

# LHCb Experiences & Plans

Nick Brook

SC3 experiences

Analysis

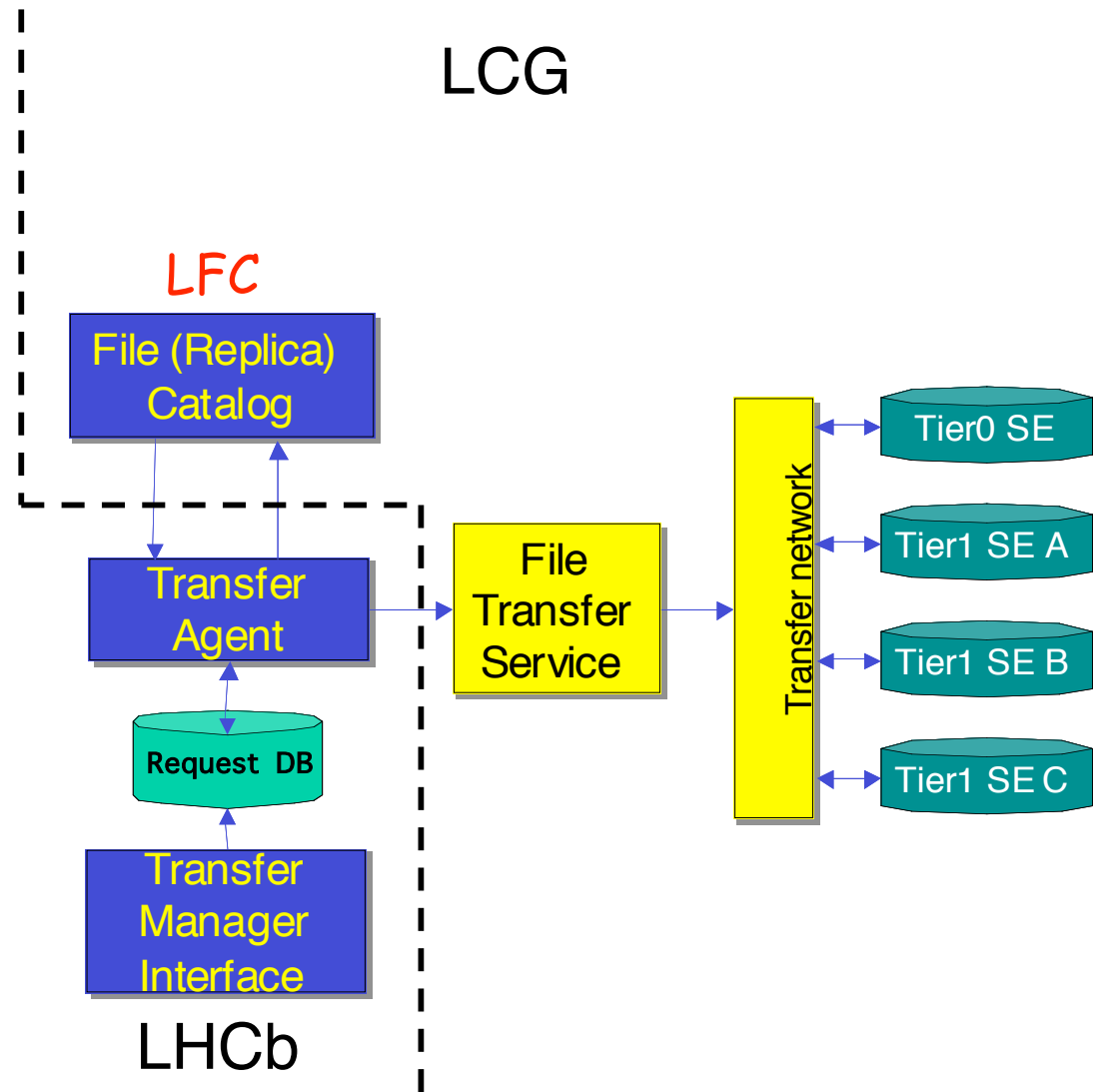
Future plans

## SC3 Aims - original planning

- **Phase 1: (Data Moving)**
  - Demonstrate Data Management to meet the requirements of the Computing Model
  - Planned: October-November
- **Phase 2: (Data Processing)**
  - Demonstrate the full data processing sequence in real time
  - Demonstrate full integration of the Data and Workload Management subsystems
  - Planned: mid-November + December

