ATLAS Distributed Data Management Operations Experience and Projection

Alexei Klimentov, Pavel Nevski
Brookhaven National Laboratory
Sergey Pirogov, Alexander Zaytsev
Budker Institute of Nuclear Physics
Stephane Jezequel
LAPP Annecy

ATLAS DDM Operation Team

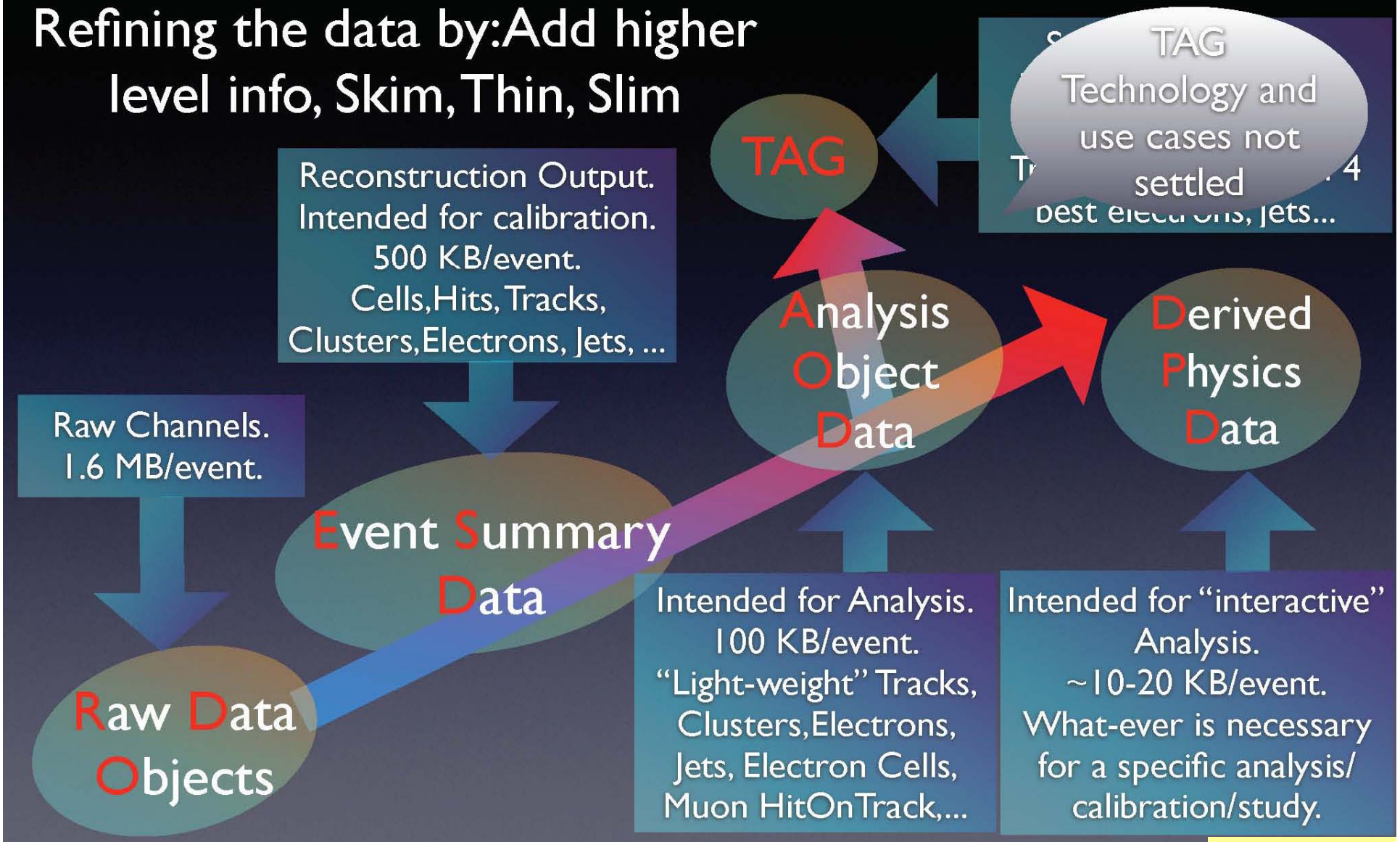
WLCG Workshop (Sep 2nd, 2007) Victoria

Outline

- ATLAS Data Model, Data Volume
- Distributed Data Management (DDM) software
- DDM Operations
  - Deployment and Operations mode
  - Main activities
    - MC production support
    - Centralized ATLAS data transfer
    - Physics Group/Users data transfer support
- Functional Tests, monitoring and analysis tools
- ATLAS cosmic run data replication (Aug-Sep 2007)
- Summary and conclusions

The Event Data Model

Refining the data by: Add higher level info, Skim, Thin, Slim



A.Farbin/UTA

The GRID

A.Farbin/UTA

- Resources Spread Around GRID

•Reprocessing of full data with improved calibrations 2 months after data taking.
•Managed Tape Access: RAW, ESD
•Disk Access: AOD, fraction of ESD

•Derive 1st pass calibrations within 24 hours.
•Reconstruct rest of the data keeping up with data taking.

40+ sites Worldwide

Tier 2

AOD

Tier 1

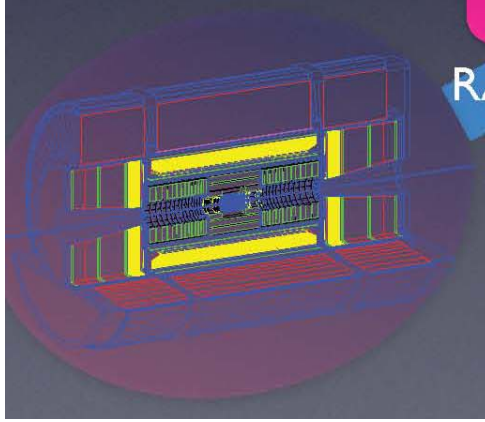
10 Sites Worldwide

RAW/ AOD/ ESD

Tier 0

RAW

•Production of simulated events.
•User Analysis: 12 CPU/ Analyzer
•Disk Store: AOD



CERN Analysis Facility

•Primary purpose: calibrations
•Small subset of collaboration will have access to full ESD.
•Limited Access to RAW Data.

ATLAS Data Management Software - Don Quijote

- The second generation of the ATLAS DDM system (DQ2)
  - DQ2 developers M.Branco, D.Cameron, P.Salgado, M.Lassnig, V.Garonne,…
  - Initial idea and architecture were proposed by M.Branco and T.Wenaus
- DQ2 is built on top of Grid data transfer tools
  - Moved to *dataset* based approach
    - Datasets : an aggregation of files plus associated DDM metadata
    - Datasets is a unit of storage and replication
    - Automatic data transfer mechanisms using distributed site services
      - Subscription system
      - Notification system
  - Production version DQ2 0.3 (since Jun 2007)
- Technicalities :
  - Global services
    - dataset repository
    - dataset location catalog
    - logical file names only, no global physical file catalog
  - Local Site services
    - File Catalog specific per Grid/site/tier. It provides logical to physical file name mapping. Implementations of this catalog are Grid specific.

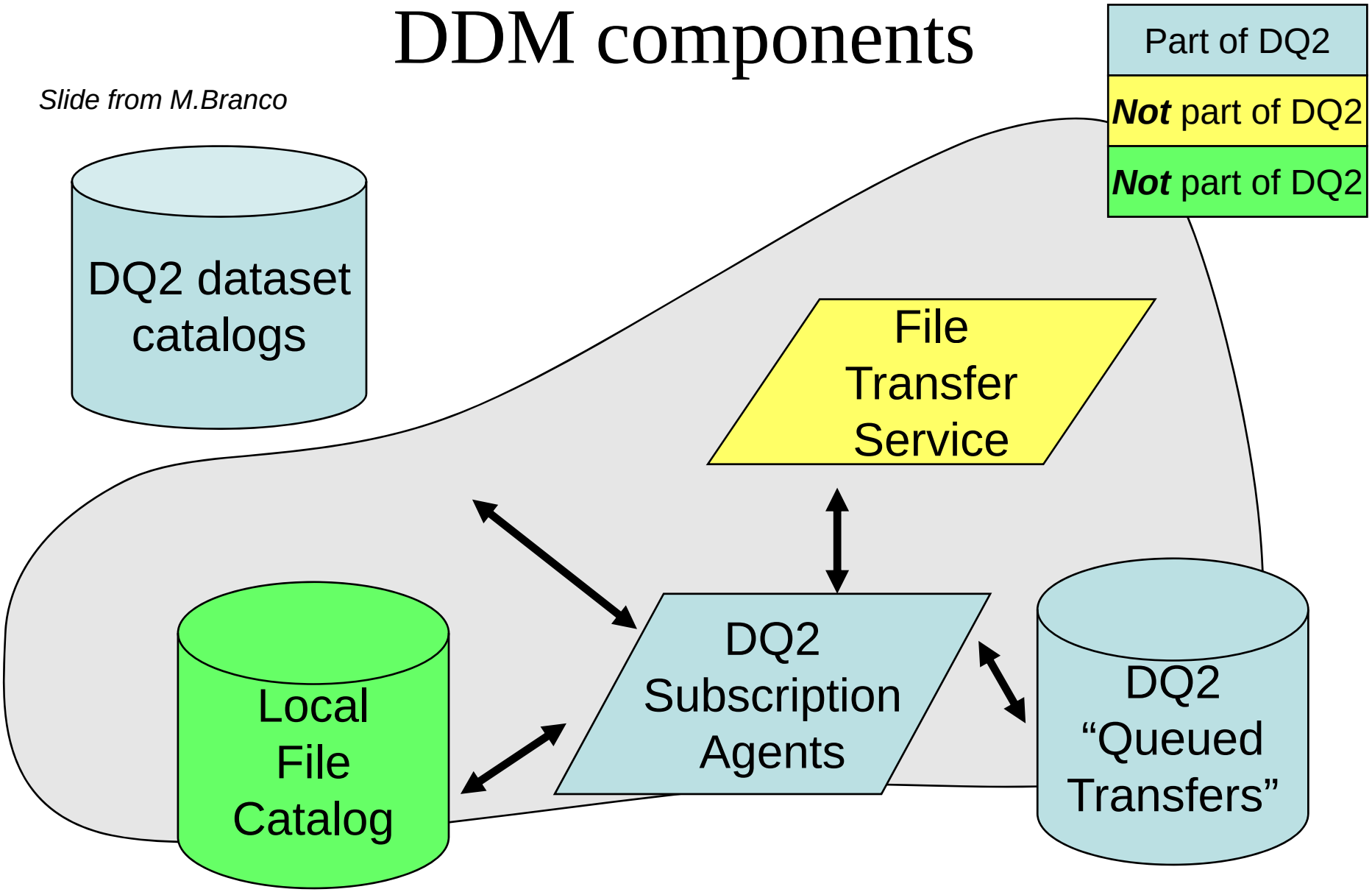
DQ2 Concepts

- ‘Dataset’:
  - an aggregation of data (spanning more than one physical file!), which are processed together and serve collectively as input or output of a computation or data acquisition process.
  - Flexible definition:
    - … can be used for grouping related data (e.g. RAW from a run with a given luminosity)
    - … can be used for data movement purposes
  - Dataset concept is extended to all ATLAS data (MC, DAQ, DB releases, etc)
- ‘File’:
  - constituent of a dataset
    - Identified by Logical File Name (LFN) and GUID
- ‘Site’
  - A computing site providing storage facilities for ATLAS
    - … which may be a federated site
- ‘Subscription’
  - Mechanism to request updates of a dataset to be delivered to a site

See M.Lassnig talk CHEP Sep 3, GM2 session

DDM components

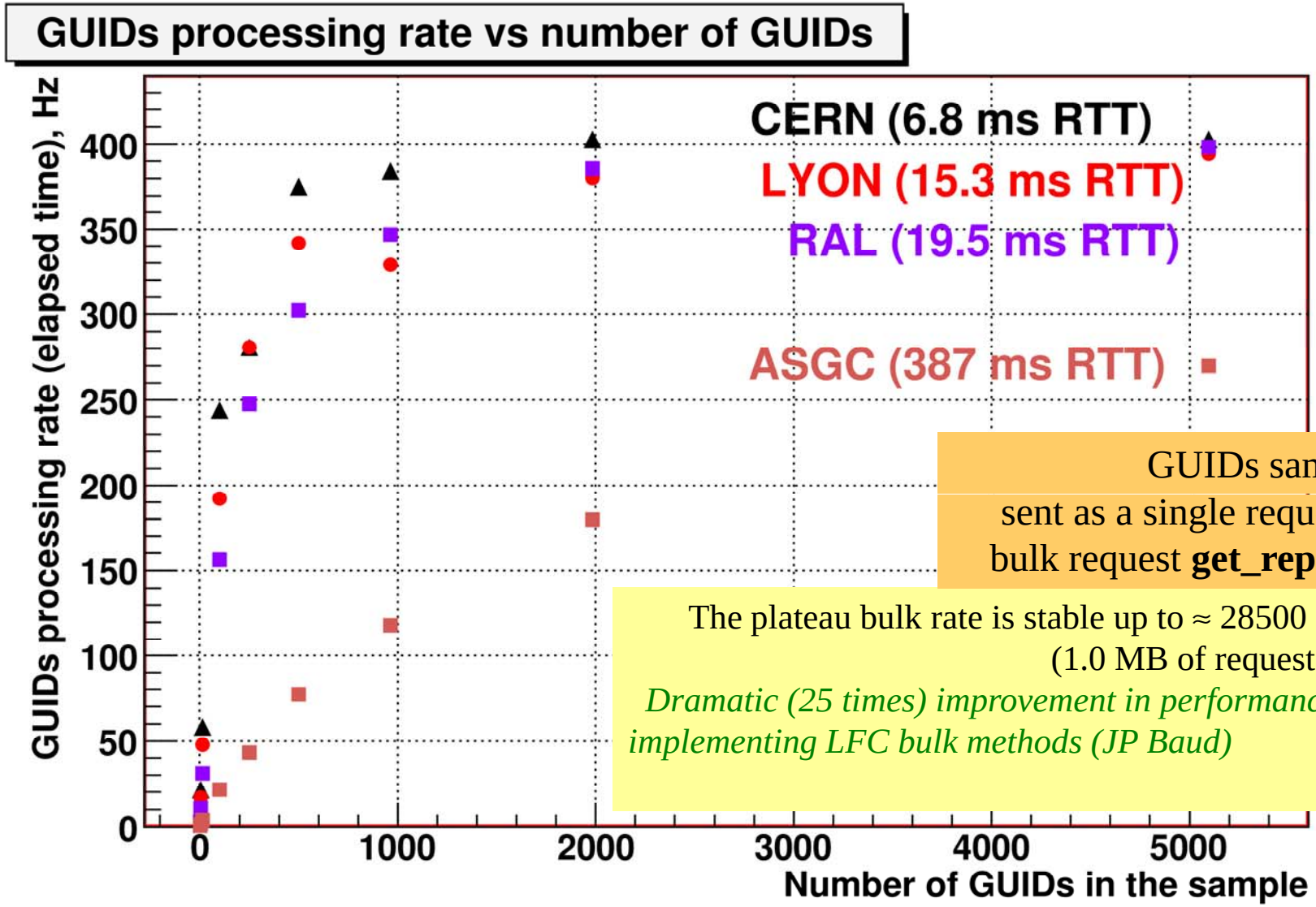
Slide from M.Branco



DDM Components. LFC

- LFC – one of vital components, out of DDM
- It is important the two work together as expected
- We organized systematic measurements to understand performance to spot and fix problems...
- Initial LFC performance was found poor : 12 Hz
- It was joint effort of ATLAS DDM Operations team, DQ2 developers, CERN ARDA and LFC Author to understand and to improve the catalog's performance.
- ATLAS performance requirement was driven by our computing and event model

Results on Recent Performance Testing of the LFC @ CERN (21–27/06/2007)



GUIDs sample is sent as a single request via bulk request **get_replicas()**

The plateau bulk rate is stable up to ≈ 28500 GUIDs (1.0 MB of request size),?

