



Enabling Grids for E-science

User experiences in data access

HEP

Dietrich Liko, IT/PSS

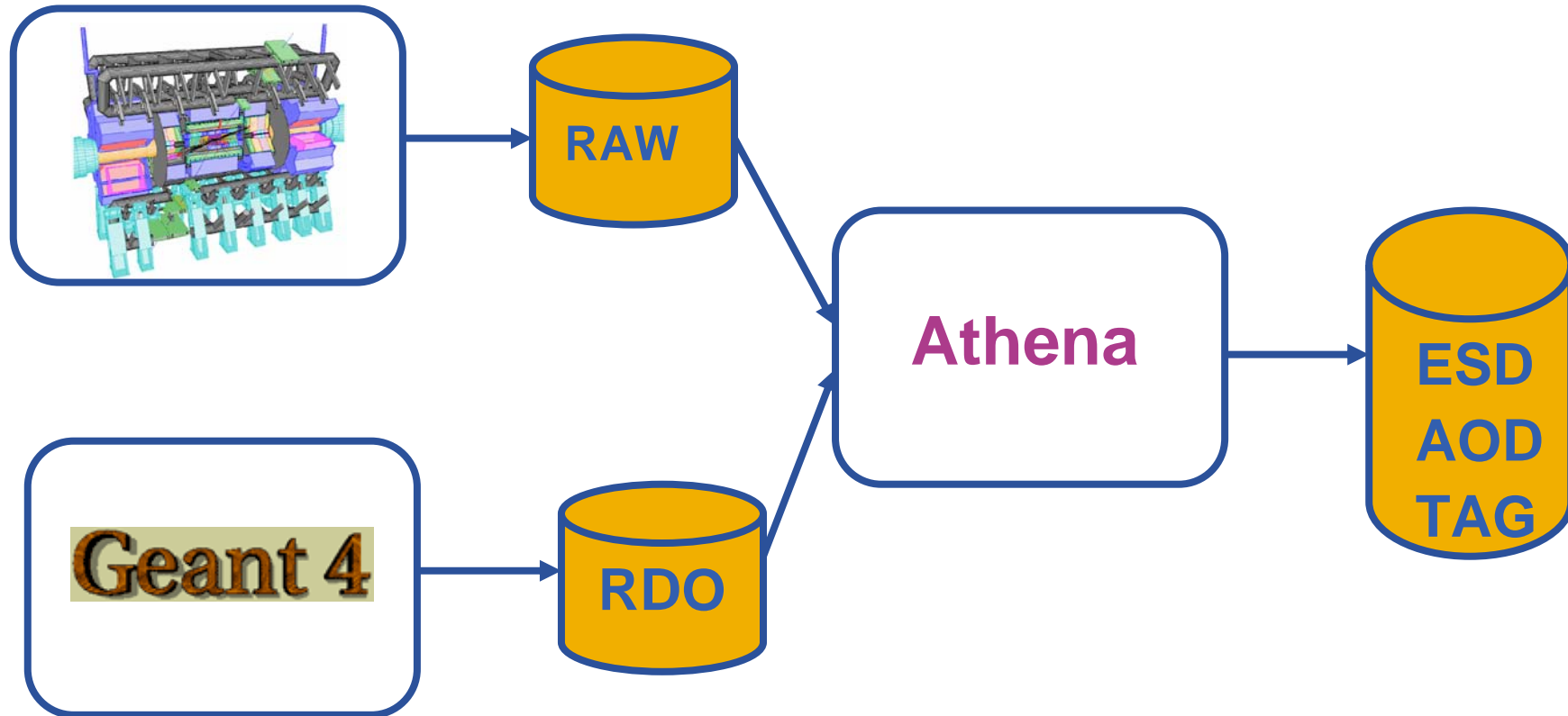
EGEE Conference, Geneva 2006

www.eu-egee.org



- **Please do not expect a review of the data management system of all HEP experiment**
 - ✓ I will present to you the ATLAS Data Management system as an example
 - ✓ Other experiments have similar requirements and choose similar solution (based on different implementations)

- **Requirements and Use cases**
- **Middleware components and their deployment**
- **The experiment layer as a wrapper on top of the middleware**
- **The first operational experiences during this year service challenge**



