

# A skimming procedure to handle large datasets at CDF



**CHEP 2006 - Mumbai, 13-17 Feb 2006**



F. Delli Paoli, D. Lucchesi, S. Da Ronco, A. Fella,  
M. Casarsa, S. Belforte, G. Compostella



- Introduction
- What is and why CDF needs the skim
- Data Handling Model at CDF
- Skimming Model
- Large Hadronic Dataset Skimming
- Final check and recovery
- Distributed Skimming
- Performances
- Conclusions

# Introduction

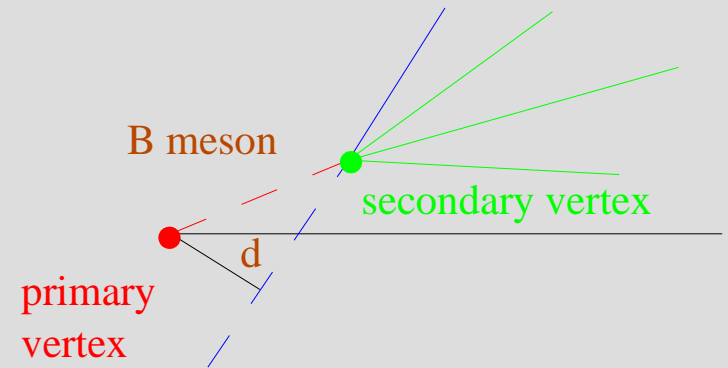
## CDF status

- CDF is an experiment running at the Tevatron, the p-pbar collider at the Fermi National Accelerator Laboratory
- The CDF experiment comprises 624 physicists from 59 institutions located around the world
- CDF has collected up today an integrated luminosity of  $\sim 1.23 \text{ fb}^{-1}$  corresponding to  $\sim 1 \text{ Pb}$  of data globally
- In the next years Tevatron will deliver to CDF  $\sim 2 \text{ fb}^{-1}$  per year
- Data are splitted in datasets based on physics requirements when they are reconstructed

# What is and why CDF needs the skim

## The hadronic dataset

- CDF has a new trigger which selects events enriched in beauty and charm mesons by cutting on the significance of the **track impact parameter ( $d$ )** respect to the primary vertex
- Sample useful for several physic measurements, the most important the determination of  **$B_s$  mixing frequency**.
- The size of this hadronic dataset is of about **20 TB** corresponding to an integrated luminosity of  $1.23 \text{ fb}^{-1}$  collected by CDF



Months of User Time to process the whole sample

Many accesses to the dataset by CDF physicists

Need to guarantee the **availability of the data for the analysis**

# What is and why CDF needs the skim

## CDF approaches

To reduce CPU time and handle dataset access CDF has two distinct approaches:

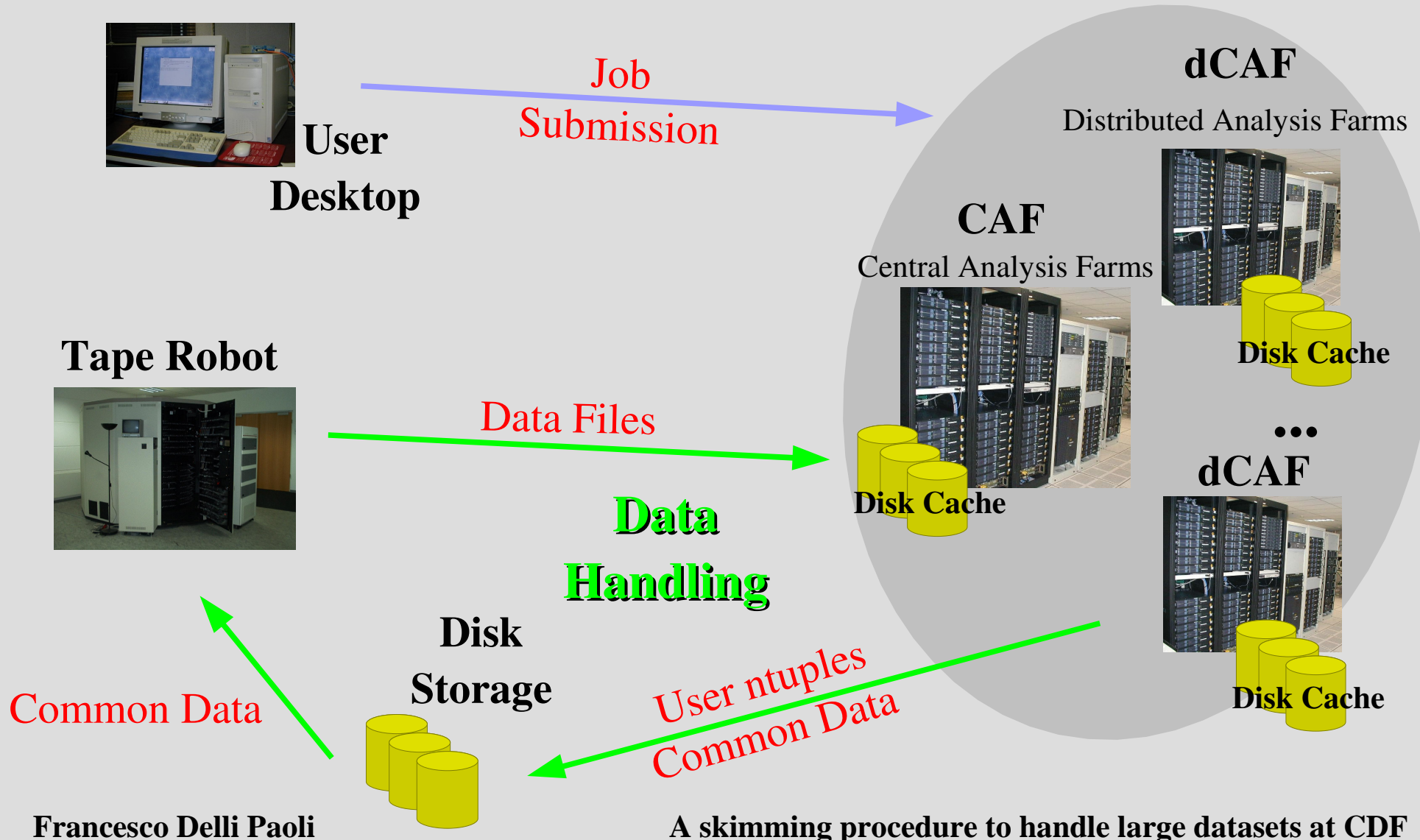
- produce standard ntuples
- reduce the dataset (**skimming**) selecting only specified physical events

## Our proposal

Implement a skimming procedure to produce smaller datasets on the basis of the **physics decay modes** useful for the **beauty and charm mesons measurements**

The procedure should **work with the CDF Data Handling Model**

# Data Handling Model at CDF



# Data Handling Model at CDF

## Computing Resources

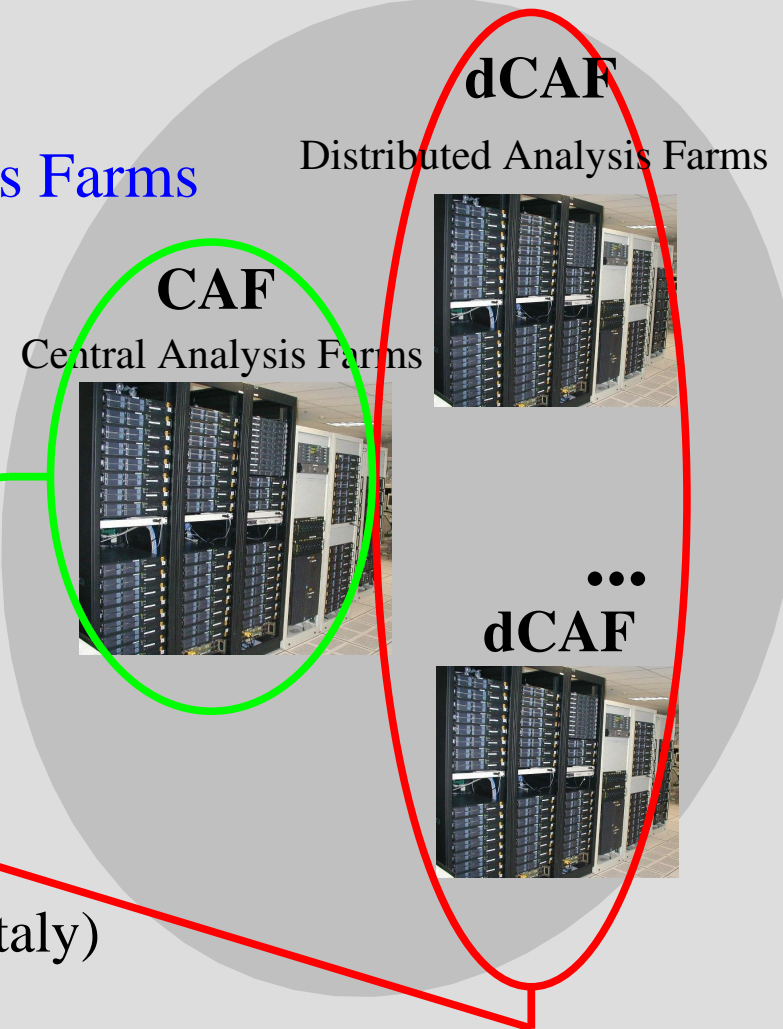
CDF has many dedicated **Computing Analysis Farms** for batch submission