Dramatic (25 times) improvement in performance after implementing LFC bulk methods (JP Baud)

DDM/DQ2 more than just s/w development

- DDM forced the introduction of many concepts, defined in the Computing Model, onto the middleware:
  - ATLAS Association between Tier-1/Tier-2s
  - Distinction between temporary (e.g. disk) and archival (e.g. tape) areas
  - Datasets as the unit of data handling
- Not all ATLAS concepts were originally supported by the GRID middleware.

See M.Lassnig talk CHEP Sep 3, GM2 session

dq2 end-user tools and DDM/DQ2 client CLI

- Insulate flavor-specific dependences from end-users
- Provide quick accesses to DQ2 datasets for end-users
- Set of programs and scripts on top of DDM/DQ2 catalogs (T.Maeno et al)
  - dq2_get, dq2_ls,dq2_register/dq2_put,dq2_cr, dq2_cleanup, dq2_sample
- DQ2 client API:
  - Interfaces to all dataset catalogs
    - in a secure way for any ‘write’ operations
  - Guarantees consistency between all the (“loosely coupled”) dataset catalogs

    e.g. a dataset in the location catalog refers to the same dataset in the content catalog..
    - Consistency now being improve (for 0.3) with initial support for transactions, etc

Deployment and Operations Model

- Dataset catalog
  - Single instance at CERN, serving complete ATLAS
    - No regional catalogues deployed
- Clouds and Tiers
  - 3 GRID flavours : LCG, NG, OSG
  - 10 clouds : CANADA,FRANCE,FZK,ITALY,NDGF,NL,SPAIN,TAIWAN,UK,US
  - LCG
    - DQ2 0.2 : *VO box* and LFC are located at Tier-1 and serve the whole cloud
    - DQ2 0.3 : *VO boxes* are located at CERN, LFCs are located at Tier-1s
    - 1 LFC instance per cloud
  - US ATLAS operates all DQ2 instances at the site-level, with 'expert' support at BNL
    - Each Tier 1/2 has VO box and LRC
    - New Operations model (TBD)
  - NordicGrid
    - 1 FC per cloud
    - 1 DQ2 VO box per cloud (at CERN)

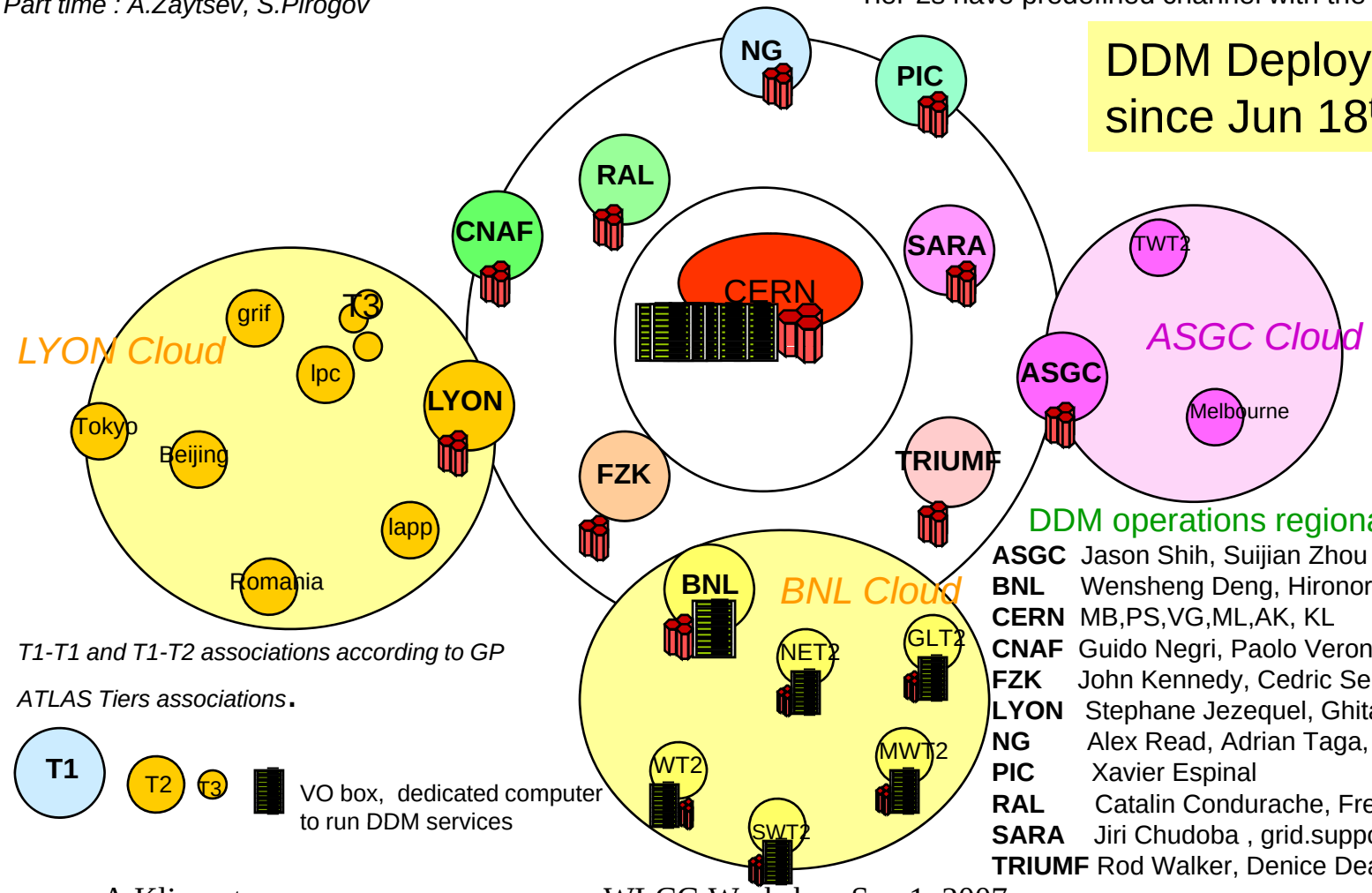
VO box – Virtual Organization Box, dedicated computer(s) on site to run ATLAS specific services

DDM Deployment and Operations Model (DQ2 0.3)

DDM operations central team
Alexei Klimentov, Pavel Nevski, *Kai Leffhalm*
DQ2 developers maintain CERN instances
Part time : A.Zaytsev, S.Pirogov

- ✓All Tier-1s have predefined (software) channel with CERN and with each other.
- ✓Tier-2s are associated with one Tier-1 and form the cloud
- ✓Tier-2s have predefined channel with the parent Tier-1 only.

DDM Deployment Model since Jun 18th, 2007



T1-T1 and T1-T2 associations according to GP ATLAS Tiers associations.

T1 T2 T3 VO box, dedicated computer to run DDM services

DDM operations regional team :

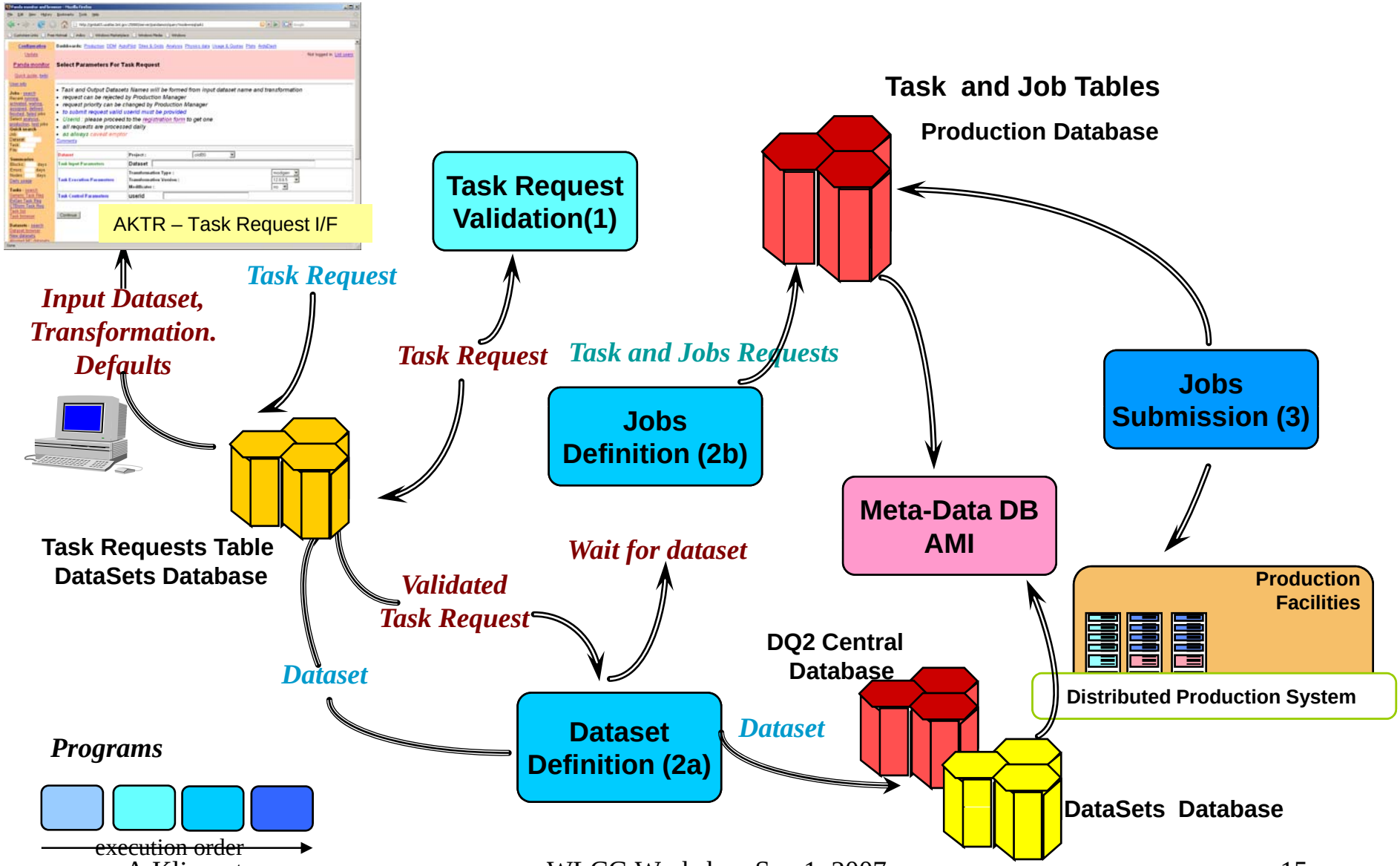
ASGC Jason Shih, Suijian Zhou
BNL Wensheng Deng, Hironori Ito
CERN MB,PS,VG,ML,AK, KL
CNAF Guido Negri, Paolo Veronesi, Giuseppe Lo Re
FZK John Kennedy, Cedric Serfon
LYON Stephane Jezequel, Ghita Rahal
NG Alex Read, Adrian Taga, D.Cameron
PIC Xavier Espinal
RAL Catalin Condurache, Frederic Brochu
SARA Jiri Chudoba , grid.support@sara.nl
TRIUMF Rod Walker, Denice Deatrich, Reda Tafirout

DDM Operations. Main Activities

DDM Operations team includes central and regional groups

- Support MC production
- Data replication to sites
- Developing monitoring and control tools
- Users and groups support
- Data integrity check
  - Data registered in DDM vs File Catalogs entries
  - Data registered in File Catalogs vs SE entries
  - ’zombies’ and ‘0’ length files
- Deletion of obsolete or invalidated data
- Data movement within Tiers

ATLAS task and dataset registration



Tasks/Datasets In Action

Run by wenaus

Jobs - search
running, activated, waiting, assigned, defined, finished, failed
Analysis jobs
Old archive

Quick search
PandaID
Dataset
Task

Summaries
Blocks: days
Errors: days
Nodes: days

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Full task list
Task browser

Datasets - search
In,out,dispatch,all
Subscriptions

Sites
Site specs
BNL BU OU UC
UTA LCG NG All

System statistics

Logging monitor

Task browser

Click for help

Current selection: STATUS=running&GRID=osg Clear selection

Click to show and select physics types

Releases: v11000301 (46) v11000302 (77) v11000303 (61) v11000304 (12) v11000305 (37) v11000306 (9) v11000307 (6) v11000308 (33) v11000401 (1)
Stages: digit (99) evgen (74) merge (18) reco (24) recon (67)
Outputs: AOD (106) CBNT (97) ESD (88) EVNT (74) HIST (9) HITS (102) RDO (102) TAG (18)
Grids: lcg (28) anygrid (1) nordic (52) osg (81) lcg-cg (120)
Status: aborted (12) done (115) finished (13) rejected (1) running (108) submitted (33)

Total selected events=2160000 jobs=21095 jobs done=16645

Task name	Task ID	Status	Grid	Total jobs	Done jobs	Events	Input files	Release	Formats
mc11.007204.singlepart_mu4.recon.v11000302	694	running	osg	2200	1498	220000	2200	11.0.3	ESD.AOD.CBNT
mc11.007430.singlepart_singlepi_pt2.digit.v11000308	667	running	osg	500	331	100000	20	11.0.3	RDO.HITS
mc11.007200.singlepart_mu2.recon.v11000303	579	running	osg	500	464	50000	500	11.0.3	ESD.AOD.CBNT
mc11.007200.singlepart_mu2.digit.v11000303									ITS
mc11.007222.singlepart_mu26.recon.v11000303									OD.CBNT
mc11.007216.singlepart_mu18.recon.v11000303									OD.CBNT
mc11.007211.singlepart_mu10.recon.v11000303									OD.CBNT
mc11.007207.singlepart_mu6.recon.v11000303									OD.CBNT
mc11.007216.singlepart_mu18.digit.v11000302									ITS
mc11.007207.singlepart_mu6.digit.v11000302									ITS
mc11.005001.pythia_minbias.recon.v11000304									OD.CBNT
mc11.005001.pythia_minbias.digit.v11000304									ITS
mc11.005800.JF17_pythia_lo[illegible]ejet_filter.recon.v1									OD.CBNT
mc11.005800.JF17_pythia[illegible]_filter.digit.v11									ITS
mc11.005055.PythiaPhotonJ[illegible] v110[illegible]303									OD.CBNT
mc11.005056.PythiaPhotonJet2.[illegible]10[illegible]303									OD.CBNT
mc11.005056.PythiaPhotonJet2.digi[illegible]									ITS

Task mc11.007216.singlepart_mu18.recon.v11000303

Datasets for task mc11.007216.singlepart_mu18.recon.v11000303

mc11.007216.singlepart_mu18.recon.ESD.v11000303
mc11.007216.singlepart_mu18.recon.AOD.v11000303
mc11.007216.singlepart_mu18.recon.CBNT.v11000303

Parameters for task mc11.007216.singlepart_mu18.recon.v11000303

Task ID	**562**
Project	mc11
Input dataset	mc11.007216.singlepart_mu18.digit.v11000302
Task name	mc11.007216.singlepart_mu18.recon.v11000303
Formats	ESD.AOD.CBNT
Transformation	csc.reco.trf
Trf Version	11.0.3.3
Release	11.0.3
Owner	i_hinchliffe@lbl.gov
CPU/event	100
Memory usage	600
First inputfile number	1
Input files	2200
Events	220000
Events/file	100
Grid	osg

