IN2P3 Scientific Council - Computing and Data Processing

Report on Low-Latency Alerts & Data Analysis

Julien PELOTON (IJCLab)

June 2022

IN2P3 has a unique position by playing a leading role in several major experiments in time-domain astronomy, and especially delivering low-latency alerts. Robust systems have been developed over the years to harness scientific potential, and the CC-IN2P3 computing centre provides the skills and resources to ensure scientific return for the community. But in the rapidly growing multi-messenger astronomy era, the key for success in joining information from different probes is to connect different communities with efficient frameworks. While the efforts have been mostly concentrated on a per-experiment basis, the challenge is now to coordinate between experiments, on data analysis as well as on the use of shared infrastructures.
In this report we first describe the real-time analysis for a set of IN2P3 experiments, and their current computation needs. We then turn our focus on what needs to be strengthened for the coming years: transverse initiatives & interoperability of tools, as well as shared infrastructures & deployment of R&D projects.

Context

The discovery of a gravitational wave signal from a binary neutron star system (GW170817) in coincidence with a short gamma-ray burst (GRB170817A) and followed by an optical kilonova (AT 2017gfo) showed that time domain and multi-messenger astronomy will largely define astrophysics in the next decade. Since then, significant efforts are being put into developing follow-up systems for turning survey alerts into data to enable scientific results. Experiments routinely survey the sky to detect transient objects as early as possible in order to identify them and obtain data before they disappear, and understand the nature of the sources and the physical processes involved. Often, follow-up observations are triggered to complement initial observations. Among several facilities in operations, let us cite the Zwicky Transient Facility, a wide-field sky astronomical survey observing in visible and infrared wavelengths, the *Fermi* and Swift satellites, performing gamma-ray/X-ray astronomy observations, or the LIGO/Virgo/Kagra observatories, designed to detect gravitational waves. These experiments produce a continuous stream of alerts originating from varying sources in the sky, that is usually publicly advertised and distributed directly to scientists. The latter then need to extract scientific information upon receival for their specific use.

This decade will see the advent of more powerful instruments being able to detect sources in the whole electromagnetic spectrum but also through high energy neutrinos, very high energy cosmic rays and gravitational waves: the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST) in the visible range [1], KM3NeT neutrino detectors [2], the extension of LIGO/Virgo and KAGRA for gravitational waves [3], the Chinese-French Space Variable Objects Monitor multi-wavelength observatory (SVOM) [4], the Cherenkov Telescope Array (CTA) [5] and many others.

To fully exploit the available data, the main challenge is the structuration of communities beyond individual experiments. For instance, inter-project coordination is necessary to constrain the source's position in particular when the resolution is very large, with sky error regions spanning hundreds of square degrees for gravitational wave instruments. In addition, the coordination must be done in a timely manner due to the transient nature of various phenomena, with time frames spanning from seconds (low-latency) to weeks. Asides from the spatial and time constraints, all communities are facing challenges to correlate the very large volume of alerts from different telescopes and detectors (real-time and post-processing), extract and classify the sources, and eventually trigger a repointing of one or several telescopes or detectors of the network. Hence the key for success for joining information of different probes with low latency is to connect them with efficient frameworks (communication and processing) developed and shared across different communities. IN2P3 with its long tradition of collaborations and large projects offer an ideal place for such a cross-community effort.

Highlight on selected IN2P3 experiments

In the following, we briefly highlight the IN2P3 contribution to some of the main experiments working in time-domain astronomy.

