



IPM

# GRID for LHC

Saeid Paktinat

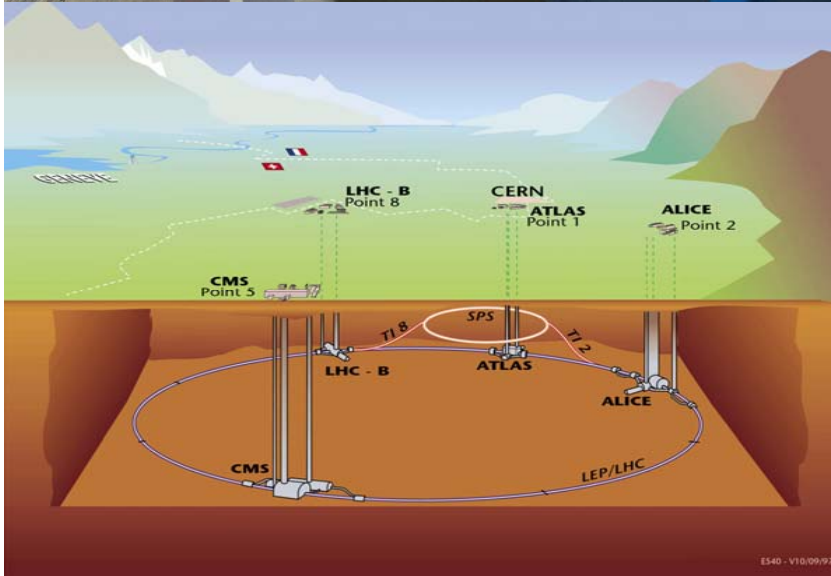
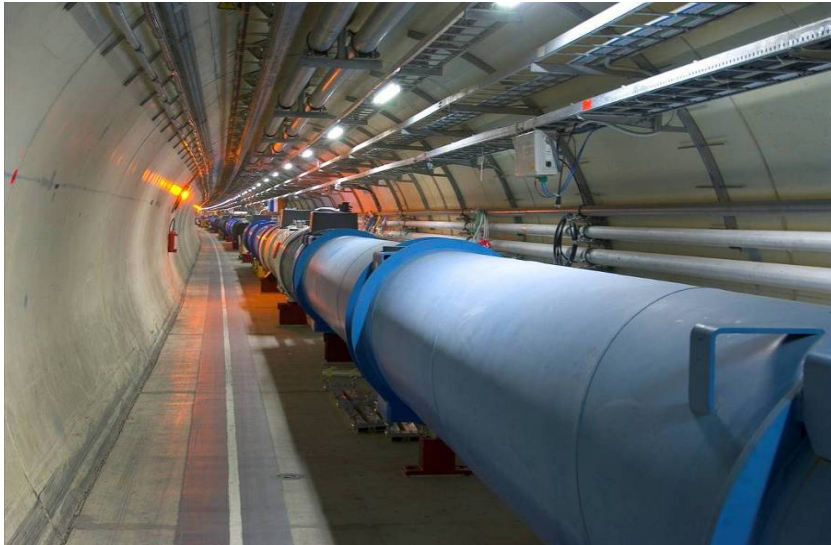
School of Particles and accelerators

IPM, Tehran

**One day seminar on GRID computing**

**Azar 14, 1387**

# The Large Hadron Collider



- 2835 + 2835 proton bunches separated by 7.5 m

- collisions every 25 ns = 40 MHz crossing rate

- $10^{11}$  protons per bunch

- at  $10^{34}/\text{cm}^2/\text{s}$   
≈ 25 pp interactions per crossing pile-up

- ≈  $10^9$  pp interactions per second !!!