Task browser

Click for help

Current selection: STATUS=running&GRID=osg

Physics type	Events
A3_Ztautau_tightfilter	20000
AlpgenJimmyW4jet	80000
Bs_Jpsi_mu6mu3_phi_KplusKminus	10000
Electron_Pt_25	20000
Electrons_e100	40000
FJ1_fwjets_e200	20000
FJ2_pythia_jetjet	100000
Gmm_500_pythia_photos	250000
H3_120_gamgam	20000
J1_pythia_jetjet	150000
J2_Pt_35_70	20000
J2_pythia_jetjet	150000
J3_Pt_70_140	20000
J3_pythia_jetjet	150000
J4_pythia_jetjet	150000
J5_Pt_280_560	20000
J5_pythia_jetjet	150000
J6_Pt_560_1120	30000
J6_pythia_jetjet	50000
J7_pythia_jetjet	30000
J8_pythia_jetjet	30000
JF17_pythia_jet_filter	3000000
JF17_pythia_loosejet_filter	1800000
LRSM_WR_1800_300	60000
M1_minbias	20000
McAtNloWenu	45000
McAtNloWmunu	60000
P5P_Single211	40000
P7P_Single211	20000
Photon_Pt_60	20000
Photons_e100	40000
PythiaH120gamgam	240000
PythiaH130zz4l	200000

Tasks define production tasks and record their associated metadata

Task Query Form

- *Text fields don't require the exact matching*
- Queries in *italic* not implemetted yet

MC Task

Project :	mc11
Input Dataset	
Transformation :	csc.reco.trf
Transformation Version :	11.0.4.1
Grid Flavour :	osg
Requested By	
Output Task Name	
Priority	
Status :	running

Datasets define and organize the task inputs and outputs

DDM Operations. Data Replication

- Centralized and automatic (according to computing model)
  - Simulated data
    - AOD/NTUP/TAG (current data volume ~1.5 TB/week)
      - Each Tier1 has a complete dataset replicas
      - Tier-2s are defined what fraction of data they will keep
        - » From 1% to 100% . Some Tier-2s are shared a complete replicas with other sites
        - » Some Tier-2s get only datasets match to the pattern
    - Validation samples
      - Replicated to CERN and 2 Tier-1s (BNL and LYON)
  - *Input data for MC production tasks*
  - Database releases
    - *replicated to all Tiers. Data volume is relatively small (~100MB)*
  - Conditions data
    - *Replicated to all Tier-1s*
  - Cosmics data
    - RAW data is replicated to Tier-1s according to T1s shares
    - ESDs are replicated to all Tier-1s and Tier-2s according to their request
    - Tier-2 can request RAW data replication via DDM Request I/F

DDM Operations. Data Replication (cont)

Panda monitor and browser - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://gridui03.usatlas.bnl.gov:25880/server/pandamon/query?mode=reqsubs0

Customize Links | Free Hotmail | indico | Windows Media | Windows | BeOnLine: связь, ко... | Windows Marketplace

User info

Jobs - search
Recent running, activated, waiting, assigned, defined, finished, failed jobs
Select analysis, production, test jobs

Quick search
Job
Dataset
Task
File

Summaries
Blocks: days
Errors: days
Nodes: days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Task list
Task browser

Datasets - search
Dataset browser
New datasets
Aborted MC datasets
Panda subscriptions
All subscriptions

Datasets Distribution
DDM Req
Req list
AODs
RDOs
Conditions DS
DB Releases
Validation Samples
Functional Tests
M4

Sites - see all
BNL BU IU OU SLAC
UC UMICH UTA LCG
NG

- **Dataset pattern :** Preferably PhysicsShort (Icase sensitivel). Wild cards (*) are allowed (f.e. *Atautau*). Don't add dataset format to the pattern (PhysicsShort is the 3rd field in dataset name)
- **Validity :** *OneTimeCopy* only existing datasets are subscribed, and subscription is done only once *Periodic* the request is checked periodically and datasets matched to the pattern are subscribed. DQ2 option *--wait-for-sources* is used for both modes
- **Userid : If you are not registrated yet, please proceed to the registration form to get one**
- **Destionation Tier :** as in ToA (Consult Tiers Of Atlas for details). Usually Tier-2s are subscribed from the parent Tier-1, if dataset is missing on parent Tier-1, then Tier-1 is subscribed to it as well.
- ***to submit request valid userid must be provided***
- **Approval process :** before subscription is started the request must be approved by Regional DDM Operations. In case of data transfer to Tier-2 the corresponding Tier-1 authorities are contacted to check available disk space and data transfer policy. Data transfer request for data volume 500+GB must be approved by Physics Coordinator
- *request can be rejected by Regional DDM Operations*
- *request priority can be changed by Regional DDM Operations*
- ***all requests are processed daily***
- **New Request** can be added by cloning the existing request. Click *Req list* , select request you want to clone, edit it and click *AddRequest*
- **Transfer Volume :** fraction of dataset (in percent) to be replicated. The particular files within the dataset are selected randomly. The Transfer Valume can be changed only if Validity is *'OneTimeCopy'*.
- **as always caveat emptor**

Comments

Subscription Parameters	Dataset Pattern	
	Software Version :	13.0.20.1
	Format :	AOD
	Destination (as in TiersOfAtlas)	(f.e. MILANO, AGLT2)
	Validity :	OneTimeCopy
	Prority :	Immediate
Control Parameters	userid	
	Req Type :	calibration
	Transfer Volume	100 %

Continue

Individual or Group or Cloud
- via DDM data transfer request I/F
- using end-users tools or/and DQ2 client

Done

User's/Group's Data Transfer Request

Panda monitor and browser - Mozilla Firefox

File Edit View History Bo

Customize Links | Free Hotmail | indico | Windows Media | Windows | BeOnLine: связь, ко... | Windows Marketplace

File

Summaries
Blocks: days
Errors: days
Nodes: days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Task list
Task browser

Datasets - search
Dataset browser
New datasets
Aborted MC datasets
Panda subscriptions
All subscriptions

Datasets Distribution
DDM Req
Req list
AODs
RDOs
Conditions DS
DB Releases
Validation Samples
Functional Tests
M4

Sites - see all
BNL BU IU OU SLAC
UC UMICH UTA LCG
NG

Applications
CHARMM

Logging monitor

ALL ASGCDISK AU-UNIMELB BNLDISK BNLTAPE CERNCAF CERNPROD CNAFDISK FZKDISK ITEP LAL LNF LYONDISK MILANO MWT2_UC
NAPOLI NDGFT1DISK NIKHEF PICDISK RALDISK ROMA1 SARADISK TOKYO

Parameters of filter: *ALL Requests for ALL tiers, Name="%", SW release="any"*

Patterns	Datasets	LastRequest	Approved	Subscribed
288	2381.0	Aug 31 08:58	Aug 31 08:51	Aug 31 09:00

Pattern	Release	ReqId	Destination	ReqStatus	Datasets	Timestamp
Zee.AOD.	v13.0.20.1	249	LAL	pending_emailed	6 [list]	Aug 31 08:58
Zee.AOD.*	v13.0.20.1	248	LAL	approved	6 [list]	Aug 31 08:51
M4.0020347.No.Streaming.RAW.*	v1.8.1	247	TOKYO	subscribed	1 [list]	Aug 29 10:50
WH120 pythia.*.AOD.*	v12.0.7.1	246	MWT2_UC	approved	2 [list]	Aug 14 03:45
pythia_minbias.NTUP.*	v12.0.6.5	245				
pythia_minbias.NTUP.*	v12.0.6.4	244				
pythia_minbias.NTUP.*	v12.0.6.3	243				
pythia_minbias.NTUP.*	v12.0.6.2	242				
pythia_minbias.NTUP.*	v12.0.6.1	241				
McAtNlo0310_JIMMY_WpWm_taunutaunu.NTUP.*	v12.0.6.4	240				
evgen.EVNT.*	v13.0.10.1	239				
evgen.EVNT.*	v13.0.10.1	238				
evgen.EVNT.*	v13.0.10.1	237				
evgen.EVNT.*	v13.0.10.1	236				
evgen.EVNT.*	v13.0.10.1	235				
evgen.EVNT.*	v13.0.10.1	234				
evgen.EVNT.*	v13.0.10.1	233				
evgen.EVNT.*	v13.0.10.1	232				
lightstop.AOD.*	v12.0.6.4	231				
lightstop.AOD.*	v12.0.6.4	230	NAPOLI	subscribed	1 [list]	Jul 19 06:49
lightstop.AOD.*	v12.0.6.4	229	MILANO	subscribed	1 [list]	Jul 19 06:49
misal1_csc11.005011.J2_pythia_jetjet.digit*.RDO.*	v12.0.6.1	228	BNLDISK	stopped	1 [list]	Jul 3 15:06
misal1_csc11.005012.J3_pythia_jetjet.digit*.RDO.*	v12.0.6.1	227	BNLDISK	subscribed	1 [list]	Jul 3 15:37
misal1_csc11.005014.J5_pythia_jetjet.digit*.RDO.*	v12.0.6.1	226	BNLDISK	stopped	1 [list]	Jul 3 15:02
misal1_csc11.005015.J6_pythia_jetjet.digit.RDO.*	v12.0.6.1	225	BNLDISK	subscribed	1 [list]	Jul 3 15:36
misal1_csc11.005017.J8_pythia_jetjet.digit*.RDO.*	v12.0.31.5	223	BNLDISK	subscribed	1 [list]	Jul 3 15:36

Done

Dataset(s) subscription and replication

DDM Operations. Monitoring (‘critical’ data)

nda monitor and browser - Mozilla Firefox

Edit View History Bookmark

ustomize Links Free Hotmail

Configuration

Update

anda monitor

Quick guide, twiki

er info

s - search
cent running,
vated, waiting,
igned, defined,
shed, failed jobs
ect analysis,
duction, test jobs
ick search

aset
k

mmaries
cks: days
ors: days
des: days
ly usage

sks - search
neric Task Req
Gen Task Req
Gsim Task Req
k list
k browser

asets - search
aset browser
y datasets
orted MC datasets
nda subscriptions
subscriptions

asets Distribution
M Req
list
Ds
Os
nditions DS
Releases
idation Samples

Dashboards: Production DDM PandaMover AutoPilot Sites & Grids Analysis Physics data Usage & Quotas Plots ProdDash DDMDash

Not logged in. List users

Database Releases Distribution

Database release replication status

- *Datasets are automatically subscribed to Tier-1s and Tier-2s.Tier-1s are subscribed from CERN. Tier-2s are subscribed only after corresponding Tier-1 has a complete dataset replicas*
- *Green - site has a complete dataset replicas (data transfer is done)*
- *Orange - if site has an incomplete dataset replicas.It also means that subscription is processed*
- *Red - the subscription is not processed*

Comments

Dataset	CANADA	CERN	FRANCE	FZK	ITALY	NDGF	NL	SPAIN	ASGC	UK	USA
ddo.000001.Atlas.Ideal.DBRelease.v040101 (Subscription Time : Wed Aug 8 10:47:55 2007)	ALBERTA MONTREAL SFU TORON TRIUMF UVIC	CERNCAF	BEIJING CPPM LAL LAPP LPC LPNHE LYON NIPNE_02 NIPNE_07 SACLAY TOKYO	CSCS CYF DESY-HH DESY-ZN FZK LRZ WUP FZU UNI-FREIBURG	CNAF LNF MILANO NAPOLI ROMA1	NDGFT1	IHEP ITEP JINR SARA PNPI SINP	IFAE PIC UAM IFIC LIP-LISBON	ASGC AU-UNIMELB TW-FTT	RAL EDINBURGH GLASGOW MANC LANCS LIV OXF CAM BHAM RHUL BRUN QMUL UCLCC UCLHEP DUR RALPP ICL LESC	BNL BU IU MWT2_IU MWT2_UC OU OUHEP SLAC SLACXRD UC AGLT2 UTA UTA_SWT2 WISC IU_BC SMU
ddo.000001.Atlas.Ideal.DBRelease.v040001 (Subscription Time : Sun Jun 10 11:34:42 2007)	ALBERTA MONTREAL SFU TORON TRIUMF UVIC	CERNCAF	BEIJING CPPM LAL LAPP LPC LPNHE LYON NIPNE_02 NIPNE_07 SACLAY TOKYO	CSCS CYF DESY-HH DESY-ZN FZK FZU LRZ WUP UNI-FREIBURG	CNAF LNF MILANO NAPOLI ROMA1	NDGFT1	IHEP ITEP JINR PNPI SARA SINP	IFAE IFIC PIC UAM LIP-LISBON	ASGC TW-FTT AU-UNIMELB	RAL GLASGOW MANC LANCS LIV OXF CAM BHAM RHUL BRUN QMUL UCLCC UCLHEP DUR EDINBURGH RALPP ICL LESC	BNL BU IU MWT2_IU MWT2_UC OU OUHEP SLAC SLACXRD UC AGLT2 UTA UTA_SWT2 WISC IU_BC SMU
ddo.000001.Atlas.Ideal.DBRelease.v030301 (Subscription Time : Mon Apr 23 05:55:25 2007)	ALBERTA MONTREAL SFU TORON TRIUMF UVIC	CERNCAF	BEIJING CPPM LAL LAPP LPC LPNHE LYON NIPNE_02 NIPNE_07 SACLAY TOKYO	CYF DESY-HH DESY-ZN FZK LRZ WUP CSCS FZU UNI-FREIBURG	CNAF LNF MILANO NAPOLI ROMA1	NDGFT1	IHEP JINR PNPI SARA ITEP SINP	IFAE IFIC PIC UAM LIP-LISBON	ASGC AU-UNIMELB TW-FTT	RAL EDINBURGH GLASGOW MANC LANCS LIV OXF CAM BHAM RHUL BRUN QMUL UCLCC UCLHEP	BNL BU BU_TEST IU MWT2_IU MWT2_UC OU OUHEP SLAC SLACXRD UC AGLT2 UTA UTA_SWT2

DDM Operations. Monitoring

Customize Links | Free Hotmail | indico | Windows Media | Windows | BeOnLine: связь, ко...

Quick guide, twiki

User info

Jobs - search
Recent running, activated, waiting, assigned, defined, finished, failed jobs
Select analysis, production, test jobs

Quick search
Job
Dataset
Task
File

Summaries
Blocks: days
Errors: days
Nodes: days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Task list
Task browser

Datasets - search
Dataset browser
New datasets
Aborted MC datasets
Panda subscriptions
All subscriptions

Datasets Distribution
DDM Req
Req list
AODs
RDOs
Conditions DS
DB Releases
Validation Samples
Functional Tests
M4

Sites - see all
BNL BU IU OU SLAC
UC UMICH UTA LCG

T1s summary | Clouds summary | **Show LAST 500 datasets** | Show A

- Datasets are automatically subscribed to Tier-1s and to preselected Tier-
eplicas.Tier-2s are subscribed from Tier-1 in the same cloud (The table of da
- *Green* - site has a complete dataset replicas (data transfer is done)
- *Orange* - if site has an incomplete dataset replicas. It also means that
- *Red* - site has 0 files
- *Magenta* - site not subscribed or subscription is not processed

Comments

ARDA Monitoring | DQ2 Dataset Browser

Total Datasets : 4613, AOD : 2165, HPTV : 377, NTUP : 165

Sites :	1	2	3	4	5	6
AOD complete replicas	582	306	157	170	190	248

Last Datasets Subscription : Mon Aug 27 05:16:45 2007 CET Last Files

- *Dataset pattern name doesn't require the exact matching (if you use %PythiaWenu%)*
- **Note: When using a filter it may take a long time your request.**

Filter

Project: any
Pattern: %
Format: any
SW Release: any
Source: ALL
Destination: ALL

Continue

LAST 500 Datasets for ALL Sites.

