How to find and store your data ?

CMS Grid Analysis Tutorial

CERN

June 25th, 2009

* Please also consult the tutorial lessons about Crab

Thomas Kress, III. Institute B

RWTH AACHEN UNIVERSITY

Role of the Tier Sites in CMS

- Tier-0 and T-1 centers used for central tasks by operation teams, no general user analysis access
- Massive MC production at the T-2 sites
- Tier-2 sites (and with restrictions CAF and CERN) only location where users are able to obtain guaranteed access to data samples
  - *Distributed primary datasets*
  - *DPG/POG/PAG groups associated with 3-5 T-2s each, 'group data' (e.g. group skims, PAT)*
  - *O(30-50) local users at every T-2*
- Tier-3 technical term is not well defined
  - *[almost no, …, full] Grid functionality*
  - *cluster is fully under local control*

Currently under review, adaption

CERN-LHCC-2005-023
CMS TDR 7
20 June 2005

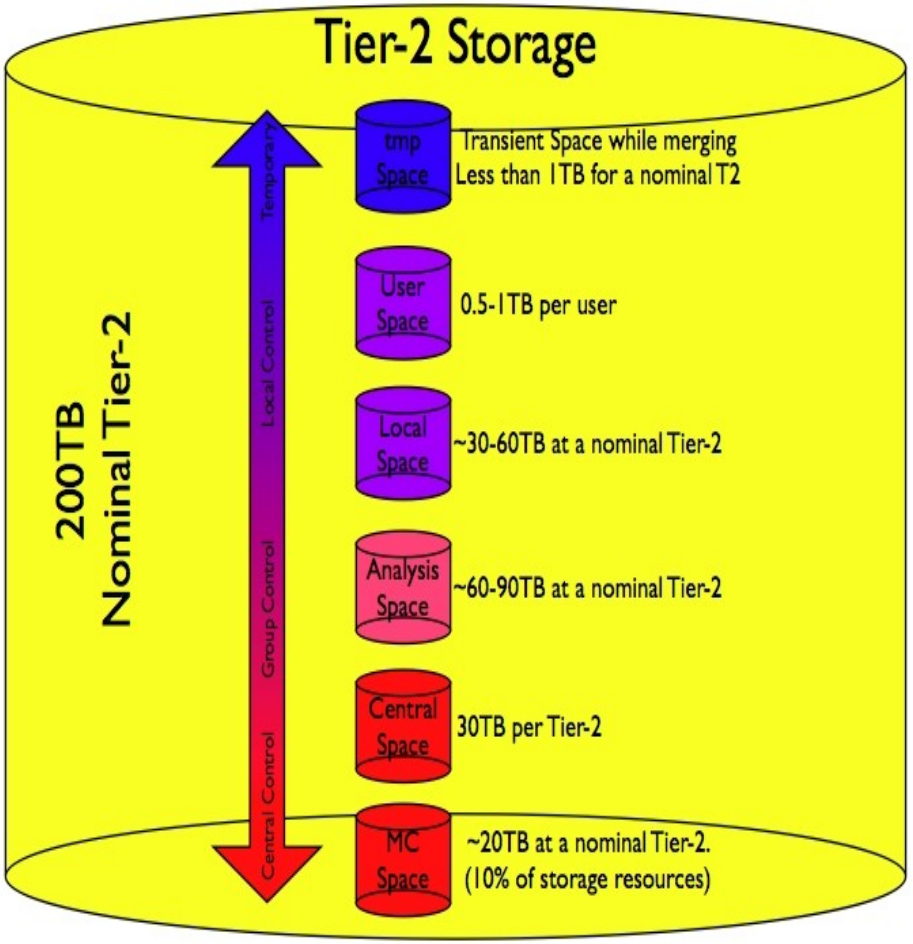
CMS

The Computing Project

Technical Design Report

Tier-2 Storage Concept 2009

- Different levels of dynamics, control, and responsibilities
- N * 0.5 -1 TB user home space
  - *Stage-out area for e.g. CRAB jobs*
  - *Additional local/national resources*
- 30 – 60 TB local/national storage
- N * 30 TB group analysis space
  - *N associated DPG/POG/PAG groups*
  - *Data of interest for the group, e.g. secondary skims, special MCs, RAW*
- 30 TB central space
  - *Primary data sets/skims, MC samples of global interest for CMS*
  - *Distributed centrally*
- Some buffers for central MCP

