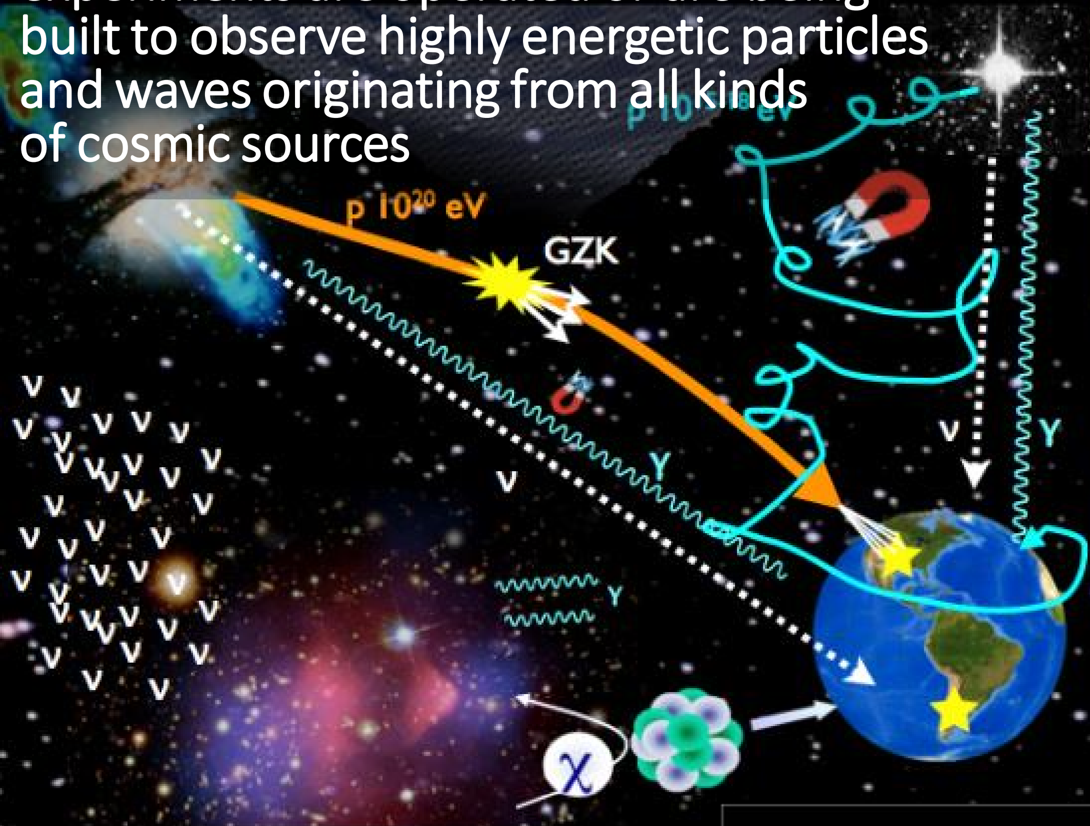
Infrastrutture di calcolo per la fisica astroparticellare nell’era delle osservazioni multi-messenger

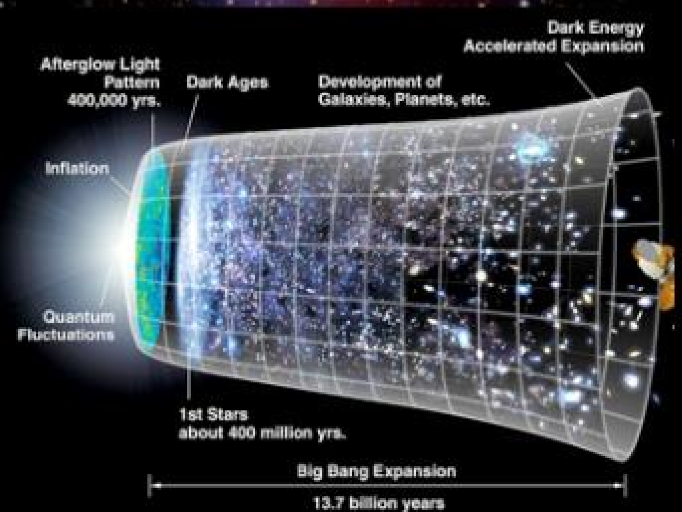
Gianluca Gemme

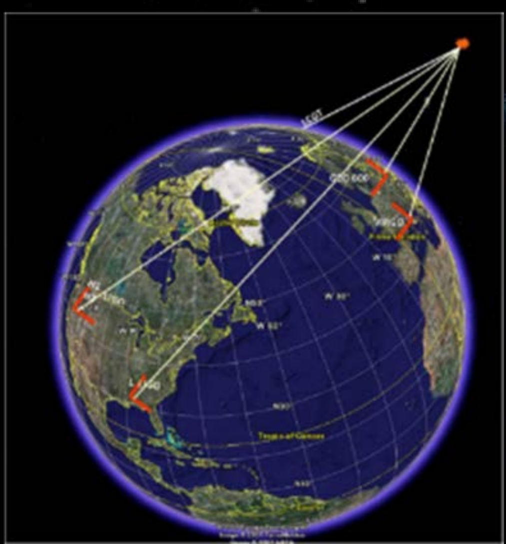
INFN - Genova

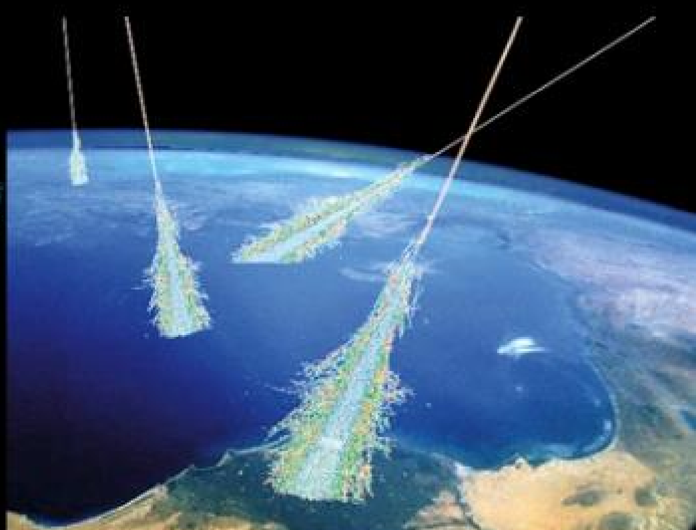
M. Branchesi, E. Brocato, E. Chassande-Mottin, G. Debreczeni,
D. D’Urso, G. Maron, M. Punturo, G. Stratta

Current and future astroparticle physics experiments are operated or are being built to observe highly energetic particles and waves originating from all kinds of cosmic sources



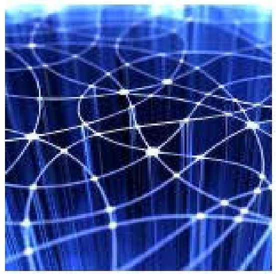







An European effort

TOWARDS A WHITE PAPER FOR COMPUTING & ASTROPARTICLE PHYSICS



A HUGE AMOUNT OF DATA TO COME

In a few years, Astroparticle Physics has grown from a field of a few charismatic pioneers, transgressing interdisciplinary frontiers to a global science activity with large infrastructures and collaborations each involving hundreds of researchers. Such large-scale projects and activities face challenging problems of data collection, data storage and data mining. For some, these computing costs will be a significant fraction of the cost of the infrastructure. The issues of computation, data mining complexity and public access are extremely challenging.

To foster coordination of efforts in these fields and estimate future requirements, three Workshops were organised in the context of ASPERA-2. The first workshop in Lyon, France, presented the computational challenges and contrasted them with the data storage and analysis models developed in neighbouring fields of particle physics (grid and cloud computing, large databases) and astrophysics (virtual observatories, public access). The second workshop in Barcelona, Spain, reviewed the current computing models developed by upcoming astroparticle observatories, including CTA, KM3NeT, Auger, VIRGO/LIGO, and LSST. The third workshop in Hannover, Germany, focused on hardware and technology.

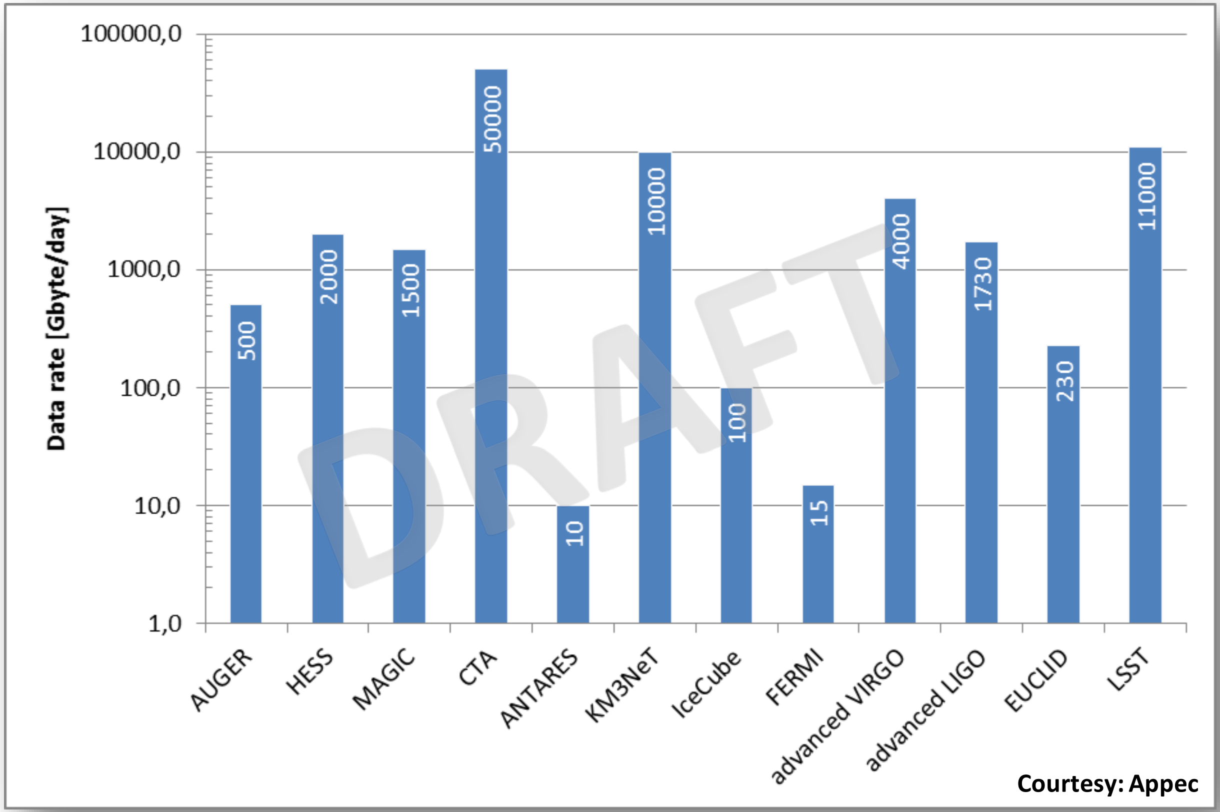
143 people in total participated in the three computing workshops. The discussions amongst them clearly demonstrated the high level of complexity of this topic and the need for further coordination between all stakeholders. Due to this complexity, a number of steps are being taken in the context of APPEC.

