ACAT 2024

Computing the Wave

Where the Gravitational Wave Community benefits from High-Energy Physics, and where it differs ?

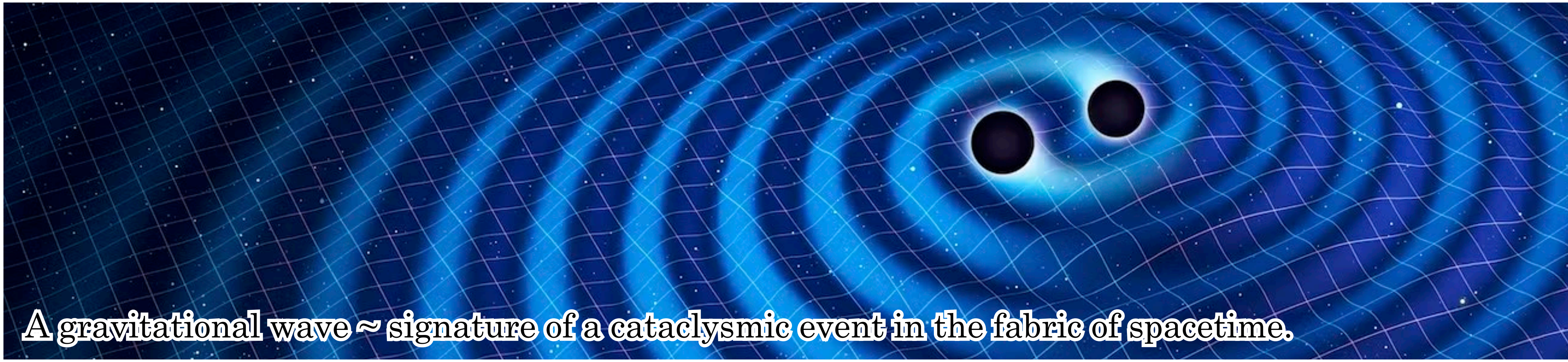
22nd International Workshop on Advanced Computing and Analysis Techniques in Physics Research
March 13th, 2024 - Stony Brook University

Marco Meyer-Conde[1,2,3], Nobuyuki Kanda[1]
Hirotaka Takahashi[2], Ken-Ichi Oohara[4], Kazuki Sakai[5]

On Behalf of KAGRA Collaboration

[1]Osaka Metropolitan University, [2]Tokyo City University,
[3]University of Illinois at Urbana-Champaign, [4]Niigata University, [5]NIT Nagaoka College

GRAVITATIONAL WAVES (GW) IN A NUTSHELL



[*e.g. two black holes (BH) or neutron stars (NS) colliding in space*]

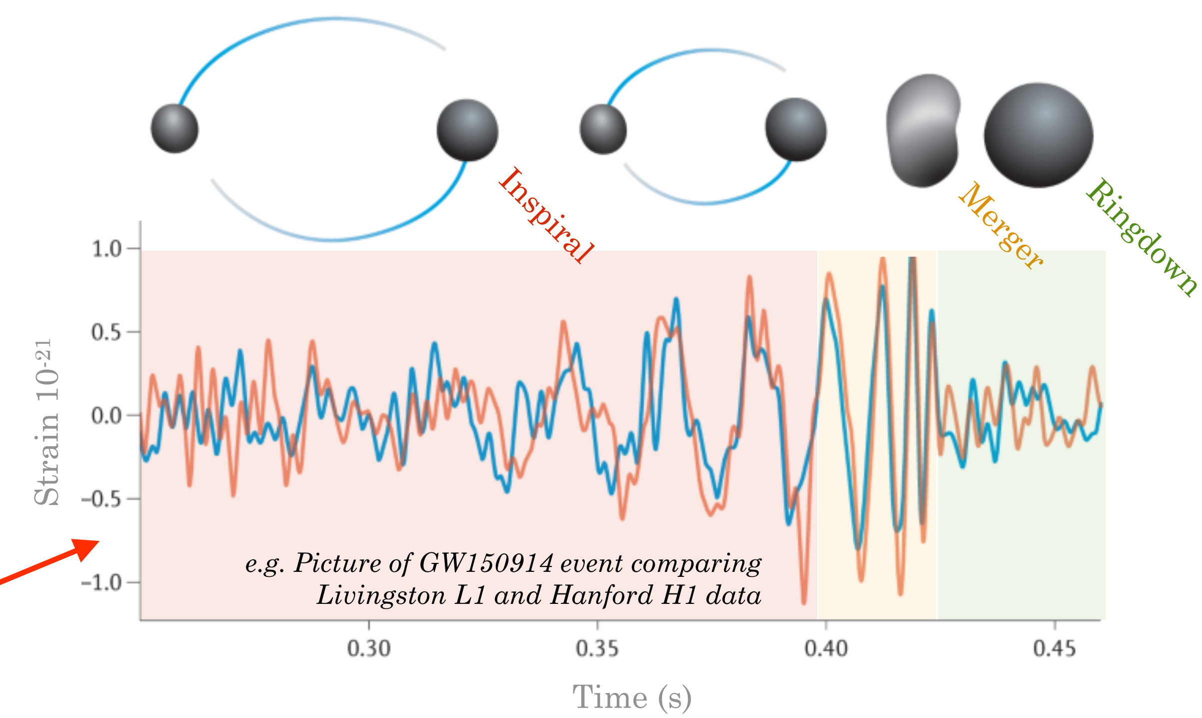
★ Key milestones in the History of GW Research:

1916 — Prediction of existence by A. Einstein

1974 — First indirect evidence: Hulse-Taylor Pulsar

90-00's — LIGO construction (Livingston & Hanford, US)

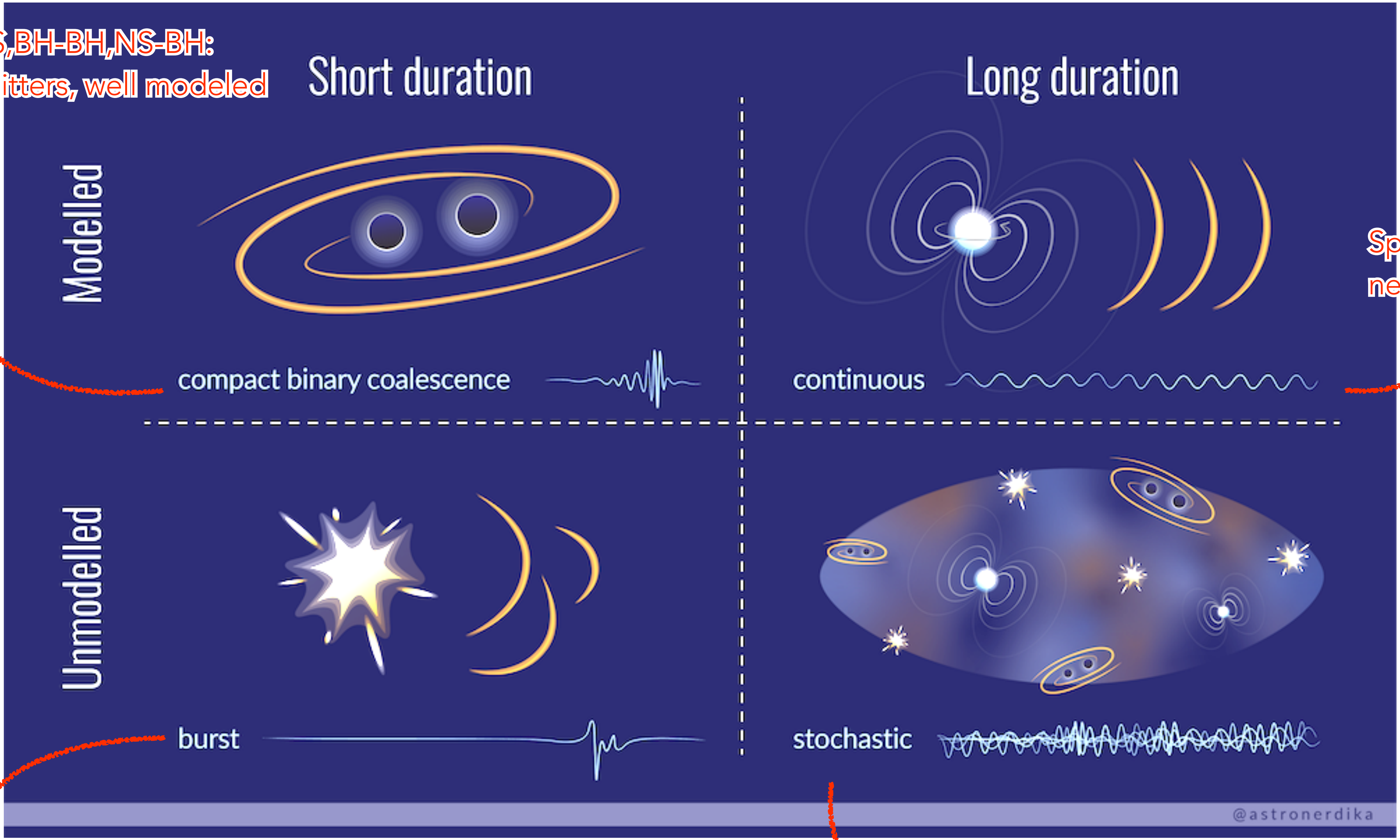
2015 — First direct BH-BH merger detection: GW150914



2017 — First NS-NS detection merger with Electromagnetic Counterpart: GW170817 / GRB170817A (New Era)