2 CAFs at Fermilab ( $\sim 2.6$  MSpecInt2K)

- user Data Analysis
- semi-coordinated activities (datasets production, MC production, common ntupes)

9 dCAFs are offsite ( $\sim 2.5$  MSpecInt2K)

- Monte Carlo Production
- One of them “dCAF at CNAF” (Bologna, Italy) also Data Analysis

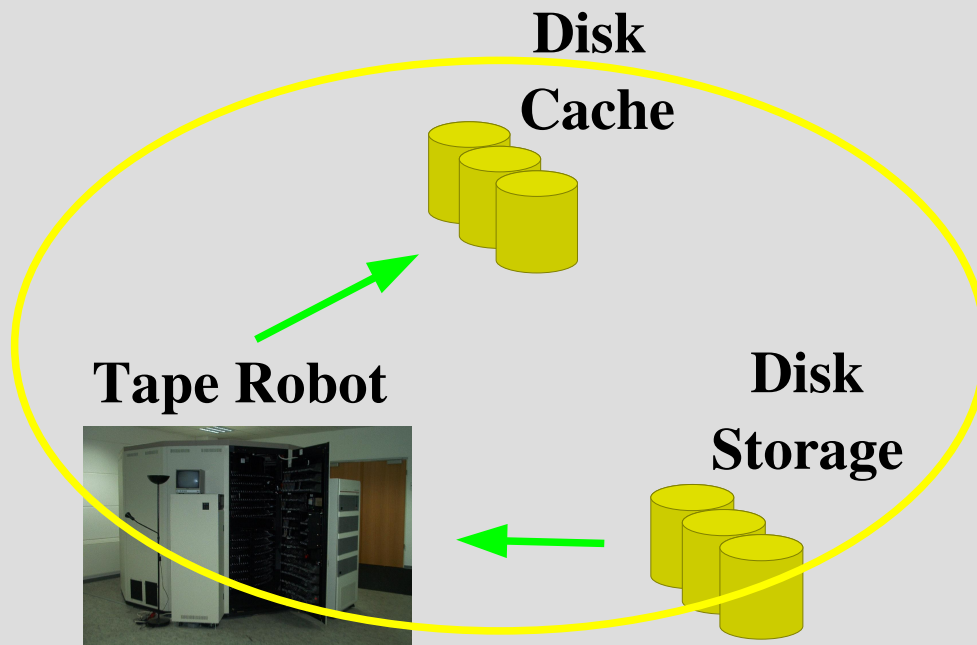


# Data Handling Model at CDF

## Data Handling

Data handling model relies on **SAM (Sequential Access via Metadata)**

- **File Catalog** for data on tape and on disk
- **File Transfers** tape-to-cache, cache-to-disk, disk-to-tape



- **File Delivery**: copy necessary files from the closest location to local cache

- **Bookkeeping**: keep track of processed and unprocessed files:  
file status: buffered, unbuffered, consumed, not consumed

# Skimming Model

## Skimming

- Selects events based on useful B decay modes from Hadronic Dataset, collects them in **smaller datasets**
- **Concatenates output** in order to reduce the number of files to be stored to facilitate the Data Handling Catalog
- Stores the skimmed datasets **on tapes and/or on storage disks** in Analysis dedicated offsite computing farm (i.e. dCAF at CNAF)

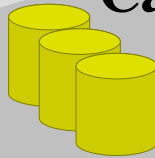


# Skimming Model

Tape Robot

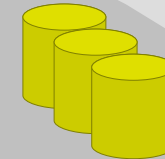


Disk  
Cache



**Skimming**

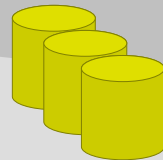
Durable  
Location



*Skimming Model*

**Storing**

Disk  
Storage



**Concatenation**

# Large Hadronic Dataset Skimming: *skimming*

## Tape Robot



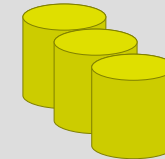
## Disk Cache



copy via *sam\_upload*

**Skimming**

## Durable Location



## Large Hadron Dataset

- ~60000 files
- ~20 Tb
- Event filtering on the basis of **8 physical decay modes**
- Selection performed **simultaneously** on each event to create 8 output datasets
- **Data delivering, bookkeeping and recovery** performed by SAM