The functional centre that is responsible for coordinating this activity is DESY. Contact: APPEC Secretariat.

Data production

The data taken by the experiments are large and expected to grow significantly during the coming years

To cope with the substantially increasing data rates of astroparticle physics projects it is important to understand the future needs for computing resources in this field



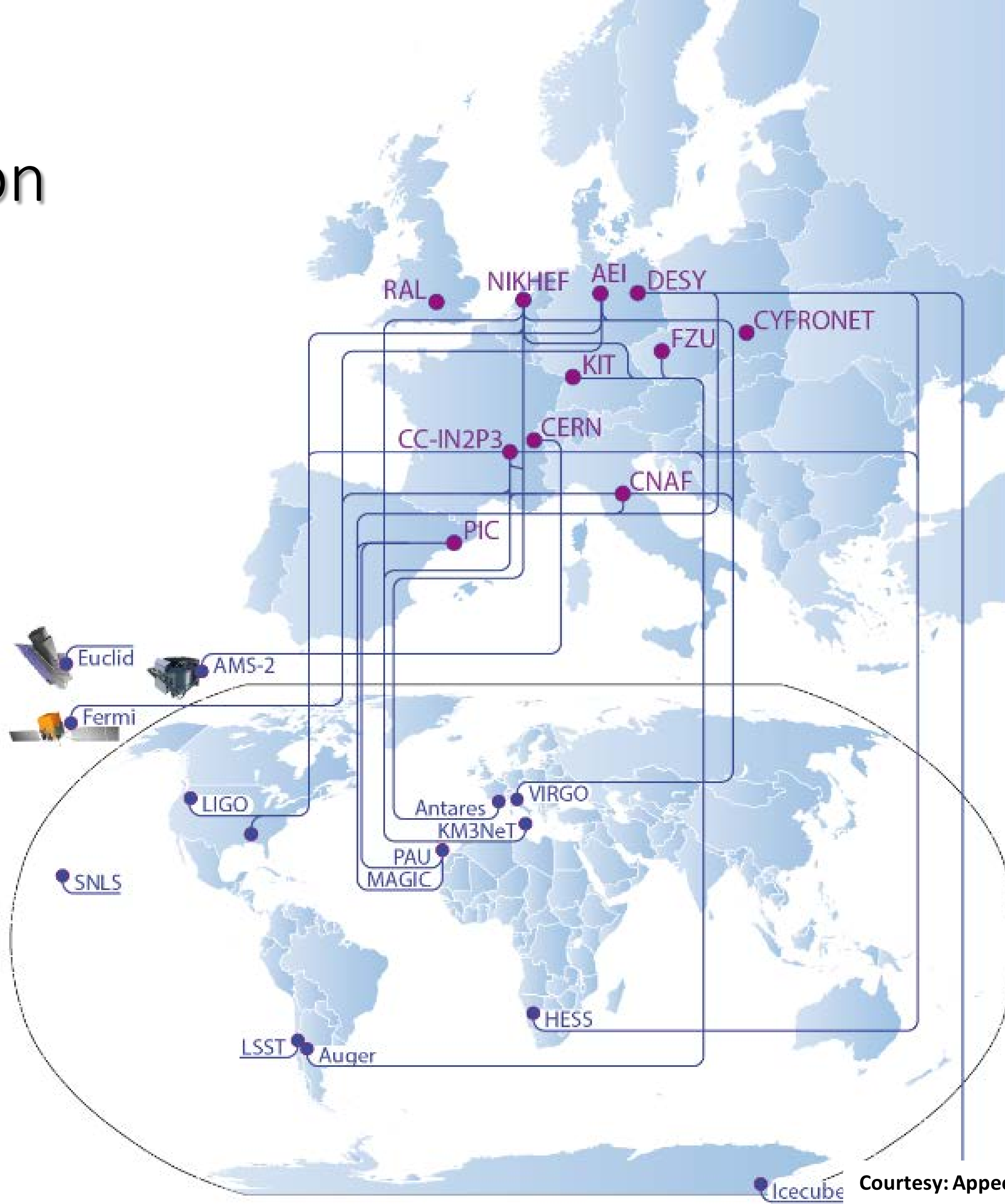
Providing these resources constitute a significant fraction of the overall running costs of future infrastructures

Data distribution

Astroparticle physics projects are often placed at remote locations often without direct access to high-speed internet

Each project is required to develop its own solution to transfer data to computing centers for processing

Data are processed in more than one cc, either all in Europe or in combination with centers in other regions

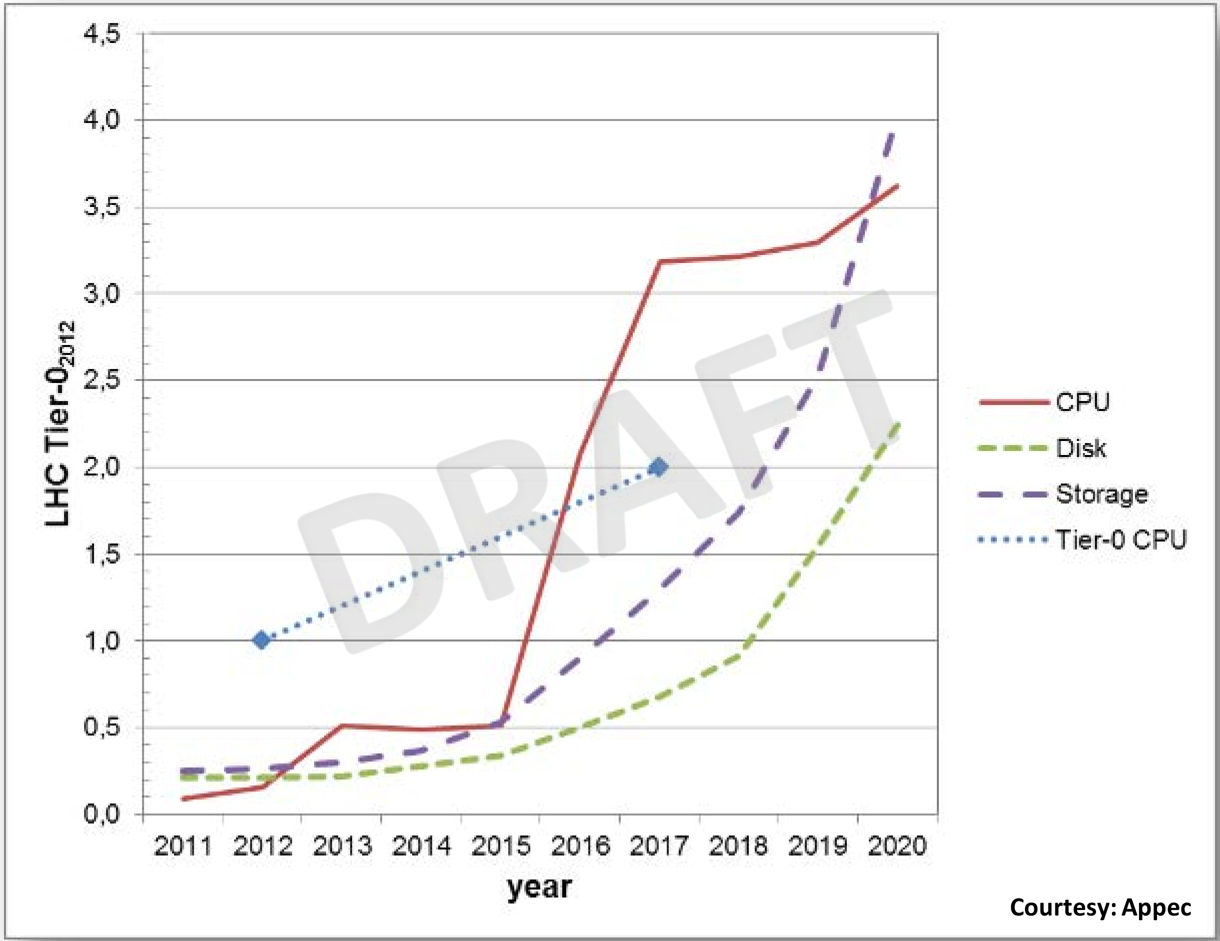


Courtesy: Appec

Data crunching

The start of the next generation of experiments there is an increasingly demand in computing power and storage space.

This requires a strong coordination between the experiment collaborations, data centers, and the funding agencies in the coming years.

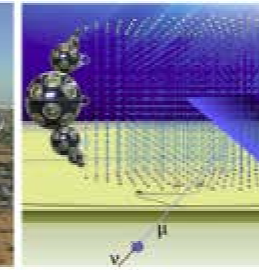


Providing these resources constitute a significant fraction of the overall running costs of future infrastructures

Astérics

Astronomy ESFRI & Research Infrastructure Cluster



Astérics :H2020 Cluster for the **interoperability** of ESFRI infrastructures

- ESFRI (European Strategy Forum on Research Infrastructures) includes **SKA, CTA, KM3NeT and E-ELT**
- The scope of Astérics extends to **Einstein Telescope** (under the umbrella of EGO) and EURO-VO (Virtual Observatory) and “pathfinders” EUCLID, LSST, **Virgo/EGO**, LOFAR, e-VLBI, HESS, MAGIC and ANTARES
- Coordinated by Mike Garrett (ASTRON)
- Interoperability = future multimessenger astronomy