TOPOLOGY OF GRAVITATIONAL WAVES

NS-NS,BH-BH,NS-BH:
Strong emitters, well modeled

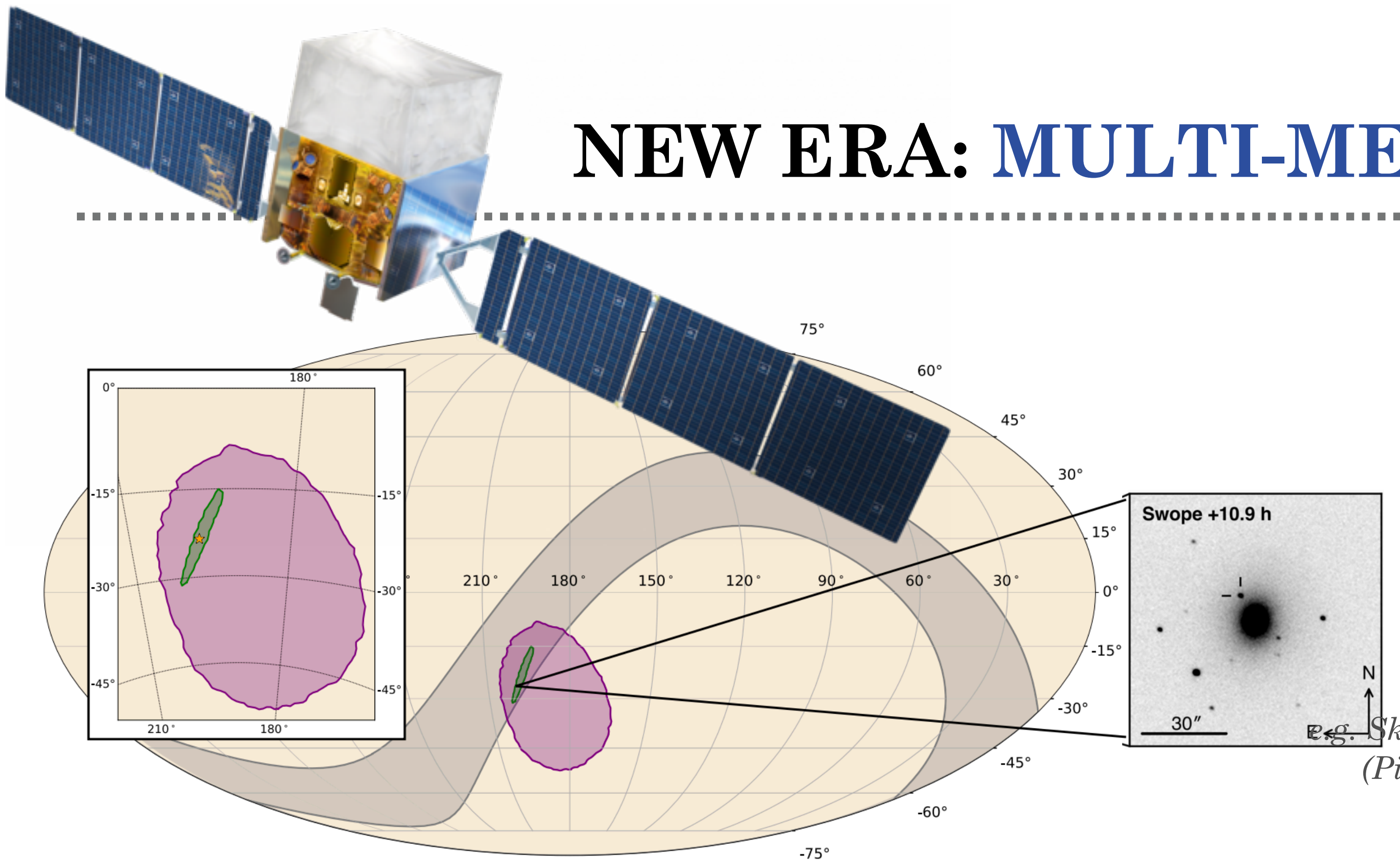


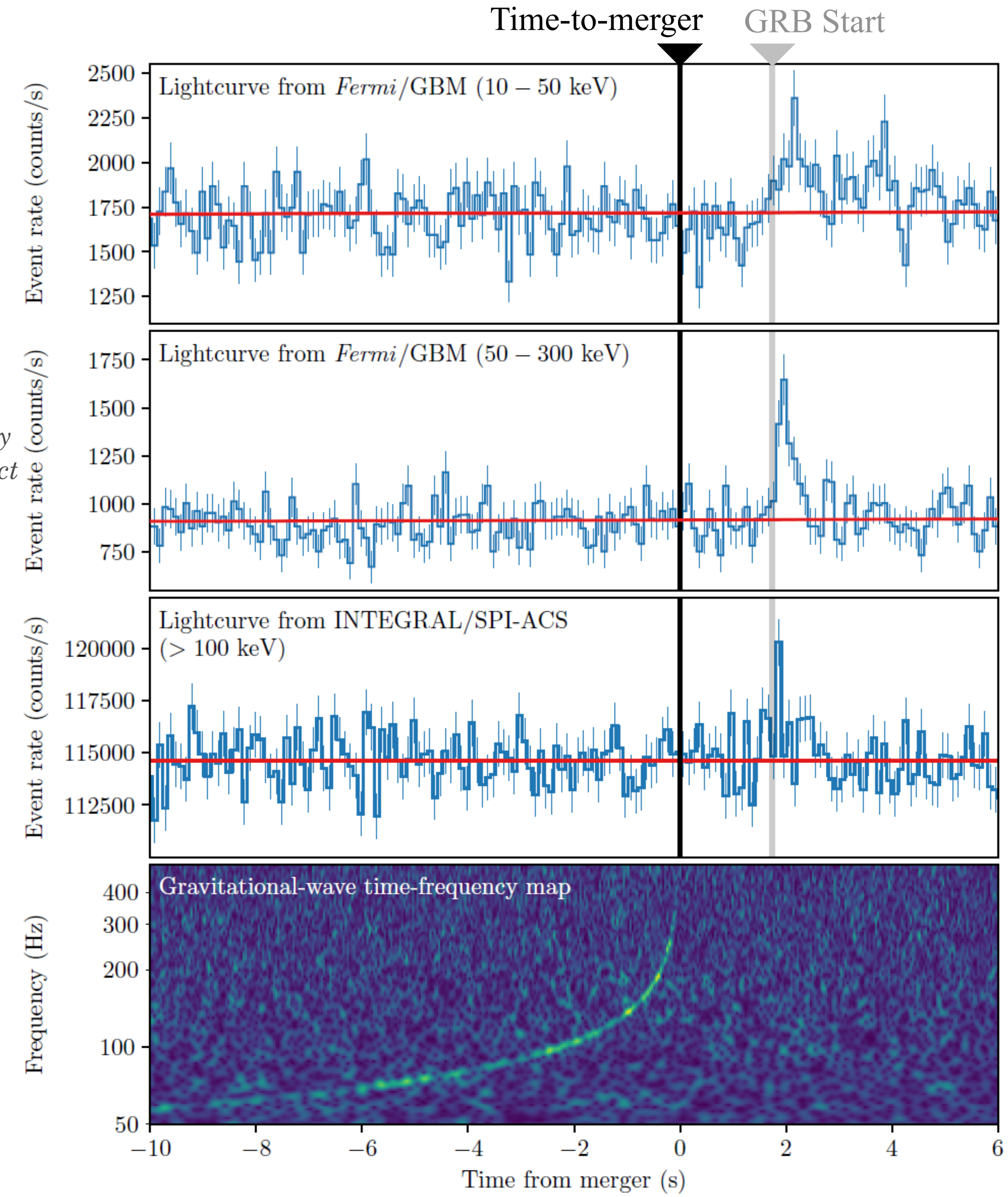
Spinning neutron star

Residual of Big-Bang, cosmic GW background, ..

