



# Organisation des données

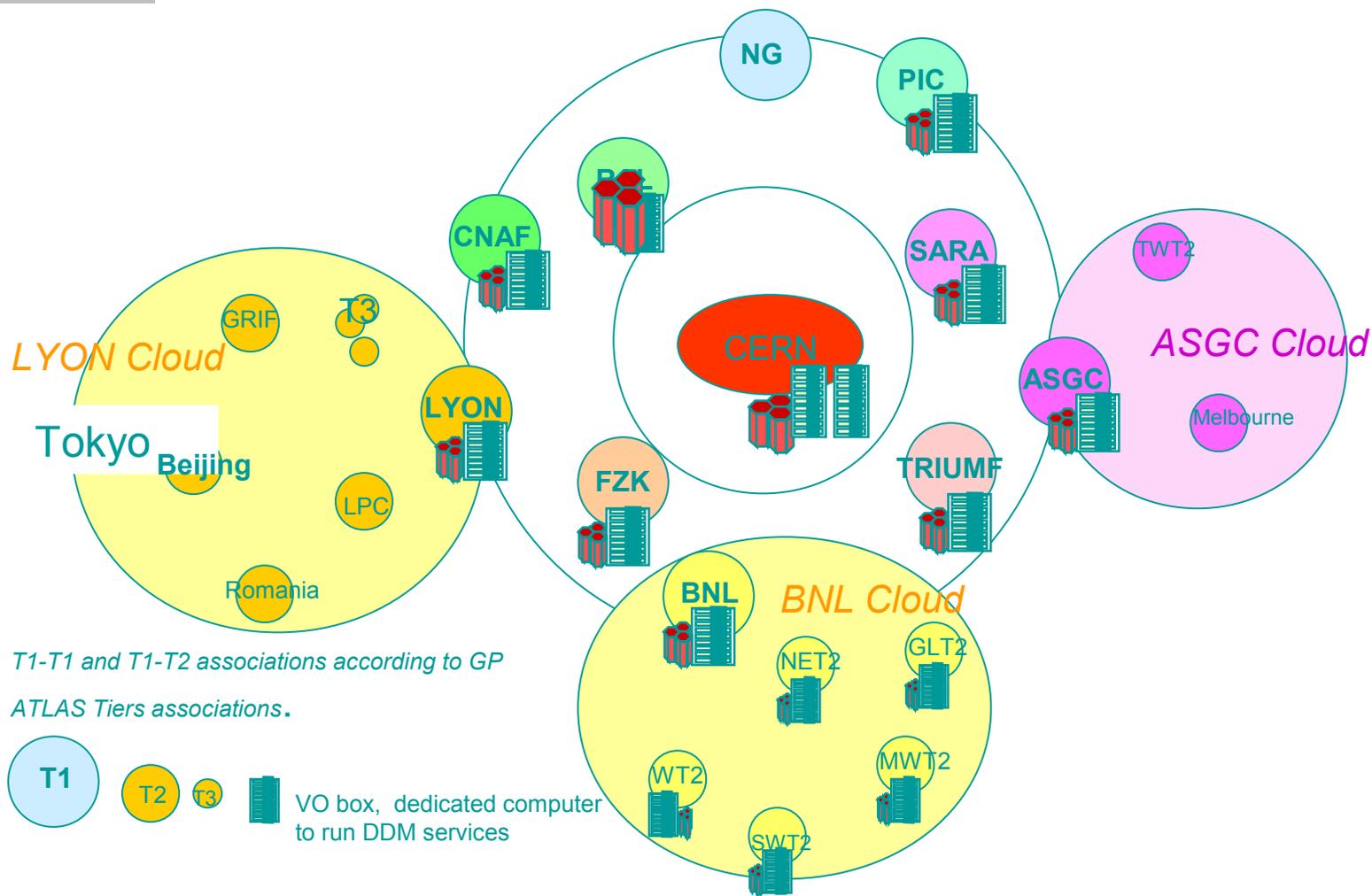
dans ATLAS

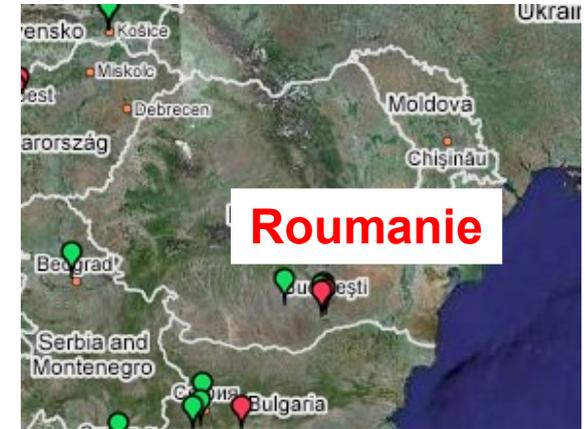
