Einstein Telescope Computing and Challenges

Stefano Bagnasco, INFN

Workshop CCR
Paestum

May 23, 2022

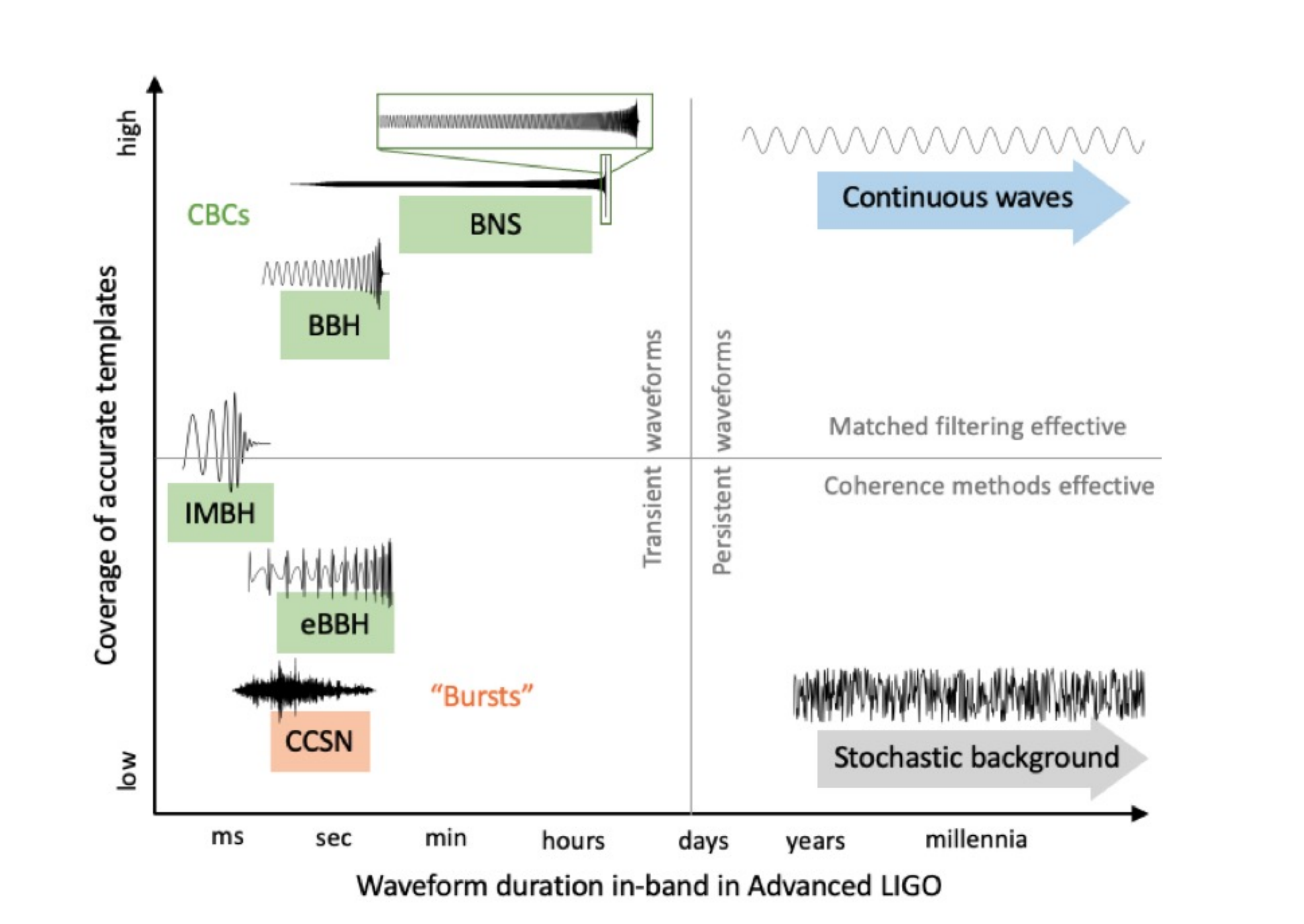
Looking for...

Burst sources:

- CBC: Compact Binary Coalescence
  - Coalescing Compact Binary Systems (Neutron Star-NS, Black Hole-NS, BH-BH): Strong emitters, well modelled for much of the parameter space
- Burst: Unmodeled transient bursts
  - Asymmetric Core Collapse Supernovae: weak emitters, not well-modelled (“bursts”), transient
  - Cosmic strings, soft gamma repeaters, pulsar glitches,...
  - Who knows?

Continuous sources:

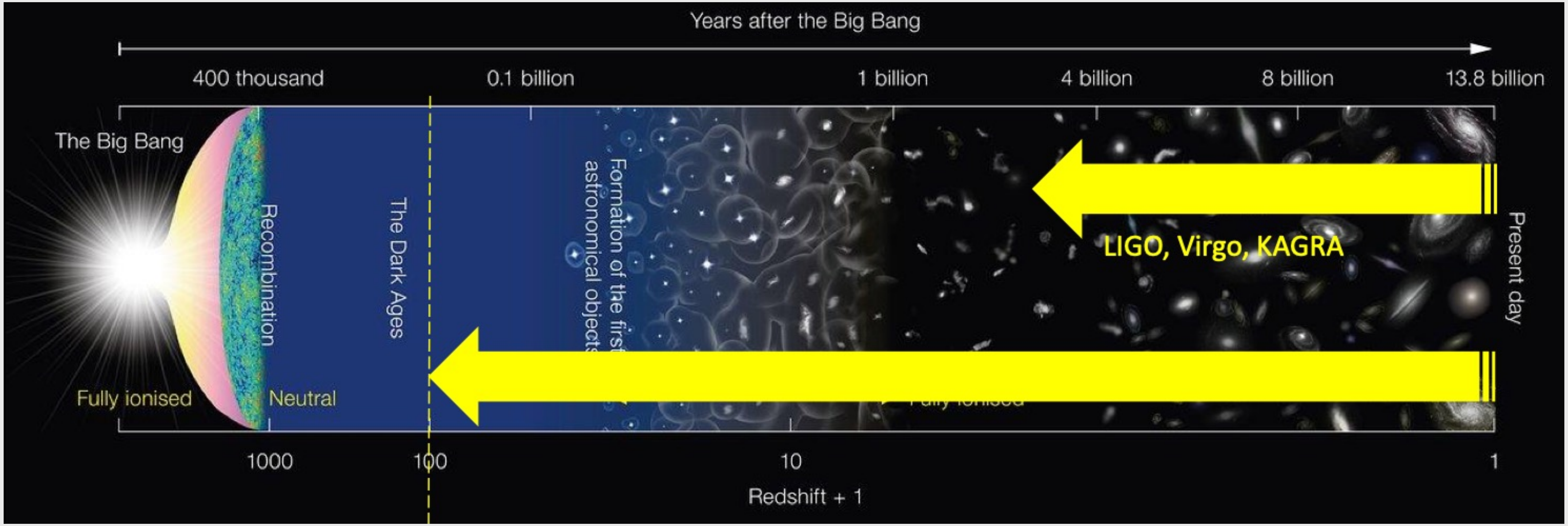
- CW: Continuous waves
  - Spinning neutron stars (known waveform, long/continuous duration)
  - All-sky and targeted searches
- SGWB: Continuous stochastic background
  - Cosmological stochastic background (residue of the Big Bang, cosmic GW background, long duration)
  - Astrophysical stochastic background



Jess McIver, D. H. Shoemaker, *Discovering Gravitational Waves with Advanced Ligo*, LIGO Document P200530-v1 (2000)

ET is the project aiming to realise the **European 3rd Generation Gravitational Wave observatory**

- A sensitivity at least 10 times better than the (nominal) 2G detectors
- Wideband, accessing the frequency band below 10Hz
- High reliability and improved observation capability