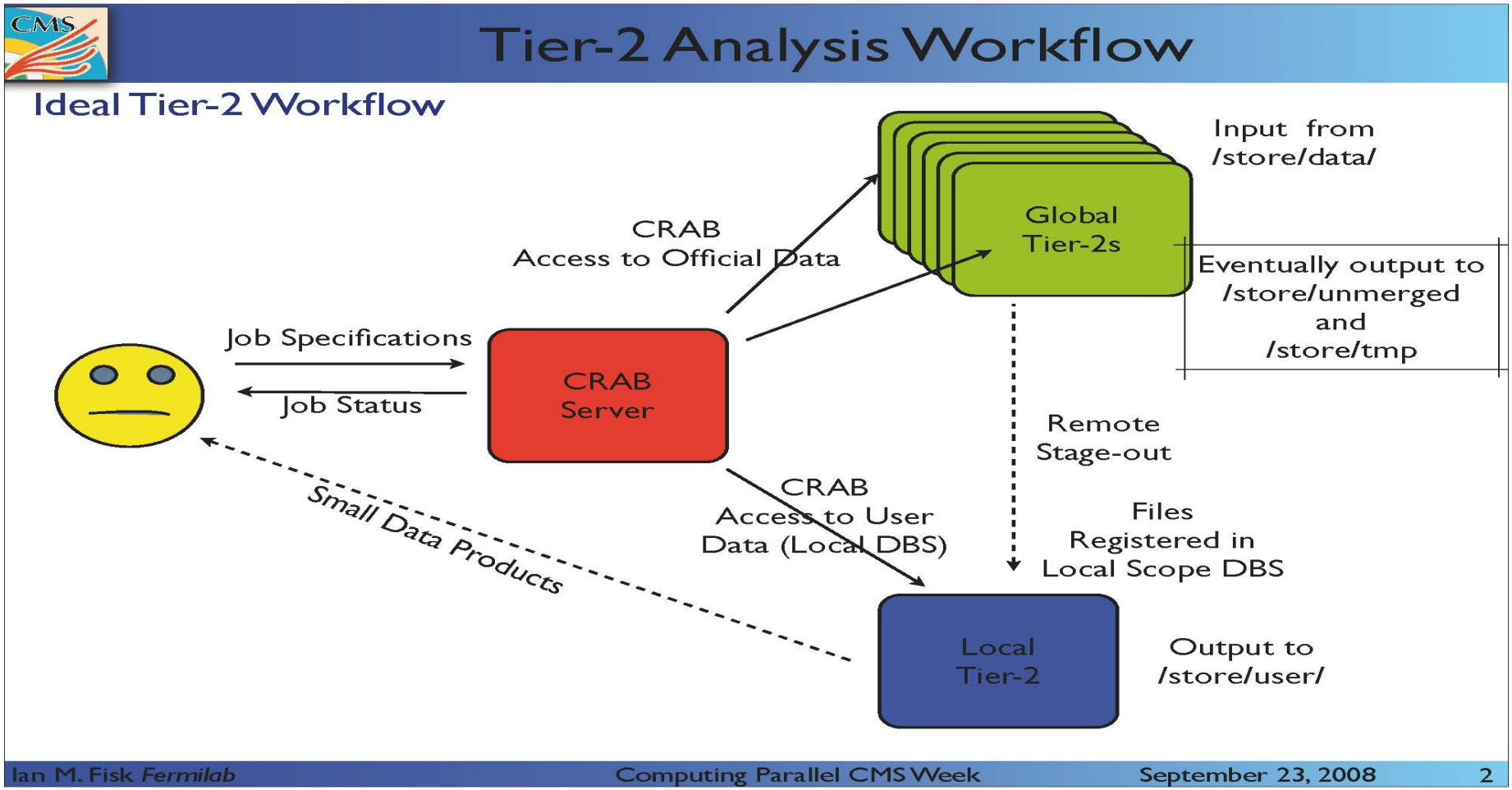


Ian Fisk

Tier-2 CPU & Batch Concept

- Generally 50% of CPU resources (of CMS pledges) are allocated to MCP, the other half is for CMS user analysis
- So far only moderate usage of VOMS groups & roles
  - *High batch fair shares for central CMS tasks (MCP, software installation, monitor jobs, …)*
  - *Some Tier-2 sites give higher fair shares to their national users (e.g. Gemany), if they authenticate by appropriate VOMS extention*
    - `To account for additional national or local resources`
    - `voms-proxy-init -voms cms:/cms/dcms (for German VOMS group)`
  - *VOMS for users associated with a (DPG/POG/PAG) detector or physics group hosted at the Tier-2 not yet implemented, but just decided to enable burst access for a small number of group users for central or high priority group tasks*

Analysis Concept



Idea: data distributed, but every user can transparently access data at every Tier-2, output to home T2

Dataset Glossary *(from DBS documentation)*

- Dataset: /PrimaryDataset/ProcessedDataset/Tier1-Tier2-...
  - A set of files representing a coherent sample. Datasets are defined primarily by processing history and event selection criteria.
- Primary Dataset
  - Data at all levels of processing pertaining to a given (HLT event classification) or common MC production (generation with same parameter) criteria.
- Proccessed Dataset
  - A slice of data from a Primary Dataset defined by the processing history applied to it. A Proccessed Dataset will correspond to a large production of data with a single major software release version, but may include multiple minor versions of small bug fixes and also may contain the output of multiple processings of some given input data.
- Data Tier (not to be confused with Grid Tiers !)
  - A set of objects to be grouped together in the output files of the processing step producing the objects.