## Intermediate Datasets

- From <1 to 5% of the original input (depending on the decay mode)
- Total saved ~20% of input

# Large Hadronic Dataset Skimming: concatenation

- A huge number of files end up **wasting time and resources**
- **Concatenation** to produce datasets with **lower number of files** which have a bigger size (1 Gb)

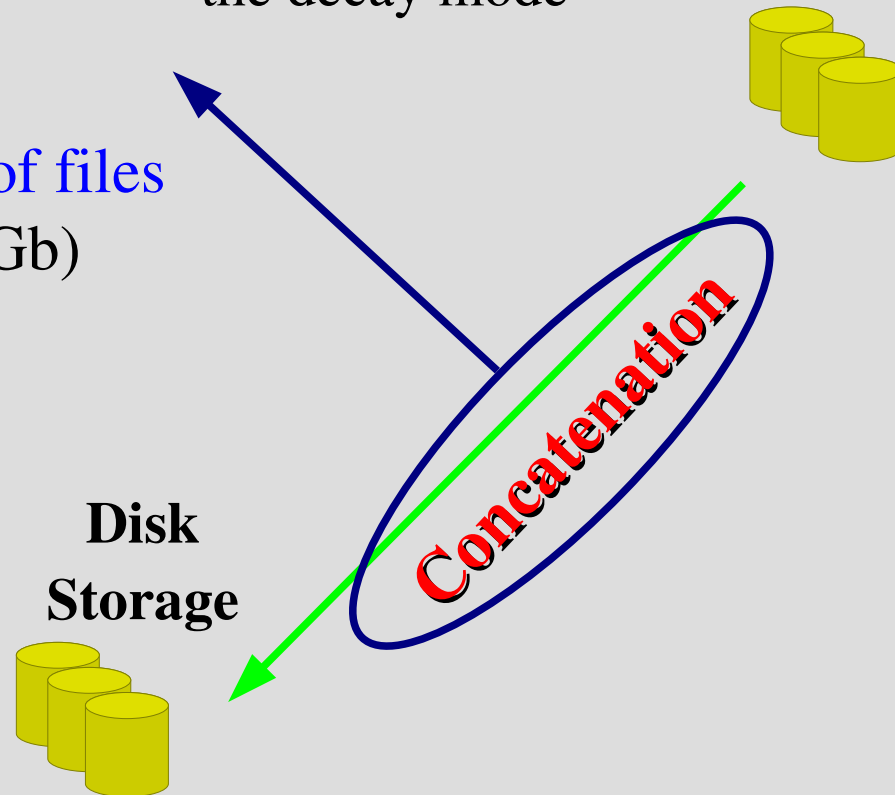
## Intermediate Datasets

- ~30000 files per dataset
- ~1-10 Mb depending on the decay mode

**Durable Location**

## Concatenated Datasets

- ~100-1000 files per dataset
- ~1 Gb per file



# Large Hadronic Dataset Skimming: storing

## Tape Robot



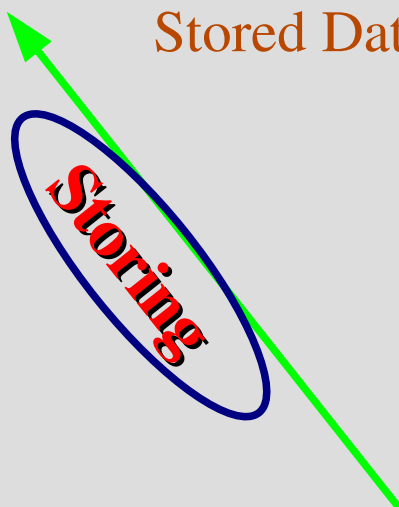
## Disk Cache



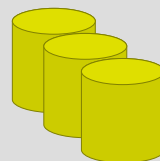
Stored Datasets

Provide **standard SAM access**

- Store datasets **on tape**
- **Cache datasets on disk** near offsite collaborations to provide fast access
- At CNAF we have at the moment **the skimmed datasets available**
- Storing performed by **sam\_upload**



