



ATLAS SC4 Service phase

Distributed Data Management

Miguel Branco

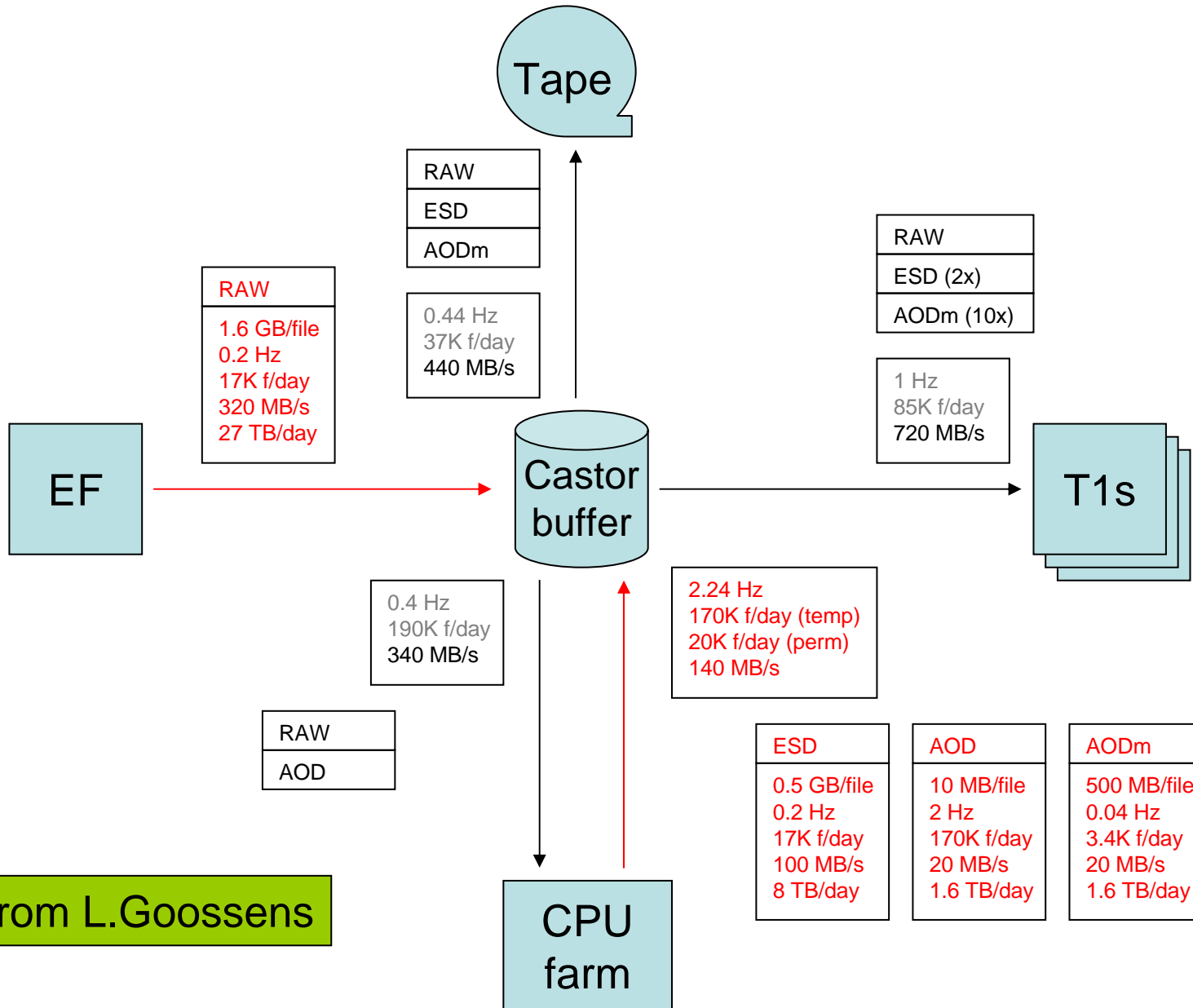
Outline

- Service phase
 - Tier0 exercise
 - DQ2
 - Status
 - Reminder of required services
- Other “SC*” issues

Service Phase

- **Main exercise: Tier0**
 - Coordinators
 - Internal Tier0: Luc Goossens; Export to T1s: DDM group
 - Repeat of last year tests
 - Involves Tier0, all Tier1s and a subset of volunteer Tier2s
 - Only goal:
 - Run complete flow @ nominal rate for all T1s
 - Very similar to SC4 throughput phase
 - But going down to T2s, proper cataloguing.. this will be new!
 - Phase 1 (19th July for 3 weeks):
 - Repetition of last year tests, with DQ2 0.2.x
 - Phase 2 (after Summer):
 - Consider 'failing T1' scenarios, cleanup of pools @ Tier0, ..

