Joint EGO-Virgo Computing Committee (JECC) meeting

The Council is invited to take note of the minutes of the JECC meeting which took place at EGO on the 21st of October 2010.

Meeting Minutes

Were present:
D. Boutigny (CC-IN2P3 Director - phone)
R. Lemrani (CC-IN2P3 - phone)
M. Morandin (CNAF- Director)
L. dell'Agnello (CNAF)
M.A. Bizouard (VIRGO Data Analysis Coordinator)
F. Ricci (VIRGO representative)
J. Colas (EGO - Director)
A. Bozzi (EGO - Head of Computing Department)

Topics:

1. Virgo use of computing center resources in 2010 and 2011 forecast (M.A. Bizoard)
2. Questions/remarks to/from the Computing Centers
3. 2010 EGO contribution to the off-line computing cost
4. Budget request for 2011
5. Mid-term future

N.B. Documents presented during this meeting are available at:
https://workarea.ego-gw.it/ego2/ego/computing/jecc/jecc-meeting-21-oct-2010/

2- Virgo use of computing center resources in 2010 and 2011 forecast (M.A. Bizoard)
The presentation is available at: https://tds.ego-gw.it/ql/?c=7857.
M.A. explains the data taking plan for 2010 with first a science run VSR3 in coincidence with LIGO and then data taking in "Astrowatch mode" VA3. In 2011 Virgo plans is to collect science data during the VSR4 run, possibly 6 months long. The amount of resources planned to be used till end 2010 and the forecast for 2011 are given. Space for storage on disk is more than adequate (more than 100TB free in each CCs), except for user space at Bologna. More detailed information is available in the "Virgo computing needs" document (https://tds.ego-gw.it/ql/?c=7819)

3- Questions/remarks to/from the Computing Centers

- Lyon. CC-IN2P3 will investigate on the few tape failures reported by M.A.
- CNAF is planning to start moving soon all data from CASTOR to the new GEMSS tape system. This system, currently successfully used by CMS, should allow access to data in a way similar to what is provided by the HPSS system in Lyon.
  Then VSR2 and VSR3 data will be copied. Most of the copy should be over by end 2010. This operation should be transparent to users.
- If the new GEMSS system satisfies the needs, Virgo storage requests to the two CCs should become similar.

- M.A. motivates the low level of computing in 2010 respect to what has been requested a year ago: for technical reasons not yet solved, some pipelines were not able to run on the CCs infrastructure and instead LIGO clusters were used; in addition, some searches have been delayed by a lack of manpower in some groups that prevented to run in parallel 2 searches.
- Both centers will have no problems providing the resources asked for by Virgo in 2011.
- To better follow the evolution of Virgo needs, F. Ricci suggests having two JECC meetings per year: one in May-June and a second one in October.
- Concerning data transfer, CNAF underlines that Virgo is the only user of "bbftppro". For better reliability and ease of maintenance, it would be wise to move to the CNAF standard tool "GRIDftp". The use of other GRID tools like lcg-utils is considered, but the Virgo collaboration and EGO have not arrived to a final decision. EGO thinks that the general Virgo context (Virgo should stop taking science data mid 2011 till sometimes in 2015) pushes to postpone the transfer tools upgrade to the Advanced Virgo era.

4- 2010 EGO contribution to the off-line computing cost

The unit cost put forward by each CCs is given in table 1.

- Cost for CPU is 23% lower at CNAF compared to Lyon (assuming that a CPU runs for 3 years one unit of HSE06 will produce 3x365=1095 HSE06.day).
- Cost for disk storage is 56% more expensive at CNAF compared to Lyon (taking into account the difference between TiB and TB).

The above figures lead to a large imbalance between the two CCs for the EGO contribution to the computing cost.

CNAF has invested in CPUs according to the Virgo computing request which turns out much higher than the actual Virgo use. On the other hand the CPU power installed in 2009 covers more than enough the 2010 needs; this excess of power may have been used by other experiments. After discussion, it is agreed that CNAF will not charge EGO for CPUs in 2010.

While still taking into account the investment cost, CNAF will study the possibility to charge every year the used computing energy (HSE06.day) instead of the increase of installed power. It turns out that Virgo requests for CPU are smaller than in previous years.

The disk storage cost drives the cost imbalance between the two CCs. To the light of the significant amount of free space on disks and of the large imbalance of available disk space between the two CCs, EGO asks if it would be possible to return some of the unused disk space at CNAF or at the minimum to postpone part of the payment to the following year. After discussion, it is agreed that for 2010, EGO will only pay for 98 TB of increased disk space. The remaining 90 TB will be paid in 2011 and a part of this space (25 TB) will be used to increase as soon as possible the user's space at CNAF. EGO also asks CNAF to study the possibility to return in 2011 some disk space in case Virgo would not need it.

Soon CNAF will move/copy data to tape. EGO proposes to contribute to the cost for tapes.

To sum up, the 2010 EGO contribution to the off-line computing cost will be:
CNAF : 87 K€ CC-IN2P3 : 78 k€

Based on requests in table 1, for 2011, a ~100k€ budget has to be foreseen.

5- Mid-term future

For 2012 one may assume requests for CPU power similar to 2011 request and a very limited need for data storage. As analysis progresses (2013-2014), the need to access raw data will diminish and only a much reduced data set will have to stay on disk (~5-10%).

CNAF and Lyon will interact with each other to define a cost model able to face the maintenance issue.

To be ready for Advanced Virgo, new investments will have to be decided in 2014.

	2010		Installed resources	2011		Installed resources	2012		Installed resources
Bologna		cost (KE)			cost (KE)			cost (KE)	
CPU (used energy in HS06.day)	280000			434000			450000		
CPU installed power(status is in HS06)		0	1240		0	1240		0	1240
Disk storage increment (TB-N)	98	69	368	90	64	458		0	458
Tape storage increment (TB-N)	180	18	325	200	20	525		0	525
Total Bologna cost		**87**			**84**			**0**	
Lyon									
CPU use (HS06.days)	150000	4		114000	2		150000	0	
Disk storage increment (TiB)	124	62	308	0	0	308		0	308
Tape storage increment (TiB)	140	12	427	200	17	627		0	627
Total Lyon cost		**78**			**19**			**0**	
Total cost 2 centers		**165**			**103**			**0**	

Unit cost	2010		2011		2012	
	Bologna	**Lyon**	**Bologna**	**Lyon**	**Bologna**	**Lyon**
CPU cost (€/HS06)	23.6					
CPU cost (€/HS06.day)		0.028		0.018		
disk cost (k€/(TB-TiB)	0.708	0.500	0.708	0.300		
Castor/HPSS cost (k€/(TB-TiB)	0.100	0.085	0.100	0.085		

EGO contribution to the cost (k€)	2010			2011			2012		
	Bologna	**Lyon**	**Total**	**Bologna**	**Lyon**	**Total**	**Bologna**	**Lyon**	**Total**
	87	78	**165**						

Table 1: 2010 used resources and 2011 forecast. Storage cost is computed on the increment. Lyon CPU cost is based on the used computing energy. The column "Installed resources" gives the total amount of resources available taking into account previous year availability. For Bologna in 2010 there is already 458TB installed, even if 90TB of these will only be paid in 2011.