Dataset Glossary ... con't

- File Block
  - A slicing of a dataset into chunks of files likely to be accessed together. The File Block us a data management packaging unit for convenience of the data location and transfer services.
- PAT is not a separate Data Tier, it appears as .../..PAT../USER
- Examples /PrimaryDataset/ProcessedDataset/Tier1-Tier2-...
  - /BB2MuMu_noMassCut/CMSSW_1_6_7-CSA07-1204465719/GEN-SIM-DIGI-RAW
  - /BB2MuMu_noMassCut/CMSSW_1_6_7-CSA07-1204787696/RECO
  - /QCDDiJetPt300to380/Summer08_IDEAL_V9_AODSIM_v1/AODSIM
  - /BeamHalo/BeamCommissioning08-PromptReco-v1/RECO
  - /Exotica_GMSB_GM1f/Summer08_IDEAL_V9_PAT_v1/USER

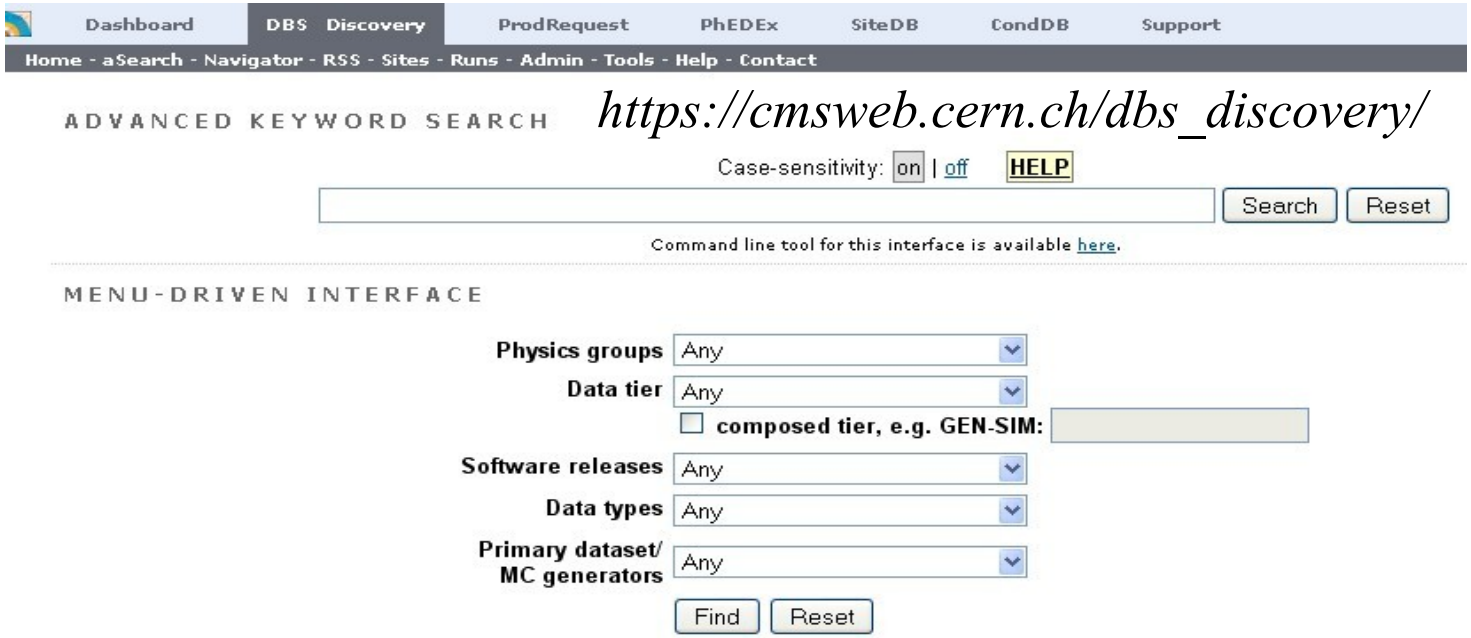
Which Dataset is Interesting for Me ?

- From the name alone, often not really obvious, neither for signal nor for bg
- Sometimes documentation from the PAG/POG/DPG groups
- Sometimes Monte Carlo production information
  - *e.g. http://mthomas.web.cern.ch/mthomas/Summer2008production.htm*
- Sometimes information about used important parameters, objects, .. stored in database
- Subscribe to your group's (and to hn-cms-physics-announcements@) Hypernews communication system !
- Not all datasets are compatible with your used CMSSW version !

DBS Discovery Page

Data Bookkeeping System

- Menu-driven browser to find datasets in database system
  - *In addition powerful command line interfaces available f. advanced functionality*
  - *Authoriative source of information about data produced and stored in CMS, with Oracle and MySQL DB back-ends. Different instances (prod-global, „local-scope“)*
  - *Meta-data system: infomation about data (names, size, creation date, sites, ...)*