LIGO/Virgo-Kagra – *Gravitational waves; 2002-; ground-based (around the globe)*

Data analysis for gravitational wave detectors, including low-latency analyses, is performed as a joint Gravitational-Wave Observatory Network (IGWN) effort[1]. The geographical separation of the detectors and the short timescale (in the order of 20 seconds from signal detection to alert generation) imply the creation of a common distributed cyber infrastructure, guaranteeing adequate storage and computing resources for detector characterization, low-latency searches and alerts generation. Regarding alert infrastructure, the consortium operates multiple on-line detection pipelines that upload candidate events to a database (GraceDB) if they have a false alarm rate (far) of less than 1/hour.
To minimise delay in the analysis, low-latency storage and computing are mainly provided by observatory computing centres. Search pipelines run on dedicated or highly-prioritised resources in an HTCondor-managed resource pool. The GraceDB production instance is currently deployed in High Availability (HA) on AWS. But to exploit support from computing centre staff and achieve cloud provider independence, an alternative HA deployment via Kubernetes is being tested on the INFN CNAF cloud for the alert generation components (GraceDB, GWCelery, LVAlert). Regarding the low-latency data distribution protocol, Apache Kafka is becoming the standard (as is done in the Rubin Observatory). Kafka is a modern high throughput stream processing software, with built-in redundancy, replication, and resilience to hardware failure.
Among the dozen pipelines reducing data from time series to generate alerts, two are led by IN2P3 scientists: MBTA (LAPP, IP2I, IPHC) and pycbc[2] (IJCLab). MBTA is currently deployed at CC-IN2P3, but the low-latency part is run at EGO to avoid delays. Pycbc is deployed at Caltech on the LIGO cluster, and another instance is being deployed at VirtualData, Université Paris-Saclay.

KM3NeT – *Neutrino; 2022-; ground-based (Mediterranean Sea)*

The computing for the Cubic Kilometre Neutrino Telescope, KM3NeT[3], is divided into Tiers. In Tier0, real-time analysis is performed in two control stations in France and Italy on small computing farms to filter the data and write to disk. The analysis pipelines consist in reconstructing events with multi-stage algorithms based on likelihood (topologies, direction, energy, etc.), and classifying their natures (background noise, atmospheric neutrinos, cosmic neutrinos, trace/shower, etc.) with machine learning algorithms. Once these events have been reconstructed and classified, several cosmic source search pipelines are applied (clustering, correlation with external alerts, sending neutrino alerts, etc.). The CPPM is in charge of the coordination of the online analyses of KM3NeT and of the implementation for the ORCA telescope (Oscillation Research with Cosmics in the Abyss), as well as responsible for the neutrino alerts, using previous expertise from the multi-messenger program of ANTARES [6]. Every night, data is then streamed from the control stations to the CC-IN2P3 computing centre for offline analysis (Tier1).

[1] https://igwn.readthedocs.io/
[2] https://github.com/gwastro/pycbc
[3] https://www.km3net.org/

Rubin – *Visible-NIR; 2024-; ground-based (Chile)*

In this landscape, the Legacy Survey of Space and Time (LSST), using the Vera C. Rubin Observatory[4], aims to survey the southern sky deeper and faster than any wide-field survey to date. During its 10 years of operations, LSST will drive the volume of alerts by enabling the discovery of an unprecedented large number of astrophysical transients, opening a new era of optical big data in astronomy. The alert system is expected to produce about 10 million alerts per night (1 TB/night). After a selection process that ended mid-2021, Rubin Data Management has retained only a few of the community broker proposals to analyse the full set of LSST alerts. Among the seven broker projects selected, one is an IN2P3 initiative: Fink (lead IJCLab & LPC Clermont). Fink (https://fink-broker.org) is a broker framework being developed to ingest the alerts from LSST and redistribute filtered alerts with added values, for studying the transient sky as a whole, from solar system objects to galactic and extragalactic science. The project includes CNRS/IN2P3 and European experts in solar system, supernovae, micro-lensing, and multi-messenger astronomy. It also houses a large number of R&D in new technologies, with a special focus on developing and applying novel cloud, big data, streaming and machine learning techniques to accommodate the paradigm change introduced by the multi-TB alert data set of LSST (Apache Spark, Kafka, HBase, Kubernetes, pytorch, tensorflow, …).
More than 30 researchers and engineers from several IN2P3 laboratories and associates co-authored the white paper [7]. A prototype has been developed and deployed at the cloud@VirtualData since 2019 on an OpenStack cloud (analysing data from the precursor ZTF), and the production service for Rubin is being deployed at CC-IN2P3, which also runs an OpenStack cloud, and will have a local copy of LSST data that can be efficiently exploited for internal cross-match needs.

CTA – *Gamma; 2020s-; ground-based (La Palma & Chile)*

The real-time CTA[5] data analyses are assured by the Science Alert Generation (SAG) system. This system provides low-latency data reconstruction, data quality monitoring, science monitoring and real-time alert issuing, with a latency of about 20 seconds. The LAPP has a large part of the leadership in the low-latency systems. This is described in more detail in the contribution of D. Boutigny (3. CTA Real Time Analysis) in this IN2P3 Scientific Council session.

