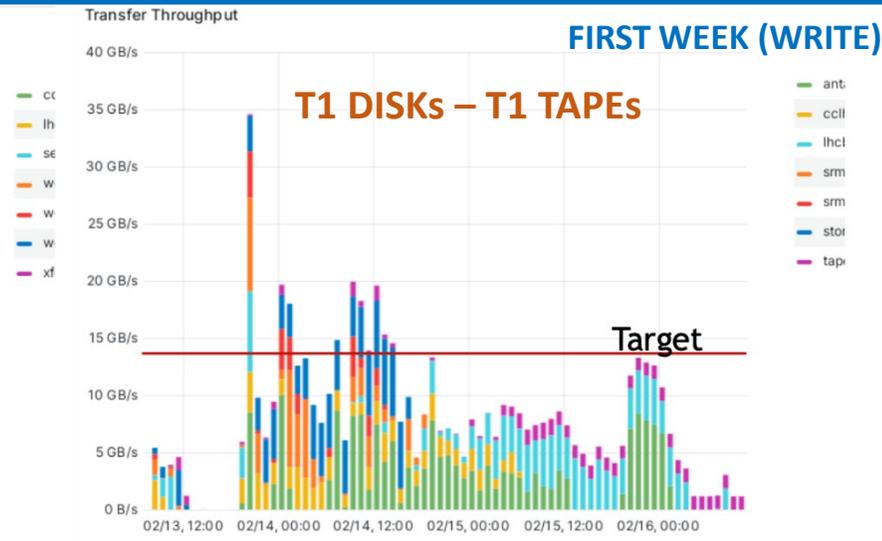
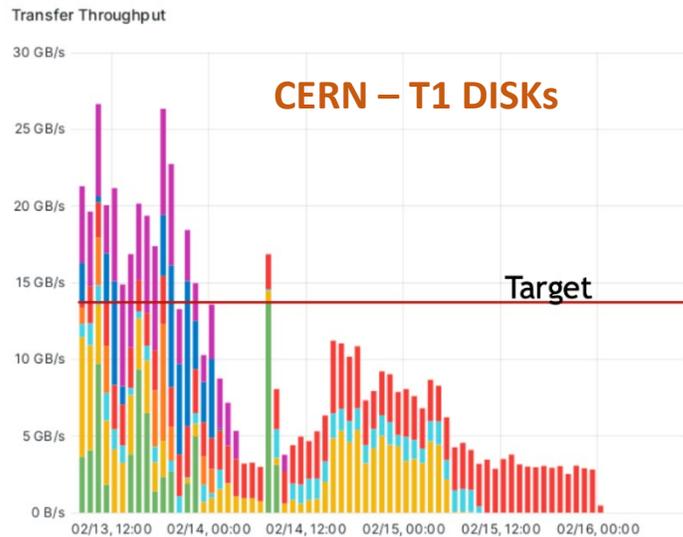


Remarques:

- Problème de performance non lié au CC-IN2P3
- Problème de surcharge des connexions de stockage qui a déclenché un arrêt de la production d'ATLAS. Donc ATLAS a dû réduire la charge de transferts du DC et donc les débits.

T1 Site	Minimal (T0→T1)			Flexible (T0→T1)			Flexible (T0+T1→T1)		
	model	reality	[%]	model	reality	[%]	model	reality	[%]
BNL-ATLAS	60.0	25.9	43	68.4	21.2	31	82.1	57.1	70
FZK-LCG2	32.0	34.1	107	39.0	13.2	34	59.4	43.2	73
IN2P3-CC	38.0	36.4	96	44.2	1.4	3	59.1	21.4	36
INFN-T1	23.0	22.0	96	28.3	8.9	31	39.4	47.6	121
NDGF-T1	15.0	0.7	5	24.4	0.0	0	52.2	0.0	0
SARA-MATRIX	15.0	17.9	119	19.3	32.8	170	36.2	84.6	234
pic	11.0	13.8	126	13.3	4.2	32	18.1	35.7	198
RAL-LCG2	38.0	12.5	33	44.4	29.7	67	56.9	48.4	85
TRIUMF-LCG2	25.0	26.0	104	29.3	12.5	43	38.6	54.0	140
Σ (no NDGF)	242.0	188.6	78	286.3	123.9	43	389.8	392.0	101

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 certaines 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 cible 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)

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.