- in each collision  
≈ 1600 charged particles produced

- enormous challenge for the detectors/Computing

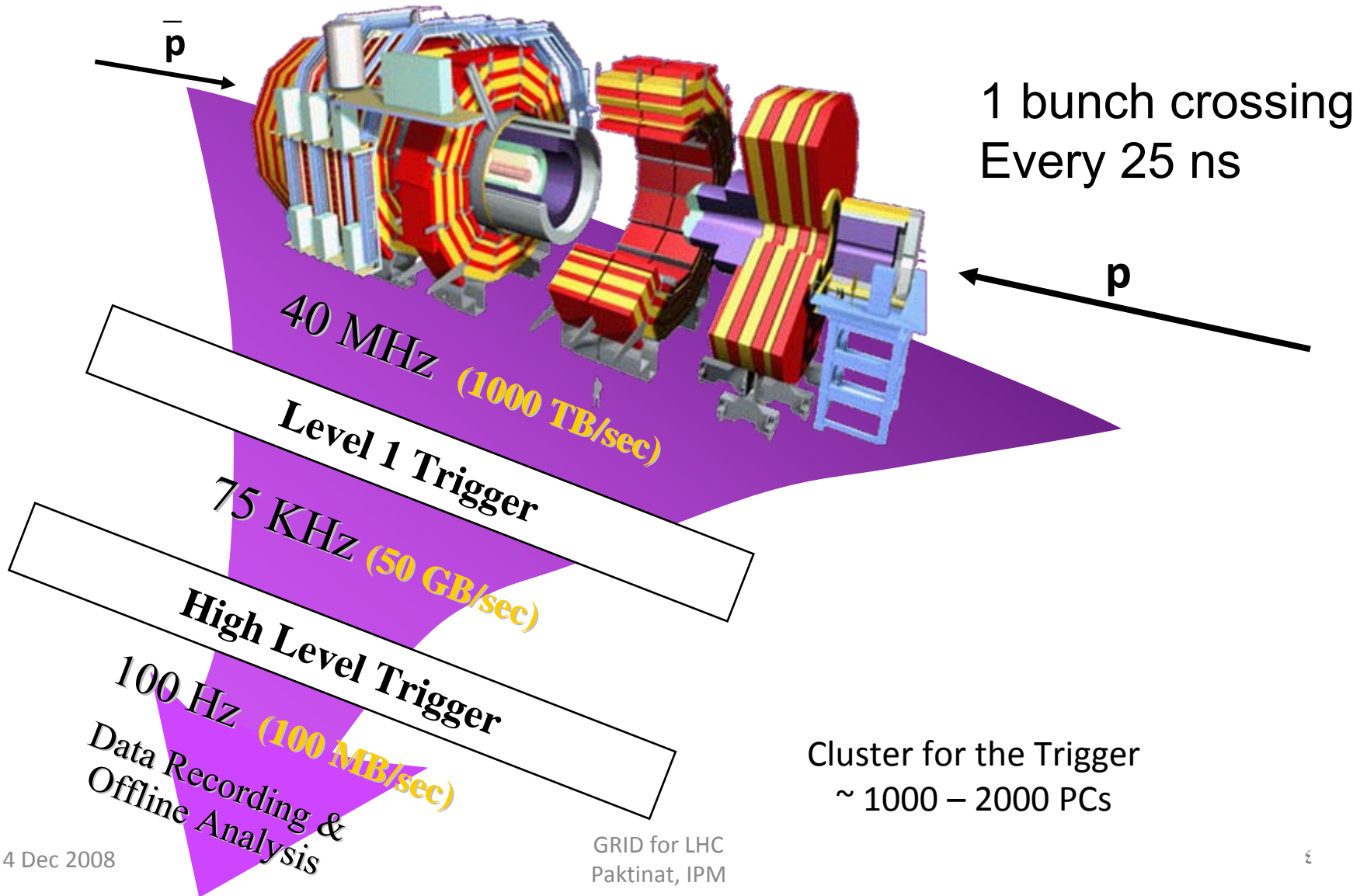
# A comparison

12-14 PetaBytes of data will be generated each year  
(20 millions CDs == 20 km)



Here we concentrate on CMS as an example

# CMS data production at LHC



# CMS Computing Model

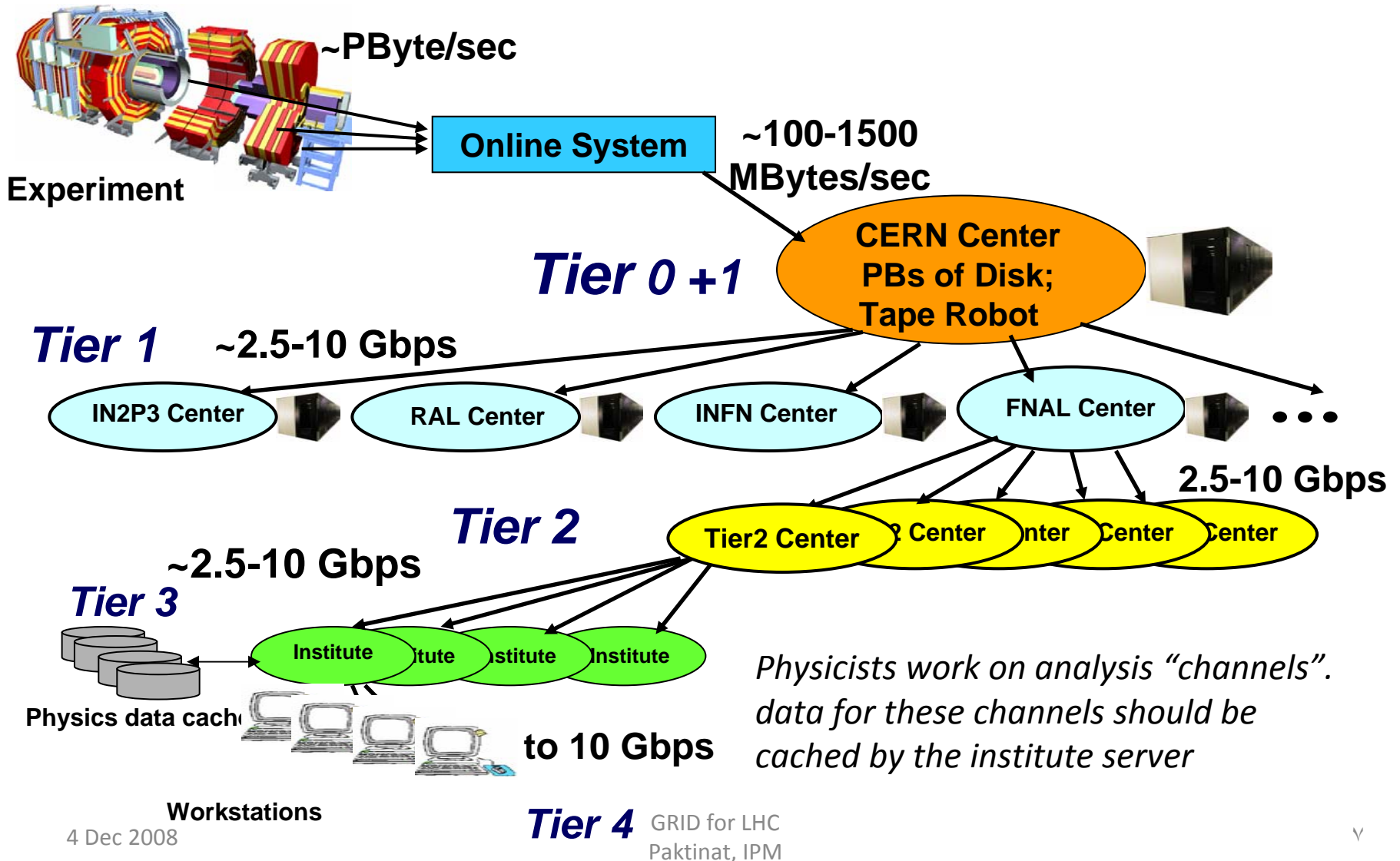
- Base ideas
- CMS computing and storage requirements would be difficult to fulfill at any one place, for both technical and funding reasons
- The CMS computing environment has been constructed as a distributed system of computing services and resources that interact with each other as Grid services.
- The computational infrastructure is intended to be available to CMS collaborators, independently of their physical locations, and on a fair share basis.