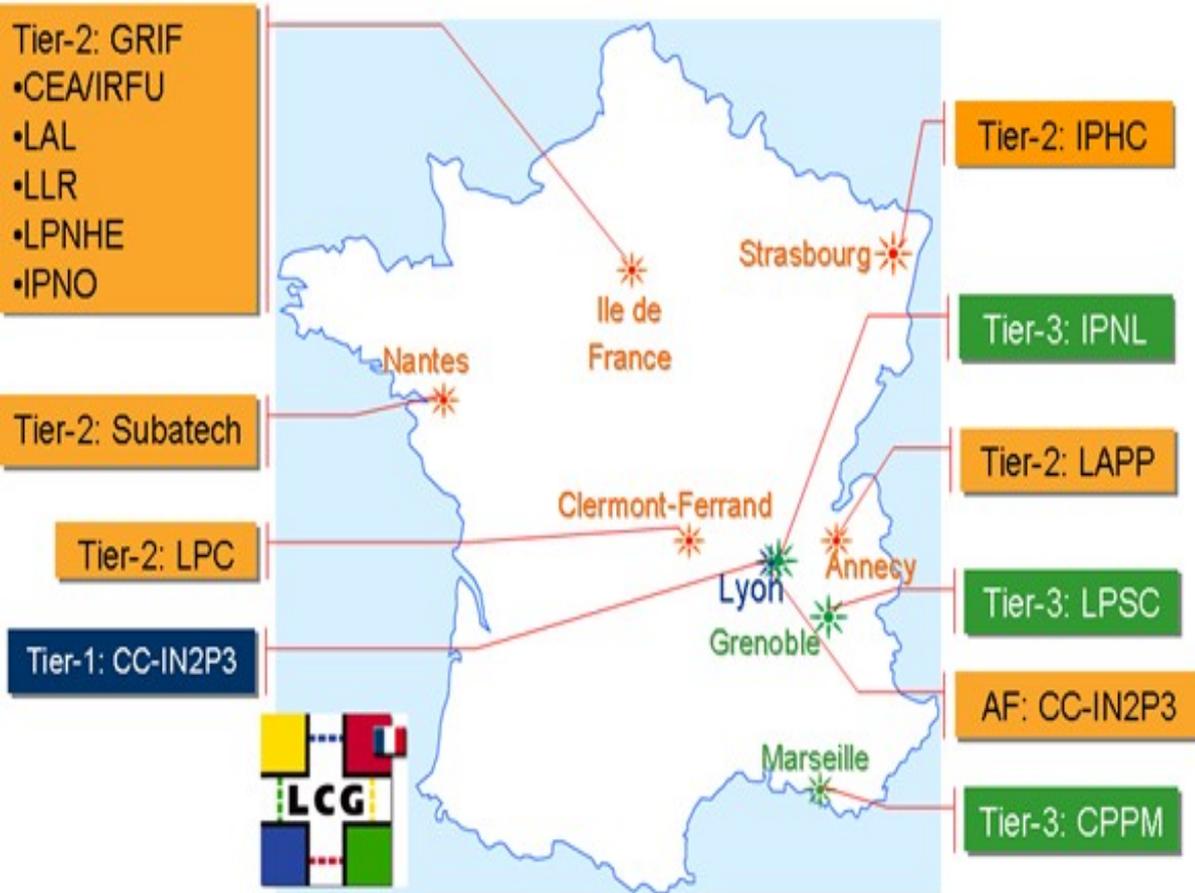
Etat des lieux

# Organisation ATLAS/Grille



“Modèle en nuage”