# LHCb Architecture for using FTS

- Central Data Movement model based at CERN.
  - FTS+TransferAgent+RequestDB
- TransferAgent+ReqDB developed for this purpose.
- Transfer Agent run on LHCb managed central machine at CERN



## Phase 1

1. Distribute stripped data Tier0 → Tier1's (1-week). 1TB
  - The goal is to demonstrate the basic tools
    - Precursor activity to eventual distributed analysis
2. Distribute data Tier0 → Tier1's (2-week). 8TB
  - The data are already accumulated at CERN
  - The data are moved to Tier1 centres in parallel.
  - The goal is to demonstrate automatic tools for data moving and bookkeeping and to achieve a reasonable performance of the transfer operations
3. Removal of replicas (via LFN) from all Tier-1's
4. Tier1 centre(s) to Tier0 and to other participating Tier1 centers
  - data are already accumulated
  - data are moved to Tier1 centres in parallel
  - Goal to meet transfer need during stripping process

# T1-T1 Channel Status

Finally configured at all sites Jan 4<sup>th</sup>

## Key

### T1 Site

FTS Server Status

Configuration of Channel Management

### RAL

FTS Server

Manage Incoming Channels

### IN2P3

FTS Server

No Incoming Channels

### PIC

No FTS Server

Channels Managed by Source SE

### SARA

FTS Server

Manage Incoming Channels

### FZK

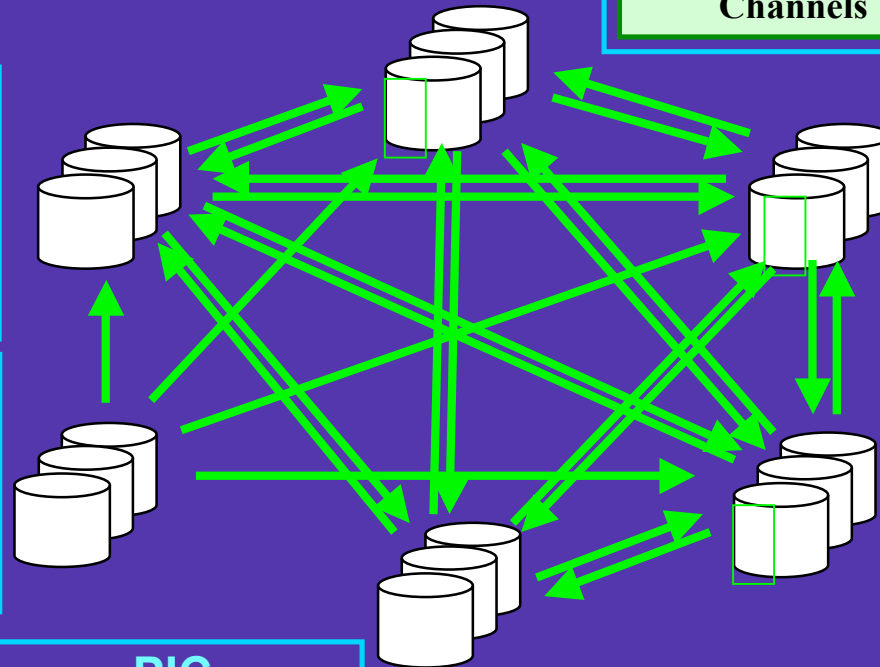
FTS Server

Manage Incoming Channels

### CNAF

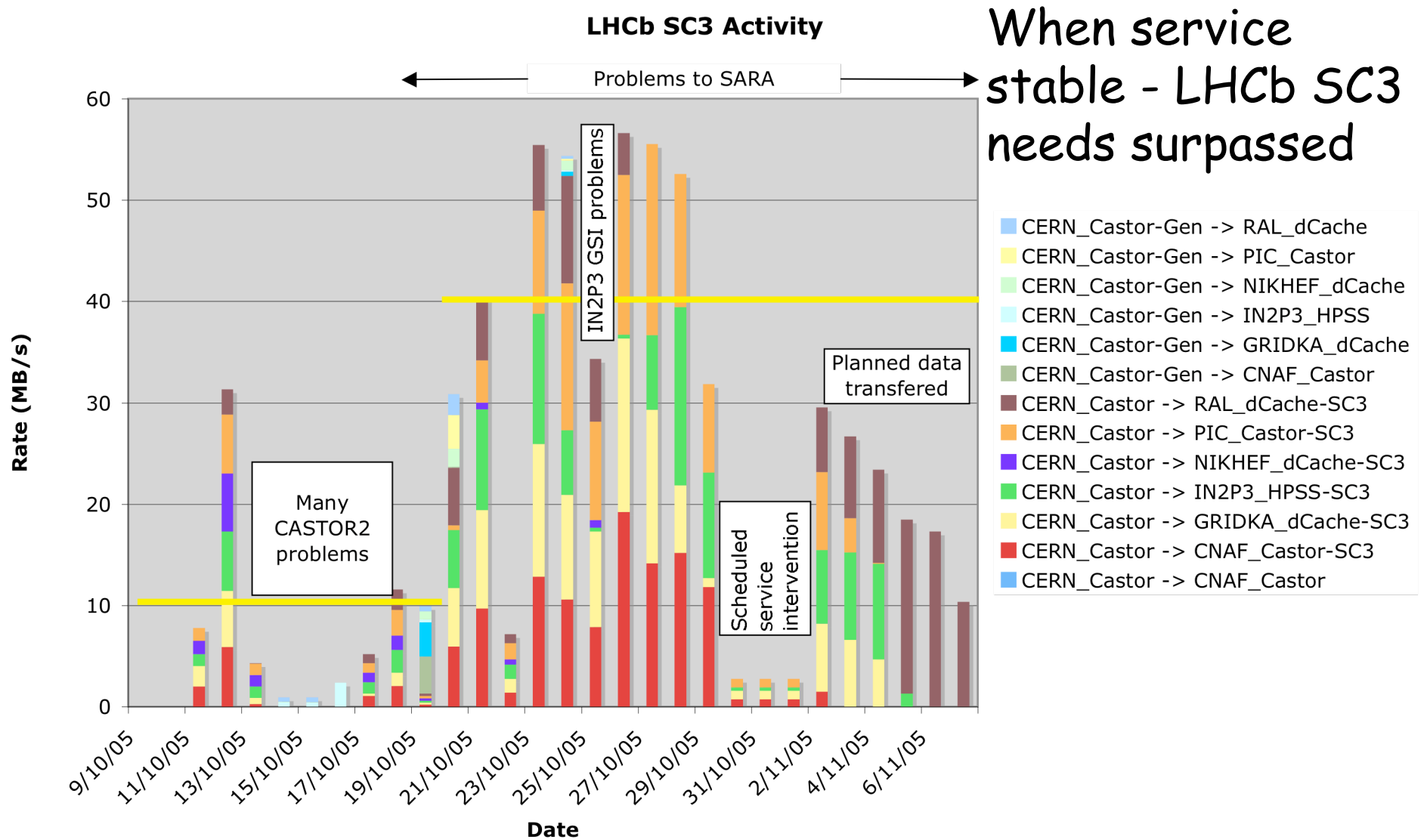
FTS Server

Manage Incoming Channels



FTS central service for managing T1-T1 matrix

# Overview of SC3 transfer activity

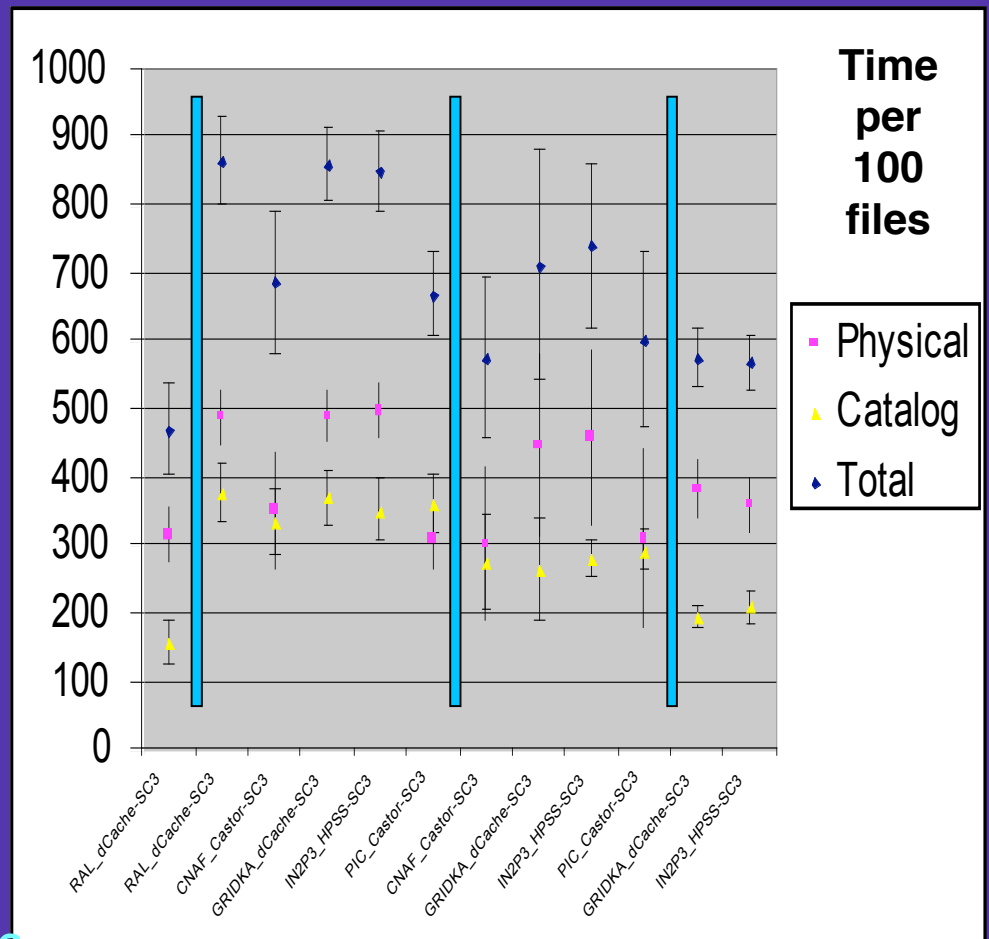


Wish to re-run T0-T1 transfers as part of SC3 re-run

# File removal

## Four phases:

- RAL only
  - -RAL, CNAF, GRIDKA, IN2P3, PIC
  - -CNAF, GRIDKA, IN2P3, PIC
  - -GRIDKA, IN2P3
- 
- 50k replicas removed in ~28 hours
    - 10k replicas at each site
    - Each site with its own agent
  - problems in the current middleware
    - Removing remote physical files is slow
      - Bulk removal operations are necessary
    - Different storage flavours are showing slightly different functionality :
      - No generic way to remove data on all the sites
    - LFC File catalog (un)registration operations are slow
      - Bulk operations within secure sessions are necessary



## Phase 1

1. Distribute stripped data Tier0 → Tier1's (1-week). 1TB
  - Succeeded but not a rate given in metrics of success
2. Distribute data Tier0 → Tier1's (2-week). 8TB
  - Succeeded - (nearly) achieved "acceptable" metric but not "success" metric
  - Like to repeat as part of the SC3 re-run
3. Removal of replicas (via LFN) from all Tier-1's
  - Failure to meet 24 hr metric Inconsistent behaviour of SRM, bulk operations needed, ...
4. Tier1 centre(s) to Tier0 and to other participating Tier1 centers
  - failure
  - FTS did not support third party transfer
  - Complicated T1-T1 matrix been set up - yet to test

