



Météo-France

Sylvain REPESSE

Agenda

- Pour la période 2014 – 2018 et quelques Statistiques.
- Projet renouvellement de notre système de stockage 2021 – 2025
- Les questions ??????

Key features

	2014	2015	2016	2017	2018
Usefull capacity (PB)	27	43	72	120	180
Number of files (Mfil)	266	345	479	745	1 130 /500
Double copy (TB)	200	240	280	465	798/1200
Cache disk (PB)	3	6	12	14	17
Max put / day (TB)	173	333	665	732	825/330
Max get / day (TB)	146	280	532	599	665
Get from disk * / 10 s	200	332	665	731	800
Get from tape * / 600 s	146	173	439	486	519

* : files of 100Mo

Architecture

	2014	2015	2016	2017	2018
FTP/NFS Servers	2	3	6	6	7
DDM Servers	3	6	6	8	8
Netapp E5500 (Datas)	3	5	8	9	10
TS 3500 HA	1	2	2	2	2
TS 4500 HA			1	2	2
TDM Servers	4	6	10	10	12
TS 1150	51	60	140	156	140
TS 1160					24
JD	3000	4880	8240	13480	17480

Statistiques quotidiennes

- 33 Tio remontés depuis les bandes tous les jours (Mars)
- 7500 montages de bandes pour les Gets (hors repack et migration)
- 28000 Fichiers remontés par jours (10%)

Stockage hiérarchisé : Bilan année 2018



**METEO
FRANCE**

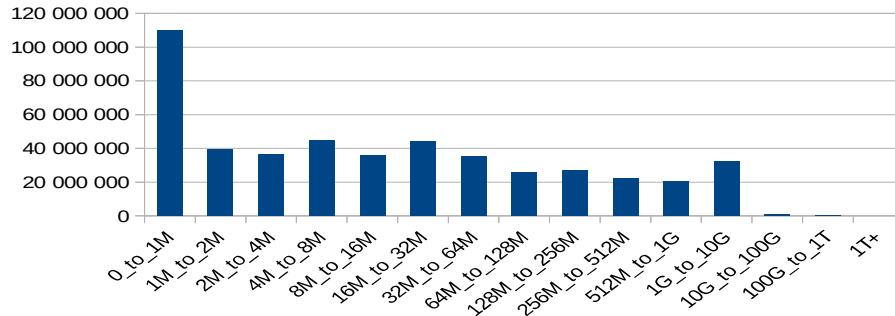
- Volume stocké le 31/12/2018 : 116 Pio pour 465 M fichiers
- Evolution durant l'année 2018 :
 - Croissance : + 32 Pio pour +100 Mfic, (+38 Pio pour +100 Mfic en 2017)
 - Soit 38 % d'augmentation sur l'année (82 % en 2017)
 - Ecritures : 64 (46,4 Pio en 2017)
 - Lectures : 23 Pio, soit 36 % des écritures (36 % en 2017)
 - Suppressions : 32 Pio soit 50 % des écritures (13,8 % en 2017)

