

BaBar Data Distribution

Adil Hasan

BaBar/SCS

Gilchriese Review 11-12 October
2000

Current Status (I)

- All Objectivity micro-level data replicated at in2p3 (France).
- All ROOT (“Kanga”) micro-level data replicated at RAL, INFN and 6 other sites.
- Sub-sets of Kanga data at ~6 other institutions.
- Exports from SLAC carried out over the network.

Current Status (II)

- Simulation regularly import to SLAC.
- All levels of simulation imported.
- 4 external sites: in2p3, Caltech, CU (Boulder), INFN (Rome).
- Simulation stored in Objectivity
- Imports carried out over the network.
- ~50% (best) simulation produced by external sites.

Data Distribution Details (I)

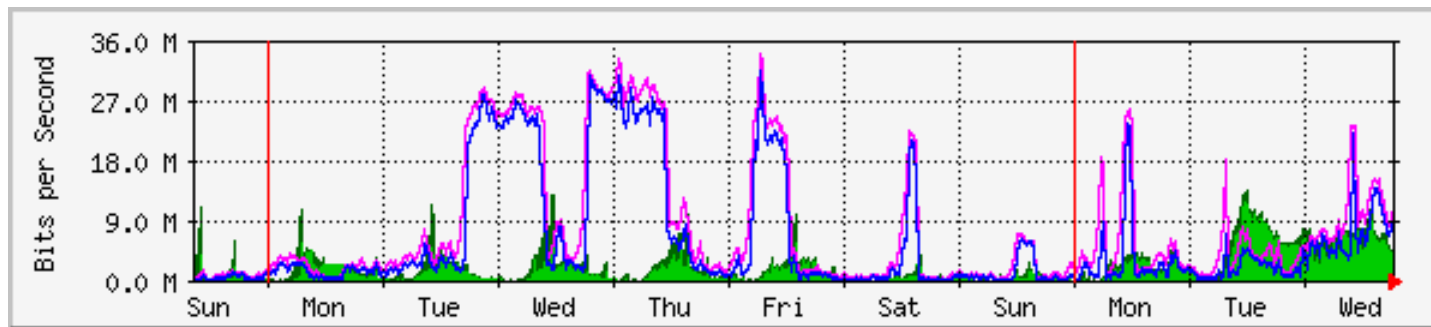
- ~300GB Objy micro-level data export to in2p3.
- Each Objy file ~2GB in size.
- Export once/week.
- >150GB Objy simulation import to SLAC.
- ~50GB Kanga data daily copied to RAL, INFN.
- European institutions copy Kanga from closest regional center.

Data Distribution Details (II)

- Have T3 (45Mbps) link to ES-net.
- Means ~35Mbps usable bandwidth.
- Frequently close to saturation:
 - Increased time to transfer data.
 - Increased packet-loss.
- Will not cope with increased demand.

Data Distribution Details (II)

Weekly network utilisation SLAC to ES-net
(30 minute average)



Problems: Objectivity (I)

- Data Distribution tools kept up with demand, but:
 - Not easy to use.
 - Not robust enough.
 - Slow in selecting files for export.
 - Very rudimentary error checking.
 - Very manual.
- Have started work on replacement tools (Java).

Problems: Objectivity (II)

- Large files (~2GB) not closed after copy:
 - Meant some files copied more than once.
 - Requires outage of federation to update files.
 - Inefficient use of the available bandwidth.
- Have started closing files after copy:
 - Reduces federation outage time.
 - Efficient use of current network resources.

Problems: Objectivity (II)

- File size studies carried out during summer:
 - Are able to account for almost all space.
 - Has led to ~20% reduction in header file sizes.
- Size study being redone with newer releases.
- Means same size files will contain more data.

Problems: Objectivity (III)

- Exporting skimmed data:
 - Contains pointers to original data:
 - Means original data must be copied.
 - Large space overhead for small skim set.
- Can copy all data for a skim to one file:
 - Only data of interest copied.
 - But, is slow to do.

Problems: Kanga

- Same event may appear in more than one stream.
 - Inefficient space usage.
 - Inefficient network usage when copying many streams.
- Current tool for copying data not scalable.
 - Now takes ~3hours to find data to copy.
 - Not easy to select a sub-set of the data.
- Working on replacement for current tool.

Future Data Distribution (I)

- 3 Tiers of sites (from Computing model):
 - Tier A: SLAC, in2p3, ...
 - will have all data (at least 30% at each).
 - Tier B: RAL, INFN (Rome), ...
 - Will have selected micro-level streams.
 - Tier C: Universities.
 - Will have sub-samples for user analysis.

Future Data Distribution (I)

- This model should improve distribution since:
 - Need only copy from nearest Tier A,B site.
 - Reduces the load at any one Tier A,B site.

Future Data Distribution (I)

- Computing Model also advocates smaller files:
 - Makes more efficient use of space:
 - Only files for stream of interest copied.
 - Tier B, C sites copy less redundant data.
 - Makes more efficient use of the network.

Future Data Distribution (II)

- Expect increase from 4 to 20 streams:
 - Each stream self-contained.
 - Quicker to find files associated to streams of interest.
 - Only streams of interest need be copied.
 - Means less data to copy.

Future Data Distribution (III)

- New distribution tools being developed need to be:
 - Site independent (or at least easily configurable).
 - Capable of being setup with very little expertise.
 - Easy to automate.

Future Data Distribution (III)

- Need to reliably keep track of data location.
 - Implies need for meta-data catalog.
 - Sharing of meta-data information between sites.
 - Research into Data Grid technologies.

Future Data Distribution (IV)

- 1999-2000 Run resulted in:
 - 20 fb⁻¹ Integrated luminosity so far.
 - ~160TB data stored (includes all reprocessing).
- Combined 1999-2001 expected to produce:
 - ~65fb⁻¹ Integrated luminosity.
 - ~400TB total data volume.
- Expect 10 months to (re)process 1999-2001 data.

Future Data Distribution (IV)

- Data distribution to in2p3.
- Average data rate ~400TB/10months ~ 100Mbps.
 - Includes all levels of data.
 - Includes all reprocessing cycles.
- Also have simulation import/export to SLAC and Tier B/C exports.
 - Adds another ~100Mbps.
- Current network capacity will not cope.

Future Data Distribution (IV)

- Need to have network capacity increased.
- Must factor in that:
 - Need enough head-room for optimum performance.
 - Cannot run at peak capacity all the time.
 - Need to have the capacity to catch-up in case of problems.
- Estimate need ~1200Mbps capacity from SLAC to other Tier A, B and C sites.

Summary

- Data Distribution has kept up with demand.
- Have demonstrated the ability to saturate the network.
- Current distribution tools are being re-worked:
 - Built on current experience.
 - Anticipating future needs.
- Crucial to have increased network capacity to handle increased data volume and demand.