

The CDF Run II Data Handling Road Map

Part 1

Robert D. Kennedy

FNAL Computing Division

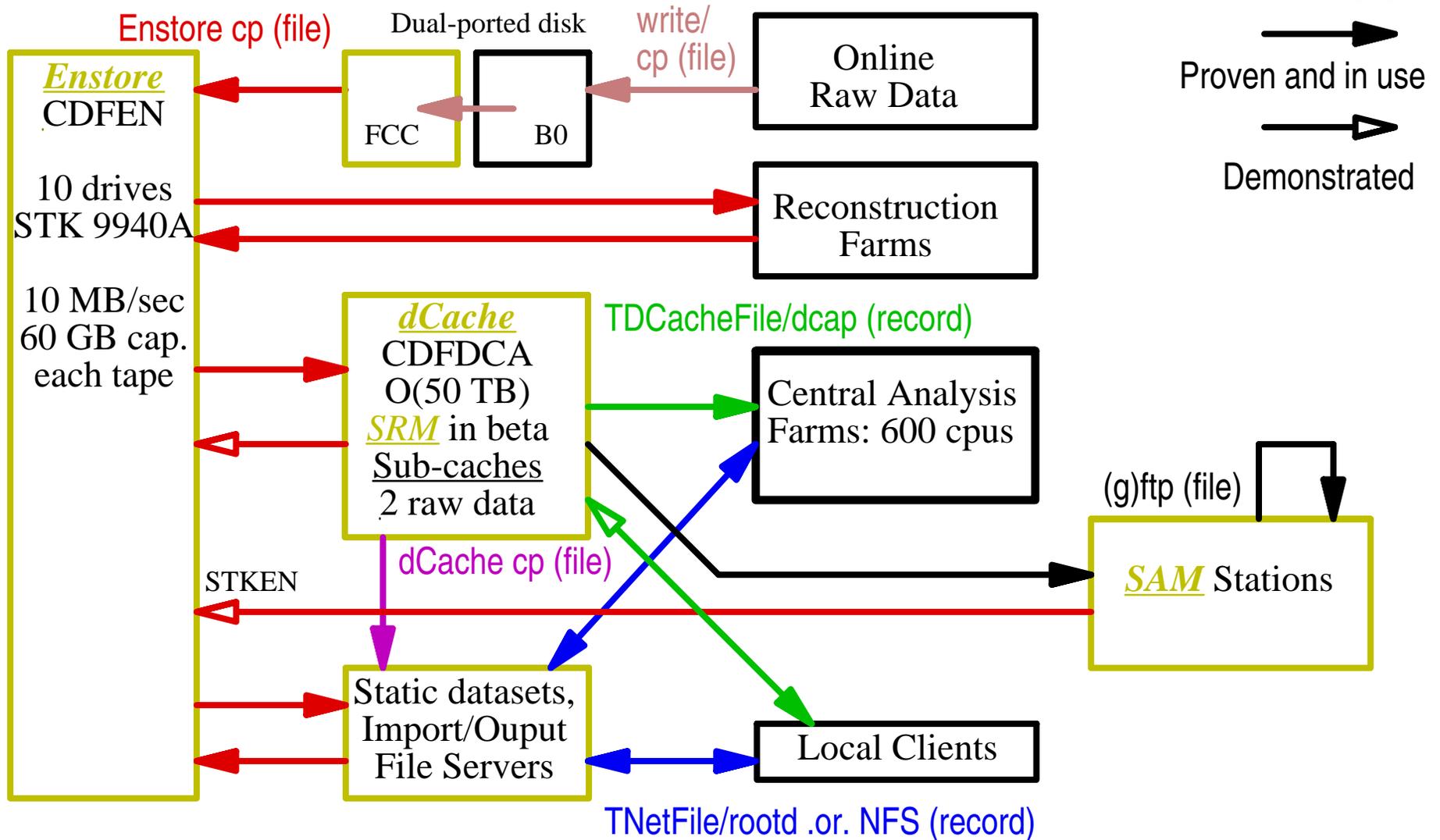
for the CDF Data Handling Group

R2D2 Meeting - 06 May 2003

v1.1

- ⇒ **Review of state of CDF - DH in transition**
- ⇒ **Overview of CDF DH Baseline Goal for FY2004**
- ⇒ **Activities break-down for Enstore, dCache, SAM, DB Layer**
- ⇒ **Part 1: Focus on Enstore & dCache, less on SAM & long-term**
- ⇒ **Part 2: Rick St. Denis' talk - More focus on SAM & long-term**