ASTROPHYSICS

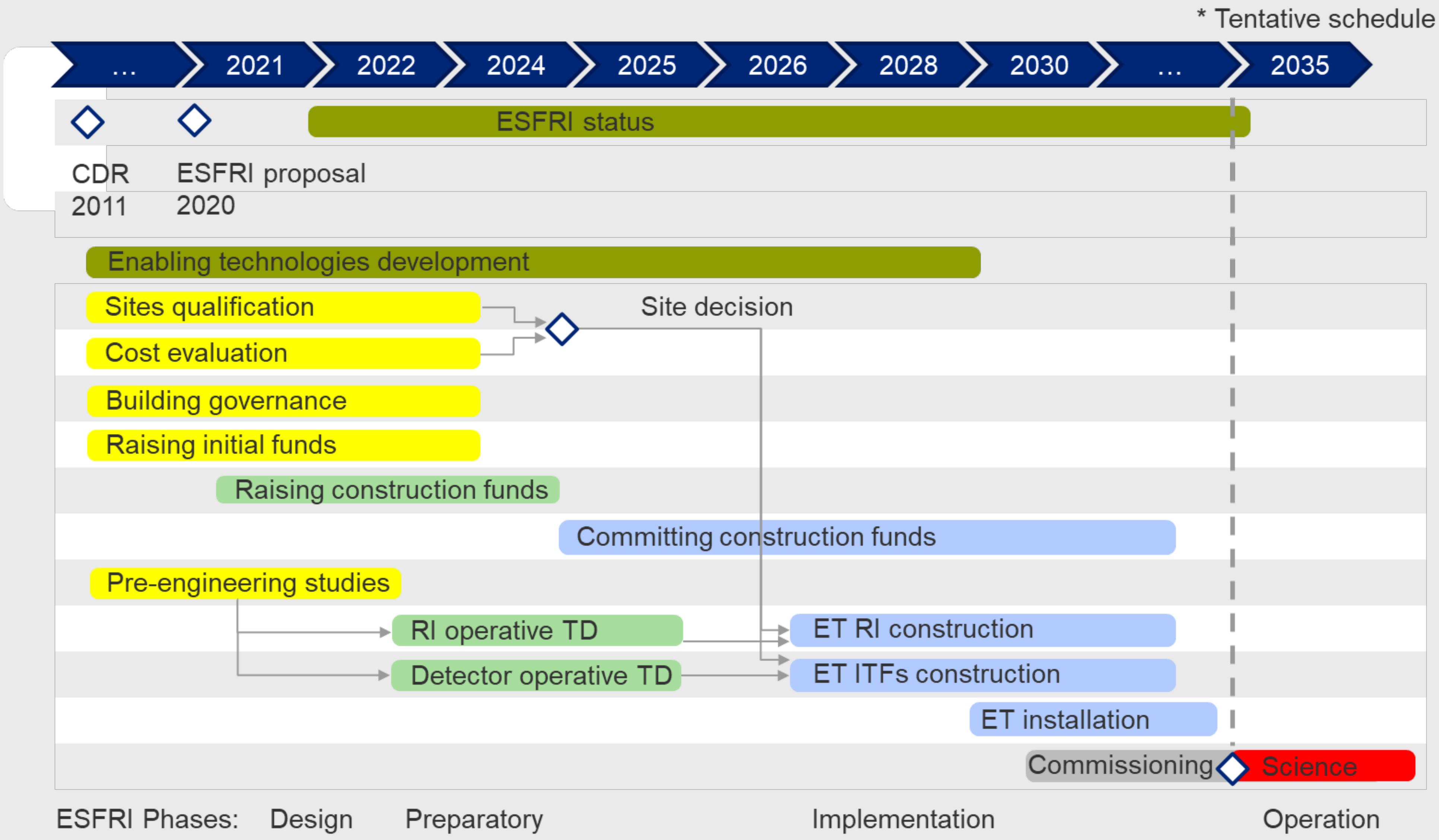
- Black hole properties
  - origin (stellar vs. primordial)
  - evolution, demography
- Neutron star properties
  - interior structure (QCD at ultra-high densities, exotic states of matter)
  - demography
- Multi-band and -messenger astronomy
  - joint GW/EM observations (GRB, kilonova,...)
  - multiband GW detection (LISA)
  - neutrinos
- Detection of new astrophysical sources
  - core collapse supernovae
  - isolated neutron stars
  - stochastic background of astrophysical origin

FUNDAMENTAL PHYSICS AND COSMOLOGY

- The nature of compact objects
  - near-horizon physics
  - tests of no-hair theorem
  - exotic compact objects
- Tests of General Relativity
  - post-Newtonian expansion
  - strong field regime
- Dark matter
  - primordial BHs
  - axion clouds, dark matter accreting on compact objects
- Dark energy and modifications of gravity on cosmological scales
  - dark energy equation of state
  - modified GW propagation
- Stochastic backgrounds of cosmological origin
  - inflation, phase transitions, cosmic strings

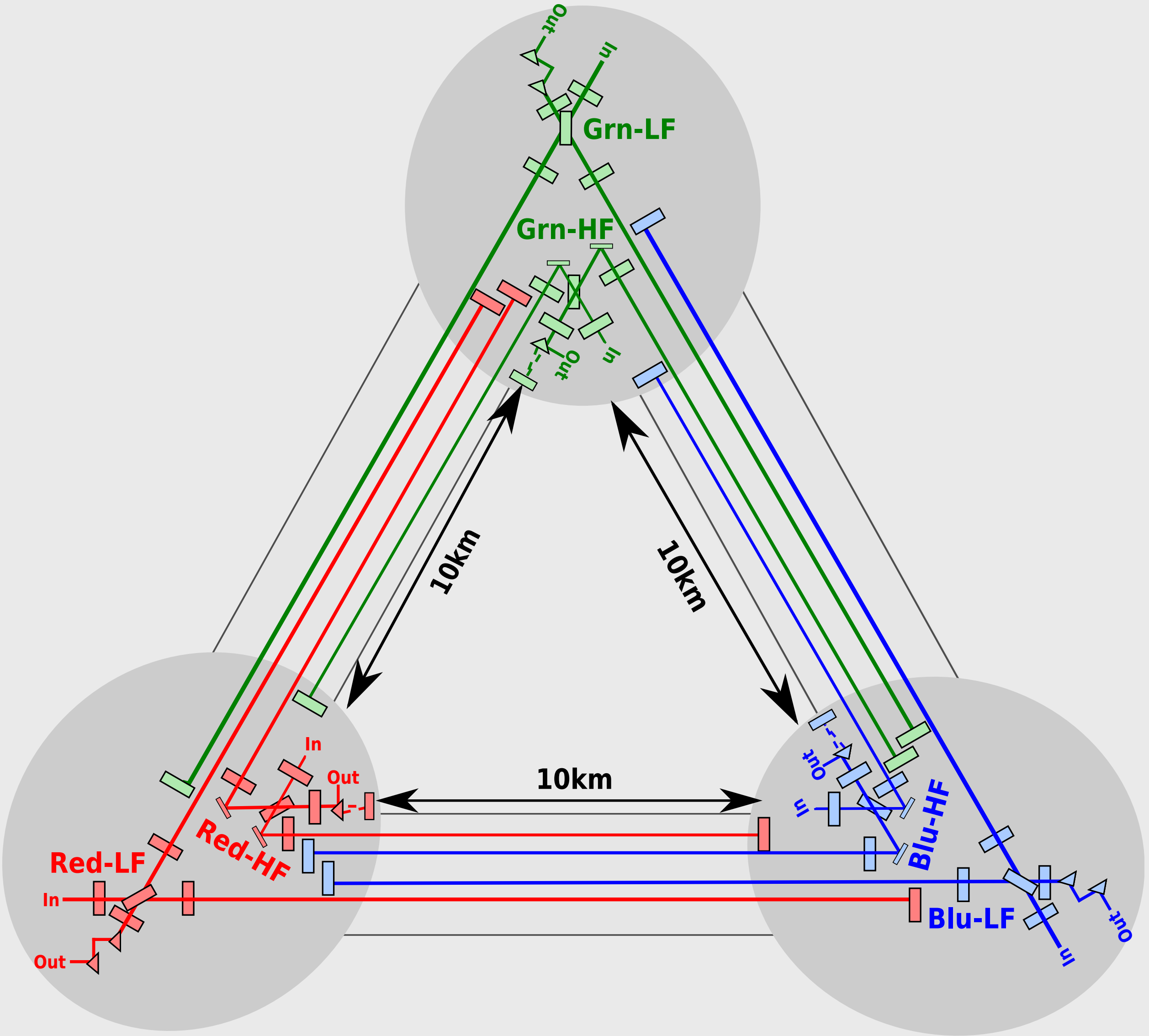
- Currently there are two candidate sites being characterized to host ET:
  - The Sardinia site, close to the Sos Enattos mine
  - The Euregio Meuse-Rhine site, close to the NL-B-D border
  - A third option in Saxony (Germany) was recently proposed and is under discussion