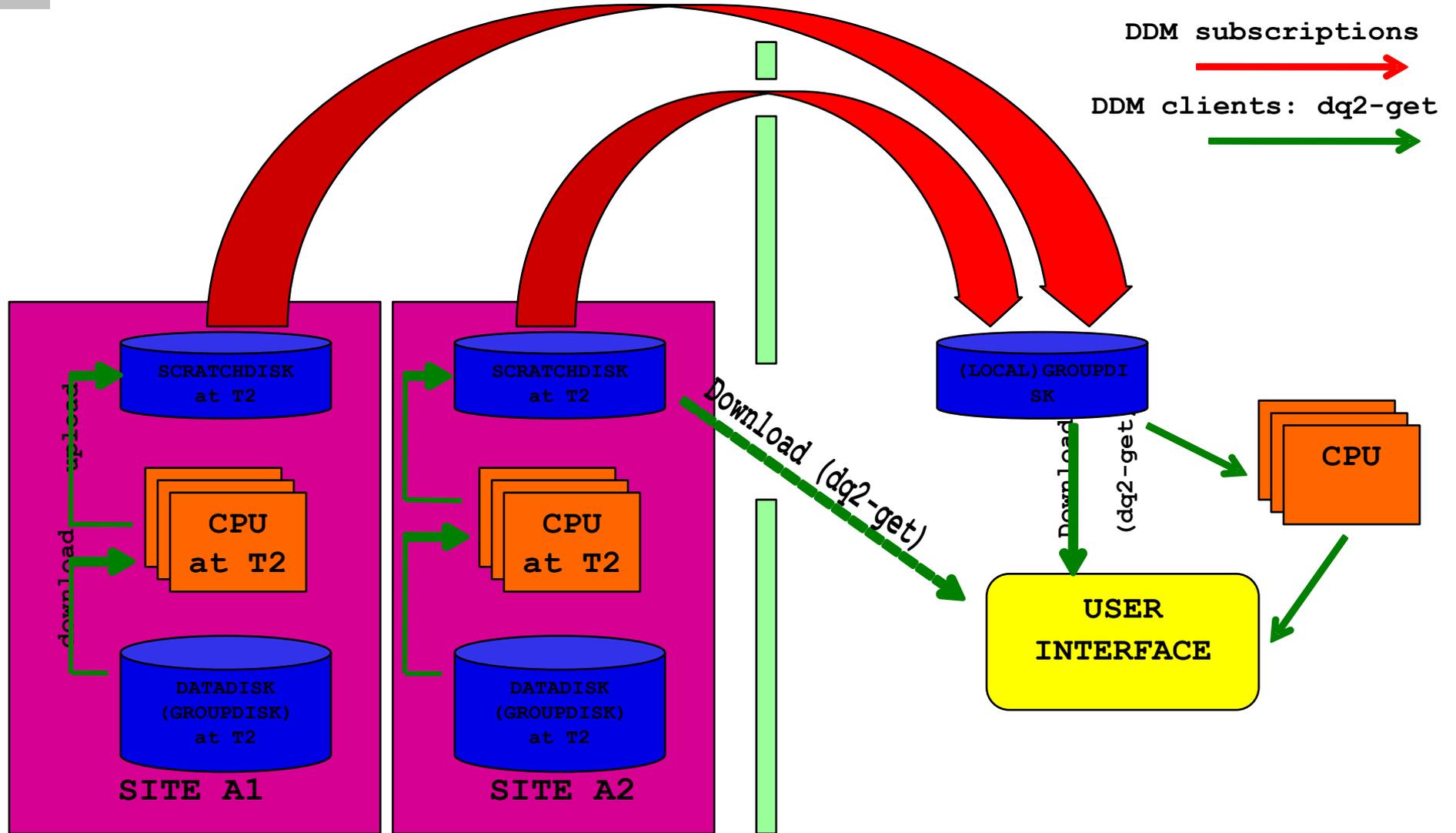


**Collaboration étroite avec Tokyo : I. Ueda**

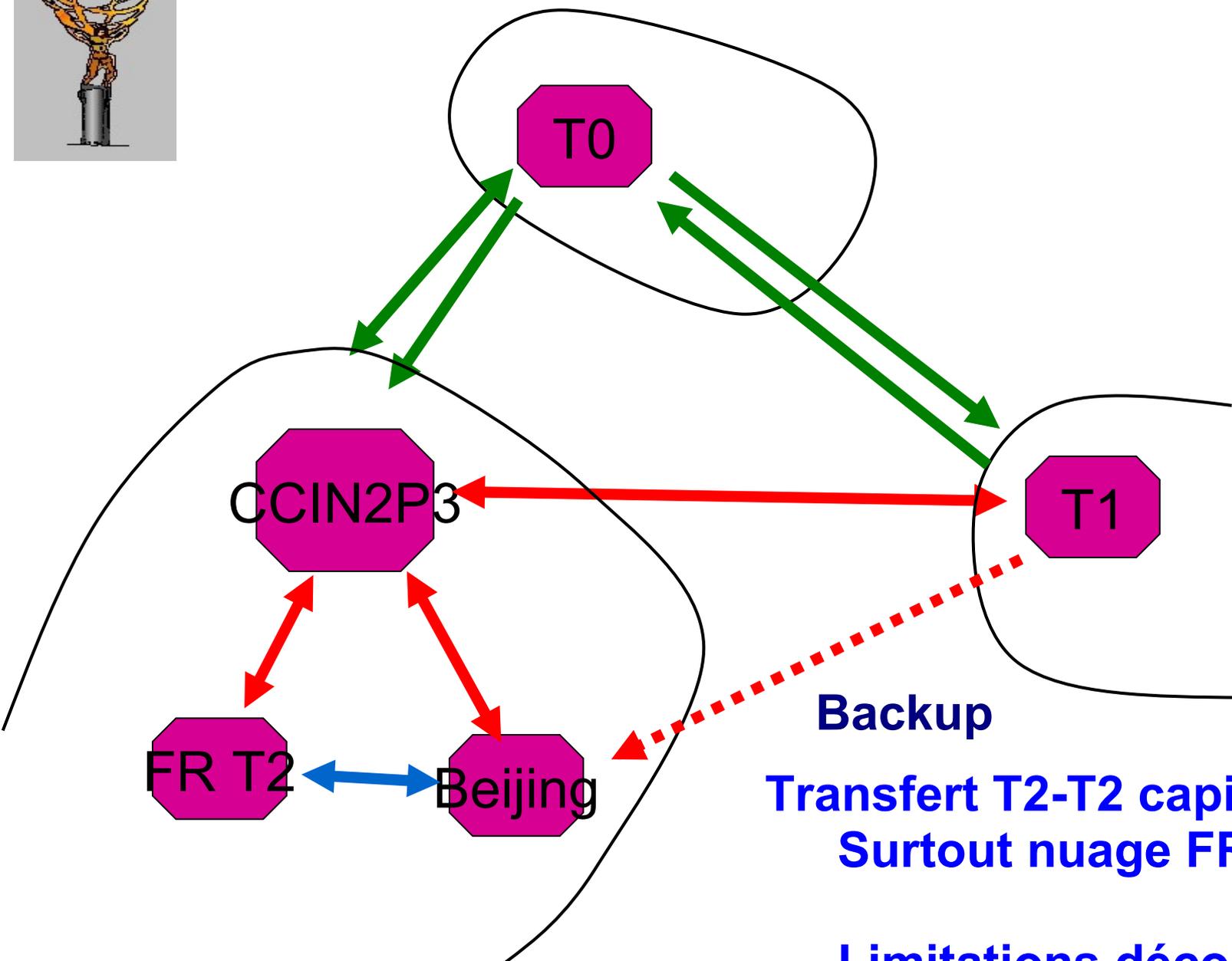
# Flux des données utilisateur



Triggered by user



