dCache

WLCG SRM 2.2 Proposal

Patrick Fuhrmann
Timur Perelmutov
Tigran Mkrtchyan

dCache.ORG

dCache.ORG

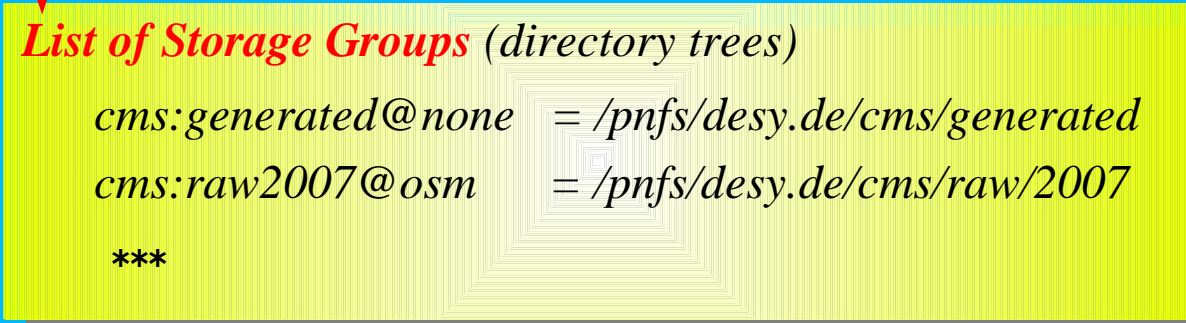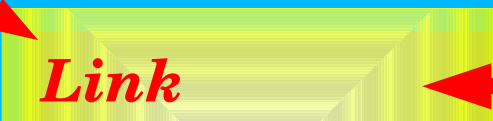dCache.ORG

Primary design Goal :

Adding Storage Classes and Space Management without reducing existing dCache functionality.

Putting SRM 2.2 just on top of dCache (as for 1.x) turned out to be too hard and would have reduced the already available functionality in dCache. So parts of the SRM functionality will move into the dCache Kernel.

DISCLAIMER : *The dCache team intends to implement the WLCG SRM 2.2 as described in this presentation. Though, at the time being, implementation details may still change, which may have consequences on the behavior of Space Reservation and Storage Classes.*

dCache.ORG

SRM Properties (VO,Role,RP,AL) in dCache internal structur

Link Group

Properties

VO

Role

(any KEY/VALUE)

Link

Link

Link

Properties

Retention Policy

Access Latency

Flow Direction

Client Subnet

Protocol

List of Storage Groups *(directory trees)*

cms:generated@none = /pnfs/desy.de/cms/generated

cms:raw2007@osm = /pnfs/desy.de/cms/raw/2007

***

Patrick Fuhrmann   WLCG SRM 2.2 Proposal   Oct 03, 2006

dCache.ORG

Some Rules

dCache internal

PRECIOUS : files which are scheduled for Tape migration only, T0Dx files are pinned

TxD0 : Space Manager is pin owner

TxD1 : dCache system is pin owner

Links get new Properties : Access Latency and Retention Policy

Link-Group Properties : VO and Role

More Rules

General rules, seen by dCache users

Pools with Space Management can NOT be mixed with pools w/o Space Management !

For non SRM transfers : directory (tree) determines Storage Class (TxDx) !

No way to use space tokens with non SRM transfers !

All Storage Classes (TxDx) may or may not reside on the same pool sets, resp. Pools

Pool-Manager can select pools based on all Storage Classes (RP and AL) !

The Space of a Space Tokens is only reserved virtually on a set of pools and is not associated to a particular pool. The actual pool selection is done on the arrival of the file, based on the well known Pool Manager rules, using Storage Group(directory), data flow direction, protocol and client subnet plus the new properties : Retention Policy and Access Latency.

For '*bring on-line*' or '*prepare to ge*t' of T1D0, or '*recover from disk space*' of T1D1, data can NOT go into Space Token Spaces.

More Rules

Space reservation in WAN / LAN pool sets : Three choices

➤ Space Reservation only on internal Pool Sets. External Pools are just transfer Pools.