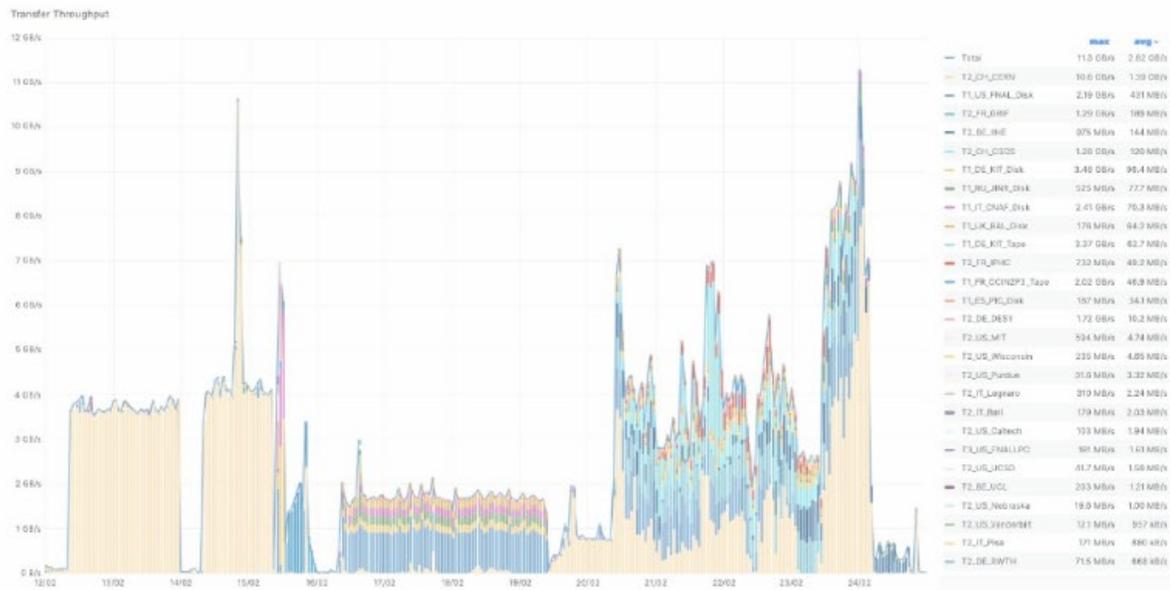
Our records say we changed the FTS config for IN2P3 twice - clearly the first time it didn't stick. Second attempt was on final day

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

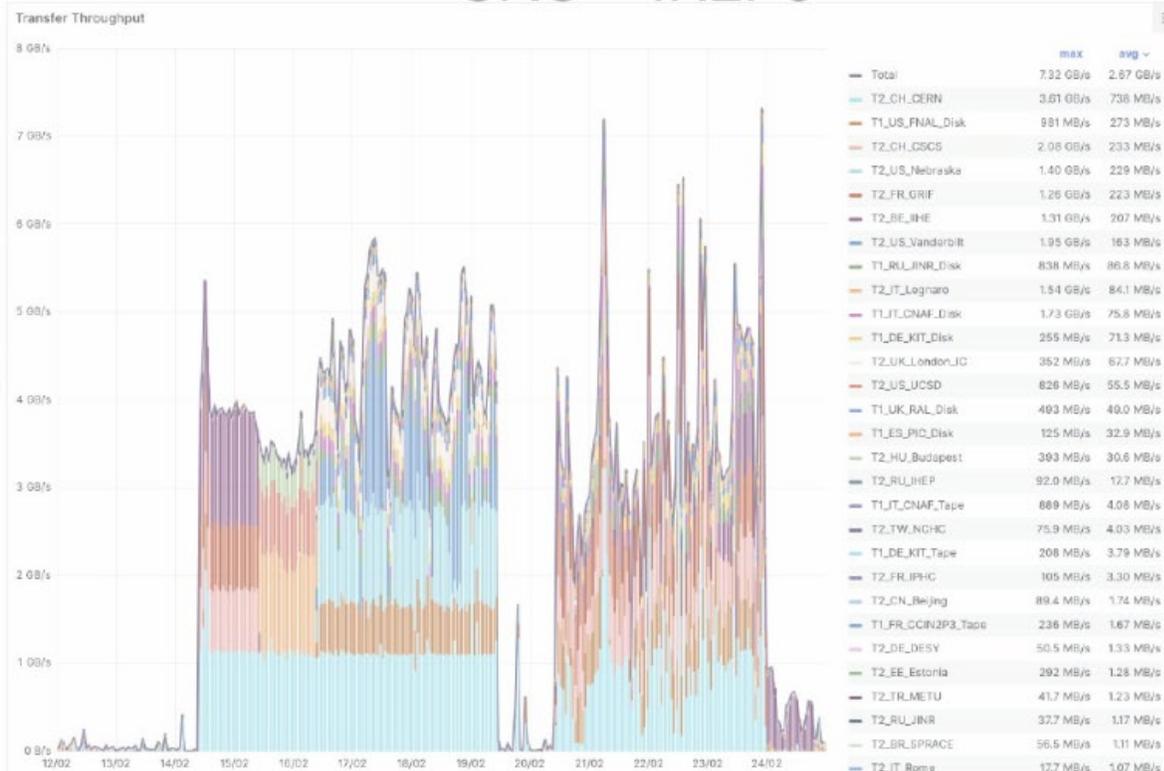
Performance here limited by FTS configuration?