# Modèle de transfert



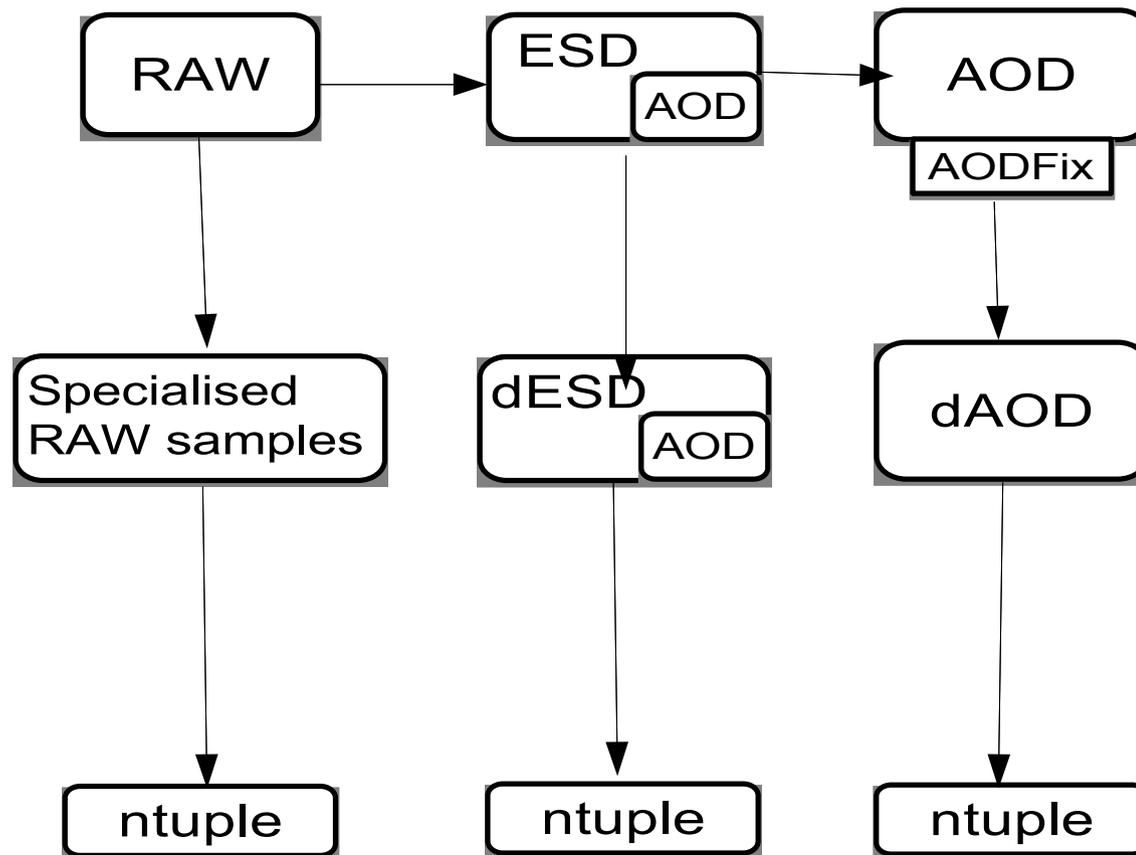
**Backup**  
**Transfert T2-T2 capital :**  
**Surtout nuage FR : FR-JP-CN**