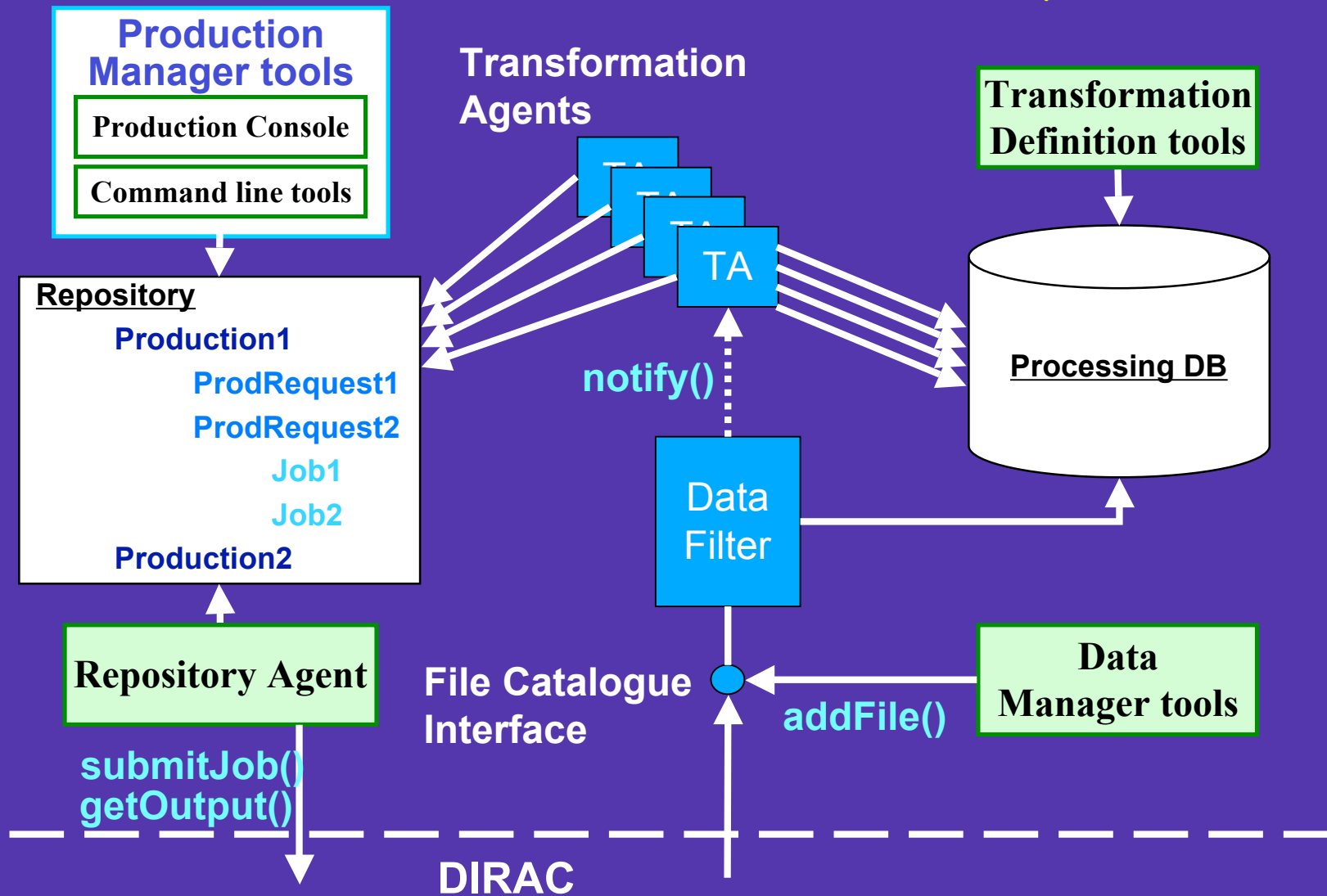


## Phase 2

- MC data production in T2 centers
  - Limit the production jobs to just T2 centers
  - In-job data transfers to the corresponding T1 centers
- This step will be done with the current Production Manager tools
- Stripping jobs submitted at T1 centers as soon as the reconstructed data become available
  - Uses automated job submission tools
- The stripped data distribution over the T1 centers
  - Done as part of the job workflow
  - Asynchronous transfers in case of failures

## Phase 2

Tests underway  
before Christmas



## General Summary of SC3 experiences

Extremely useful for shaking down sites, experiment system & WLCG

- Many new components used for the 1<sup>st</sup> time in anger
- Need for additional functionality in services
  - FTS, LFC, SRM, ...

Reliability seems to be the major issue

- CASTOR2 - still ironing out problems, but big improvements
- Coordination issues
- Problems with sites and networks
  - MSS, security, network, services...

