

JECC meeting EGO, 2013 Dec., 4th

Pia Astone for the Virgo collaboration

VIR-0505A-13

- Some considerations on the AdV CM, in view of next years organization, for which the work is ongoing now
- Next year computing needs

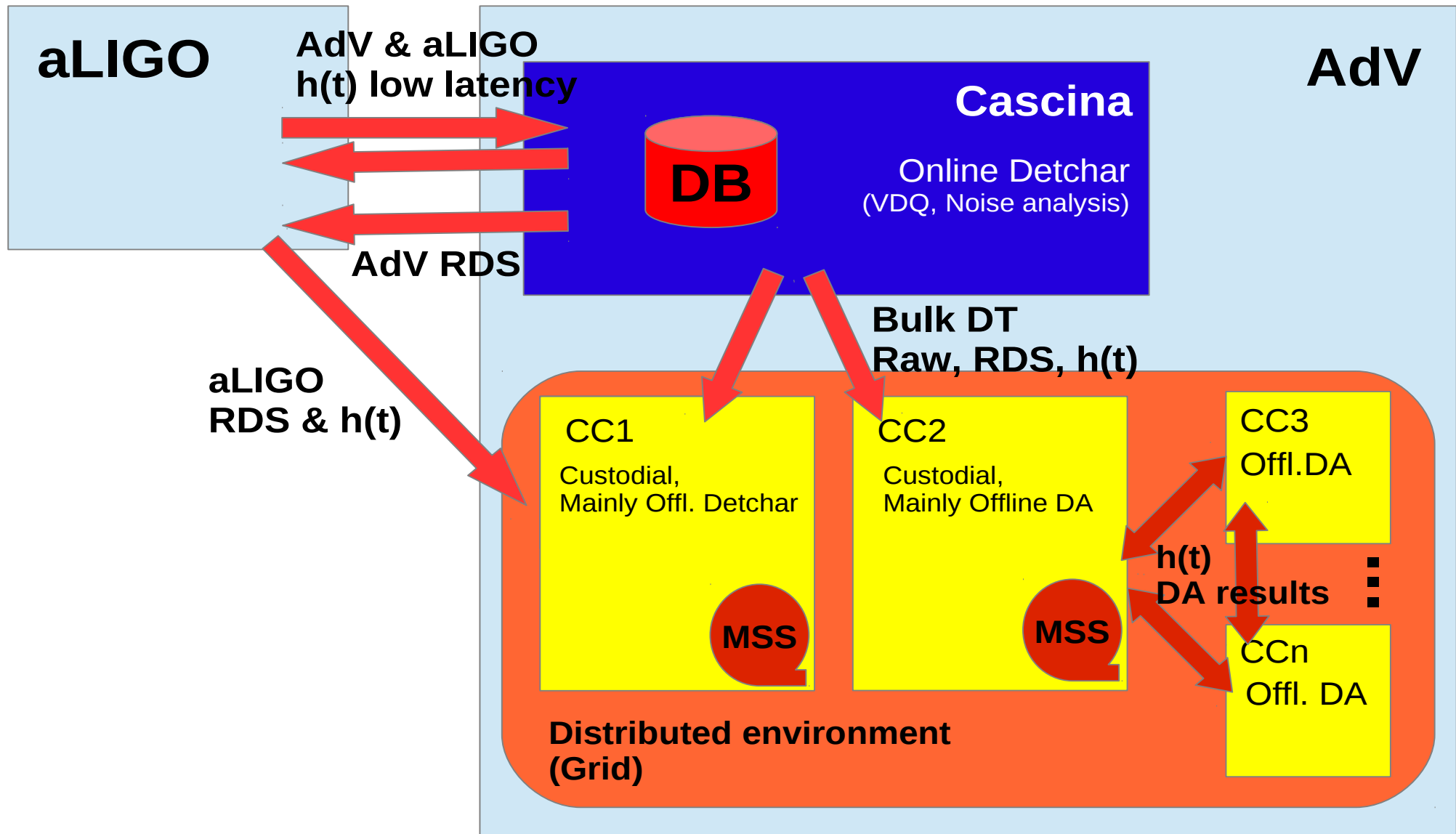
The AdV CM : VIR-0129E-13

- I. Computing and DA (“Data analysis”) Workflows.
- II. Data Model.
- III. Data management, distribution and access.
- IV. Software description and management. This section contains details on milestones and responsibilities for each project.
- V. Computing Facilities resource requirements (Storage and CPU foreseen needs for next years)

Model for data production and distribution

- ◆ Cascina (EGO) hosts the Tier-0
- ◆ The instrument “Primary data” are distributed to Tier-1s: CNAF and CCIN2P3 (one full copy in each CC), maximum latency of 1 day
- ◆ CNAF (indicated as CC2 in next slide), CCIN2P3 (CC1) and LIGO clusters are the main places where the offline science analyses have been done.