**Limitations découvertes  
en collaboration japonais**

# Modèle de calcul 1ere année



- ▶ 2009 : Groupe de travail pour définir le modèle de calcul pour 2009-2010
  - ▶ A partir de l'état des lieux, définir le mode de fct pour 2010
  - ▶ Rapport presque final : Atlas week (Oct. 09)
  - ▶ Redéfinition de la nomenclature : DPD → dAOD ou dESD





- ◆ **Fournir un service de qualité aux utilisateurs de la Grille ATLAS**
- ◆ **Diminuer le nb de replicas de datasets parmi T1 et T2 (25 → 15)**
- ◆ **Forte pression du managment ATLAS pour implementer rapidement**
  - **Préparer les utilisateurs à un environnement**  
**avec grandes quantités de données**



- ◆ **12 Oct 09: Fermeture de l'accès en lecture/écriture à Castor/Tape pour les utilisateurs ATLAS sans responsabilité de groupe centrale**
- ◆ **Pas d'accès en lecture aux données RAW sur bande:**
  - **Analyse à partir des copies sur disque sur la Grille**
- ◆ **Mise à disposition au CERN de 0.5 TB sur disque par utilisateur**
  - **Va contraindre utilisateurs français à migrer vers la France**



# Mesure de la popularité des datasets

- ◆ **Besoin: Optimiser la répartition des données sur la Grille**
- ◆ **Classifier les datasets en fct popularité**
  - ◆ Grands nb de réplica pour les datasets populaires
  - ◆ Petits nb de réplica pour les autres
  - ◆ Identifier les données jamais accédées
- ◆ **Avoir une indication du taux de succès de l'accès aux données**
- ◆ **A long terme:**
  - ◆ Gestion plus dynamique de la répartition des données
  - ◆ Permettra aussi de mesurer la popularité des sites

(complémentaire avec le monitoring des jobs d'analyse)