# CMS Dataformats

- RAW: Detector data after High Level Trigger (HLT) processing. Large (1.5MB/ev)
- RECO (Reconstructed Data): Reconstructed objects (tracks, vertices, jets, electrons, muons, etc.) and reconstructed hits/clusters. Small (400kB/ev)
- AOD (Analysis Object Data): Subset of RECO. Reconstructed objects (tracks, vertices, jets, electrons, muons, etc.). Possible small quantities of very localised hit information.

Mainly used for physics analysis. Very small (50kB/ev)

# CMS computing model



# Tier-0

The first tier in the CMS model, for which there is only one site, CERN, is known as Tier-0 (T0)

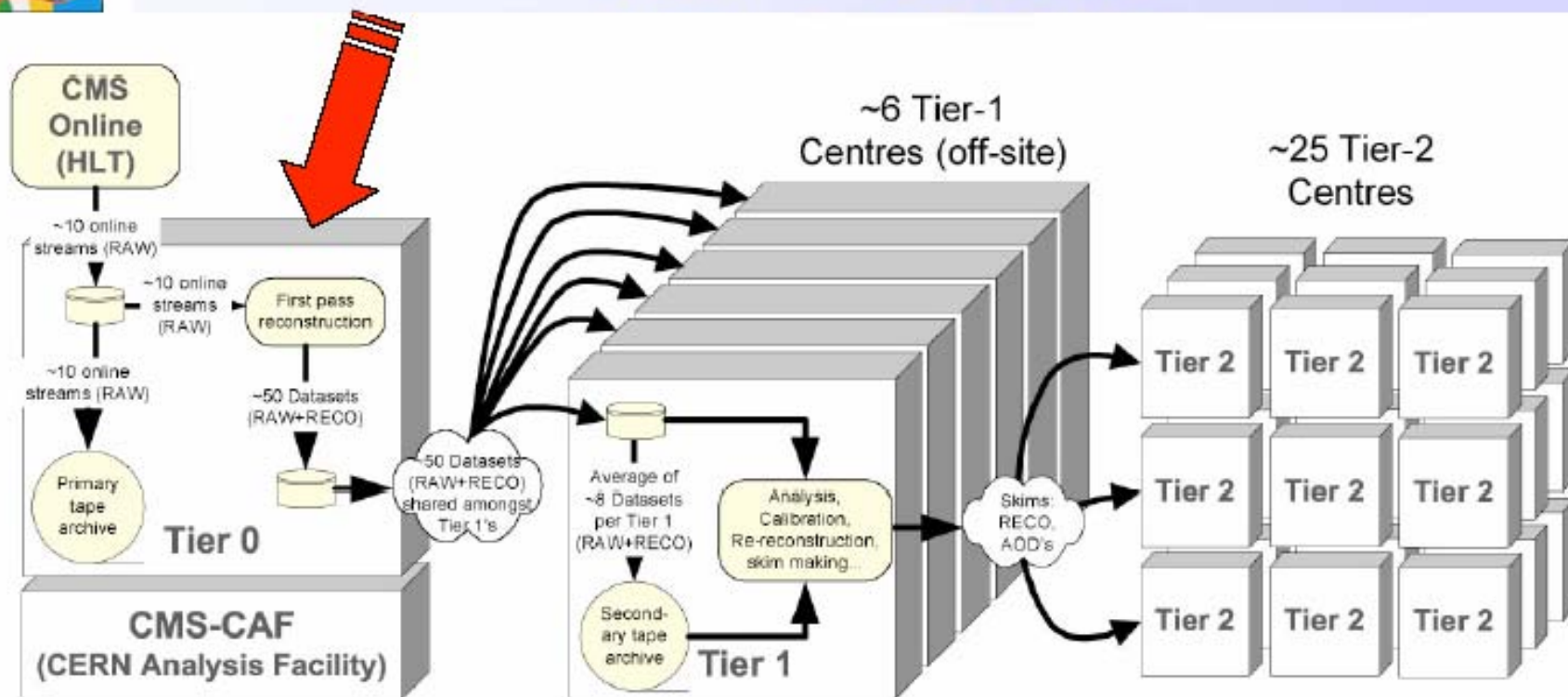
1. accepts RAW data from the CMS Online
2. repacks the RAW data into primary datasets
3. distributes RAW data sets among the next tier stage resources (Tier-1) so that two copies of every piece of RAW data is saved, one at CERN, another at a Tier-1.
4. distributes the RECO datasets among Tier-1 centers, such that the RAW and RECO match up at each Tier-1.
5. distributes full AOD to all Tier-1 centers.