- **RAW data is recorded by the experiment and has to be recorded permanently at Tier-0 and at a Tier-1**
- **Simulated data (RDO) is produced by the GEANT4 simulation distributed on the Grid. It has to be collected**
- **Analysis data (AOD, ESD, TAG) is produced by the reconstruction program Athena. It has to be distributed according to schema on the grid to be analyzed by our users**

➤ Tier-0

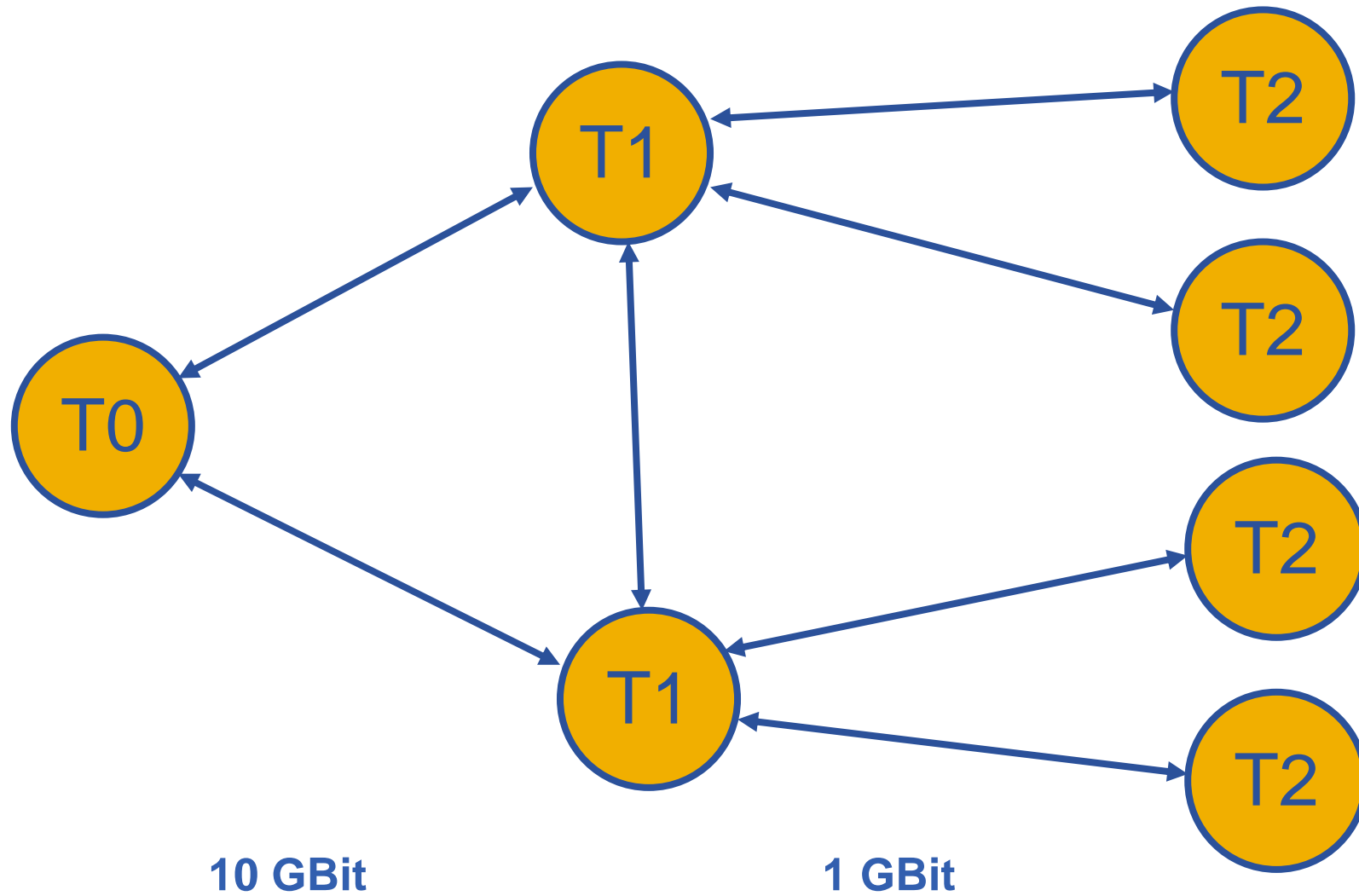
- ✓ Experiment will take data
- ✓ Permanent record of experiment data
- ✓ Export of all experiment data to the Tier-1

