disclaimer
All proposals or numbers in these slides are for internal ATLAS usage
and no conclusions should be drawn outside the ATLAS scope of things

Tier-2's
in the ATLAS Computing Model

GDB@CERN, October 8 2008
IT-Aud.

Kors Bos, CERN & NIKHEF

ATLAS Computing Model for Tier-2's

- Very hierarchical
  - T0 communicates with T1's
  - T1's communicate with T2's in the same cloud
  - only T1-T1 communication between clouds
  - no T2-T2 communication between clouds
  - exception: muon calibration T2 sites get data from T0
- Monte Carlo Simulation in T2's
  - production of HITS
  - T1's also used when free CPU available
- Group and User Analysis in T2's
  - group analysis also done in T1's
  - some T1's even exclude user analysis

ATLAS Tier-2 Space Tokens

token name	storage type	used for	@T2	@T1	@T0
ATLASDATATAPE	T1D0	RAW data, ESD, AOD from re-proc		X	X
ATLASDATADISK	T0D1	AOD, DPD from data	X	X	X
ATLASMCTAPE	T1D0	HITS, AOD from GEANT4, AOD from ATLFAST		X	
ATLASMCDISK	T0D1	AOD, DPD from MC	X	X	X
ATLASPRODDISK	T0D1	Buffer for in-and export	X		
ATLASGROUPDISK	T0D1	Group analysis data	X	X	X
ATLASUSERDISK	T0D1	User analysis ata; Scratch	X	X *)	X
ATLASLOCALGROUPDISK	T0D1	Local User Data @T3			X

**) Although there is officially no user analysis at the Tier-1's, many Tier-1's have a Tier-2 component for which this USERDISK is needed (to be decided by each T1 individually)*

Computing Model for MC Simulation

GEANT4 full simulation

- EVNT data downloaded from T1 to T2 as input for G4
- G4 HITS are produced in the T2's and uploaded to the T1
- At the T1 HITS are archived on tape and stay on disk
- Digitization, Pile-up and reconstruction is done in the T1
- AOD's are archived on tape and stay on disk and are exported to all other T1's to stay on disk
- All AOD's from all other T1's are imported

ATLFAST simulation

- ATLFAST AOD's are produced in the T2's and uploaded to the T1
- At the T1 these AOD's are archived on tape and stay on disk and are exported to all other T1's to stay on disk
- All AOD's from all other T1's are imported