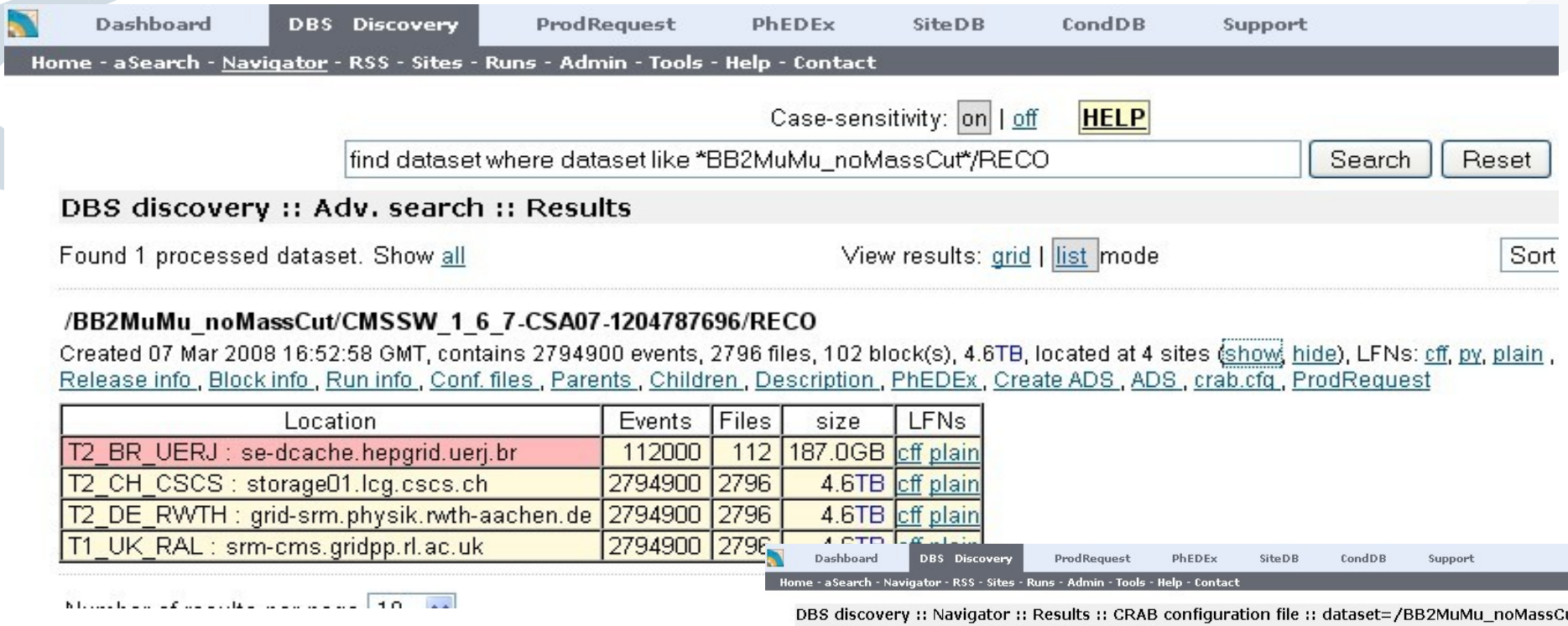


view

Pull-down menu to select global ... local instances

DBS Query Language

- find <key1, key2, ..> where <key> <op> <value> and/or …
  - Complex queries possible with little effort
  - Supports many keywords (+ attributes): dataset, site, run, lumi, …
  - Supports boolean expressions and comparison operators
  - Supports patterns and regular expressions
  - Supports date stamps
- Some examples:
  - find dataset where site = T2_DE_RWTH
  - find dataset where dataset like /BB2MuMu_noMAssCut*
  - find dataset where dataset like *Cosmics* and release > CMSSW_2_0
  - find sum(file.size) where dataset like /BeamHalo/*/RAW



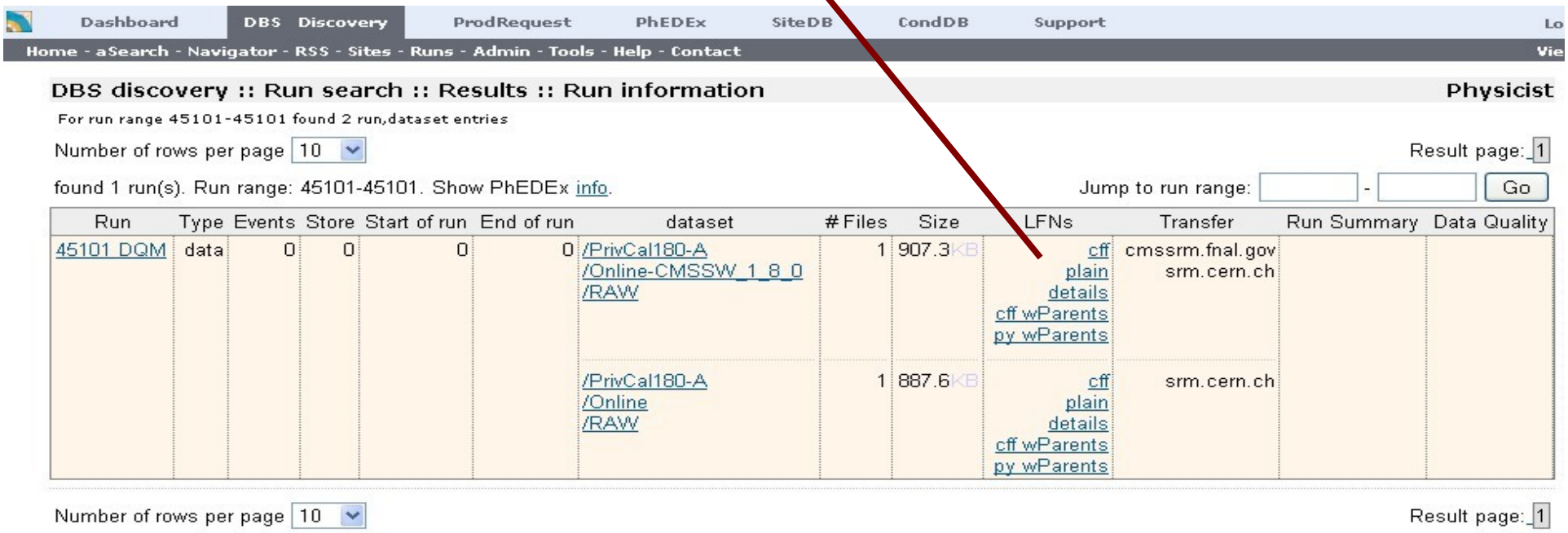
Show: table with site specific information
Conf.files: important information about e.g. used MC parameters, objects, …
crab.cfg: generates Crab config file

