IN2P3 Scientific Council - Computing and Data Processing

Report on Low-Latency Alerts & Data Analysis

Julien Peloton, IJCLab

Scientific context & technical challenges

The sky changes all the time

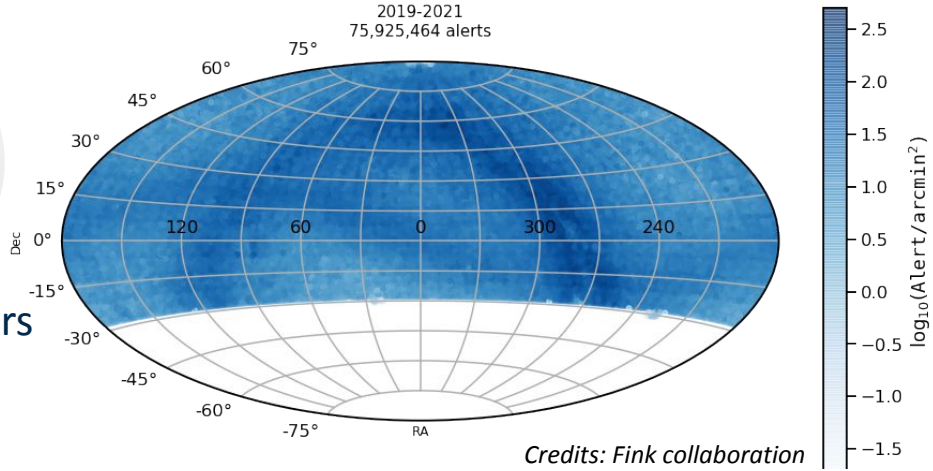
- Objects evolve (transients, variables)
- From Solar System to high redshifts
- Different timescales: from sub-second to years

2020+ new generation observatories

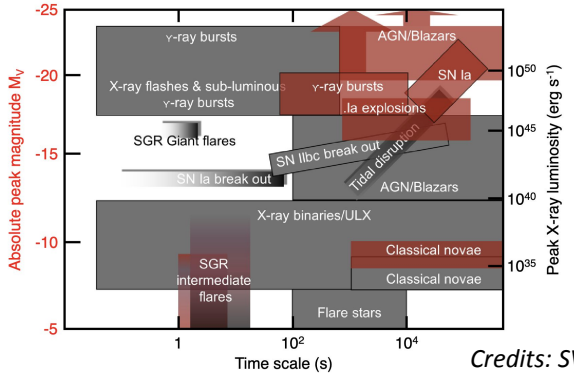
- Faster, wider, deeper
- Price: Communication & big data challenges

Each observatory will not characterise all of its events

- Additional observations will be necessary
- Often within a short time delay after initial discovery



Credits: Fink collaboration



Credits: SVOM

Developments at IN2P3: Highlights for 2022+

IN2P3 plays a leading role in several of the main experiments working in time-domain astronomy:

- LIGO/Virgo-Kagra – Gravitational waves; 2002-; ground-based (around the globe)
  - Reducing data from time series to generate alerts: MBTA (LAPP, IP2I, IPHC) and pycbc (IJCLab)
- KM3NeT – Neutrino; 2022-; ground-based (Mediterranean Sea)
  - Expertise from multi-messenger program of ANTARES (APC)
  - Tier0 (online): CPPM is in charge of the coordination of the online analyses (on-site)
  - Tier1 (offline): analysis performed at CC-IN2P3
- Vera Rubin Observatory – Visible-NIR; 2024-; ground-based (Chile)
  - Fink broker (lead: IJCLab & LPC), deployed at VirtualData & CC-IN2P3
- CTA – Gamma; 2020s-; ground-based (La Palma & Chile)
  - LAPP has a large part of the leadership in the low-latency systems (DL3, on-site)
- SVOM – Visible-X-Gamma; 2023-; space & ground-based
  - SVOM is relying on the services provided by the VirtualData cloud (IJCLab, APC)
  - CC-IN2P3 will be used as a production centre when the satellite is launched.
- CMB-S4 – Millimetre; 2020s-; ground-based (Chile & South Pole)

And the list goes on…

Working across different communities

Assuming internal challenges will be overcome, to fully exploit the available data, the remaining challenge is the structuration of communities beyond individual experiments.

Credits: S. Antier

GW170817 **Objective**

Interoperability of tools

Interoperability is possible thanks to the definition and adoption of standards which set the common language and technology between services and tools.

- The Virtual Observatory (VO)
  - Science-driven organisation that builds the technical standards
  - Framework for astronomical datasets, tools, and services to work together in a seamless way.
- VO standards widely adopted nowadays by major experiments (although not perfect!)
  - Multi-wavelength & messenger approach, planning for follow-up observations and coordination
  - Easy visualisation & navigation through multiple data sets
- But there are new 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?

Transverse initiatives at IN2P3 1/2

Scaling the alert processing: Fink (Rubin Observatory+)

- How to classify 10,000,000 alerts per night, every night? How to extract the few of interest for a broad range of science? How to make the system flexible enough to enable its evolution over a decade, while keeping all the information from the past available?
- Fink is the gateway for all experiments working with Rubin alert data, enabling real-time classification and data reduction on a 1TB stream each night. The object database will be 3PB after 10 years of operation.
- Interconnection of experiments through services (streaming, container, compute, web)

Network of telescopes: GRANDMA

- How to coordinate multiple telescopes to follow-up a target? How to optimise the coverage, and uniformise the data reduction?
- GRANDMA brings heterogeneous set of 30 already-existing telescopes that operate in a coordinated fashion as a single observatory.
- Unified web platform to report and plan observations.

Transverse initiatives at IN2P3 2/2

Unification of interfaces: ASTRO-Colibri

- We cannot rely (only) on serendipity… Many tools and platforms are available but there is a real need for automatisation and common interfaces to minimise communication delay.
- Astro-COLIBRI is an automatic pipeline providing: easy access to multi-wavelength and multi-messenger transient detections from various experiments, different interfaces (web-based, Android, iOS), and a central API with publicly available endpoints.

Online data analysis: Multi-Messenger Online Analysis

- How to move away from the “static” version that offer papers? How to deal with the complexity of multiple experiments?
- The Multi-Messenger Online Data Analysis (MMODA) is a web interface to explore and extract analysis results in an automatic way. The French part is deployed on OpenStack infrastructure provided by France Grilles.

Shared infrastructures and the need for R&D

Access to CC-IN2P3 is instrumental to deploy production-scale systems. But in a rapidly changing & growing environment, R&D projects are also crucial to design solutions.

- Access to smaller and more flexible but yet reliable and powerful enough infrastructure is not always guaranteed.
- Mesocentres, and local expertise in laboratories, should play a bigger role.
- Cloud computing opens a possibility to provide telescope data analysis as a service.

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.
- Promoting shared solutions when possible, e.g. database, web front-end, cluster of containers.

Training, networking & community initiatives

Larger interconnection between different communities

- Need to foster interdisciplinary activities
  - Data management, reduction, storage and distribution; machine learning basics; new paradigm of programming techniques
  - Example: Cosmostatistics Initiative (COIN)

At the highest level, inter-institute efforts should be organised.

- Some recent examples:
  - Transient Sky 2020 workshops organised by the Programme National Hautes Energies (2017-2021)
  - Low-latency alerts and data analysis for Multi-messenger Astrophysics workshop (2022)

Existence of these efforts in the long term must be protected by official inter-institute agreements ensuring long-term logistic support and funding.