The T0 does not provide analysis resources.

CERN has CERN Analysis Facility (CAF) for that



# Tier-0 Centre



- ➔ Prompt reco (24/200); FEVT storage; data distribution
- ➔ Provided by IT Division, formal QSA.
- ➔ CPU 4.6MSI2K; Disk 0.4PB; MSS 4.9PB; WAN > 5Gb/s

# Tier-1

There is a set of (currently) seven Tier-1 (T1) sites. Each T1 center is associated with and provides data to a group of smaller Tier-2 centers.

1. provides substantial CPU power for:

- \* re-reconstruction
- \* calibration
- \* AOD extraction
- \* and other data-intensive analysis tasks.

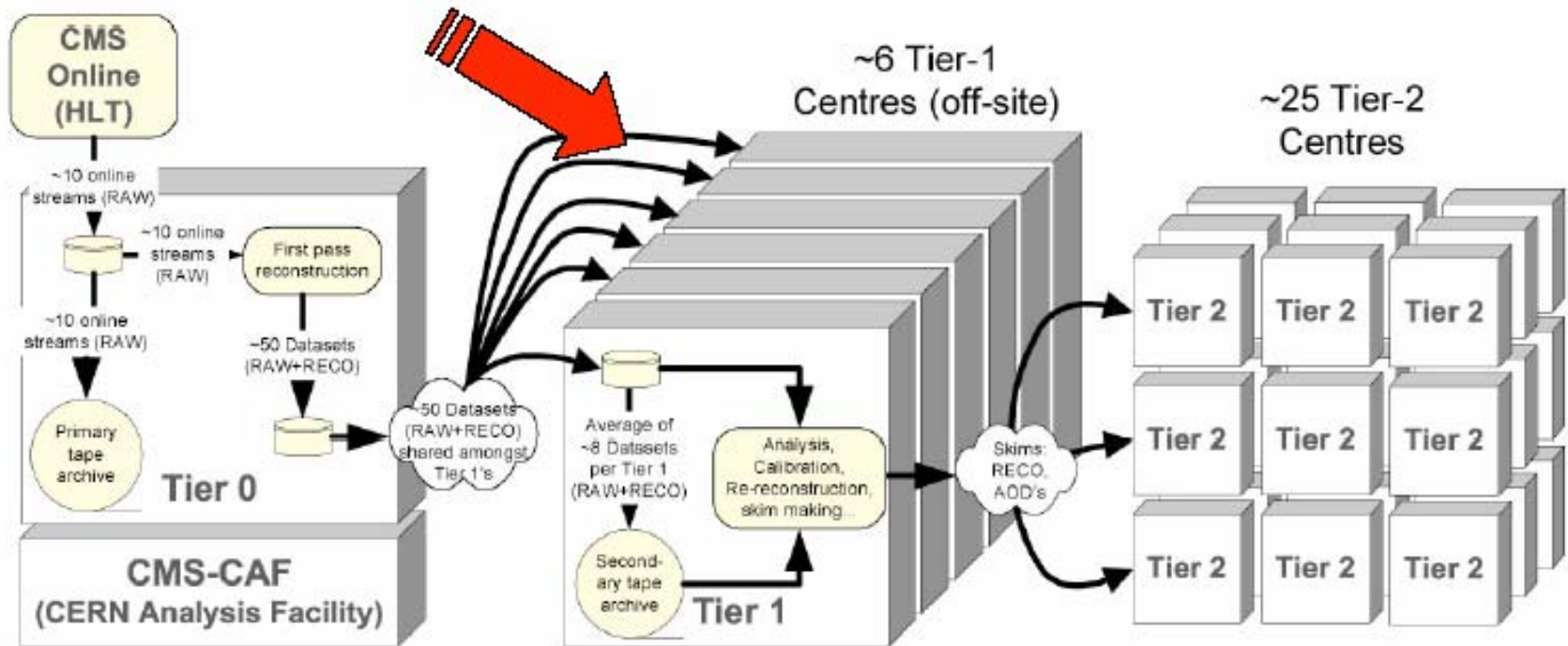
2. distributes RECOs and AOD to its associated group of next tier stage resources (Tier-2)

3. provides secure storage for MC events generated by the T2's

List: FNAL, FZK, ASGC, PIC, CNAF, RAL, IN2P3



# Tier-1 Centres



- ➔ Data curation; 'heavy lifting' (re - reco, skim, AOD extraction); raw data access; Tier-2 support
- ➔ Likely to include: ASCC, CCIN2P3, FNAL, GridKA, INFN-CNAF, PIC, RAL
- ➔ Nominally: CPU 2.5MSI2K; Disk 1.2PB; MSS 2.8PB; WAN >10Gb/s

