

SCIENTIFIC COMPUTING AND TRAINING FOR THE VIRGO EXPERIMENT: users point of view

Pia Astone, Cristiano Palomba, Michele Punturo
INFN, Virgo collaboration

Naples, 2018/03/27

WLCG/HSF Workshop

- P. Astone is the former (2012-2014) Virgo computing coordinator. Main author of the "Advanced Virgo Computing Model".
- C. Palomba is the contact person of the Virgo collaboration towards CNAF
- M. Punturo is the actual Virgo computing coordinator.

MOTIVATIONS IN TRAINING ACTIVITIES-I

- Virgo is a gravitational wave (GW) experiment, mainly funded by INFN (Italy) and CNRs (France).
- We work as part of the much larger LIGO/Virgo (LV) community. LIGO is mainly funded by the NSF (National Science Foundation), US.

We think it is important to spread as much as possible the knowledge about Virgo-related software/computing. Organizing training activities for our students/PhDs/new and also old members will be crucial.

1. Most of us have developed computing/programming skills specific to Virgo software/computing by self-learning or thanks to interactions with more expert colleagues
2. Lack of systematic and organic training activities for the specific issues of Virgo.

Very partial exceptions:

- A one-day training on developing Virgo-specific application on the GRID was carried on ~10 years ago (prehistory) and never repeated
- Python courses have been delivered at the Virgo site (mainly on interferometer controls)
- ASTERICS-OBELICS (Observatory E-environments LINKed by common ChallengeS) schools. One the project leaders is a Virgo member (Eric Chassande Mottin). Next school will be held next June at LAPP, Annecy. Topic: Python for project development in astroparticle physics.
- A two days course on GPUs on Nvidia will be organized in June, following what recently done in LIGO, and with a remote connection with an expert, who has experience with LIGO software, in Seattle

MOTIVATIONS IN TRAINING ACTIVITIES-II

1. A part of the Virgo collaboration uses software fully developed in Virgo and typically uses computing power available in Italy (e.g. CNAF, Bologna) , Netherlands (Nihkef CC) or France (e.g. CCIN2P3, Lyon);
2. LIGO, from this point of view, has a completely different organization. An important part of the software used in the analysis and of the computing power used by Virgo resides in LIGO clusters (e.g. Caltech, Hannover). The training related issues for this part of the collaboration are completely different and in the hands of LIGO. To our knowledge, Virgo members running LIGO-based software do not strongly feel the need to modify what they are doing.

Virgo is now in the European proposal for Scientific Cloud “ESCAPE” (call INFRA-EOSC-04-2018). In this frame we expect to be able to organize some training activities.

COMPUTING FIELDS WE ARE INTERESTED IN

1. Generally speaking, Virgo software is based on C, C++, Python, Matlab (often compiled, sometime using MEX files), ROOT (in few cases).
 2. New developments are typically driven by the interest of single Virgo groups.
 - E.g. in the Rome group we have recently started – through the work of a Physics master student - to do the porting of the computationally heaviest analysis pipeline we have to **GPUs** (work is in progress to run on a cluster of GPUs) using **TensorFlow**;
 - With another master student we have started to work with **Machine Learning** (ML) techniques for signal detection.
- **It's important to say that both porting to GPUs and ML are now considered of great interest in the whole LV collaboration. So schools/seminars/online tutorials with example of applications would receive interest from many, also US based.**
 - Lack of manpower to drive a strong training effort is a big problem in Virgo.

ACTIVITY STRATEGIES

- Some activities, mainly those more LIGO-oriented, already well documented via online wiki-pages.
- An example is here: <https://www.lsc-group.phys.uwm.edu/lscdatagrid/> , but then quite soon a LV password is needed.
- Also the LIGO software is typically well documented. See <https://losc.ligo.org/software/>
 - One of the reasons is the fact that this software is public—as part of the NSF strategy, quite different from what e.g. INFN typically requires. This, together with a much larger number of involved people (~ 1000 vs ~ 200 respectively in the LIGO and Virgo collaborations) has made a big difference.

- On the Virgo side we have wiki pages, sometime not updated and not covering everything.
- A more systematic effort in writing **well documented wiki** on the software but also on the specific modalities to run on our computing resources , e.g. at CNAF, Nihkef, etc. is needed
- **Dedicated schools** on relevant tasks like running analyses on the **GRID**, or on the use of submission framework like **Dirac**, or to learn how to optimize a code vs new architectures, like **GPUs**, probably would help in pushing new people to adopt these tools.
- We are started a test activity on Dirac in touch with CNAF and EOSC-Hub. We are thinking to a Dirac dedicated school driven by Dirac experts.
- A significant effort has been done by LIGO people and Virgo CC admins to run some of the LIGO Condor-based pipelines on our CC via **OSG**. This also must be documented and training sessions would be useful.
- Cascina (Virgo site) would be a perfect place for any training activity.

SUGGESTIONS

- We need an effort to produce reliable and complete wiki documentation on software and computing.
- We think, however, that lectures and face2face meeting with experts (internal at Virgo – if exist - or external), where we can pose our problem, is the perfect formula for training people.
- Virgo (and LIGO) have very specific data analysis problems. For us, the formula "across experiments/fields " does not always work. Of course, basic training on widely used tools/frameworks (e.g. Python, Dirac etc.) could be shared with other experiments.
- To conclude, the final answer is YES. We strongly feel the Virgo community would greatly benefit from a systematic training activity on several aspects. This would help to increase the impact of Virgo in the software and computing aspects related to the worldwide effort to study gravitational waves sources.