

Exceptional Computing Committee (JECC) meeting

22nd October 2008

Meeting Minutes

version 2

Presents :

B. D'Ettore (INFN), S. Katsanevas (CNRS, by telephone)

D. Boutigny (CC-IN2P3 Director, by telephone)

L. dell'Agnello (CNAF)

J. Colas (EGO)

F. Fidecaro (Virgo/EGO)

Stefano Cortese (EGO Head of Computing Department)

G. Guidi (VIRGO Data Analysis Coordinator)

C. Palomba (VIRGO representative)

Topics:

1. Virgo use of computing center resources and forecast (G. Guidi)
2. Cost estimates (Computing centers)
3. Financing

Note : further information about cost of computing were given after the meeting. They are integrated within the minutes. The conclusions reached during the meeting are untouched.

1– Virgo use of computing center resources and forecast (G. Guidi)

Based upon LIGO and Virgo plans for scientific runs and upon the physics analysis performed by the Virgo collaboration, G. Guidi explained the present use of resources in Bologna and Lyon and the Virgo needs for 2009 and beyond (see note VIR-088A-08 attached to these minutes).

The first phase of data analysis requires fast access to raw data which thus have to be on disk; the long term repository is on tape (two copies, in Bologna and in Lyon). The hardware used to store data differs in the two centers. The tape storage system will evolve soon in CNAF. In Lyon, the mass-storage system (HPSS) comes with a front-end cache based on disks; in the following this cache is loosely referred to as disk storage to ease a common presentation of the two centers.

The reduced stream of data (h_{rec}) is used for the GW searches and it runs on four interferometers (3 LIGO and Virgo). Data quality, vetoes, and follow-up studies need access to raw data.

Use of CNAF should notably increase before the end of the year; nevertheless the total amount of CPU used will be less than expected at the beginning of the year. The reverse applies for CC-IN2P3 where Virgo is planning to compute twice as much as anticipated.

The request, summarized in the table below, calls for a sizeable increase in the use of computing resources both in CNAF and in CC-IN2P3. The forecasts for 2010 and 2011 cannot be considered as reliable as for the year 2009.

2– Cost estimates

During the meeting, the following cost estimates are taken based upon information by the computing centers:

cost of CPU : CNAF 0,48 €KSI2K.days in 2008 CC-IN2P3 : 0,25 €KSI2K.days
CNAF 0,27 €KSI2K.days in 2009

Note : These costs are lower than the estimate given at the previous JECC meeting (0,53€KSI2K.days see minutes of 16/4/08 JECC meeting)

For CC-IN2P3 the cost of CPU is assumed identical for 2009. For 2010 and 2011, for both centers, a 30% yearly reduction is assumed (see JECC 16/4/08).

cost of disk storage : CNAF 1.25 €GB CC-IN2P3 : 1 €GB in 2008
0.5 €GB in 2009

Note : The cost reduction in CC-IN2P3 for 2009 is linked to a change in technology.

In the following summary table no further cost reduction is assumed in the following years

cost for tape storage : CNAF : 0.1 €GB CC-IN2P3 : 0,17 €GB in 2008
0.074 €GB in 2009

Note : In the following summary table no cost reduction is assumed in the following years

B. D'Ettorre mentions that, as done for other experiments, CNAF should only charge for the cost of hardware and not for personnel, infrastructure or other running costs. Thus, for CPU, rather than giving a cost in unit of energy, he prefers to give a cost for computing power i.e. €KSI2K.

After the meeting, B. D'Ettorre gives the following costs for CNAF (VAT included):

	2008	2009	2010	2011
cost of CPU (€KSI2K)	0.16	0.09	0.07	0.054
cost of disk (€GB)	1.25	0.96	0.71	0.53
cost for tape (€GB)	0.17	0.10	0.088	0.063

As for disk, only the incremental (or replacement) computing power (new CPUs to be bought) has to be charged.

CC-IN2P3 charges the cost of hardware, based upon results of the latest tenders and the cost of electricity (~40% of the total cost). CC-IN2P3 prefers to charge for computing energy units i.e. KSI2K.days. Prices given during the meeting are confirmed.

The following summary table is thus compiled (corrected after the meeting to take into account the new figures for CNAF):

Virgo off-line computing needs	2008		2009		2010		2011	
Data production (Virgo+LSC) (TB)	12		190		192		300	
Bologna (Incremental)		cost (K€)		cost (K€)		cost (K€)		cost (K€)
CPU (KSI2K.days)	60000	20	390000	81	390000	3	390000	7
Disk storage (TB)	12	15	82	79	156	111	169	90
tape storage (Castor/HPSS) (TB)	55	9	108	11	36	3	131	8
Total cost Bologna		44		171		117		104
Lyon (Incremental)								
CPU (KSI2K.days)	240000	60	545000	136	545000	95	545000	67
Disk storage TB	125	125	44	22	156	78	169	85
tape storage (HPSS) TB	12	2	190	14	192	14	300	22
Total cost Lyon		187		172		187		173
Total cost 2 centers		231		343		304		277

Note : For CC-IN2P3, only numbers for 2008 and 2009 were given during the meeting; numbers for following years have been extrapolated as explained above. In CNAF in years 2010 and later, only old CPUs will have to be replaced.

3– Contribution to the cost of the Virgo off-line computing from the EGO budget

The above cost estimates, called for the following comments:

- Rather than an explosion of the Virgo computing cost, we should speak of a phase transition between a state with very little data to analyze and the phase of a running experiment.
- Nevertheless, Virgo should try to optimize its computing strategy to minimize costs. This might apply for instance to the use of disk storage which, in both centers, is much more expensive than storage to tape.
- Unit cost for computing and storage may decrease faster than assumed here.

After discussions it is agreed:

- EGO will set aside in its 2009 budget, 200k€ for Virgo off-line computing with, if the budget allows, an additional 100k€ as contingency. A priori, as the use of the computing centers is balanced, the sum would be shared about equally between them.
- for 2008, only 120k€ have been reserved for offline computing in the EGO budget. This amount will be shared between the two computing centers according to the use of resources.