## Disk Storage



Concatenated Datasets

A skimming procedure to handle large datasets at CDF

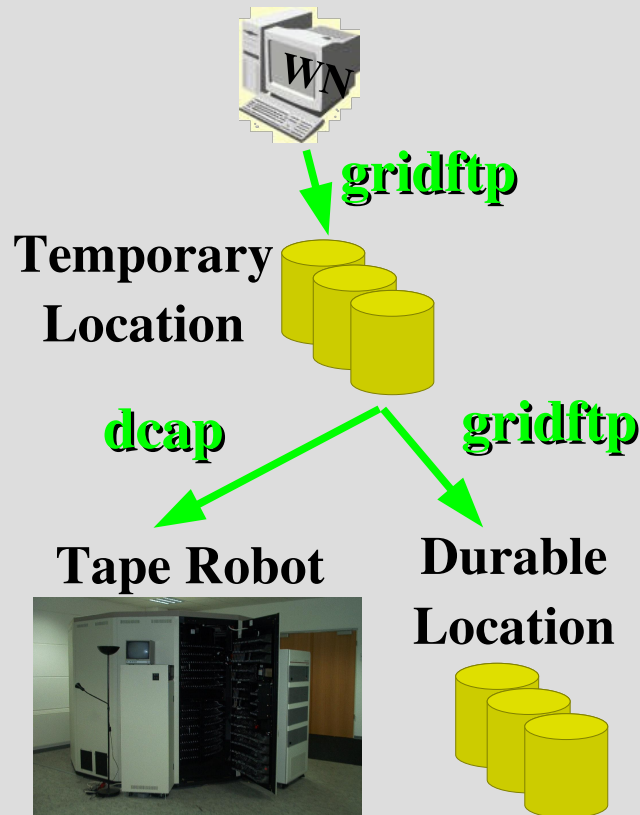
# Large Hadronic Dataset Skimming: `sam_upload`

## `sam_upload`

Performs transfer and storage of files:

`sam_upload` client

- **GSI Authentication** (Globus Security Infrastructure)

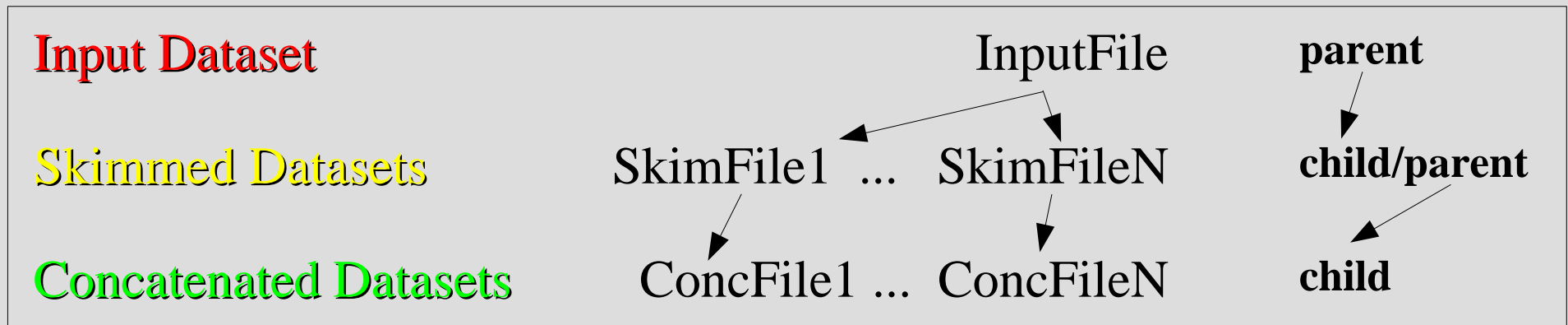


- Transfer layer to a temporary location is an **enqueued gridftp**
- File to be transferred are **catalogued with SAM**
- Store **on tape** (via a File Storage System around dcap protocol) or **on durable locations** (via gridftp)
- Manage **retry procedure** in case of lack of services

# Final check and recovery

## Recovery and Checking

- Not delivered files are recovered by SAM bookkeeping system: file status not consumed
- The skimming procedure allows to keep track of every input file in the skimming chain thanks to the **file lineage** of the SAM catalog

