# dCache

by

Michael Ernst

Patrick Fuhrmann

Tigran Mkrtchyan

at

Chep March 2003 , San Diego

# *Overview*

dCache basics

*Modes of operation*

*Attraction Model*

*Dataset Location Management*

*dCache Components (Overview)*

*Native dCache access method (dCap)*

The GRID Storage Resource Manager (srm)

*The Goal of the Srm Initiative*

*The Storage System Abstraction*

*An SRM initiated transfer example*

*Major SRM Topics*

# dCache : Modes of Operation

Used as an HSM frontend, the dCache provides standard caching mechanisms to optimize tape accesses :

Transfer speed adaption

Tunable deferred HSM stores (space , time )

Automatic staging

Continuous garbage collection (no tresholds)

Fetch ahead (from Hsm) [in preparation]

# dCache : Modes of Operation

- dCache Pools without HSM backend can hold :

  Precious datasets

  *Files are never automatically removed.*
  *System can run out of disk space.*

  Volatile datasets

  *Unused files are automatically removed.*
  *System won't run out of disk space.*

- The dCache can be operated in hybrid mode, running HSM and NON - HSM pools.

# dCache : The Attraction Model

File (resp. store/retrieve requests) are attracted by pools based on :

Statically configured parameters, e.g. :

*Client host IP or subnet numbers*

*HSM groups*

*Subdirectorytrees*

Dynamically taken parameters from life system, e.g. :

*Pool CPU cost (number of active movers)*

*Pool Space costs (space left, age of datasets)*

# dCache : Dataset Location Management

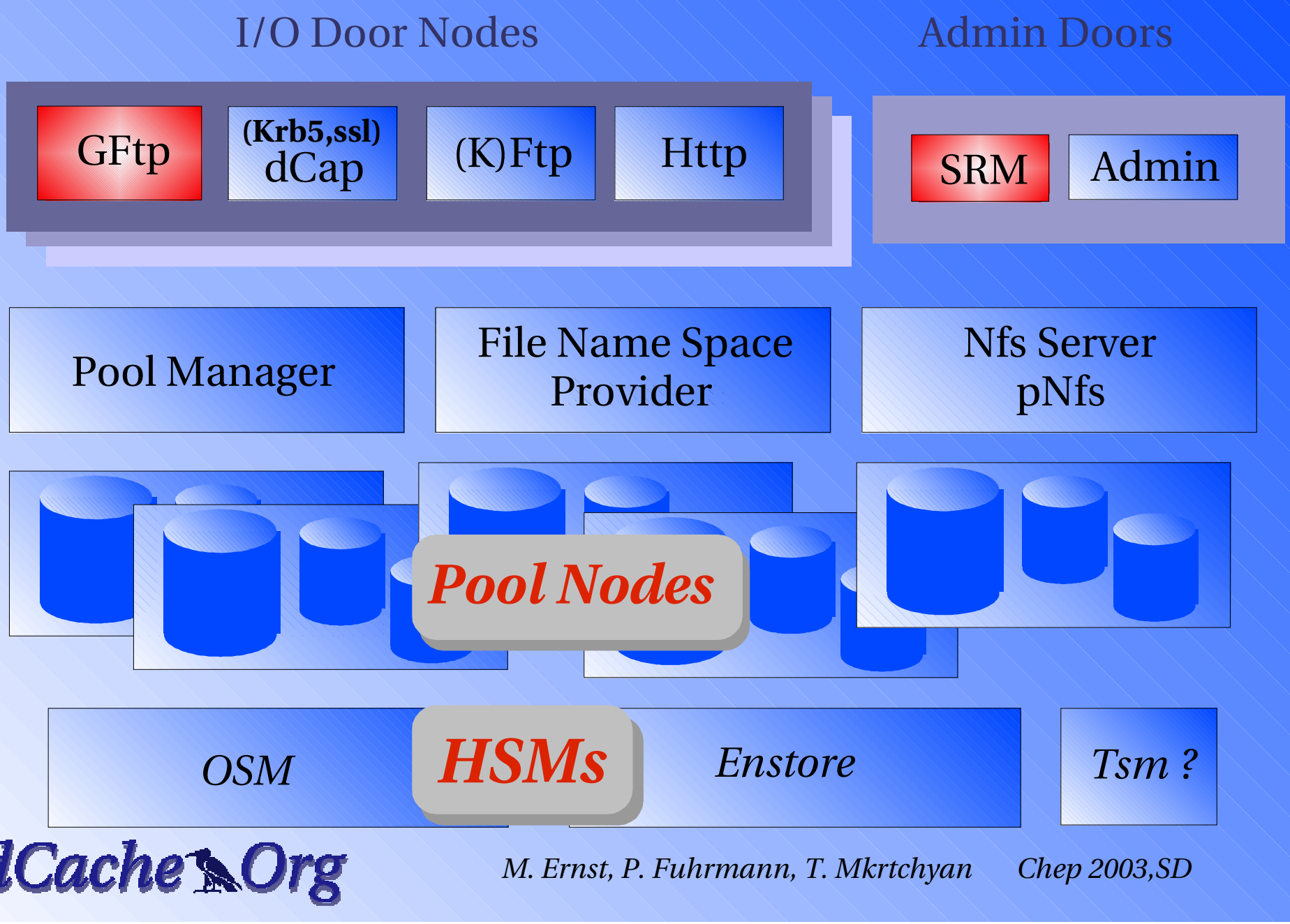
On increasing load of pool nodes, the cache creates dataset duplicates on decently used nodes to smoothen hot spots.