Not well modeled, transients, weak signal, soft gamma-repeater,..

NEW ERA: MULTI-MESSENGER ASTRONOMY

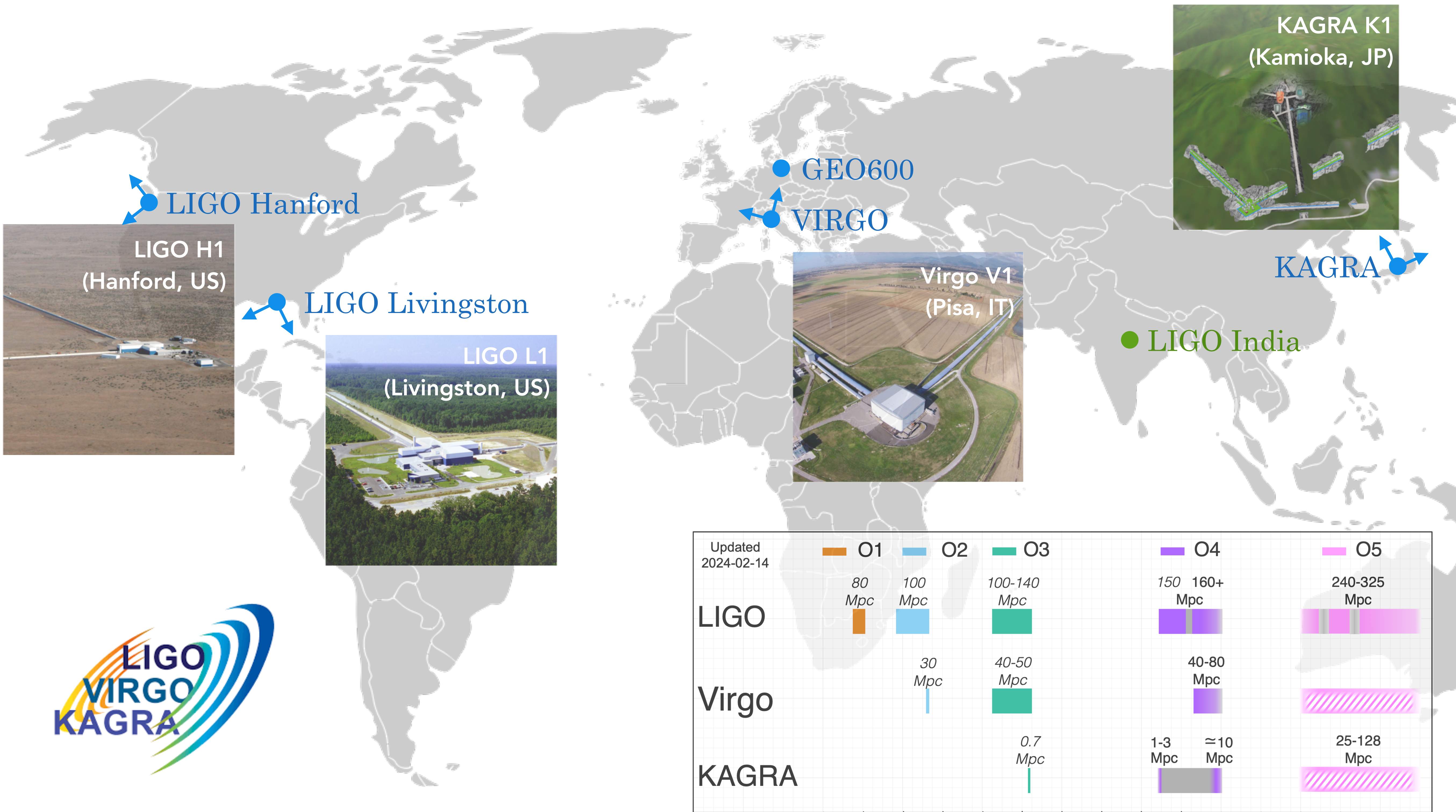




First NS-NS Detection with Electromagnetic Counterpart
(Landmark event: GW170817 / GRB170817A)

