

Bilan du CCRC08, phase 1, pour ATLAS et le CC

-

Activités présentes et futures

Catherine Biscarat (CC-IN2P3)
Réunion LCG-France Tier-1 & Analysis Facility,
au CC-IN2P3, à Lyon, le 13/03/2008

Merci à S. Jézéquel, E. Lançon et G. Rahal
pour leur aide à la préparation de cette présentation

RECAP DU PLAN INITIAL DE ATLAS

Présentés par CB le 13 février dans cette réunion

CCRC08-01 dure 4 semaines (4-29 février)

- **Semaine #1 : test T0 + test SRM v2.2 (+ transfert SRMv2 vers T1 et T2)**
- **Semaine #2-4: transfert vers T1 et T2 (& nettoyage) + MC prod. (FDR2)**
- **2 T1s: analyse par les groupes de physique aux T1 (DPD) (sem. #3 et #4)**
- **2 T2s: analyse par les end-users sur les T1s et les T2s (sem. #3 et #4)**
- **2 T1s: Reprocessing des données muons cosmiques « M5 » (semaine #4)**

A la suite de quoi

- **Semaines #5 (mars) : prise de données cosmiques M6 + MC prod.**
- **Semaine #7 : tests de fonctionnalités + MC prod. (-> semaine #10)**

ATLAS

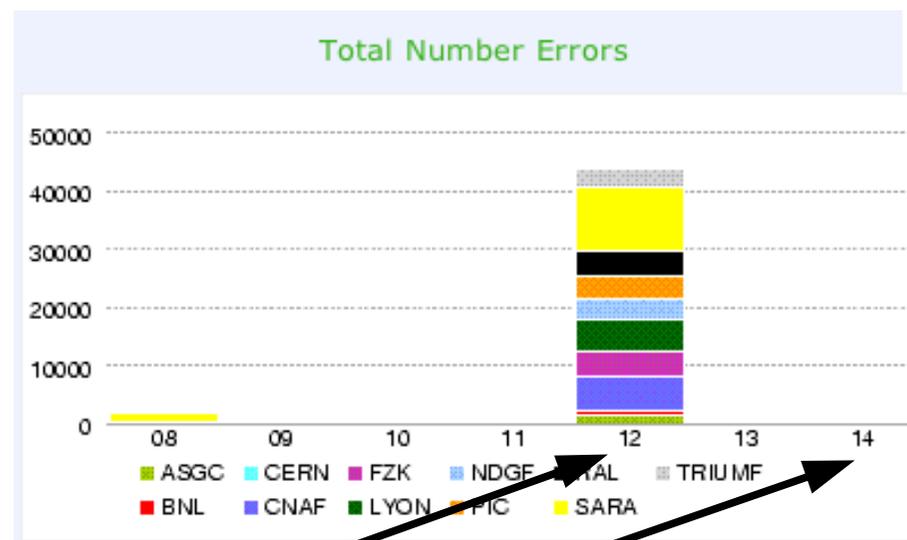
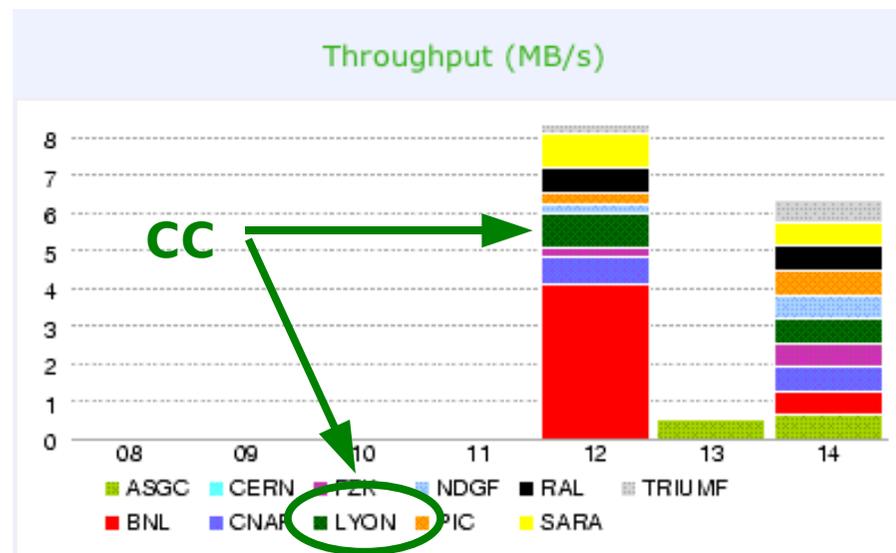
- **Mise en place et test** de l'infrastructure
 - Nouvelles zones de stockage avec space token (droits en ecriture et lecture)
 - Délégation de proxy dans FTS
 - Test des end-points SRMv2.2
- **Aucun export** vers les T1
 - L'optention du dataset prévu tarde