Datasets	Transfer Completed	Transfer InProgress	Datasets
500	15	100	385

Dataset

trig1_misal1_mc12.005941.McAtNlo0310_JIMMY_WpZ_Inull.
(Subscription Time : Mon Aug 27 05:16:45 2007)

trig1_misal1_mc12.005941.McAtNlo0310_JIMMY_WpZ_Inull.recon.NTUP.v12000604_tid013277

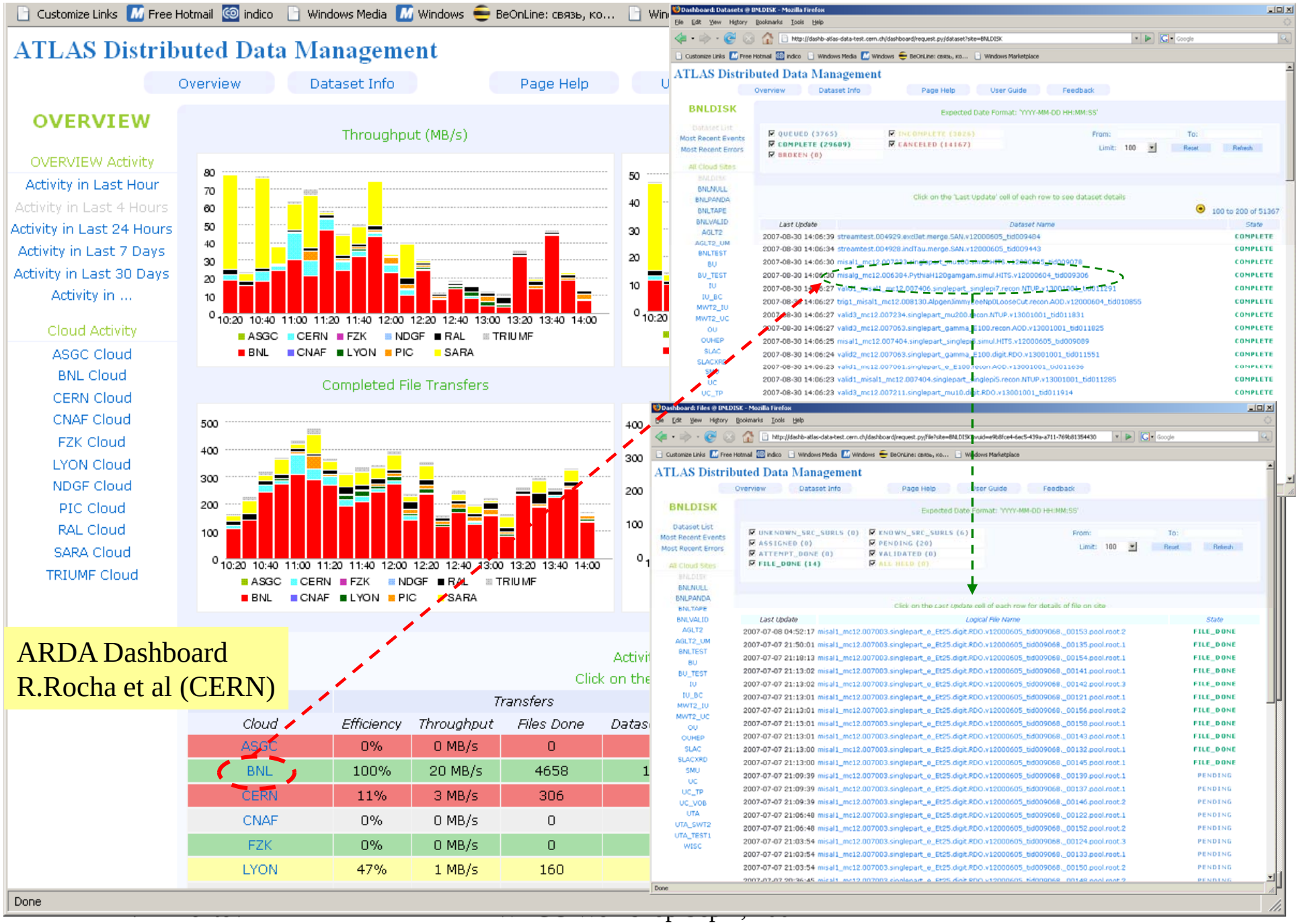
Done

Replication Status to Tiers

Conditions data replication status

Datasets	Total Files in datasets	Last Subscription	LFC Checked	Last Transfer
17	1049	Jul 20 09:16:42	Aug 30 06:30:40	Aug 29 16:30:21

Dataset	ASGC	BNL	CNAF	FZK	LYON	NDGF	PIC	RAL	SARA	TRIUMF
cmccond.000001.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)
cmccond.000001.conditions.simul.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(22/21)	(22/22)	(22/22)	(22/22)	(22/22)	(22/22)	(22/21)	(22/22)	(22/22)	(22/22)
comcond.000001.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(11/10)	(11/11)	(11/11)	(11/11)	(11/11)	(11/11)	(11/10)	(11/11)	(11/11)	(11/11)
comcond.000001.lar_conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(310/271)	(310/309)	(310/282)	(310/296)	(310/265)	(310/306)	(310/292)	(310/309)	(310/306)	(310/309)
comcond.000002.lar_conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(70/26)	(70/70)	(70/64)	(70/70)	(70/62)	(70/67)	(70/50)	(70/70)	(70/70)	(70/70)
oflcond.000001.bfield_conditions.simul.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(3/0)	(3/3)	(3/3)	(3/3)	(3/3)	(3/3)	(3/2)	(3/3)	(3/3)	(3/3)
oflcond.000001.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(13/12)	(13/13)	(13/13)	(13/13)	(13/12)	(13/13)	(13/12)	(13/13)	(13/13)	(13/13)
oflcond.000001.conditions.simul.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(13/10)	(13/13)	(13/13)	(13/13)	(13/12)	(13/13)	(13/10)	(13/13)	(13/13)	(13/13)
oflcond.000002.conditions.simul.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(58/56)	(58/58)	(58/58)	(58/58)	(58/58)	(58/58)	(58/57)	(58/58)	(58/58)	(58/58)
oflcond.000003.conditions.simul.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(41/28)	(41/41)	(41/28)	(41/40)	(41/33)	(41/33)	(41/32)	(41/41)	(41/41)	(41/41)
tbcond.000001.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(296/292)	(298/298)	(298/298)	(298/296)	(296/278)	(296/296)	(296/295)	(296/296)	(296/296)	(296/296)
tbcond.000002.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(5/4)	(5/5)	(5/4)	(5/5)	(5/4)	(5/5)	(5/4)	(5/5)	(5/5)	(5/5)
tbcond.000003.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(368/280)	(368/368)	(368/368)	(368/368)	(368/348)	(368/368)	(368/328)	(368/368)	(368/368)	(368/368)
tbcond.000004.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(5/4)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/4)	(5/5)	(5/5)	(5/5)
tbcond.000005.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)	(5/5)
testcond.000001.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(9/7)	(9/7)	(9/7)	(9/7)	(9/5)	(9/7)	(9/7)	(9/7)	(9/7)	(9/7)
testcond.000005.conditions.recon.pool.v0000 (Subscription Time: Jul 20 2007 09:16)	(21/18)	(21/21)	(21/21)	(21/21)	(21/21)	(21/21)	(21/19)	(21/21)	(21/21)	(21/21)



ARDA Dashboard
R.Rocha et al (CERN)

MC AOD and NTUP Replication Status (Tier-1s)

- ***MC data replication period Feb-Aug 2007***
- ***Total data volume : 4600+ datasets, 800+Kfiles, 30.5+TB***

Export →

Import ↑

to \ from	ASGC	BNL	CERN	CNAF	FZK	LYON	NG	PIC	RAL	SARA	TRIUMF
ASGC											
BNL											
CERN											
CNAF											
FZK											
LYON											
NG											
PIC											
RAL											
NIKHEF											
TRIUMF											

95+% | 90-95% | 80-90% | 60-80% | 25-60% | <25%

Data Transfer Functional Test

- The goal is to have a regular ***system*** test
  - A coherent data transfer test between Tier-1 and Tier-2s for all *clouds*, using existing SW to generate and replicate data and to monitor data flow.
    - Data generated by ATLAS MC production SW
    - Organized in DQ2 datasets
    - Test different file size (from 0.5 to 4.5GB)
- We consider functional test as a part of ATLAS DDM operations activity
- 4 Tests have been conducted since Sep 2006
- All Tier-1s and 40+ Tier-2s (Tier-3s) are participated in tests

DDM Functional Test 2006. Summary Table

Tier1	Tier2s	Sep		Oct		Nov	
ASGC	IPAS, Uni Melbourne		Failed within the cloud		Failed for Melbourne		T1-T1 not testd
BNL	GLT2, NET2,MWT2,SET2, WT2		done		done		2+GB & DPM
CNAF	LNF,Milano,Napoli,Roma1		65% failure rate		done		
FZK	CSCS, CYF, DESY-ZN, DESY-HH, FZU, *WUP*		Failed from T2 to FZK		dCache problem		T1-T1 not testd
LYON	BEIIJING, CPPM, LAPP, LPC, LPHNE, SACLAY, TOKYO		done		done, FTS conn =< 6		
NG			Not tested		not tested		T0-T1,T1/T1
PIC	IFAE, IFIC, UAM		Failed within the cloud		done		
RAL	CAM, EDINBOURGH, GLASGOW, LANCS, MANC, QMUL		Failed within the cloud		Failed for Edinbrg.		done
SARA	*IHEP*, ITEP, SINP		Failed		IHEP not tested		IHEP in progress
TRIUMF	ALBERTA, TORONTO, UniMontreal, SFU, UVIC		Failed within the cloud		Failed		T1-T1 not testd

New version of DDM SW 0.2

The Most Critical Problems identified after 2006 Functional Tests

- Lack of monitoring
  - *Solved in 2007* : ARDA monitoring + DDM operations dedicated set of programs
- Instability of DDM/DQ2 central DB
  - *Improved in 2007* : New DQ2 version (0.3) and new DB backend (ORACLE)
- Poor LFC performance – 12 Hz
  - *Solved in 2007* : Bulk methods implementation current performance ~400 Hz
- Site services stability and subscription processing
  - Not solved yet.

Data Transfer Functional Tests

(Aug 2007)

Tests Scope :

- Data transfer from CERN to Tiers for datasets with average file size 0.5GB and 4GB. This step simulated the flow of ATLAS data from CERN to the sites
  - Step 1 : Data transfer from CERN to Tiers
    - a : Data transfer from Tier-0 to Tier-1
    - b : Data replication within clouds (Tier-1 to Tier-2s)
  - Step 2 : MC production data flow simulation. Data transfer from Tier-2s to Tier-1.
- Data transfer between regional centers
  - Tier-1/Tier-1 data transfer
  - from Tier-2 to Tier-1 in another cloud
  - from Tier-1 to Tier-2 in another cloud

Functional Test Aug 2007

Datasets Replication Status within *clouds*

Panda monitor and browser - Mozilla Firefox

http://gridui03.usatlas.bnl.gov:25880/server/pandamon/query?mode=listFunctionalTests&testType=T1toT2s

RAL

Dataset	Source Host	Files on Source Host	Number of CpFiles on CAM	GLASGOW	LANCS	MANC	OXF	QMUL
ddmf4.001036.RAL.CAM.Step1.v00041101	RALDISK	31	31				31	
ddmf4.001040.RAL.QMUL.Step1.v00041101	RALDISK	30						30
ddmf4.001033.RAL.GLOBAL.Step1.v00041101	RALDISK	30	30	30	30	30	30	30
ddmf4.001035.RAL.MANC.Step1.v00041101	RALDISK	30				30		
ddmf4.001039.RAL.GLASGOW.Step1.v00041101	RALDISK	30		30				
ddmf4.001034.RAL.LANCS.Step1.v00041101	RALDISK	30			30			

SARA

Dataset	Source Host	Files on Source Host	Number of CpFiles on IHEP	ITEP	JINR	PNPI	SINP
ddmf4.001042.SARA.ITEP.Step1.v00041101	SARADISK	31		31	24		
ddmf4.001043.SARA.IHEP.Step1.v00041101	SARADISK	29	30				
ddmf4.001096.SARA.Global.Step1.v00041101	SARADISK	20	0	2	0	0	0
ddmf4.001033.RAL.GLOBAL.Step1.v00041101	RALDISK	30	30		30		
ddmf4.001044.SARA.SINP.Step1.v00041101	SARADISK	29				29	0
ddmf4.001001.BNL.GLOBAL.Step1.v00041101	SARADISK	45	45	45	39		
ddmf4.001041.SARA.GLOBAL.Step1.v00041101	SARADISK	30	30	30	22	27	0

TRIUMF

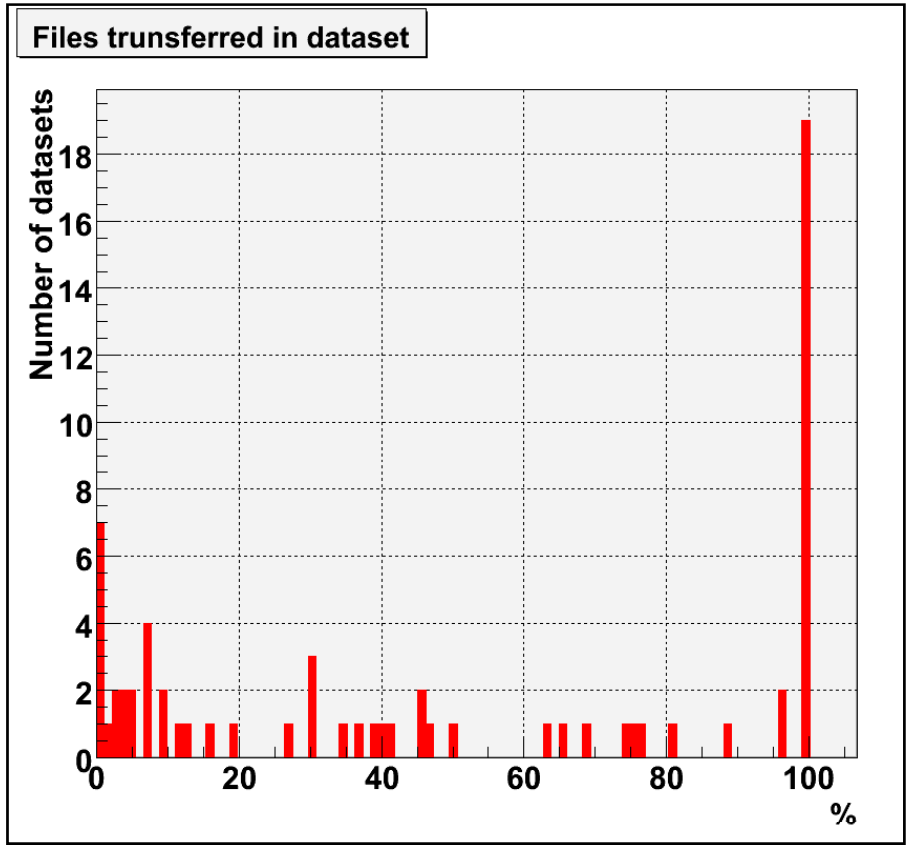
Dataset	Source Host	Files on Source Host	Number of CpFiles on ALBERTA	MONTREAL	SFU	TORON	UVIC
ddmf4.001049.TRIUMF.SFU.Step1.v00041101	TRIUMFDISK	30			0		
ddmf4.001050.TRIUMF.TORONTO.Step1.v00041101	TRIUMFDISK	30				30	
ddmf4.001001.BNL.GLOBAL.Step1.v00041101	BNLDISK	45	45				
ddmf4.001098.TRIUMF.Global.Step1.v00041101	TRIUMFDISK	19	19		0	19	0
ddmf4.001002.BNL.GLT2.Step1.v00041101	CERNCAF	52	52				
ddmf4.001053.TRIUMF.ALBERTA.Step1.v00041101	TRIUMFDISK	32	32				

