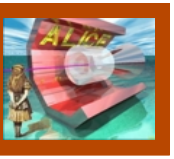
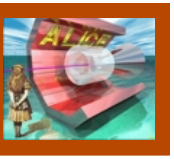


ALICE Analysis Use Cases at Tier-2 Centers

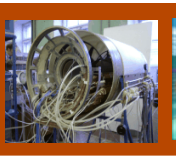
Andreas-Joachim Peters

Andreas.Joachim Peters CERN IT-PSS
for the ALICE experiment
LCG-Tier 2 Workshop 13.06.06 CERN-Geneva

Overview

- The ALICE Computing & Analysis Model
- Infrastructure for Analysis at Tier-2 Sites
  - LCG/AliEn Tier-2 'batch' analysis setup
  - LCG/AliEn Tier-2 'interactive' setup with PROOF
- Summary



ALICE Computing & Analysis Model

- Description of the ALICE computing model and the Tier-2 involvement in the TDR Chapter 7:
  http://aliceinfo.cern.ch/static/Documents/TDR/Computing/Chap7/chap7.pdf
- Detailed description of the Analysis model in the TRD Chapter 6:
  http://aliceinfo.cern.ch/static/Documents/TDR/Computing/Chap6/chap6.pdf

⇒ Short summary of Analysis and Computing Model following.

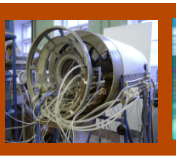
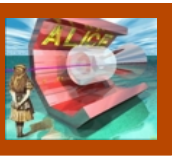
Details about expected data rates etc. can be found in these documents and are omitted in this presentation.

ALICE Computing Model

- Data Types
  - **RAW**
    recorded during the data taking
  - **ESD**
    event summary data produced by reconstruction
  - **AOD**
    physics analysis object data (like ntuple)
  - **TAG**
    event tags for event selection
- Software Framework
  - **ROOT / AliROOT**

ALICE Computing Model

- **Tier 0**
  - Permanent storage of raw data
  - Distribution of raw data to Tier 1
  - Calibration and Alignment and 1st reconstruction
- **Tier 1**
  - Permanent copy of raw data
  - Subsequent reconstruction passes
  - Scheduled reconstruction & Analysis of Pb/Pb MC
  - Long term storage of T1/T2 processed data
- **Tier 2**
  - Generate and reconstruct MC data
  - 'Chaotic' Analysis

ALICE Computing Policies

- Jobs are assigned where data is located
- Resources are shared
  - No 'localization' of groups
  - Equal Group/Site Contribution and Consumption will be regulated by accounting system
- Data access only through the GRID
  - No backdoor access to data
  - No 'private' processing on shared resources

ALICE Analysis Model

- Two types
  main difference: data access patterns, storage, code change frequencies
  - **Scheduled**
    - Analyses all data of a given type
    - Centralized – like data filtering for 'Sub-Analysis'
    - Output typically ESD/AOD (+ control histograms)
  - **Chaotic**
    - Focused on single physics tasks
    - Based on filtered data
    - Many iterations on 'random' subsamples of data
    - Output typically histogram files + event lists

Tier 0/1 Activity!

Tier 1/2 Activity!

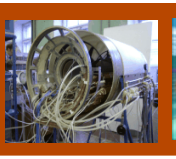
ALICE Analysis Model

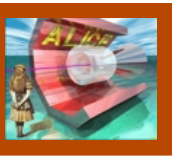
How to create a set of data files for Analysis?

⇒

File Catalogue and Meta Data Model

Andreas.Joachim Peters CERN IT-PSS
for the ALICE experiment
LCG-Tier 2 Workshop 13.06.06 CERN-Geneva





Creation of an Input Data List

- Created initially by queries to different database instances:
  - **Run Meta Data**
    - The Run Meta Data is stored as (Directory) Meta Data in the AliEn File Catalogue. Contains parameters describing equal conditions over a longer period of data taking.
  - **Event Meta Data**
    - Event Meta Data is stored in Tag Databases. The Tag Database is created for each ESD file after reconstruction and contains key/value pairs to apply selection criterias on physics variables
  - **File Meta Data**

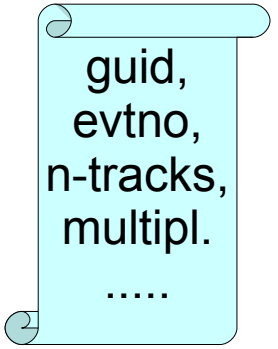
Run Meta Data

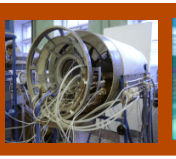
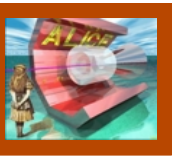
List under discussion ...

tag name	data format/`possible values`	data source
run comment	text	log book
run type	**`physics`**, **`laser`**, **`pulser`**, **`pedestal`**, **`simulation`**	log book
run start time	yyyymmddhhmmss	log book
run stop time	yyyymmddhhmmss	log book
run stop reason	**`normal`**, **`beam loss`**, **`detector failure`**, …	log book
magnetic field setting	**`FullField`**, **`HalfField`**, **`ZeroField`**, **`ReversedHalfField`**, **`ReversedFullField`**, …	DCS
collision system	**`PbPb`**, **`pp`**, **`pPb`**, …	DCS
collision energy	text, e.g **`5.5TeV`**	DCS
trigger class		log book
detectors present in run*	bitmap: `0`=not included, `1`=included	log book
number of events in this run		log book
run sanity	flag bit or bit mask, default `1`=OK	manually

Event Meta Data -Tag DB

- Tag DB
  - Implemented in ROOT trees
  - Registered in the AliEn File Catalogue
  - Structure
    - Tag DB with one Tag file per ESD file
    - Catalogue files for all ESD files of certain periods/conditions
  - Queried by the client or at the beginning of a running analysis job to create event lists for an Analysis



File Meta Data

- File Meta Data

  ⇒ contains no physics information:

  - Sanity of Files
  - Accessability of Files (permissions)
  - Location of Files

  ⇒ stored in the AliEn File Catalogue!