DBS Run Search

/store/data/ (or /store/mc/) LFN is mapped to transparent internal Tier-2 directory structure mapping info in site's trivial file catalogue TFC

LFN (logical file name):
/store/data/PrivCal180/A/000/000/000/RAW/0000/
E80B1ACE-9E2D-DD11-AF0C-000423D6B358.root

/store/mc/ for MC



Desired Input Dataset Discovered

- Datasets at the Tier-1 sites can not be accessed by Crab user analyses to protect central work-flows and tape systems
- For the CERN analysis facility (CAF) and other CERN resources there are / will be soon access restrictions
- For a standard user analysis the sample has to be at one/more Tier-2 sites (or at „your CMS/Grid-enabled T-3“, if existing and you have access)
  - *If it is available at least at one Tier-2, Crab will find it automatically*
  - *Otherwise it has to be replicated from a Tier-1 to a Tier-2 site*

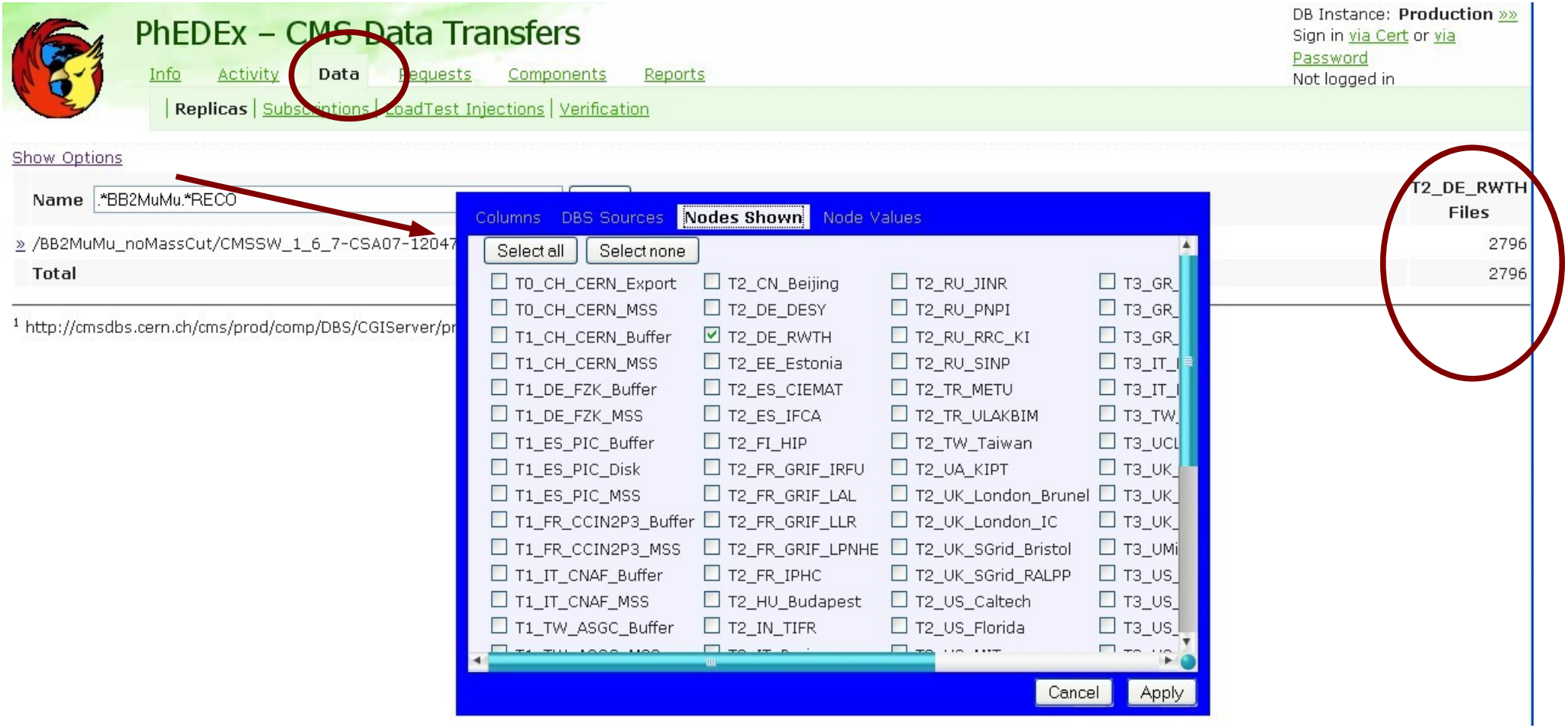