Decreasing load marks dataset duplicates for removal in case space is running short.

Already staged files can be pinned to disk, overwriting dynamic location management.

Datasets can be defined 'sticky', independetly of its status, CACHED, DUPLICATED or ON TAPE ONLY.

# dCache Components



# dCache : Native Access Method (dCap)

Beside supporting Ftp, GFtp, KeberosFtp and Http, dCache defines a native access protocol (dCap), allowing posix like file operations.

dCache provides a dCap c-language implementation

*As shared object or preload library*

*For linux, solaris and irix OS.*

*Supporting automatic reconnect on network or server problems*

*Providing security tunnels for Kerberos and ssl.*

*Interfacing ROOT*

*Win dCap .dll support*

*e.t.c.*

# *The Storage Resource Manager Initiative*

In order to make site local storage resources like disk space, tertiary storage space and large quantaties of HEP information globally available, e.g. in the GRID context, an initiative has been setup by JLAB, FermiLab, LBNL and CERN, defining some kind of abstract storage system, covering methods for :

*Storing and retrieving datasets*

*Obtaining status information about datasets*

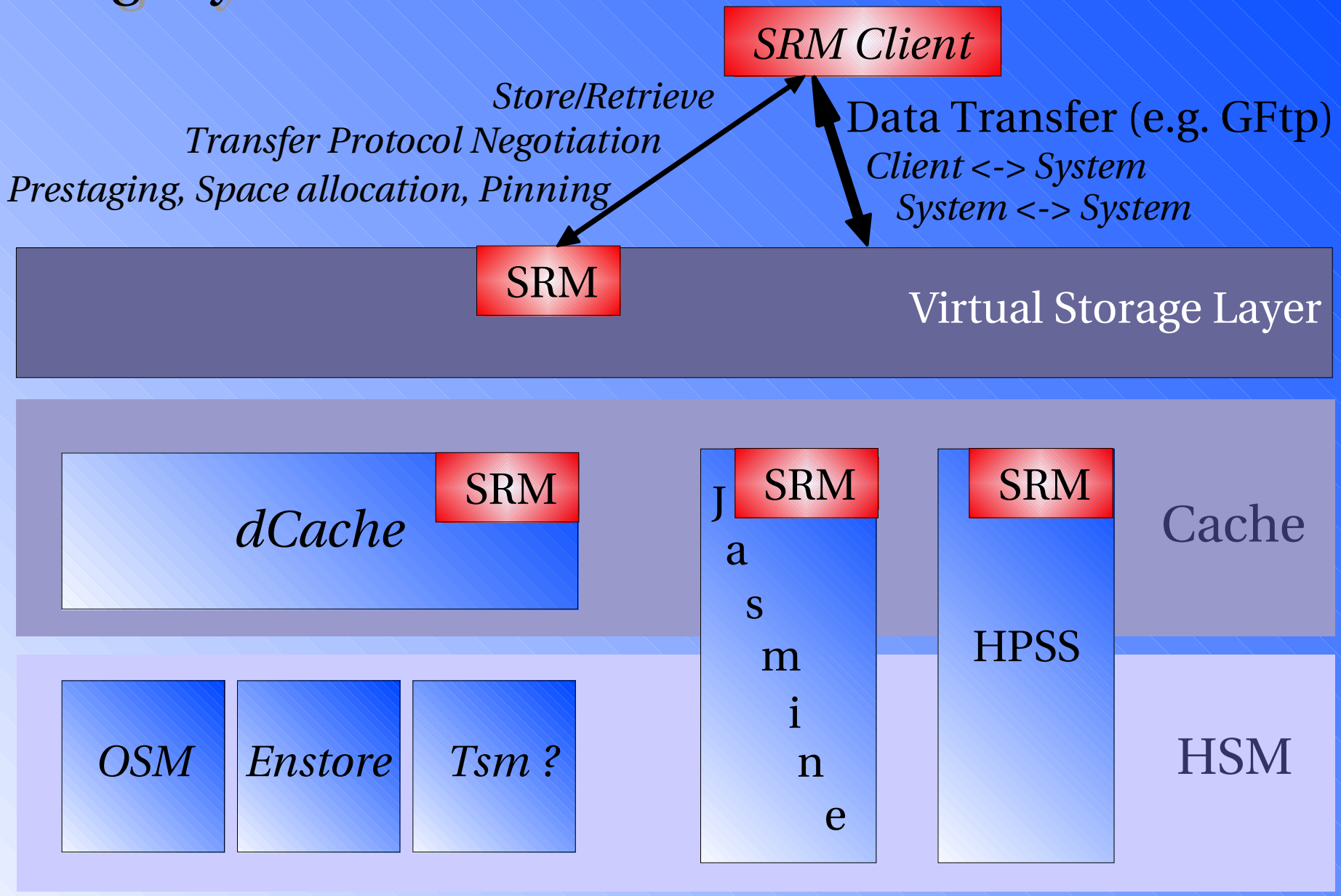
*Pinning datasets (quaranty of availability)*

*Negotiating data transfer protocols.*

*Defining dataset lifetimes and possible transitions.*

HEP sites, intending to access remote storage resources, are assumed to implement the SRM protocol into their local storage system. For the dCache, FermiLab took over this task.

# Storage System Abstraction



# SRM Initiated Transfer

map LFN -> SURL

*CERN-dataset/user/p/file1*

REPLY (SURL)

*srm://srm.fnal.gov/cern/user/p/file1*

GET (SURL)

*srm://srm.fnal.gov/cern/user/p/file1*

REPLY (Request ID)

*#-12345*

STATUS (Request ID)