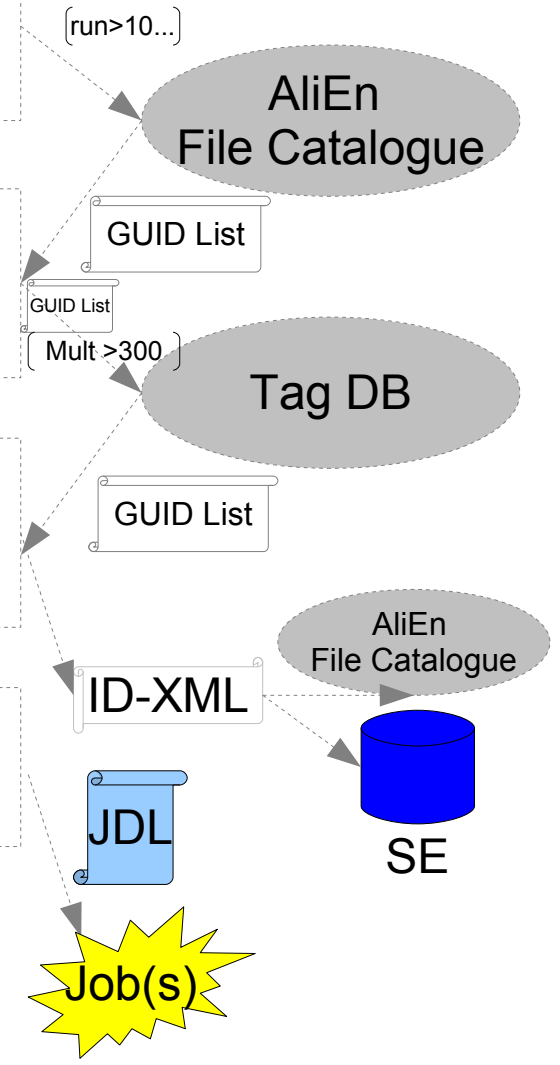
Example to create Analyis Input Data for a GRID job

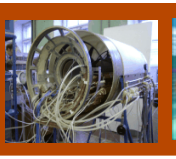
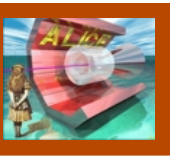
Namespace query
to the File Catalogue for a certain Run period and Run conditions

Tag DB query for retrieved GUID list for events fullfilling certain physics criteria