➤ Use *Transfer Parameters* in the *reserve Space* SRM call which results in reserving Space either inside OR outside of firewall. -> files have to move from space token space into non-space-token space or other space token spaces in order let it cross the firewall border.

➤ One space token may be associated to pools inside and outside of the firewall

Mixed Space Solution

Available Space

Token Space

Used Space
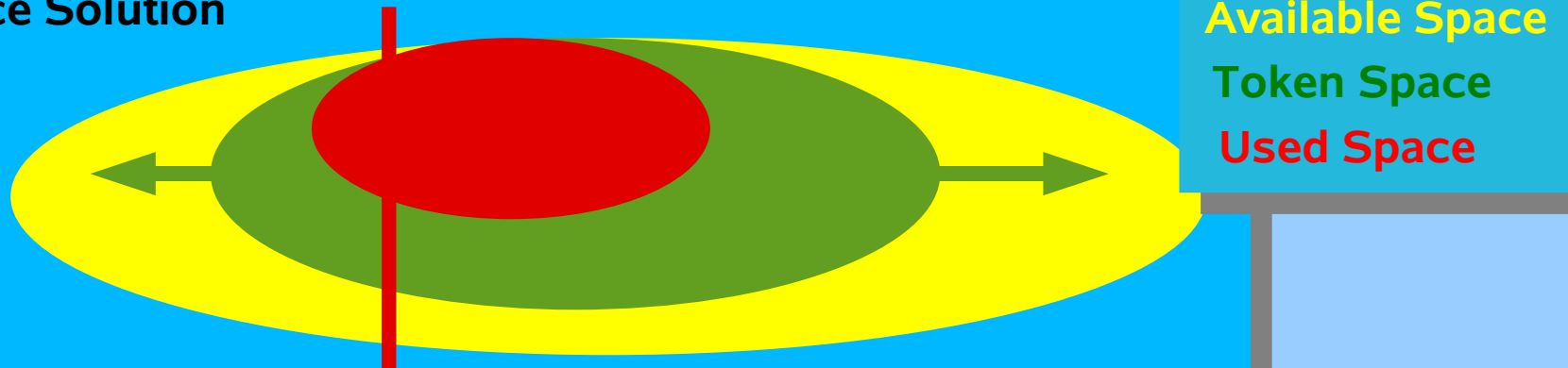
* **Token Space** may move between WAN and LAN, but neither **Used** nor **Available** space can

* If Token Space doesn't entirely fit into either WAN or LAN, reservation promise was a lie.

Remark on Storage Groups and Directories

dCache has one level of indirection between the directory structure and the unit, directory based selection is working on, which is called :

Storage Group : cms:generated[@HSM]