CC

- difficultés avec les space token;
- Mais les **transferts ont marché à temps** (jeudi 7 février);
- interfacage de SRMv2/dCache délicat :
 - T2-FR dCache : problèmes d'interfacage (BEIJING)
 - T2-FR DPM : prêts pour recevoir les fichiers du CERN, sauf RO-07-NIPNE qui a prit un peu de retard

CCRC08 : SEMAINE #2 : transferts T0->T1

- **T0 -> T1s -> T2s** (données FDR)
- **Peu de stress**
 - Petit dataset de RAW (< 2TB)
 - Petit fichiers (~100 MB)
 - Partage des données selon le modèle (share RAW+share ESD + 100% AOD)



Mauvaise configuration : déclenche migration pool à pool dans CASTOR, finalement corrigée

CCRC08 : SEMAINE #2 : transferts T1->T2

T1->T2

AOD Replication to ATLAS Tiers (Feb 08. Full Dress Rehearsal)

ASGC	AU-ATLAS	TW-FIT															
BNL	AGLT2	BU	IU	OU	SMU	SWT2	SLAC	UMICH	WISC	UC							
CNAF	LNF	MILANO	NAPOLI	ROMA1													
FZK	CSCS	CYF	DESY-HH	DESY-ZN	FREIBURG	FZU	HEPHY-UIBK	LRZ	WUP								
LYON	BEIJING	CPPM	LAL	LAPP	LPC	LPNE	NIPNE_02	NIPNE_07	SACLAY	TOKYO							
NDGF	IBT2																
PIC	IFAE	IFIC	LIP-COIMBRA	LIP-LISBON	UAM												
RAL	BHAM	BRUN	CAM	DUR	EDINBURGH	ECDF	GLASGOW	LANCS	LIV	MANC	OXF	SHEF	QMUL	ICL	RALPP		
SARA	IHEP	ITEP	JINR	NIKHEF	PNPI	SINP											
IUMF	ALBERTA	MCGILL	SFU	TORONTO	UVIC												

arrêt réseau

A. Klimentov,
ATLAS SW week,
Feb' 08

done part failed No test

10 Tier-1s and 56 "Tier-2s"
Metrics for T1 success : 100% data transferred (from CERN, from Tier-1s and to Tier-2s)
Metrics for T2/T3 success : 95+% data transferred (transfer within cloud)
Metrics for cloud success : 75% of sites participated in the test and 75% passed the test

A.Klimentov

ATLAS SW week. Feb, 2008

11

CCRC08 : SEMAINES #3-4 : transferts T0->T1 ->T2

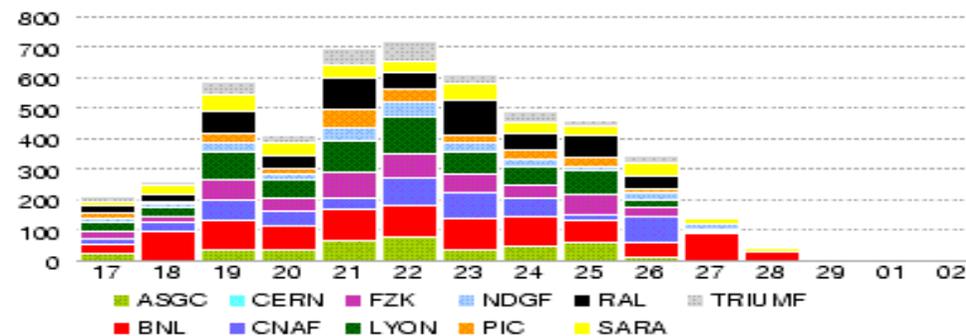
- Génération de données « random »
- taille de fichiers réaliste (≤ 3.6 GB)
- T0 throughput \sim cible 900 MB/s
- T0 -> T1 (17 fév. - 2 mars compris):

Cloud	Efficiency	Throughput
ASGC	39%	25 MB/s
BNL	57%	63 MB/s
CERN	0%	0 MB/s
CNAF	33%	37 MB/s
FZK	21%	35 MB/s
LYON	51%	45 MB/s
NDGF	52%	18 MB/s
PIC	65%	19 MB/s
RAL	61%	41 MB/s
SARA	15%	27 MB/s
TRIUMF	59%	21 MB/s

