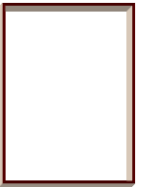# PD2P Planning

Kaushik De

Univ. of Texas at Arlington
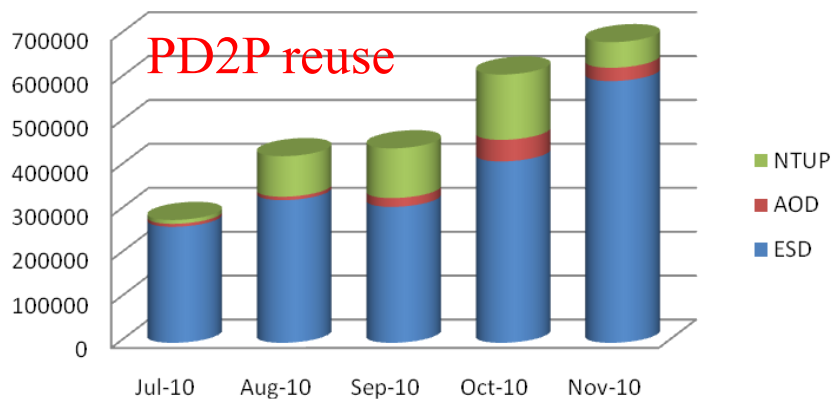
S&C Week, CERN

Dec 2, 2010

What Next?

- Small changes coming soon
  - Change the name – Stephane and Ikuo will make suggestion!
  - Better space checking – Tadashi will implement in ~1 week
  - We may replicate on 2nd use (instead of 1st use) – will reduce ~50% of subscriptions (but not really necessary yet, keep in back pocket)
  - Check, and if necessary fine tune subscription rate to empty sites
  - Group datasets – maybe wait till space token consolidation in spring
  - Some PandaDB schema changes to improve monitoring
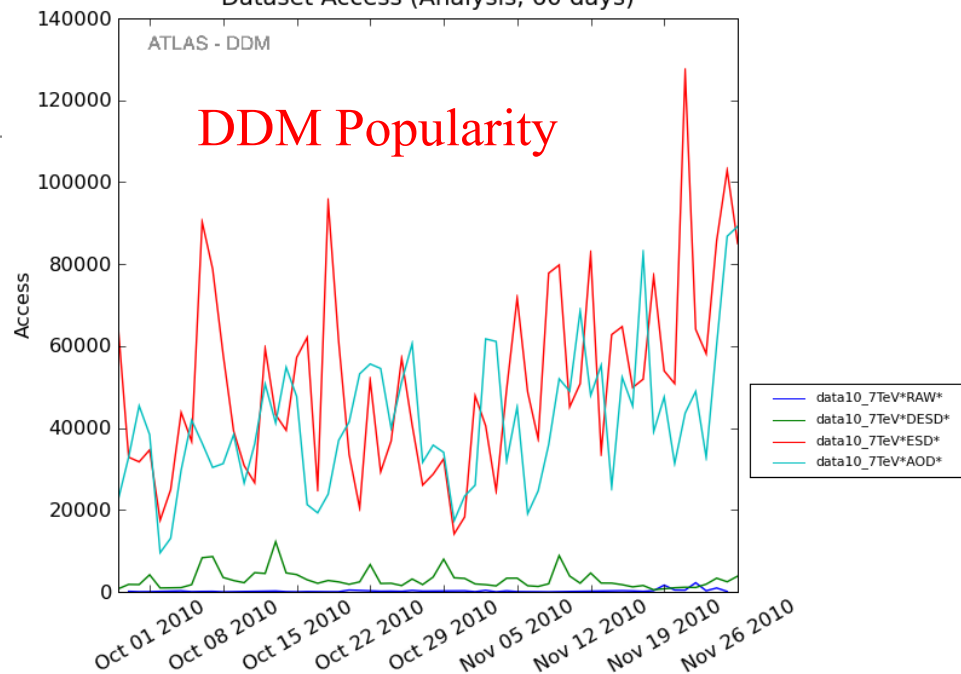- Big change proposed here
  - Start PD2P for Tier 1's