Creation of an Input Data XML File from the GUID list

Submission of a JDL file referencing the Input Data XML File as Input Data

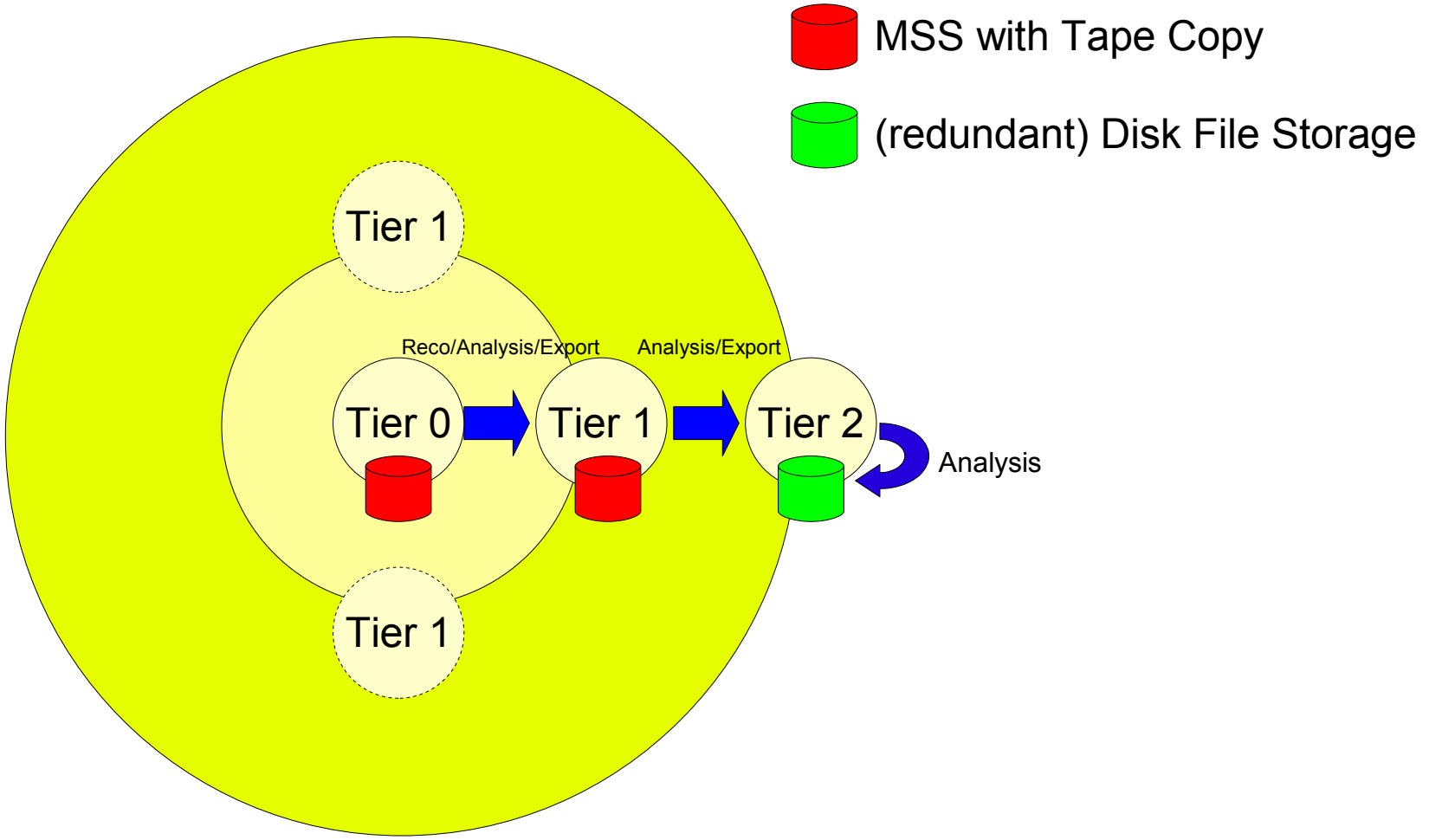


Data for Analysis at Tier-2 Centers

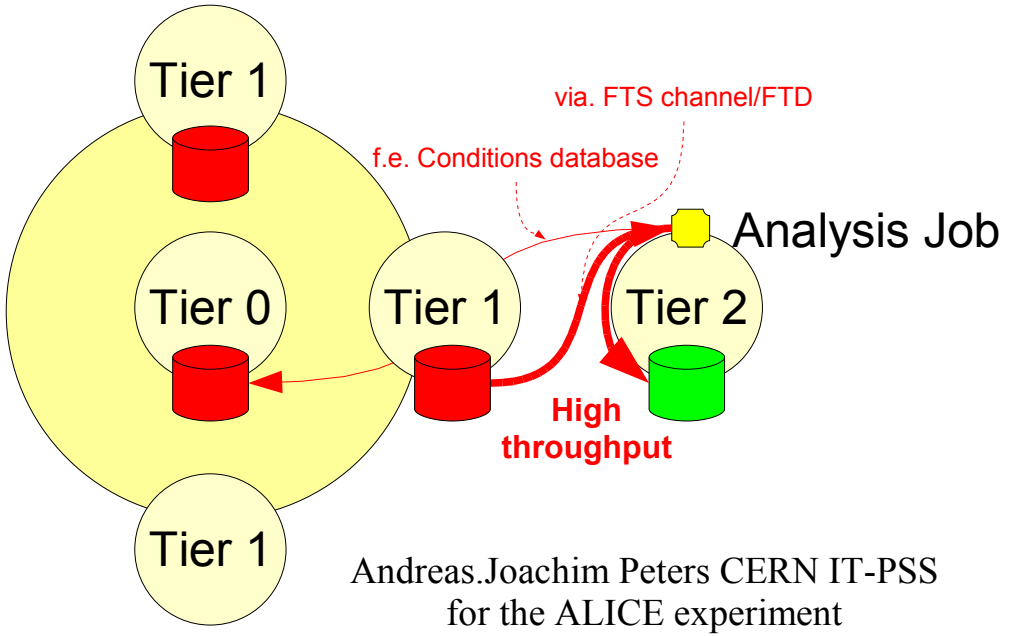
- Locally stored data:
  - Subsets of filtered ESDs/AODs for a specific physics analysis
  - MC data fo a specific physics analysis channel
  - Subsets of 'hot' data (cache of Tier-1 data)
  - No single replica data at Tier-2 centers
- Data transfer channel from assigned Tier-1:
  - Larger ESD/AOD samples stored at Tier-1 centers

Data Flow to Tier-2 Centers



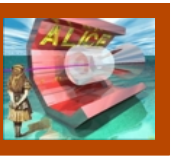
ALICE Model for Data Access

- Tier(N±1) to Tier(N):
  - scheduled transfers with high throughput
- Site A to Site B: remote access with low throughput possible (calibration files, Tag DB files, macro files, etc.) via xrootd protocol



1 data is moved between Tier centers via FTS/SRM

2 data is accessed from applications inside the Tier-2 exclusively using the xrootd protocol



Infrastructure for Analysis at Tier-2 Centers

Andreas.Joachim Peters CERN IT-PSS
for the ALICE experiment
LCG-Tier 2 Workshop 13.06.06 CERN-Geneva

Infrastructure for Analysis at Tier-2 Centers - Storage

- Data are stored in a tactical storage element:
  - Single Disk Server with xrootd protocol interface
    - xrootd server ⇒ see in **xrootd Tutorial!**
    - SRM/dCache with xrootd protocol (*under developement*)
    - SRM/DPM with xrootd protocol (*under developement*)
  - Multiple Disk Server with xrootd protocol interface
    - xrootd backend (redirector setup) ⇒ see **xrootd Tutorial!**
    - SRM/dCache backend (redirector setup)
    - SRM/DPM backend (redirector setup)
  - AliEn Storage Element service running on a VO box
  - AliEn File Transfer Daemon – LCG FTS + LFC