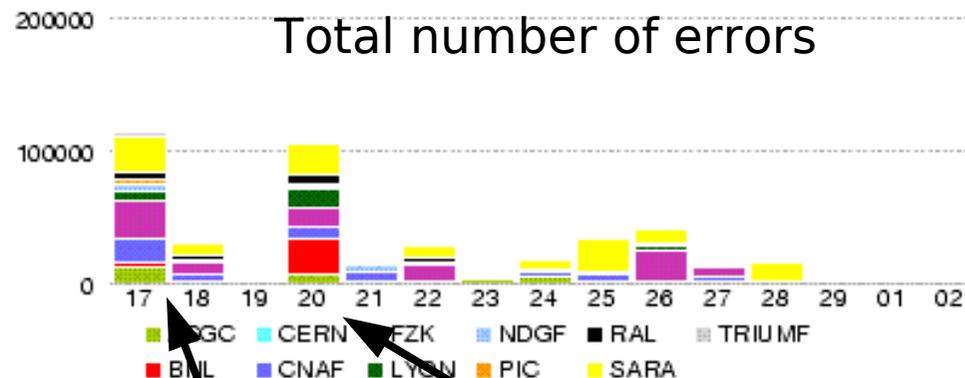
Eff: prend en compte essais ratés, ie 10 essais pour 1 fichier donne 10%.

C. Biscarat, Réunion LCG-France T1 & AF, Lyon, 13-Mars-08

Throughput (MB/s)



Total number of errors



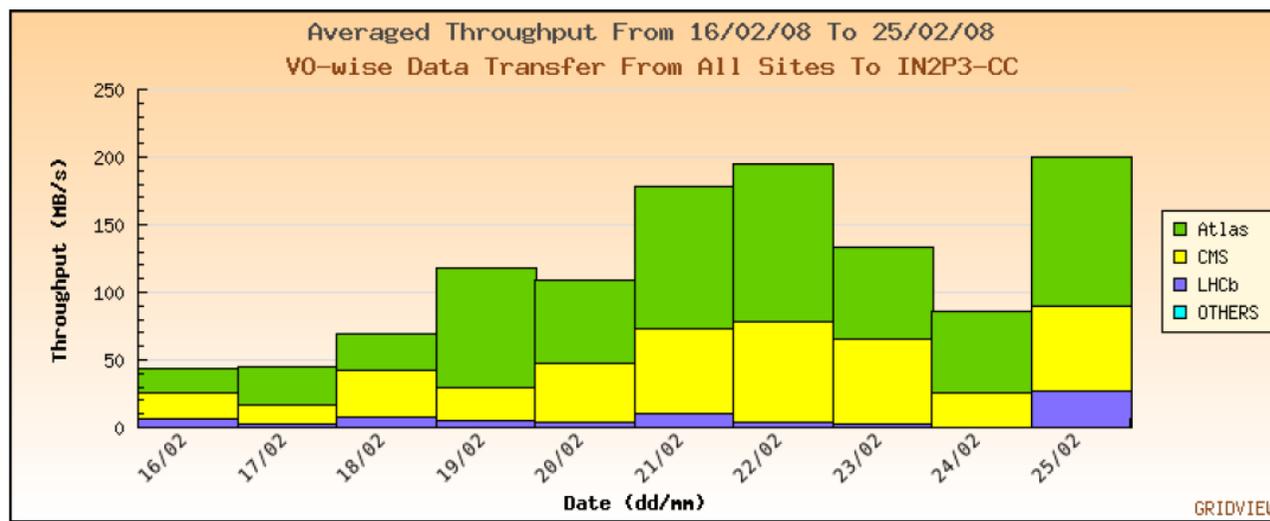
Problème de disk (hardware) sur CASTOR au CERN

Mismatch clé privée et publique dans FTS délégation de proxy

CCRC08 : SEMAINES #3-4 : transferts

ANYWHERE->CC

VO LHC en concurrence



- **vu par ATLAS** : « **Quite smooth. Occasional SRM frontend downtimes.** »
- Mode de stress, erreurs rares apparaissent :
 - **'File exists'** au premier essai de copie d'un fichier (2 origines)
 - Erreur dCache corrigée et déployée le 12 fév. au CC
 - Erreur DDM, corrigée pas déployée