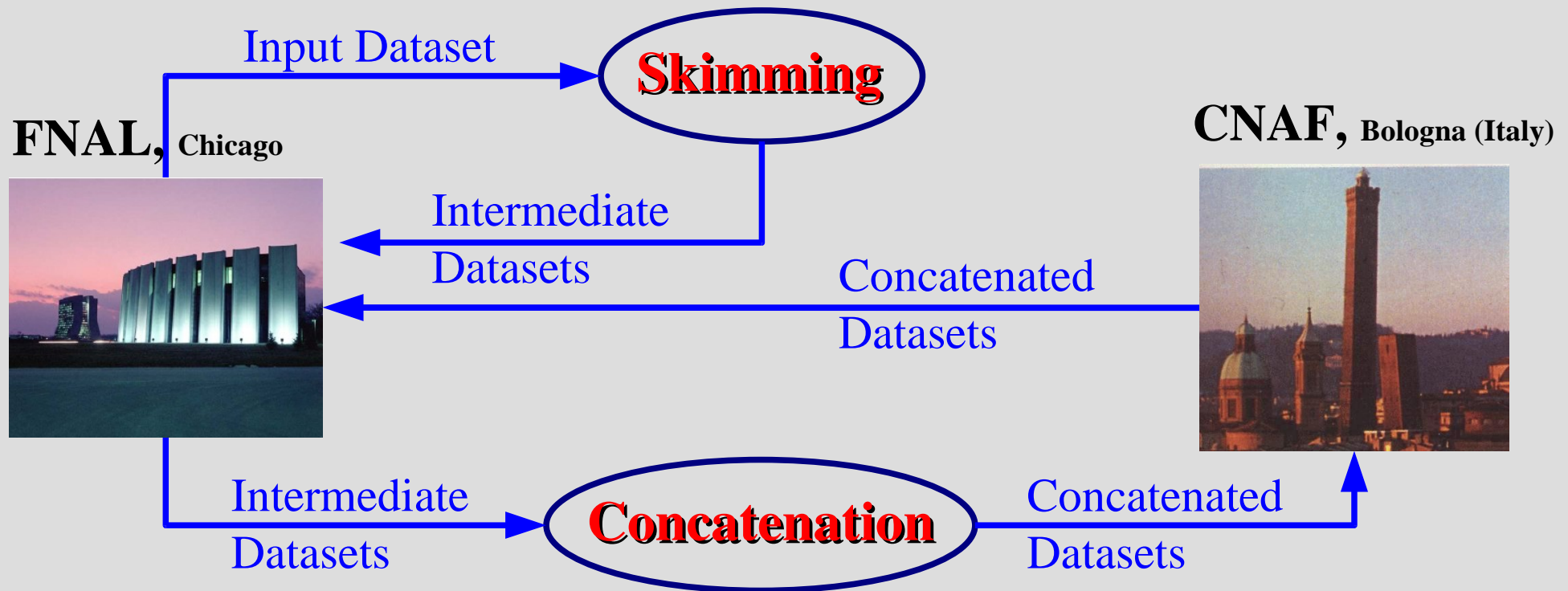


- It is possible to check if a file is already skimmed or concatenated:
  - avoid and find input file duplication
  - avoid input file loss

# Distributed skimming

CDF data collected until Dec 2004 (12.96 Tb):