Tier0



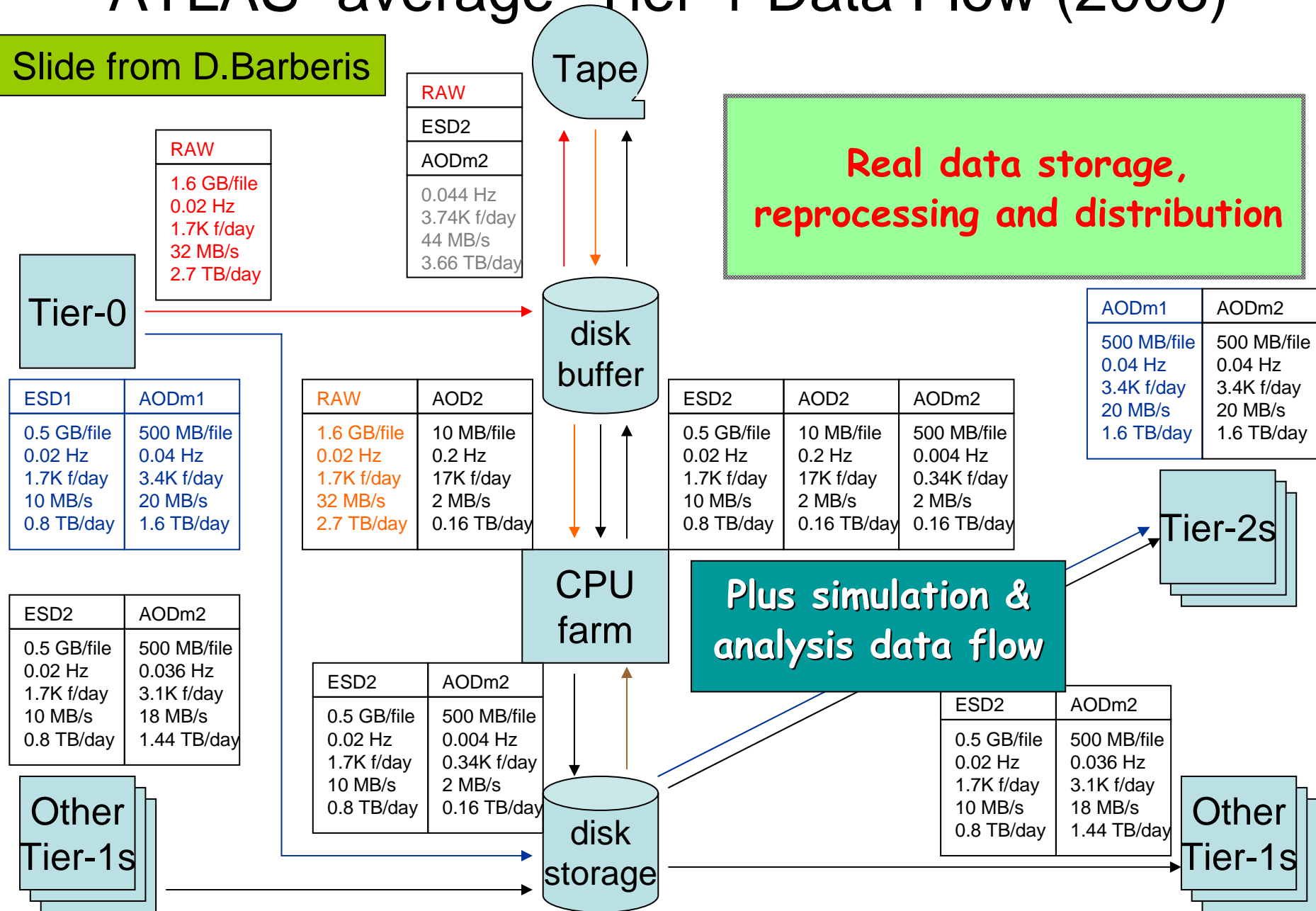
Slide from L.Goossens

Service Phase

- Remaining exercises:
 - Distributed Production
 - Distributed Analysis
 - Reprocessing
- More details to be sent out soon!!...
 - (ATLAS CSC/SC4 coordination to sent out details for remaining exercises)