➤ Tier-1

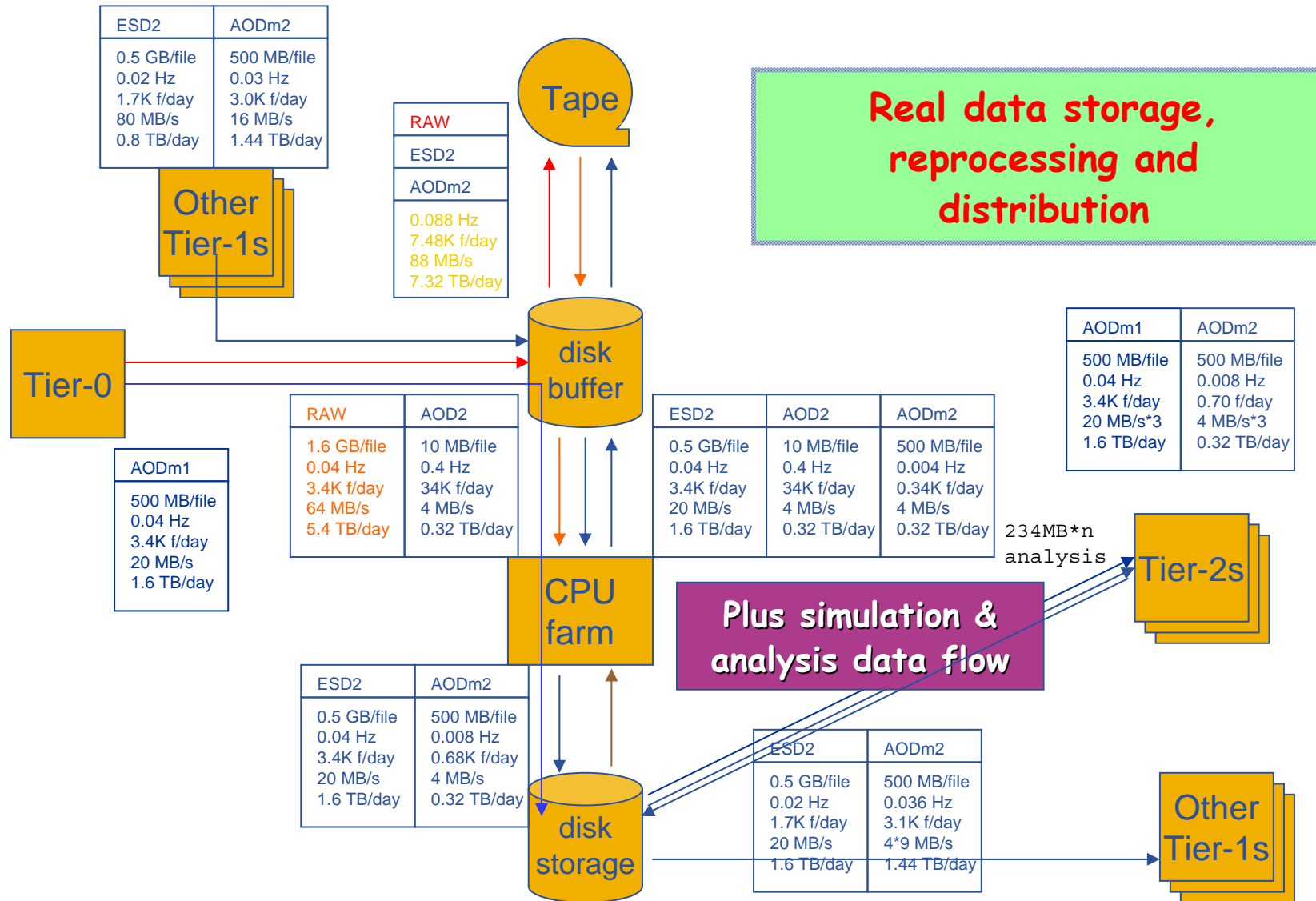
- ✓ Permanent record of experiment data
- ✓ Reprocessing of experiment data
- ✓ Collection of simulation data
- ✓ Distribution of reconstruction data