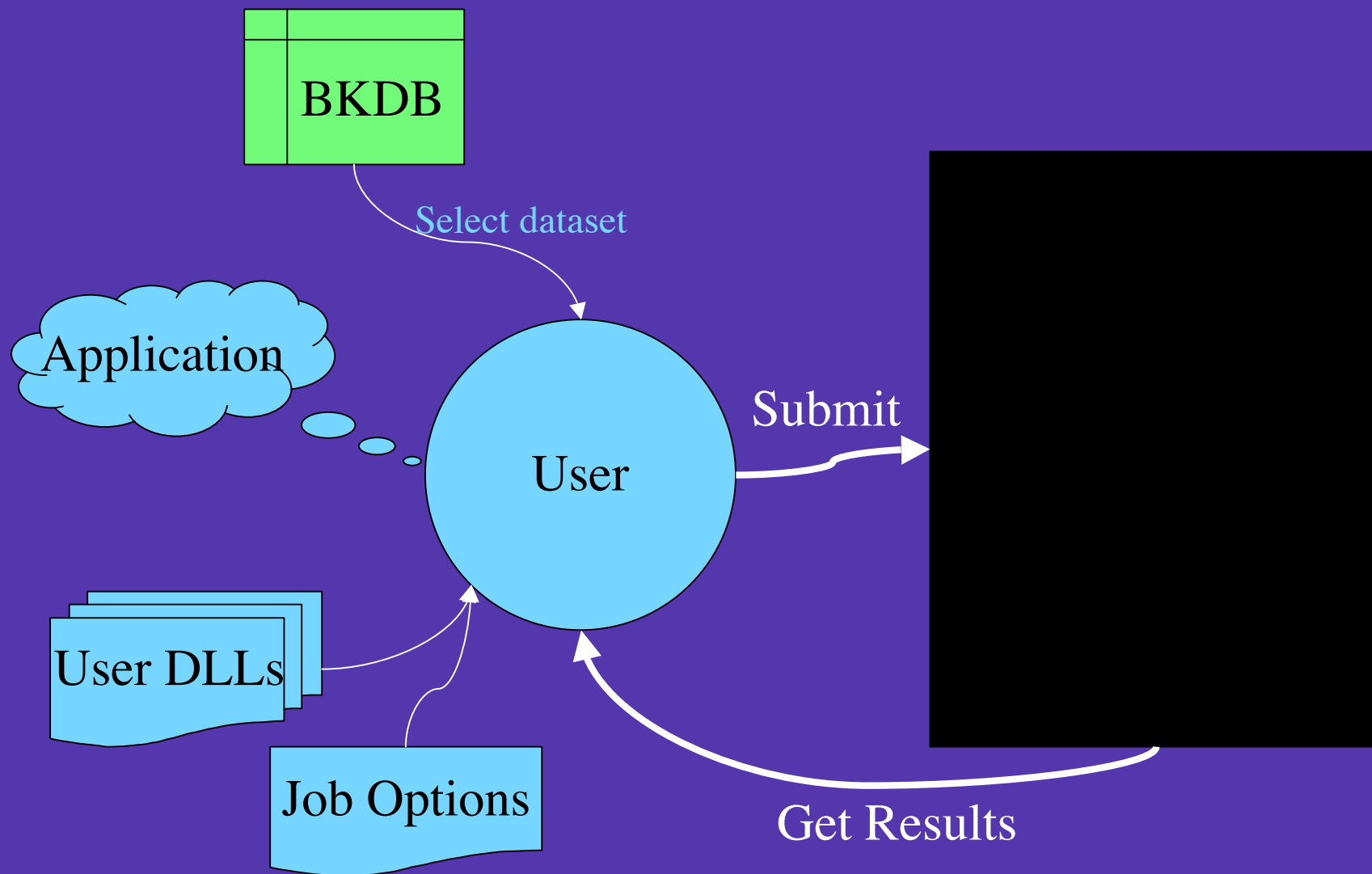
FTS:

- For well-defined site/channels performs well after tuning
- Timeout problems dealing with accessing data from MSS

