

Data Distribution

Tim Adye
Rutherford Appleton Laboratory

BaBar Collaboration Meeting
15th February 2000

Objectivity Export

Tools

Exports

KanGA Export

Techniques

Bulk exports

Objectivity Export Tools

- **Book-keeping** database up and running (Teresa Barillari)
 - Maintains a catalogue of exports
 - Accessible from the web
- Red Queen's race to maintain **2 Gb file** support in the latest releases (Dominique Boutigny, TJA)
 - 8.2.x OK.
 - Problems with 8.3.x – 8.4.x .
 - 8.5.x (+ BdbAccess tag) OK
- Large (>**2 Gb**) file exports (TJA)
 - SLAC/IN2P3 staging has 2 Gb limit
 - Workaround implemented

Objectivity Export Tools - ongoing work

- Use collection database (**collldb**) (Moreno Marzolla, Artem Trunov)
 - Replace (slow) Objectivity Federation scan with (fast) Oracle query
- HPSS support (TJA)
 - Bypass disk-resident Federation
 - Maybe faster. Maybe slower.
 - Reduce impact on/from other users
 - Being tested.
- **Graphical User Interface** (Cristina Bulfon)
 - Simplifies export/import process
 - First version looks good!
 - *Needs work for production use*

Objectivity Export Tools - The Next Generation

- **Redesign** of BdbDistTools using Java/JNI being investigated (Jean-Noel Albert and Yemi Adesanya)
 - Present design is quite **byzantine**
 - *Perl, TCL, sh, C++*
 - Continue with present tools in parallel

Objectivity Exports from SLAC

- Dominique Boutigny and Cristina Bulfon have continued making regular exports to **IN2P3**, **RAL**, and **CASPUR**.
 - ~1 export / 2 weeks
 - **Analysis** and **SP2** mini/micro data
 - *No digis or reco data*
 - **isPhysicsEvents** and **Analysis Working Group (AWG)** skims

Future Objectivity Exports

- Last export **811 Gb** on 17 DLTs
 - If this continues, will very soon **overwhelm** regional centres
- Propose to
 - Stop export of **ESD** (“miniDST”)
 - *not currently usable*
 - Suspend export of **AWG skims**
 - *Need to understand why these are so large*
 - *One large component (AIO) can probably be dropped immediately*

Hopefully this will reduce future exports by a factor of **3**.

KanGA Exports

KanGA Exports

- Exporting **KanGA** (née NOTMA) from SLAC is much simpler than Objectivity
 - Ordinary Unix files and directory structure
- Only network transfers so far
 - Eg. SLAC -> RAL -> CASPUR
 - *Full 25 Gb EventStore*
 - Greatly simplified by mirroring tool, **rsync**
 - *Any site with sufficient network bandwidth can use this **now***
 - *See Data Distribution web page for details*

KanGA Tape Export

- Still need to develop tape export tools
 - Catalogue what's been transferred
 - Handle SLAC/IN2P3/direct tape access
- Initial version will be based on existing **BdbDistTools**
 - Already does much of the same job
 - Use same sort of catalogue
 - *TDF and web*

DLT Drive

- Currently all DLT exports via DLT **stacker**
 - Single drive; only 39 slots
 - attached to **tapeserv2**
- New DLT **robot** awaiting installation
 - 3 drives; 300 slots
 - will be directly attached to our export machine, **datamove3**

Bulk KanGA Exports

- Won't do bulk KanGA export before **reprocessed** data is available
 - 2nd half of March
 - Export all reprocessed KanGA data
- Bulk exports on **DLT-7000** only
 - IN2P3 excepted: Redwood/Eagle
- Invite expressions of interest to receive bulk KanGA tape exports

Site requirements:-

- DLT-7000 drive
 - *NB. Not DLT-4000*
 - *Own DLT tapes (~\$64/tape)*
- Sufficient disk space (>100 Gb?)
- Manpower to make imports

Bulk KanGA Exports

- Encourage some groups to “share” tapes
 - DLTs sent from one institute to another
 - + Less work overall
 - + Cheaper (tapes, postage)
 - Groups must coordinate
 - Some won't get data as rapidly
- Could still use network copy from nearby institute
 - Requires very good network connection
 - Still require manpower and disk space

Bulk KanGA Exports

- Let me know if you want to receive **bulk** KanGA export:-
 - Contact me directly or speak up at **CCG Meeting**
 - What sort of export?
 - *DLT or network?*
 - *Is there anyone else you can share with?*