T1->T2

- Réplication vers les T2 francais <nom>_datadisk

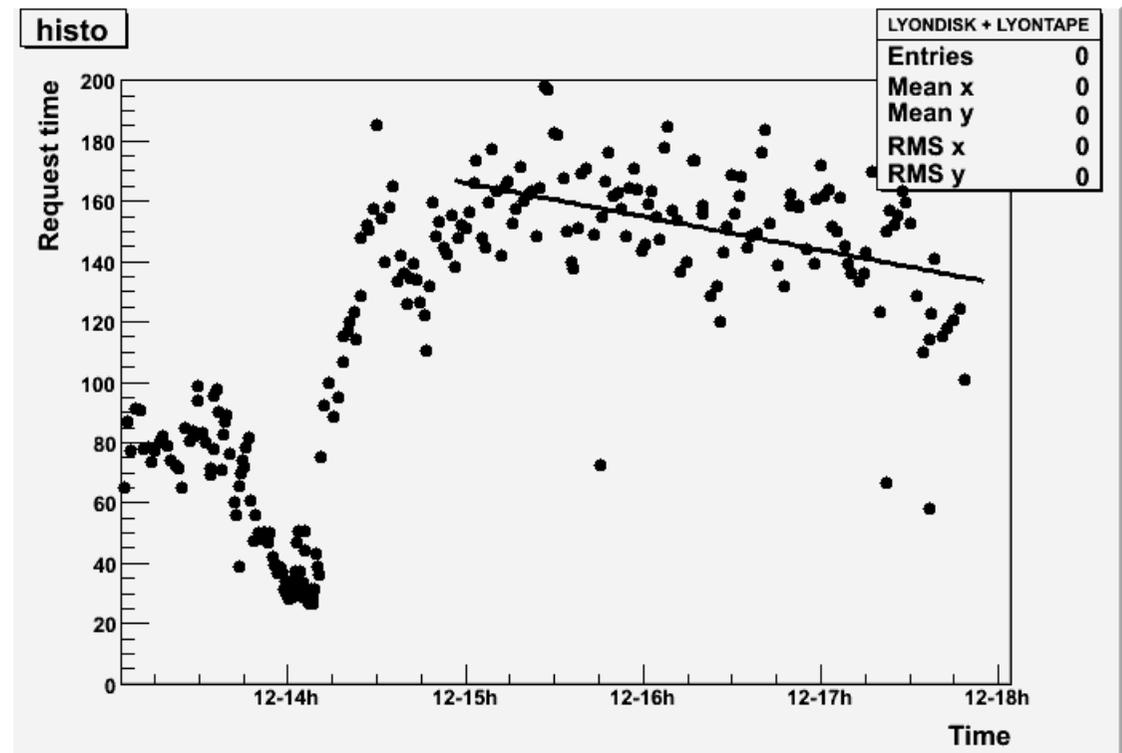
CCRC08 : SEMAINES #3-4 : transferts T1->T1

from/to	BNL	FZK	INFN	IN2P3	NDGF	NIKHEF	PIC	RAL	TW	TRIUMPH
BNL	X	100	100	20	100	25	15	20	100	100
FZK	X	X	100	30	100	40	50	30	100	100
INFN	100	90	X	0	100	25	30	X	100	100
IN2P3	X	100	100	X	100	35	20	30	100	100
NDGF	100	100	100	0	X	40	50	30	100	100
NIKHEF	0	0	0	0	0	X	0	0	0	0
PIC	95	100	95	3	100	15	X	5	90	95
RAL	100	100	100	10	0	30	10	10 (?)	100	100
TW	100	100	100	25	0	40	70	80	X	100
TRIUMPH	100	100	30	20	100	100	30	30	X (?)	100 (?)

- toutes les souscriptions sont faites à la fois (qq-soit état du SE, canal FTS)
- résultats **compliqués à analyser** (sauf les problèmes évidents)
- **CC export mieux que CC import**
- test répété maintenant T1 par T1, dans un environnement contrôlé

CCRC08 : SEMAINE #3-4 : effacement de données

- Nécessité d'**effacer massivement des données** régulièrement (MC obsolète)
- Outil **centralisé** (SRMv2), testé
- **Ralentissement de pnfs** observé au CC (bande et disque) cette semaine
 - Vitesse de 1,5 s/fichier
- Action transparente dans d'autres sites (FZK, Nikhef)
- dCache masters en contact avec développeurs ATLAS
- Explication possible :
 - effacement de fichier asynchrone (SRMv2 et pnfs)
 - Backlog dans pnfs



Temps de réponse d'une requête d'effacement (100 fichiers) en fonction du temps écoulé; V. Garonne et C. Serfon, le 12-mars-08

CCRC08 : QUELQUES REMARQUES

TRANSFERTS

- CCRC08 **retardé au T0**
- seulement **2 semaines de tests effectifs** aux tiers 1 et 2
- **programme augmenté** de nouveaux tests : **conjoint ATLAS+CMS**
 - *Voir la présentation suivante (Farida pour CMS)*
- **en passe d'être complété** : **reprocessing**

SIMULATION ECHANTILLON POUR FDR-2

- Cette activité **n'a pas eu lieu**
- aucune activité de simulation
- démarrage prévu bientôt

REPROCESSING : SEMAINE #4-#8 (3 mars - 2 avril)

- **Re-processing** = (1) descendre des données de HPSS, (2) les traiter avec de **nouvelles constantes de calibration**, (3) distribuer les outputs

Raison d'être principale des T1 dans ATLAS

- Exercée récemment dans ATLAS à BNL, TRIUMPH, NDGF
- Effort de ATLAS pour tester le reprocessing dans tous les T1, **avant le 2 avril**
 - **données de muons cosmiques** de octobre '07 (M5)
 - **mém. jusqu'à 2.7 GB** (échantillonnage fin des détecteurs, non standard)
 - **10 datasets de 250 fichiers** de ~2 GB
- CB et E. Lançon chapotent cette activité au CC

- Release installée : OK
- Modifications manuelles fichiers de calib. et release (pb release) : OK
- Test de qq jobs (problème dCache avec checksum time out): OK
- Mise en place d'une queue mappée sur classe J : en bonne voie
- Staging 'massif' (peu etre contrôlé en amont) : pas encore abordé

- **Nous sommes dans la course** (ni les premiers ni les derniers).