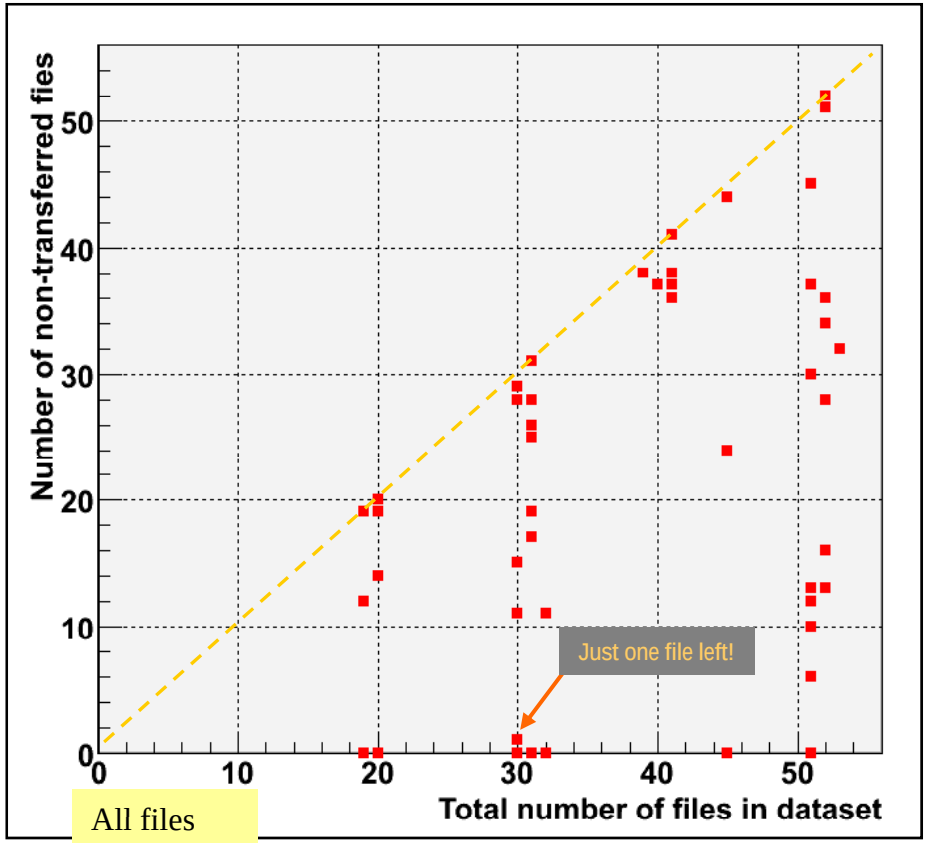
Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)

10 datasets – no files transferred (not analyzed)

66 datasets with at least one file transferred

~ 25 % of test datasets were completed

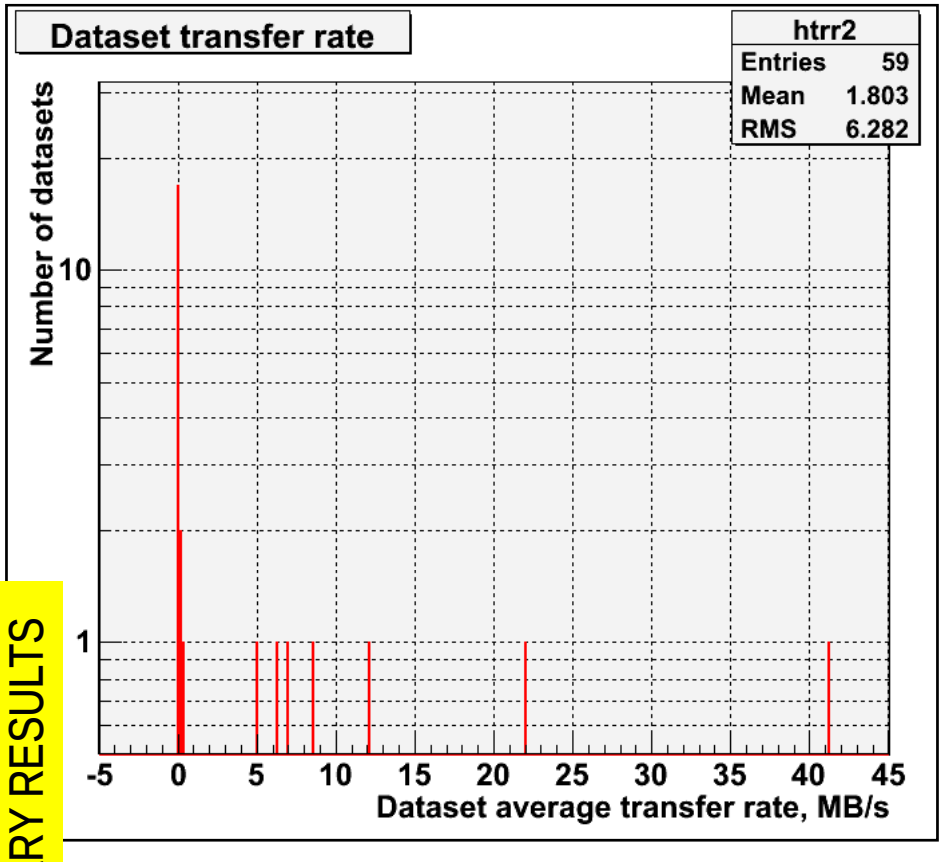


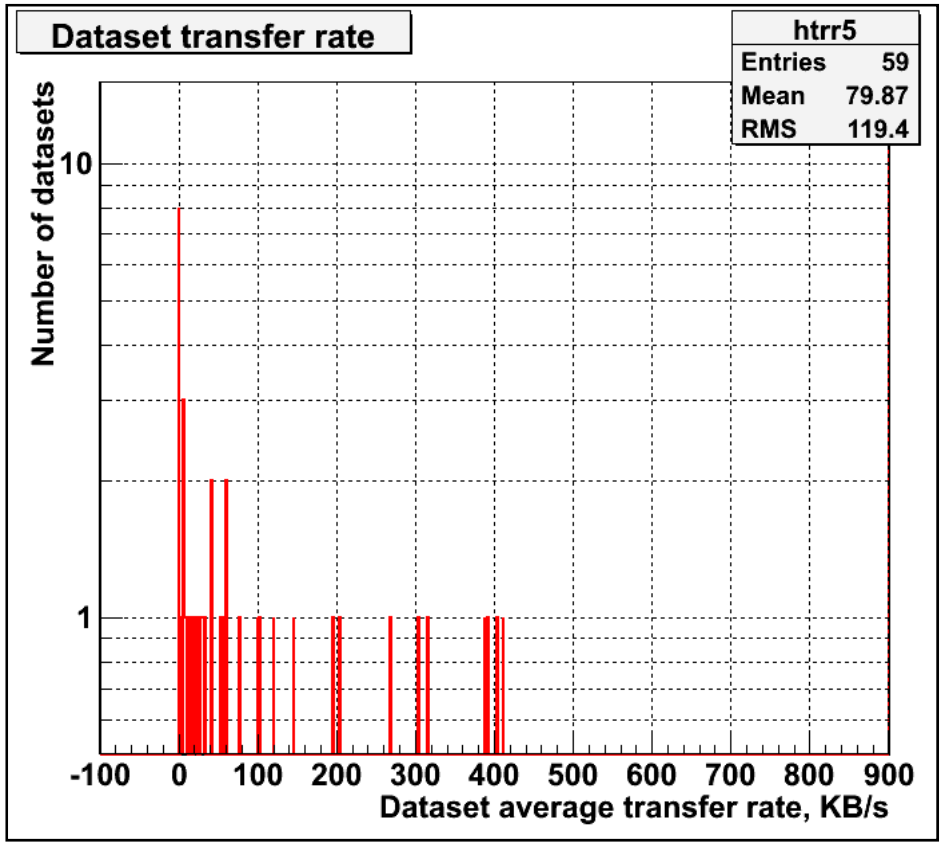


All files transferred

Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)

Average transfer rate imply software and hardware induced delays and the impact of the over stability of the system!



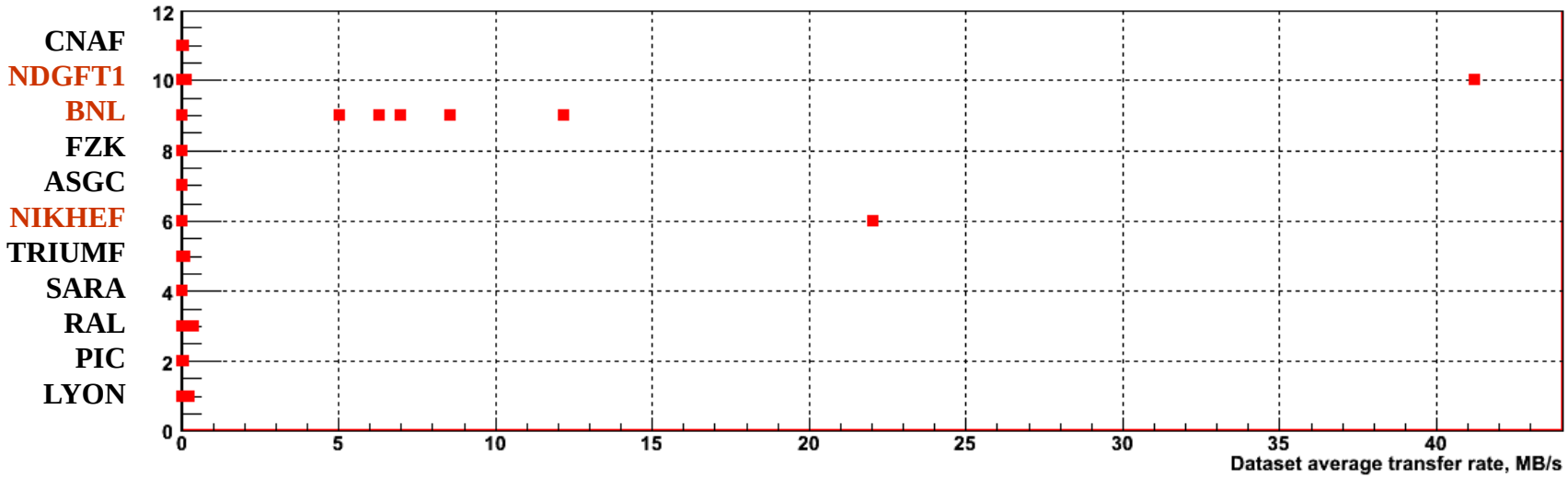


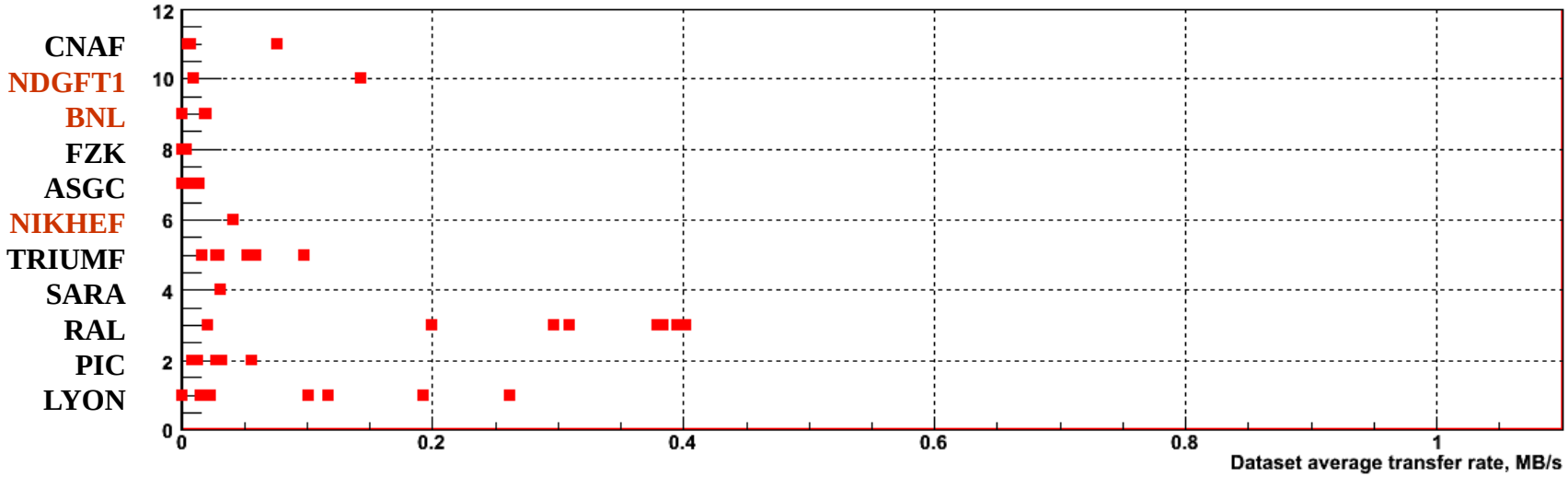
< 1 MB/s region

PRELIMINARY RESULTS

Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)

Average transfer rate imply software and hardware induced delays and the impact of the over stability of the system!