SVOM – *Visible-X-Gamma; 2023-; space & ground-based*

The SVOM[6] observation program will be divided in three parts: the Core Program (~30-40 GRBs/yr with prompt emission over 3 decades), the General Program where SVOM will be an open observatory and observations will be awarded by a Time Allocation Committee, and the Target of Opportunity (ToO) program where alerts will be sent from the ground to the satellite (initially one ToO per day focussed on time domain astrophysics including multi-messengers). For the Core Program, the real-time scientific products (under the supervision of the Burst

[4] https://www.lsst.org/
[5] https://www.cta-observatory.org/
[6] https://www.svom.eu/en/

Advocates) will be made public as soon as they are available (similar to Swift or *Fermi*-GBM). All the other scientific products are public six months after the data production. Concerning the General Program (GP), there will be semester calls for proposals (in association with a SVOM Co-I). All the SVOM data will be distributed to the Responsible Co-I (one year of proprietary period before all the scientific products become public). The ToO Program data policy is still under discussion. SVOM will also use the Fink broker developed for the Rubin Observatory which has the capacity to filter large volumes of alerts. Fink will be used to trigger SVOM's ground segment to enrich promising candidates from Rubin/SVOM with additional data and eventually trigger a ToO for the satellite.

SVOM is relying on the services provided by the VirtualData cloud (Université Paris-Saclay), both for the ground-based and the space-based components. In particular, the cloud is used as the integration platform for the ground-based component software, for the development of the alert messaging system and for running the emulator of the on-board processor that will be used by the MXT instrument on the satellite. The CC-IN2P3 will be used as a production centre when the satellite is launched.

CMB-S4 – *Millimetre; 2020s-; ground-based (Chile & South Pole)*

CMB-S4 will provide a unique platform to conduct a wide-field time-domain survey in the millimetre wavelength band, covering over half of the sky every day. In this waveband, the time-variable sky is largely unexplored, with only shallow surveys or those limited in area or scope. Based on the experience from the already established transient sky community, discussions have started at IN2P3 to develop tools to enable low-latency alert science.

Transverse initiatives & interoperability of tools

In this section, we discuss several existing initiatives working towards structuration of communities beyond individual experiments. The list is not exhaustive, but their efforts should be seen as drivers or examples for future actions or contributions within IN2P3.

Network of telescopes: GRANDMA example

GRANDMA is a world-wide network of telescopes with the primary scientific goal of discovering and characterising electromagnetic counterparts of gravitational waves [8]. It brings together an heterogeneous set of already-existing telescopes that operate in a coordinated fashion as a single observatory. Within the network, there are wide-field imagers that observe large areas of the sky, looking for optical counterparts; there are also narrow field of view instruments that do targeted searches of a predefined list of possible host galaxies, and larger telescopes that are devoted to the characterisation and follow-up of the identified counterparts.

The GRANDMA network is managed through a central system that distributes alerts and the desired observing sequences to each observatory in a coordinated way, to maximise the coverage. The collaboration provides software for data reduction and analysis that allow observatories to identify counterparts and report them in real time to the system. Further

follow-up observations can be then performed in a way that optimises the use of resources and avoids duplications. The network also includes a collaboration with amateur astronomers which can also receive alerts and report their observations, as was done for example with the Fink broker during the Kilonova campaign [9].
GRANDMA was created in April 2018 by APC/IJCLab and Observatoire de la Côte d'Azur (PI. S. Antier, Artémis & Co-PI. A. Klotz, IRAP). As of 2022, it gathers more than 80 scientists from 30 institutes and groups worldwide (18 countries), operating about 30 observatories.

Interoperable tools: Virtual Observatory example

Interoperability is possible thanks to the definition and adoption of standards which set the common language and technology between services and tools. Ideally, each observatory should provide standardised, high quality, compliant data products allowing the multi-wavelength & messenger identification of a source. But this is not always the case.
The Virtual Observatory[7] (VO) is a framework for astronomical datasets, tools, and services to work together in a seamless way. The International Virtual Observatory Alliance is a science-driven organisation that builds the technical standards, offering a place for discussing and sharing VO ideas and technology to enable science. There are clear benefits in using VO standards [10]: multi-wavelength & messenger approach, planning follow-up observations and coordination, easy visualisation & navigation through the data that leads to capability to discover and merge multiple data sets. In summary, there is a growth in the scientific return of data. But there are also challenges: how to add new projects coming up, especially PB scale missions? How to support science platforms with analysis close to data? How to support new data-types driven by growth in size and complexity of data sets? IN2P3 could have a role to play here: supporting meetings between technical and scientific communities to tackle specific questions, enabling training for interoperability, and providing resources to enable these developments within each experiment.