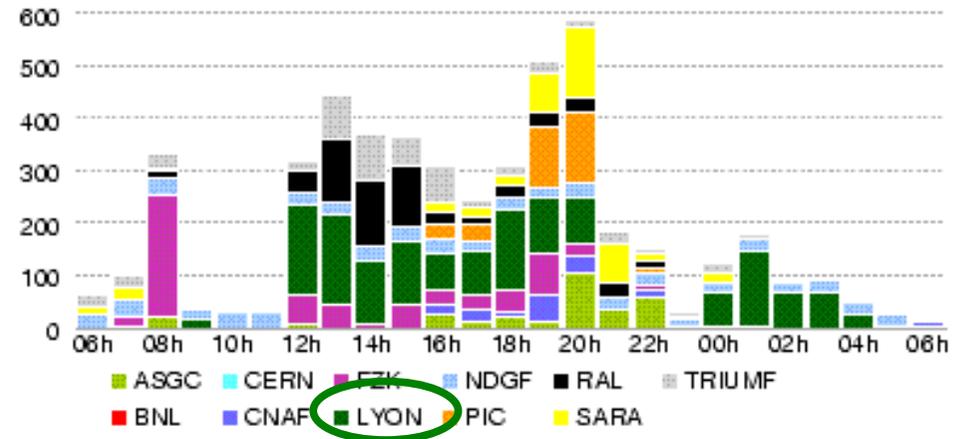
MUONS COSMIQUES M6 : SEMAINE #5 (3-10 mars)

- Total : 38 TB de bonnes données RAW (25k fichiers), ~4 TB ESD
- Distribuée depuis le vend. 7 fév., selon le share (T1-CC=15%)

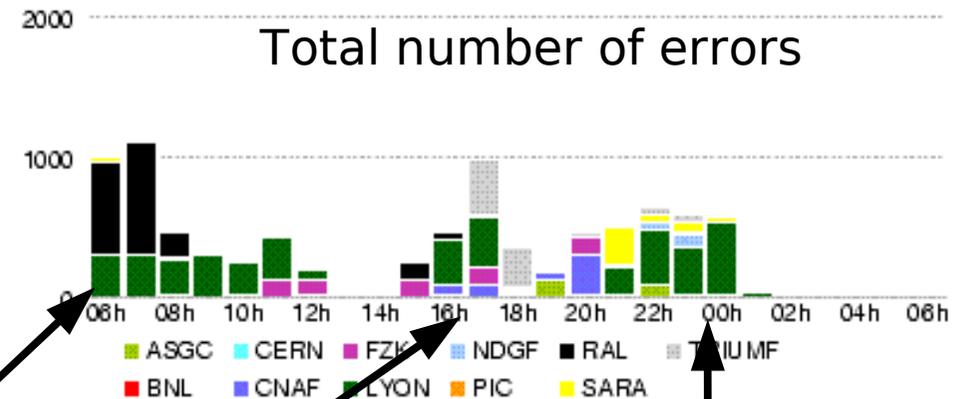
T0 -> CC

- Qq **problèmes dCache au CC**
- Mais aussi **problèmes au CERN**
- **Bilan globalement bon**
- **pics à 70-80 MB/s** sur qq heures

Throughput (MB/s)



Total number of errors



CC serveur
pnfs engorgé

CERN VO box
problématique

CC problème indéterminé
(aucune intervention)

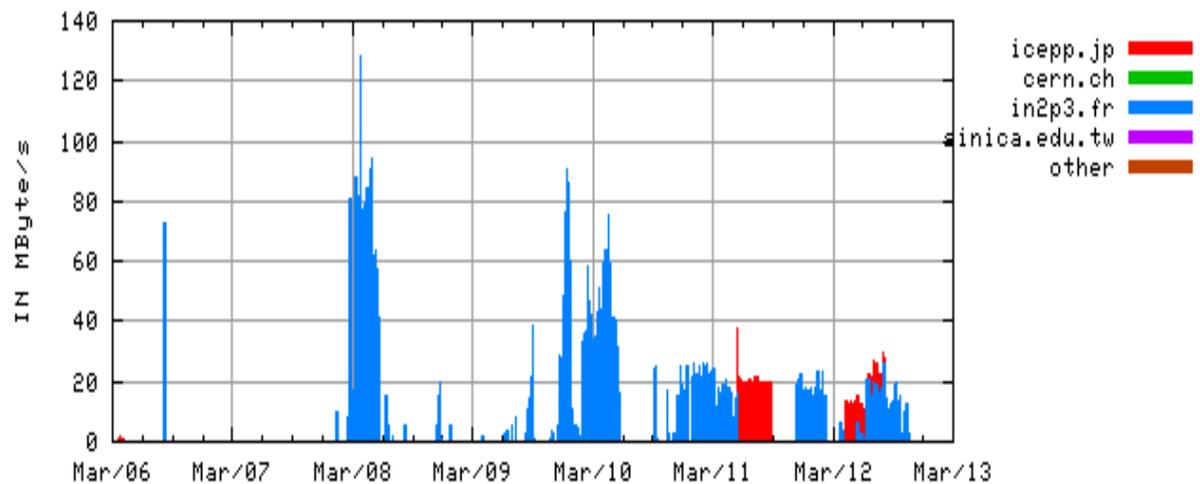
Dimanche-lundi (9-10 mars)