Infrastructure for Analysis at Tier-2 Centers – Computing/Data Access

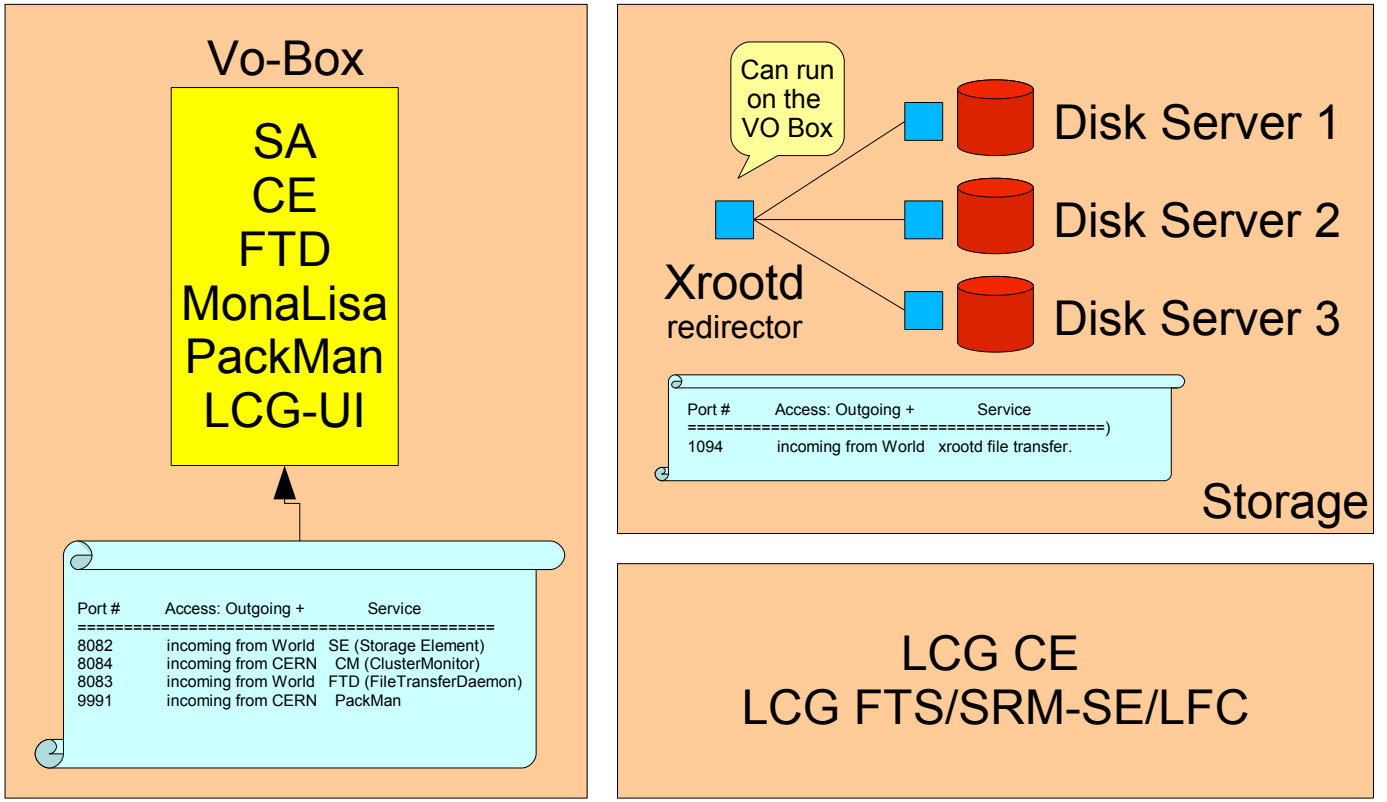
- Analysis Jobs run via an LCG Computing Element in a Tier-2 batch system
  - Jobs are submitted only via AliEn (GRID) tools
    submit ⇒TaskQueue ⇒ AliEn CE ⇒LCG CE⇒Tier-2 Batch System
  - Data is accessed only via AliEn (GRID) tools
    - All files are owned by a priviledged SE user and access is granted based on file catalogue permissions
    - No backdoor access without FC authorization
    - Data access only via xrootd protocol
- Monitoring via MonaLisa Service

Infrastructure for Analysis at Tier-2 Centers - VO-Box

- A Tier-2 Center needs to setup an ALICE LCG-VO Box providing:
  - AliEn Storage Adaptor and Computing Element
  - Package Management Service
  - MonaLisa Server
  - LCG UI
- Instructions with further requirements and the setup procedure can be found under:

  http://alien.cern.ch/twiki/bin/view/AliEn/HowToInstallAliEnSite

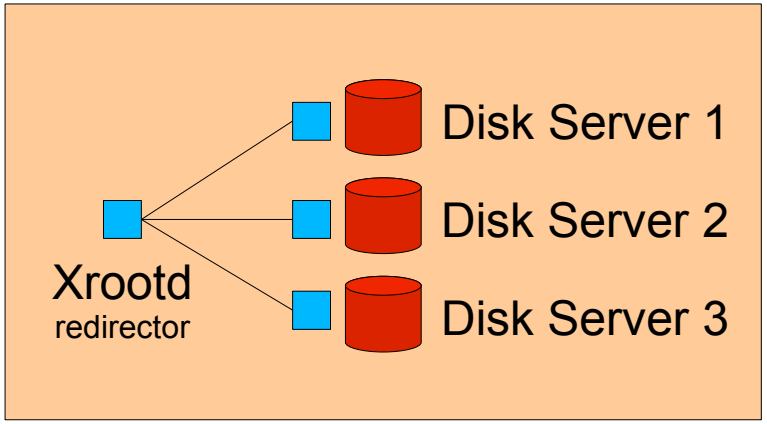
Tier-2 Infrastructure/Setup Example



Workernode configuration/requirements are equal for batch processing at Tier0/1/2 centers (2 GB Ram/CPU – 4 GB local scratch space)

Tier-2 Infrastructure/xrootd Setup

1



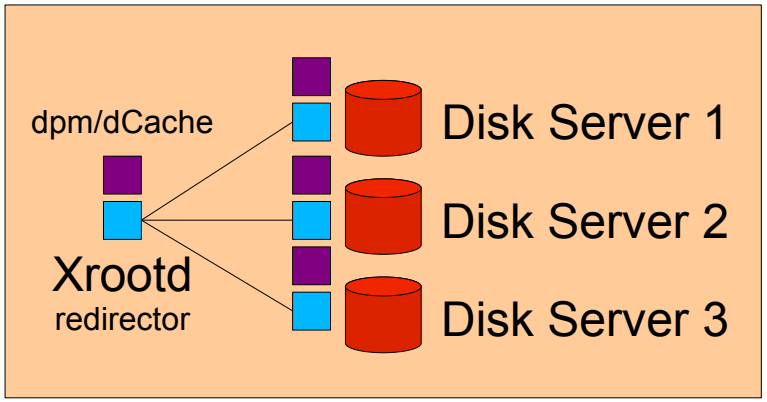
Xrootd native setup

The xrootd setup needs some ALICE customization. Read:
http://alien.cern.ch/twiki/bin/view/AliEn/HowToInstallXrootd

Alice run's an OFS plugin for authorization:

- every file access has to be granted based on information in the FC by an API service.
- Authorization is enforced by 'Token Envelopes' which are encoded/decoded with two key pairs
- The setup procedure includes the xrootd configuration and the installation of two keys for token decoding