 - **But** important milestone is success of **Tier0 exercise**:
 - **Validation of DQ2 and WLCG data management infrastructure**, which is now being done as part of Tier0 exercise;
 - Remaining exercises depend on this outcome..

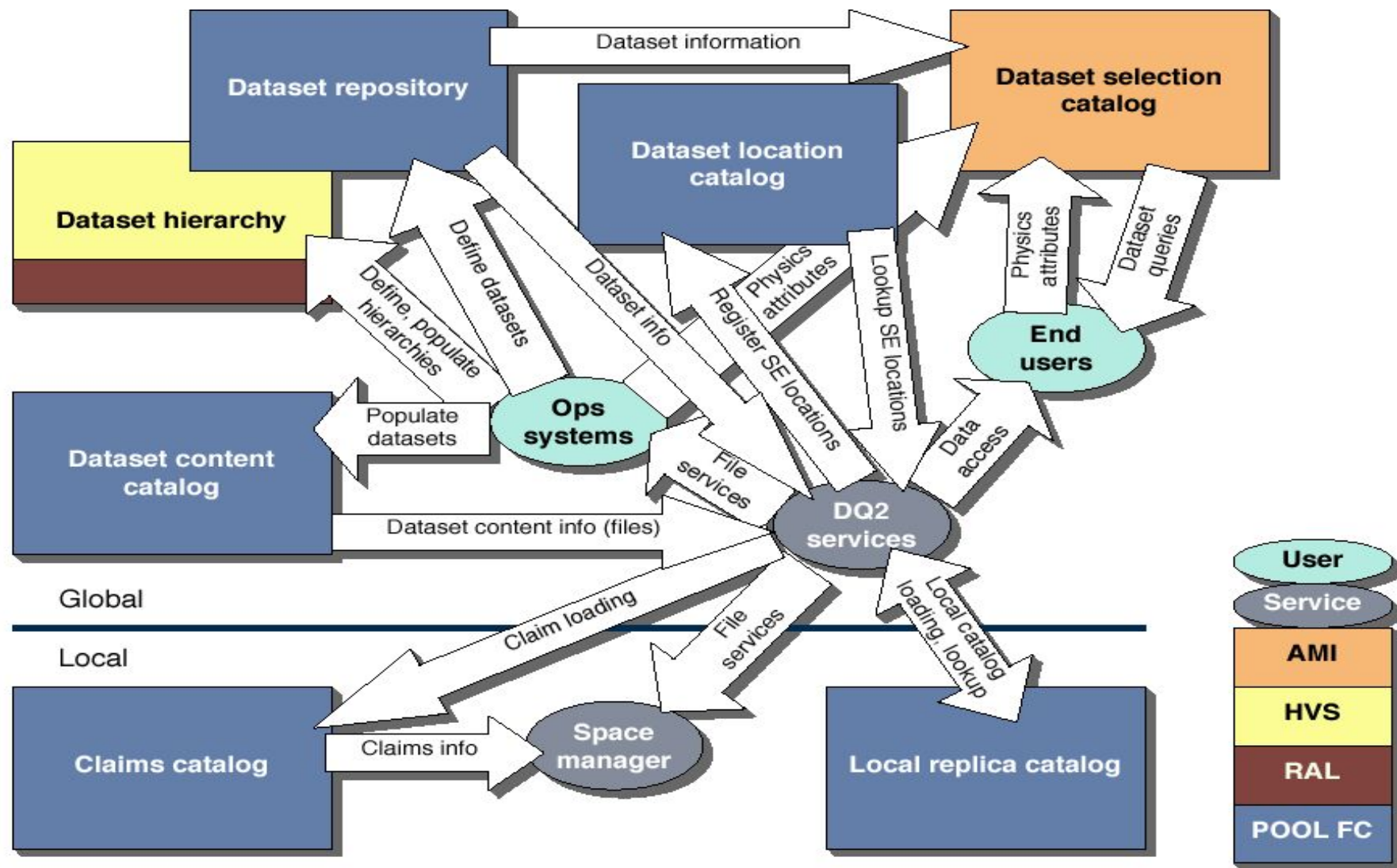
ATLAS "average" Tier-1 Data Flow (2008)

