



**Low-latency alerts & Data analysis  
for Multi-messenger Astrophysics Workshop**

13-14 January 2022, Paris

# Report from MMA Paris 2022 Workshop

MMAW @ EGO | 10 October 2022

- Some pickups from workshop reports
- Setting the stakes: LVK (IGWN)
- Moving stakes higher: Rubin
- Interoperability of tools: Virtual Observatory
- Unification of interfaces: ASTRO-Colibri
- Unification on Alerts: SCiMMA HOPSKOTCH
- Existing Transverse Initiatives: GRANDMA
- Multi-Messenger Online Analysis: MMODA
- Conclusions




# Some pickups from workshop reports

- Fermi-LAT: “**Common standards** for cross-detectors analysis. Long-term strategy for the release of these data to the broader community (similarly to e.g. g-ray, gravitational waves..)”
- Swift: “Coordination between different facilities and **rapid dissemination** of information is important”
- XMM-Newton: “Each observatory should provide **standardised**, high quality, virtual observatory compliant, data products allowing the multiwavelength/messenger identification of a Source”
- AGILE: “Coordination is mandatory. Enhance **cooperation with the community**, e.g. on sub-threshold events”



# Some pickups from workshop reports

- SNEWS: “Collaborating with: **SCIMMA (using Hopskotch)**”
- INTEGRAL/ISDC: “Let’s interoperate based on **standards and tools!**”
- ENGRAVE: “how to work efficiently and fairly together within and between large (and small) collaborations.[?]”
- Gammapy: “Multi-messenger analysis requires joint multiinstrument analysis. This requires **common and open high level data formats and tools.**”
- AMON: “**Updating to SCIMMA** cyber-infrastructure / Recently transitioned from PSU servers to **AWS servers**”
- ZTF: “**open source is** , **Constructive community feedback is essential for success**”

# Some pickups from workshop reports

- TNS: “The system resides on the **AWS cloud**, increasing its high-availability capacity and scalability.”
- ALeRCE: “Modular pipeline in **AWS**, using Docker **containers** and scaling with **Kubernetes**.”
- LVK: “GraceDB production instance currently deployed in High Availability (HA) on **AWS**. Alternative HA deployment via **Kubernetes** is being tested on INFN CNAF Cloud.”



# Setting the stakes: LVK (IGWN)

- Raw Data: 4TB/day
- GW data: 15GB/day/detector
- Prompt Data Products
  - Alerts (O4 expectations): 1 per day (latency in the few tens of seconds)
- The common distributed cyber infrastructure must guarantee:
  - Adequate storage and computing resources for detector characterization, low-latency searches and alerts generation: dedicated or highly-prioritised resources in an HTCondor-managed resource pool / **Cloud (AWS)**
  - Low-latency data distribution among the different observatories and computing clusters for low-latency searches: Apache **Kafka**
  - An ubiquitous and uniform running environment on dedicated resources and heterogeneous infrastructures: **CVMFS + Conda**



# Moving stakes higher: Rubin

- Raw Data: 20TB/night:
  - Sequential 30s images covering the entire visible sky every few days
- Prompt Data Products
  - Alerts: up to 10 million per night (60s via nightly alert streams to the community brokers)
- Community alert brokers will be vital to identifying EM counterparts

Rubin, with its multi-TB alert data set introduce a paradigm change and brokers are developing and applying novel cloud, big data, streaming and machine learning Techniques to deal with it.



# Interoperable tools: Virtual Observatory

- Framework for astronomical datasets, tools, services to work together in a seamless way.
- IVOA: A science driven organisation that builds the technical standards, a place for discussing and sharing VO ideas and technology to enable science and promoting and publicising the VO
- Clear benefits : Growth in the scientific return of data. Capability to discover and fuse multiple data sets. Application of the VO in planning new observations and observing strategies.

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





# Unification of user interfaces: ASTRO-Colibri

- 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.
- Relying on a modern cloud-based architecture enabling for example real-time smartphone notifications informing a large range of people



# Unification on Alerts: SCiMMA HOPSKOTCH

- HOPSKOTCH is a scalable, high-throughput low-latency platform for handling real-time data streams for MMA applications, which:
  - provides a cloud-based instance of **Kafka** data streams to support MMA applications via a publish-subscribe paradigm
  - allows for “at least once” delivery semantics
  - offers extensive identity and access management controls and fine-grained **Kafka** topic permission configuration to respect data right



# Existing Transverse Initiatives: GRANDMA

- GRANDMA is a world-wide network of telescopes with the primary scientific goal of discovering and characterising electromagnetic counterparts of gravitational waves.
- It brings together an heterogeneous set of already-existing telescopes that operate in a coordinated fashion as a single observatory

The collaboration provides the software for data reduction and analysis that allow 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 duplicities.



# Multi-Messenger Online Analysis: MMODA

- A system that helps multi messenger analysis and extracts data analysis results in automatic way
- Catalogues, images, spectra, and lightcurves are adjustable to the needs of specific multi-instrument and multi-messenger tasks using an intuitive web interface.
- Services of MMO are deployed in **containers** (Docker, Singularity), easy and fast to create and remove. The containers are managed by **Kubernetes** (k8s), container orchestrator system.



- Technological convergence over common multi-collaboration infrastructure is ongoing based on mainstream tools:
  - Computing resources in HA: Cloud deployment based on orchestrators (**Kubernetes**)
  - Low-latency data distribution: Apache **Kafka**
  - Running environment: **Conda+CVMFS, Containers**
- Definition and adoption of standards and unification trends are ongoing
- Capability to connect different communities greatly enhanced via the development of common (open source) software libraries
- Leveraging on the technological convergence and on the above trends some pilot projects are paving the way for the high stakes solutions we need



This is short (and biased) summary. Take a look to the full story at:

[https://indico.in2p3.fr/event/25290/attachments/66712/102484/MMA130122\\_contributions\\_report.pdf](https://indico.in2p3.fr/event/25290/attachments/66712/102484/MMA130122_contributions_report.pdf)