2

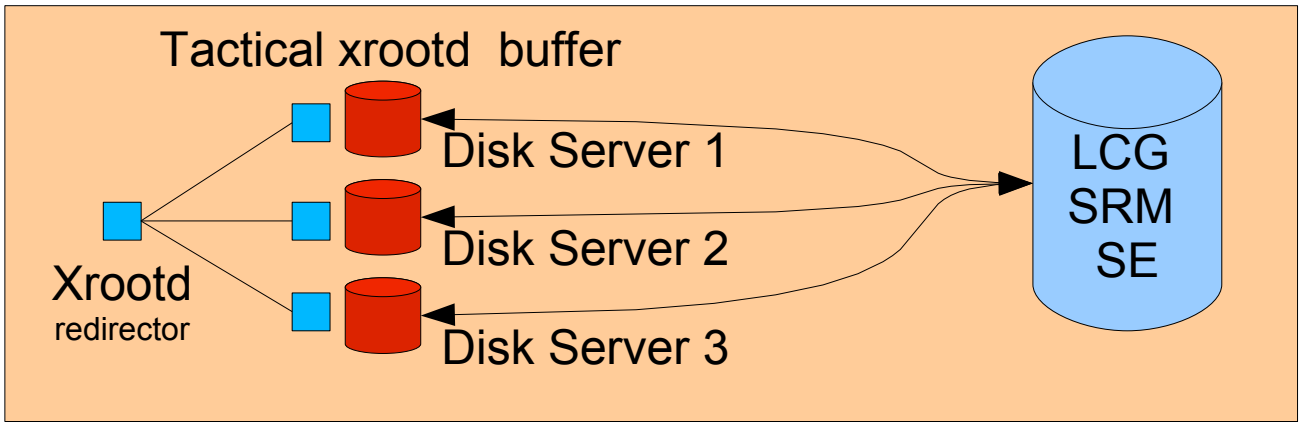


Xrootd overlay setup on LCG SRM storage

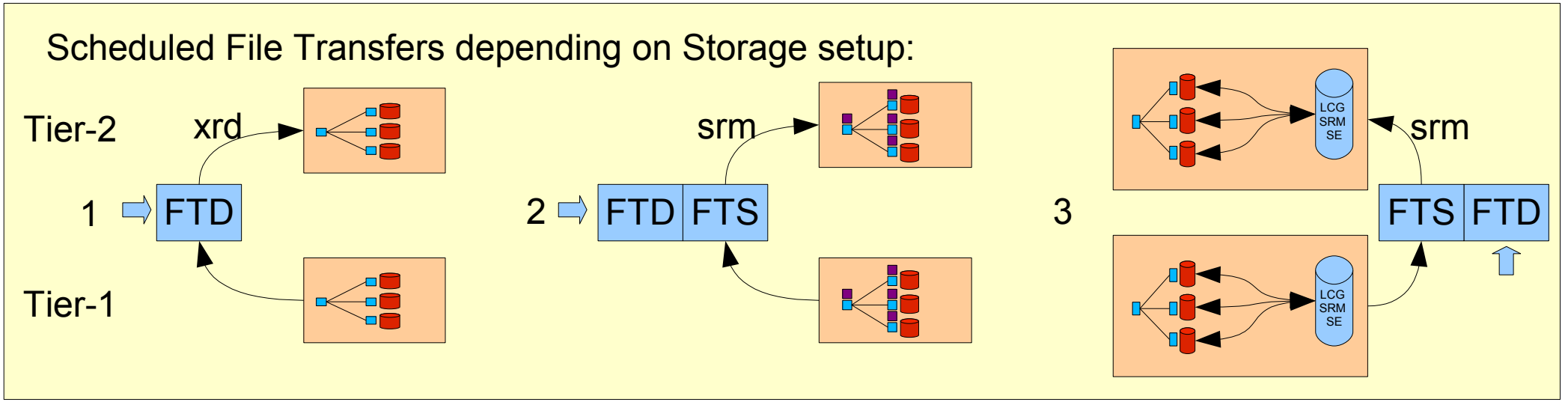
The xrootd interface to dpm/dCache is not yet ready for installation.

Tier-2 Infrastructure/xrootd Setup

3

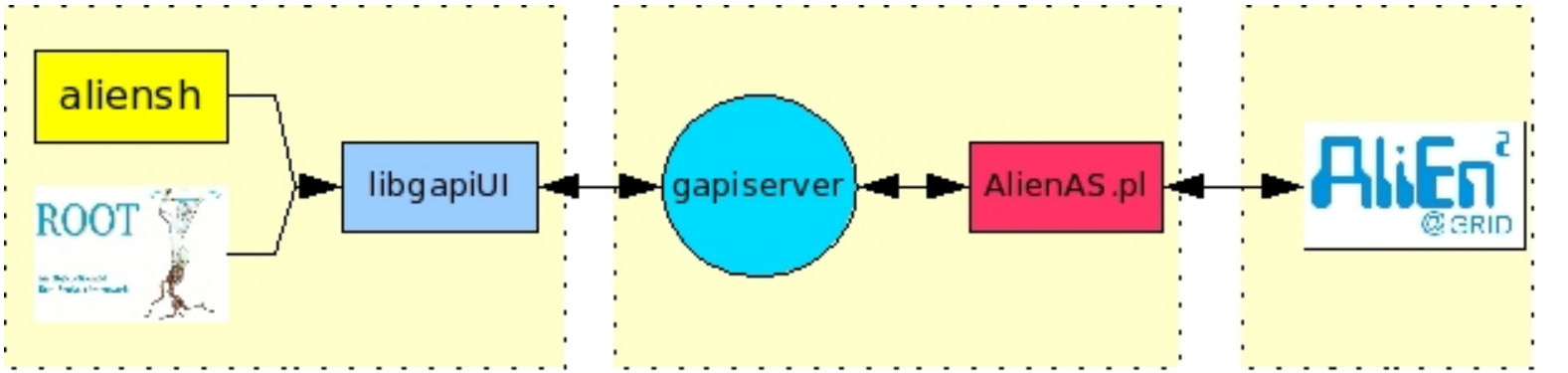


Xrootd tactical buffer in front of SRM SE



Analysis User Interfaces

- Two user interfaces provided using AliEn Application Interface Services:



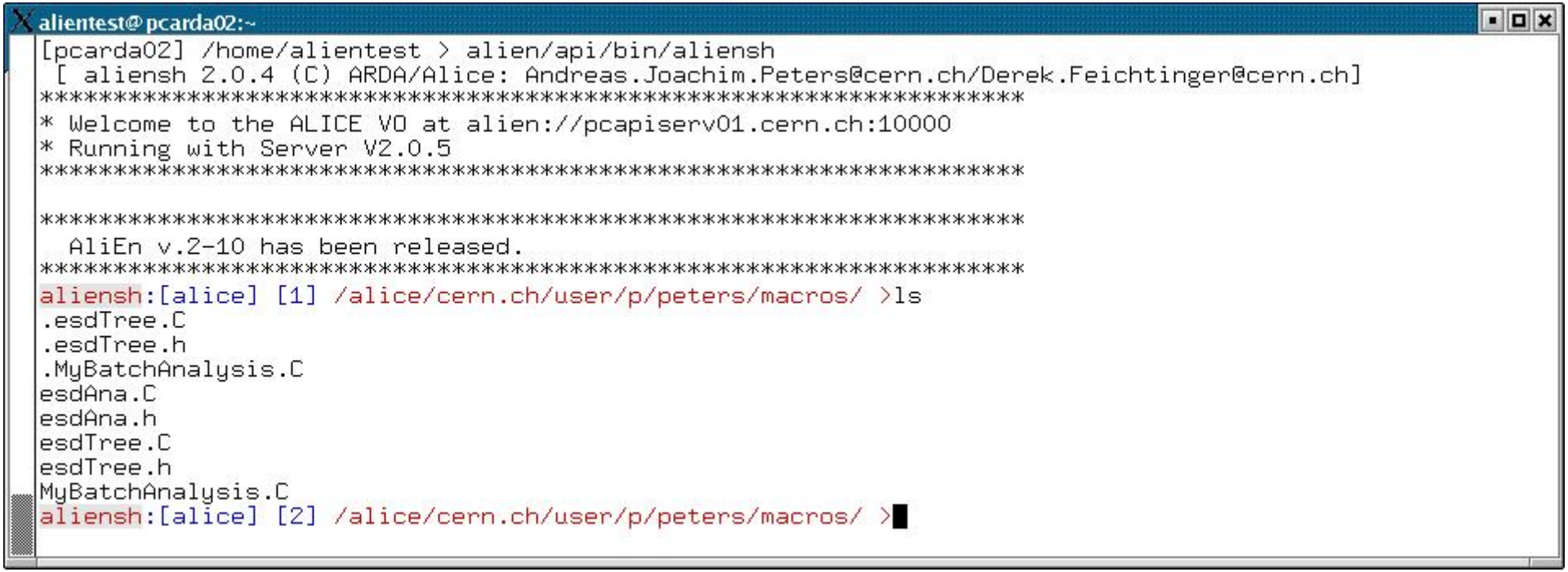
Running at Tier 0/1/2

Central redundant services

Analysis User Interfaces

aliensh

- Shell-like interface to ALICE GRID resources
  (File Catalogue/Storage+Computing Resources etc.)