Phedex

Physics Experiment Data Export *http://cmsweb.cern.ch/phedex/*

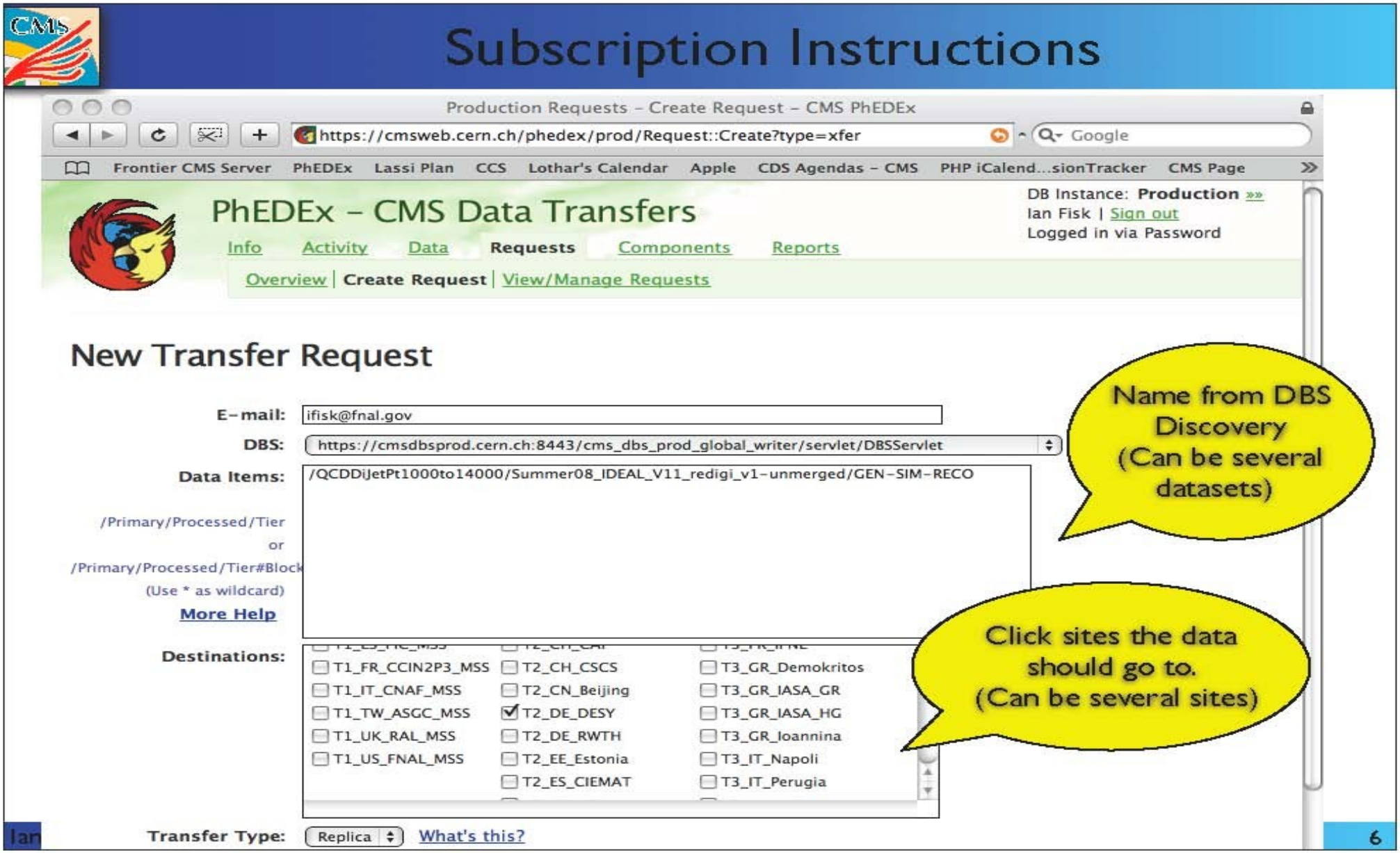
- Phedex (based on FTS & SRM) moves/copies/registers/deletes Grid data
- Similar to DBS, it has location, information, … query functionality
- E.g. to replicate data from T-0/T-1s to Tier-2 sites for analysis purposes
- Someone requests, local Tier data managers approve or deny
- Mainly used by responsibles for Tier data transfers of global interest
  - *Central data area → DataOps team*
  - *Physics & Detector group area → group's data managers*
  - *Local T-2/3 space area → local/national T-2/3 community data managers*
- Also you can request a dataset to be replicated into a Tier-2 group or local T-2 (or sometimes T-3) storage area, after prior consultation with:
  - *Local/national Tier representatives*
  - *Your associated detector or physics group's data managers or convenors*