Data workflow for data analysis and detector characterization activities



Comments on the workflows


- Virgo's approach to the CM was to maintain the existing (heterogeneous) shared resources, but their interconnection has really never been satisfying
- Our goal is to organize a layer between the user and the computing infrastructures which will hide the different underlying technologies. But this work will require additional man-power and skills.
- To face the huge computational demands of g.w. searches in ADE, there will be the need to gather the resources of many CCs into a homogeneous distributed environment (either Grids and/or Clouds) and to adapt the science pipelines to run under such distributed environment.
- A big jump towards the unification of LSC and Virgo resources and towards a more effective usage would be to set up Grid interfaces on all Virgo sites and enable LSC compliant submission entry points on a subset of them, if needed

Data	CNAF [TB]	CCIN2P3 [TB]
Raw data	745	745
AdV RDS	11	11
LIGO RDS	22	22
Trend data	1.5	1.5
Minute trend data	0.25	0.25
AdV h(t) and status flags	3	3
MDC h(t)	9	9
Calibration output	1	1
Omieron triggers	-	4
DQ veto production data	-	negligible
Spectrogram data	-	1
MonitoringWeb data	-	0.8
DQ developments data	-	0.5
DQ segment	negligible	negligible
NoEMi data	12	-
BURST	15	3
CBC	4.5	0.5
CW	25	-
STOCHASTIC	-	3.6
Total	849.5	802.8

Storage @Ccs (tapes and cache disks)

850 TB/yr CNAF
803 TB/yr CCIN2P3

 Transferred from Cascina

 Produced at the CCs

Need to better organize the usage of tapes, to reduce the data access times

10

Data Transfer. Part III of the CM

- 1) AdV data are transferred from Cascina to CCs (2-4 TB/day)
- 2) aLIGO RDS and $h(t)$ data are transferred from one aLIGO cluster to CCs (tests ongoing). 60 GB/day
- 3) AdV $h(t)$ are transferred from Cascina to aLIGO and aLIGO $h(t)$ are transferred from aLIGO to Cascina following different rules, defined to guarantee the low-latency workflows for these searches (few seconds !).

These are ~ 8 GB/day from Cascina to aLIGO;
 ~ 16 GB/day from aLIGO to Cascina

Data Access (local and remote) @ CCs

- Among the data in the table, the ones for which we have the need of parallel fast Data Access are a subset of those transferred from Cascina or from aLIGO clusters, which have dimensions of the order of (1-1.6) GB each.
- We will use DataBases, for metadata information (LVDB) and for file bookkeeping (lfc ? See next slide). Need to understand scalability and synchronization issues.
- We have plans to organize a File Locator Service to extend/replace the lfc functionalities
- We wonder if an alternative for remote DA might be possible with “Storage Federations”, at least for some use case: e.g. only scientific analysis using the reconstructed g.w. channel (3 TB/year/detector)

lcg tools and lfc

- There are ongoing discussions on the future of LCG tools and LFC;
- We use LCG tools for data transfer and access and we have started to use LFC.
- Thus we wonder if:
 - lcg commands will have an analogous in gfal
 - Will lfc be maintained ?
 - Are other experiments interested in a common file catalog in GRID ?
 - What about DIRAC , a tool which should manage workloads and data ? It has been suggested to us by CNAF and by CCIN2P3. We have started some workload management tests.

Pipeline	SVN Vers Vers.	Responsible and Collaborators	Status of the project (July 2013)	Goal (Jan. 2015)
Scientific pipelines	TBD	TBD	TBD	TBD
Others, e.g. commissioning, detchar	TBD	TBD	TBD	TBD

*...too many
To Be Defined
here.*

Table 5.29: Summary Table for Data management tools, having divided the item into two main sub-parts, as explained in the text, to be able to describe detailed milestones for each of them. TBD stands for “to be defined”.

Data management project: we need to define Responsibilities and reasonable milestones for the project. The Computing Coordinator, Gergely Debreczeni, is now working to the Implementation Plan

The Data Management project will be unitary and will guarantee a transparent access to the data, but here we have divided the work into its two main sub-parts, as they will have different milestones.

One regards the access to the data by the scientific pipelines, where the data to be accessed are not a huge quantity, $O(10 \text{ TB})$, there is no need for graphical monitoring tools and the access to the data is through tools like GRID/CLOUD, that is not interactive.

