

BaBar Data Distribution

Adil Hasan

BaBar Computing Review

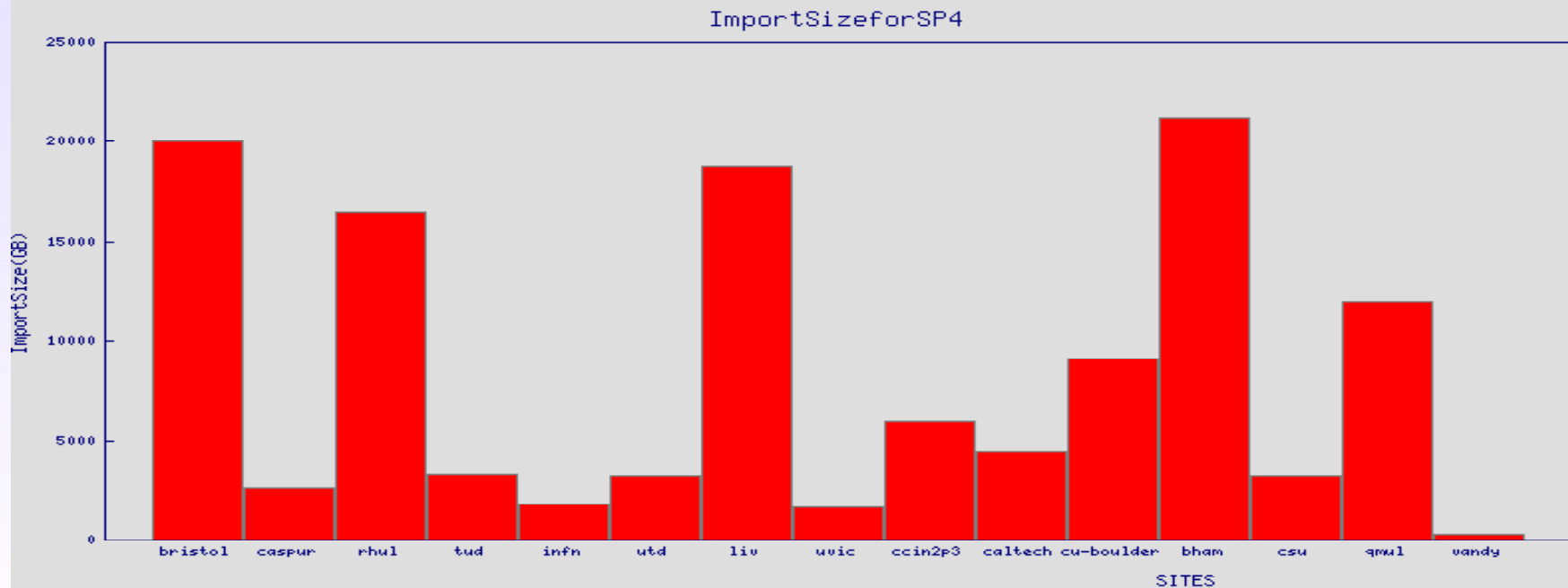
April 11th 2002

SLAC

Brief History (Objy)

- *SP4: To date 127 TB of data import from external sites.*
- *Both Tier A and Tier C shipped data to SLAC.*
- *Sites located in Germany, UK, France, Italy, USA, Canada - truly distributed effort!*

Snapshot of Total size of imports from external sites



Brief History (Objy)

- *Export from SLAC to in2p3:*
 - *~75TB export from SLAC to ccin2p3.*
 - *Regularly ship 300GB in < 1 day.*
 - *In addition ccin2p3 have also export data to SLAC*
- *General tools used for data distribution:*
 - *BdbServer and JImport (SLAC -> in2p3).*
 - *MocaEspresso (SP sites -> SLAC).*
 - *All use BdbDistTools components and bbftp or bbcp to copy data.*
 - *All tools mainly written in C++, Perl.*
- *Automated Import procedure at SLAC to cope with increased growth of SP production sites.*
 - *Book-keeping of each sites import in Oracle.*
 - *Import procedure has automatic integrity checking of databases.*
 - *Recovery procedures are necessarily manual.*
 - *Automated tools also at in2p3 with very similar capabilities*

Brief History (Kanga)

- *Kanga collections made by running prod jobs off Objy collections at SLAC.*
- *Shipped ~8TB Kanga data to RAL since beginning 2001*
- *Data to transfer selected using skimTools.*
 - *In bulk copy mode new data is found by mirroring mySQL database at SLAC and RAL.*
 - *Recently Alessandra Forti worked on providing a tool so mySQL database update after each sweep from OPR to analysis.*
- *Use bbftp for transfers (a bit inefficient as opens a connection for each file - lots of small files).*
- *Currently ~12 sites import Kanga data:*
 - *US sites take Kanga from SLAC.*
 - *European sites take Kanga from RAL.*