Slide from D.Barberis



DQ2

- DQ2, our Distributed Data Management system which builds on top of Grid data transfer tools, is based on:
 - A hierarchical definition of datasets
 - Central dataset catalogues
 - Data blocks as units of file storage and replication
 - Distributed file catalogues
 - Automatic data transfer mechanisms using distributed services (dataset subscription system)
- DQ2 allows the implementation of the basic ATLAS Computing Model concepts, as described in the Computing TDR (June 2005):
 - Distribution of raw and reconstructed data from CERN to the Tier-1s
 - Distribution of AODs (Analysis Object Data) to Tier-2 centres for analysis
 - Storage of simulated data (produced by Tier-2s) at Tier-1 centres for further distribution and/or processing

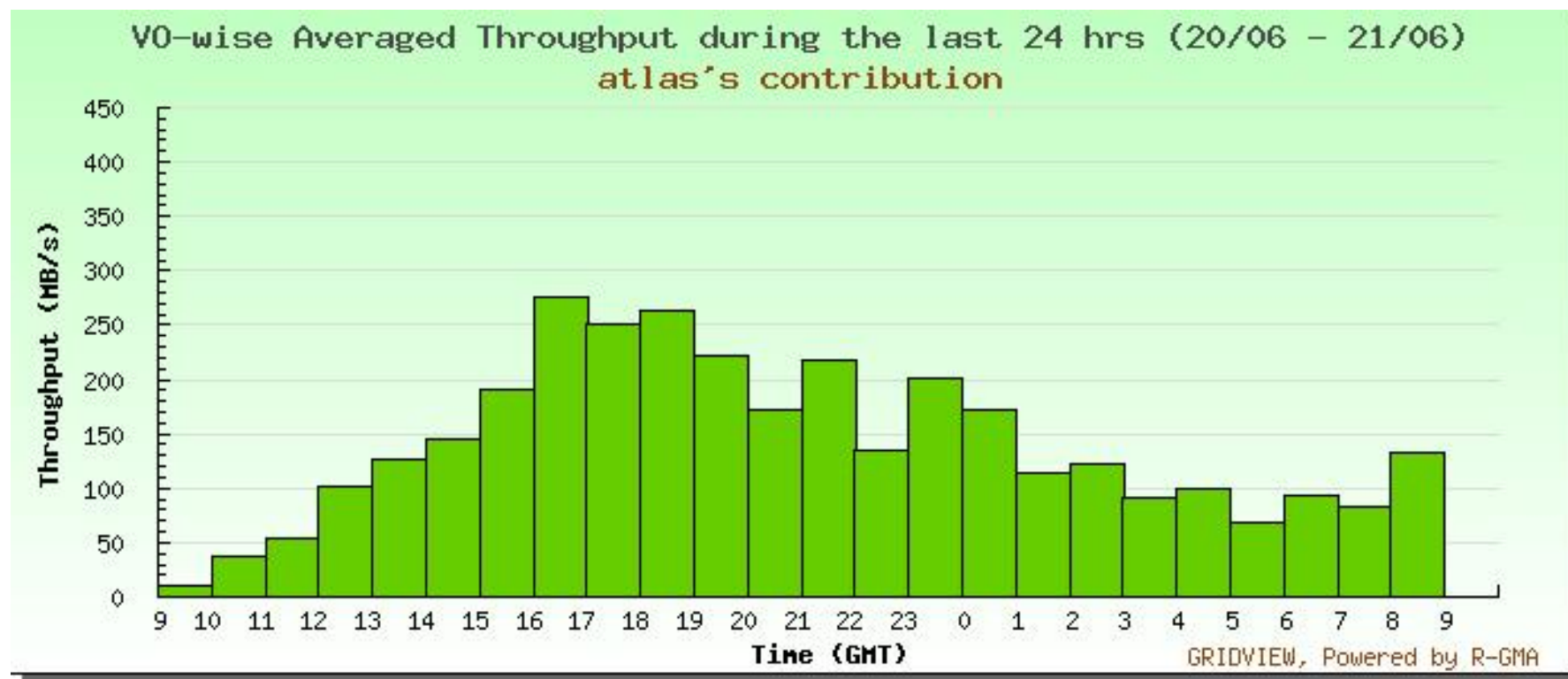


Central vs Local Services

- DQ2 has now a central role with respect to ATLAS Grid tools
- One fundamental feature is the presence of distributed file catalogues and (above all) auxiliary services
 - Clearly we cannot ask every single Grid centre to install ATLAS services