T1_FR_CCIN2P3



DEST = IN2P3

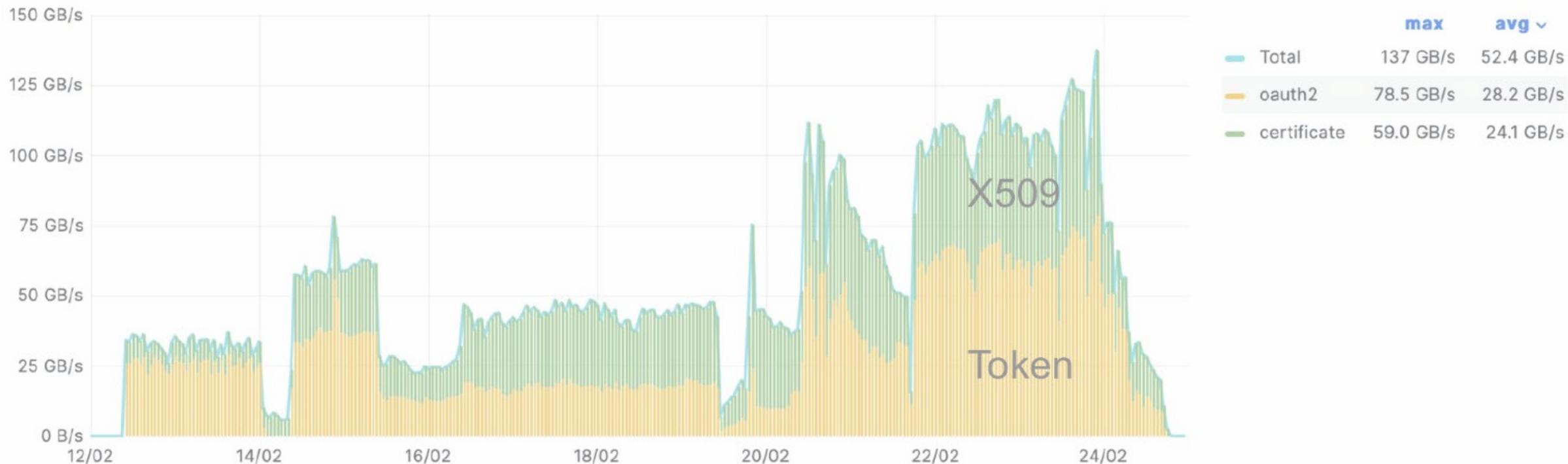
SRC = IN2P3



Token usage in DC24

- 19 sites (T1 + quelques T2) étaient « token »-compatibles au démarrage du DC24
- 6 ont été ajoutés la seconde semaine
- Utilisation au final minimale, aucun stress/problème détecté