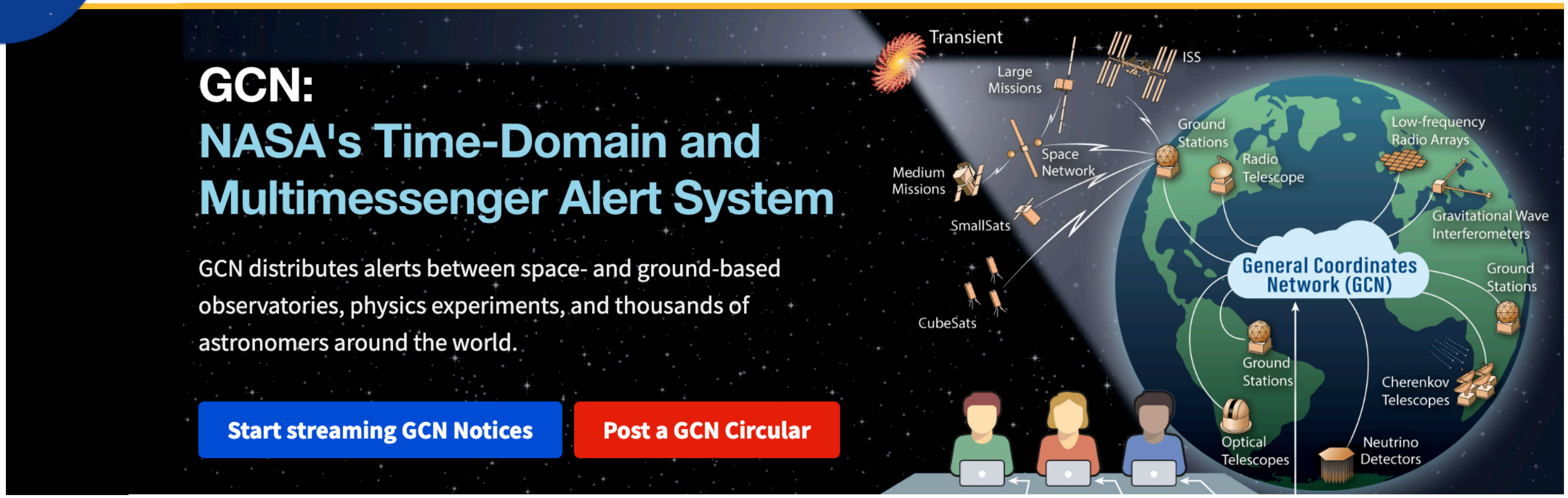
- Multi-messenger Astronomy (MMA) Era (started in 2017)
  - — IGWN: Ground-based coordination for low-latency alerts
  - — NASA: Coordination between space and astronomers.
  - — This challenge is about speed (< 60s) not size (~1Mb/s)
- Stereoscopic sky localization to be provided
  - — The more data, the better the localization is.
  - — Goal: GW signal with negative latency
    (meaning.. detection prior time-to-merger)
- Interest: connecting GW and HEP fields.
  - — Gravitational wave observatories
  - — Electromagnetic telescopes *(e.g. Fermi)*

INTL. GRAVITATIONAL WAVE NETWORK (IGWN)



e.g. Future projects: Einstein telescope, LISA interferometer, etc.

GENERAL COORDINATES NETWORK (GCN)

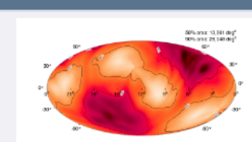
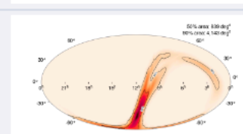
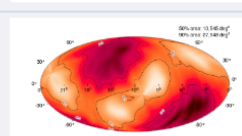
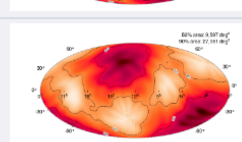


GraceDB

LIGO/Virgo/KAGRA Public Alerts .. March 1st, 2024

O4 Significant Detection Candidates: **81** (92 Total - 11 Retracted)