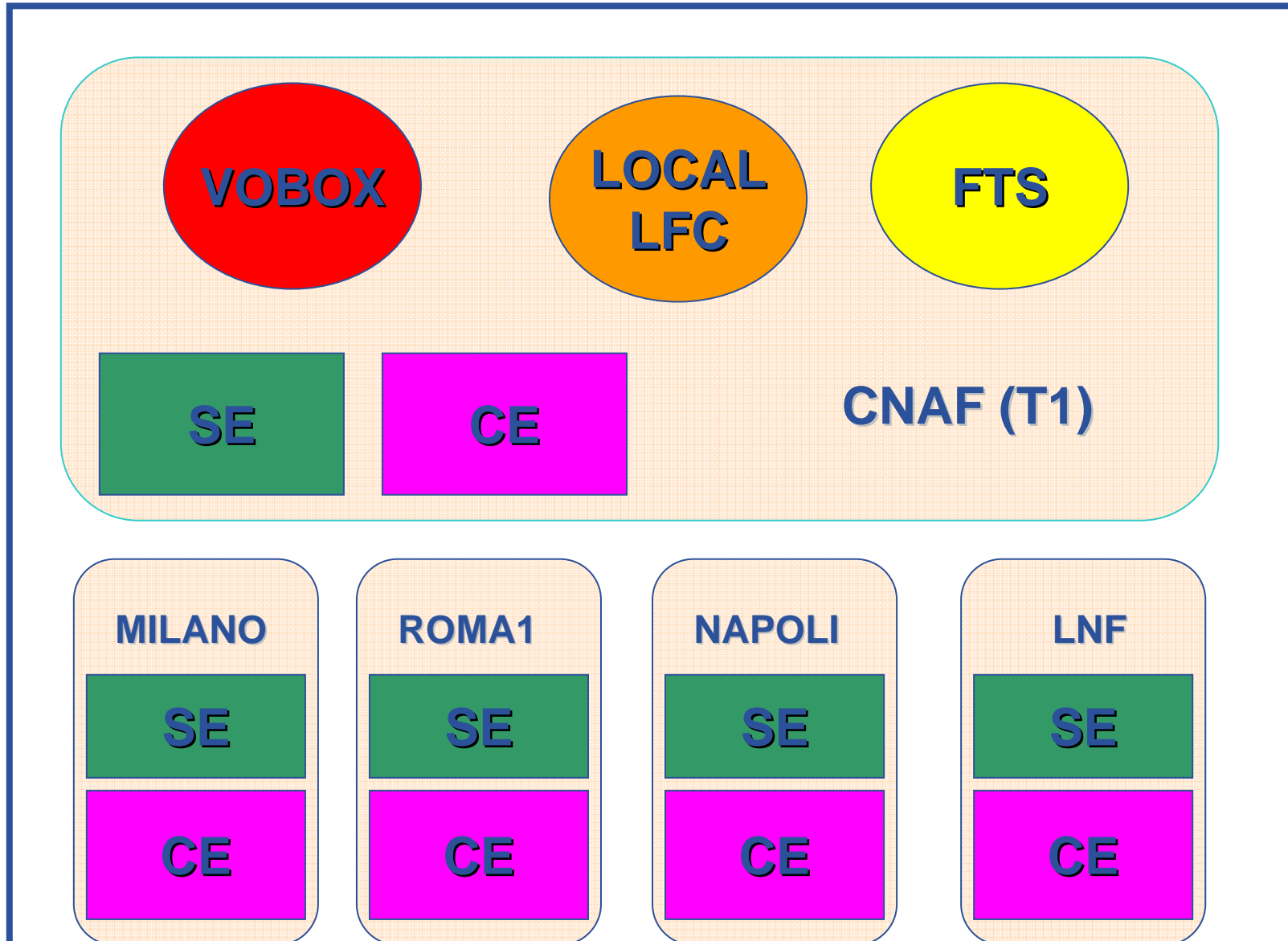
➤ Tier-2

- ✓ Production of simulation data
- ✓ Data Analysis



- 1. Distribution of data from Tier-0 to Tier-1**
- 2. Distribution of data from Tier-1 to Tier-2**
- 3. Collection of data from Tier-2 to Tier-1**
- 4. Distribution of data between Tier-1 and other Tier-1**
- 5. Local access to data for analysis**





- **Experiment specific layer on top of the middleware**

- **Dataset concept**
 - ✓ Collection of files
 - ✓ Location on a site
 - ✓ Hierarchical concept
 - ✓ Metadata are associated to a dataset
 - ✓ Aims to increase the scalability of the system