ET: when

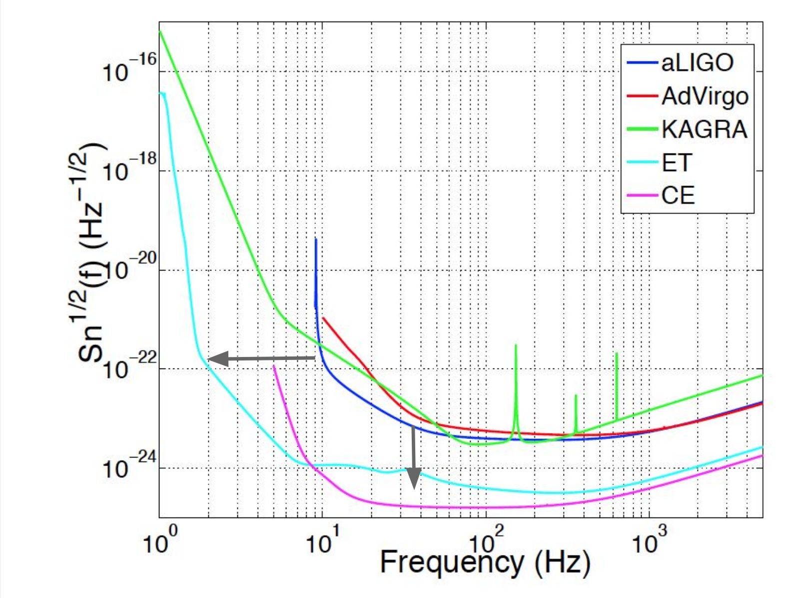

* Tentative schedule
…
2021
2022
2024
2025
2026
2028
2030
…
2035
ESFRI status
CDR
2011
ESFRI proposal
2020
Enabling technologies development
Sites qualification
Site decision
Cost evaluation
Building governance
Raising initial funds
Raising construction funds
Committing construction funds
Pre-engineering studies
RI operative TD
ET RI construction
Detector operative TD
ET ITFs construction
ET installation
Commissioning
Science
ESFRI Phases:
Design
Preparatory
Implementation
Operation

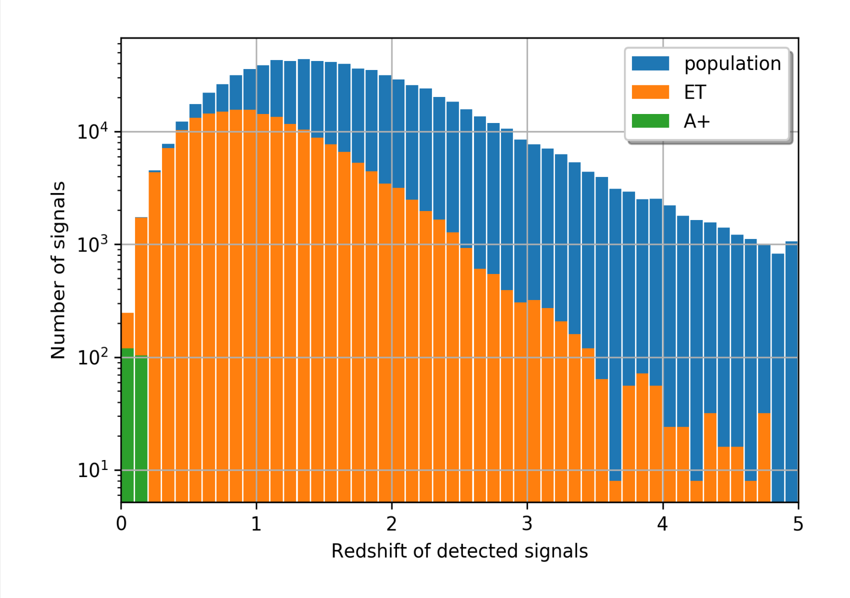
A triangular xylophone

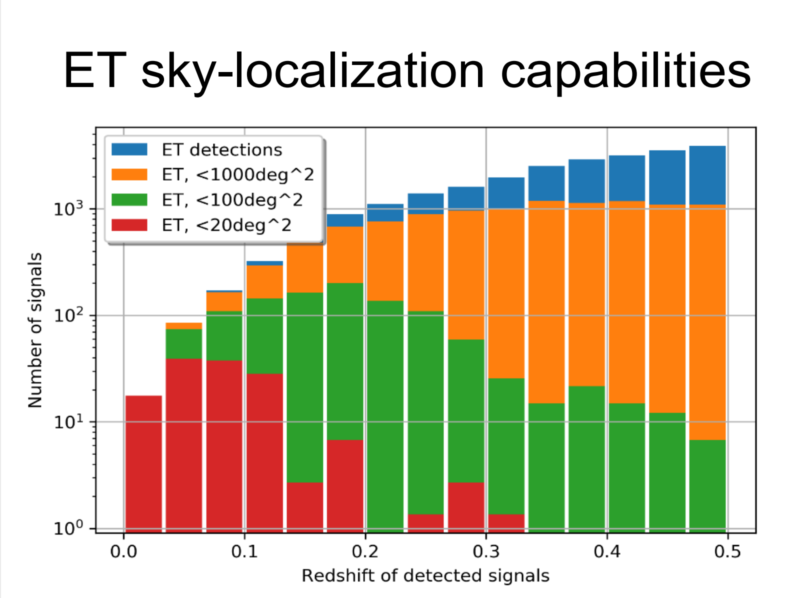
- Three detectors in a **triangular** structure
  - Closed geometry allows the use of the null data stream
  - The third detector makes up for 60° angle
- Each detector (red, green and blue) consists of **two** Michelson interferometers
  - High-frequency and (more challenging) Low-frequency