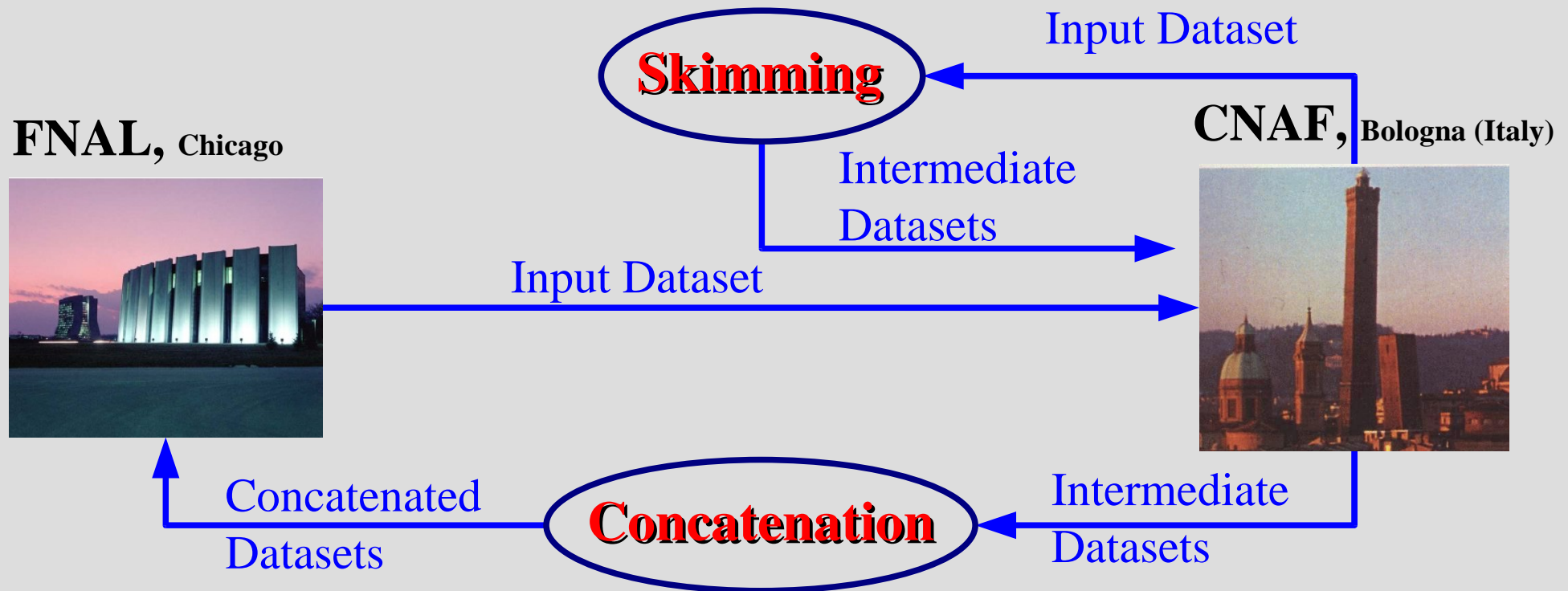
Input Dataset at FNAL, Intermediate Datasets written on FNAL Durable Location, Concatenated Datasets stored on CNAF disks. A copy saved to FNAL's tapes.



# Distributed skimming

CDF data collected from Dec 2004 till now (9.19 Tb)

Input Dataset copied at CNAF, Intermediate Datasets written on CNAF Durable Location, Concatenated Datasets stored at FNAL disks and on tape





# Performances

## Skimming Performances

The procedure is in production. The total failure rate has been **< 1%**

The errors are due to **SAM delivery problems**:

- Not all the files of the input dataset are delivered. Need to resubmit job to process missing files.
- File delivered are actually not processed because of Worker Node problems (2-3 times over all the skimming procedure).

# Status and Performances

## Benchmark

- To skim the Hadronic Dataset:
  - All CAF onsite resources (2.6 MSpecInt2K): 2 weeks
  - Standard Users Team (4 users on ~600 KSpecInt2K): 2 months of User Time
  - Standard Users Team onsite + CNAF resources (~900 KSpecInt2K): 1 month of User Time
- User Analysis
  - Running on the Hadronic Dataset: ~2 months of User Time
  - Running on a skimmed dataset: 24 h of User Time

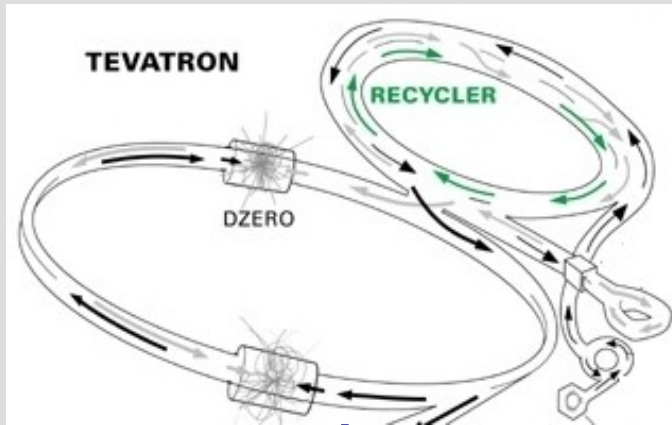
# Conclusion

## Conclusions

- New CDF trigger produces a Large Hadronic Dataset: 20 Tb of data
- In the future CDF will collect  $\sim 2 \text{ fb}^{-1}$  per year
- Need to guarantee data availability
- Running over this dataset takes now  $\sim 2$  months of User Time on CAF
- We propose a skimming procedure, working with the above requirements, to reduce User Time for analysis to 24 h
- Skimmed datasets are now available on SAM and mirrored in offsite institutions like CNAF