Distribution Nombres et Volumes Fichiers



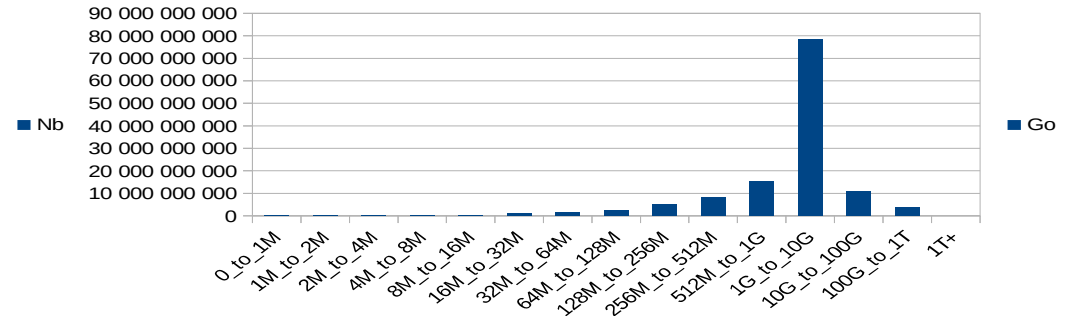
Distribution Nombre Fichiers Total

475 M



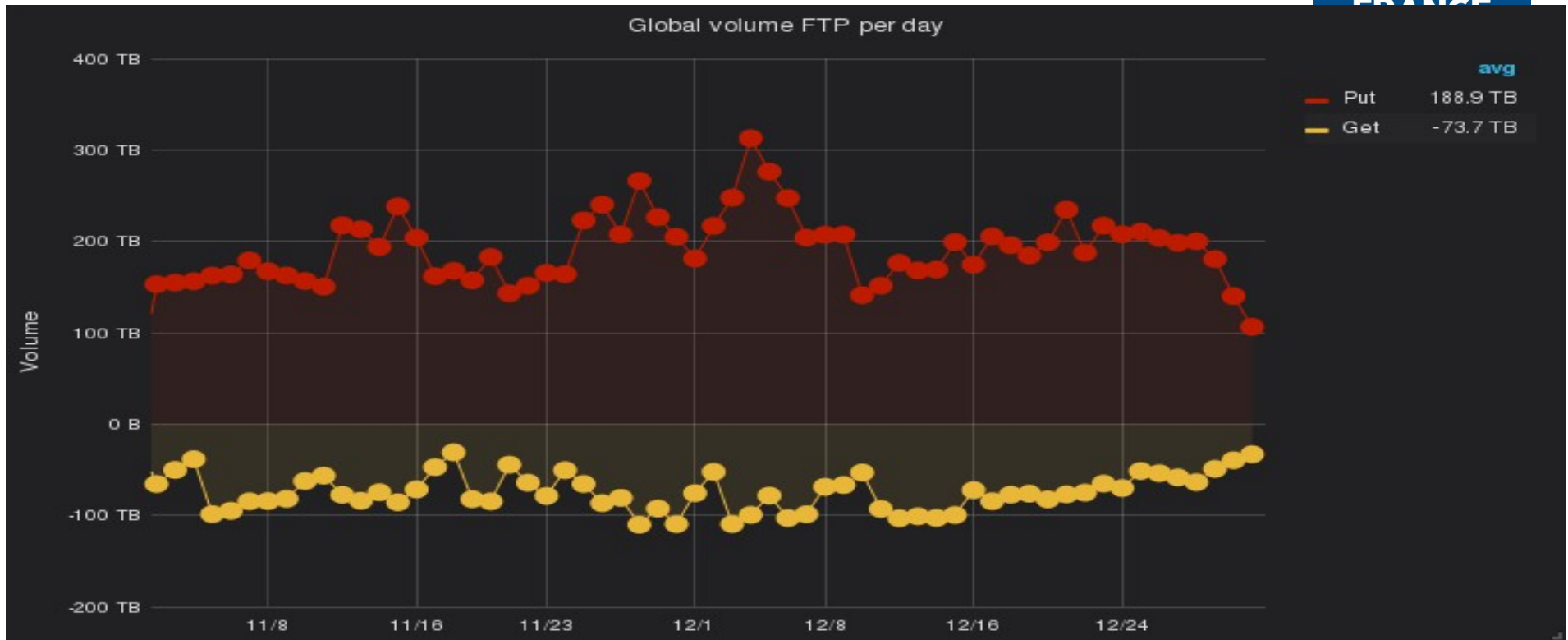
Distribution Volume Fichiers Total

118 Pio



- 84 % de fichiers de taille < 256 Mo pour 9 % du volume
- 7 % des fichiers de taille [1G-10G] pour 61 % du volume
- *Taille moyenne le 31/12/2018 : environ 267 Mo*

Flux FTP Novembre-décembre 2018



Projet 2021-2025

- Gros travail pour identifier l'utilisation faite des fichiers.
 - Statistiques sortie depuis 2 tables :
 - HPSS.BITFILE
 - HPSS.NSOBJECT
 - Plus de 500Million d'entrées dans ces deux tables
 - 5 jours de requêtes sur db2 (hors prod sur un serveur de restauration de notre base)
 - Recopie des données dans une table Postgres, avec optimisation (index, et index partiels) 8 minutes pour exécuter la même requête.
- 82 % des fichiers créés avant le premier Janvier 2018 n'ont jamais été relus.
- nettoyage/suppression en 2019 – 2020 au moins 29 Pio
- Modification des durées de rétention des données par service.

Les questions qu'on se pose

- Qrep et la réplication en temps réel.
 - Un arrêt de production est-il nécessaire ?
 - Quel en est la raison ?
 - Combien de temps ?
 - Quels sont ces checks fait par IBM ?
- Hardware
 - Quel retour pouvez vous nous faire sur les librairies spectralogic ?
 - Lecteurs, quelle technologie envisagez vous dans le futur ?
 - Quel taux de panne avez vous sur les lecteurs LTO ?
 - Quelles sont les performances en Lecture/ecriture sur LTO que vous obtenez ?
 - Quelle est votre taux de relecture du bande.