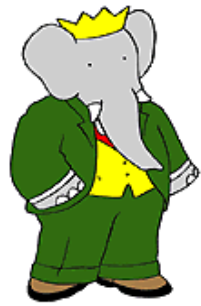




CCRLRC
Rutherford Appleton Laboratory



TM and © Laurent de Brunhoff

Data Distribution and Management

Tim Adye

Rutherford Appleton Laboratory

Computing Model 2 Implementation Workshop

24th June 2003

Three types of “Publishing”

1. “Registration”: Collection entry created in the (global) dataset bookkeeping database
 - All production on this collection at PR/SP/skim site is done
 - All merging has taken place: any subsequent merging (eg. user skims) will *always* create a new collection
 - ROOT File Integrity (RFI!) checked
 - Ready for export to QA site (probably SLAC)
2. “Location”: Collection is available at QA site
 - Some QA sites may also require archive to be complete before marking as located.
 - Could also allow detector studied where bad runs are allowed
3. “Publication”: Collection available for physics use
 - QA has signed off on the collection
 - May be ~week later

Data Distribution

- Location export (Padova/SP/skim→SLAC) is push
- SLAC→Tier A→Tier C is all pull
- Export may be from
 - dedicated host (datamover)
 - analysis disks via NFS (or XRootd?)
 - MSS (eg. HPSS)

but access method is up to each Tier A

- may be trivial (NFS) or not (bbftp-to-xrootd?)

Data Management

- Artem has a long list of use-cases and requirements for the XRootd
 - We added to the list
 - Artem will collate and send to Andy
- XRootd will only be used for reading
 - At least for the moment
- Schedule:-
 - XRootd-enabled release (**XNetFile**) will be available this week
 - Jean-Yves (ccI N2P3) and Chris Brew (RAL) will start testing
 - “Full” version of XRootd available by end July