Enhanced sensitivity





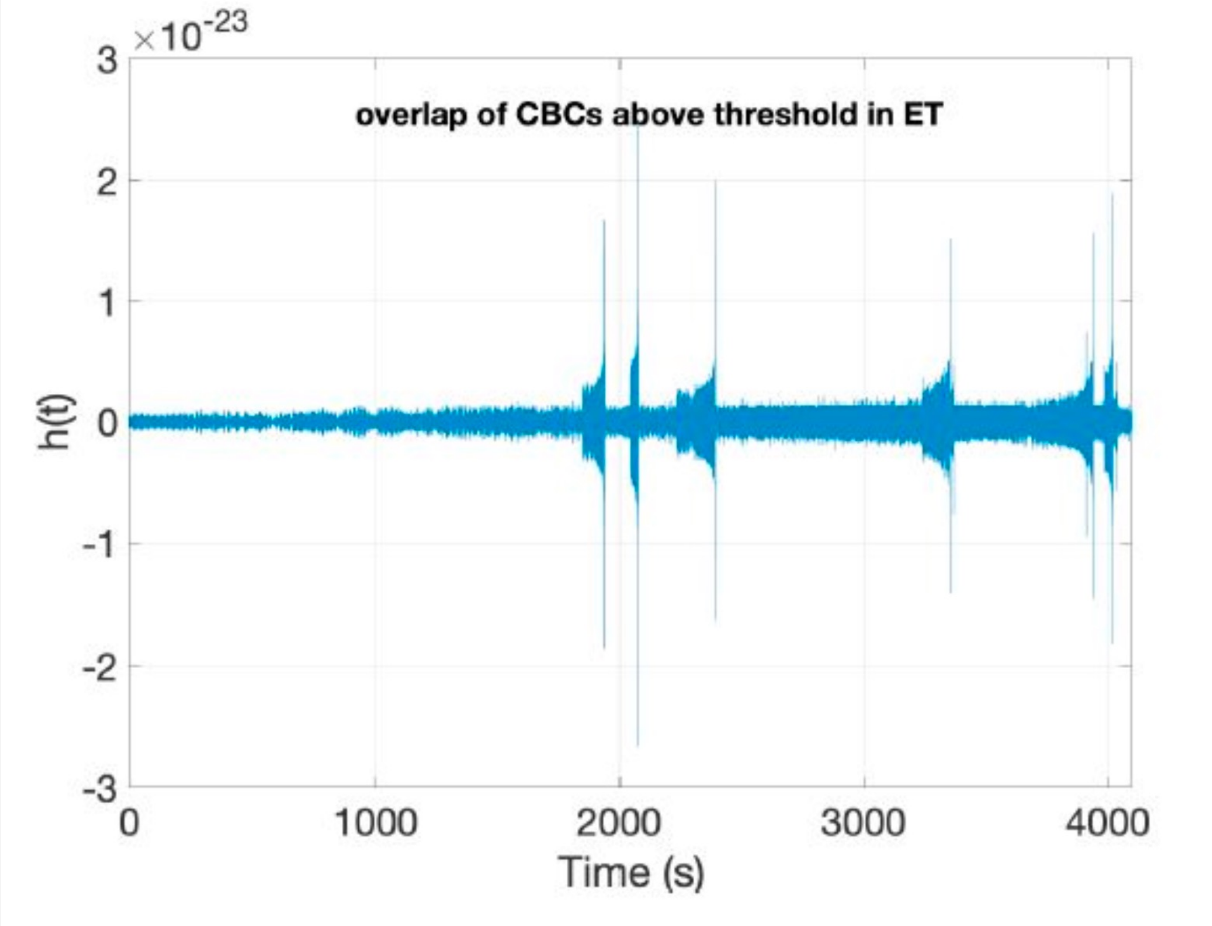


marica.branchesi@gssi.it

- Lower frequencies (down to 1-5Hz)
- 10^5 BBH detections per year
- o(100) detections per year with <20 deg^2
- Much higher sensitivity
- 10^5 BNS detections per year
- Early warning by minutes (hours)

Some challenges

- High alert rates
- Overlapping signals
- FAR estimate in the presence of a strong foreground
- Environmental correlated noise (non-independent colocated detectors)
- Long duration waveform for CBCs (and a moving detector), meaning large increase in computing needs
- ...



Data and more data

- Luckily, raw interferometer data don't grow much with increasing instrument sensitivity
  - We're not exploding like HL-LHC was!
  - Current detectors write o(2PB)/year of raw data per detector
  - h(t) (or "strain", the physics channel) + o(10^5) control channels
  - Pre-processed data for final user analysis is more than 1 order of magnitude smaller
  - In ET we expect about few tens of PB of raw data per year (baseline 6-interferometer design, more control channels,...)
  - No big deal today, piece of cake by 2035

- However, the amount of useful scientific information encoded in the data does grow a lot
  - And the computing power needed to wring it out (mostly from CBC Parameter Estimation)
  - It's a difficult task in itself to precisely estimate the computing power needs

1/10th of an LHC experiment

- Current computing needs of the entire GW network are roughly o(10%) of an LHC experiment of today

- In ET the event rate will be 10^3 -10^4 times the current one
  - Analysis of the “golden” events (EM counterparts, high SNR or “special” events) would already be within reach using current technologies
    - O(500) events per year = 12.5MHS06-y per year, the same order of magnitude of a LHC experiment in Run 4
    - Target: 1/10th of an LHC experiment in Run 4

But: low-latency!

Three computing domains

On-site infrastructure

Online

- Data acquisition and pre-processing
- Instrument control
- Environmental monitoring
- ...

Plain old HTC (and some HPC)

Offline

- Deep searches
- Offline parameter estimation
- (Template bank generation)
- ...

Here's the fun

Low-latency

- Candidate search
- Sky localization
- LL parameter estimation
- Alert generation and distribution

Early warning

The Astrophysical Journal Letters, 910:L21 (7pp), 2021 April 1 https://doi.org/10.3847/2041-8213/abed54

First Demonstration of Early Warning Gravitational-wave Alerts

Ryan Magee[1,2,3,27], Deep Chatterjee[4,5,6,27], Leo P. Singer[7,27], Surabhi Sachdev[2,3,27], Manoj Kovalam[8,9], Geoffrey Mo[10,11], Stuart Anderson[1], Patrick Brady[6], Patrick Brockill[6], Kipp Cannon[12], Tito Dal Canton[13], Qi Chu[8,9], Patrick Clearwater[8,14], Alex Codoreanu[8,14], Marco Drago[15,16], Patrick Godwin[2,3], Shaon Ghosh[17], Giuseppe Greco[18,19], Chad Hanna[2,3,20], Shasvath J. Kapadia[21], Erik Katsavounidis[10,11], Victor Oloworaran[8,9], Alexander E. Pace[2,3], Fiona Panther[8,9], Anwarul Patwary[8,9], Roberto De Pietri[22,23], Brandon Piotrzkowski[6], Tanner Prestegard[6], Luca Rei[24], Anala K. Sreekumar[8,9], Marek J. Szczepańczyk[25], Vinaya Valsan[6], Aaron Viets[6], Madeline Wade[26], Linqing Wen[8,9], and John Zweizig[1]

[1] LIGO, California Institute of Technology, Pasadena, CA 91125, USA
[2] Department of Physics, The Pennsylvania State University, University Park, PA 16802, USA
[3] Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA
[4] Center for Astrophysical Surveys, National Center for Supercomputing Applications, Urbana, IL 61801, USA
[5] Illinois Center for Advanced Studies of the Universe, Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[6] Leonard E. Parker Center for Gravitation, Cosmology, and Astrophysics, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
[7] Astroparticle Physics Laboratory, NASA Goddard Space Flight Center, Mail Code 661, Greenbelt, MD 20771, USA
[8] Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), Australia
[9] Department of Physics, University of Western Australia, Crawley WA 6009, Australia
[10] LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[11] Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[12] Research Center for the Early Universe, The University of Tokyo, 113-0033, Japan
[13] Université Paris-Saclay, CNRS/IN2P3, IJCLab, F-91405 Orsay, France
[14] Gravitational Wave Data Centre, Swinburne University, Hawthorn VIC 3122, Australia
[15] Università di Roma "La Sapienza," I-00185 Roma, Italy
[16] INFN, Sezione di Roma, I-00185 Roma, Italy
[17] Department of Physics and Astronomy, Montclair State University, Montclair, NJ 07043, USA
[18] Università degli Studi di Urbino "Carlo Bo", I-61029 Urbino, Italy
[19] INFN, Sezione di Firenze, I-50019 Sesto Fiorentino, Firenze, Italy
[20] Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA
[21] International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bangalore 560089, India
[22] Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, I-43124 Parma, Italy
[23] 7INFN, Sezione di Milano Bicocca, Gruppo Collegato di Parma, I-43124 Parma, Italy
[24] INFN, Sezione di Genova, I-16146 Genova, Italy
[25] University of Florida, Gainesville, FL 32611, USA
[26] Department of Physics, Kenyon College, Gambier, OH 43022, USA
Received 2021 February 25; accepted 2021 March 10; published 2021 April 6

Abstract

Gravitational-wave observations became commonplace in Advanced LIGO-Virgo's recently concluded third observing run. 56 nonretracted candidates were identified and publicly announced in near real time. Gravitational waves from binary neutron star mergers, however, remain of special interest since they can be precursors to high-energy astrophysical phenomena like γ-ray bursts and kilonovae. While late-time electromagnetic emissions provide important information about the astrophysical processes within, the prompt emission along with gravitational waves uniquely reveals the extreme matter and gravity during—and in the seconds following—merger. Rapid communication of source location and properties from the gravitational-wave data is crucial to facilitate multimessenger follow-up of such sources. This is especially enabled if the partner facilities are forewarned via an early warning (pre-merger) alert. Here we describe the commissioning and performance of such a low-latency infrastructure within LIGO-Virgo. We present results from an end-to-end mock data challenge that detects binary neutron star mergers and alerts partner facilities before merger. We set expectations for these alerts in future observing runs.