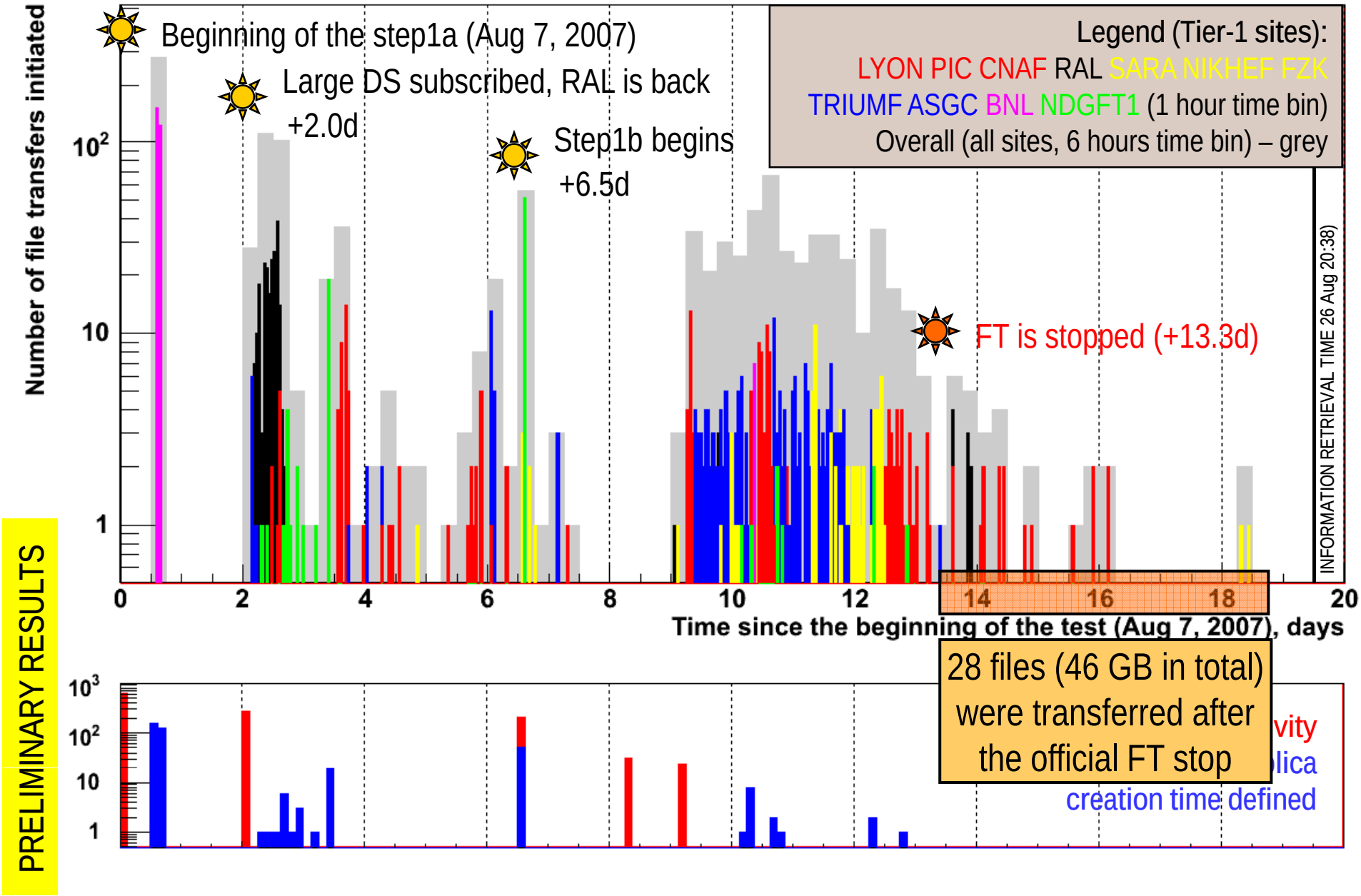


PRELIMINARY RESULTS

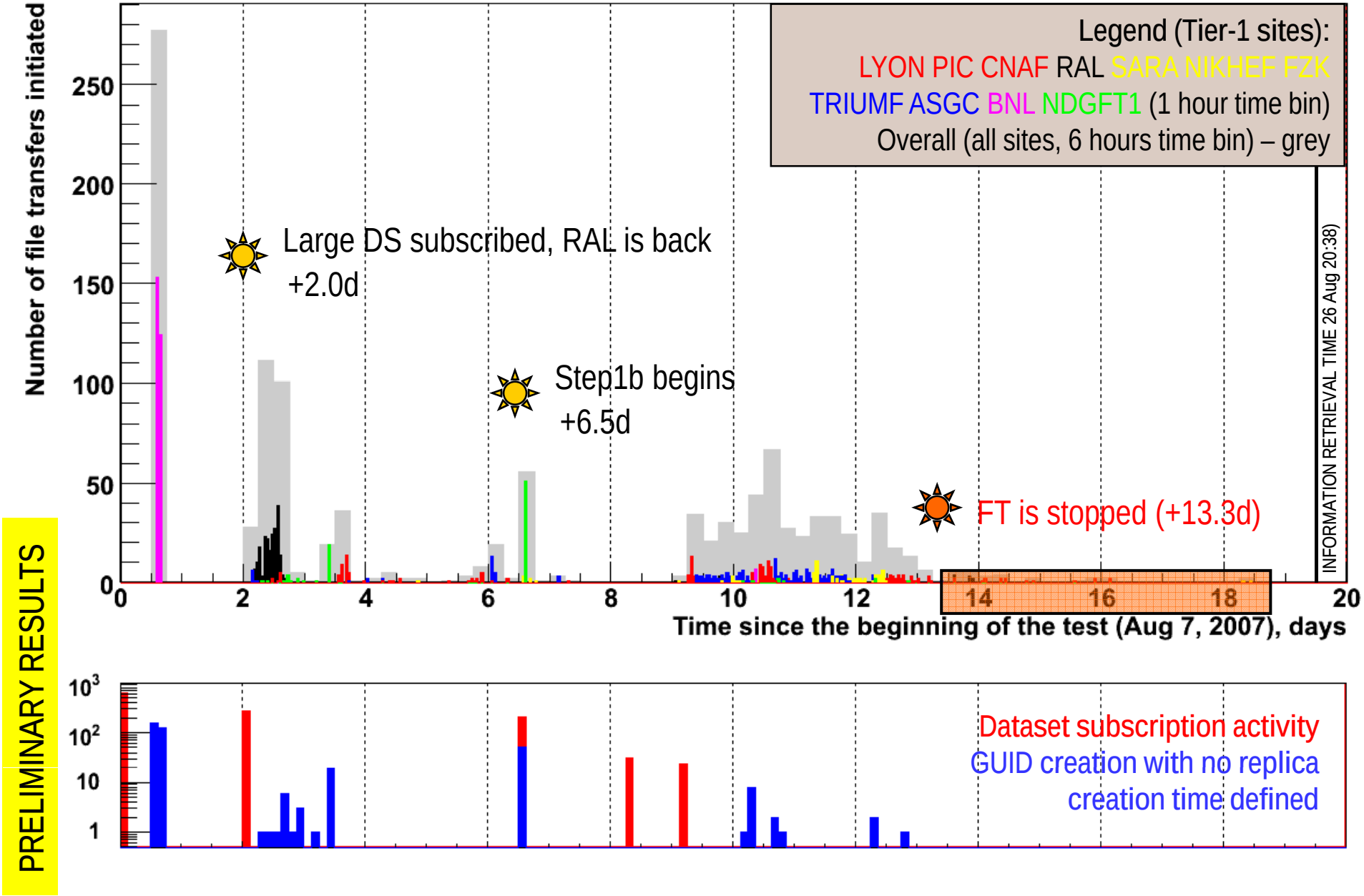
DDM/DQ2 Functional Tests (Aug 2007)

- No visible correlation for:
  - ✓ Data set transfer rate and total dataset transfer time
    vs
    number of files in dataset, size of the dataset, average file size within the dataset, initial delay between the subscription and time of transferring the first file;
  - ✓ initial delay between the subscription
    vs
    time of transferring the first file and time of subscription;
  - ✓ etc.

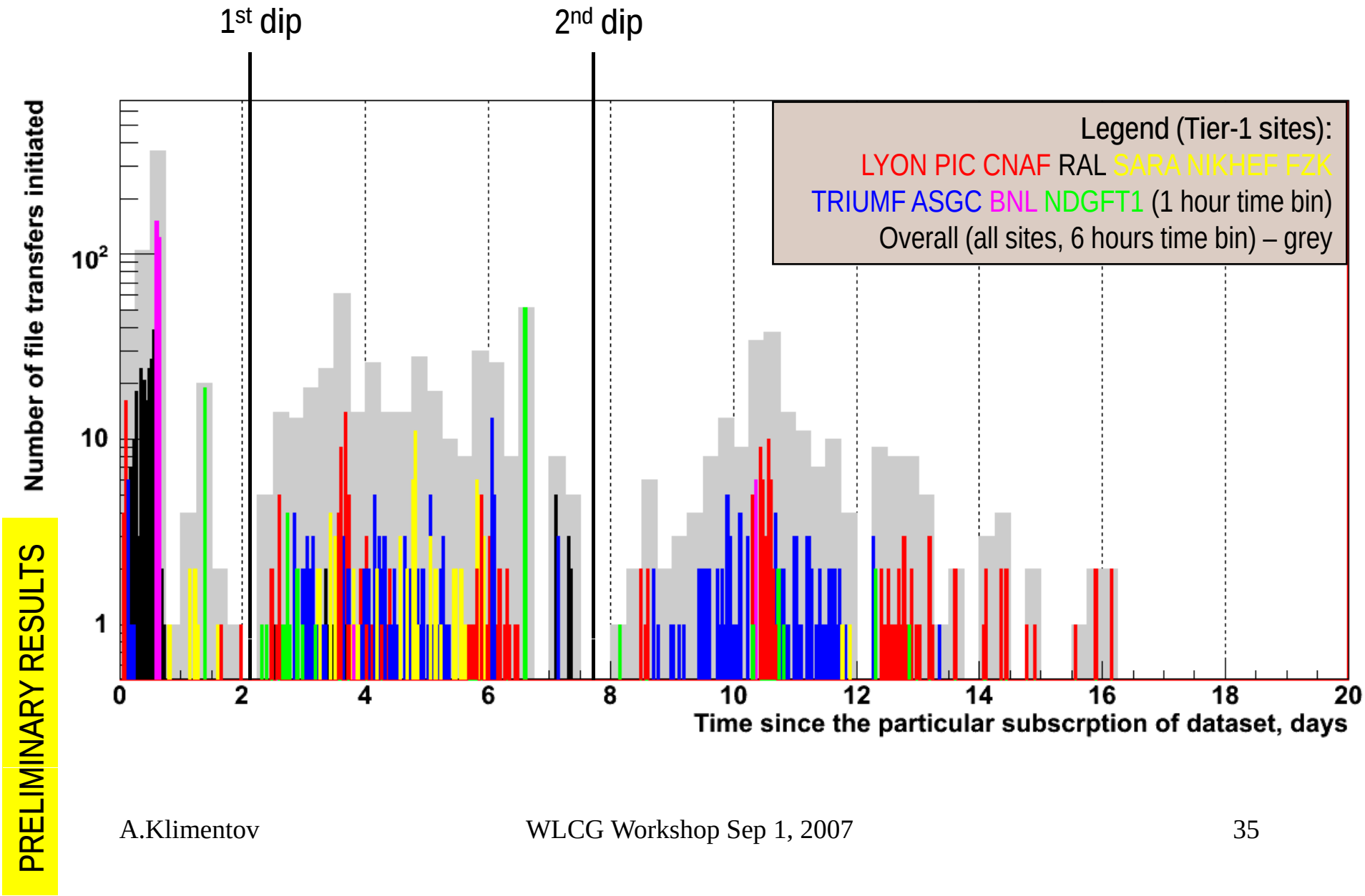
FT Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)



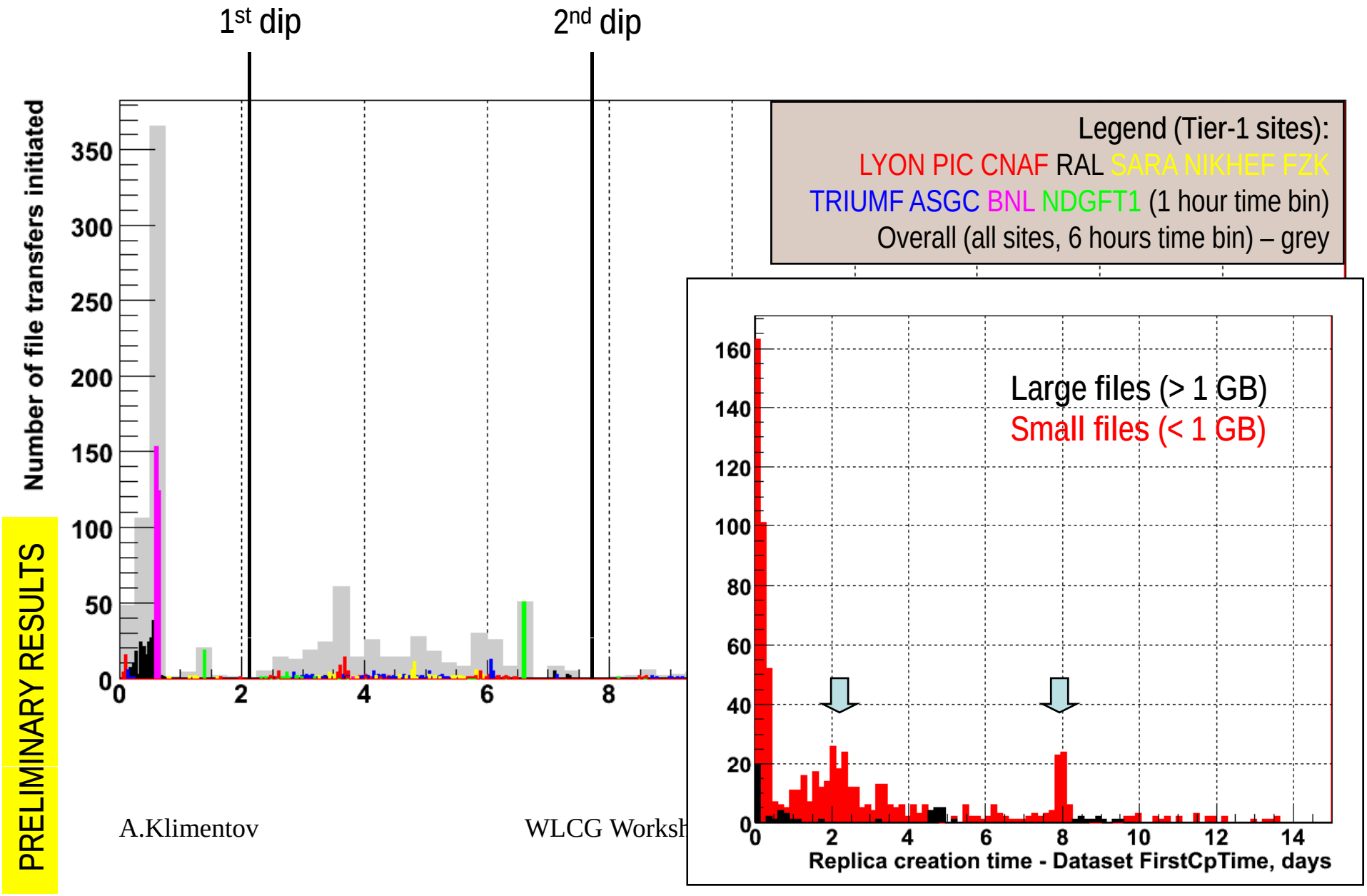
FT Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)



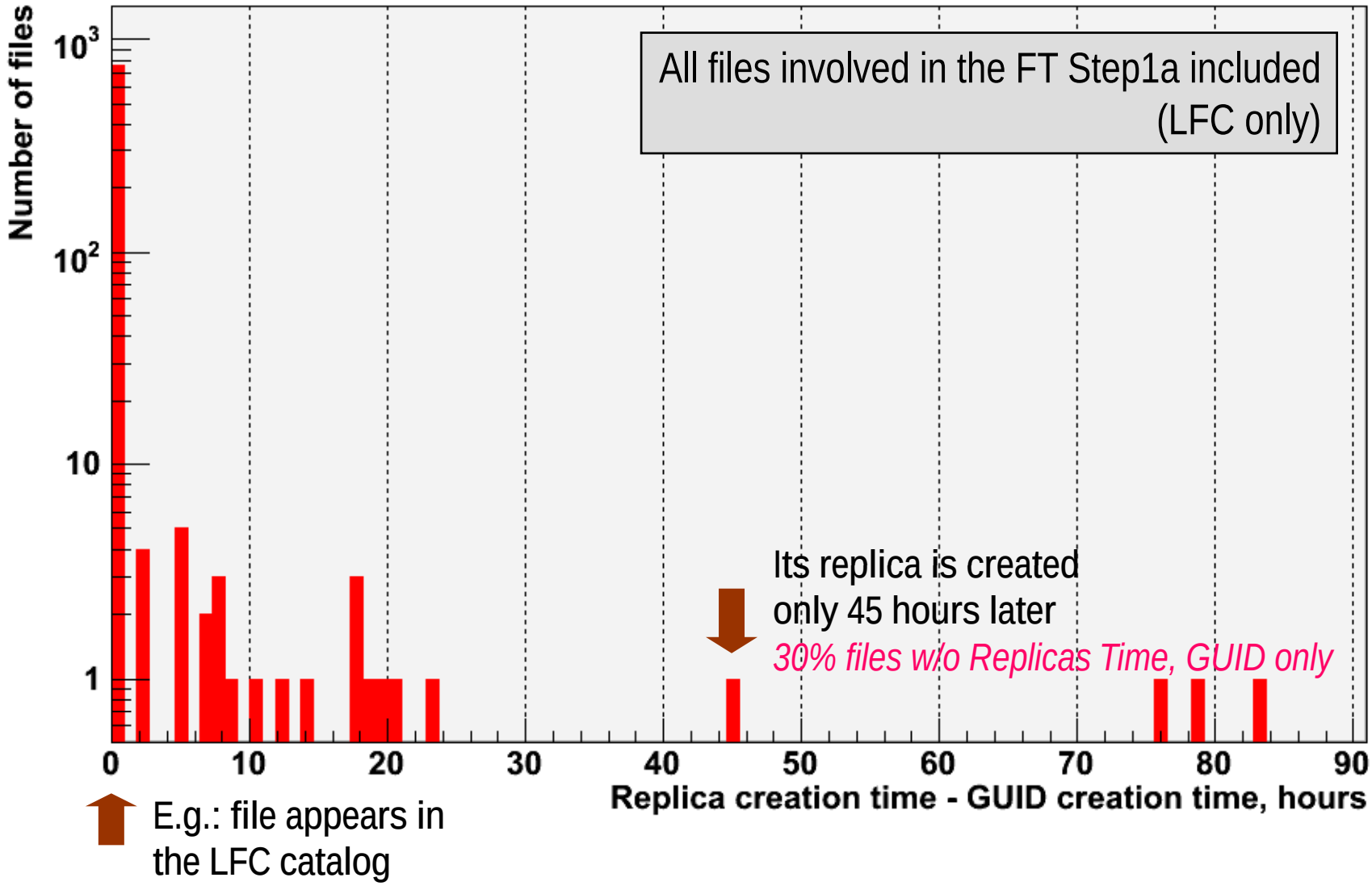
FT Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)



FT Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)



FT Step 1a (Tier0-Tier1s: 76 DS, 1334 CpFiles)



PRELIMINARY RESULTS

DDM/DQ2 Functional Tests (Aug 2007)

- Functional test status summary:
  - ✓ FT was conducted for 14 days
  - ✓ Data analysis tools were developed,
  - ✓ Step1a analysis preliminary results are available,
  - ✓ Step1b (Tier2s << | >> Tier1s) analysis is being performed
  - ✓ Several crashes of sites services have been observed.
  - ✓ Most of the problems have been fixed/understood by DQ2 developers within hours.
  - ✓ ARDA monitoring is improved since last FT. Only one outages happened (24h) during 14 days of running. Some checks still can be done only on the level of site services logs.
  - ✓ NO problem with central services
  - ✓ Number of 'active' subscriptions after certain level caused performance degradation
  - ✓ Number of files DDM/DQ2 has to treat creates the limitation on the transfer speed in TB.