Credit: E. Chassande-Mottin

Astérics

Astronomy ESFRI & Research Infrastructure Cluster

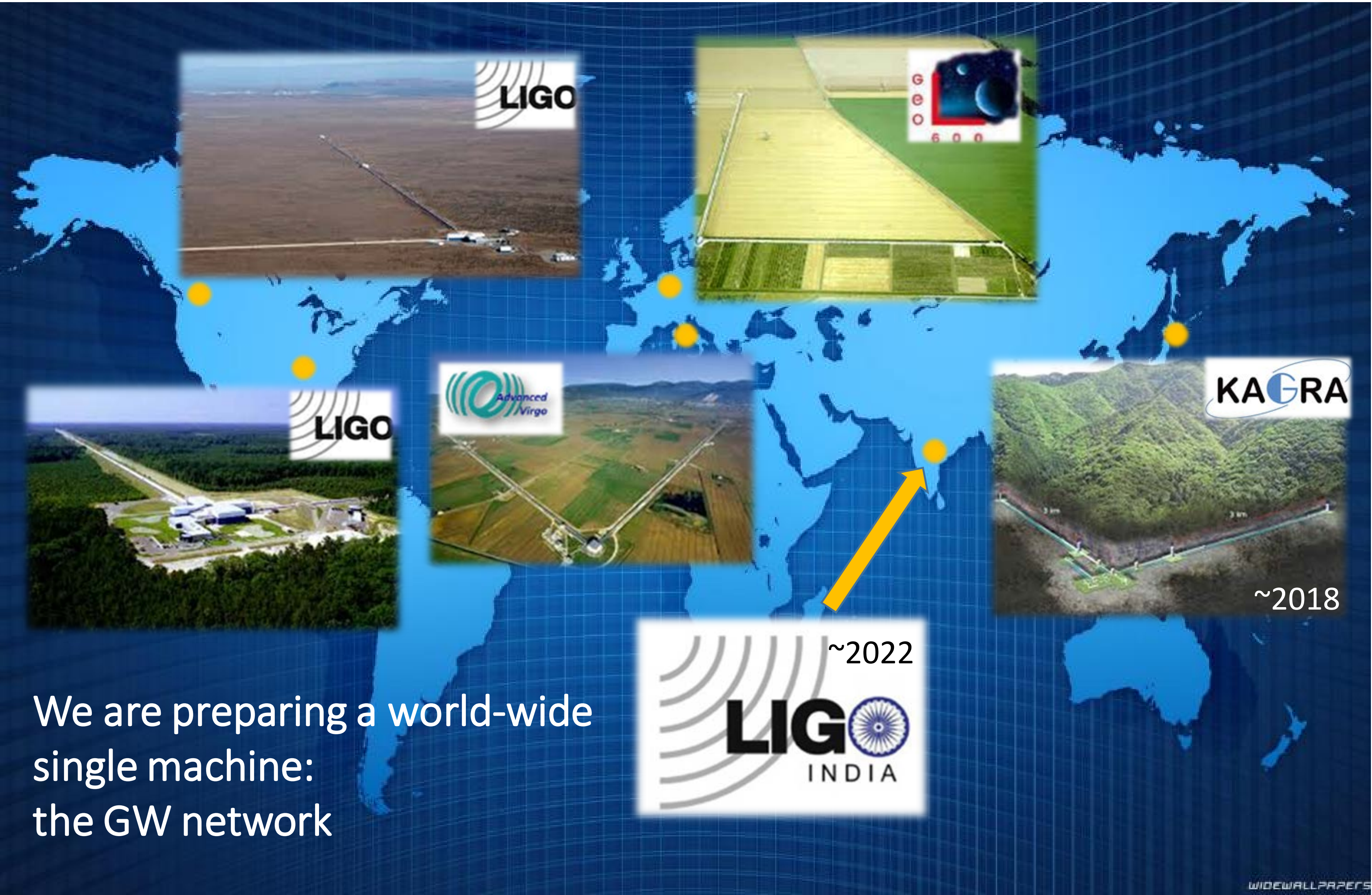
Astérics : H2020 Cluster for the interoperability of ESFRI infrastructures

- **Cross-matching** between observations from various observatories
- Connection with the **Virtual Observatory** (astronomical catalogs and surveys – transients and galaxy)
- Eric Chassande-Mottin participated to the definition of Astérics
- On-going **collaboration with CDS** (Centre de Données astronomiques), Strasbourg – Vizier and Simbad database
- Approved recently. Kick-off meeting, end of May.

Credit: E. Chassande-Mottin

Complexity

Advanced gw detector network 2015-2022



We are preparing a world-wide single machine:
the GW network

AdV computing model

G. Debreczeni, yesterday

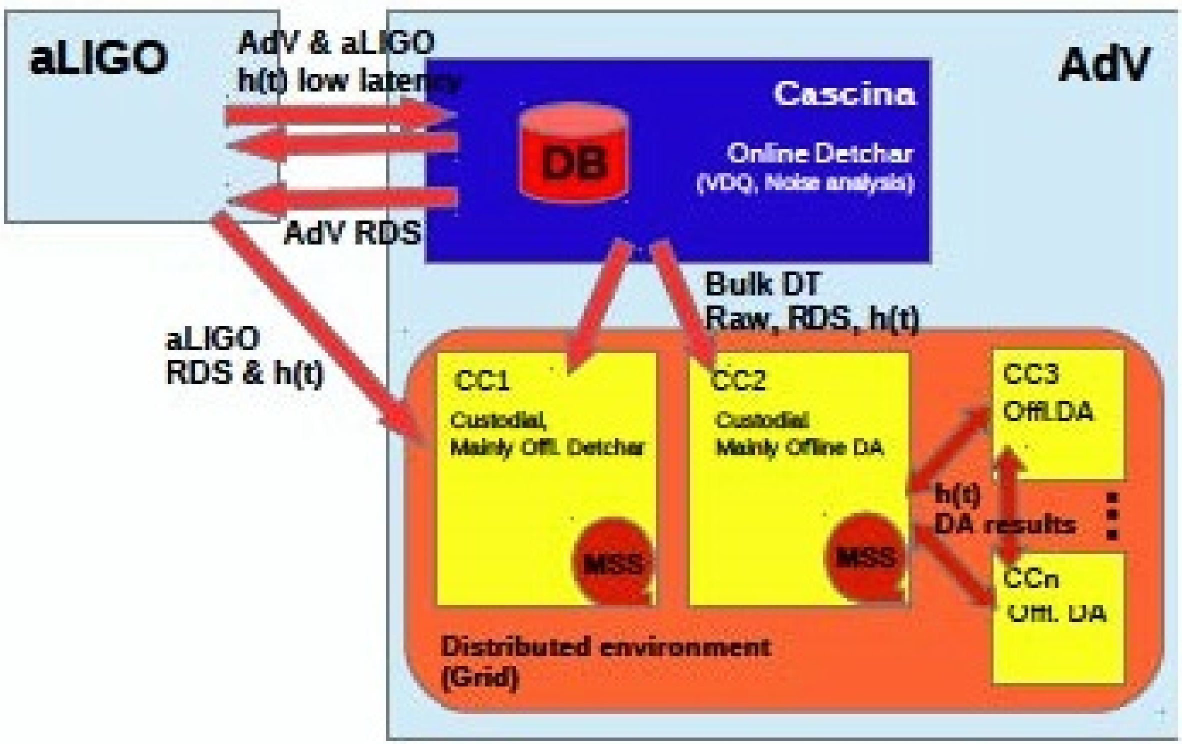
EGO site at Cascina hosts the Tier-0

"primary data" are distributed to Tier-1s, CNAF and CCIN2P3 with a maximum latency of 1 day

CNAF, CCIN2P3 and LIGO clusters are the main places where offline analysis are done.

EGO farm is fully dedicated to

- Data production
- Commissioning
- Detector characterization
- Low-latency searches (em follow-up)



Possible usage of other computing resources under investigation

Wigner (Hungary)

Holland, Poland

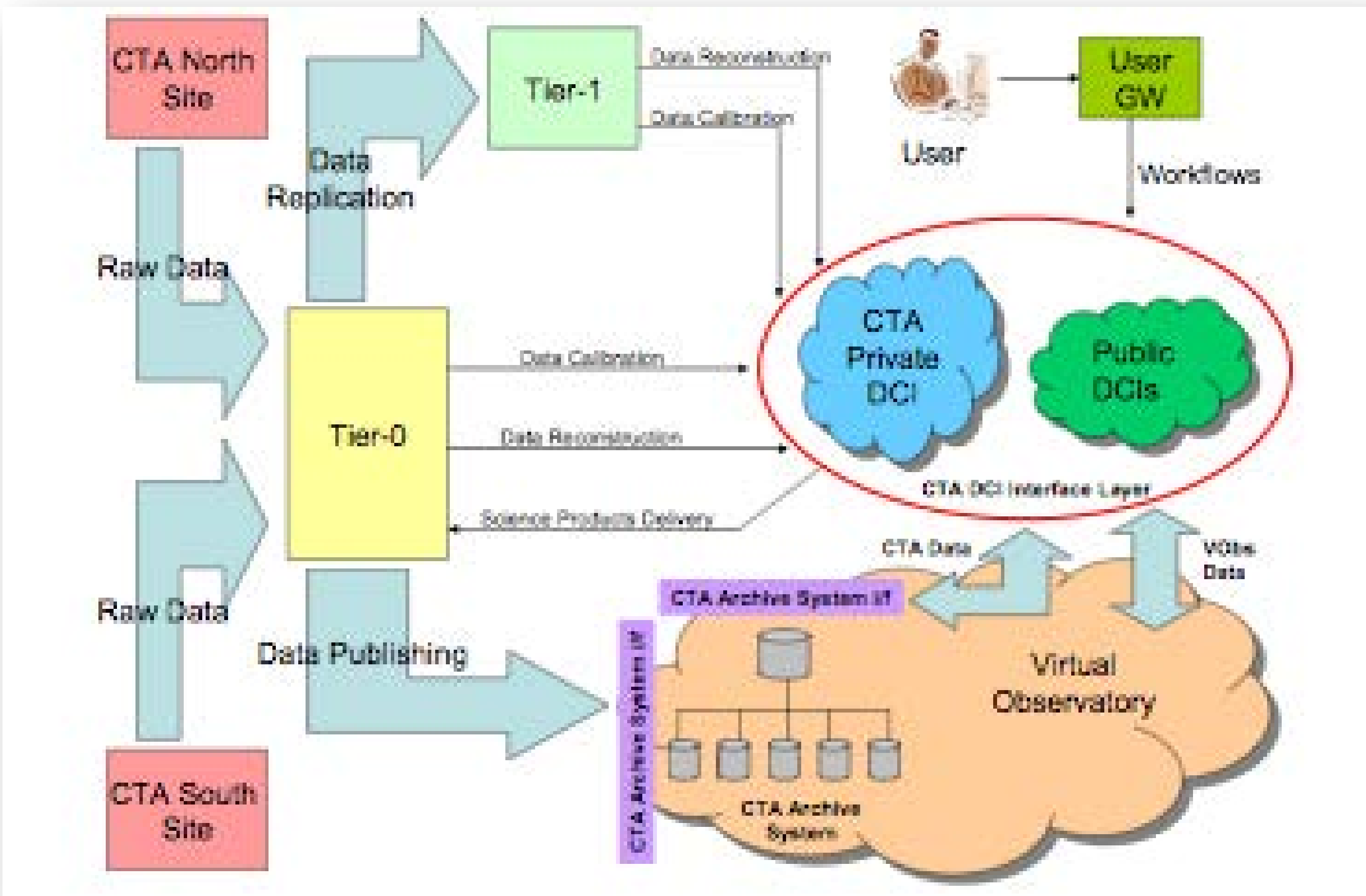
CHERENKOV TELESCOPE ARRAY
potential site locations