- **Subscription concept**
 - ✓ A dataset can be moved by a subscription of this dataset to a site
 - ✓ The system will translate the subscription command into actions on the middleware

- **System is operational on EGEE since the summer**

- **LFC Server**
 - ✓ At Tier-0 (hosting Tier-0 and CAF entries) and at Tier-1s (hosting Tier-1 and associated Tier-2 entries)

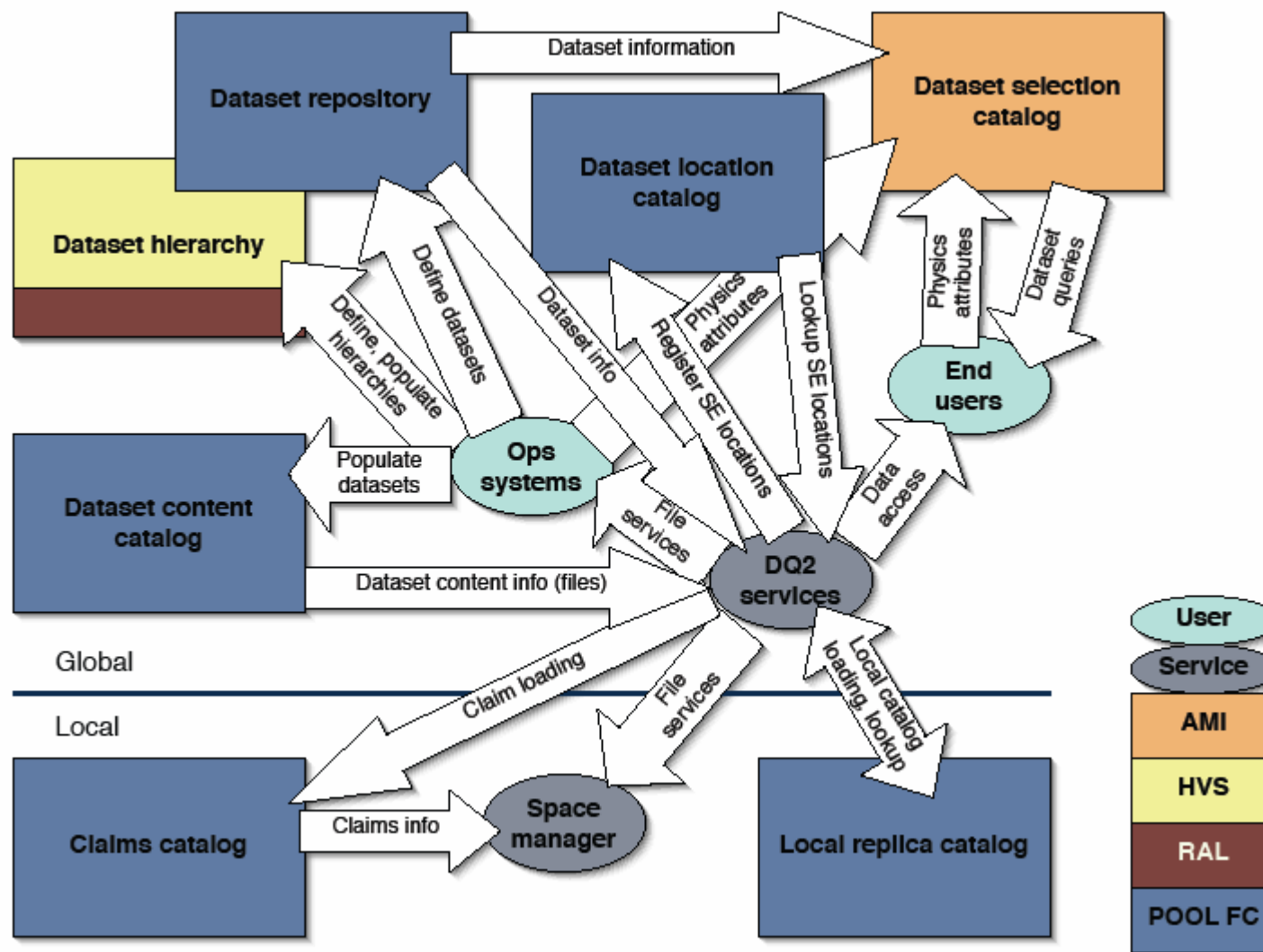
- **FTS Server**
 - ✓ At Tier-0 and at Tier-1s
 - ✓ Channels from other Tier-1s, from/to associated Tier-2s and source wildcard channel