Phedex Dataset Query

- An alternative to a DBS query

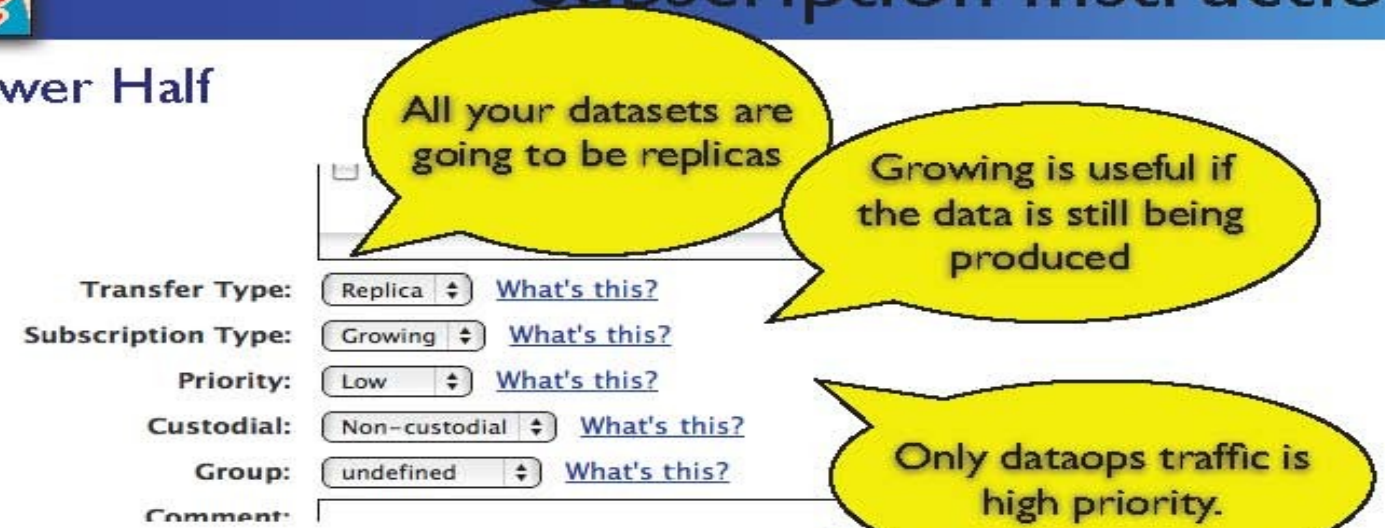


Subscription Instructions



Subscription Instructions

Lower Half

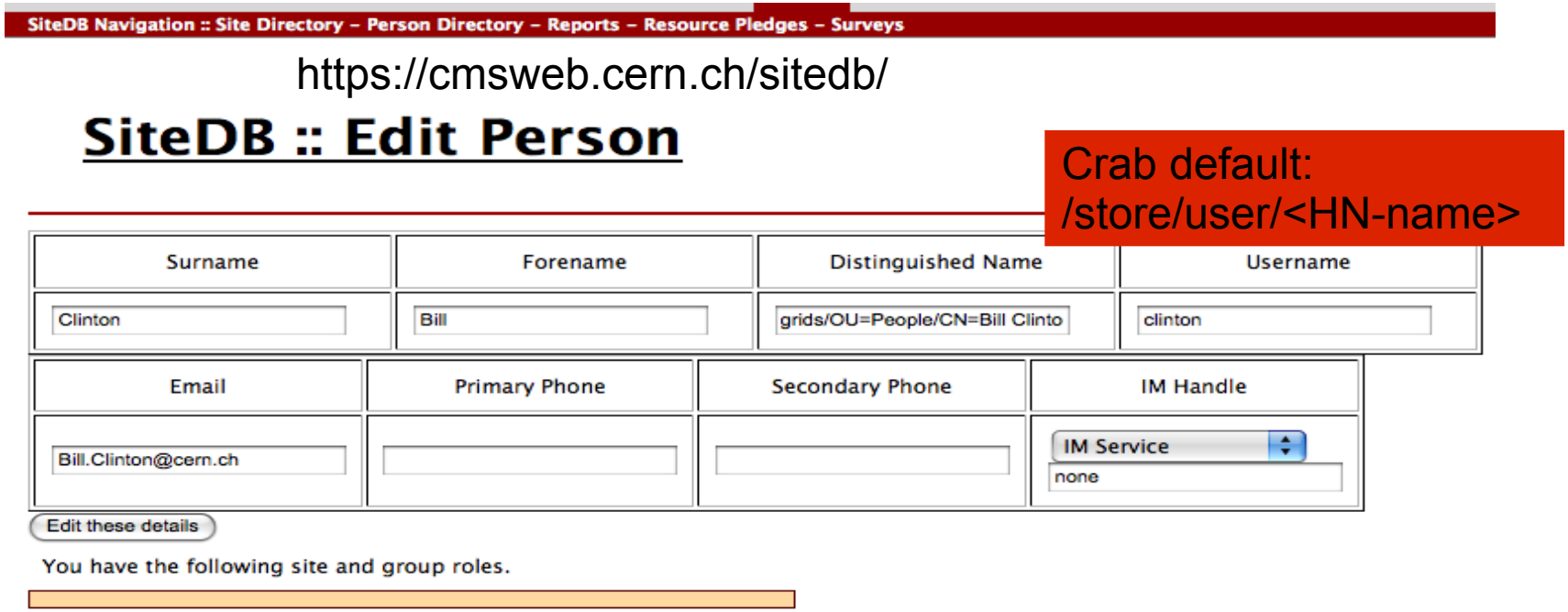


Groups are listed in pull down menu.

- DataOps is central data
- The other categories are for physics groups
  - Will ask for a group called "Local"

Where to Store Your Output ?