www.cta-observatory.org

CTA computing model

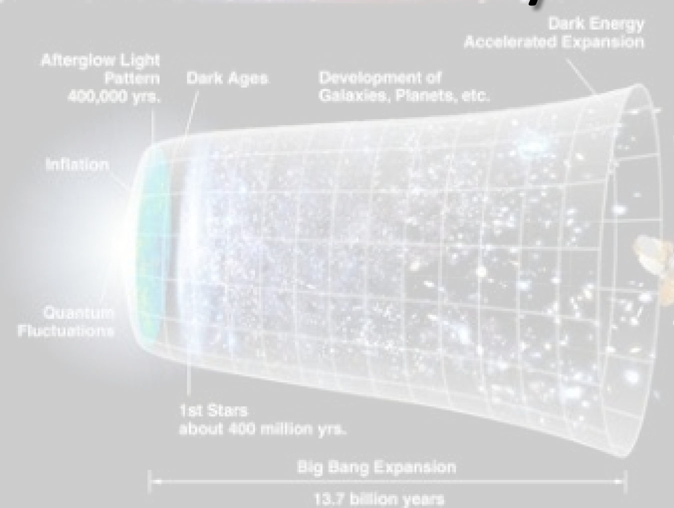
D. D'Urso, yesterday

CTA will produce a huge amount of data
that have to be Archived, Transferred, Processed and Accessed



Diversity


p 10^15-18 eV
p 10^20 eV
GZK
ν
γ


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
400,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.7 billion years

Astroparticle physics data can be categorized as events, time-series and images

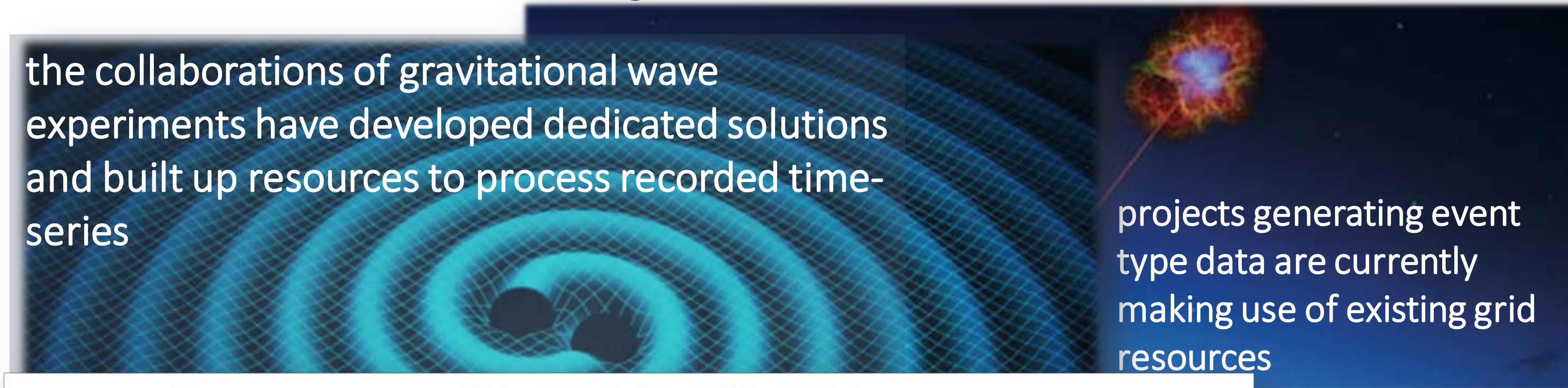
Each type of data requires a different type of analysis method

Astroparticle physics data can be categorized as events, time-series and images

Each type of data requires a different type of analysis method

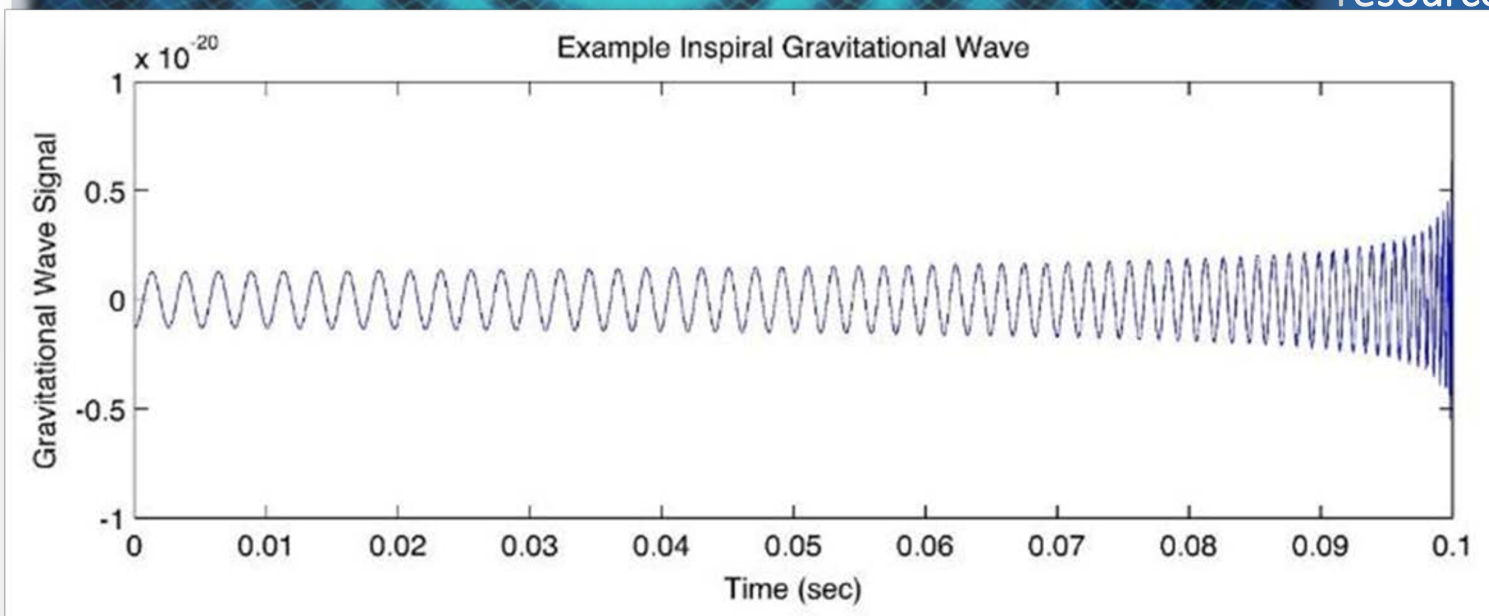
Astroparticle physics data can be categorized as events, time-series and images

Each type of data requires a different type of analysis method



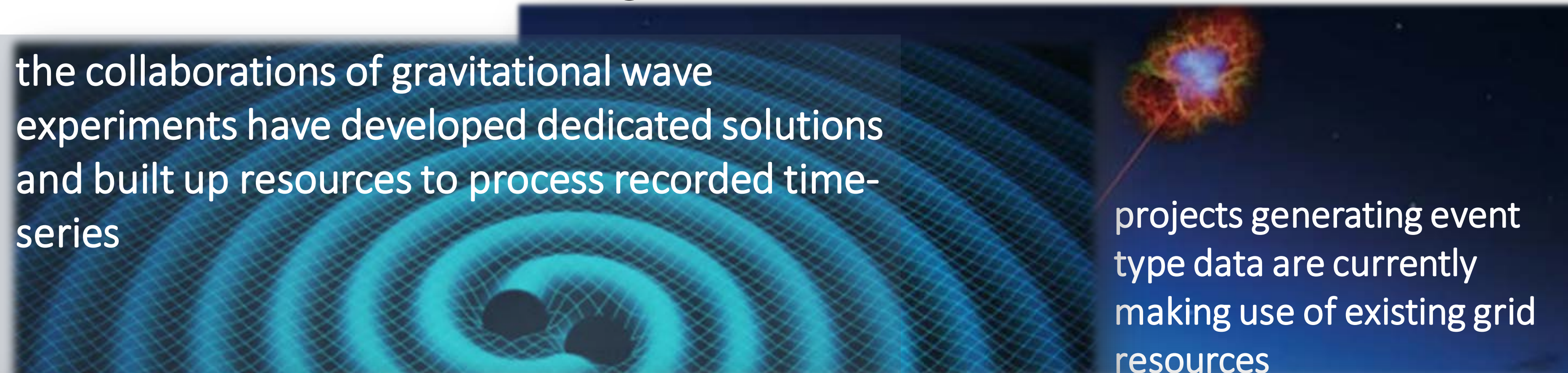
the collaborations of gravitational wave experiments have developed dedicated solutions and built up resources to process recorded time-series

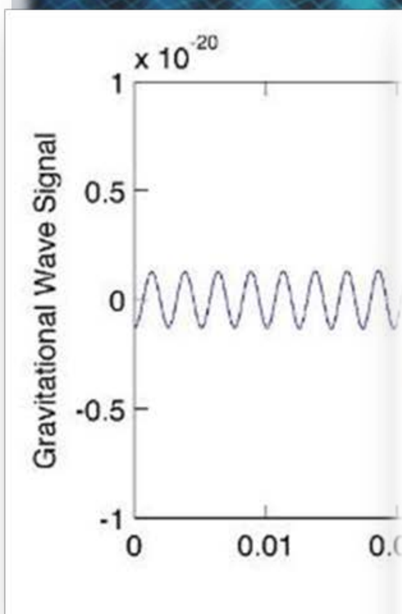
projects generating event type data are currently making use of existing grid resources