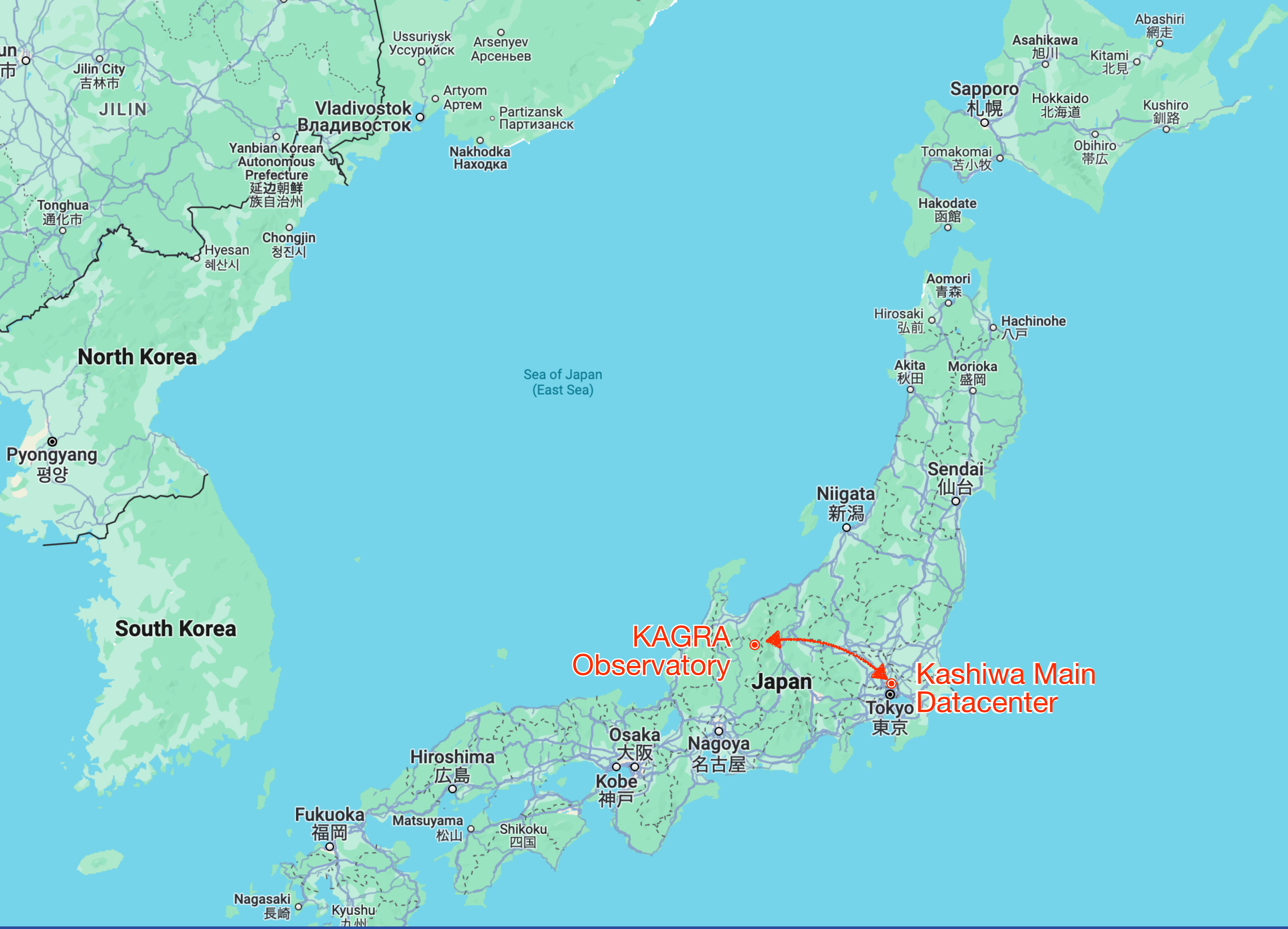
O4 Low Significance Detection Candidates: **1610** (Total)

Event ID	Possible Source	Significant	UTC	GCN	Location	FAR
S240109a	BBH (99%)	Yes	Jan. 9, 2024 05:04:31 UTC	GCN Circular Query Notices | VOE		1 per 4.3136 years
S240107b	BBH (97%) Terrestrial (3%)	Yes	Jan. 7, 2024 01:32:15 UTC	GCN Circular Query Notices | VOE		1.8411 per year
S240104bl	BBH (>99%)	Yes	Jan. 4, 2024 16:49:32 UTC	GCN Circular Query Notices | VOE		1 per 8.9137e+08 years
S231231ag	BBH (>99%)	Yes	Dec. 31, 2023 15:40:16 UTC	GCN Circular Query Notices | VOE		1 per 3.7932e+06 years