Transfer Throughput

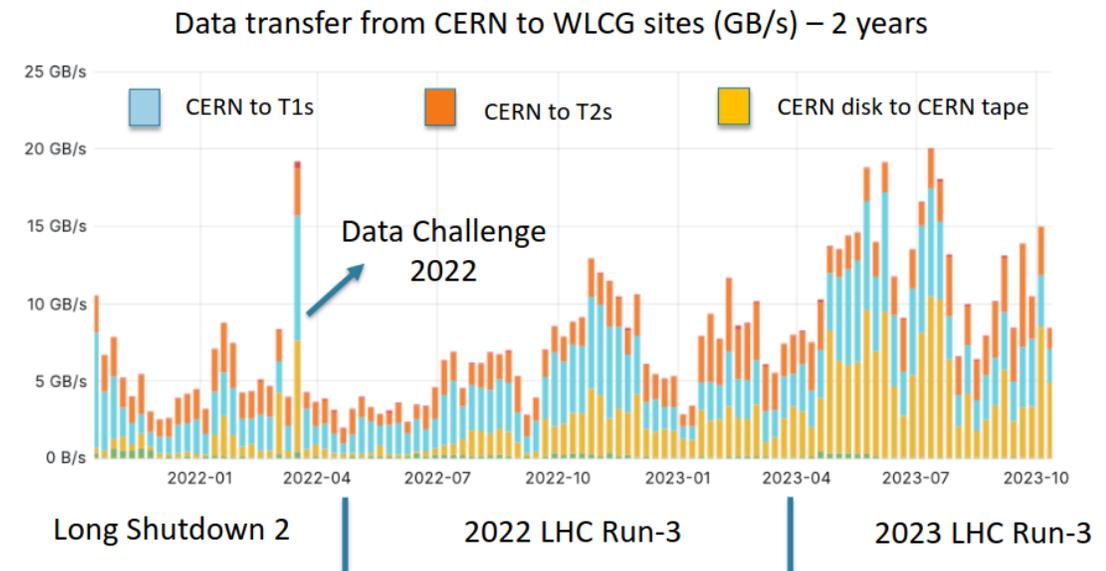


- Les résultats du DC24 sont dans l'ensemble (très) satisfaisants du point de vue des Tier-1s.
 - Tous les sites ont atteints (ou presque) les valeurs cibles attendues
 - On est (encore) parmi ceux qui ont donné les meilleurs résultats
- Quelques soucis, essentiellement du côté du “computing central”
 - Côté FTS, avec des mises-à-jour à plusieurs reprises lors de l'exercice
 - 2 fois la seconde semaine du DC24 pour nous
 - Expiration des token pendant la première semaine
 - Côté Rucio (qui gère les données localement), qui n'effaçaient pas suffisamment vite, et certains stockages se sont trouvés remplis
 - Impact \pm important, mais visible sur ~tous les tier-1s
 - Côté CMS, on a surtout vu des limitations côtés outils, et non pas côté sites (Tier-1s)

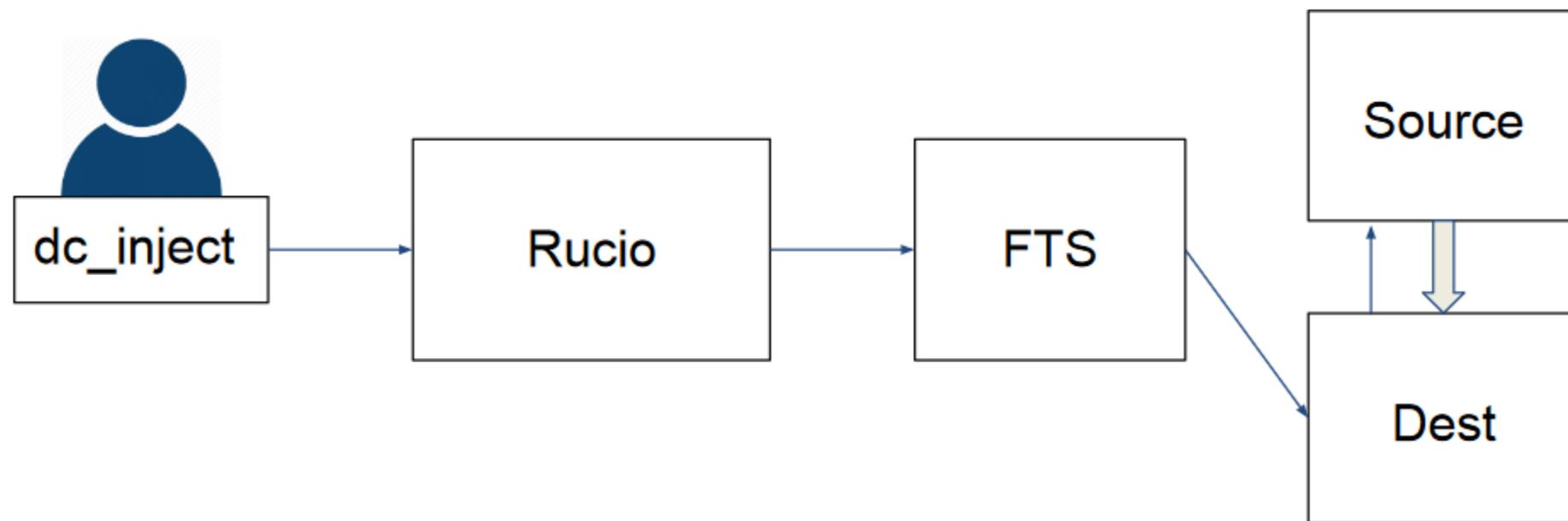
- **Data Challenge : DC24**

- 12 Février – 23 Février
- Objectifs :
 - Valider la montée en charge des capacités de transfert pour HL-LHC (et autres expériences) : 25 % du HL-LHC
 - Augmenter les activités réseaux et volumes échangés.
 - Valider les nouvelles technologies/solutions mises en places.
- Les DC 24 consistent essentiellement à superposer à l'activité normale (**QUI RESTERA EFFECTIVE DURANT LA PERIODE**) des transferts de vrai data selon les modèles de distribution de données de chaque expériences.