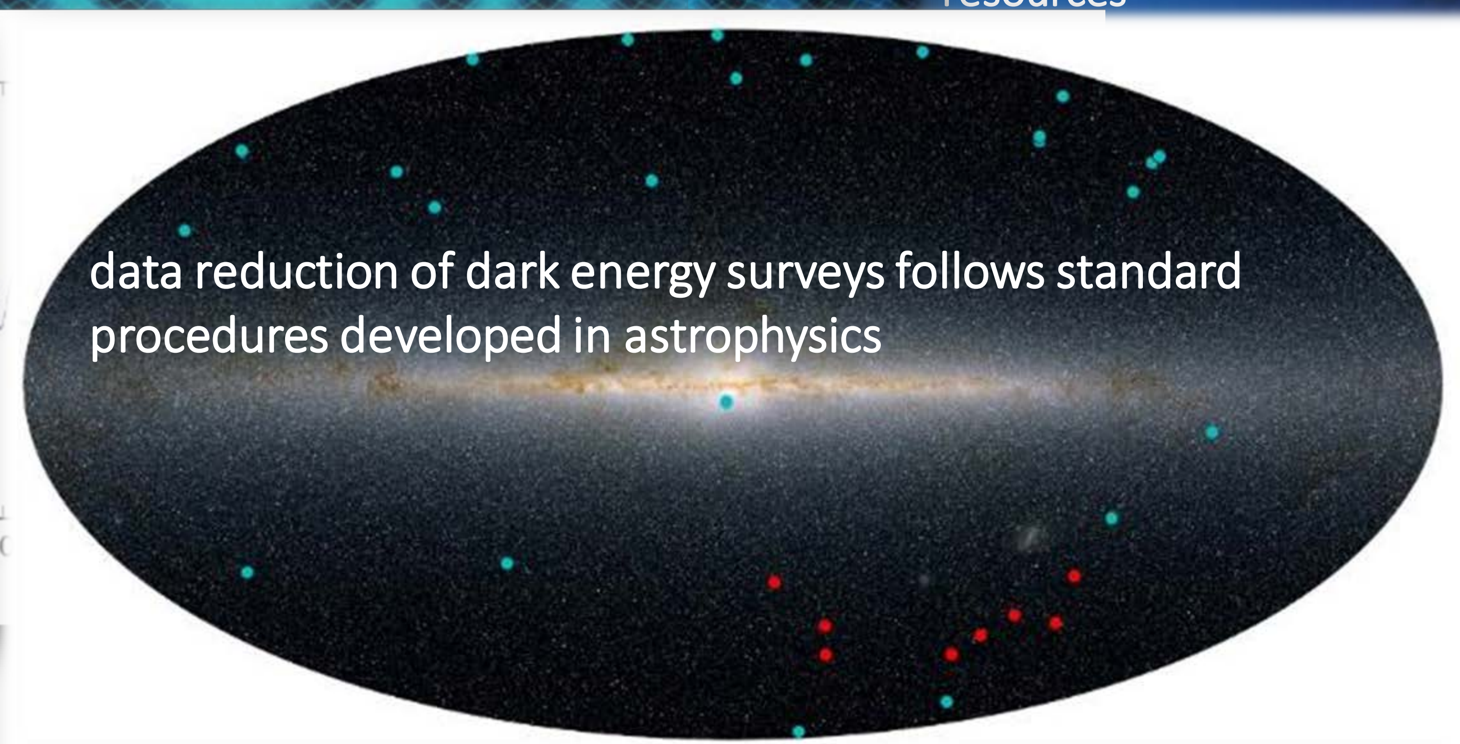


Astroparticle physics data can be categorized as events, time-series and images

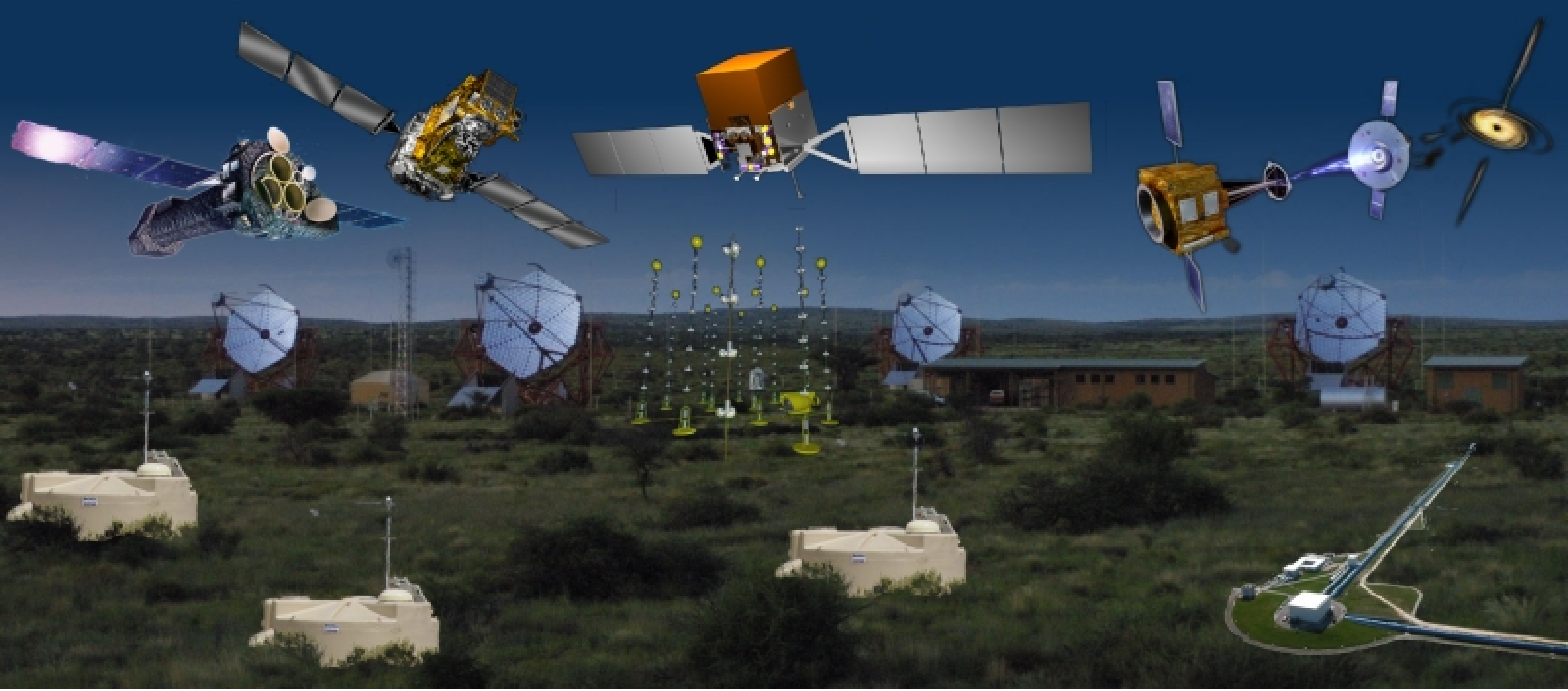
Each type of data requires a different type of analysis method







Multi-messenger



Astroparticle physics has much to gain from the **multi-messenger approach**

Cross-correlate data not only from experiments of the same type (e.g. from several neutrino detectors) but also across different domains

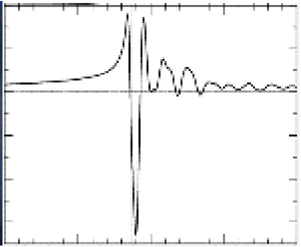
The **diversity** of data and its analysis makes astroparticle physics a formidable test-bed for new and innovative computing techniques concerning hardware, middleware, analysis software and database schemes

GRB prompt emission → TRIGGERED GW SEARCH

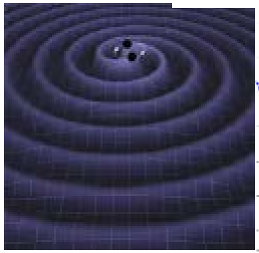
Known **GRB event time** and **sky position:**

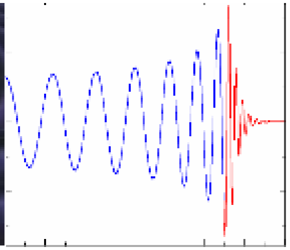
→ **reduction in search parameter space**
→ **gain in search sensitivity**

GW transient searches

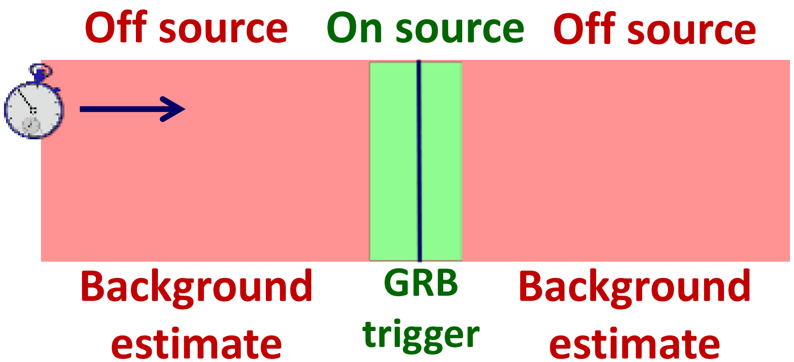


Unmodeled GW burst
(< 1 sec duration)
Arbitrary waveform
→ **Excess power**





Compact Binary Coalescence
Known waveform
→ **Matched filter**



Analyzed 154 GRBs detected by gamma-ray satellites during **2009-2010** while 2 or 3 LIGO/Virgo detectors were taken good data

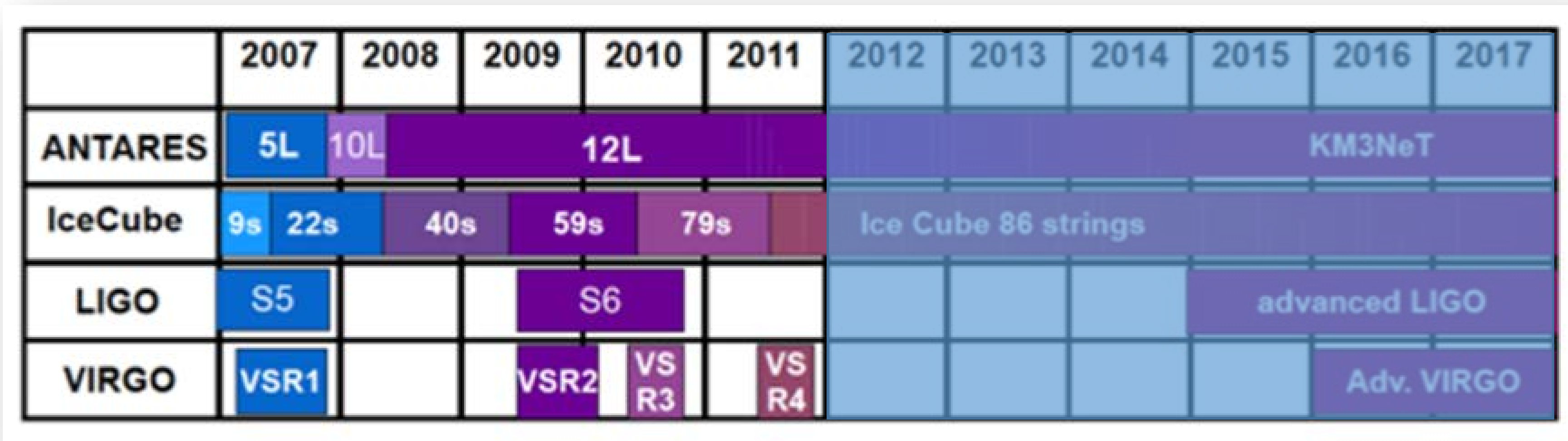
No evidence for gravitational-wave counterparts