CDF DH Today (legacy systems not shown)



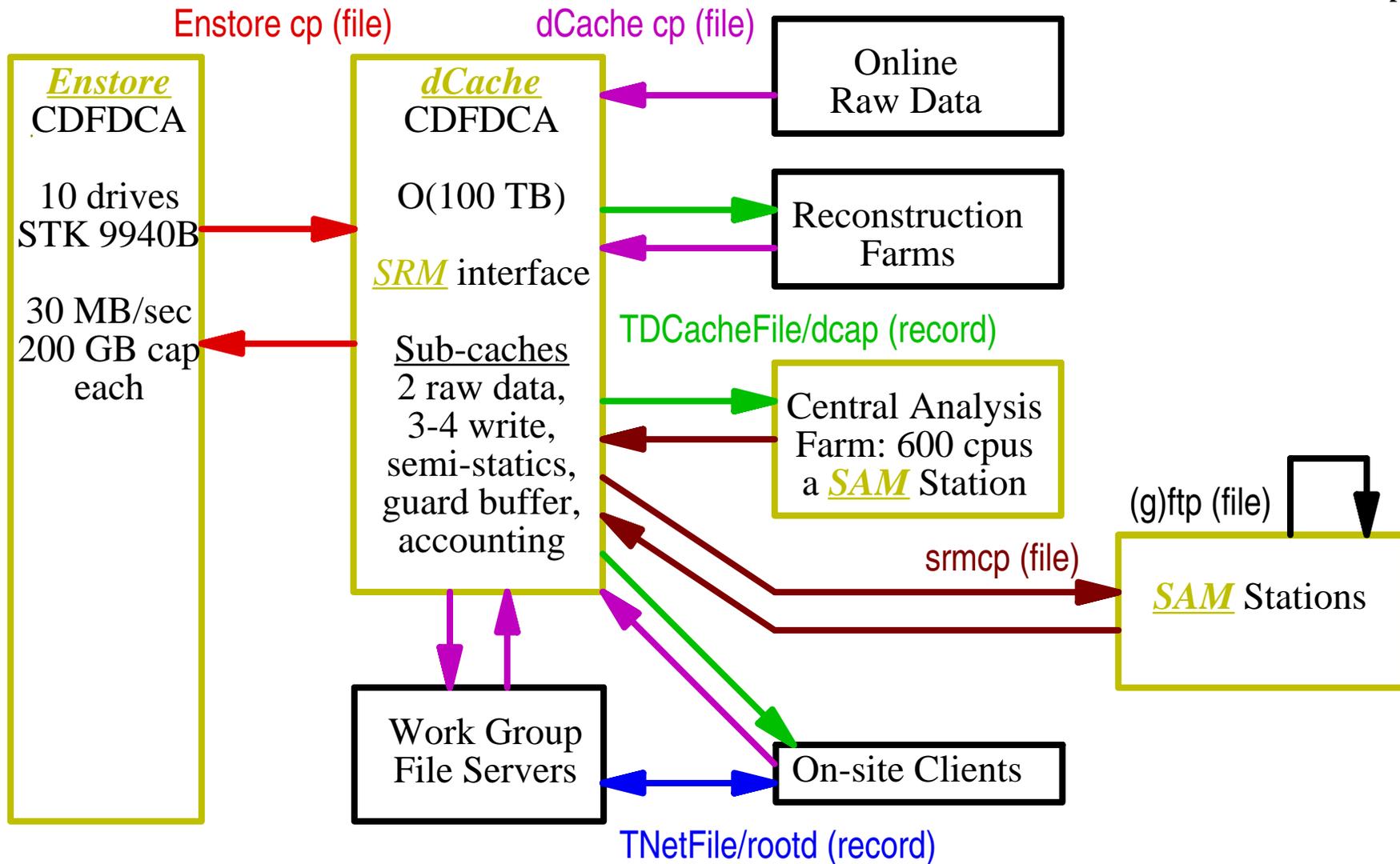
Underlying *Meta-data*: PNFS (Enstore/dCache), CDF Data File Catalog, SAM meta-data