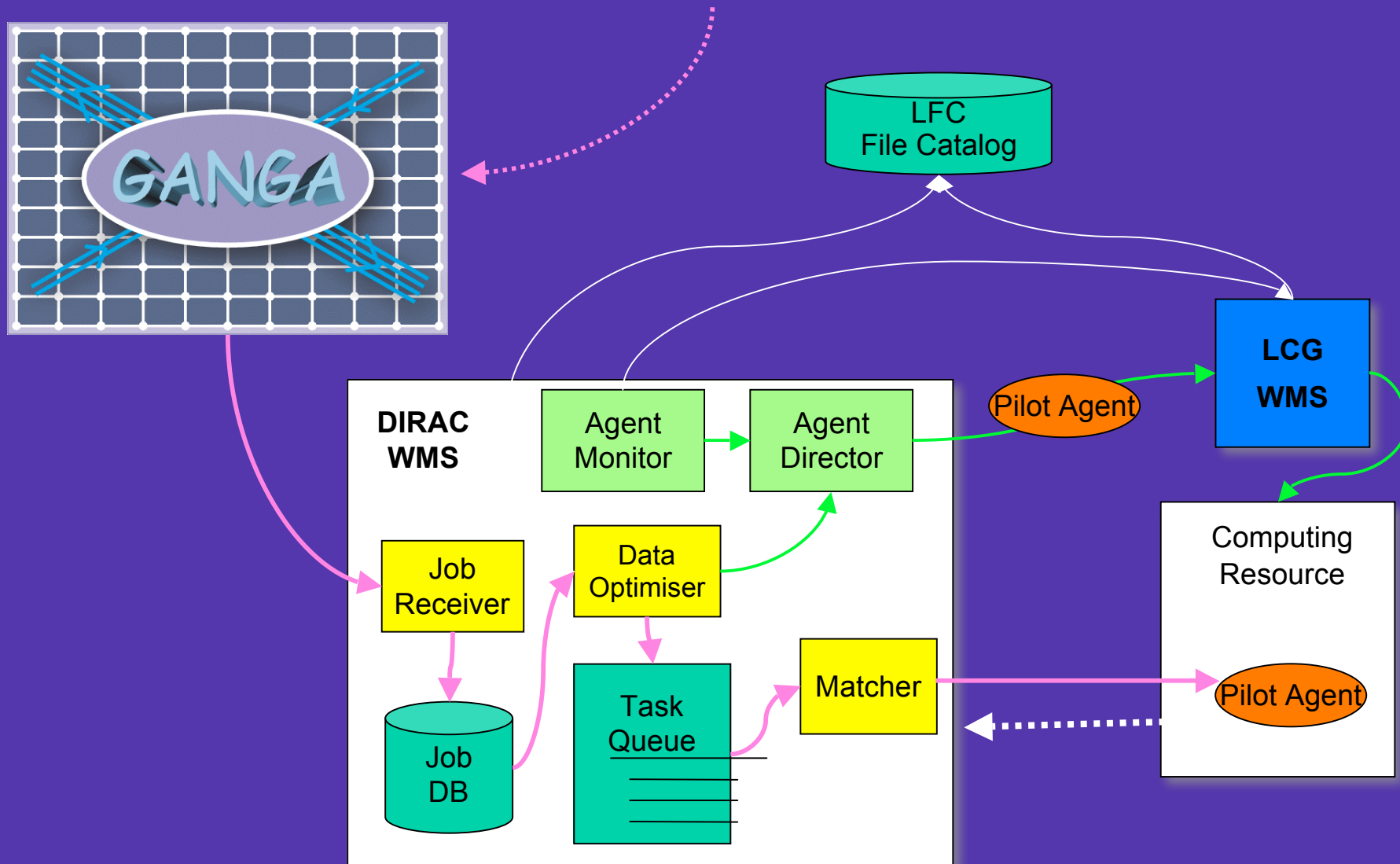
SRM:

- Limitations/ambiguity (already flagged) in functionality

# Distributed analysis:user viewpoint



The stripped DST will be used for analysis for general LHCb users



## Analysis status

- Many special enhancements to DIRAC for analysis:
  - Secure job submission (using Grid proxy)
  - Deployment of user DLL and setting environment
  - Automatic pilot-agents submission to LCG
    - Running with user's credentials
    - Capability to run several jobs (faster response time)
- Complete job wrapper
  - XML File Catalog creation
  - Output data transfer and registration
- Caveat
  - Files can be accessed by rfiio: or dcap:
    - Not optimised for staging - computing model envisaged disk SE not currently true at all sites
- More than 20 users so far in LHCb
  - Very valuable feedback
  - Often users tackle from the start very complex use cases (thousands of jobs, Gauss running, etc)

