

Notes on the BaBar Common Fund and Computing

John Womersley

8/22/02 (*this version edited 9/28/02*)

Here I attempt to summarize information on the BaBar Common Fund and how offsite computing is handled. Information comes from Richard Mount via telephone and e-mail, August 2002. (All mistakes are probably mine!)

Below are the logical steps in numerical order:

0. Operations Common Fund

There is an Operations Common Fund (OCF) to cover the cost of operating the experiment = \$2.9M in FY01. This includes online computing. The sharing fractions are currently SLAC(US) 51%, Italy 18%, UK 12%, France 10% + others at 3% or less.

The Computing Contribution is then handled as below. This is all done in an Excel spreadsheet by Richard, and currently includes estimates by year out to FY05. It is updated annually, I believe.

1. Estimate Luminosity

This drives the calculation (how much data to be handled). They have added 20% contingency over accelerator estimates. SLAC originally promised to handle 30fb⁻¹ per year. What is needed to be covered by the computing common fund is whatever it takes to process more data than this. (Note that SLAC still pays the US share, 51%, of this extra computing).

2. Estimate needs for Production processing

CPU and disk required for prompt (first pass) reconstruction, for re-reconstruction, and for simulation. BaBar plan to re-reconstruct all their old data each year, and since their luminosity is increasing fast this makes sense for them. Numbers are scaled from actual CPU used in '01.

3. Estimate needs for Analysis Processing

CPU and disk required for analysis based on accessing datasets N times.

4. Unit Costs

Based on SLAC's in-house trends analysis.

5. Subsystem Costs

The approach used is to cost a unit of CPU or unit of disk "encumbered" by all the other stuff that is needed to make it useful (servers, network, tape drives, mass storage, etc.) This way you can

simply scale the number of SpecInts and TB of disk at any center to get a SLAC equivalent cost, without needing to worry about different architecture choices at the different centers. The "encumbering" is then based on the way SLAC sets things up.

6. Annual Costs

Costs each year based on what needs to be bought to expand capacity, plus what needs to be replaced (when 3 years old). Personnel and maintenance not included! (They avoided getting into that discussion as they felt they could not have converged otherwise).

7. Regional Center valuations

Based on disk and CPU purchased, times unit costs in #5.

8. Cost Sharing

The algorithm below is used. I gather it originated with Guy Wormser. The numbers quoted below are for FY2001 just to show how it works; the example is straightforward because there was only one remote center at that time, contributed by IN2P3.

1. calculate total computing need = \$7.6M
2. subtract what is needed to handle 30fb^{-1} (this is done because 30fb^{-1} was the original promise from SLAC)
3. the result is the "sharable need" = \$2.67M
4. evaluate the remote center contribution = \$955k from IN2P3
5. this contribution is then divided into two equal parts that are applied differently:
6. one half reduces the sharable need still to be found (which is therefore \$2.19M)
7. the other half directly is applied to reduce the IN2P3 assessment
8. apportion the \$2.19M that is still to be found among all nations
9. calculate contributions by fraction (IN2P3's reduced as in #7; this leaves a net \$257k credit for IN2P3)
10. finally, add the computing to the Operations Common Fund (OCF) numbers for each nation to get a net contribution. (This results in \$40k being assessed to IN2P3)

9. Scrutiny

A Computing Steering committee, consisting of Physicists plus Computer Center directors from BaBar, verifies the spreadsheet. The BaBar finance committee (IFC) approves it. The IFC separately looks at operations costs (OCF).