- **SRM at all sites**

- **VO BOX at Tier-0 and at Tier-1s**
 - ✓ Requirements: <https://uimon.cern.ch/twiki/bin/view/Atlas/DDMVoBoxRequirements>

- **Overall DQ2 setup:**
 - ✓ Central DQ2 dataset catalogs:
 - ❖ Hosted at CERN, developed and managed by ATLAS
 - ✓ Site services, hosted in each VO BOX
 - ❖ multiple DQ2 installation on each VO BOX (one 'installation' per ATLAS Tier-1 and Tier-2 site, serving SC4 and Production simultaneously)

The Distributed Data Management



➤ T0 Exercise

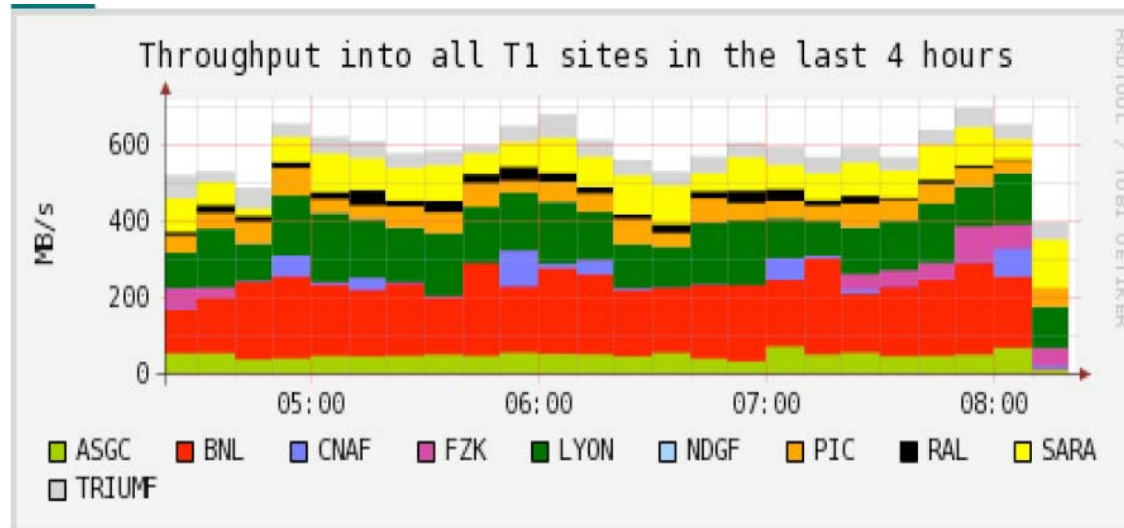
- ✓ Distribution of data

➤ Simulation production

- ✓ Collection of data
- ✓ Distribution of data between the sites

➤ Distributed Analysis

- ✓ Local access to data on the sites



- Run a full-scale exercise, from EF, reconstruction farm, T1 export, T2 export
 - ✓ Realistic data sizes, complete flow

- Maximum export rate (per hour) ~ 700 MB/s
 - ✓ Nominal rate should be ~ 780 MB/s (with NG)

- **Data collection from Tier-2 to Tier-1**
- **Data distribution between Tier-1's**
- **Data distribution to Tier-2s**

- **Many issues found**
 - ✓ No stable operation yet
 - ✓ Datasets are not moving fast enough
 - ✓ Manual operations are still necessary

- **We are working in trying to understand the issues**
 - ✓ Middleware
 - ✓ Services
 - ✓ Our application
 - ✓ Our operation

- **We offer distributed analysis at sites where we have data accumulated**
 - ✓ BNL, CERN, Lyon, FZK
 - ✓ Other Tier-1 will come later
 - ✓ And Tier-2 in the future

- **Local access by analysis jobs using GANGA**
 - ✓ Many options
 - ✓ Rfio, dcap
 - ✓ SRM
 - ✓ GFAL
 - ✓ Xrootd

- **We have started to evaluate the different options**

- **Data management for HEP experiments is a complex operation**
- **ATLAS choose to build a data management system on top of the EGEE middleware and other components**
- **We have put the system in production over the summer**
- **We could achieve already some of the aims, other will still need work**
- **Distributed Analysis has started to locally access data**