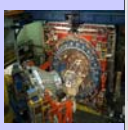
A New CDF Model for Data Movement Based on SRM

Manoj Kumar Jha (INFN Bologna)
Gabriele Compostella, Donatella Lucchessi, Simone P. Griso
(University and INFN Padova)
Doug Benjamin (Duke University)

Being a large international collaboration established well before the full development of the Grid as the main computing tool for High Energy Physics, CDF has recently changed and improved its computing model, decentralizing some parts of it in order to be able to exploit the rising number of distributed resources available nowadays. Despite those efforts, while the large majority of CDF Monte Carlo production has moved to the Grid, data processing is still mainly performed in dedicated farms hosted at FNAL, requiring a centralized management of data and Monte Carlo samples needed for physics analysis. This rises the question on how to manage the transfer of produced Monte Carlo samples from remote Grid sites to FNAL in an efficient way; up to now CDF has relied on a non scalable centralized solution based on dedicated data servers accessed through rcp protocol, which has proven to be unsatisfactory. A new data transfer model has been designed that uses SRMs as local caches for remote Monte Carlo production sites, interfaces them with SAM, the experiment data catalog, and finally realizes the file movement exploiting the features provided by the data catalog transfer layer. We describe here the model and its integration within the current CDF computing architecture.

Introduction

❖ CDF :
- ❖Large multipurpose Particle Physics Experiment at Fermi National Accelerator Lab (Fermilab) Started collecting data in 1988
- ❖Data taking will continue until at least Oct. 2009 (with desire to extend another year) also in 2010.

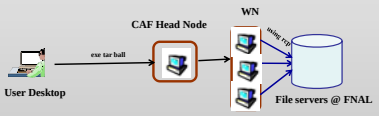


❖Great Tevatron Performance
- ❖ Close to 5fb-1 delivered

❖Need Monte Carlo(MC) data for detector understanding and physics analysis.
- ❖Not feasible to produce all the MC data onsite due to limitation on computing resources.
- ❖ Only option left is to utilize resources at Tier-I and Tier – II remote grid sites using the CDF Analysis Framework (CAF) and bring back the MC output at Fermilab.

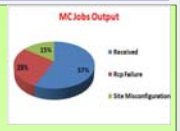
CDF Analysis Framework(CAF)

❖CAF developed as a portal.
- ❖A set of daemons (submitter, monitor, mailer) accept requests from the users via kerberized connections.
- ❖Requests are converted into commands to the underlying batch system.
- ❖ A user job consists of several sections. Each section has a wrapper associated with it.
  - ❖ Task of wrapper is to setup security envelop & prepare environment before the actual user code starts.
  - ❖ When the user code finishes, it will also copy whatever is left to a user specified location using rcp.



Demerits of CAF

❖ Presently, CDF is relying on rcp tool for transfer of output from WN to destination file servers. It leads to



- ❖ Worker nodes are sitting idle just because another WN is transferring files to the fileservers.
  - ❖Loss of output from the WN in most of the cases.
  - ❖Wastage of CPU and network resources
- ❖ No mechanism to deal with the sudden arrival of output from WN at file servers. Especially, happens during the conference period.
  - ❖Overloading of available file servers in most of the cases.
    - ❖ Users have to resubmit the job.

❖ No catalogue of MC produced files on user basis.

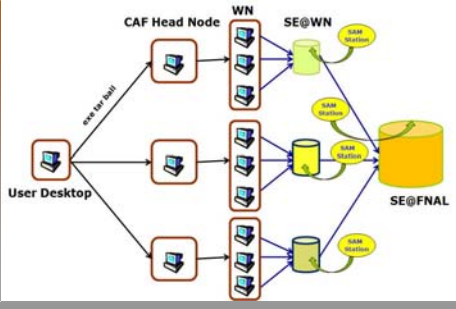
Data Transfer Model

❖Output from WN will be temporarily collected in the SE closer to WN.
- ❖Transfer or wait time for MC output data on WN will be considerably reduced due to large bandwidth between WN and its SE.