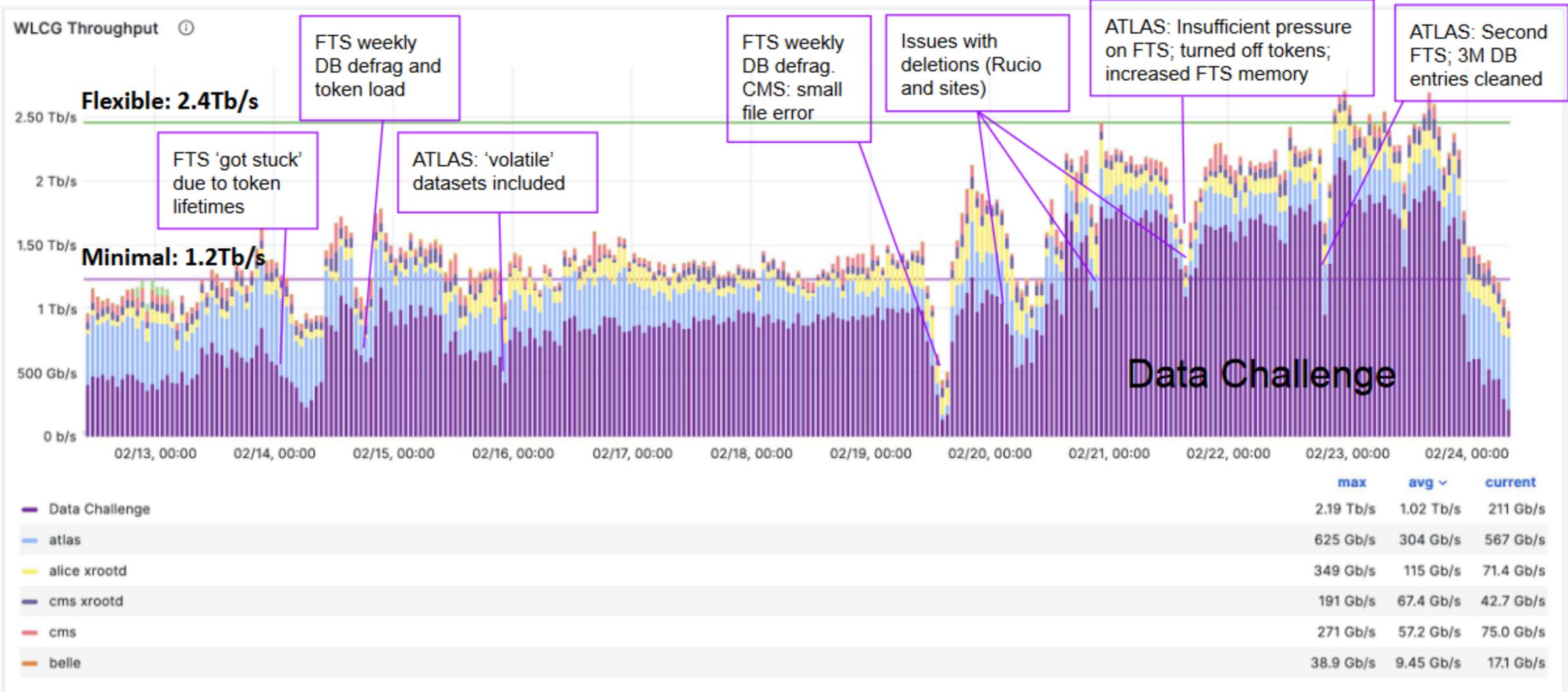
TO EXPORT



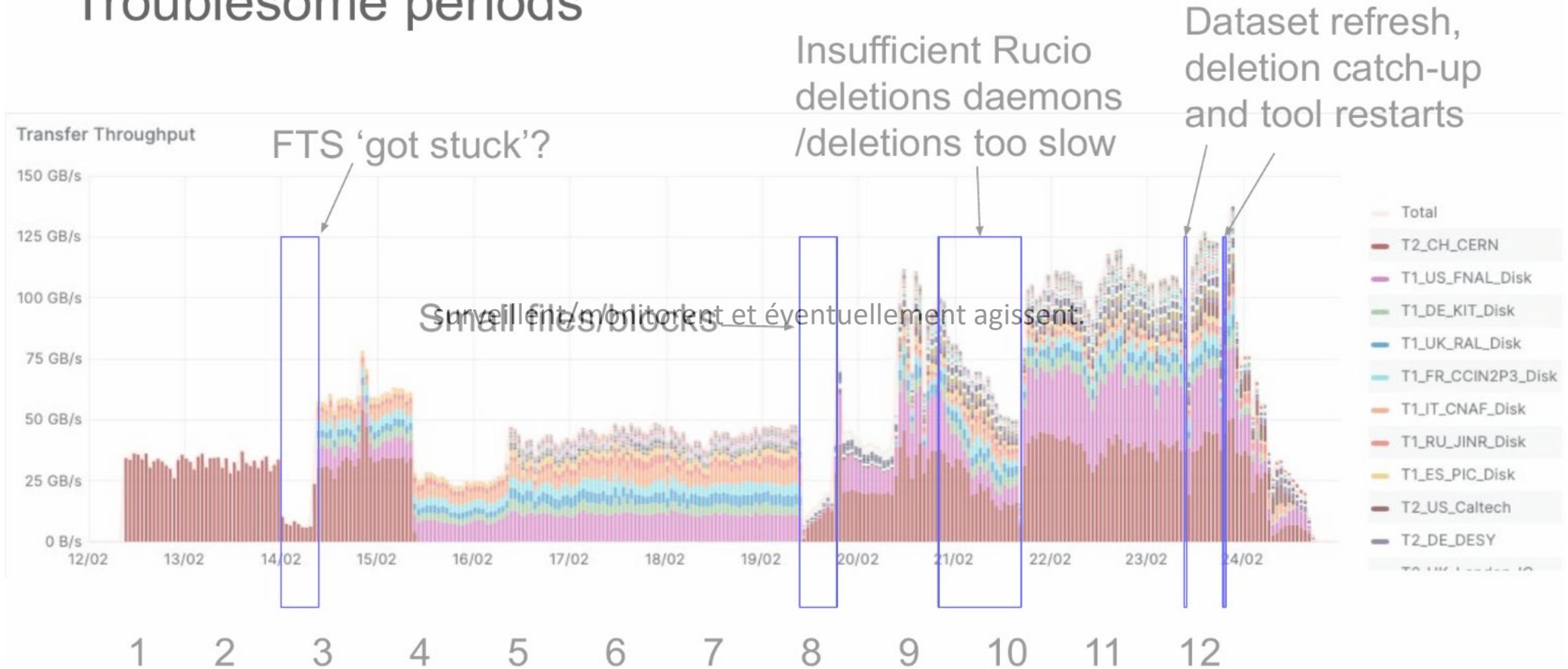
ATLAS and CMS workflow



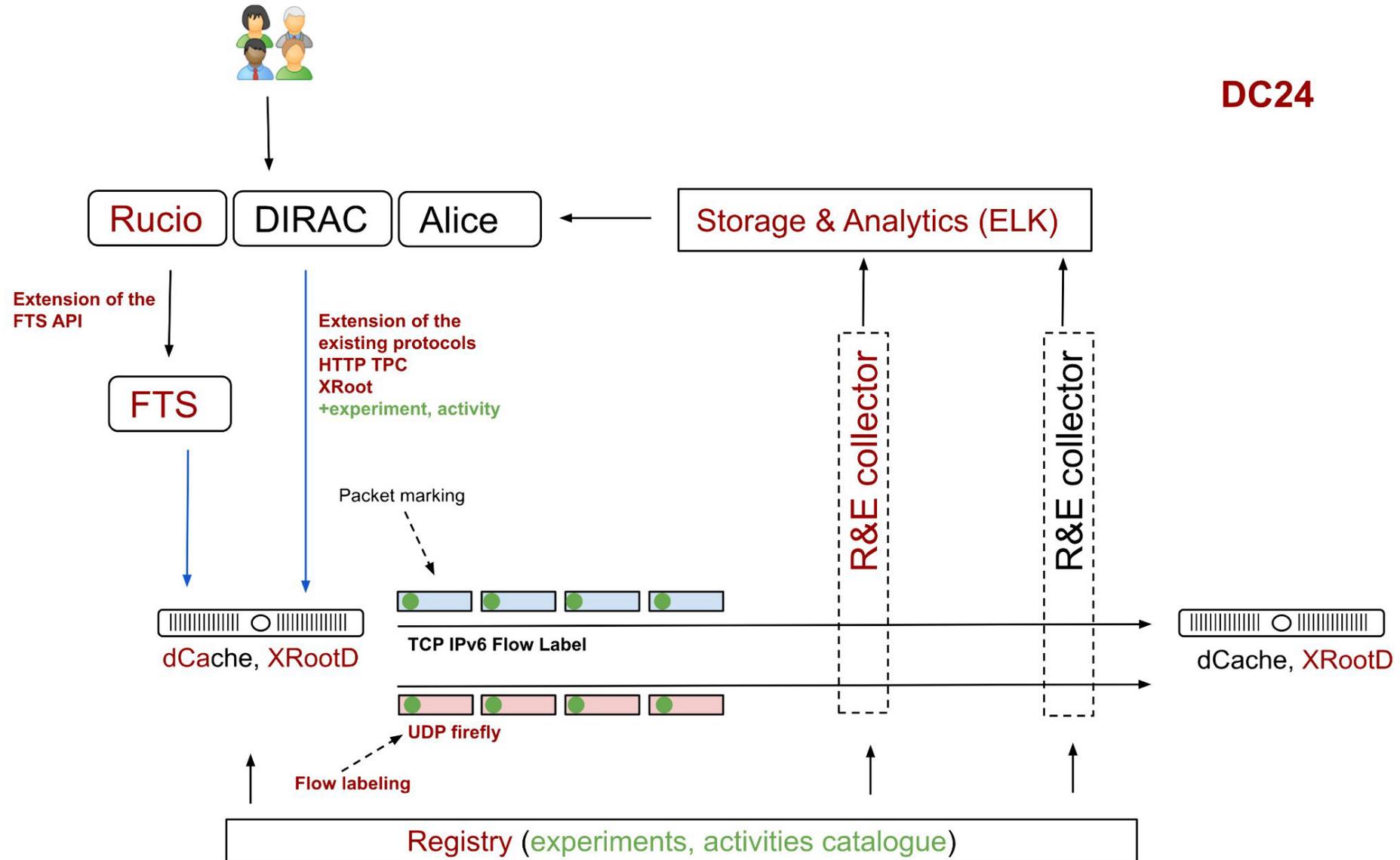
Notable issues



Troublesome periods



- **Le packet marking**
 - Être en capacité d'identifier au niveau réseau les différents flux (cf à quelle expérience ils appartiennent)