Overview: Baseline Goal FY2004

CDF Run II DH Roadmap Pt1
Robert D. Kennedy
06 May 2003
page 3

- ⇒ **CDF Enstore fully migrated to use of 9940B tape drives**
- ⇒ **CDF dCache fully implemented, 100 TB+, with sub-caches**
- ⇒ **All Enstore access goes through dCache (read .and. write)**
- ⇒ **Tapeless data paths for produced (and raw?) datasets**
- ⇒ **Retire "legacy" systems, once baseline proves reliable**
- ⇒ **Fully adapt to SAM schema; drop DFC schema, keep API**
- ⇒ **CAF adapted to become a SAM station w/direct dCache access**
- ⇒ **Simplified "entry-level" SAM input for users (carrot, not stick)**
- ⇒ **Robust operations, doc'd procedures, user experience issues**
- ⇒ **Groundwork: CAF output catenation, unified data processing**

CDF DH Baseline Goal FY2004



Underlying *Meta-data*: PNFS (Enstore/dCache) and SAM meta-data

Activities: Enstore

CDF Enstore: depends on PNFS service
maintained by CD, activities in coll. w/CD-ISA

En1) 9940A to 9940B Format Migration: achieves 2 PB capacity
Status: in progress; target completion ? (<= 01 Sep'03?)

- a) Migrate already-stored data from A format tapes to B format tapes
- b) Recycle tapes whose data have been migrated
- c) Switch Reconstruction Farms to write to B format tapes - CDF Farms
- d) Switch secondary data sets creators to write to B tapes - CDF "Strippers"
- e) Switch raw data logging to write to B format tapes - CDF CSL

En2) Migrate SAM File Storage to write to CDFEN, not STKEN
-) See SAM project activities list

Activities: dCache 1

CDF dCache: depends on PNFS service, Enstore service
file servers - CDF; future specs in coll. with CD
service hosts acquired, maintained by CDF Task Force
application/service maintained by CD, dev. w/DESY

Dc1) Raise CDF dCache service to production level (w/CCF)

Status: in progress; target completion \leq 1 month

- a) Manage (done), investigate bit-errors observed when restoring files
- b) Restore test stand system: admins are spares for production
- c) Use only shared-lib external product client interface to dCache (done)
- d) Upgrade admin nodes (done), later to task force-blessed Dell 2650s
- e) Investigate performance, bugs in TDCacheFile access: almost all resolved
- f) Expand CAF access to dCache - now 50% (302 cpus), more later
- g) Increase Nservers in dCache - now 25, soon 35 when bit-error tests done.
Later, 50 fs (= 100% of CAF) when static dataset file servers absorbed.
- h) Define, load "golden" semi-static sub-cache with "static datasets"
- i) Update documentation; advertise "pre-staging" service
- j) Perform 20 TB/day load test on production system: read only, read w/staging

Activities: dCache 2

Dc2) Improve/expand SAM access to dCache

Begin: in progress; target completion \leq 1 month

- a) Improve/expand weakly-auth ftp access to dCache (more doors)
- b) Test alternative access methods: srmcp
- c) Validate large N simultaneous transfers to larger SAM stations

Dc3) Re-enable Pool-to-Pool copies (needed for tape-less paths)

Begun, on hold; target completion \leq 1 month

- a) Implement file integrity checks in internal "java" file copies
- b) Tune to avoid p2p thrashing, but still allow load-sharing replication
- c) Tune again to allow semi-static sub-caches to minimize replication