# Tier-2

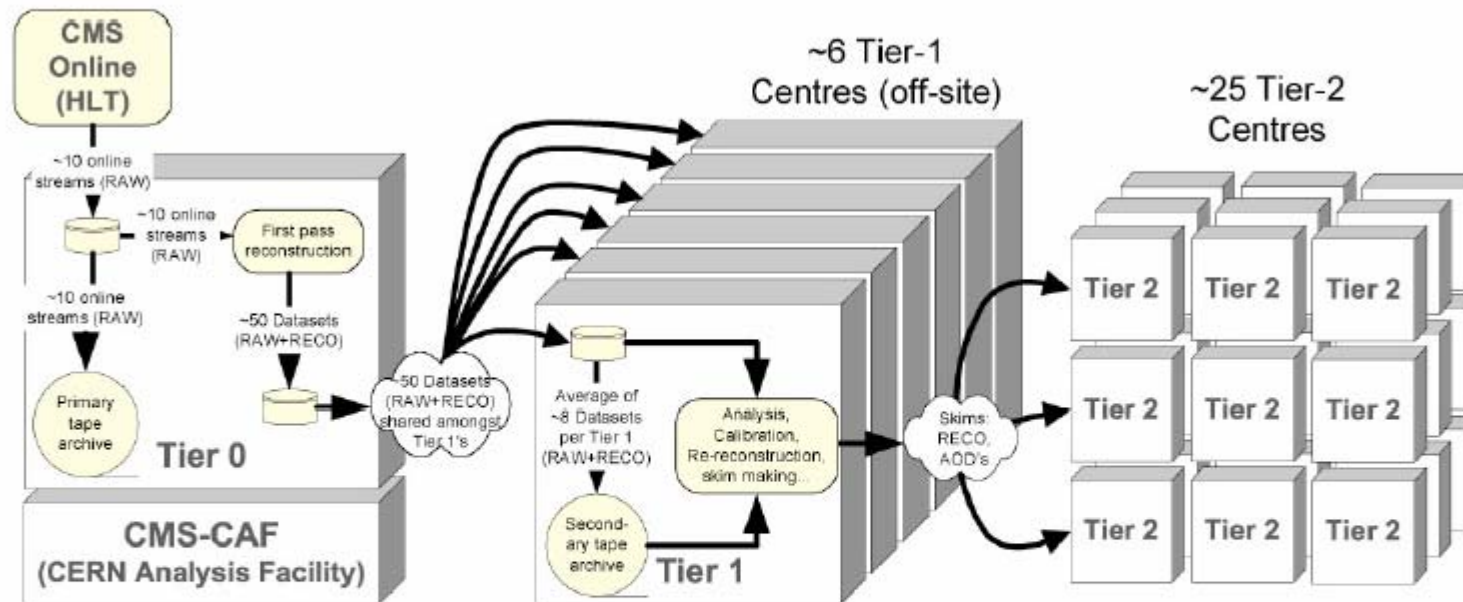
A more numerous set of smaller Tier-2 (T2) centres

- provide capacity for user analysis, calibration studies, and Monte Carlo production.
- T2 centers provide limited disk space, and no tape archiving.
- T2 centers rely upon T1s for access to large datasets and for secure storage of the new data (generally Monte Carlo) produced at the T2.

All T2	2007	2008	2009	2010	
CPU	10.4	20.8	35.5	56.3	MSi2k
DISK	2.7	5.4	10.8	16.2	PB
All T1	2007	2008	2009	2010	
CPU	7.4	14.9	21.4	40.5	MSi2k
DISK	3.9	7.8	11.8	17.7	PB
TAPE	5.9	17.7	29.6	41.4	PB