Unified Astronomy Thesaurus concepts: Gravitational waves (678); Gravitational wave astronomy (675); Neutron stars (1108); High energy astrophysics (739)

1. Introduction

The field of gravitational-wave astronomy has exploded in the years following the first direct observation of gravitational waves (GWs) from a binary black hole (BBH) merger (Abbott et al. 2016). Since then, LIGO-Virgo have published 49 candidate events, many of which were identified in low-latency;[28] these include two binary neutron star (BNS) and two neutron star–black hole (NSBH) candidates (Abbott et al. 2020a). The detection of GWs from compact binaries, especially from BBHs, has become routine. GWs from BNS and NSBH mergers, however, remain rare. BNS and NSBH mergers are of special interest due to the possibility of counterpart electromagnetic (EM) signals. For BNS mergers, in particular, it has long been hypothesized that the central engine (post merger) can launch short gamma-ray bursts (SGRBs; Lattimer & Schramm 1976; Lee & Ramirez-Ruiz 2007), kilonovae (Li & Paczynski 1998; Metzger et al. 2010), and radio waves and X-rays post merger (Nakar & Piran 2011; Metzger & Berger 2012). In the special case of the

[27] These authors contributed equally to this work.
[28] Some of the 56 have not yet appeared in a LIGO-Virgo publication.

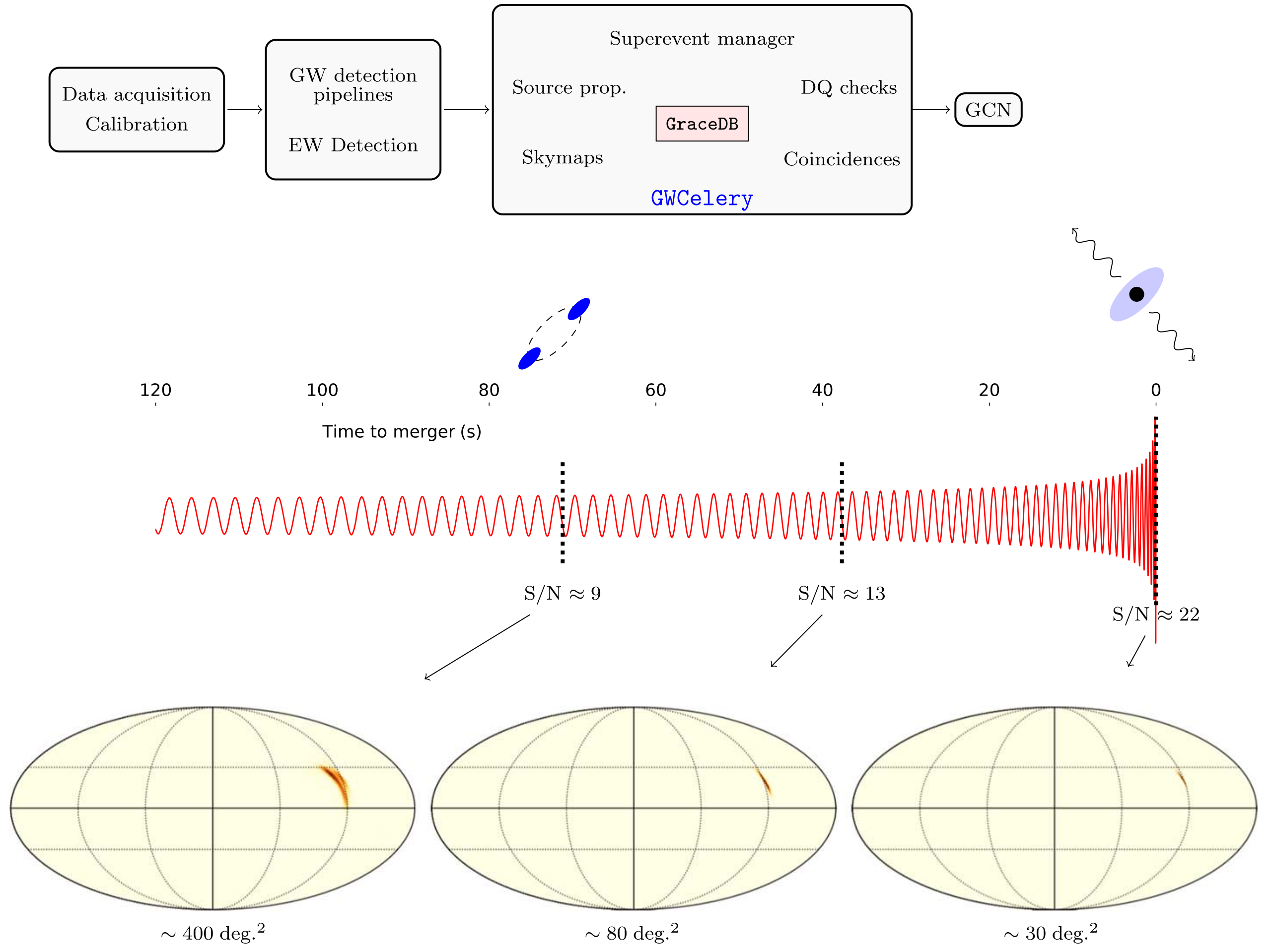
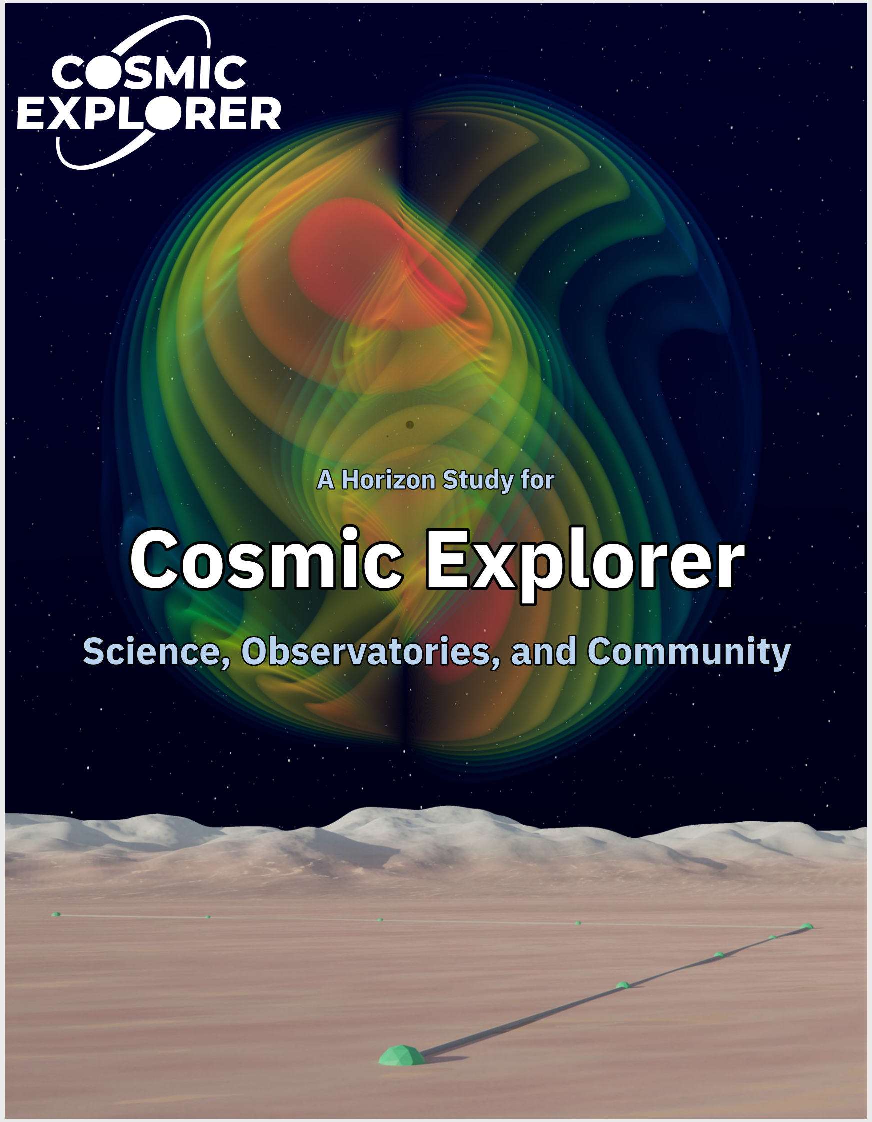