## Performance

- Test with Ganga 4.0.2 and DIRAC back-end
  - Based on 200 jobs submitted
  - 197 completed successfully, running at RAL, CERN, PIC, NIKHEF, Karlsruhe (FZK)
  - 3 remaining jobs succeeded after first re-submission
    - Clone them and run simple script
  - Job submission still slow (~25 s per job)
    - Will improve in next version (job splitting)
- However there is a learning period for users...  
most complex use cases tend to come first!

# LHCb DC'06

2 data challenges envisaged in 2006

Aim to start production for 1st of these challenges end of Feb'06 - with a finalised event model

~200M events for 1st DC

## LHCb LHCC computing milestones

2005		
Example of sub-detector alignment	September	
Analysis at all Tier-1's	November	
End of review of reconstruction software, including event model	December	
2006		
Final alignment strategy	March	
Start data processing phase of DC'06 (i) Distribution of RAW data from CERN (ii) Reconstruction/stripping at Tier-1's including CERN (iii) DST distribution to CERN & other Tier-1's	May	New version of SRM at all Tier-1's LCG services/tools from baseline service working group deployed at all Tier-1's
Alignment/calibration challenge - participation of all Tier-1's (i) Align/calibrate detector (ii) Distribute DB slice - synchronize remote DB's (iii) Reconstruct data	October	DB service to support COOL at all Tier-1's
2007		
Permanent Monte Carlo production mode ready for data taking	January	
Production system and software ready for data taking	April	

Table 6-3: High-level milestones for the computing project



## LHCb DC'06-1

Challenge (using the production services):

- a) Distribution of RAW data from CERN to Tier-1's
- b) Reconstruction/stripping at Tier-1's including CERN
- c) DST distribution to CERN & other Tier-1's

Pre-production before DC'06-1

- Event generation, detector simulation & digitisation
- For 200M B-physics events:
  - ~ 7.5 MSI2k.months will be required (over ~2 months)
  - ~ 300 TB of storage: data originally accumulated on MSS at CERN for DC06 stage

## LHCb DC'06-1

### Distribution of RAW data from CERN

~ 300 TB of storage replicated over Tier-1 sites (each file on only one Tier-1)

CERN MSS SRM → local MSS SRM endpoint (would want to pin in cache at destination until processed)

### Reconstruction/stripping at Tier-1's including CERN

~400 kSI2k.months needed to reconstruct & strip events (over 1 month period)

#### Output:

- rDST (from recons): ~100 TB at Tier-1 sites where produced, on MSS
- DST (from stripping): ~2TB in total on disk SE

### DST distribution to CERN & all other Tier-1's

- 2 TB of DST will be replicated to a disk based SE at each Tier-1 & CERN

## Data Challenge in October'06 (DC06-2)

Reconstruction using Conditions data (COOL-based)

Essential ALL LHCb Tier-1 centres are running database service supporting COOL & 3D

Timescales presented at last GDB allows months of pre-testing of services before use in anger

- GDB proposal for March'06 deployment fully supported by LHCb
  - to understand access performance
  - replication issues ...
- Preliminary tests are already ongoing

**IMPORTANT:** data will be accessed directly from SE through protocols supported by ROOT/POOL & NOT only by GridFTP/srmcp

This is NOT a trivial statement ! It has become apparent some sites had not appreciated this

# Summary

- Last few months dominated by SC3
  - Useful in debugging system/site and understanding of needs
  - Functionality & stability for SC4 essential
- Analysis activity has commenced & critical part of studies in 2006
  - Changes in DIRAC in place, system seems to be functional
  - Data access needs to be addressed - both in LHCb & provision at centres
- Planning for DC'06 underway
  - Major stress test of computing model - need stability & reliability of services
  - Reconstruction of "realistic" events in a distributed environment - testing of COOL/3D projects