Abadie et al. 2012, ApJ, 760
Aasi et al. 2014, PhRvL, 113

Credit: M. Branchesi

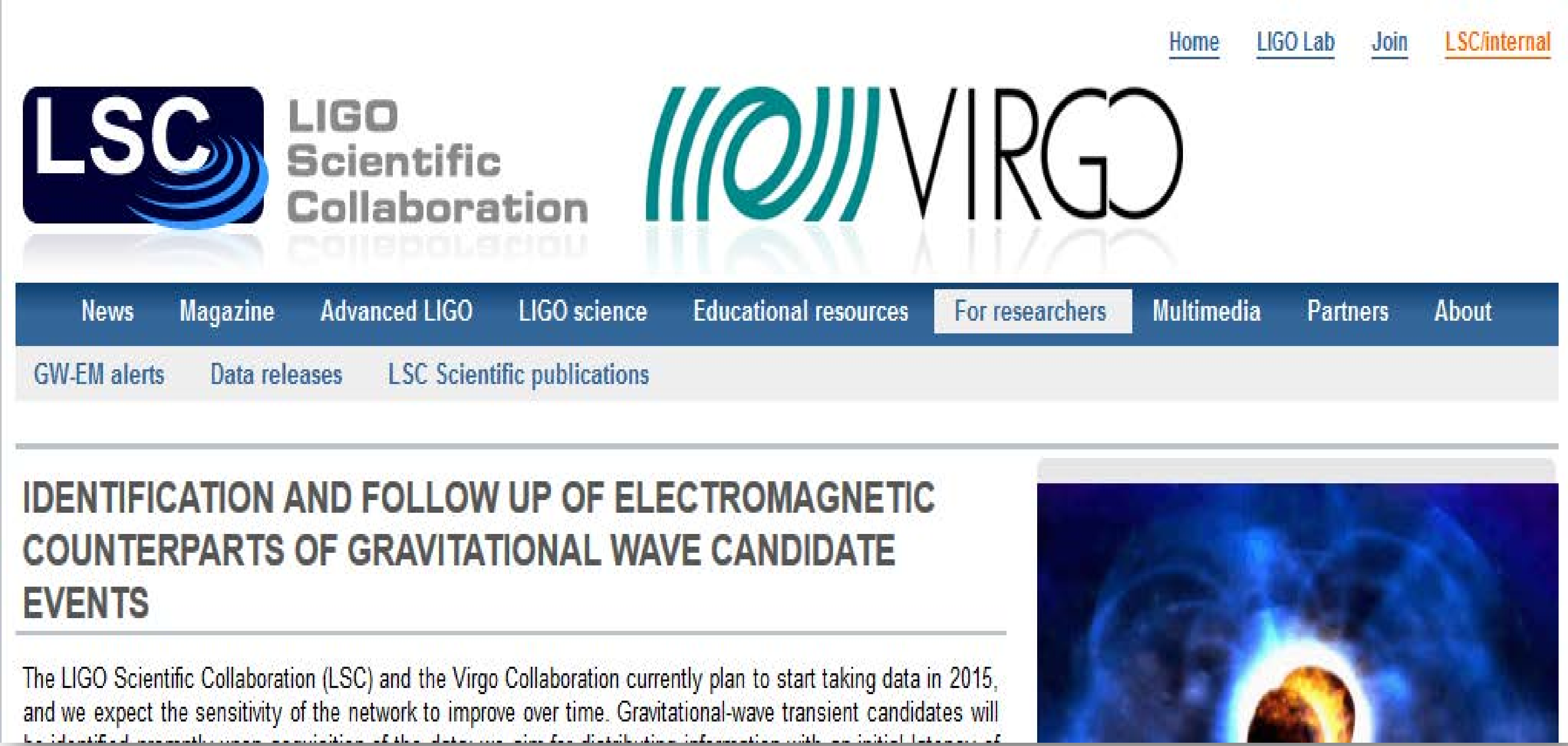
From ν to GW

Credit: G. Stratta



Search for coincident signals from LIGO and Virgo and ν detectors
No significant coincident event

Rate < 10^{-2} Mpc^{-3} yr^{-1}
This rate upper limit does not constraint current astrophysical models
(Aartsen et al. 2014 (Icecube), Adrian-Martinz et al. 2013 (ANTARES))



External triggers (e.g. from GRBs, neutrinos) drive GW data analysis providing trigger time and position in the sky

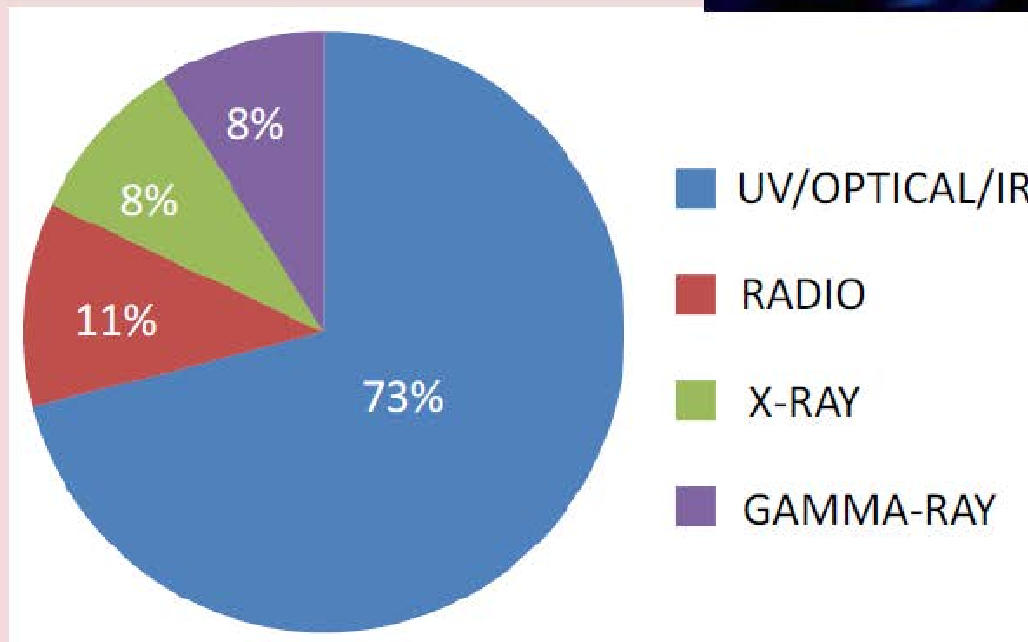
GW triggers above a certain threshold are released after few minutes to main observatories activating EM follow-up

LVC GW-EM follow-up program

~~Sixty~~ ~75 MoUs involving

- ➢ **150 instruments**

(satellites/world-wide ground-based)
covering the full spectrum
from radio to
very high-energy gamma-rays!



- ➢ **Astronomical institutions, agencies and large/small groups of astronomers** from 19 countries

Some of them https://gw-astronomy.org/wiki/LV_EM/PublicParticipatingGroups

Workshop just concluded (23rd of April) at EGO/Virgo

Credit: M. Branchesi

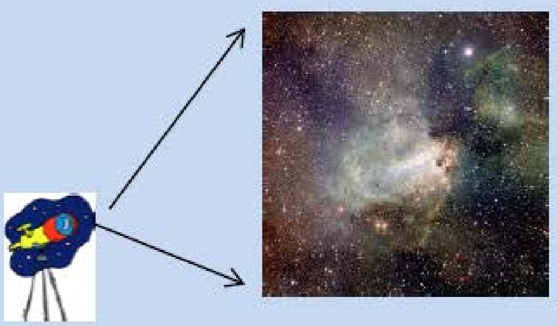
Challenges of the EM follow-up

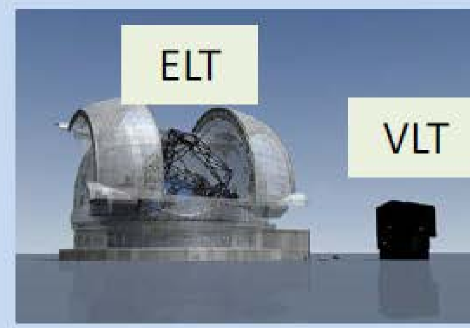
➢ Find **fast faint transient counterparts over** wide sky-localization areas of the GW candidates of order of **hundreds of square degrees**

This requires:

- **Hierarchical EM search**

Wide-field facilities ➡ **Analysis** to rapidly classify and select a sample of candidate counterparts ➡ **Larger telescopes** to **characterize** the nature of candidates