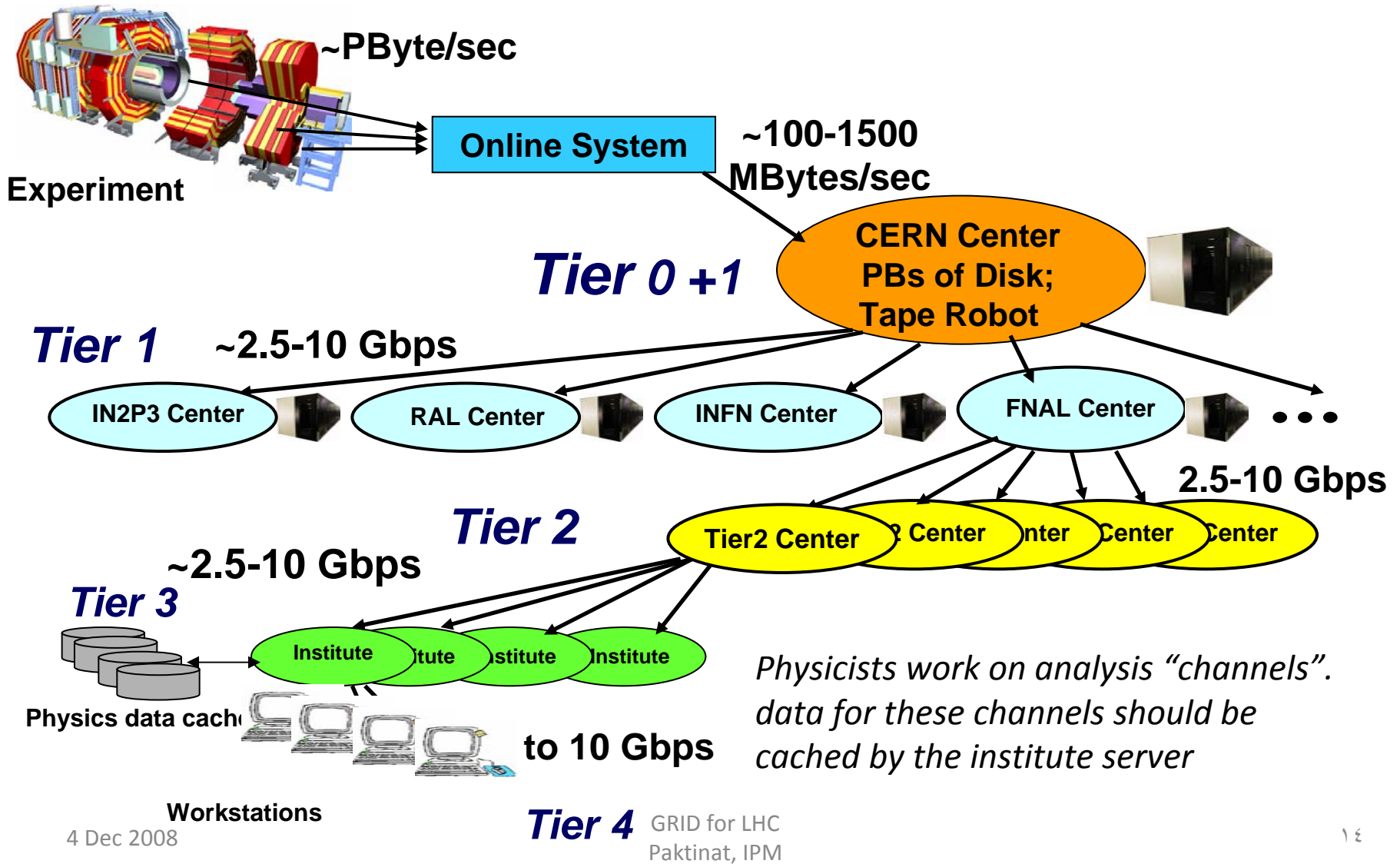
# Tier-2 Centres



- Analysis; MC generation; specialised support tasks
- Supplied by (federated) Institutes; local + common use
- Nominal: CPU 0.9MSI2K; Disk 200TB; No MSS; WAN > 1Gb/s



# CMS computing model



# CMS User tools

Normal users should only come in contact with

- DBS discovery page
- CRAB (CMS Remote Analysis Builder) Large scale analysis Tries to hide most Grid complexity Depends on DBS, Jobs are sent to where the data is to avoid data copying <http://mon.ihe.ac.be/trac/t2b/wiki/Crab>
- ProdAgent Handles larger productions Takes care of merging, publishing

DBS discovery :: Navigator :: Results

Physicist

Physics group: Any
Data tier: Any
Software release: Any
Data type: \*
Primary dataset: TauolaTTbar
Site: Any
Date: Any

Number of found datasets: 5
Full list of datasets: show
Datasets summary tables: show | hide

Number of datasets per page 10

Result page: 1

/TauolaTTbar/Summer08\_IDEAL\_V9\_AODSIM\_v1/AODSIM

Created 18 Oct 2008 08:40:11 GMT, contains 147000 events, 20 files, 1 block(s), 21.4GB, located at 7 sites (show, hide), LFNs: cff, py, plain, (L=N/A)
Release info, Block info, Run info, Conf. files, Parents, Children, Description, PhEDEx, Create ADS, ADS, crab.cfg

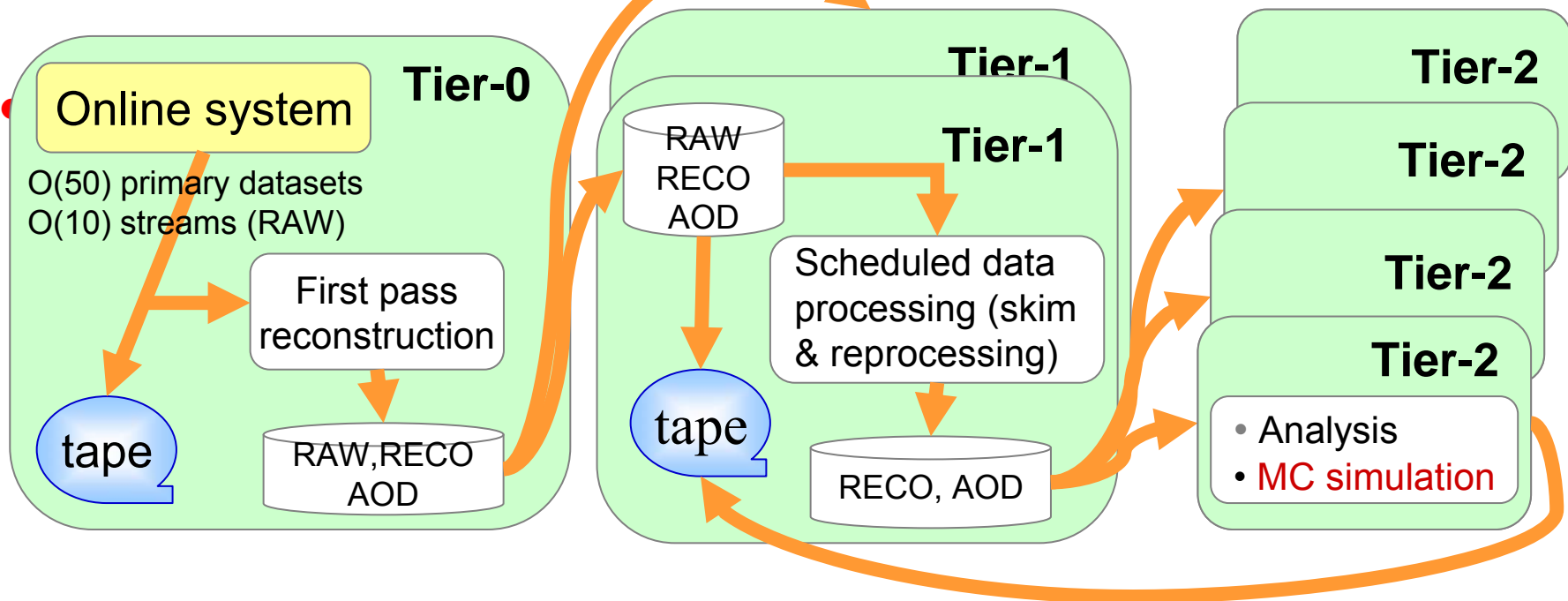
Table with 5 columns: Location, Events, Files, size, LFNs. Rows list various site locations like phedex.geol.uniovi.es, gridka-dCache.fzk.de, maite.iihe.ac.be, etc.