MUONS COSMIQUES : SEMAINE #5 (3-10 mars)

CC -> TOKYO (T2)

- Challenge dans le cadre de **recherche de financement**
- **Du vendredi au lundi matin** (6:30 heure FR, conférence de presse)
- **Très bon résultats de transferts** (~3 fois mieux que il y a un mois)
- Toujours qq problèmes de stabilité

- pics à ~ 80 MB/s
- Taux soutenus à 40 MB/s



Moniteur Tokyo, réf. Ueda

AGENDA POUR LES 3 MOIS A VENIR

FT means Functional Test, nominal rate, follow precisely data movements
TT means Throughput Tests, push to a maximum of 200% of the nominal rate
 T1-T1 tests will be done during FT weeks
 Clean-up during the weekend.

March 06 **Deadline for atldata pool to be ready**
Deadline for T0 – T2 for Michigan, Munic, Roma, Napoli
Deadline for definition of directories for tape families

March 03 – 07 **M6**
10 – 14 **FT** **T1 – T1 for Lyon and PIC**
17 – 20 **TT**

CB, en passe de commencer (downtime, effacement massif)

March 20 **Deadline for SRMv2@CERN**
21 – 24 **Easter**
25 – 28 **FT** **T1 – T1 for CNAF and SARA**

April 02 **Deadline for re-processing at all T1's of M5 data**
Deadline for decision on how to do file merging in the T0
Deadline for SRMv2@all T2's

April 31 – 04 **TT**
07 – 11 **FT** **T1 – T1 for RAL and NDGF**
14 – 16 **TT**
21 – 25 **FT** **T1 – T1 for BNL and FZK**
28 – 30 **TT**
01 – 04 **Ascension**

WLCG workshop
 ATLAS S&C

May 02 **Deadline for re-processing of FDR-1 data at all T1's**
Deadline for all T1 – T1 channels tested

May 05 – 09 **CCRC** **T1 – T1 ASGC**
13 – 16 **CCRC**
19 – 23 **CCRC** **Possibly also FDR-2**
26 – 30 **CCRC**

Plan discuté au sein de ATLAS, édité par K. Bos

CCRC08-2 encore indéterminé

CONCLUSIONS – REMARQUES FINALES

ATLAS

- mise en place de SRM v2 et des space token a bien marché
- CCRC08 écourté de part les problèmes rencontrés au CERN
- Transferts se sont bien passés
- Tests plus fins en cours : T1-T1

CC

- **Bonnes performances générales**
- **Certaines utilisations de dCache sont encore problématiques**
 - Différence de site à site, à comprendre
 - Réactivité des 'masters' reconnue et appréciée
- **Démarrage d'une activité fondamentale au CC : le reprocessing**
 - La partie staging est la partie restante à étudier
 - travail conjoint avec ATLAS (production centralisée) et le CC

BACKUP SLIDES

CCRC08 : SEMAINE #3 et #4 : effacement de données

- Premiers resultats de vitesse tres epars entre les T1:

CNAF-STORM 3240 files in 376 second : 8.6 Hz

NIKHEF-DPM 2840 files in 14661 second:) 0.2 Hz

TOKYO-DPM 1050 files in 562 second : 1.9 Hz

RAL-CASTOR 8502 files in 1163 second: 7.3 Hz

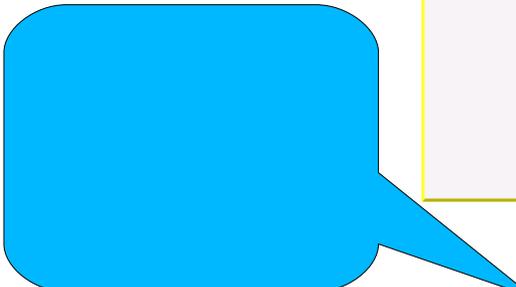
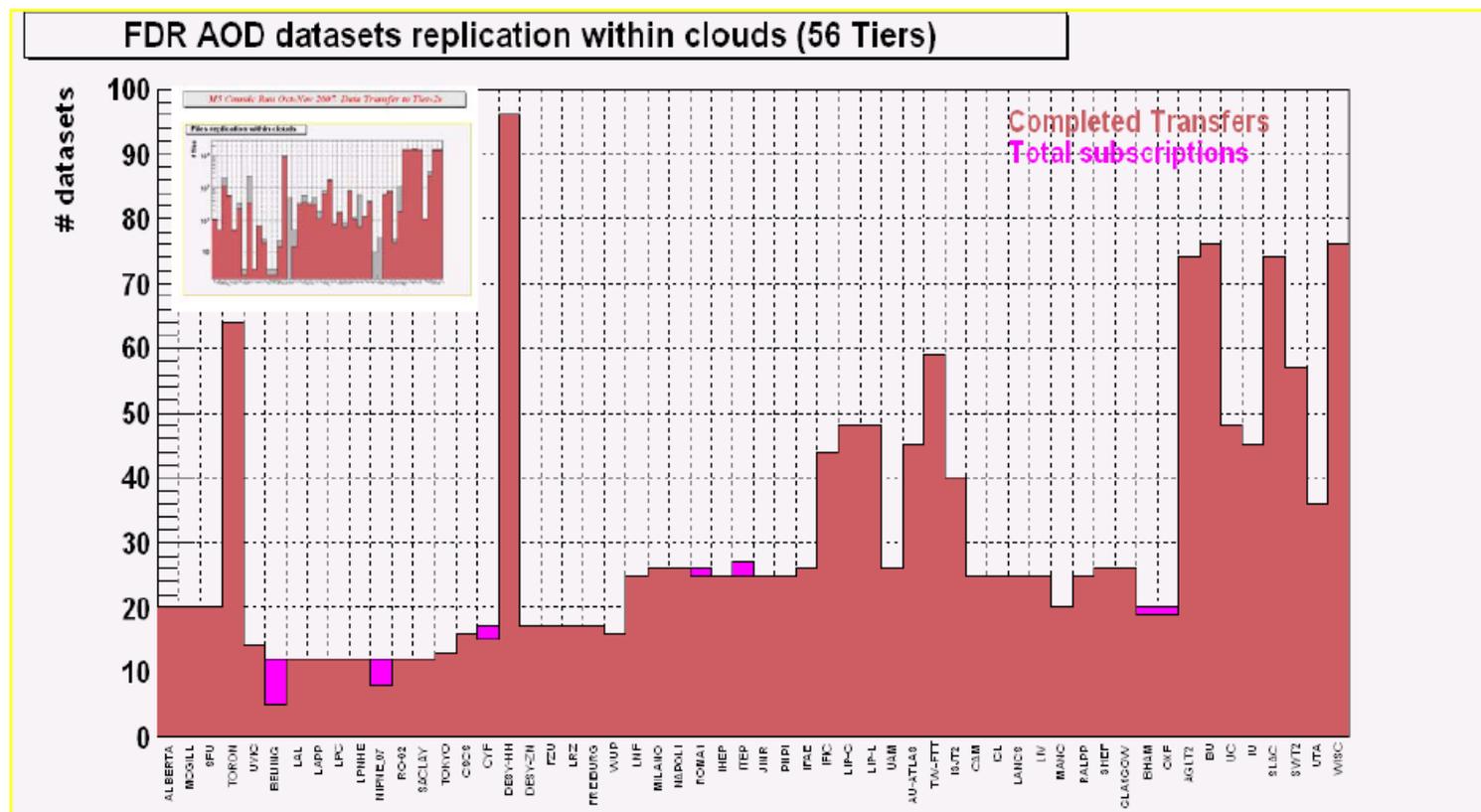
PIC-DCACHE 5523 files in 1894 second : 2.9 Hz

TRIUMF-DCACHE 6793 files in 2525 second : 2.7 Hz

CC-DCACHE 5242 files in 3222 second : 1.6 Hz

SEMAINE #2 (FRANCE)