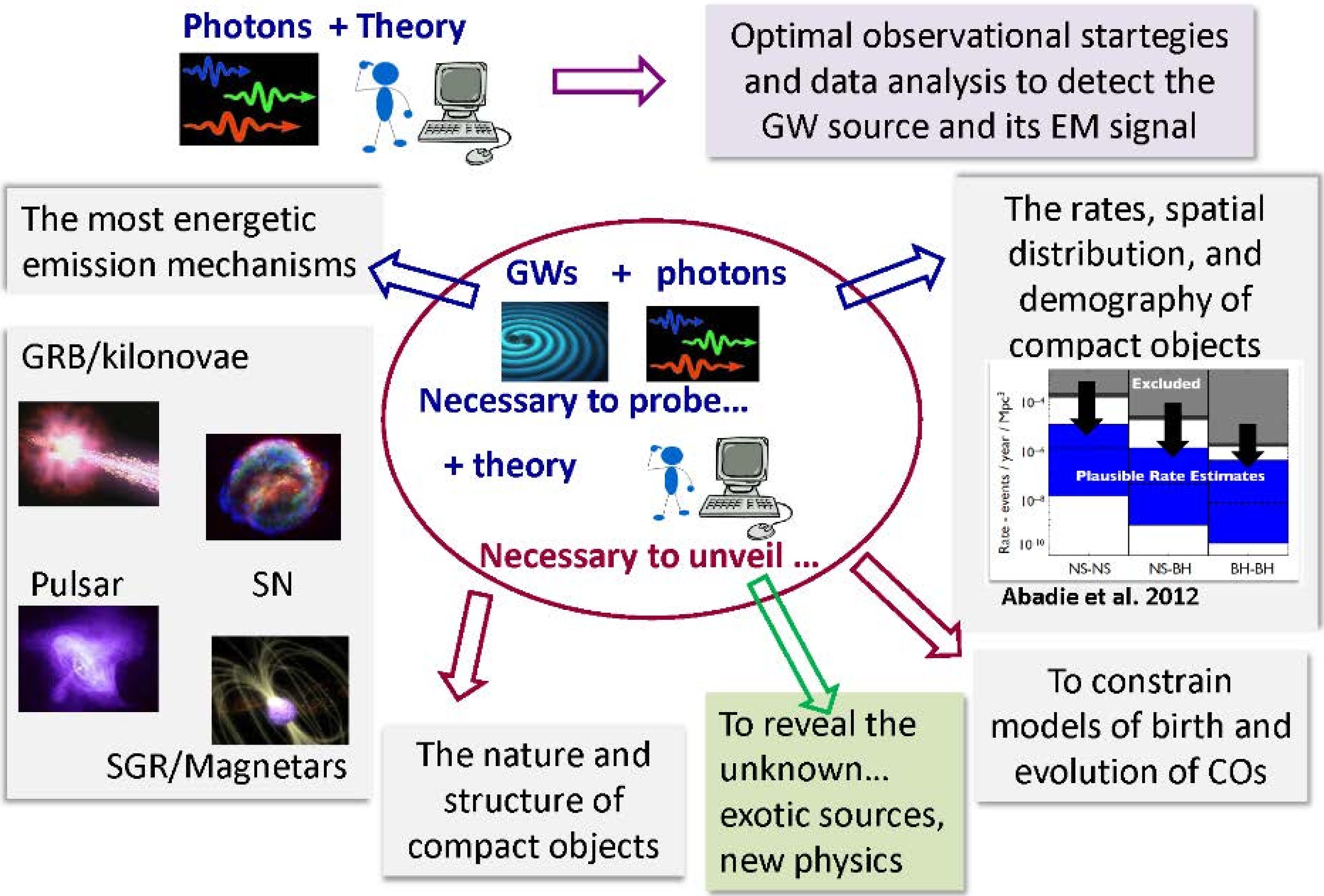




- **A tight network and collaborative effort between the gravitational-wave and astronomical communities**

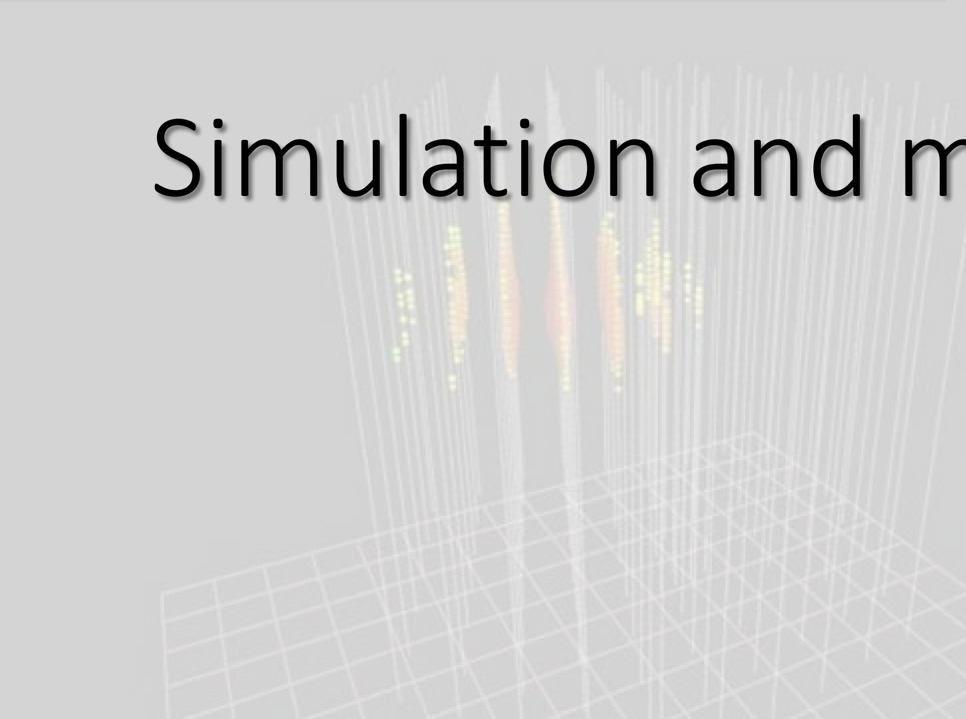
Credit: M. Branchesi

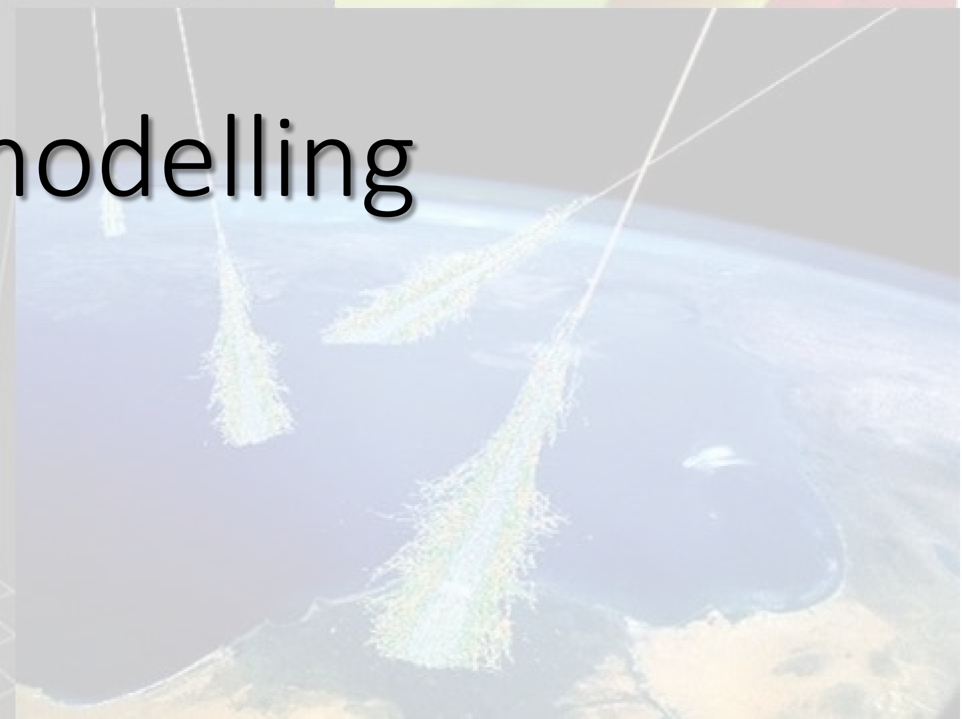
The multi-messenger photon and GW astronomy