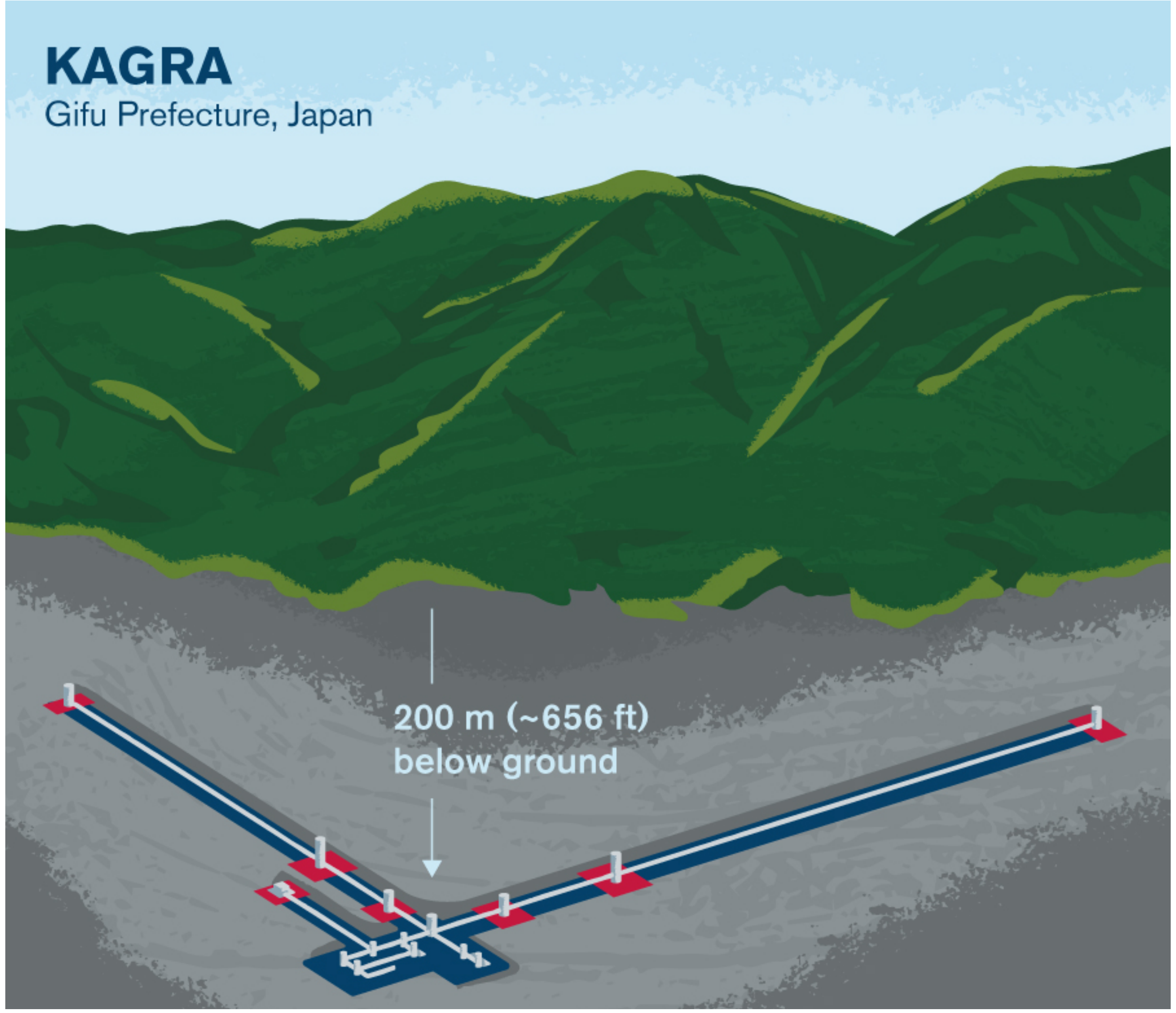
. . .

Gravitational-Wave Candidate Event Database (GraceDB)

- GraceDB — Centralized API for public alerts and sending GCN Circular Query
- Celery message broker — Full orchestration performed with GWCelery package
- Early public warning alerts — for rapid comms. *(https://gcn.gsfc.nasa.gov/)*


KAGRA
Observatory
Kashiwa Main
Datacenter
Japan
Tokyo
東京
Sea of Japan
(East Sea)
North Korea
South Korea
Pyongyang
평양
Jilin City
吉林市
JILIN
Vladivostok
Владивосток
Ussuriysk
Уссурийск
Arsenyev
Арсеньев
Artyom
Артем
Partizansk
Партизанск
Nakhodka
Находка
Yanbian Korean
Autonomous
Prefecture
延边朝鲜
族自治州
Tonghua
通化市
Hyesan
혜산시
Chongjin
청진시
Sapporo
札幌
Asahikawa
旭川
Abashiri
網走
Kitami
北見
Hokkaido
北海道
Kushiro
釧路
Obihiro
帯広
Tomakomai
苫小牧
Hakodate
函館
Aomori
青森
Hirosaki
弘前
Hachinohe
八戸
Akita
秋田
Morioka
盛岡
Sendai
仙台
Niigata
新潟
Osaka
大阪
Nagoya
名古屋
Kobe
神戸
Hiroshima
広島
Fukuoka
福岡
Matsuyama
松山
Shikoku
四国
Nagasaki
長崎
Kyushu