- Usually you run your analysis by Crab
- For most purposes the Sandbox is too small, so use Crab's „stage out“ !
- For using the Crab default, store your DN in the SiteDB, this maps your job certificate to /store/user/<HN-name> for the stage out

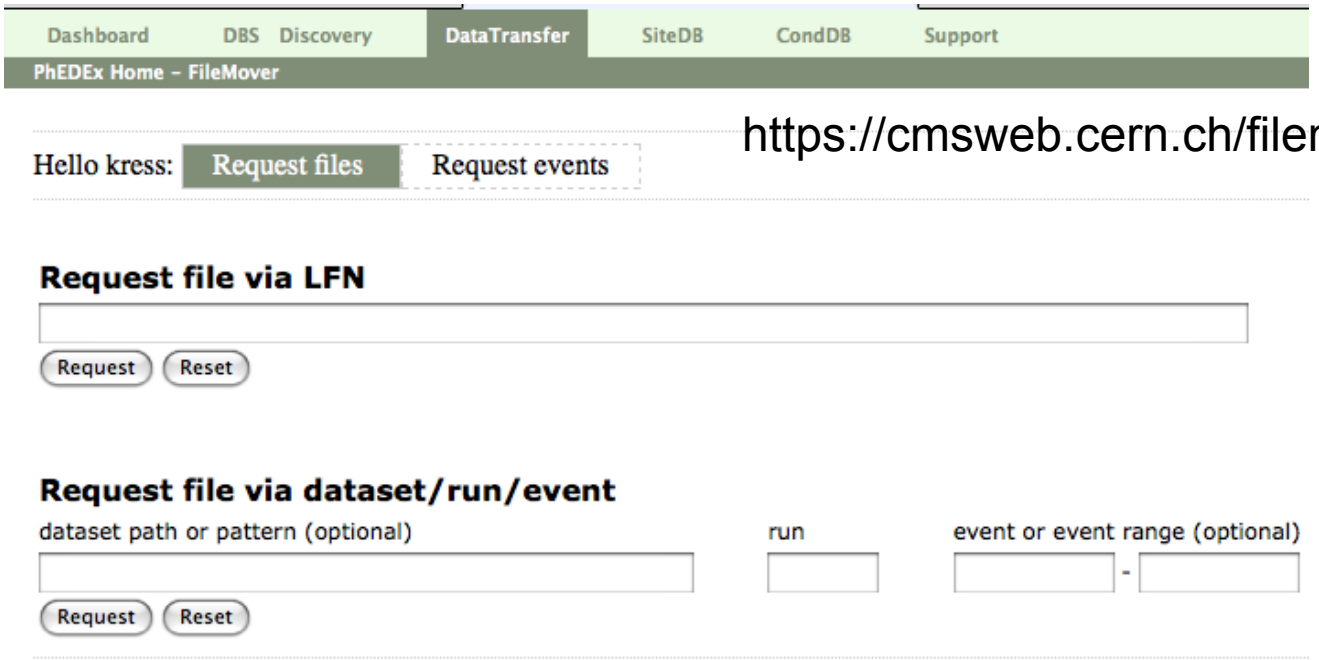


Grid User Home Storage Space

- Every user gets Grid home storage space /store/user/<HN-name>
  - (For 2009) 0.5 - 1 TB (t.b.d.) from the CMS pledges
  - Usually at his/her local/national Tier-2
  - Could also be a „CMS Grid capable“ Tier-3 (but only best effort support !)
  - Special arrangements for CERN users and users without a national T-2 or fully functinal T-3 are currently under discussion
  - The 0.5 - 1 TB /store/user/ are the guaranteed resources, some users will have more by additional national or local funds, might be spread over several sites
- Contact your local/national Tier representative; for some sites the access to /store/users/ has to be granted individually
- So far storage systems do not have quotas, so Tier-2 admins will review /store/users/ regularly, but do not overbook without consultation !
- In the personal SiteDB a default „home“ Tier-2 is not stored, you have to specify it (or another site you have access to) in the Crab config file

Access to /store/user/ Data

- If Crab registration option is used, data are registered in local-scope DB, can then be accessed by you and others via further Grid/Crab jobs
- Some T-2 sites allow direct read access e.g. from local desktops
- Not for extensive usage, data can be copied over to local desktop/notebook by Grid copy commands or by using FileMover



https://cmsweb.cern.ch/filemover

/store/user/ Data Deletion

- Crab dataset registration in „local-scope“ DBS (= local instance of CERN DBS or an DBS instance at a Tier-2) and storing the data at a Tier-2 is almost fully automized
- A deletion of /store/user/ datasets and cleaning of the corresponding local-scope DBS entries is however not yet possible in a same convenient way
  - *Needs knowlege of the Tier-2 local directory structure and the use of Grid commands (e.g. srm-rm) from an UI to delete the data on file system level*
  - *Needs priveledges and knowlegde how to de-register files from a local-scope DBS*
- Data at a Tier-2/3 is usually saved on hardware level (e.g. disk Raid) but usually no backup (or user access to e.g. tape systems at T1 sites)
  - *It is possible that /store/user/ data is lost in case of storage system crashes*
  - *Opposite to Phedex data (central, groups), for /store/user/ there is no tool available to synchronize local-scope DB with storage area*