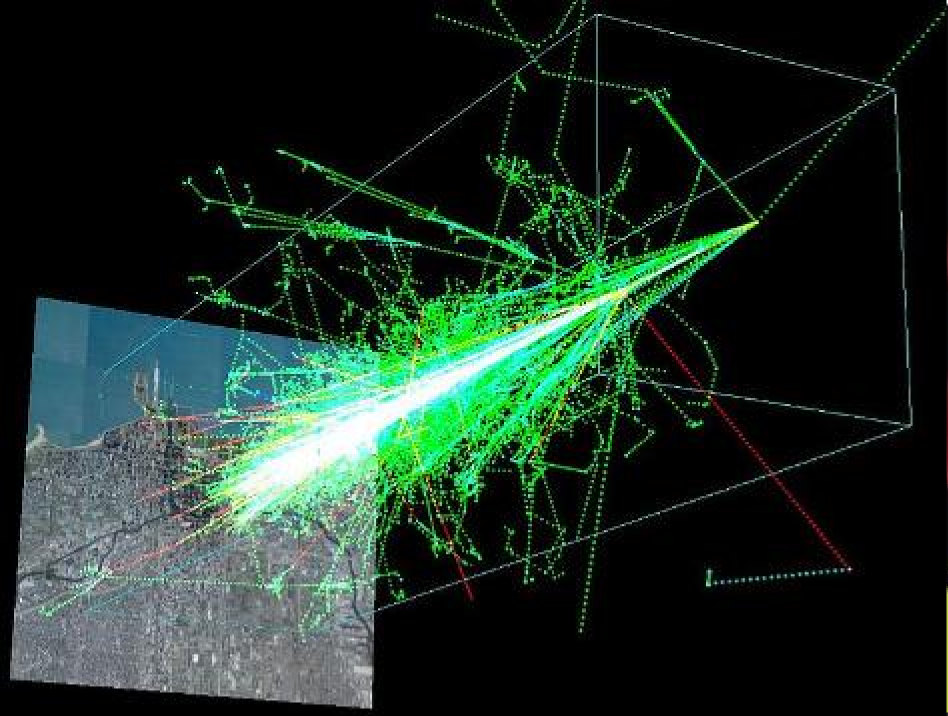


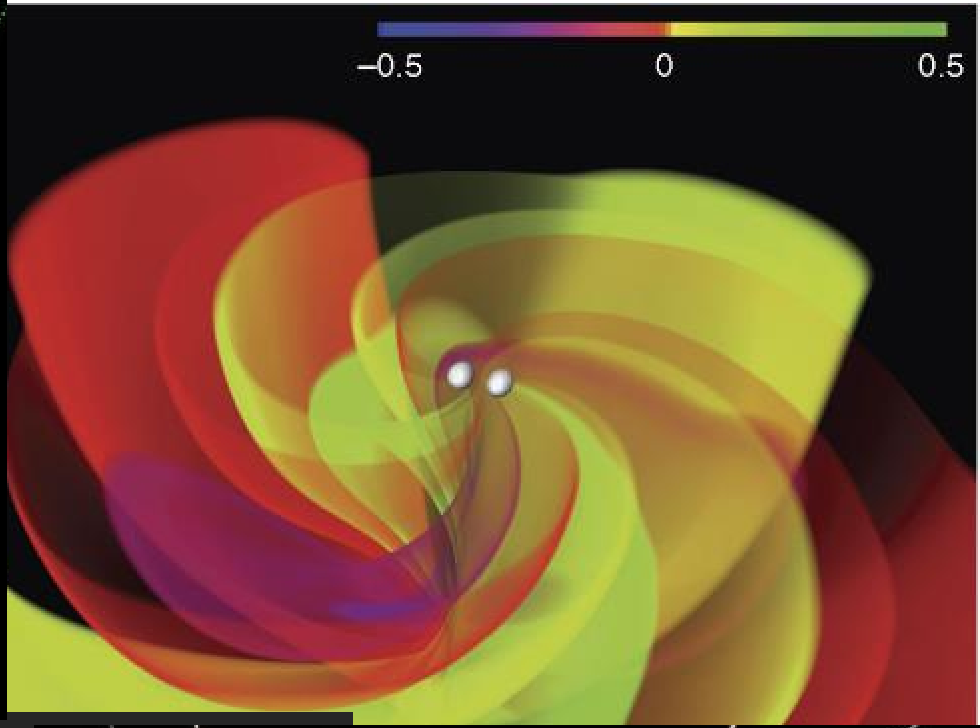
Credit: M. Branchesi

Simulation and modelling

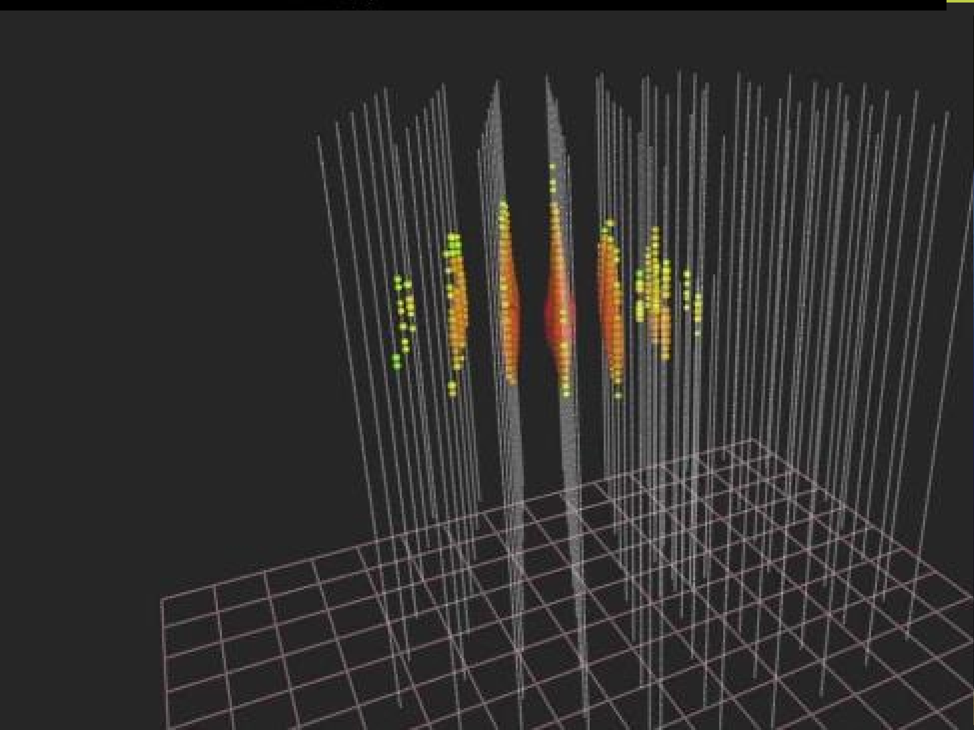


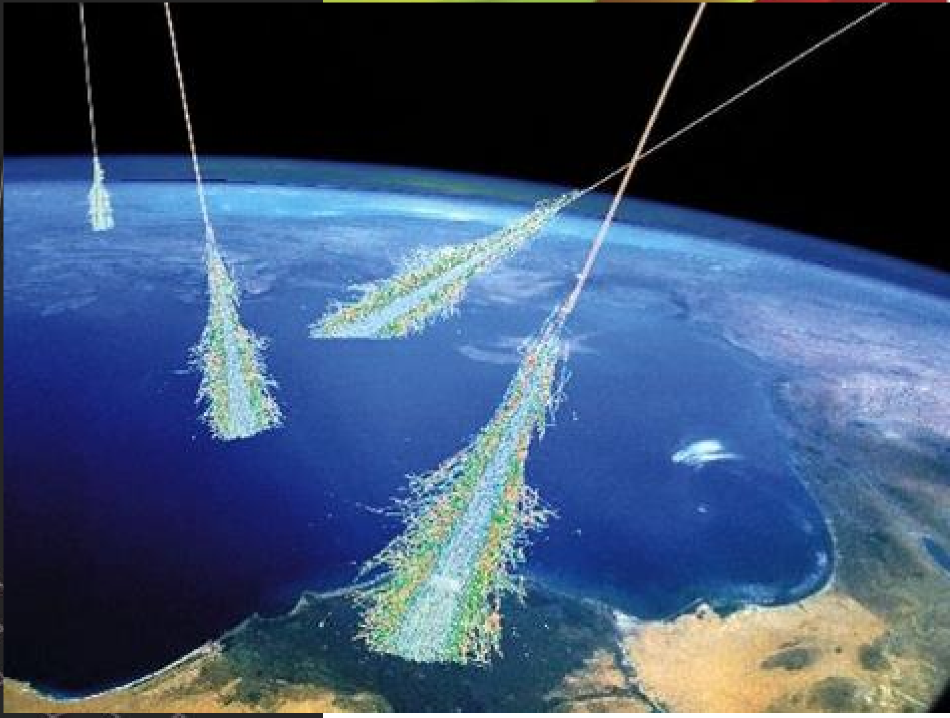






−0.5
0
0.5





GPUs

Simulations and modelling are often more CPU-consuming than the capture and analysis of real data

The gravitational wave community has been first in using GPU bases systems for their data analysis. Other communities are now using GPU based systems as well, e.g. for detailed extensive air showers simulations

GPUs for analysis

- Many search algorithm can be accelerated by making use of operation level paralellizability offered by various many-core hardwares such as GPUs. Such examples are:
  - FFT, vector operations, reduce, max finding, clustering **in CBC analysis pipelines**
  - FFT, 2D thresholding, differential Hough map creation, integration, peak finding, etc.. **in CW analysis**
- There are multiple tool developed to allow easier use of GPUs by less advanced programmers, such as:
  - GWTools - An OpenCL based templates C++ generic algorithm library for GW searches
  - pyCBC - CUDA based set of Python algorithm used in CBC analysis
  - CB - Compute Backend - offers a unified host code for CUDA and OpenCL, so there is no need to write the code twice for NVidia and AMD cards
- GPUs will play crucial role in the following years probably even for the discovery
- Typical full-pipeline **accelerations** experienced are ranging from **x30 to x120**

INFN Istituto Nazionale di Fisica Nucleare

Conclusions

- CTA Data model is going to defined. Final design with the first detector prototype
- Pipelines for Data and MC processing are under optimization (test on GPU for Real Time Analysis)

Credit: D. D'Urso

Cornell University Library

arXiv.org > physics > arXiv:1411.3968

Physics > Instrumentation and Detectors

Parallel Neutrino Triggers using GPUs for an underwater telescope

Bachir Bouhadef, Mauro Morganti, Giuseppe Terreni

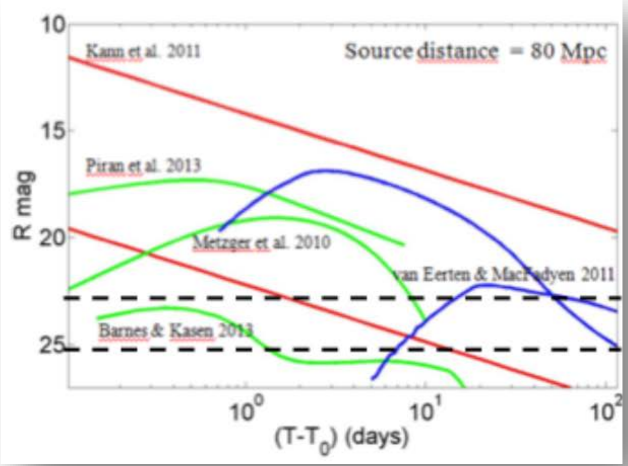
(Submitted on 14 Nov 2014 (v1), last revised 9 Mar 2015 (this version, v2))

Graphics Processing Units are high performance co-processors originally intended to improve the use and the acceleration of computer graphics applications. Because of their performance, researchers have extended their use beyond the computer graphics scope. We have investigated the possibility of implementing online neutrino trigger algorithms in the KM3Net-It experiment using a CPU-GPU system. The results of a neutrino trigger simulation on a NEMO Phase II tower and a KM3-It 14 floors tower are reported.

Trattamento di immagini di grande campo

Ricerca di oggetti transienti su campi estesi (decine gradi quadri)
Esempio: VST – 500 img/night@268 Mpx (300 GB/night)

- tempo (ricerca di transienti su oltre 500 immagini in tempi dell'ordine di un ora)
- smartness (know--how su metodi x detection di transienti + machine learning)



Questa rapidità è critica per comunicare in breve tempo le esatte coordinate dei più promettenti candidati ai grandi telescopi (classe 8m) e ottenere così la identificazione (caratterizzazione) definitiva con misure spettroscopiche

L' utilizzo della tecnologia GPU e la revisione dei codici sono in grado di guadagnare un fattore ~80 nei tempi di analisi delle immagini a grande campo

Credit: E. Brocato

Ministero dell'Istruzione, dell'Università e della Ricerca

Idee per un progetto premiale

Queste questioni vanno affrontate coinvolgendo una comunità ampia, e richiedono un forte coordinamento tra le collaborazioni scientifiche i centri di calcolo e le agenzie finanziatrici

La vera sfida è mettere insieme comunità con un background culturale differente e trovare un linguaggio comune per poter collaborare

- stato dell'arte dei modelli di calcolo e outlook degli sviluppi futuri. Ruolo dei centri di calcolo locali/regionali
- stato dell'arte e sviluppo prevedibile dell'hardware (computing e storage). Collaborazioni con l'industria?
- stato dell'arte e sviluppo del middleware (GRID, Cloud,...) Strategie comuni?
- data access policy (modelli di accesso ai dati, infrastrutture, outreach), coinvolgimento di altre comunità (fisica dell'ambiente, geofisica-INGV)

Partecipanti

- INFN
  - CNAF
  - CTA
  - FERMI
  - KM3Net
  - VIRGO
  - …

- INAF
  - CTA
  - …

Finanziamento

~ 2 M€

Programma

- istituire un gruppo di lavoro che determini le sinergie fra le diverse metodiche di analisi in ambito multi-messenger, in modo da sviluppare una comune architettura hardware e software
- investire su infrastrutture pilota per alcune tecnologie mirate (GPU, …)
- inserire giovani ricercatori nell'ambito di questa ricerca
- raccordarsi ad iniziative europee (ASTERICS)