Contributions à la prospective IN2P3 – GT09

Eric Chassande-Mottin, Simone Mastrogiovanni

Computing models/Infrastructure needs

The future of astroparticle physics includes a range of large-data science experiments, such as CTA, LSST and the third generation of gravitational-wave (GW) observatories (Einstein Telescope in Europe). The amount of data storage and computing that those experiments will need is orders of magnitude larger than what is dealt with today. It is clear that a common international data management and computing infrastructure capable of supporting all of these would be mutually beneficial to the experiments, the computing facilities, and the funding agencies. The high-luminosity LHC experiments face similarly-daunting computational scaling problems over the same time frame. There are thus obvious synergies to exploit in this area.

The demand for data analysis computing for third-generation gravitational wave detectors (our focus here) will be driven by the high number of detections (up to hundreds per day compared to few per week today) and a expanded search parameter space (0.1 solar masses to 1000+ solar masses) due to the larger observational frequency bandwidth going down to few Hertz. GW signals will last hours to days in the detector (as opposed to seconds/minutes currently). These longer-duration signals will strain physical memory abilities and lead to a steep increase in the computing demand.

In a recent review [1], it is estimated that ~300 million CPU core-hour per hour are required to complete the science associated with the current GW detectors (second generation). More than 3 order of magnitude more CPU and RAM will be required for the third generation barring an algorithmic breakthrough in the way the analysis is performed. Assuming a conservative 10 %/year computing performance/cost improvement at constant cost during the next decade (following projections of WLCG), it is almost certain that the current approach will not scale.

This calls for significant changes in the computing infrastructure (going from a centralized system around few large to mid-scale computing centers to a more distributed system) and software infrastructure (optimization of the methods, codes and workflows; hardware portability and compatibility to heterogeneous platforms).

The review [1] provides recommendations for the funding agencies to anticipate the necessary future changing by encouraging more domain-science and computer-science interactions, and by creating discussion forums for exchange know-how between the HEP and astroparticle and GW communities that will face similar challenges.

[1] Bird, I *et al*, *Gravitational-Wave Data Analysis Computing Challenges in the 3G Era*, 3G subcommittee reports, 2019. https://gwic.ligo.org/3Gsubcomm

Contributions to other sections of the report

Artificial intelligence and Machine learning

As in many fields, machine learning and AI are likely to be transformative in astroparticle physics. This naturally includes gravitational wave astrophysics where a number of very promising results have been recently obtained in the area of noise background rejection and source parameter estimation. The COST Action CA17137 - A network for Gravitational Waves, Geophysics and Machine Learning that gathers GW and data scientists provides a good model to accelerate the dissemination of those techniques to the production pipelines. PhD co-tutorship is also particularly efficient. IN2P3 should consider to offer fellowship specifically for thesis co-tutored by two domain-science and data-science experts.

Big data

Open data would deserve a specific section in this chapter. The political context at the national [1] and european [2] level is rapidly evolving with major decision that will change the way we (IN2P3) do science in the next decade. Soon, there will be a legal obligation to publicly-funded research to release the data to the public.

Globally speaking, the current funding scheme of major experiments does include a provision for opening the data, while this activity includes many tasks such as data preparation, curation, documentation and release that have a significant cost and should be anticipated in the budget. The role of data curator is currently missing in the type of careers that can be considered at IN2P3.

[1] Plan national pour la science ouverte, 2018.
http://cache.media.enseignementsup-recherche.gouv.fr/file/Actus/67/2/PLAN_NATIONAL_SCIENCE_OUVERTE_978672.pdf

[2] Open science policy platform, 2019.
https://ec.europa.eu/research/openscience/index.cfm?pg=home

Accelerators

Gravitational-wave astronomy is one of the fields that can largely benefit from GPU computing. LIGO and Virgo have carefully investigated the use of parallel GPU and MIC architectures for its most compute-intensive searches; CUDA GPUs have been the most successful and cost-effective, and were deployed at scale for the first time in science run O3 (2019). GPU allow a faster generation of spectrograms and noise classification, and thus to reduce the latency to release gravitational-wave alerts. Long-term software development costs may be higher for rapidly-evolving parallel hardware platforms (GPU, MIC, AVX512, etc.) than for traditional CPUs, given that parallel programming interfaces are less stable targets. Single-threaded CPU codes have worked on new hardware with minimal modifications for 30+ years. Data analysis on more distributed, non-dedicated computing grid/cloud platforms will require ongoing computing infrastructure investment.