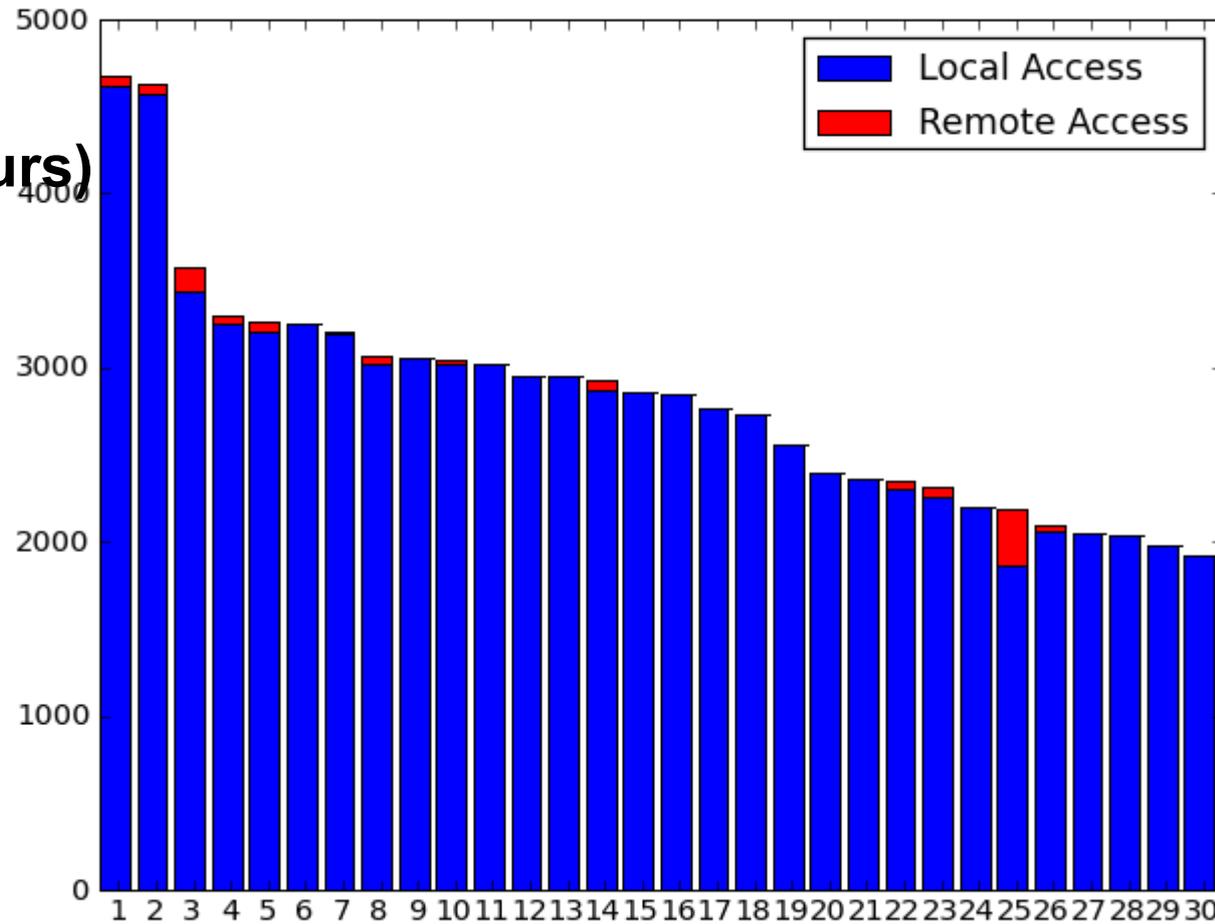
**Mesurer le nombre d'accès aux datasets par job/commandes Grille:**

**Pour chaque accès, callback à un serveur central (CERN)**

**(les jobs non Grille n'envoient pas d'infos )**



**Nb accès  
(30 derniers jours)**



**Classement dataset**



http://voatlas64/missing.py?dataset\_filter2=mc08%25merge.AOD%25\_tid%25&exclude2=%25\_sub%25&peri - Windows Internet Explorer

http://voatlas64/missing.py?dataset\_filter2=mc08%25merge.AOD%25\_tid%25&exclude2=%25\_sub

File Edit View Favorites Tools Help

http://voatlas64/missing.py?dataset\_filter2=mc08%2...

## Unused datasets (200 days)

[New Search](#)

<u>Dataset Name</u>	<u>Replicas</u>	<u>Last Read</u>	<u>Creation Date</u>
mc08.209726.WH115lnbb_Herwig.merge.AOD.e419_a84_t53_tid068645	27		2009-06-01
mc08.209725.ZH130nbb_Herwig.merge.AOD.e419_a84_t53_tid068644	28		2009-06-01
mc08.209721.Jimmy_ttbar_1Lepton20_ptMax150.merge.AOD.e419_a84_t53_tid068640	25		2009-06-01
mc08.209719.Jimmy_Zj_1Lepton15_ptMin150.merge.AOD.e419_a84_t53_tid068638	25		2009-06-01
mc08.209718.Jimmy_Wj_1Lepton20_ptMin150.merge.AOD.e419_a84_t53_tid068637	26		2009-06-01
mc08.209717.AcerMC_Wt_1Lepton20.merge.AOD.e419_a84_t53_tid068636	25		2009-06-01
mc08.209714.Jimmy_Zj_MET100_ptMin150.merge.AOD.e419_a84_t53_tid068633	27		2009-06-01
mc08.209643.PythiaB_bb_J3_1lepPt15.merge.AOD.e415_a84_t53_tid068081	23		2009-05-29
mc08.209641.PythiaB_bb_J1_1lepPt15.merge.AOD.e415_a84_t53_tid068079	22		2009-05-29
mc08.209532.SherpaW3jetstotauunu_1lepPt15.merge.AOD.e399_a84_t53_tid071427	27		2009-06-24
mc08.209531.SherpaW3jetstomunu_1lepPt15.merge.AOD.e408_a84_t53_tid066138	32		2009-05-15
mc08.209522.SherpaZ3jetstotautau_1lepPt15.merge.AOD.e415_a84_t53_tid068076	22		2009-05-29
mc08.209500.SherpabbAtautaulhMA150TB20.merge.AOD.e399_a84_t53_tid060950	27		2009-04-24
mc08.209221.SherpabbAtautaulhMA200TB20.merge.AOD.e399_a84_t53_tid060919	26		2009-04-23
mc08.209198.Jimmy_ttbar_1Lepton20_ptMin150.merge.AOD.e419_a84_t53_tid068655	29		2009-06-01
mc08.209196.Jimmy_Wj_1Lepton15_ptMin150.merge.AOD.e419_a84_t53_tid068653	28		2009-06-01