➢ ATLAS Cosmic Run (M4) is on going. We apply the same technique to analyze M4 data transfer

Cosmic data replication within clouds (new replication policy)

History Bookmarks Tools Help

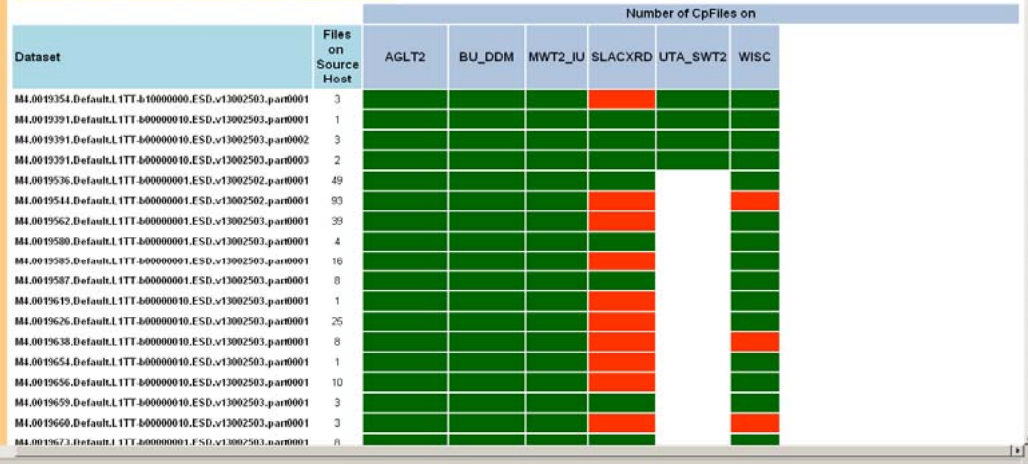
http://gridui03.usatlas.bnl.gov:25880/server/pandamon/query?mode=listM4#LYON

Free Hotmail | indico | Windows Media | Windows | BeOnLine: связь, ко... | Windows

Tier2	Datasets	Total Files in datasets	Total CpFiles in datasets	Completed	Transfer	Subscribed
EIJING	50	370	30	4	10	36
CPPM	11	73	7	1	2	8
LAL	98	857	60	1	39	58
LAPP	11	184	25	0	5	6
LPC	21	159	23	3	7	11
LPNHE	98	1205	101	0	36	62
IPNE_02	11	65	16	1	5	5
IPNE_07	11	112	18	5	3	3
ACLAY	98	547	28	0	18	80
TOKYO	291	3768	156	0	65	226

		in datasets	in datasets			
AGLT2	652	7015	7015	652	0	0
BU_DDM	652	7015	6955	639	13	0
MWT2_IU	652	7015	7015	652	0	0
SLACXRD	651	7004	2334	261	4	386
UTA_SWT2	102	848	848	102	0	0
WISC	652	7015	5263	462	0	190

100% of data requested by 5 T2s



ataset	Files on Source Host	BEIJING	CPPM	LAL	LAPP	LPC	LPNHE	NIPNE_02	NIPNE_07	SACLAY	TOKYO
4.0019354.Default.L1TT-b10000000.ESD.v13002503.part0001	3		1								
4.0019391.Default.L1TT-b00000010.ESD.v13002503.part0001	1										
4.0019391.Default.L1TT-b00000010.ESD.v13002503.part0002	3										
4.0019391.Default.L1TT-b00000010.ESD.v13002503.part0003	2										
4.0019536.Default.L1TT-b00000001.ESD.v13002502.part0001	49				8						
4.0019544.Default.L1TT-b00000001.ESD.v13002502.part0001	93				14						
4.0019562.Default.L1TT-b00000001.ESD.v13002503.part0001	39					8			8		
4.0019580.Default.L1TT-b00000001.ESD.v13002503.part0001	4			1				1			
4.0019585.Default.L1TT-b00000001.ESD.v13002503.part0001	16			4				4			
4.0019587.Default.L1TT-b00000001.ESD.v13002503.part0001	8			1					1		
4.0019619.Default.L1TT-b00000010.ESD.v13002503.part0001	1										
4.0019626.Default.L1TT-b00000010.ESD.v13002503.part0001	25			2					2		
4.0019638.Default.L1TT-b00000010.ESD.v13002503.part0001	8			1							
4.0019654.Default.L1TT-b00000010.ESD.v13002503.part0001	1										
4.0019656.Default.L1TT-b00000010.ESD.v13002503.part0001	10			2							
4.0019659.Default.L1TT-b00000010.ESD.v13002503.part0001	3					2			2		

100% of data shared by T2s within the cloud

Summary.

- ATLAS DDM Operations team successfully works for more than a 1 year in a close contact other ATLAS groups, especially with DQ2 developers
- Currently support from developers is essential, especially for non-trivial cases with DDM SW
- A set of monitoring and control tools are developed and integrated into ATLAS Operations work
- The set of Functional Tests and T0 export tests addressed and helped to improve two important issues
  - System performance
  - System functionality

Summary (cont).

Now on the agenda

- Critical data replication
- Operational model at Data Taking .
- system scalability and subscriptions processing
  - Many datasets with small number of files vs datasets with 1K files
  - 10K (and more) files are not well managed by existing GRID tools and SW .
  - Number of opened subscriptions after certain level causes performance degradation
- Site services stability
- Number of files DDM/DQ2 has to treat creates the limitation on the transfer speed in TB.
  - We need to do both : increase file size and to decrease number of files
- Integration of sites, FTS, etc monitoring with ARDA data transfer monitoring

Backup Slides

Event Data Model

LHC parameters and trigger rate

- E : 14 TeV (two 7 TeV proton beams)
- L
  - $0.5*10^{33}$ $cm^{-2}s^{-1}$ in 2007
  - $2.0*10^{33}$ $cm^{-2}s^{-1}$ in 2008-2009
  - 10^{34} $cm^{-2}s^{-1}$ in 2007

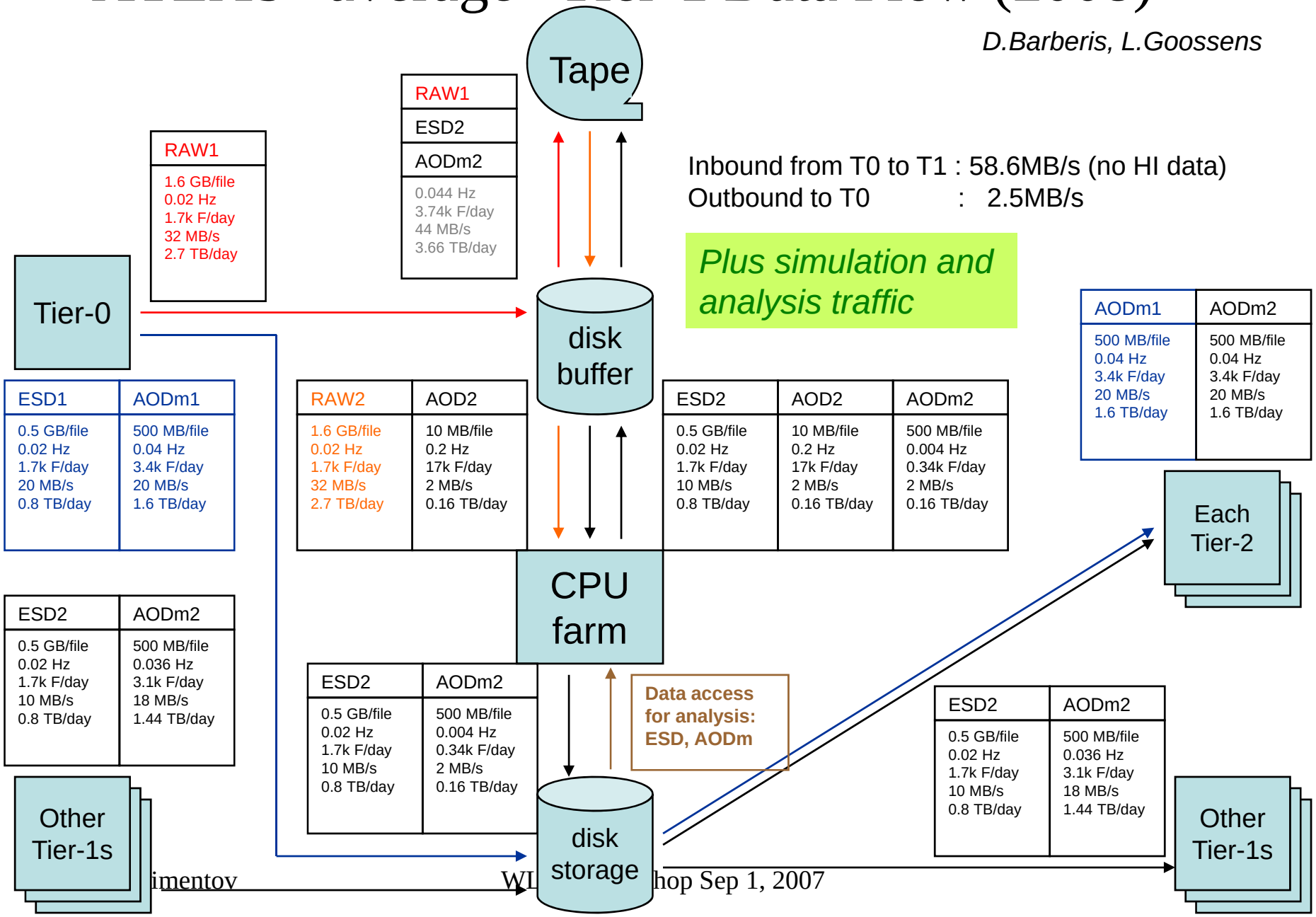
  σ : 100 mb
- Collisions : $L*\sigma = 10^9$ Hz p-p at designed luminosity
- Trigger rate : 200 Hz

Inbound from T0 to T1 : 58.6MB/s (no HI data)
Outbound to T0 : 2.5MB/s

Data Type	Size (KB)	Accessibility
Raw Data Objects	1600	Tier 0/1s
Event Summary Data	500	Tier-0/Tier-1s
Analysis Object Data	100	Tier-0, Tier1/2 At least 1 complete copy per cloud. Tier-3 - subset
Derived Physics Data	10	Tier-3s (eg your laptop)
TAG	1	All Tiers

ATLAS “average” Tier-1 Data Flow (2008)

D.Barberis, L.Goossens



More about datasets status and availability

Configuration | Dashboards: Production DDM AutoPilot Sites & Grids Analysis Physics data Usage & Quotas Plots ArdaDash

Update

anda monitor

uick guide, twiki

er info

s - search
cent running,
vated, waiting,
igned, defined,
shed, failed jobs
ect analysis,
duction, test jobs
ck search

set

mmaries
cks: days
rs: days
es: days
ly usage

sks - search
neric Task Req
Gen Task Req
Bsim Task Req
k list
k browser

tasets - search
aset browser
w datasets
orted MC datasets
nda subscriptions
subscriptions

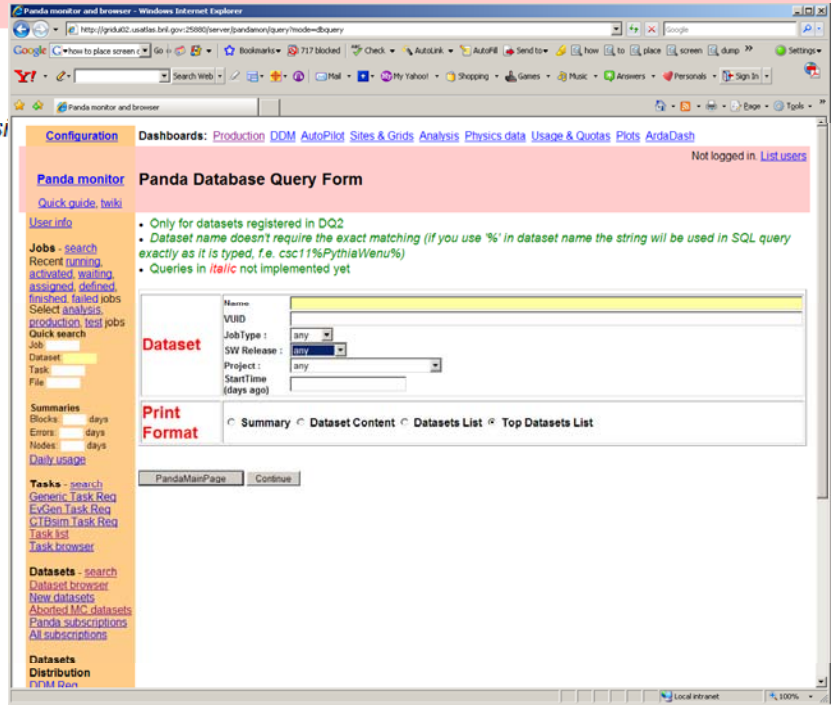
DQ2 dataset browser

Click for help
Dataset lists last updated 407 min ago

Select a project: Go

Or (the old way) select a dataset category *Counts are totals, exclusi*

Category	Count	Description
All	150419	All datasets
T0	0	Tier 0 test
conditions	29	Datasets for conditions data files
csc	4389	Computing system commissioning production
ctb	88	Combined testbeam production
dc2	6	Data Challenge 2 production
destination	42340	Panda destination sub-blocks
dispatch	25708	Panda dispatch blocks
larg	53	LAr commissioning
mc	4308	MC validation production
other	40934	Everything else
rome	210	Rome physics workshop production
testpanda	1315	Panda test datasets
tile	52	Tilecal commissioning
user	21144	User datasets
validation	642	Validation samples (testIdeal* etc)



Choose a site if you want to restrict dataset listings to site-resident datasets

CANADA	CERN	FRANCE	GERMANY	ITALY	NL	SPAIN	TAIWAN	UK	US
ALBERTA	CERNCAF	BEIJING	CSCS	CNAF	IHEP	IFAE	ASGC	RAL	BNI
MONTREAL	CERNINTEGRATIONTEST1	CPPM	CYF	CNAFDISK	ITEP	IFIC	ASGCDISK	RALDISK	BN
SFU	CERNPROD	LAL	DESY-HH	CNAFTAPE	JINR	IFICDISK	ASGCTAPE	RALTAPE	BN
TORON	TIER0DISK	LAPP	DESY-ZN	LNF	NIKHEF	IFICTAPE	AU-UNIMELB	UKTIER2S	BN
TRIUMF	TIER0TAPE	LPC	FZK	MILANO	PNPI	PIC	TW-FTT	EDINBURGH	BN
TRIUMFDISK		LPNHE	FZKDISK	NAPOLI	SARA	PICDISK		GLASGOW	BN
TRIUMFTAPE		LYON	FZKTAPE	ROMA1	SARADISK	PICTAPE		MANC	BNI
UVIC		LYONDISK	FZU		SARATAPE	UAM		LANCS	BU
		LYONTAPE	LRZ		SINP			LIV	BU