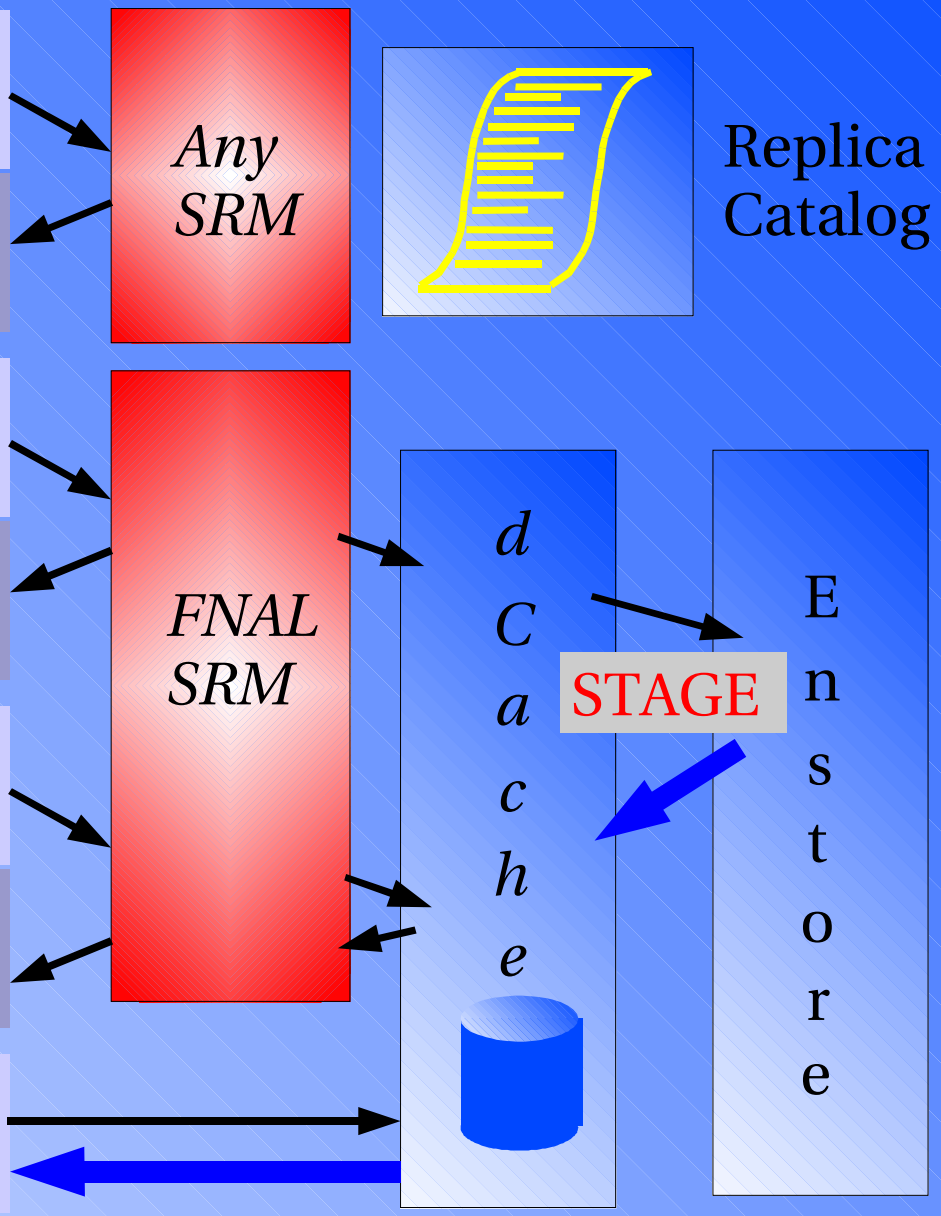
*#-12345*

REPLY (TURL)

*gridftp://pnfs/fnal.gov/cern/user/p/file1*

GET (TURL)

*gridftp://pnfs/fnal.gov/cern/user/p/file1*



# SRM Topics (Srm 2.1)

Storing and retrieving datasets.

Transfer Protocol Negotiation.
*(includes direct I/O methods rfio,dCap globus-xio)*

Obtaining dataset status information.

Pinning datasets (make DS available)

File Space Allocation / Reservation

Dataset / space lifetime definitions (*volatile, durable, perm.*)

Srm to Srm third party transfers.

Directory support (*mkdir / rmdir* )

Security ( *srm will support gsi over http*)

# *For Details*

***dCache*** *www.dCache.org*

***SRM*** *http://sdm.lbl.gov/srm-wg/*