Brief History: Problems

- *Data dist problems:*
 - *Unreadable databases:*
 - *mainly due to outstanding transactions on extracted databases.*
 - *Few problems because of bad disks corrupting databases.*
 - *Missing databases:*
 - *Problems in MocaEspresso which are fixed now.*
- *A majority of these problems are picked up by the integrity checking when importing to SLAC and in2p3.*
 - *However, there are some problems (~2%) that pass the integrity checks.*
 - *Improvements to the existing QA tools should catch the remaining problems.*
- *Data dist problems occupy up to 30% FTEs (depending on the experience of the person).*

Collection Export: BdbServer

- *BdbServer is the main tool for shipping Objy data from in2p3 and SLAC to external sites.*
 - *originally designed at ccin2p3 for bulk data transfers.*
 - *Uses BdbDistTools components.*
- *BdbServer has easily been extended at ccin2p3 to ship individual collections (either pointer or deep-copy collections).*
- *It was recognised that the distrib of pointer collections means that undesired as well as desired data is export.*
- *BdbCopyJob allows the user to make a deep copy of a pointer collection of interest, thus eliminating the undesired data.*

Collection Export: BdbServer

- Ccin2p3 are currently making a deep copy of Stream17 events for Caltech using BdbCopyJob.*
- So far, ~25% of current total has been copied using BdbCopyJob and export using BdbServer to Caltech.*
- Have experienced very few problems with the copy or extraction of data.*

Future: Data production

- *SP data distribution (to end of next year):*
 - *Expect approx same data volume as now (not shipping raw or sim, but adding more sites). I.e. ~1.5 - 2 TB/week shipped to SLAC.*
- *Remote reprocessing in Padova (starting July):*
 - *Plan is to carry out reprocessing in Padova.*
 - *Will require a dedicated data server to send/receive data at SLAC.*
 - *Will require slight modifications to the automatic import tools at SLAC.*
 - *MocaEspresso (modified) will be the main tool to distribute data from Padova to SLAC.*
 - *Initial plan is to ship all data to SLAC.*
 - *It's possible that in2p3 will take data directly from Padova in the future.*

Future: Tools

• BdbDistTools:

- Originally developed at Orsay by Jean-Noel Albert.*
- Has provided the foundation applications for many of the Objy data distribution tools.*
- But, is getting more difficult to maintain.*
- Have embarked on re-design of BdbDistTools keeping in mind:*
 - Ease of maintenance.*
 - Ability to plug in Grid applications.*
 - Ease of use at Tier A, C sites (i.e. shouldn't require any additional s/w outside of what's needed for a BaBar release).*
- The re-design is expected to be completed by end August.*
- Factors that have caused time slippage are:*
 - Developers for new system have to maintain old system*
 - Developers for new system also have to administer the Objy database*
- Manpower for BdbDistTools re-design:*
 - Wilko Kroeger (~30%)*
 - Adil Hasan (~20%)*

Future: Tools

- *MocaEspresso:*

- *This is the tool used by all external SP sites to distribute data.*
- *There are still a number of improvements that need to be implemented.*
- *The package maintainer (and architect) is moving off to help support the Padova reprocessing farm.*
- *We need help to find a person willing to take on the responsibility for maintenance and future development of this critical tool!*
- *Ideally someone (or more than one person) from one of the SP producer sites.*

Future: Tools

• BdbServer++:

- The goal is to allow users to create and extract deep copies of pointer collections through BdbServer.*
- Work done so far:*
 - List of use cases and requirements collected for the system.*
 - Manual walk through of the separate components of the system at in2p3.*
- Work to be done:*
 - Integration of BdbCopyJob with BdbServer*
 - Stage III of bridge federations (see Jacek's talk).*
 - RDBMS containing collection-database mapping.*
- It is expected that a working version of the BdbServer++ will be ready on the timescale of Run3 (late this year).*
- Manpower:*
 - Dominique Boutigny + others at in2p3*
 - Akram Khan, Alasdair Earle Edinburgh, UK.*

Future: Tools

- *Kanga server (export mechanism):*
 - *As with Objectivity, Kanga pointer collections imply an inefficient space usage for recipient sites.*
 - *There is an effort underway to follow the model of BdbServer and produce a "KangaServer" capable of creating a deep copy of collections of interest and ship them to the interested site.*
 - *This work is underway in the UK (see Ulrik's presentation).*

Future: Tools

• QA tools:

- *Better QA tools should reduce the time DB Importers spend tracking down problems.*
- *Current work:*
 - *Compiling a list of outstanding problems.*
- *To be done:*
 - *Create a set of tools that are sensitive to these problems and incorporate them in a package usable at producer sites.*
 - *Combine all the existing tools used at all sites into this package.*
- *This will be an iterative (but convergent) process.*
- *Manpower:*
 - *Tofigh Azemoon (~20% increasing to ~50% in next few weeks).*
 - *Data Dist contributors from all sites.*

• Collection to database mapping:

- *A critical component of the Objy data distribution is the production and maintenance of a metadata store containing the mapping between collections and databases.*

Future: Tools

• Collection to database map

• Current:

- Are scanning all data and SP federations to produce the mapping which will be stored in Oracle at SLAC.*
- Are up to date with SP, 1 month behind with data.*

• To be done:

- Tools to extract and populate similar stores (not necessarily using Oracle) at remote sites.*
- Improvements/optimization of the scanning tool.*
- Integration with Grid tools.*
- Will leverage expertise in the PPDG (and other Grid projects)*

• Manpower:

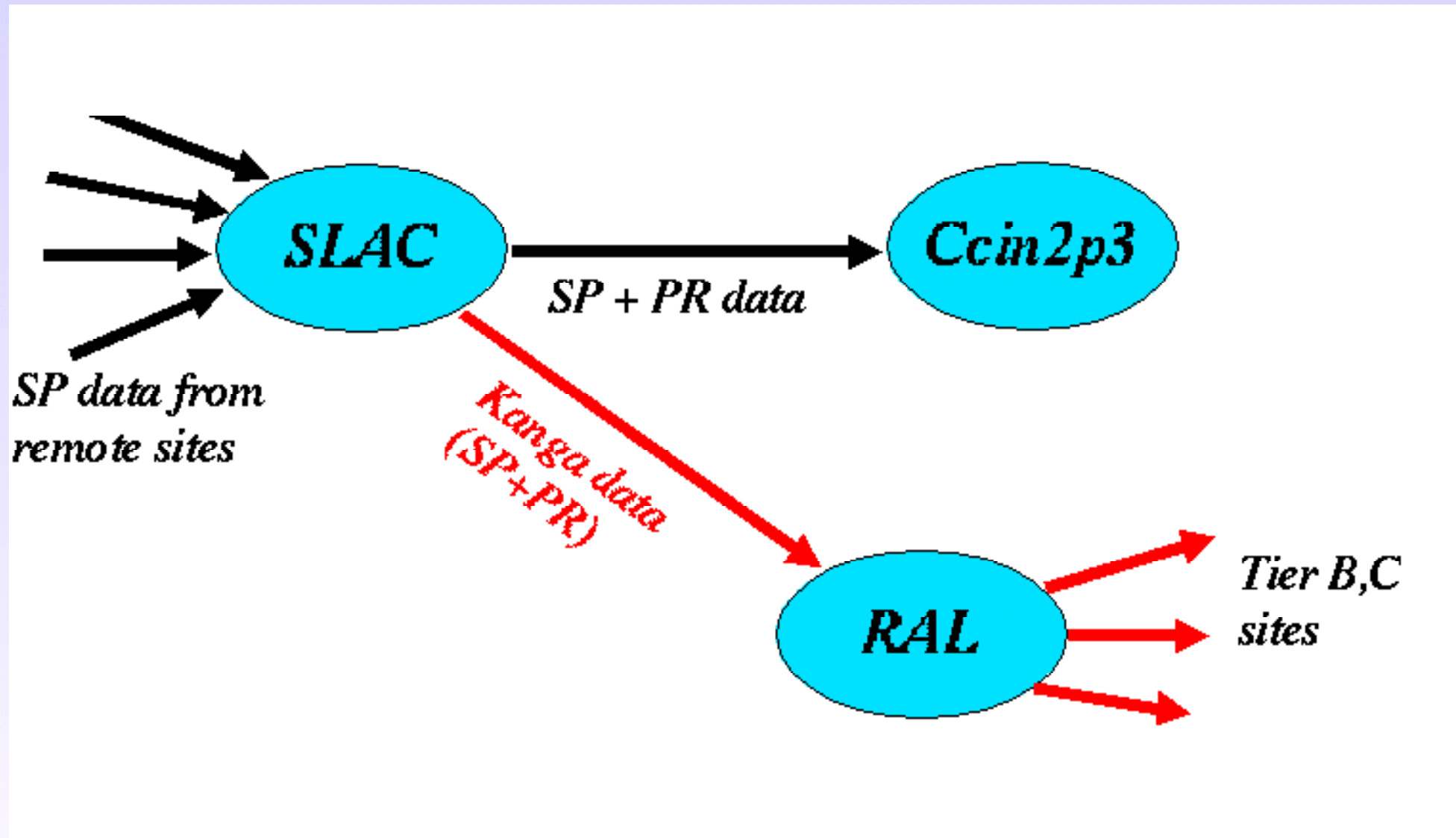
- Wilko Kroeger (~30%)*
- Liliana Martin (~50%)*
- Adil Hasan (~20%)*
- 1 Stanford RA (~50%)*
- + others from Tier A's in the near future*

Future: Requirements

- *Software:*

- *The data distribution plans consider the need to burden external sites with as little extra software as possible of utmost importance.*
- *Kanga:*
 - *MySQL database (currently at Tier A, B and C).*
 - *BaBar software and Root (including bbftp and bbcp).*
- *Objy:*
 - *BaBar software and Objectivity (including bbftp and bbcp).*
- *Both data formats will require collaborators to possess Grid certificates in order to use the Grid-enabled tools.*

schematic dataflow



Summary

- *BaBar Data Distribution faces interesting times...*
 - *Implementing the distributed nature of computing model.*
 - *Incorporating Grid tools into data distribution.*
- *The successful implementation of the computing model depends on the continued work and close collaboration of all members of the data distribution group in all BaBar institutes.*