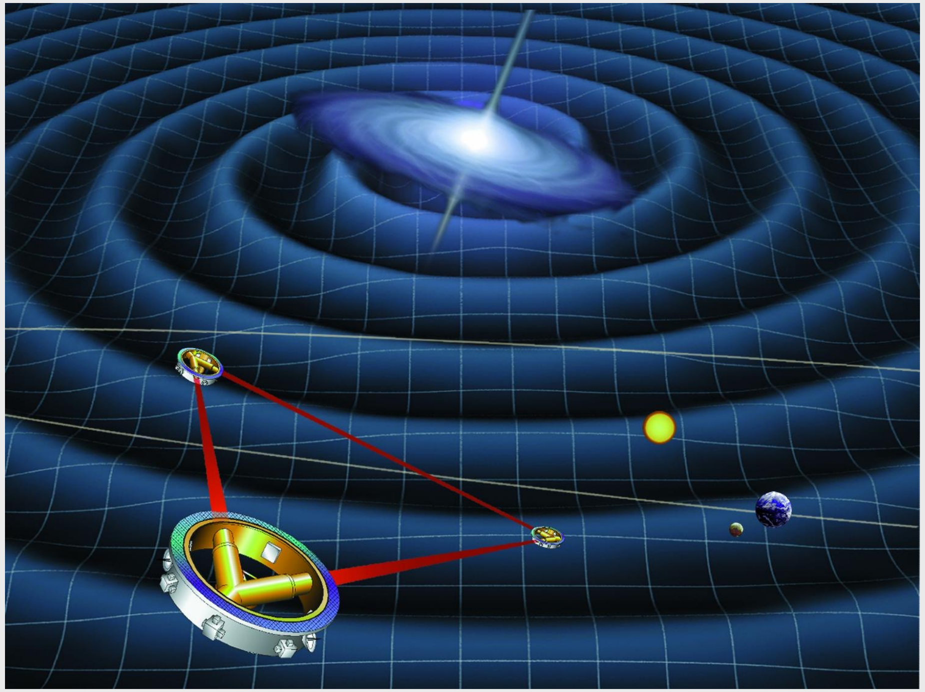
Figure 1. The upper half of the figure illustrates the complete pipeline and interaction of the various (sub)systems, mentioned in Section 2, responsible for disseminating early warning alerts. The waveform evolution with time is shown in the bottom half along with the dependence of the sky-localization area on the cutoff time of the early warning templates and the accumulated S/N during the binary inspiral. The waveforms, time to merger, S/N, and localizations in this figure are qualitative.

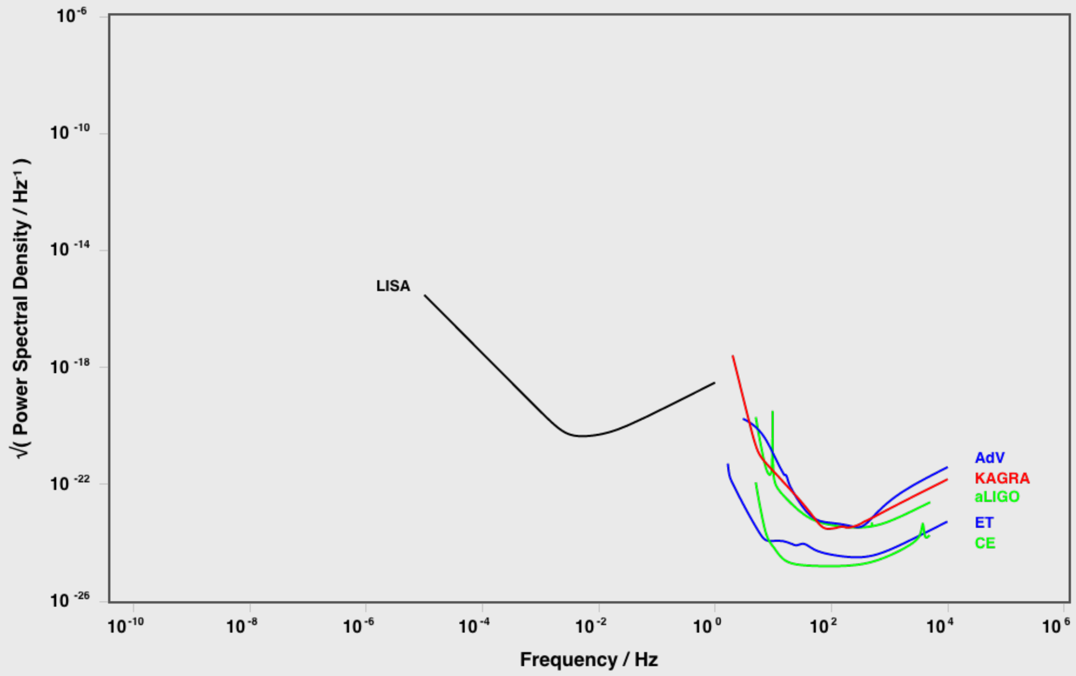
Meanwhile, in the US...



- Cosmic Explorer: proposed 3G facility in the US
- More conventional (and less risky) surface L-shaped design
- 40km + 20km arms
- Better sensitivity in mid-frequency range
- Higher (worse) low-frequency cutoff
- We'll have a 3G ground-based network!

...and farther away: LISA





- Space-based (solar orbit) triangular laser interferometer
- 2.5 million km arms, launch planned in 2034
- Very low frequencies: complementary to ground-based interferometers
- Very massive (astrophysical) black holes, very early alerts for BNS, ultra-compact binaries, extreme mass ratio inspirals, precision tests of GR,...

The multimessenger ecosystem

A large number of existing and future facilities will produce and consume alert triggers. For example, ESFRI-only:

Groud-based optical telescopes, like the Extremely Large Telescope: 5-mirror 39m optical telescope for the ESO on Cerro Armazones, Chile.

Large radiotelescope arrays, like the Square Kilometer Array: huge multi-band radiotelescope arrays in Africa and Australia

Facilities for cosmic ray astronomy, like the Cherenkov Telescope Array: Cherenkov telescopes for highest-energy gamma-ray astronomy, in the Canary Islands and Chile

Neutrino detectors, like KM3NeT: underwater network of neutrino detectors in the Mediterranean

...and many more.