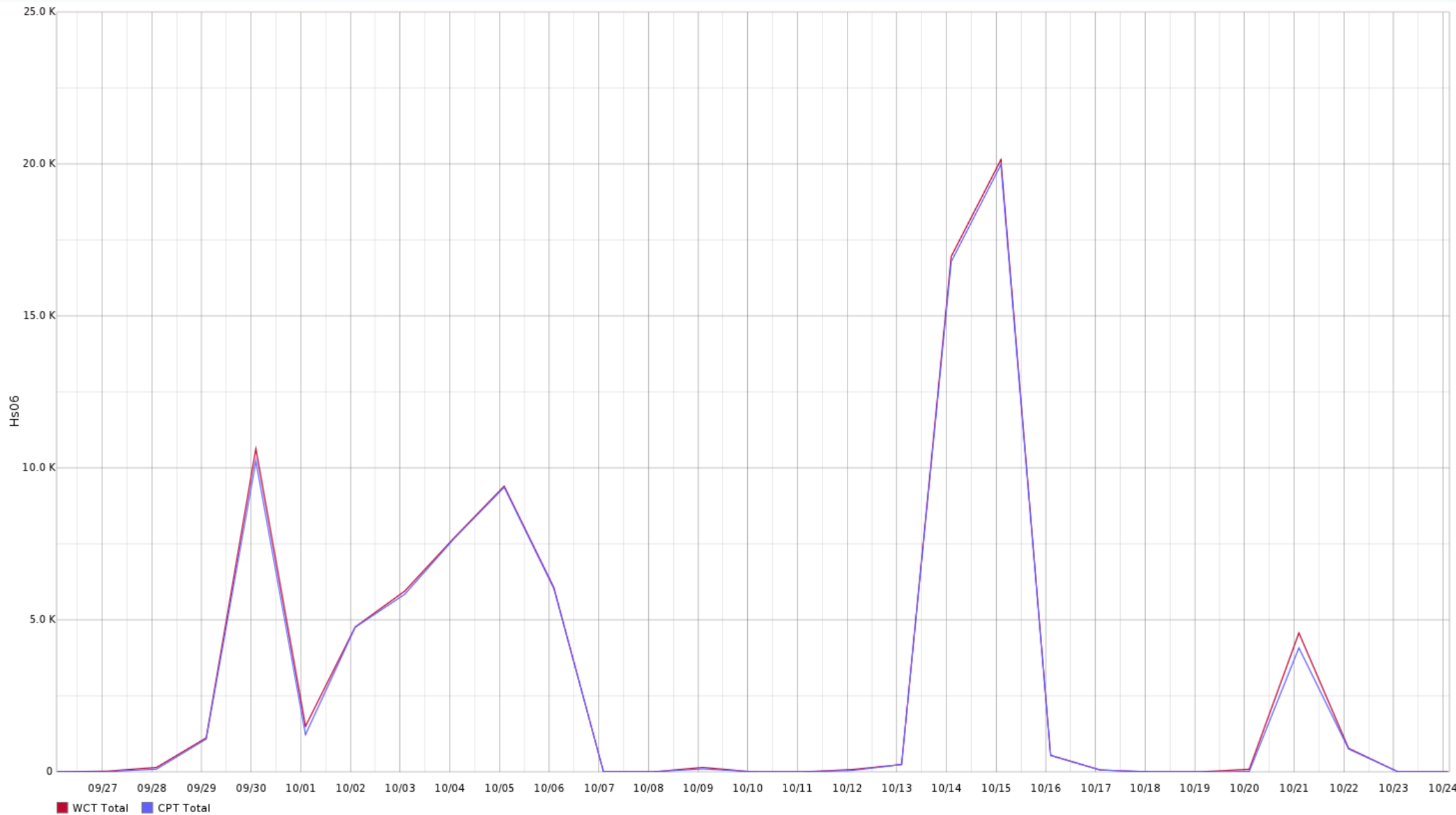
The second regards all the other activities, e.g. commissioning, detector characterization, where graphics tools are very important and the data accessed are typically huge $O(100 \text{ TB})$.

A comment on Resources needed in 2014 at CNAF

- An important effort has been made in the last year to solve some architecture constraints which were preventing us to run our Parameter Estimation pipeline (CBC) on our CC. This is a highly demanding pipeline, mainly developed by Virgo people, so the principle of entirely rely on on Architecture LSC constrained is clearly wrong. The very original request for next year was to double the actual computing power at CNAF, thus reaching roughly 750 cores but then, as the porting work has progressed and after having received some detailed evaluations of needs from the Virgo CBC group, we have changed the request, to reach a minimum of 1000 cores.
- This is to guarantee the needed benchmarks for the CBC parameter estimation work (especially the computationally expensive pipeline for testing GR), but also to allow us to continue running the pending CW searches and some other tests on low-latency CBC pipelines and data pre-conditioning work needed for the Burst All-Sky searches

Tests for CBC are now running at CNAF

(we have got additional CPUs to run these tests up to the end of this year)



Part V: Computing facility resource requirements.