Storage Groups may be attached to any directory and are inherited by subdirectories unless overwritten.
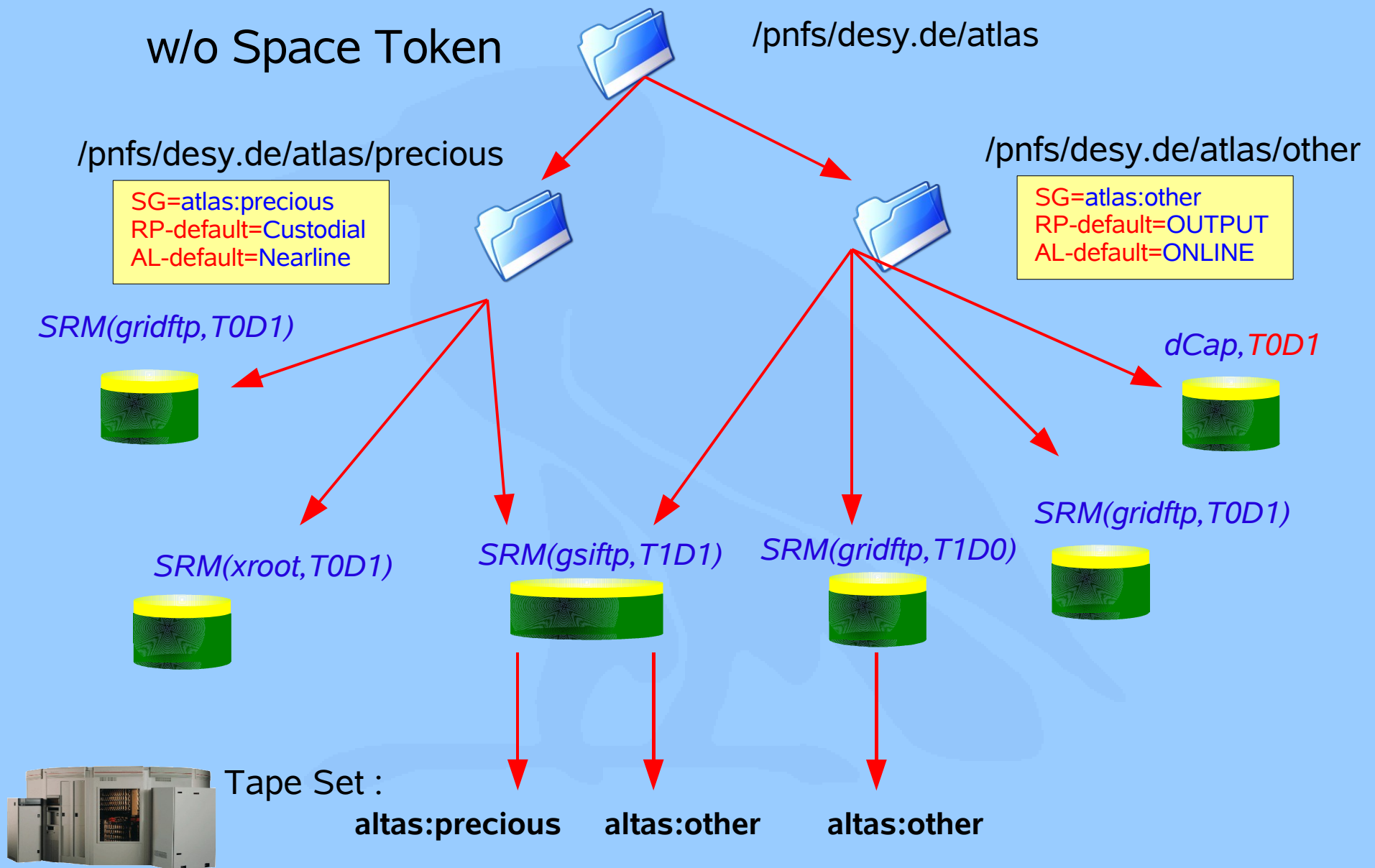
Storage Groups are used to :
- select disk pools
- select Tape Sets in case of Tape Back-ends (and T1)

The Storage Group of a file is identical to the Storage Group of it's parent directory in which it has been initially created.