First – Answer Question from Yesterday

Reuse by Data Type
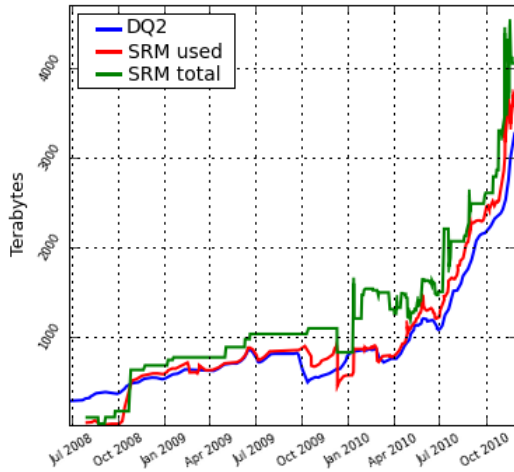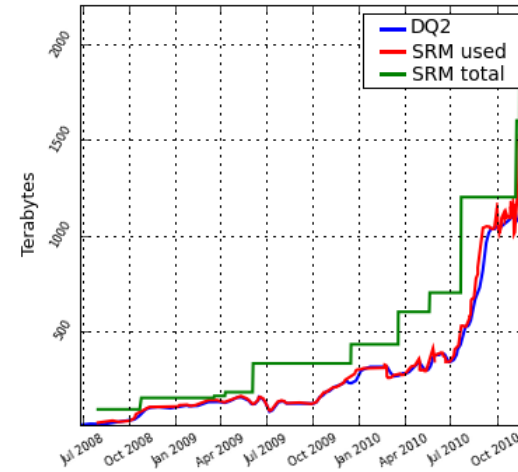
PD2P reuse

DDM Popularity

Dataset Access (Analysis, 60 days)

Second – How bad is the Situation?
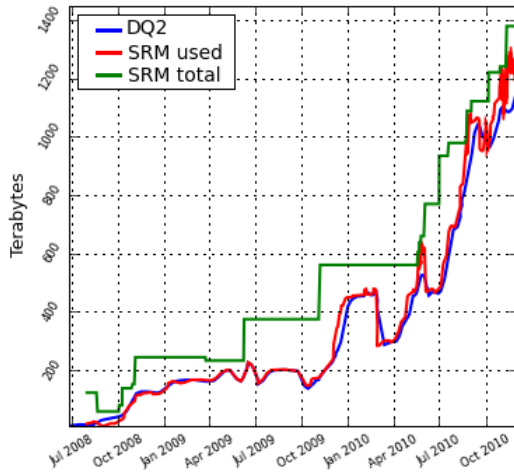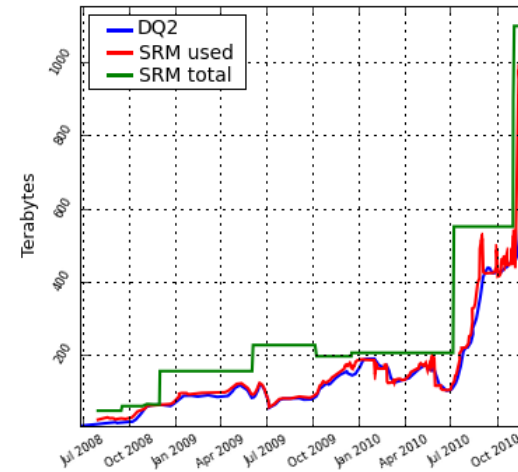
Third – What Can be Reduced?