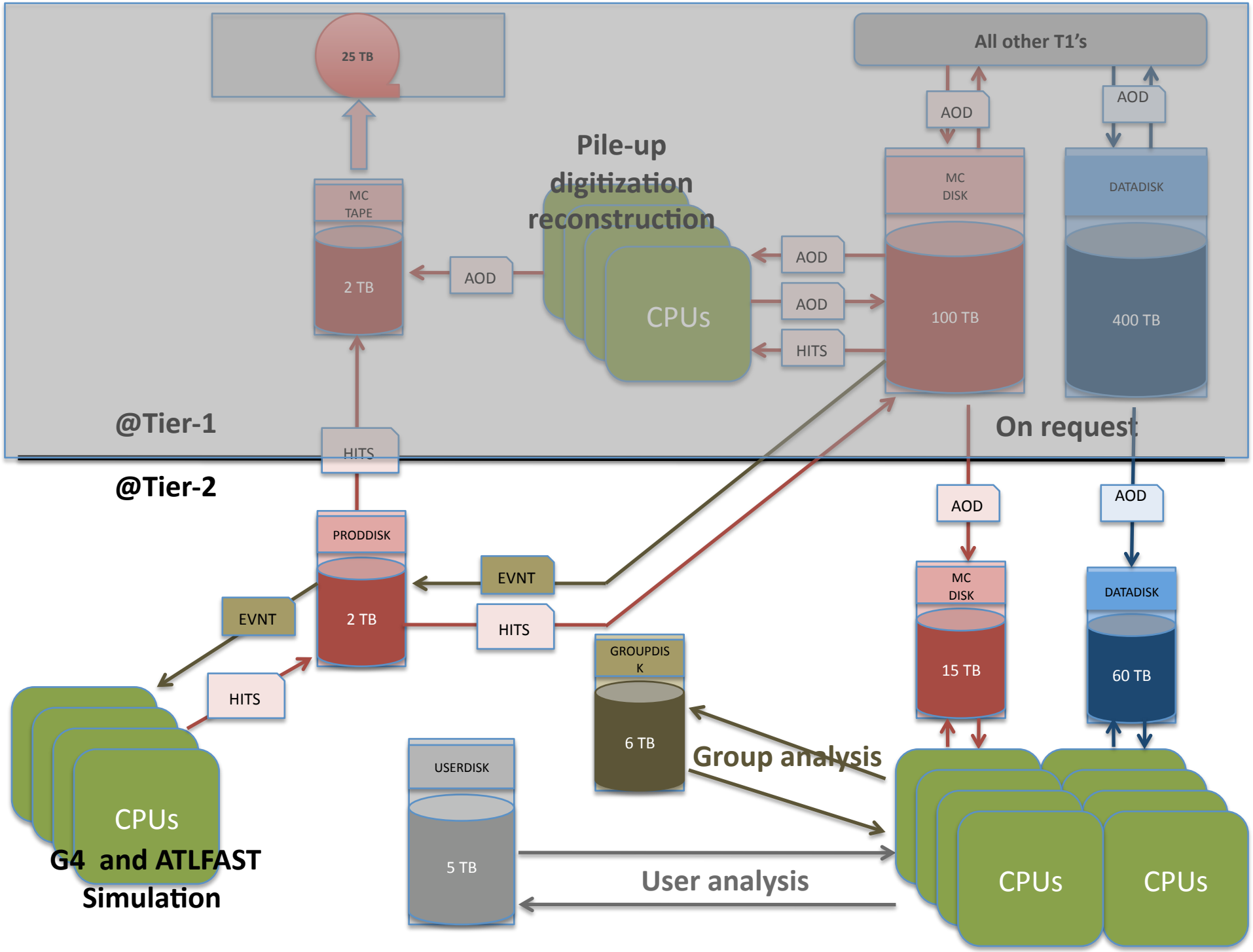
Computing Model for Analysis at Tier2's

- **Any analysis** will be primarily done from AOD and DPD's
- Within a cloud one may decide how to distribute AOD's and DPD's
- But within the cloud (sum of all T2's and T1) there must be one full copy
- This is true for Simulated Data (on MCDISK) and Detector Data (on DATADISK)
- Other data types may also be download on request (always from/via the T1)
- This is true for group analysis and user analysis

- **Group analysis** uses (will use) GROUPDISK as storage pool
- Directory defined per physics group (but no quota's)
- Only limited number of people have write access

- **User analysis** uses (will use) USERDISK as storage pool
- Scratch space, no guaranteed lifetime for data

- **Local T2/T3 users** use (will use) LOCALGROUPDISK as storage pool
- not ATLAS pledged resource but locally owned (T3)


25 TB
All other T1's
AOD
AOD
MC
TAPE
2 TB
AOD
Pile-up
digitization
reconstruction
CPUs
AOD
AOD
HITS
MC
DISK
100 TB
DATADISK
400 TB
@Tier-1
On request
HITS
@Tier-2
AOD
AOD
PRODDISK
2 TB
EVNT
EVNT
HITS
HITS
MC
DISK
15 TB
DATADISK
60 TB
GROUPDIS
K
6 TB
Group analysis
USERDISK
5 TB
User analysis
CPUs
G4 and ATLFAST
Simulation
CPUs
CPUs

Storage Pools @Tier-2

for a typical T2 with ~500 CPU's and ~100 TB disk

- **PRODDISK (2 TB)**
  - for EVNT data downloaded from the T1
  - for HITS and AOD before uploading them to the T1 of the cloud
  - space requirement depends on CPU capacity of the T2
  - but ~2 TB should almost always be enough
- **MCDISK (15 TB)**
  - to download from the T1 on request AOD and DPD from simulation
  - requested volume depends on analysis (CPU) capacity of the T2
  - left to the cloud to determine distribution
- **DATADISK(60 TB)**
  - to download from the T1 on request detector AOD and DPD
  - requested volume depends on analysis (CPU) capacity of the T2
- **GROUPDISK (6 TB)**
  - For physics group analysis
  - more may be needed (?)
- **USERDISK (5 TB)**
  - scratch space for users

Functional Tests for Tier-2's

- in DDM FT:
  - subscription of AOD's from T1 to T2
  - also subscription of calibration data from T0 to T2
  - tests DATADISK (could also test MCDISK, not done now)
- in MC FT:
  - run standard task in each cloud
  - also tests HITS subscription from T2 to T1
  - tests PRODDISK
- would also like to have Analysis FT:
  - still needs to be defined in detail
  - need metrics for succes
  - group analysis is sort of user analysis with results of general interest
  - would test GROUPDISK and LOCALGROUPDISK
- Mostly DPM but also many dCache systems and 1 Storm

analysis tools

- GANGA is our official analysis tool
  - Pathena for the US cloud
  - but more tools are used
  - runs within the cloud where the user is (too restrictive?)
- runs in T2's within a cloud
  - finds the data and splits the task accordingly
  - sends grid jobs to where data is
  - may also run in T1 (except in the UK)
  - input data mostly in MCDISK and DATADISK
  - output preferably to LOCALGROUPDISK
- final (ntuple) analysis locally in Tier-3
  - in and output on local disk
- still many unknowns, need more users!

Ticketing Tier-2s

- ATLAS urgently needs to be able to send tickets directly to Tier-2 sites
  - This requirement was raised in February and accepted by the GGUS team, but is not yet implemented
  - We really need this now (especially for our calibration Tier-2 sites)
- Team ticket interface also needs to support sensible searches (e.g., all open tickets)