Dc4) Use of SRM as dCache access manager

Begun, on hold; target completion FY2004

Consider re-orienting to be more inline with CDF-SAM

- a) Re-start beta-testing of SRM-dCache interface - load tests, etc.
- b) Adapt DHInput - send whole list of input files to SRM (also useful for SAM)
- c) Given SAM features: what access policies useful: files-in-cache first, ...
- d) When all tests passed, make this the default dCache access method in AC++

Activities: dCache 3

Dc5) Implement/verify dCache write-pools, tapeless data paths **Begin: was in progress/on hold; completion FY2004**

- a) Use all-write-pool file servers to prove fileset contiguity on tape.
- b) Test file family-affinity with write pools to steer files to pools "near" Enstore
- c) Kerberos-authenticating doors for write access (read access - unique user ids)
- 6 mos d) File integrity checking in dccp, TDCacheFile (writes: whole files only)

- 12 mos e) User access policy - who can write how much and how managed
- f) Hardware specification for write-pool file servers - in conjunction w/CD
- g) Establish write pools for Reconstruction output

- h) Establish tape-less path for Reconstruction output to consumers

- not yet defined i) ??? Establish write pools for Raw Data Logging - may still use dual-ported disk
- j) ??? Adapt Raw Data Logging to dCache write pools directly from B0
- k) ??? Establish tape-less path for Raw Data to Reconstr Farms (size/protection)

Dc6) Retire legacy caching mechanisms **Begin: 2 weeks after dCache prd; completion 01 Sep 2003**

- a) End static datasets availability, absorb file servers into dCache
- b) Reduce DIM data space from 15 TB to 9 TB, warn users of system life span
- c) Proposed end of DIM support: 01 Sep 2003

Activities: SAM

SAM: (at CDF) SAM Meta-data, dCache, Enstore
SAM is crucial, required DH layer for resource mgmt
See Rick's talk. Here - a sketch of items affecting baseline

Sc1) Adapt CDF SAM to new meta-data schema; in progress

***) Long list of tasks - outlined in Jeff's plan.**

Activity is proceeding slower than than plan expected

***) Drop CDF DFC tables and CDF use of predator to fill "SAM" tables.**

Sc2) SAM consumer using direct access to dCache datafiles

Status: Andrew has station code; Rob K. to consult

***) Goal: Where network is "perfectly reliable", allow consumers to access data files in dCache. Reduces nBytes transferred (ROOT multi-branch)**

Sc3) Adapt CDF SAM and CAF to each other; target 1 Oct 2003

***) Many tasks, see Rick's list**

Sc4) "Entry-level" SAM input for CDF users; target 1 Oct 2003

***) Goal: Simple data access does not require paradigm change for casual users.**

SAM is a (mostly) implementation detail, until users want added functionality

Activities: Database Layer

Database Layers: PNFS, DFC, SAM Meta-data

PNFS - CD

DFC - CDF DH and DB groups, based on Oracle DB

SAM - Joint SAM team, based on Oracle DB

Db1) Re-implement CDF DFC API to use SAM meta-data

***) Can proceed once new SAM meta-data schema available and populated**

***) List of tasks yet to be defined. Estimate only 1-2 weeks of work for completion.**

Db2) Ground-work for future: improve (remote) access

Replication, multi-tier structure, free-ware databases

Db3) Various infrastructure improvements:

Planned PNFS upgrade to assure very high reliability service

Activities: All Other

All Other: General themes in short-term, prep for longer-term

Ot1) Stabilize, improve robustness of DH operations

- *) Look for weaknesses before users find them, address user reports**
- *) Tune DH systems for better performance, reduce resource utilization**
- *) Document procedures, formalize shifts, etc.**
- *) Educate users how to best use, and what use-cases are supported.**

Ot2) User experience: How to catenate CAF output?

- *) N sections produce N histograms, ntuples, and/or data files**
- *) Recon Farms have a model; no model defined yet for CAF.**
- *) Opportunity for SAM adaptation to help?**

Ot3) Unification of data processing chain

- *) There are now separate approaches used in each link in data processing chain**
- *) Is it productive to plan now to adopt a single approach after 2004 for
Raw Data Logging, Reconstruction Farms, 2ndary data sets, CAF, desktops?**