The multimessenger ecosystem

- Several new, large EM and astroparticle facilities coming of age in roughly the same time frame
  - Several neutrino facilities (JUNO, DUNE, Hyperkamiokande,...)
  - Space and groud-based EM instruments (LOFAR, QTT, Vera Rubin, James Webb, Euclid, Nancy Roman, gamma and x instruments on satellites,...)
  - They will all have very similar low-latency alert requirements, as producers or consumers (or both)
  - High rates will imply extreme automation in the generation and selection of triggers, and sophisticated scheduling algorithms
  - How will the 2030's heir to today's NASA GCN work?
  - Will there be a MM-specific (virtual) shared infrastructure like the WLCG?

What will ML look like 10 years from now?

nature astronomy

ARTICLES

https://doi.org/10.1038/s41550-021-01405-0

Check for updates

Accelerated, scalable and reproducible AI-driven gravitational wave detection

E. A. Huerta [1,2], Asad Khan [3], Xiaobo Huang[3], Minyang Tian[3], Maksim Levental[2], Ryan Chard[1], Wei Wei[3], Maeve Heflin[3], Daniel S. Katz[3], Volodymyr Kindratenko[3], Dawei Mu[3], Ben Blaiszik[1,2] and Ian Foster[1,2]

The development of reusable artificial intelligence (AI) models for wider use and rigorous validation by the community promises to unlock new opportunities in multi-messenger astrophysics. Here we develop a workflow that connects the Data and Learning Hub for Science, a repository for publishing AI models, with the Hardware-Accelerated Learning (HAL) cluster, using funcX as a universal distributed computing service. Using this workflow, an ensemble of four openly available AI models can be run on HAL to process an entire month's worth (August 2017) of advanced Laser Interferometer Gravitational-Wave Observatory data in just seven minutes, identifying all four binary black hole mergers previously identified in this dataset and reporting no misclassifications. This approach combines advances in AI, distributed computing and scientific data infrastructure to open new pathways to conduct reproducible, accelerated, data-driven discovery.

Gravitational waves were added to the growing set of detectable cosmic messengers in the fall of 2015 when the advanced Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors reported the observation of gravitational waves consistent with the collision of two massive, stellar-mass black holes[1]. Over the last five years, the advanced LIGO and advanced Virgo detectors have completed three observing runs, reporting over 50 gravitational wave sources[2,3]. As advanced LIGO and advanced Virgo continue to enhance their detection capabilities and other detectors join the international array of gravitational wave detectors, it is expected that gravitational wave sources will be observed at a rate of several per day[4].

An ever-increasing catalogue of gravitational waves will enable systematic studies to advance our understanding of stellar evolution, cosmology, alternative theories of gravity, the nature of supranuclear matter in neutron stars, and the formation and evolution of black holes and neutron stars, among other phenomena[5–11]. Although these science goals are feasible in principle given the proven detection capabilities of astronomical observatories, it is equally true that established algorithms for the observation of multi-messenger sources, such as template-matching and nearest-neighbour algorithms, are compute-intensive and poorly scalable[12–14]. Furthermore, available computational resources will remain oversubscribed, and planned enhancements will be outstripped rapidly with the advent of next-generation detectors within the next couple of years[15]. Thus, an urgent rethink is critical if we are to realize the multi-messenger astrophysics program in the big-data era[16].

To contend with these challenges, a number of researchers have been exploring the application of deep learning and of computing accelerated by graphics processing units (GPUs). Co-authors of this article pioneered the use of deep learning and high-performance computing to accelerate the detection of gravitational waves[17,18]. The first generation of these algorithms targeted a shallow signal manifold (the masses of the binary components) and required only tens of thousands of modelled waveforms for training, but these models served the purpose of demonstrating that an alternative method for gravitational wave detection is as sensitive as template matching and significantly faster, at a fraction of the computational cost.

Research and development in deep learning is moving at an incredible pace[19–37] (see also ref. [38] for a review of machine-learning applications in gravitational wave astrophysics). Specific milestones in the development of artificial intelligence (AI) tools for gravitational wave astrophysics include the construction of neural networks that describe the four-dimensional (4D) signal manifold of established gravitational wave detection pipelines, that is, the masses of the binary components and the z component of the three-dimensional spin vector in (m_1, m_2, s_1^z, s_2^z). This requires the combination of distributed training algorithms and extreme-scale computing to train these AI models with millions of modelled waveforms in a reasonable amount of time[30]. Another milestone concerns the creation of AI models that enable gravitational wave searches over hour-long datasets, keeping the number of misclassifications at a minimum[39].

In this article, we introduce an AI ensemble, designed to cover the 4D signal manifold (m_1, m_2, s_1^z, s_2^z), to search for and find binary black hole mergers over the entire month of August 2017 in advanced LIGO data[40]. Our findings indicate that this approach clearly identifies all black hole mergers contained in that data batch with no misclassifications. To conduct this analysis we used the Hardware-Accelerated Learning (HAL) cluster deployed and operated by the Innovative Systems Laboratory at the National Center for Supercomputing Applications. This cluster consists of 16 IBM SC922 POWER9 nodes, with four NVIDIA V100 GPUs per node[41]. The nodes are interconnected with an EDR InfiniBand network, and the storage system is made of two DataDirect Networks all-flash arrays with SpectrumScale file system, providing 250 TB of usable space. Job scheduling and resource allocation are managed by the SLURM (Simple Linux Utility for Resource Management) system. As we show below, we can process data from the entire month of

[1]Data Science and Learning Division, Argonne National Laboratory, Lemont, IL, USA. [2]University of Chicago, Chicago, IL, USA. [3]University of Illinois at Urbana-Champaign, Urbana, IL, USA. e-mail: elihu@anl.gov

ML is not yet a mainstream "tool of the trade" in GW, but a huge lot of R&D is already ongoing

- Efficiency & speed
  - Signal Classification
  - Parameter estimation
  - Noise glitch hunting
  - (Template bank generation)
- Technology exploitation
  - Use advanced hardware (GPU, TPU...)
  - FPGAs / custom hardware
- Automatization
  - Automatize standard procedure for Data Quality
  - Automated de-noising with synthetic noise from GANs