# Backup

# Data Production Flow



## Production Farm

Reco Data Size:  
120 Kb/evt



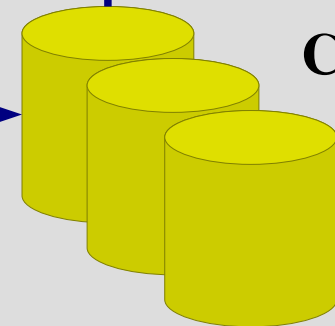
Raw Data Size:  
150 Kb/evt

## Tape Robot



Level 3 trigger: 100 Hz

## Disk Cache



# Hadronic Trigger Selection

	B_CHARM_LOWPT	B_CHARM	B_CHARM_HIGHTPT
L 1	$p_T^1 > 2 \text{ GeV}/c$ $p_T^2 > 2 \text{ GeV}/c$ $\Delta\phi_{1-2} < 90^\circ$	$p_T^1 > 2 \text{ GeV}/c$ $p_T^2 > 2 \text{ GeV}/c$ $p_T^1 + p_T^2 > 5.5 \text{ GeV}/c$ $\Delta\phi_{1-2} < 135^\circ$ opposite charge	$p_T^1 > 2.6 \text{ GeV}/c$ $p_T^2 > 2.6 \text{ GeV}/c$ $p_T^1 + p_T^2 > 5.5 \text{ GeV}/c$ $\Delta\phi_{1-2} < 135^\circ$ opposite charge
L 2	$p_T^1 > 2 \text{ GeV}/c$ $p_T^2 > 2 \text{ GeV}/c$ $0.012 < d_0^1 < 0.1 \text{ cm}$ $0.012 < d_0^2 < 0.1 \text{ cm}$ $2^\circ < \Delta\phi_{1-2} < 90^\circ$ $L_{xy} > 0.02 \text{ cm}$	$p_T^1 > 2 \text{ GeV}/c$ $p_T^2 > 2 \text{ GeV}/c$ $p_T^1 + p_T^2 > 5.5 \text{ GeV}/c$ $0.012 < d_0^1 < 0.1 \text{ cm}$ $0.012 < d_0^2 < 0.1 \text{ cm}$ $2^\circ < \Delta\phi_{1-2} < 90^\circ$ $L_{xy} > 0.02 \text{ cm}$ opposite charge	$p_T^1 > 2.6 \text{ GeV}/c$ $p_T^2 > 2.6 \text{ GeV}/c$ $p_T^1 + p_T^2 > 5.5 \text{ GeV}/c$ $0.012 < d_0^1 < 0.1 \text{ cm}$ $0.012 < d_0^2 < 0.1 \text{ cm}$ $2^\circ < \Delta\phi_{1-2} < 90^\circ$ $L_{xy} > 0.02 \text{ cm}$ opposite charge
L 3	$p_T^1 > 2 \text{ GeV}/c$ $p_T^2 > 2 \text{ GeV}/c$ $p_T^1 + p_T^2 > 4.0 \text{ GeV}/c$ $0.008 < d_0^1 < 0.1 \text{ cm}$ $0.008 < d_0^2 < 0.1 \text{ cm}$ $2^\circ < \Delta\phi_{1-2} < 135^\circ$ $L_{xy} > 0.02 \text{ cm}$	$p_T^1 > 2 \text{ GeV}/c$ $p_T^2 > 2 \text{ GeV}/c$ $p_T^1 + p_T^2 > 5.5 \text{ GeV}/c$ $0.008 < d_0^1 < 0.1 \text{ cm}$ $0.008 < d_0^2 < 0.1 \text{ cm}$ $2^\circ < \Delta\phi_{1-2} < 135^\circ$ $L_{xy} > 0.02 \text{ cm}$ opposite charge	$p_T^1 > 2.6 \text{ GeV}/c$ $p_T^2 > 2.6 \text{ GeV}/c$ $p_T^1 + p_T^2 > 5.5 \text{ GeV}/c$ $0.008 < d_0^1 < 0.1 \text{ cm}$ $0.008 < d_0^2 < 0.1 \text{ cm}$ $2^\circ < \Delta\phi_{1-2} < 90^\circ$ $L_{xy} > 0.02 \text{ cm}$ opposite charge
	$ \eta  < 1.2 \ \& \  \Delta z_0  < 5 \text{ cm}$	$ \eta  < 1.2 \ \& \  \Delta z_0  < 5 \text{ cm}$	$ \eta  < 1.2 \ \& \  \Delta z_0  < 5 \text{ cm}$