Scaling alert processing: Fink example

The daunting volume of alerts sent by LSST is one of the major differences with respect to previous experiments, making traditional streaming and processing tools inoperative. The broker Fink is an initiative to overcome this barrier, and to enable low-latency analyses at scale (millions of transients per night). The processing is based on the cluster computing framework, Apache Spark[8], which has proven successful in many contexts to efficiently analyse large amounts of data. To operate efficiently, the broker operates in two stages: the first part provides fast, scalable, fault-tolerant solution to decode the stream, and archive it regardless of the underlying volume of data, while the second part provides fast downstream access to the alert database for Artificial Intelligence applications, astrophysics software programs, and other services [7].
In order to ensure that the broker will continue to take advantage of new technologies and theoretical developments, it is imperative to maintain a high level of collaboration and

[7] https://ivoa.net/
[8] https://spark.apache.org/

coordination between the different scientific communities involved in the study of the transient sky. Given the variety of detectable astronomical events, the strategy of Fink is to study the transient sky as a whole, from solar system objects to galactic and extragalactic science. As a result, Fink includes experts from more than a dozen scientific domains that could not do otherwise under the constraint of data volume, and this has already led to tight collaborations with the SVOM mission and the GRANDMA network, and the broker is expanding to incorporate more actors within the project. Furthermore, the infrastructure heavily relies on widely used services and tools, to benefit from standards in place and ensure interoperability with the different multi-messenger actors. On the one hand, Fink users (individuals, telescopes, collaborations…) receive alerts tailored to their specific needs thanks to a series of community-built filters and a tight integration with services developed and maintained at the Centre de Données astronomique de Strasbourg, including for instance the latest information from crossmatched surveys. On the other hand, interoperability between alerts generated by Fink managed surveys and others is ensured by Fink by using standard protocols largely used by the community such as VOEvents and VOEvent transport protocol.

Online data analysis: Multi-Messenger Online Analysis example

Scientific publications conventionally report “static” results (e.g. table of objects with their properties at the time of publication). While this was justified when publications were published on paper, this rigidity limits the potential of their re-use or updates for a large variety of analyses. In addition, individuals cannot master all possible telescopes at once and, contrary to the high-level data products, raw telescope data may not be directly publicly exposed. One would rather need a system that helps multi-messenger analysis and extracts analysis results in an automatic way.

In this context, the Multi-Messenger Online Data Analysis[9] (MMODA) was created (lead A. Neronov, APC). Catalogues, images, spectra, and lightcurves are adjustable to the needs of specific multi-instrument and multi-messenger tasks using an intuitive web interface. The platform uses a mixture of HPC and cloud resources. The French part is deployed on OpenStack infrastructure provided by France Grilles. Cloud computing opens a possibility to provide telescope data analysis as a service. Services of MMODA are deployed in containers (Docker, Singularity), easy and fast to create and remove, and the containers are managed by Kubernetes, a container orchestrator system.

Unification of interfaces: ASTRO-Colibri example

Detecting events in two experiments is not a guarantee of getting the correlation between the two. For example, in 2017 it took 6 days to realize that there is a flaring blazar (TXS 0506+056) within the localisation uncertainty of a high-energy neutrino detected by Icecube. Many tools and platforms are available but there is a real need for automatisation and common interfaces. Astro-COLIBRI[10] [11] is an example of such an automatic pipeline providing: easy access to multi-wavelength and multi-messenger transient detections from various experiments, different

[9] https://www.astro.unige.ch/mmoda/
[10] https://astro-colibri.com

interfaces (web-based, Android, iOS), and a central API with publicly available endpoints. Astro-COLIBRI has been developed at IRFU/CEA Paris-Saclay and Ruhr-University (Bochum, Germany). It is relying on a modern cloud-based architecture enabling for example real-time smartphone notifications informing a large range of people (from burst advocates at large observatories to amateur astronomers) about newly detected transient sources.