Local intranet 100%

Tasks are aborted, or data is obsolete
the list is defined by Physics Coordinator
Data deletion is sites responsibility
DDM operations deletes datasets from central catalog
and maintains the list of deleted tasks/datasets

ration
te
onitor
e, twiki

Dashboards: Production DDM AutoPil

Aborted Datasets

List of the datasets is generated using

- ***log datasets*** are not listed, but they will be deleted. Datasets of empty and test tasks are not listed
- ***Timestamp*** indicates time when dataset/task info was updated
- ***Status :***
- aborted/failed - task is aborted/failed and dataset will be deleted w/o notification
- deleted - dataset is deleted from DQ2 catalog
- toBeErased - dataset will be deleted soon
- waitErased - dataset will be deleted after timestamp + 1 week (sites notified)

It is sites responsibilities to delete files from SE, Files Catalog entries and stop subscriptions.The recommended procedure for sites , for LCG sites (TBD)

Only recent 500 datasets are listed. The whole list is here

Dataset	Task ID	Status	Timestamp
trig1_misal1_csc11.005862.A10_Atautau_filter.recon.ESD.v12000601_tid0010494	10494	waitErased	Jun 15 16:16:28
trig1_misal1_csc11.005862.A10_Atautau_filter.recon.AOD.v12000601_tid0010494	10494	waitErased	Jun 15 16:16:28
trig1_misal1_csc11.005862.A10_Atautau_filter.recon.NTUP.v12000601_tid0010494	10494	waitErased	Jun 15 16:16:28
misal1_csc11.005862.A10_Atautau_filter.digit.RDO.v12003103_tid0010493	10493	waitErased	Jun 15 16:16:28
misal1_csc11.005862.A10_Atautau_filter.simul.HITS.v12003103_tid0010493	10493	waitErased	Jun 15 16:16:28
trig1_misal1_mc12.006346.A10_600_Atautau_filter.recon.ESD.v12000604_tid0010439	10439	waitErased	Jun 15 16:16:28
trig1_misal1_mc12.006346.A10_600_Atautau_filter.recon.AOD.v12000604_tid0010439	10439	waitErased	Jun 15 16:16:28
trig1_misal1_mc12.006346.A10_600_Atautau_filter.recon.NTUP.v12000604_tid0010439	10439	waitErased	Jun 15 16:16:28
misal1_mc12.008097.PythiaPhotonJet3_FIXED.digit.RDO.v12003103_tid0010368	10368	waitErased	Jun 15 16:16:28
misal1_mc12.008097.PythiaPhotonJet3_FIXED.simul.HITS.v12003103_tid0010368	10368	waitErased	Jun 15 16:16:28
mc12.007263.singlepart_mu_p5.evgen.EVNT.v12000604_tid0010365	10365	waitErased	Jun 15 16:16:28

Panda monitor and browser - Windows Internet Explorer

http://gridui02.usatlas.bnl.gov:25880/server/pandamon/query?mode=listRDOReplications

Configuration
Update
Panda monitor
Quick guide, twiki
User info

Jobs - search
Recent running, activated, waiting, assigned, defined, finished, failed jobs
Select analysis, production, test jobs
Quick search
Job
Dataset
Task
File

Summaries
Blocks: days
Errors: days
Nodes: days
Daily usage

Tasks - search
Generic Task Req
EvGen Task Req
CTBsim Task Req
Task list
Task browser

Datasets - search
Dataset browser
New datasets
Aborted MC datasets
Panda subscriptions
All subscriptions

Datasets Distribution

Dashboards: Production DDM AutoPilot Sites & Grids Analysis Physics data Usage & Quotas Plots ArdaDash

RDO Datasets Replication Status

RDO datasets replication status (Mon Jun 18 11:58:21 2007 CET)

- *Datasets are automatically subscribed to CERNPROD*
- *Green - site has a complete dataset replicas (data transfer is done)*
- *Cyan - 90% files (or more are replicated).*
- *Orange - if site has an incomplete dataset replicas.It also means that subscription is processed*
- *Red - Dataset has 0 files*
- *Magenta - CERNPROD is the only registered location, but there are no files at CERN*

Comments

misal1%..%.RDO.v120031%_tid%*

5001 5002 5009
5010 5011 5012 5013 5014 5015 5016 5017
5030 5031 5032 5033 5034 5035 5036 5037
5100 5101 5103 5117 5118
5144 5145
5200 5204
5802 5806
8078
8095 8096 8097 8098 8099

List of patterns is defined by Physics Coordinator

Total Datasets : 73 . Files : 524K/430K (53.855 TB)
Last Checked : Mon Jun 18 02:52:51 2007, Last Transfer : Mon Jun 18 02:40:26 2007
RDO consolidation : 73 524K/430K 53.855TB.

Dataset	Files	Copied	GBs	From
misal1_csc11.005002.pythia_diffractive.digit.RDO.v12003108_tid005902	1971	608	53	ASGCDISK,AU-UNIMELB,TW-IPAS-T2
misal1_mc12.005802.JF17_pythia_jet_filter.digit.RDO.v12003105_tid005481	28429	14329	1810	BNLPANDA
misal1_mc12.005802.JF17_pythia_jet_filter.digit.RDO.v12003105_tid005015	15630	18662	2356	BEIJING,CPPM,LAL,LAPP,LPC,LPNHE,LYON
misal1_csc11.005117.JimmyZeeLowM_onelep.digit.RDO.v12003105_tid004947	3993	3919	508	BEIJING,LPNHE,LYONDISK,SACLAY,TOKYO
misal1_mc12.008078.PythiaPhotonJet6_FIXED.digit.RDO.v12003108_tid004847	3982	3383	434	IFAE,IFICDISK,PICDISK,PICTAPE,UAM
misal1_mc12.005204.TTbar_FullHad_McAtNlo_Jimmy.digit.RDO.v12003108_tid004787	2016	2016	300	BNLPANDA
misal1_mc12.005200.T1_McAtNlo_Jimmy.digit.RDO.v12003108_tid004783	12259	12188	1754	BNLPANDA,LYONTAPE

Done

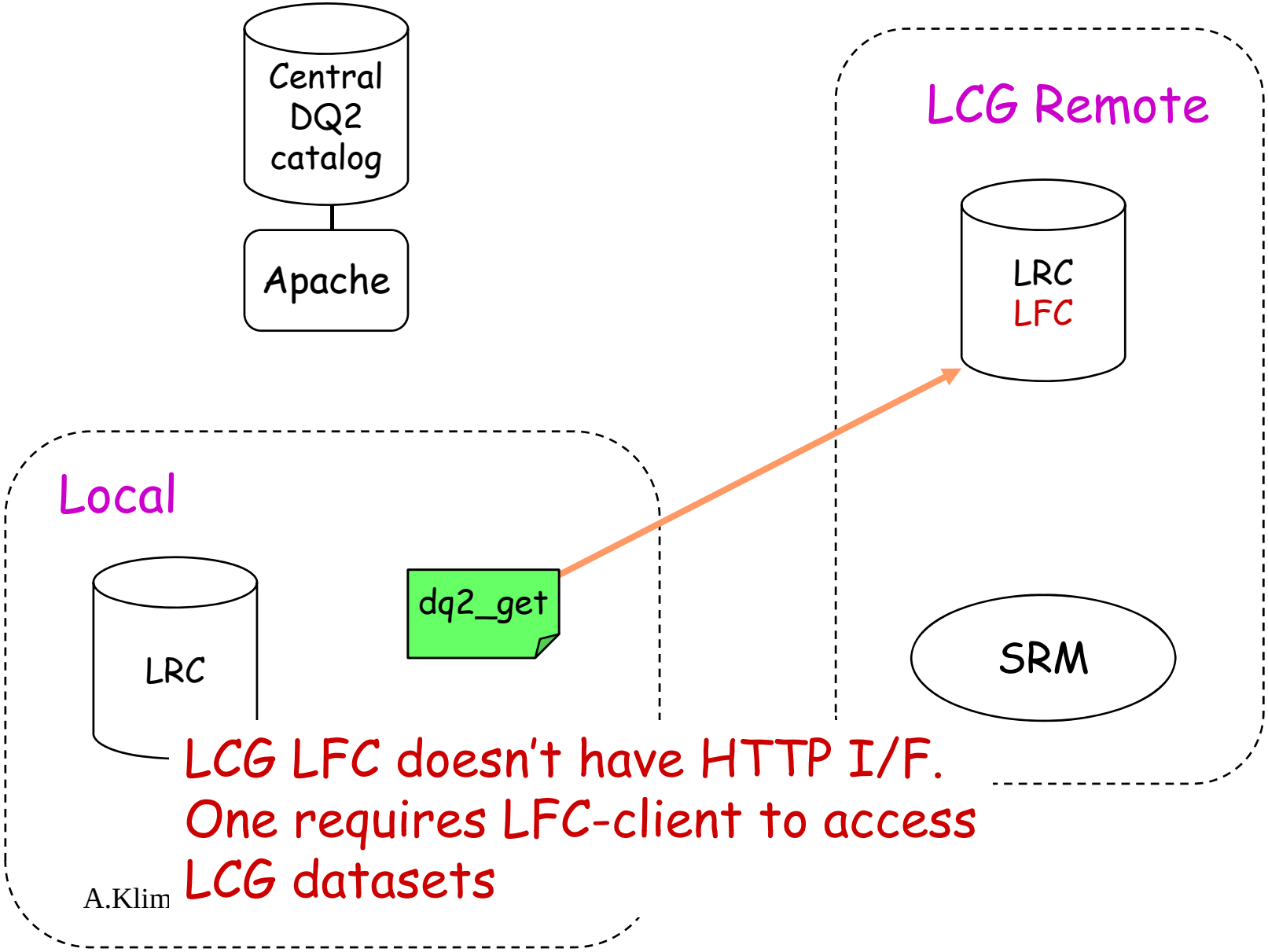
dq2 end-user tools

- Insulate flavor-specific dependences from end-users
- Provide quick accesses to DQ2 datasets for end-users
- Several tools
  - dq2_get
  - dq2_ls
  - dq2_register/dq2_put
  - dq2_cr
  - dq2_cleanup
  - dq2_sample

dq2_get

- Copy files over the grid
  - DQ2 subscriptions → asynchronous
  - dq2_get → synchronous
- Use grid tools internally
  - srmcp, globus-url-copy, lcg-cp, …
  - Users can configure dq2_get to use appropriate tool according each site policy
    (e.g., firewall setting)
- Dependence on python-bindings of LFC-client which is not included in VDT
- Doesn't register files to LRC
  - Users may copy same files many times
    → redundant traffics

Access to LCG datasets



LCG LFC doesn't have HTTP I/F.
One requires LFC-client to access LCG datasets

dq2_cr

- dq2_get + LRC registration
- dq2_cr scans LRC first, copies only missing files, and then register them in the LRC
  - Avoid duplicated transfers
- Authorized person should use it instead of end-users

dq2_ls

- dq2 provides a functionality to shows files in a dataset
  - dq2-list-files-in-dataset xyz
  - However, dq2 is a central catalog client
    $\rightarrow$ No info on LRC
- dq2_ls provides info on LRC
  - E.g., which files in a dataset a site holds
  - No corresponding functionality in dq2

dq2_register/dq2_put

- Register dq2 datasets
- dq2_register
  - Copies local files to a remote site and then registers a dataset to the site
  - Works for LCG
- dq2_put
  - Registers a dataset for local files
  - Works for LCG and OSG
- Will be merged

dq2_cleanup and dq2_sample

- dq2_cleanup
  - Delete files from an LFC and an SE
  - Works for LCG
  - Requires local-installation of mysql-python (e.g., doesn't work on lxplus)
- dq2_sample
  - Create a small sample dataset from an existing dataset

User Support

- dq2-ddm-ops Savannah

  https://savannah.cern.ch/projects/dq2-ddm-ops/

  – Each site problem
    → each site damin

  – dq2_* problems
    → developers

- Documentation

  https://uimon.cern.ch/twiki/bin/view/Atlas/UsingDQ2

DQ2 Client

- DQ2 client API:
  - Interfaces to all dataset catalogs
    - in a secure way for any ‘write’ operations
  - Guarantees consistency between all the (“loosely coupled”) dataset catalogs

    e.g. a dataset in the location catalog refers to the same dataset in the content catalog..
    - Consistency now being improve (for 0.3) with initial support for transactions, etc

M.Branco.& P.Salgado DDM WS at CERN Nov 2006

Dataset subscription using DQ2 client

Read : How to access CSC data using DQ2 end-user tools

https://twiki.cern.ch/twiki/bin/view/Atlas/UsingDQ2

I want to subscribe dataset to my Tier center using DQ2 client

My favorite physics channel : Zmumu, **Data format** : AOD, **SW version** : 12000601

(and I prefer CLI to Web Browser)

dq2-list-datasets **'*PythiaZmumu.*AOD.v12000601'**
trig1_misal1_csc11_V2.005145.PythiaZmumu.recon.AOD.v12000601

dq2-get-number-files **trig1_misal1_csc11_V2.005145.PythiaZmumu.recon.AOD.v12000601**
1962
dq2-list-dataset-replicas **trig1_misal1_csc11_V2.005145.PythiaZmumu.recon.AOD.v12000601**

INCOMPLETE: AGLT2,ASGCDISK,BNLPANDA,CNAFDISK,DESY-HH,FZKDISK,
LYONDISK,NDGFT1DISK,NIKHEF,RALDISK,TORON,TRIUMFDISK,WUP
COMPLETE:

dq2-register-subscription
trig1_misal1_csc11_V2.005145.PythiaZmumu.recon.AOD.v12000601 CERNCAF
dq2-list-subscription-info
trig1_misal1_csc11_V2.005145.PythiaZmumu.recon.AOD.v12000601 CERNCAF

Dataset subscription using DQ2 client (Cont.)

I subscribed dataset to my Tier center. How can I know the status

The easiest ways : check SE, use dq2_ls commands

Recommended : ARDA DDM monitoring

http://dashb-atlas-data.cern.ch/dashboard/request.py/site

Some subscription recommendations :

use the exact dataset location option (if you know)
–source=BNLPANDA (in previous example)
use –wait-for-sources, it helps if files or file catalog isn't available, or files are not copied yet to the 'source' of the subscription

Useful links

- DDM Operations Savannah
  https://savannah.cern.ch/projects/dq2-ddm-ops

- DDM User's Work Book (Horst Severini, Borut Kesevan)
  https://twiki.cern.ch/twiki/bin/view/Atlas/WorkBookDDM

- DDM operations FAQ
  https://twiki.cern.ch/twiki/bin/view/Atlas/DDMoperationFAQ

- DDM Operations TWiKi
  https://twiki.cern.ch/twiki/bin/view/Atlas/DDMOperationsGroup
  http://www.usatlas.bnl.gov/twiki/bin/view/Projects/ATLASDDM

- From panda browser
  - Datasets : browser, search, replication, transfer request

- DQ2 TWiKi
  https://twiki.cern.ch/twiki/bin/view/Atlas/DistributedDataManagement

- DQ2 end users tools and tutorial
  https://twiki.cern.ch/twiki/bin/view/Atlas/DDMEndUser
  https://twiki.cern.ch/twiki/bin/view/Atlas/UsingDQ2

- ARDA Monitoring
  http://dashb-atlas-data-test.cern.ch/dashboard/request.py/site

Acknowledgements

- Thanks to D.Barberis, M.Branco, A.Farbin, T.Maeno, R.Rocha, P.Salgado and T.Wenaus for their slides