Storage space evolution at T1 DATADISK acc. to DQ2

Legend:
- ToBeDeleted
- default
- secondary
- primary
- custodial
- SRM total

Don't Forget MCDISK

Storage space evolution at T1 MCDISK acc. to DQ2

Legend:
- ToBeDeleted
- default
- secondary
- primary
- custodial
- SRM total
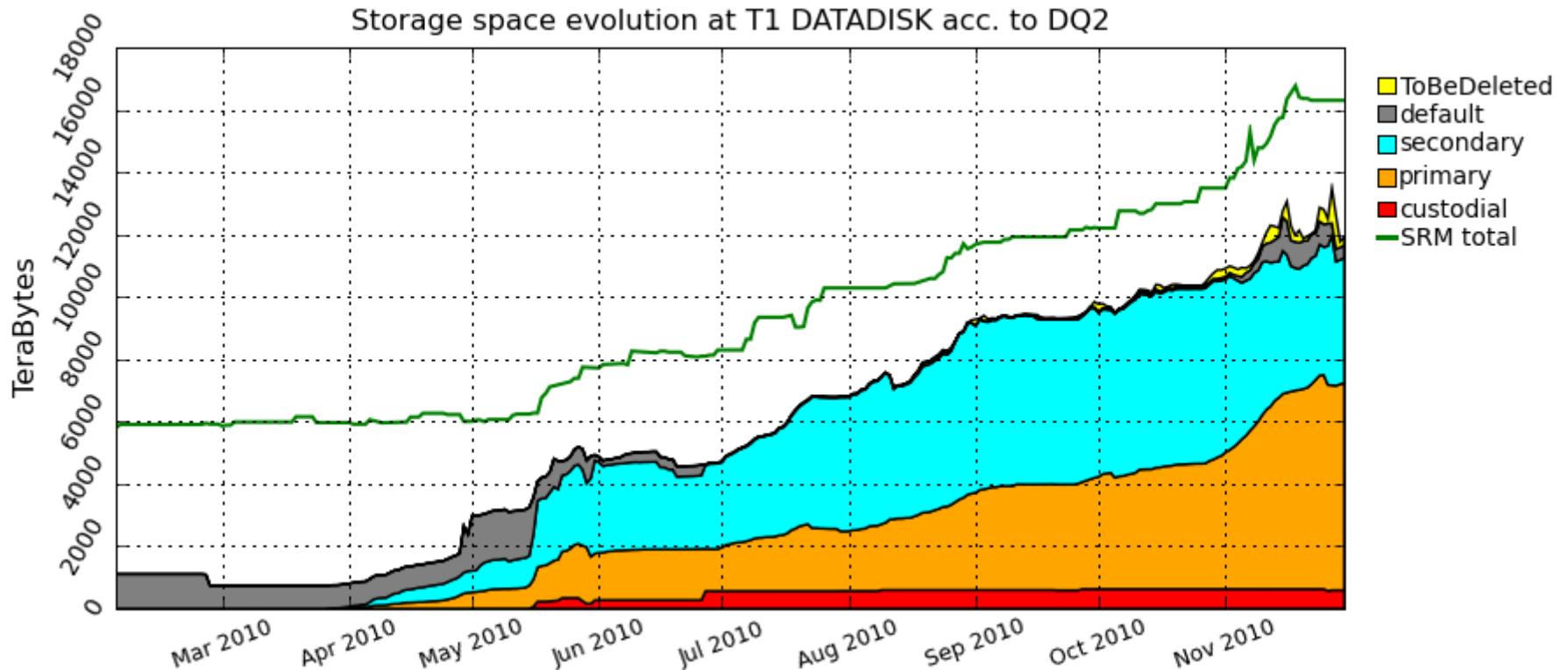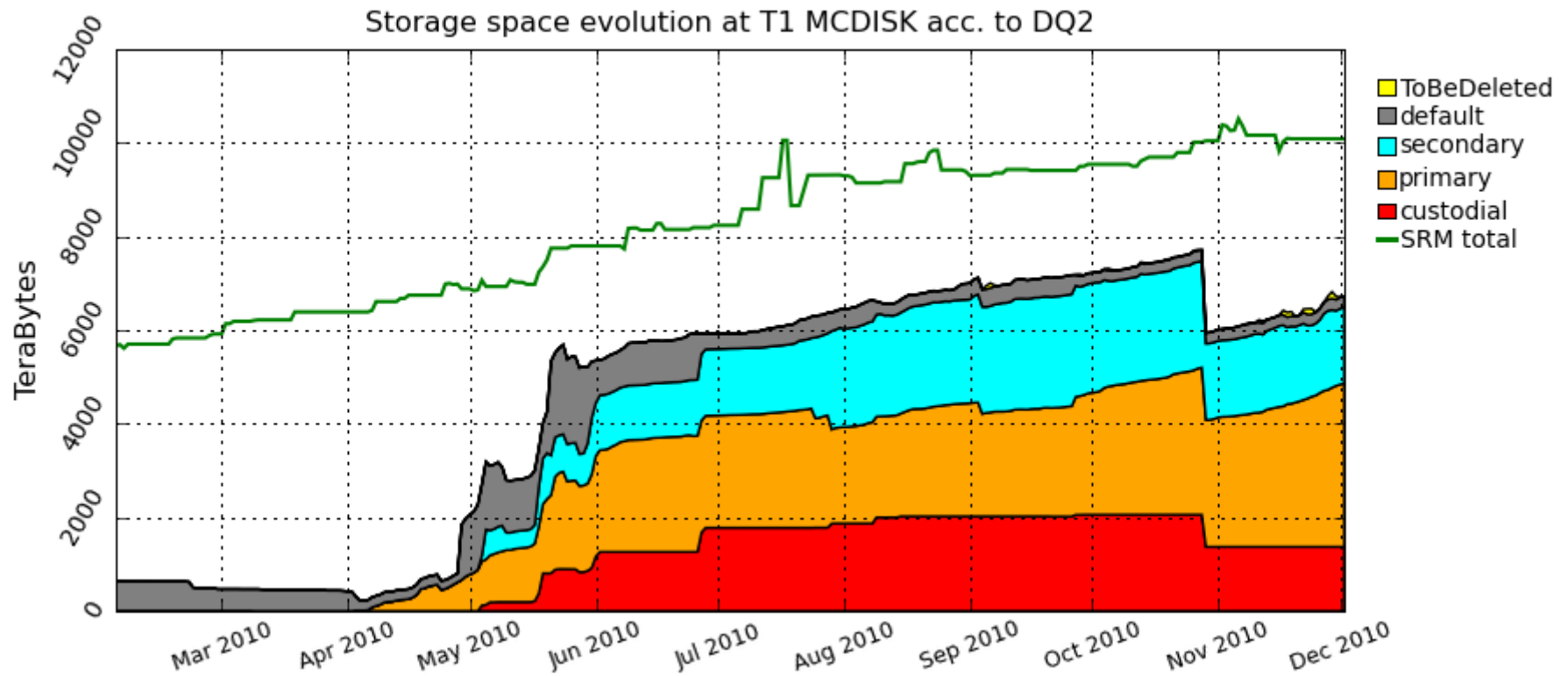
A Possible Solution