The Storage Group of an existing file can't be changed but may be different from the Storage Group of its current parent directory, due to 'file move' operations. (This is different from Access Latency and Retention Policy, which may be theoretically changed)
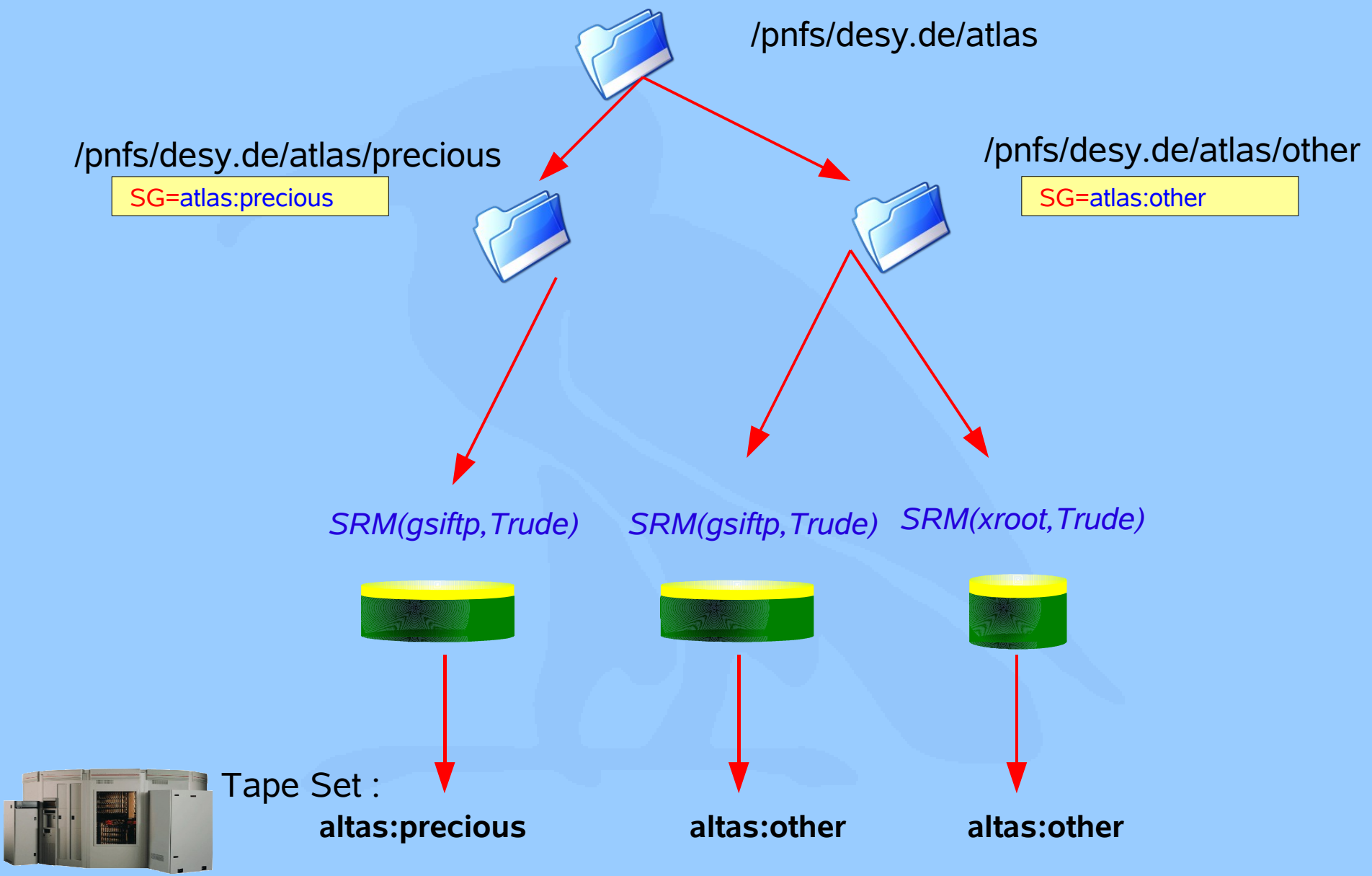
dCache.ORG

File location depending on directory location and Storage Class

w/o Space Token   /pnfs/desy.de/atlas

/pnfs/desy.de/atlas/precious

SG=atlas:precious
RP-default=Custodial
AL-default=Nearline

/pnfs/desy.de/atlas/other

SG=atlas:other
RP-default=OUTPUT
AL-default=ONLINE

SRM(gridftp,T0D1)

dCap,T0D1

SRM(xroot,T0D1)

SRM(gsiftp,T1D1)

SRM(gridftp,T1D0)

SRM(gridftp,T0D1)

Tape Set :

altas:precious   **altas:other**   **altas:other**

dCache.ORG   dCache.ORG   dCache.ORG

Space Token : *Trude := (VO=altas,Role=none,Custodial,Nearline)*

dCache.ORG

/pnfs/desy.de/atlas

/pnfs/desy.de/atlas/precious

SG=atlas:precious

/pnfs/desy.de/atlas/other

SG=atlas:other

SRM(gsiftp,Trude)   *SRM(gsiftp,Trude)*   *SRM(xroot,Trude)*

Tape Set :

altas:precious   **altas:other**   **altas:other**

Patrick Fuhrmann   WLCG SRM 2.2 Proposal   Oct 03, 2006

Storage Class Transitions

T1 D1 $\longrightarrow$ T1 D0   Ok, agreed

T0D1 $\longrightarrow$ T1Dx   Agreed to be non existent

T1D0 $\longrightarrow$ T1D1   Transition done on access of
                                                the data only.