 - We decided to install “local” catalogues and services at Tier-1 centres
 - Then we defined “regions” which consist of a Tier-1 and all other Grid computing centres that:
 - Are well (network) connected to this Tier-1
 - Depend on this Tier-1 for ATLAS services (including the file catalogue)
- We believe that this architecture scales to our needs for the LHC data-taking era:
 - Moving several 10000s files/day
 - Supporting up to 100000 organized production jobs/day
 - Supporting the analysis work of >1000 active ATLAS physicists

Status: Service Phase

- ATLAS 'slot' started June 19th
- 1st ramp 'attempt' yesterday:



Status: Service Phase

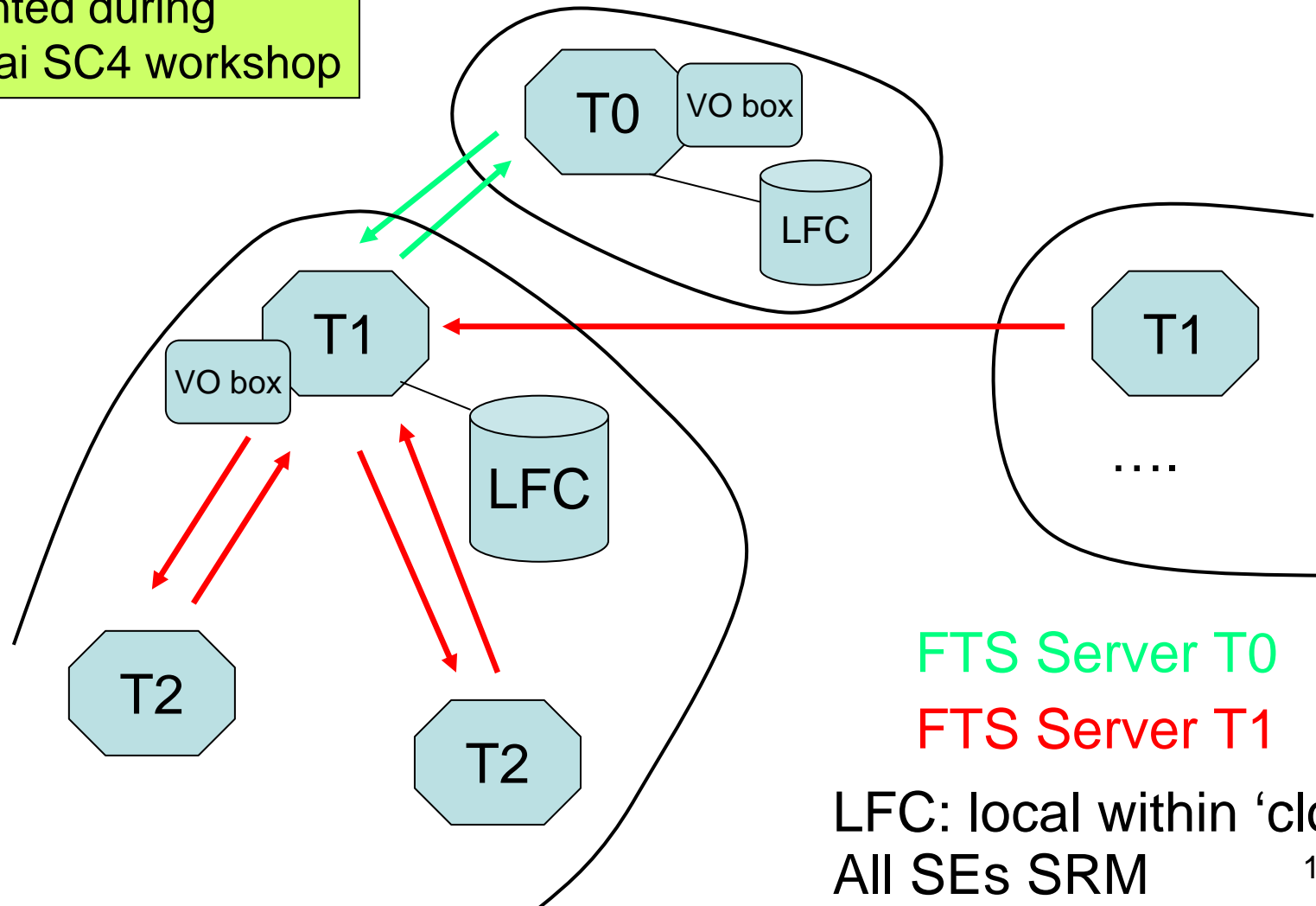
- Missing:
 - Enabling DQ2 monitoring again
 - Understanding impact of grid transfer layers on data organization
 - e.g. what to do if we transfer 90% of each dataset, but fail to transfer last 10% of the files ?
 - Adding T2s:
 - Requires volunteer contribution from T2 + strong connection with associated T1
 - Just now started on this, by establishing ATLAS ‘coupling’ of T1<->T2s;
 - Expect to join T2s when T0/T1 runs stably