❖ Transfer sequentially the collected output to the destination SE.
- ❖ How to manage the files on SE ?
  - ❖Use SRM managed SE
- ❖ How to transfer file b/w SEs ?
  - ❖Use Sequential Data Access via Meta-Data (SAM) features for transfer of files between its stations.

❖ Model uses the SAM SRM interface for data transfer.

Proposed Transfer Model



SAM SRM Interface

❖SAM-SRM interface apply SAM data handling policies over the generic storage system managed by SRM protocol.
- ❖Supports operation transferring file, removing file, building directory listing and retrieving metadata as common for POSIX file systems.

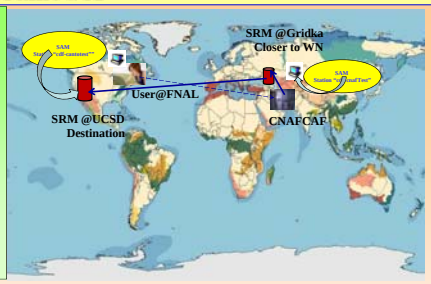
❖SAM disks are used as root locations to place data files .
- ❖Logical disks and SAM file replacement policy constitute SAM cache that brokers resources between SE and SE clients.

❖More information on SAM to SRM mapping can be found in specification document [2].

Test Framework

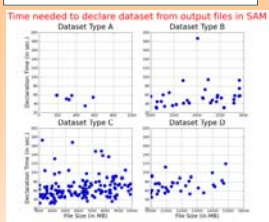
❖Hardware:
- ❖CAF: CNAF Tier –I
- ❖ SE: dCache managed SRMs
  - ❖ 500 GB of space at Gridka, Karlsruhe, 10 channels per request (closer to WN)
  - ❖ 1 TB of space at University of California, San Diego(UCSD), 50 channels per request (destination SE)
- ❖ SAM station:
  - ❖station "cdf-cnafTest" at "cdfsam1.cr.cnaf.infn.it" (closer to WN)
  - ❖ station "canto-test" at "cdfsam15.fnal.gov" (destination SE)

❖Setup:
- ❖A cronjob submits a job of variable no. of segments every 10 minutes at CNAF CAF. Maximum segments per job is 5.
  - ❖Each segment creates a dummy file of random size which vary between 10 MB to 5 GB.
- ❖A cron process running at station "cdf-cnafTest" for creation of datasets from the files in SRM closer to WN (Gridka).
- ❖ Another cron job running at station "canto-test" for transfer of dataset between two stations.

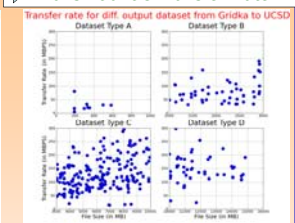


Performance Parameters

Dataset creation Time



Trans Atlantic Transfer Rate



❖Model proposes two types of dataset corresponding to a JID.
- ❖ Output dataset: Collection of all the output files corresponding to a JID
- ❖Log dataset: Collection of all the log files corresponding to a JID

❖Dataset creation time is of the order of around 60 seconds.

■ Using general purpose wide area network for file transfers.

■ Trans Atlantic transfer rate is of the order of around 200 MBPS.
- ■ It should be noted that conventional rcp rate is of the order of few hundreds of kilo bytes per second.

■ Tremendous increase in transfer rate as compare to the present CAF model.

Future Plan

Our next goal is to use this model for movement of data from production sites to the WN at remote site for further processing of real data.

Useful Links:
1. CDF Homepage, http://www-cdf.fnal.gov
2. SAM-SRM Design Document, https://plon3.fnal.gov/SAMGrid/Wiki/SAM-SRM-Design.doc/download
3. SRM Working Group Homepage, http://sdm.lbl.gov/srm-wg/
4. SAM Homepage, http://d0ora1.fnal.gov/sam/
5. CAF Homepage, http://cdfcaf.fnal.gov

ACKOWLEDGMENTS:
Thanks to *Frank Wuerthwein and Abhishek Singh Rana (from UCSD), Thomas Kuhr (from Karlsruhe), Brian Bockelman (from Nebraska), Jon Bakken (from FNAL), for providing resources on SRMs !*