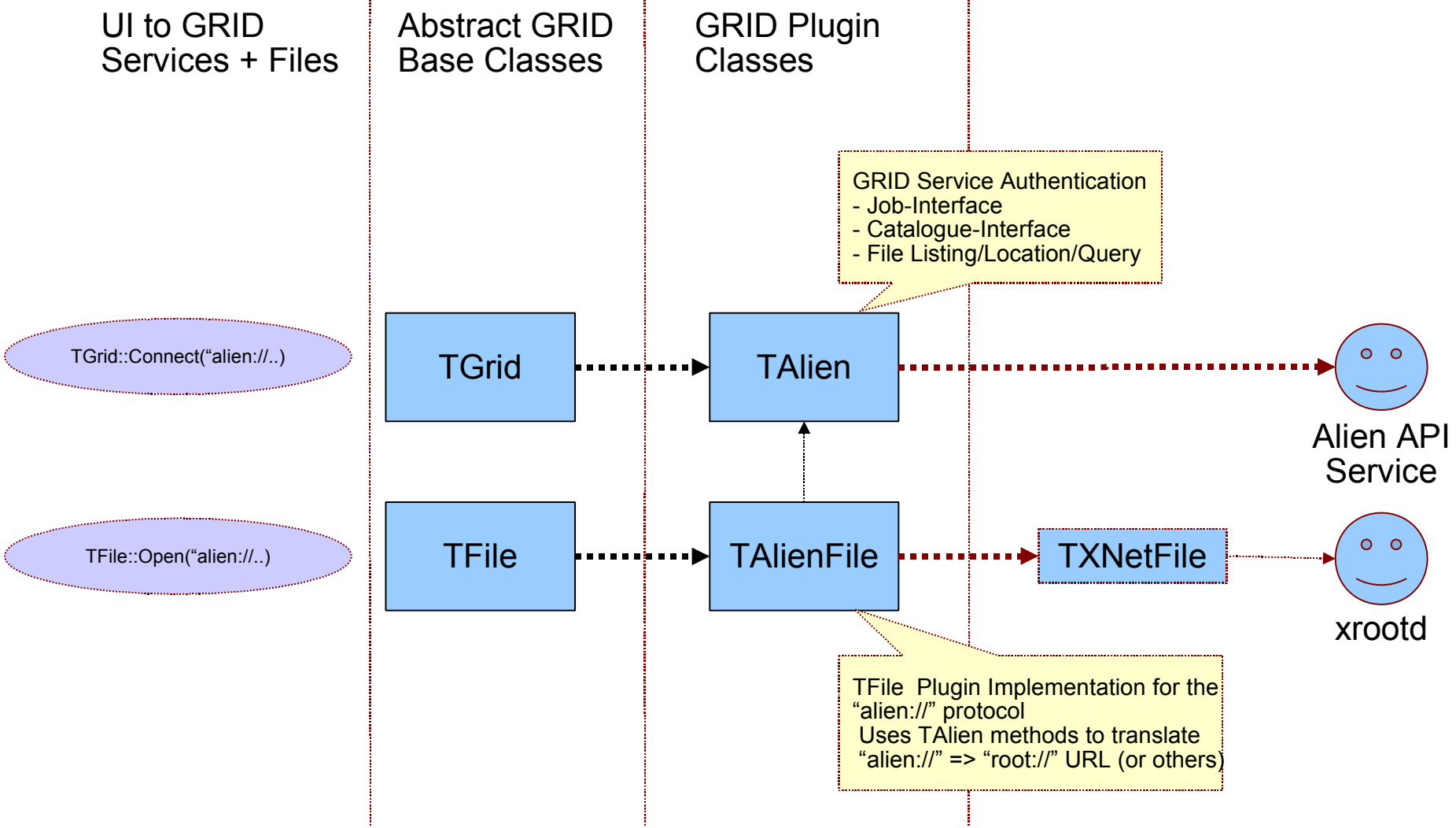


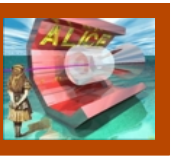
Analysis User Interfaces

ROOT TGrid Interface

- exports all aliensh commands into the ROOT C++ application
- exports all files registered in the AliEn FC via the “alien:” protocol

Analysis User Interfaces – ROOT AliEn interface

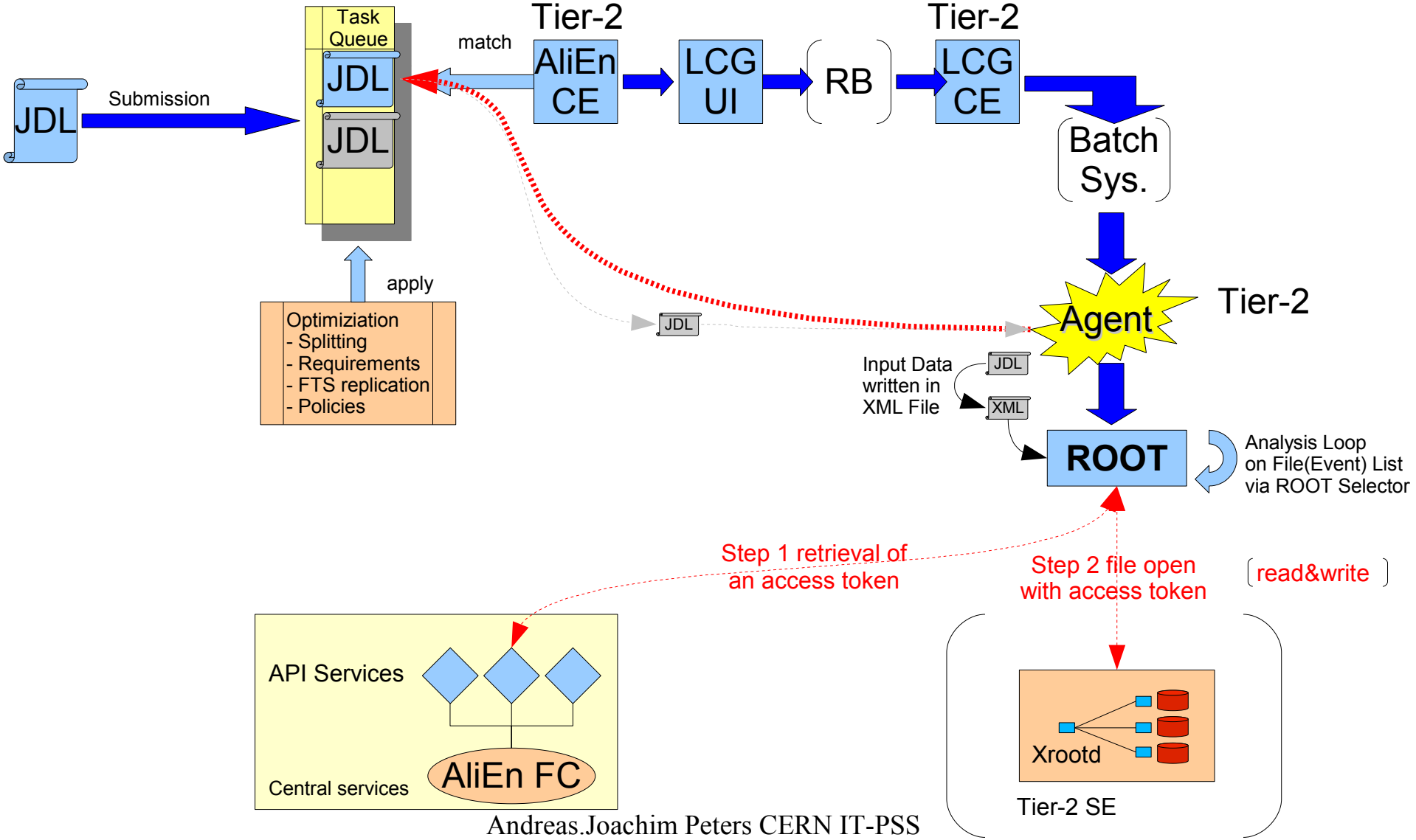


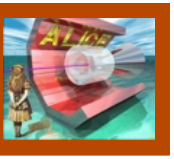


Tier-2 Batch Analysis

Andreas.Joachim Peters CERN IT-PSS
for the ALICE experiment
LCG-Tier 2 Workshop 13.06.06 CERN-Geneva