Contact points for information

- Always a problem...
 - Error up in the chain (eg. T0 reserved LSF cluster) leads to less throughput to sites or to decision to stop export...
 - Difficult to communicate these events to the sites
- Distributed Data Management (DDM) log:
 - <https://uimon.cern.ch/twiki/bin/view/Atlas/DDMSc4>
- Mailing list:
 - atlas-t1-ddm-oper@cern.ch
- GGUS
 - + direct emails
- DQ2 Monitoring:
 - (currently being worked on due to scalability problems... will send link soon to the lists)
 - Allows monitoring of overall dataset transfers, single file transfer or cataloguing errors, etc

“Reminder”: ATLAS plans for using FTS

Presented during
Mumbai SC4 workshop



LFC: local within 'cloud'
All SEs SRM

“Reminder”: FTS Channels

- Tier-0 FTS server:
 - Channel from Tier-0 to all Tier-1s: used to move "Tier-0" (raw and 1st pass reconstruction data)
 - Channel from Tier-1s to Tier-0/CAF: to move e.g. AOD (CAF also acts as "Tier-2" for analysis)
 - “Star”-channel for all remaining traffic [new: low-traffic]
- Tier-1 FTS server:
 - Channel from all other Tier-1s to this Tier-1 (pulling data): used for DQ2 dataset subscriptions (e.g. reprocessing, or massive "organized" movement when doing Distributed Production)
 - Channel to and from this Tier-1 to all its **associated** Tier-2s
 - Association defined by ATLAS management (along with LCG)
 - “Star”-channel for all remaining traffic [new: low-traffic]

“Reminder”: Required services

- VO BOX per Tier1 and Tier0
 - Done. [trying to understand load on machine...]
- LFC server per Tier1
 - Done. [please make sure LFC is alive! ...]
- FTS server per Tier1 and Tier0
 - Done. Not all channel associations in place and validated but very close to it
 - A lot of **very recent progress** on this!
- Additional disk-only area on MSS
 - For all sites with mass storages, a SRM base path must be set as not garbage collected
 - Few sites have it

Other “SC*” issues

- Clean up of SC4 data?
 - Depends on exercise. Ask atlas-t1-ddm-oper@cern.ch
 - **For Tier0 exercise** we agreed sites delete data and we delete catalog entries
 - Data going to <SRM SC4 PATH>/sc4tier/<month>/<day>/
- SRM SC4 entries vs production (non-SC4) entries vs ‘tape’ vs disk basepaths?
 - Inform atlas-t1-ddm-oper@cern.ch of any changes. We manually retrieve endpoints from emails, wiki pages, LCG info sys, etc..
- Any other questions?