/TauolaTTbar/Summer08\_IDEAL\_V9\_v1/GEN-SIM-RAW

Created 22 Sep 2008 08:03:38 GMT, contains 183000 events, 229 files, 4 block(s), 347.4GB, located at 2 sites (show, hide), LFNs: cff, py, plain, (L=N/A)
Release info, Block info, Run info, Conf. files, Parents, Children, Description, PhEDEx, Create ADS, ADS, crab.cfg

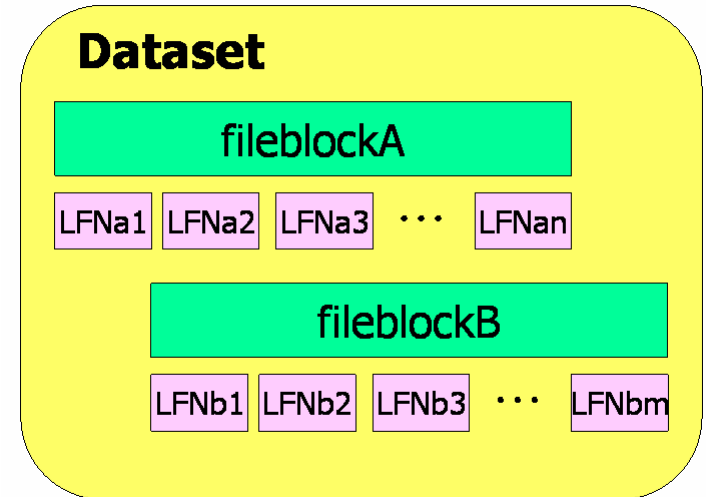
# The CMS Computing Model

- Distributed computing model for data storage, processing and analysis
- Grid technologies (Worldwide LHC Computing Grid, WLCG)
- Tiered architecture of computing resources
- Several Petabytes of data (real and simulated) every year



# Data Organisation

- Total volume of data
  - ~ 10 PB/y
- Datasets
  - ~ 0.1 – 100 TB each
  - size driven by physics
  - composed of
- Fileblocks
  - ~ 1 – 10 TB each
  - composed of
- Files
  - ~ 1 GB each
  - avoids scaling issues for MSS and catalogues



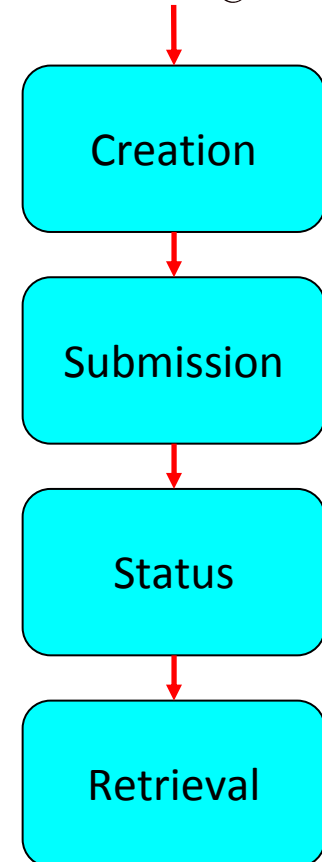
Data Tier	Description
GEN	Event kinematics
SIM	Simulation with detector material
DIGI	Simulation of detector readout
RAW	Raw detector output
RECO	Reconstructed event
AOD	High level analysis objects
USER	Final user files