External CCs

- At regime, the storage needed for 1 year of data (including raw) will be $<\sim 1\text{PB}$;
- Computing needs, at regime:

Pipeline needs in kHS06 power	local	GRID/CLOUD
Detchar Data Quality	1	–
Detchar Noise studies	1	1 ?
BURST	negl	3
CBC	–	33+
CW	–	60+
STOCHASTIC	negl.	negl.
TOTAL	2 ?	97+

Units here are kHS06(power).

The energy kHS06.day needed for 1 yr is $\text{power} \times 365$

Table 3: Summary Table: Estimation of the computing needed locally in the CCs and under GRID/CLOUD at a regime situation (2018+), under certain hypotheses on the parameter space covered. Units are power in kHS06. The “+” indicates that this is the minimal request, with more resources we could cover a wider parameter space

Considerations on next years computing needs

- In the next years, from 2014 to 2017, a fraction of the computing power will be needed to complete the analysis of the Virgo data and to do tests in preparation for the Advanced detector expected sensitivity
- We have considered in the CM only the needs of the most demanding pipelines and they are clearly our best estimations as of today. In particular, while it is clear what will be needed to carry out CW searches over a given parameter space, the CBC needs will strongly depend on the number of the triggers found.
- A new aLIGO scenario for computing (XSEDE based) might lead to some increase to our request: it seems now important also to run on our clusters the main CBC pipeline (“ihope”) and the burst All-Sky, All times pipeline (“cWB”).
- New, unexpected results might clearly vary the scenario and hence the computing needs.
- The results of the ongoing work with GWTOOLS and GPUs could significantly affect the requests: thus this work is urgent.

In conclusion, the AdV CM needs to:

- guarantee adequate storage and computing resources at Cascina, for commissioning, detector characterization and low-latency searches;
- guarantee fast communications between Virgo applications at Cascina and aLIGO CCs/other detectors for low-latency searches;
- guarantee reliable storage and computing resources for off-line analyses in the AdV CCs;
- push towards the use of geographically distributed resources (Grid/Cloud), whenever appropriate;
- push towards a homogeneous model for data distribution, bookkeeping and access.

Status of the Storage in external CCs

Table 4: Storage at CNAF since 2009.

Year CNAF	gpfs4 [TB] used / available Virgo	gpfs3 [TB] used / available Virgo	Castor or GEMSS [TB]	Castor disk [TB] used / available all exp.
2009	190 / 256	9 / 16	145 (Castor)	(+)
2010 (Oct. 1)	261 / (256+186)=442	16 / 16	163 (Castor)	17 / 36
2011	345 / 384	26 / 32	750 (**)	0
2012 (Oct. 29)	325 / 368	33 / 48	826	0
2013 (Nov. 18)	254 / 379	47 / 48*	826	0

Table 4: Storage at CNAF since 2009. (+) means that we don't know the exact number. In 2011 data from Castor have migrated to GEMSS, which uses gpfs_virgo4 as cache disk. (*): 19 TB of the 30 TB we have asked for 2013 are going to be assigned and thus the available quota on gpfs3 will be 67 TB. (**) by the end of 2011, ~ 80 TB more data were stored in GEMSS, but it was decided that these tapes cost would be reported to 2012 budget.

CPUs used in external CCs

year	CNAF (WCT) [kHSE06.day]	CCIN2P3 (WCT/CPU) [kHSE06.day]
2007	60	+ / 91
2008	240	+ / 740
2009	453	+ / 388
2010	162	+ / 130
2011	674	+ / 142
2012	669	103 / 80
2013 (Nov. 20 th)	850	115 / 63♣

Table 7: Evolution since 2007 of the CPU used at the CCs. + means that wall clock time numbers are not known for all years before 2012. ♣ 2013 CCIN2P3 numbers are derived from end of october consumption extrapolated to 12 months.

Units here are kHS06.day

Requests for next year

- No additional storage is needed for next year at CNAF;
- +5 TB on the user 's disk space at CCIN2P3 and
- +10 GB on the disk that hosts the software at CCIN2P3

CPU: energy in kHSE06.day

Group	CNAF [kHSE06.day]	CCIN2P3 [kHSE06.day]
Burst	0	300
CBC/CW	3600	0
SGWB	0	0
detchar	O(360) *	100
TOTAL	3600	400

Table 8: Computing needs for 2014 in kHSE06.day. the request at CNAF amounts to a number of cores O(1000) assigned to Virgo for the whole year. (*) = not added to the total as they will be taken from the cores principally assigned to CBC/CW

At CNAF: the request in the table is the total energy we will need. Which means to add to our pledge the number of cores needed to reach a total energy¹⁸ of 3600 kHSE06.day (O(1000) cores for 1 year)