Tier-2 Batch Analysis Workflow Example





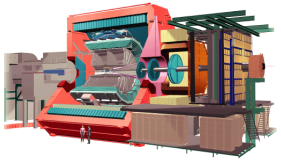
Tier-2 Interactive Analysis

Andreas.Joachim Peters CERN IT-PSS
for the ALICE experiment
LCG-Tier 2 Workshop 13.06.06 CERN-Geneva

Infrastructure for interactive Analysis at Tier-2 Centers - PROOF Cluster

- Analysis in a Tier-2 PROOF cluster
  - Installation of a PROOF cluster with access to (a) GRID enabled storage element(s)
  - First installation now at CERN CAF (CERN Analysis Facility)
    - Requires xrootd with special configuration file to be deployed on cluster nodes
    - After successful evaluation, experience can be used to setup T2AF
  - Analysis code is equal to GRID job analysis
  - Faster response – faster analysis cycles

CAF PROOF Setup



Experiment

Disk Buffer

Tier-1 data export

Tape storage

Sub set (moderated)

CAF computing cluster

Proof node | local disk

Proof node | local disk

Proof node | local disk

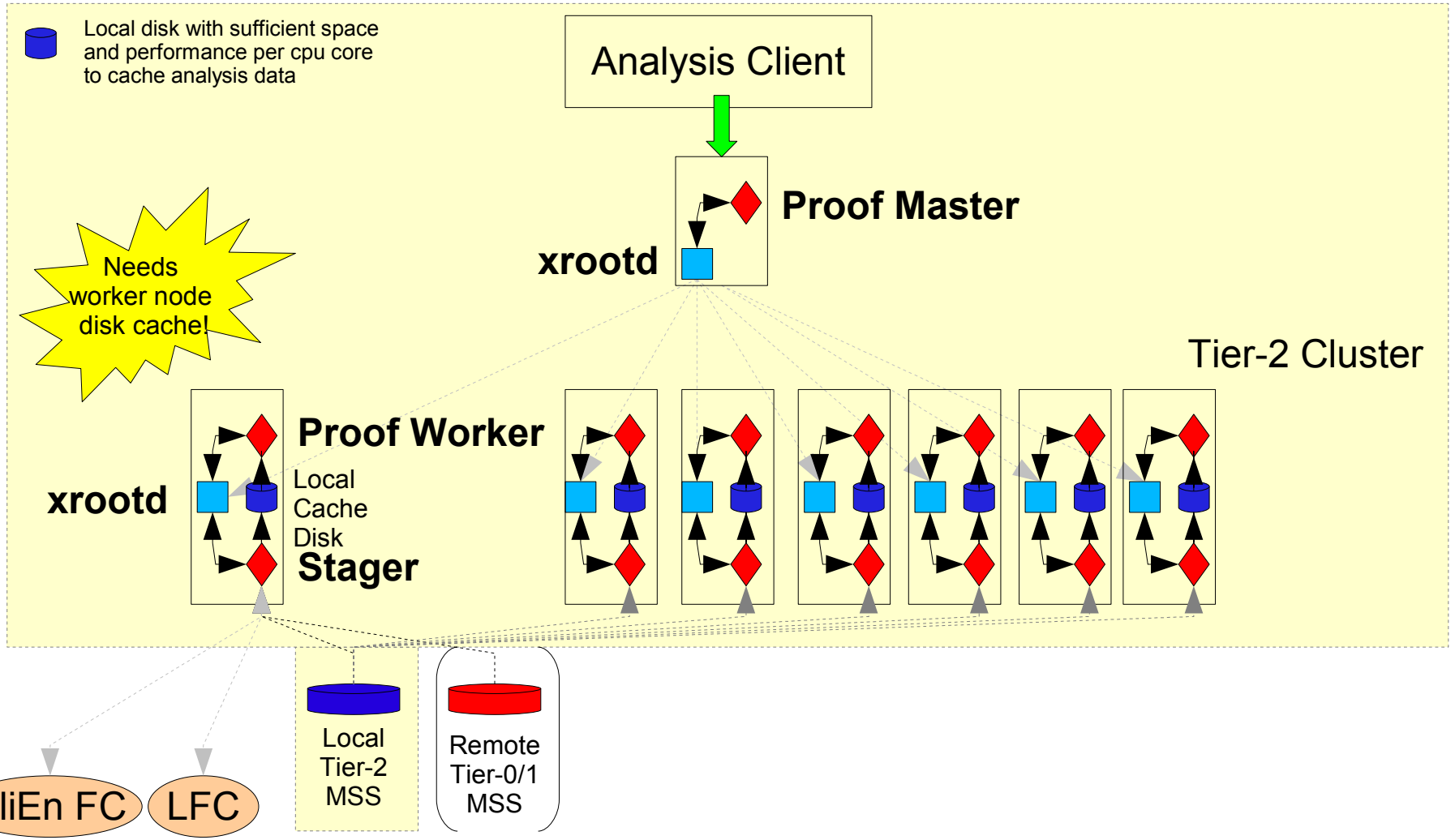
...

Proof node | local disk

Proof node | local disk

Proof node | local disk

Tier-2 PROOF Setup



Tier-2 Requirements from ALICE User Model

- File Ownership is virtualized in the AliEn FC and authorization enforced via the xrootd plugin
- Users are identified by their proxy certificate
- Jobs are running under a generic ALICE account and will hopefully be sandboxed into impersonal identities using some 'glexec'-like functionality soon
- Therefore at the moment only **two accounts** needed
  - 1 identity owning files
  - 1 identity running job agents/user jobs

Summary

- All user analysis will be running at Tier-2 centers
- All Tier-2 resources are shared among users
- Tier-2 Site Requirements for Analysis:
  - LCG-CE/FTS/LFC + ALICE VO-Box Setup
  - Tactical Storage element with xrootd-protocol interface (difference to Tier-1 setup)
  - ALICE encourages to setup a Tier-2 analysis PROOF cluster as soon as CAF expertise available
- Tier-2 Analysis is scheduled starting autumn 2006 !