# Data analysis

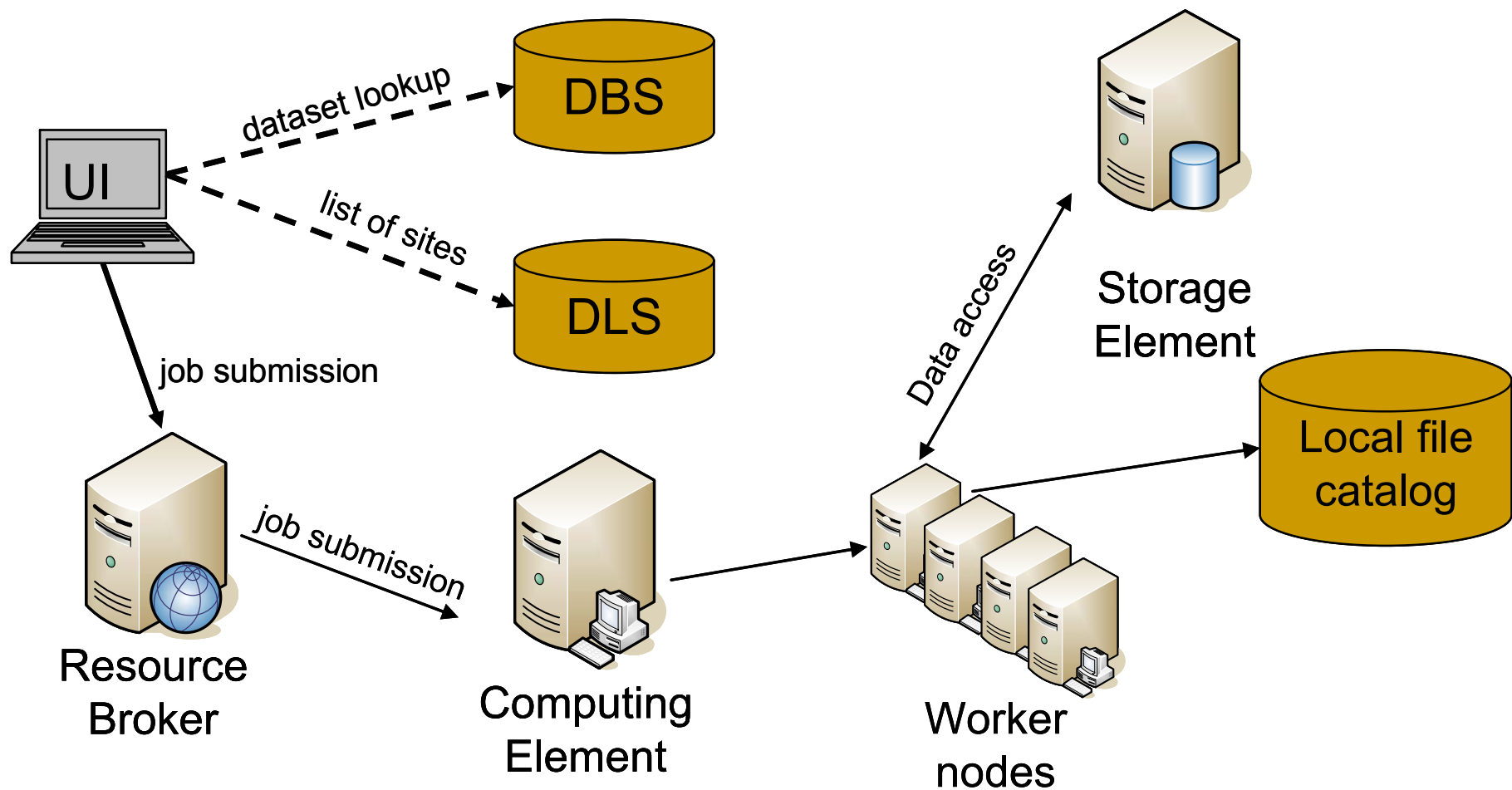
- Definition
  - The process of extracting information from real or simulated data and writing it in a form more convenient for physics studies
- Classification
  - Organized analysis
    - official data processing (RAW→RECO→AOD)
    - Reprocessing, skimming
    - managed by a few people
  - Chaotic analysis
    - RECO, AOD → root files
    - done by any CMS member
- Scale
  - O(10) PB/year
  - O(2000) users
  - O(50) sites

# Typical analysis workflow

- The user develops his analysis code and builds the executable and the libraries
  - done on a UI
- The user defines which data wants to analyse
  - data registered in DBS
- The user submits an analysis task to the Grid (or a local farm)
  - transparent access to different Grids
- The task is divided into many parallelised jobs
  - jobs allocated to run near the data
- The user checks the status of his task
- When the task is finished, the user collects the results
  - stored on user space or a SE



# Analysis workflow with CRAB





```
SSH Secure Shell 3.2.9 (Build 283)
Copyright (c) 2000-2003 SSH Communications Security Corp - http://www.ssh.com/

This copy of SSH Secure Shell is a non-commercial version.
This version does not include PKI and PKCS #11 functionality.
```

```
*****
*                               *
*       The LXPLUS Public Login Unix Service       *
* http://cern.ch/ComputingRules : Govern the use of CERN computing facilities *
*****
```

```
Last login: Tue Nov 11 12:30:09 2008 from 194.225.68.66
*****
*                               *
*       The LXPLUS Public Login Unix Service       *
*       (Scientific Linux SLC 4.6 x86_64)          *
*                               *
* A web page containing information about this Linux version on LXPLUS:   *
* https://cern.ch/plus/SLC4.html                                           *
* For a list of known issues, please check:                               *
* https://cern.ch/plus/issues_SLC4.html                                   *
* In case of problems, please contact the helpdesk: tel 78888            *
* If you have any feedback not already included there please send it to: *
* it-dep-fio-lxslc4@cern.ch                                              *
*                               *
* In                                                                           *
* https://cern.ch/plus           : Information on the usage of LXPLUS/LXBATCH *
* https://cern.ch/ComputingRules : Govern the use of CERN computing facilities *
*                               *
*****
```

[lxplus229] ~ >