- **Step 1: no change to custodial data tagging**
  - Only TAPE copies should be labeled custodial
  - Need to clean up MC datasets currently tagged as custodial

- **Step 2: minimize the number of primary datasets**
  - Always first copy should be primary, if it is not custodial
  - If data is kept only on disk, second copy is also primary

- **Step 3: secondary replication is based on usage**
  - Additional copies requested by CREM, distributed by AKTR
    - But this should be minimized
  - Usually, secondary replicas made only by DaTri and PD2P
  - Can be deleted by Victor if running out of space

Proposal – What is Primary?

- ## First copy at any Tier 1
    - Could come from Tier 0 – distributed according to MoU share
        - Distribution is handled by Santa Claus
        - Maximum number of primary copies should be set to one
    - Data produced at Tier 1 (reprocessing, prompt processing, MC…)
        - Set by AKTR

- ## Second copy
    - Data produced at Tier 1, copied to another Tier 1
    - This data distribution handled by AKTR
    - Note: this means at a minimum 2 primary copies if not on TAPE
    - This prevents blocked data when Tier 1's become unavailable

Secondary Replicas at Tier 1
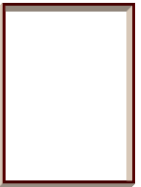
- Any additional distribution requested by CREM
  - But we should minimize/eliminate this
  - Do we need 10 copies of AOD? 10 copies of DESD?
  - Proposal – let PD2P handle all extra copies
- Secondary replicas made by PD2P
  - Check number of waiting jobs
  - If too many waiting (threshold to be decided), and no copies made by PD2P already, start replication
  - Use MoU share to decide which Tier 1 gets extra copy
  - If still too many waiting (that is, higher threshold), make 2nd copy
- What about copies at Tier 2?
  - No changes right now – keep current algorithm (brokerage)
  - We may need to increase number of extra copies to two?