Training, networking & community initiatives

In order to allow this unprecedented interconnection between different communities, it is important to foster interdisciplinary activities in all stages of scientific training. This includes: summer/winter schools detailing a range of astronomical probes and their theoretical backgrounds; joint gatherings tackling scientific aspects of current and future experiments and activities focusing on the construction of a common language across technical subjects (data management, reduction, storage and distribution; machine learning basics; new paradigm of programming techniques) and scientific applications. An example of this type of event, in the international level, which has received strong support from IN2P3 and the French community in the last few years is the Cosmostatistics Initiative[11].

The same strategy should also guide the development of long term activities of the permanent research staff. This includes specifically designed venues where social and professional connections can be established between researchers working on Science, Engineering, Technology and Maths related subjects - in academia as well as in the industry sector. Once these connections are established, there will be a natural avenue for the adaptation of state-of-the-art techniques to the many different fields within the IN2P3 community involved in multi-messenger astronomy.

At the highest level, inter-institute efforts are already in place such as the series of workshops Transient Sky 2020[12] organised by the Programme National Hautes Energies or more recently the Low-latency alerts and data analysis for Multi-messenger Astrophysics workshop[13]. For a few years, these events successfully gathered actors from different communities, and concrete collaborations and partnerships often emerged from the discussions. Additional events should be made focusing on technical issues related to interoperability and new technologies, similar to the ones organised by the Virtual Observatory, and via the *réseaux métiers* in place at IN2P3. But the existence of these efforts in the long term must be protected by official inter-institute agreements ensuring long-term logistic support and funding. Despite the undeniable effort necessary to put these structures in place, or sustain the continuation of the activities already under development, such initiatives hold the key to ensure the long term scientific impact of interdisciplinary activities developed within IN2P3.

[11] https://cosmostatistics-initiative.org
[12] See e.g. https://indico.in2p3.fr/event/19471/, https://forge.in2p3.fr/projects/ts2020
[13] https://indico.in2p3.fr/event/25290/

Shared infrastructures and the need for R&D

Many experiments rely on the computing expertise built over years within IN2P3, and the access to CC-IN2P3 is instrumental to deploy production-scale systems. However, we note that the support to develop R&D should be enforced. The techniques and frameworks evolve quickly, driven by the fast growing computing and big data industries. While our teams constantly survey the emerging new techniques, and try to adapt computation architectures to take advantage of recent developments (both software and hardware), the access to smaller and more flexible but yet reliable and powerful enough infrastructure is not always guaranteed. In this context, mesocentres, and local expertise in laboratories, should play a bigger role. As an example, Fink benefited for many years from the VirtualData cloud to emerge and to demonstrate capabilities to handle massive amounts of data in real-time from emerging technologies and cloud computing techniques. In return, the techniques developed in this context are generalizable to other projects, and the project is now being deployed at CC-IN2P3 for its full-scale production.

Finally, many systems from the low-latency alerts ecosystem are often competing for the same infrastructure grants, and efforts should be put in removing barriers between them when necessary, promoting shared solutions when possible (database, web front-end, cluster of containers, …).

References

[1] Bellm E., et al., 2019, LDM-612, Plans and Policies for LSST Alert Distribution, https://ls.st/ldm-612

[2] Adrían-Martínez, et al., 2016, Letter of intent for KM3NeT 2.0, Journal of Physics G Nuclear Physics

[3] Abbott B. P., et al., 2018, Living Reviews in Relativity, 21, 3

[4] Wei J., et al., 2016, arXiv e-prints, arXiv:1610.06892

[5] Actis M., et al., 2011, Experimental Astronomy, 32, 193

[6] M. Ageron et al, Astropart. Phys. 35 (2012) 530–536

[7] Möller, et al., 2021, MNRAS, 510, 3, 3272-3288

[8] S. Agayeva, et al., 2020, 6th Workshop on Robotic Autonomous Observatories, 198

[9] V. Aivazyan, et al., 2022, arXiv:2202.09766 [astro-ph.HE]

[10] A. Nebot, et al., PoS Asterics2019 (2020) 056

[11] P. Reichherzer et al., 2021 ApJS 256 5