 - LHCOPN/LHCONE représentent 40% de tout le trafic Recherche Education
 - Deux façons de faire
 - Insertion de Scitag dans l'entête du paquet
 - IPV6 seulement
 - Méthode parfaite , on rate rien
 - Insertion de paquet UDP dans le flux de transfert : UDP Fireflies
 - Marche avec ipv4 et ipv6
 - Possible perte de paquet UDP et donc mesure « fausse »
 - Nécessite que les middleware soient compatibles



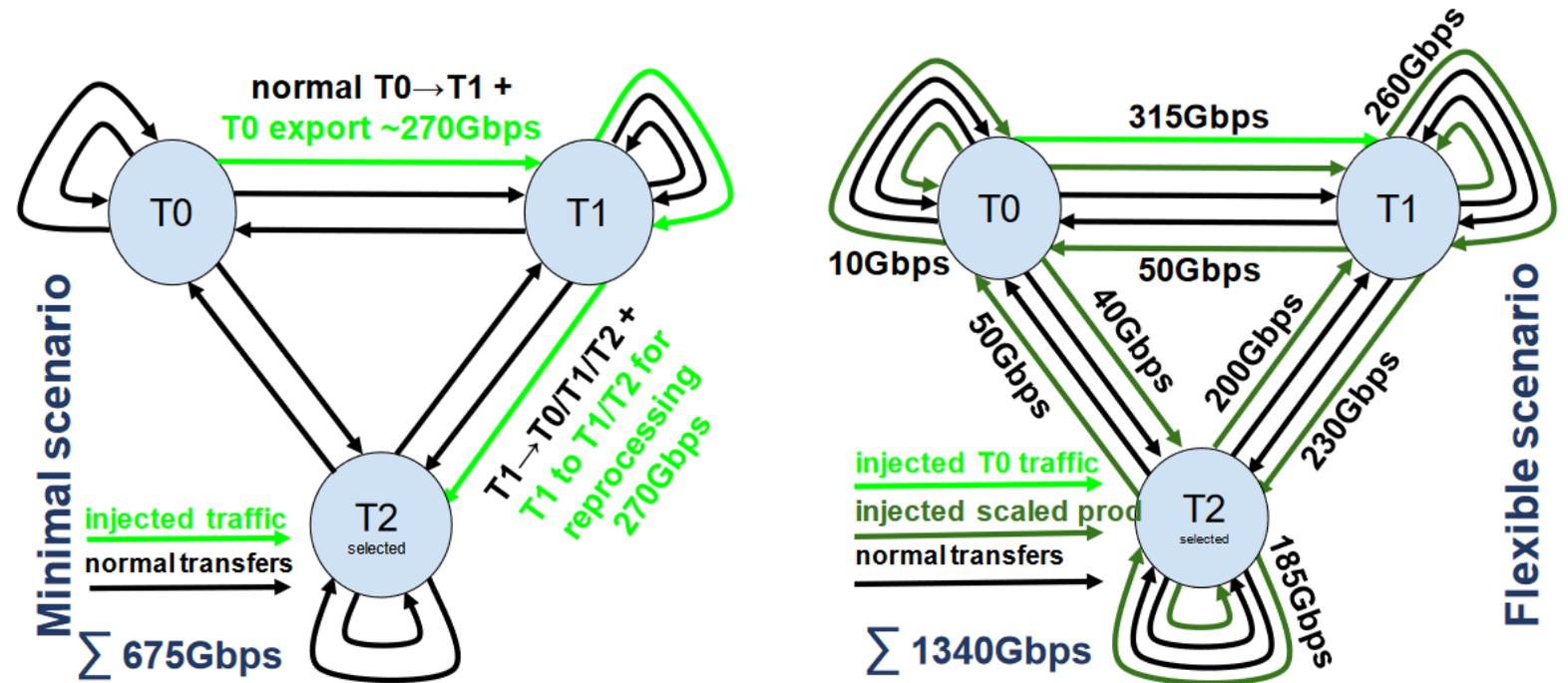
- **Alice**
 - Cas le plus simple, ils ont besoin de rien à part atteindre les objectifs chiffrés
 - Soit pour le CC : 3,2 Gb/s en ingress