dCache.ORG
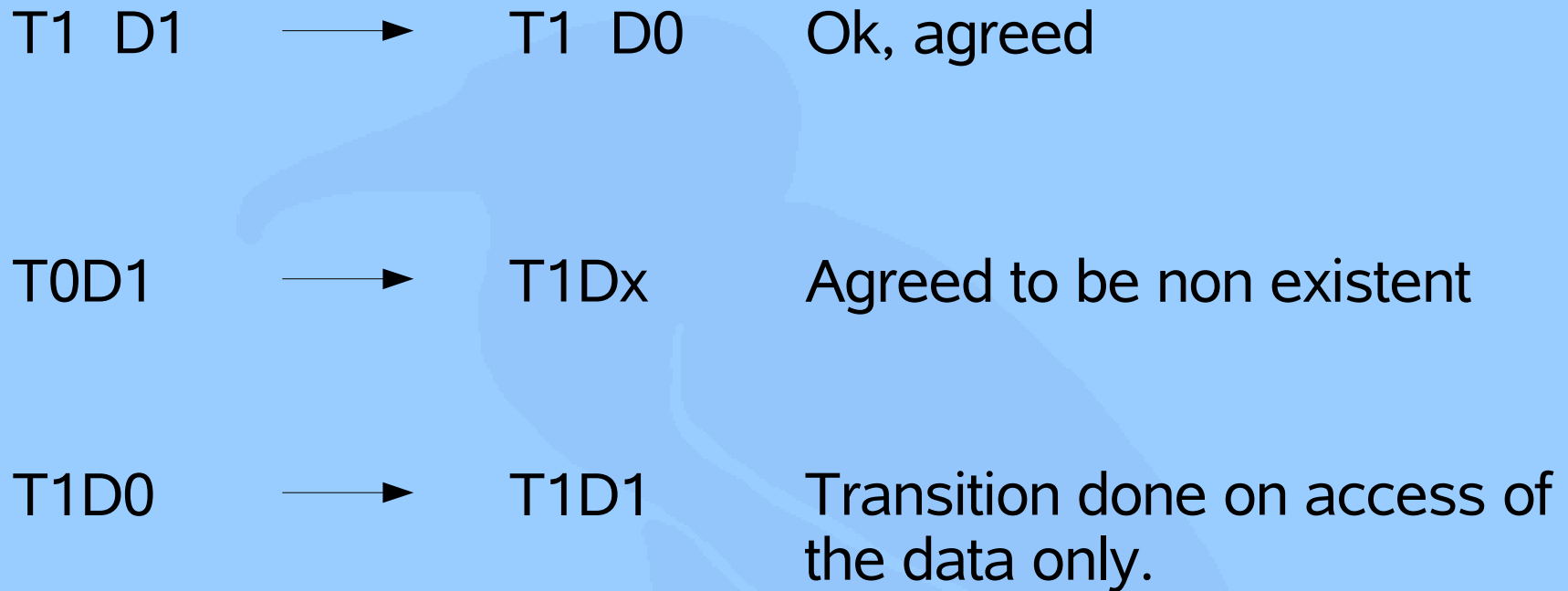
DCache 1.7.0 supports Extended Certificates (voms-proxy-init)
 including VO's and Roles

Migration path

  dcache.kpwd on each host

  With 1.7.0

    dcache.kpwd (centrally)

    extended gridmap file (centrally)

      DN may be in different VO's

      DN may have a role

      All DN of a VO allowed (exclude list)

xRoot protocol support

xRoot natively supported by dCache (including load balancing)
(Has been positively tested at various sites already)

Supports ALICE security model (Catalog based authorization)
(DESY tests ok, CERN tests will be done soon)

gsi xRoot : we need more information on this

passive dCap

dCap protocol now supports passive transfers to handle NAT and firewalls

dCache.ORG

dCache, the Book

www.dCache.ORG

need specific help for you installation or help
in designing your dCache instance.

support@dCache.ORG

dCache user forum

user-forum@dCache.ORG