The mandatory slide with boxes and arrows

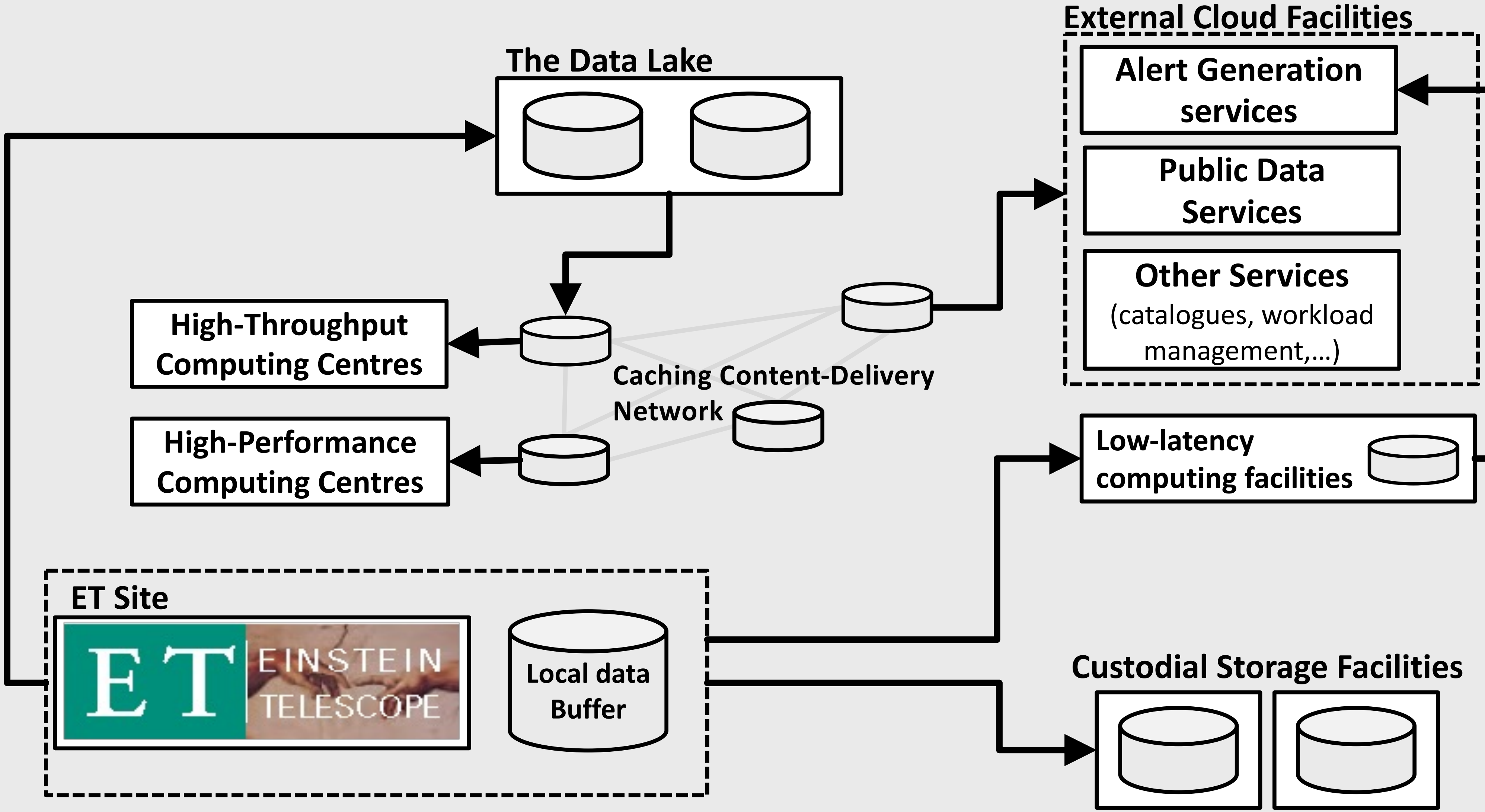


Figure from the ET ESFRI proposal

The mandatory slide with boxes and arrows

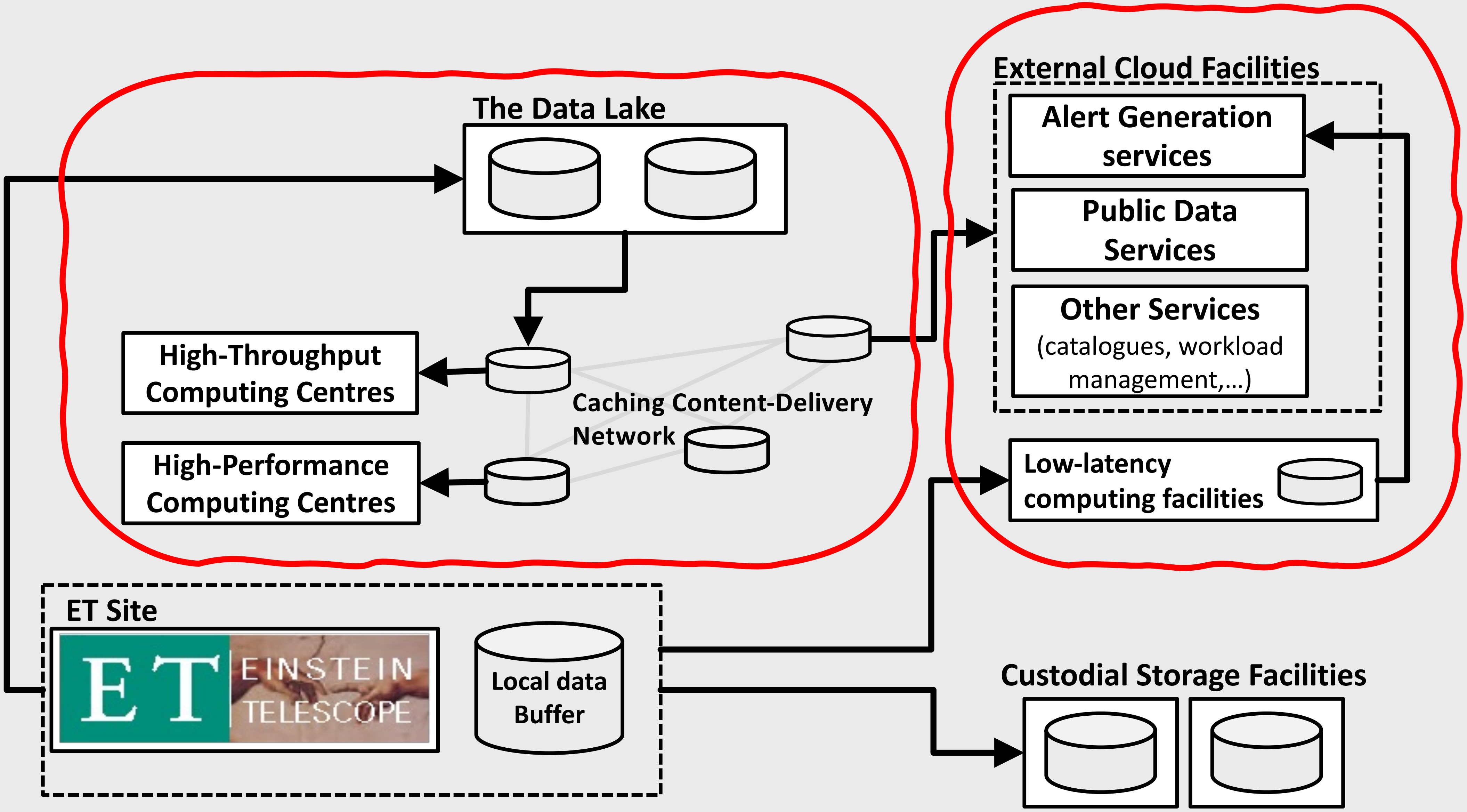


Figure from the ET ESFRI proposal

Lots of work ahead

- Down-sampling of the data stream for long duration events
- Hierarchical methods and “decimation” strategies
- Technology tracking of leading-edge technologies
  - Artificial Intelligence and Machine Learning
  - CUDA GPUs and HPC (FPGA and fancier architectures such as TPUs still to be tested)
  - Role of HPC is expected to grow with the SNR (Numerical Relativity, template bank production) and role of ML
  - Quantum computing!
- Early Mock Data Challenges to develop and validate everything
- Low-latency alert management and exchange system
- …

The ET e-Infrastructure Board

Division 1: Software, frameworks, and data challenge support

Define the software frameworks for ET computing workflows, the middleware for infrastructure, workload and data management. Develop software quality best practices and support their adoption with training and enforcement policies. Support code development in all computing domains. Provide computing support for mock-data challenges.

Division 2: Services and Collaboration Support

Define and provide all the IT services needed for the administrative management of the Collaboration. Define and provide all the IT services needed for communication and collaboration within the Collaboration and outside.

Division 3: Computing and data model, Resource Estimation

Develop the Einstein Telescope Computing Model. Provide a running estimate of the computing resources needed for all computing domains.

Division 4: Multimessenger alerts infrastructure

Design and develop the infrastructure needed for multi-messenger triggers management and distribution. Follow the development of software tools for low-latency computing.

TTG: Technology Tracking working Group

Follow the evolution of hardware and software computing technologies. Organize regular occasions for inter-division updates.

Summary

- Einstein Telescope will provide greatly enhanced sensitivities to GW detection, enabling a very rich scientific program in astrophysics and fundamental physics
- ET has to tackle many challenges in the fields of computing and algorithms development in order to meet its analysis needs
- ET needs to get involved in the discussion of common tools and services for Low Latency Alerts & Multi-Messenger Analysis
- The project timescale also allows to start innovative R&D to develop advanced technologies for the 2030's