Centre	Target rate GB/s	Achieved rate GB/s
CNAF	0.8	2 (250%)
IN2P3	0.4	0.8 (200%)
KISTI	0.2	1 (500%)
GridKA	0.6	2 (300%)
NDGF	0.3	0.4 (133%)
NL-T1	0.1	0.9 (900%)
RRG-KI	0.4	0.53 (128%)
RAL	0.1	0.7 (700%)
<i>CERN</i>	<i>10</i>	<i>20 (200%)</i>

Target 2.5GB/s (T1s) + 10GB/s (T0)

- **ATLAS**

- T0, T1 et T2 seront sollicités
- Veulent se placer dans des conditions « réalistes » donc le modèle flexible.



- **CMS**
 - Pour le CC : 45 + 36 Gb/s
 - **Des pré test et mini challenges seront mis en œuvre entre novembre et février**
 - En octobre on a déjà eu un mini test qui a occupé pour le CC 12,5 Gbps du lien lhcone

Tier 1 total rate goals (all 4 scenarios summed)

RSE	Ingress (GB/s)	Egress (GB/s)
T0_CH_CERN_Disk	0.000	31.250
T1_DE_KIT_Disk	5.588	4.472
T1_ES_PIC_Disk	2.333	1.887
T1_FR_CCIN2P3_Disk	5.576	4.494
T1_IT_CNAF_Disk	7.108	5.658
T1_RU_JINR_Disk	8.465	5.118
T1_UK_RAL_Disk	4.427	3.551
T1_US_FNAL_Disk	22.076	31.811

- **LHCb**
 - Pour le CC : 12Gb/s en ingress
 - Staging : 9,6 Gb/s
 - Si ils peuvent avoir les token c'est bien sinon fera sens

Site	shares	Data written (TB)	Export speed	Staging Speed (GB/s)	Staging duration (hours)
CERN		2117.00	14.00	2.57	48.39
CNAF	14.61%	309.36	2.05	1.60	53.77
GRIDKA	19.56%	414.01	2.74	1.66	69.13
IN2P3	10.93%	231.38	1.53	1.20	53.77
NCBJ	7.30%	154.64	1.02	0.89	48.39
PIC	3.64%	77.13	0.51	0.40	53.77
RAL	28.26%	598.29	3.96	2.40	69.13
RRCKI	0.00%	0.00	0.00	0.00	0.00
SARA	8.24%	174.40	1.15	0.80	60.49
Beijing	7.45%	157.79	1.04	0.63	69.13
Total Tier1s	100.00%	2117.00	14.00	9.58	

- **Belle 2**

- Les DC 24 porteront sur le trafic KEK vers T1s
- Va permettre de stresser leurs infras et services (Rucio,FTS,SE,Network)
- Pour cela ils vont appliquer un facteur 5 à leur besoin qui est de 40TB par jour (3,7Gb/s)

- Souhaite si possible tester le packet marking

- Souhaite si possible faire des transferts site à site notamment parce qu'un certain nombre de leurs sites sont sur LHCOPN

- **Belle 2**

- Pour le CC : 2,8 Gb/s en ingress
- Des pre tests dès novembre

Site	Country	#5G Files	Replica Factor	Total TB	Ingress (Gb/s)	Egress (Gb/s)
KEK	JP	8000	5,0	200	0,0	18,5
BNL	US	2400	5,0	60	5,6	0
CNAF	IT	1600	5,0	40	3,7	0
DESY	DE	800	5,0	20	1,9	0
KIT	DE	800	5,0	20	1,9	0
IN2P3CC	FR	1200	5,0	30	2,8	0
UIVc	CA	1200	5,0	30	2,8	0
Napoli	IT	TBD	TBD	TBD	TBD	TBD
SIGNET	SL	TBD	TBD	TBD	TBD	TBD

- **DUNE**

- Le CC n'est pas impliqué (DB tu confirmes)

Code couleur des réseaux à prendre avec précaution.

Considérer aussi le staging LHCb.

Considérer que contrairement au DC22, les tests seront fait de façon concomitante.

Bien sur cela suppose que les infra (Rucio, FTS, serveurs de tokens ...) suivent les débits

ATTENTION : Pas mal de pré test (pas de performance j'espère) entre novembre et Février

	Ingress (Gb/s)	Exgress (Gb/s)
ALICE	3,2	0
ATLAS	96	66
CMS	45	36
LHCb	12	0
BELLE 2	2,8	??
SOMME	159	102

Sur LHCOPN, sur LHCGNE, sur LHCGNE/LHCOPN