Done Local intranet 100%

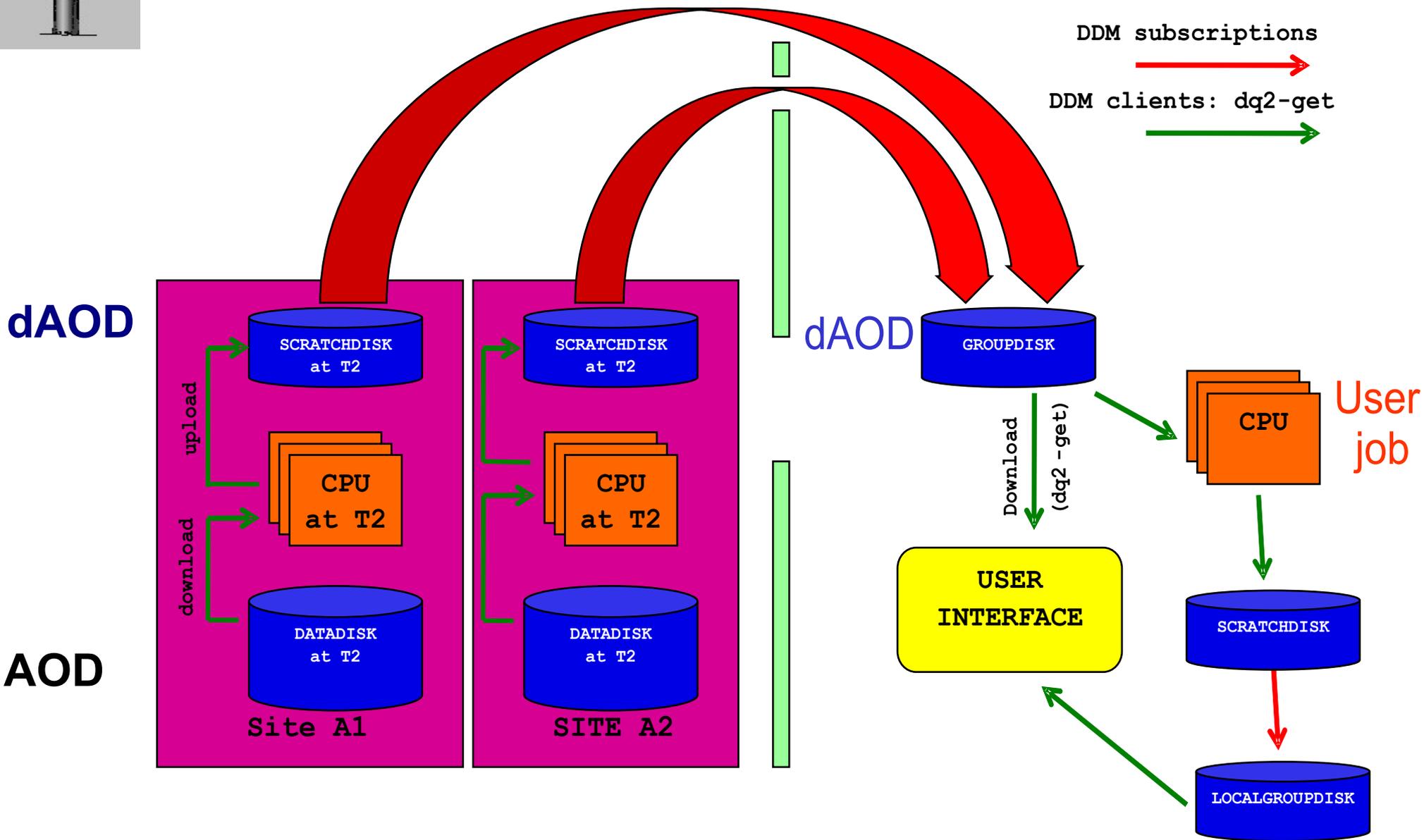


# Implémentation des zones de groupe

- ◆ **Zones de stockage dédiées à des sous communautés d'ATLAS**
  - ◆ 6 groupes de performances de detecteur
  - ◆ ~10 groupes de physique
- ◆ **Sites ATLAS ont déployé les zones disques pour des tests**
- ◆ **But actuel:**  
**Poussé les groupes à mettre en place leur chaine de traitement**