FDR Feb 2008. Data Transfer to Tier-2s

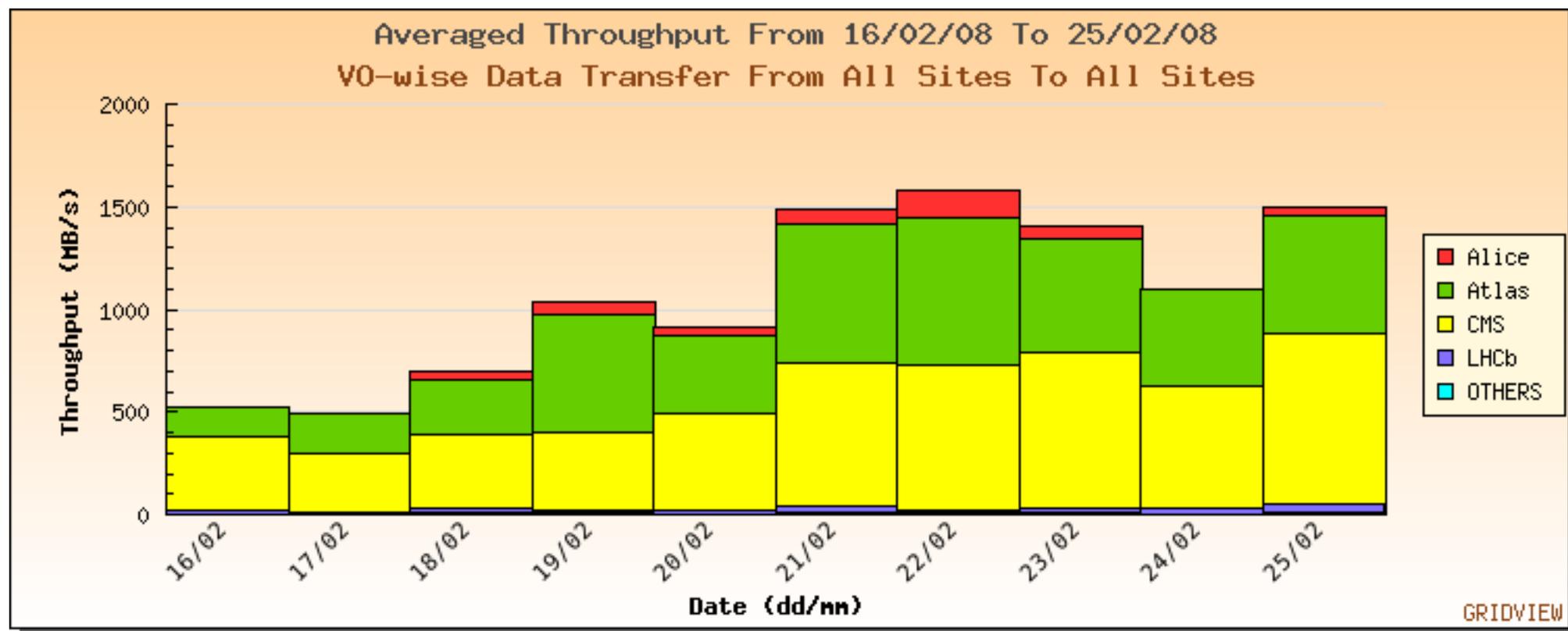


A.Klimentov

ATLAS SW week. Feb, 2008

10

CCRC08 : SEMAINES #3-4



TEST RECO ATLAS+CMS : SEMAINE #4 (24 fév - 2 mars)

- test de **reconstruction conjoint avec ATLAS et CMS**
- ATLAS en général plus demandeur de mémoire que CMS,
- But : exécuter **des jobs gourmands en mémoire** 1 contre 1 sur la meme machine, **mesurer la gène** (le swap)
- Organisation: **CC et PIC** participent; **Ghita coordonne (et analyse)**
- Run: 2 jours au CC; pas du tout a PIC
- Problème: peu de jobs, ATLAS a des petits jobs (mem. ~ 700 GB), période de recouvrement avec 4+4 jobs de seulement qq heures
- Conclusion :
 - test **délicat à mettre en oeuvre** (le premier en son genre)
 - la **production MC a du être stoppée** (FR et ES)
 - Mesures de mémoire consommée (<mémoire max de la machine, 16GB) et de CPU (< CPU disponible, 8 unités) **non concluantes**

MUONS COSMIQUES : SEMAINE #5 (3-10 mars)

ATLAS

Datasets	Total Files in datasets	Last Subscription	LFC Checked	Last Transfer
1002	41355	Mar 11 12:36:36	Mar 11 22:42:08	Mar 11 22:41:28

Tier1	Datasets	Total Files in datasets	Total CpFiles in datasets	Completed	Transfer	Subscribed
BNL	479	23100	13438	280	35	164
FZK	395	17892	18437	392	1	1
IN2P3	437	20354	21542	432	0	4
INFN	396	18321	18321	396	0	0
NDGF	377	19931	10238	158	8	211
NIKHEF	388	17368	17368	388	0	0
PIC	358	15985	15985	358	0	0
RAL	437	20097	20253	436	0	0
SARA	44	2529	2529	44	0	0
TAIWAN	373	17564	17564	373	0	0
TRIUMF	399	18932	18932	399	0	0

ATLAS

- Préparation de l'échantillon MC pour le Full Dress Rehearsal, phase 2 (test de stress de toute la chaîne, ie. SFO -> analyses de physique)
- à plus haute luminosité que FDR-1 (représentatif des premières collisions, 10^{32} et 10^{33} /cm²/s)
- génération pas encore commencé, elle devrait durer ~ 8 semaines.
 - EVGEN: T1 -> autres T1s
 - DIGI: T2 -> T1
 - RECO: T1 (20->1) -> BNL
- Maintenant on valide la release

CC

- Pics de jobs : validation