# Flux des données de groupe

Triggered by data manager





- ◆ **Définie par l'équipe Computing Operation d'ATLAS pour minimiser l'aspect politique/émotionnel en liaison**
  - ◆ **Un site (SE) rapide : Traitement des données > 10 Hz**
  - ◆ **Un site robuste pendant HammerCloud : Efficacité > 90 %**
  - ◆ **Un site réactif aux problèmes dans les derniers mois**
  - ◆ **Un site suffisamment gros : > ~150 TB**
  - ◆ **Accepte les jobs utilisateurs (important pour les T1)**
  - ◆ **Maximum de 100 couples site-groupe avec <= 4 groupe par site**
    - ◆ **~25-30 sites**
  - ◆ **1 site par % de contribution du nuage (nuage FR : ~ 13 %)**
    - ◆ **Participation importante Tokyo (2-3 paires)**

# Monitoring des zones de groupe



- Zones de groupes déjà définies pour validation par les groupes

## PERF-JETS

Group	Site	Cloud	Used(TiB)	Booked(TiB)
PERF-JETS	CERN-PROD_PERF-JETS	CERN	0.64	1.1
	TOKYO-LCG2_PERF-JETS	FRANCESITES	0.281	1.1
	TOTAL	-	0.921	2.2

## PERF-MUONS

Group	Site	Cloud	Used(TiB)	Booked(TiB)
PERF-MUONS	IN2P3-CC_PERF-MUONS	FRANCESITES	0	2.2
	INFN-NAPOLI-ATLAS_PERF-MUONS	ITALYSITES	0.274	1.1
	INFN-ROMA1_PERF-MUONS	ITALYSITES	0.425	5.5
	TRIUMF-LCG2_PERF-MUONS	CANADASITES	4.499	5.5
	TOTAL	-	5.198	14.3

## PHYS-BEAUTY

Group	Site	Cloud	Used(TiB)	Booked(TiB)
PHYS-BEAUTY	HEPHY-UIBK_PHYS-BEAUTY	FZKSITES	0	1.1
	UKI-NORTHGRID-LANCS-HEP_PHYS-BEAUTY	UKSITES	0.888	2.2
	TOTAL	-	0.888	3.3

Terminé Maintenant : Dégagé, 11 °C Mar : 20 °C Mer : 21 °C

# Répartition du stockage



- ◆ Répartition entre space -tokens définie récemment
- ◆ T1 avec zone de groupe (cas du CCIN2P3) :
  - 25 % DATADISK, 25 % MCDISK, 25 % GROUPLISK
  - 50 TB SCRATCHDISK, reste pour buffer TAPE
- ◆ T2 :
  - 25 % DATADISK, 25 % MCDISK
  - 25 % GROUPLISK (sera réaffecté si pas de groupe)
  - 50 TB SCRATCHDISK+PRODDISK



	/atlas/Role=production	/atlas, /atlas/*	Others
ATLASDATADISK	Read/Write	Read	No access
ATLASPRODDISK	Read/Write	Read	No access
ATLASMCDISK	Read/Write	Read	No access
ATLASDATATAPE	Read/Write	No access	No access
ATLASMCTAPE	Read/Write	No access	No access
ATLASGROUPDISK	Read/Write	Read	No access
ATLASSCRATCHDISK	Read/Write	Read/Write	No access
ATLASLOCALGROUPDISK	Read/Write	Read (recommended)	No access

## Protection du name space : Requête ATLAS



**ATLAS demande le déploiement des ressources 2009  
disponibles dans les tous les sites  
pour le 1er Novembre 2009**



token name	used for	@T3	@T2	@T1	@T0
ATLASDATATAPE	RAW data, ESD, AOD from re- proc			X	X
ATLASDATADISK	ESD, AOD from data		X	X	X
ATLASMCTAPE	HITS from G4, AOD from ATLFAST			X	
ATLASMCDISK	AOD from MC		X	X	X
ATLASHOTDISK	Hot files		X	X	
ATLASPRODDISK	buffer for in-and export		X		
ATLASGROUPDISK	DPD		X	X	X
ATLASSCRATCHDISK	Temp. User Data	X	X	X*	X
ATLASLOCALGROUPDISK	Perm/Local User Data @T3	X			X

Central operation management  
Phys/Perf group data  
User data