FOCUS ON KAGRA EXPERIMENT



KAGRA Observatory has been established in 2016 in Kamioka, Gifu prefecture

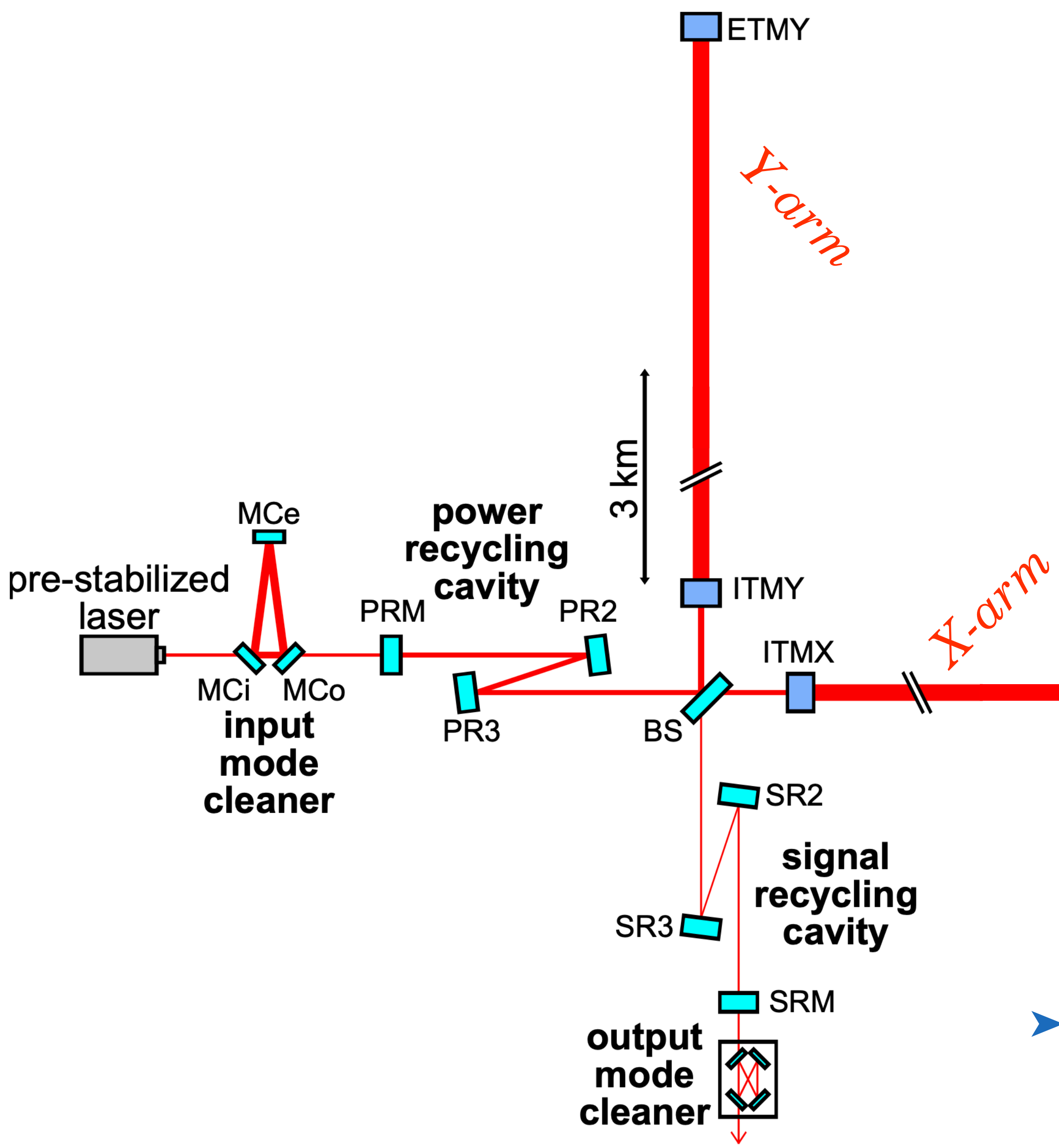
(same place as Kamiokande neutrino experiment)

The telescope is a 3 km Fabry-Perot interferometer
with 20K Cryogenic Sapphire Mirrors in Mt. Ikenoyama.

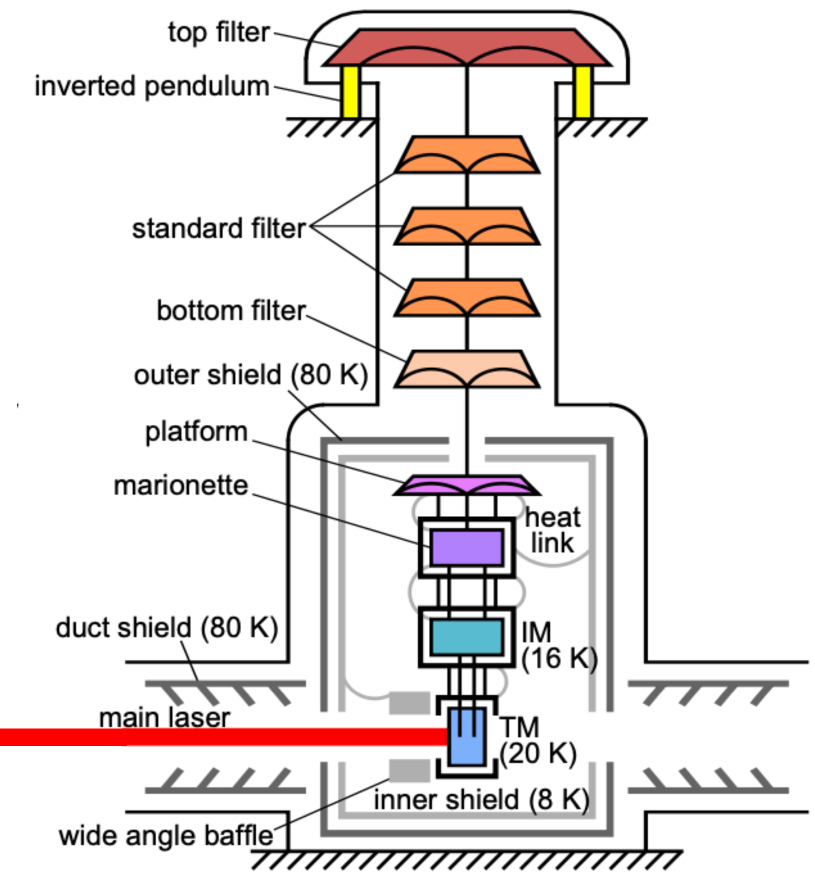
Field Racks (Sensors) → µs → Electronics Room (ADC converter) → ms → [ ] → [ ]

T. Akutsu *et al* 2020 *J. Phys.: Conf. Ser.* **1342** 012014
https://www.ligo.org/science/Publication-S6PE/index.php

ON-SITE DATA ACQUISITION (AT KAGRA)



Gravitational wave signal output



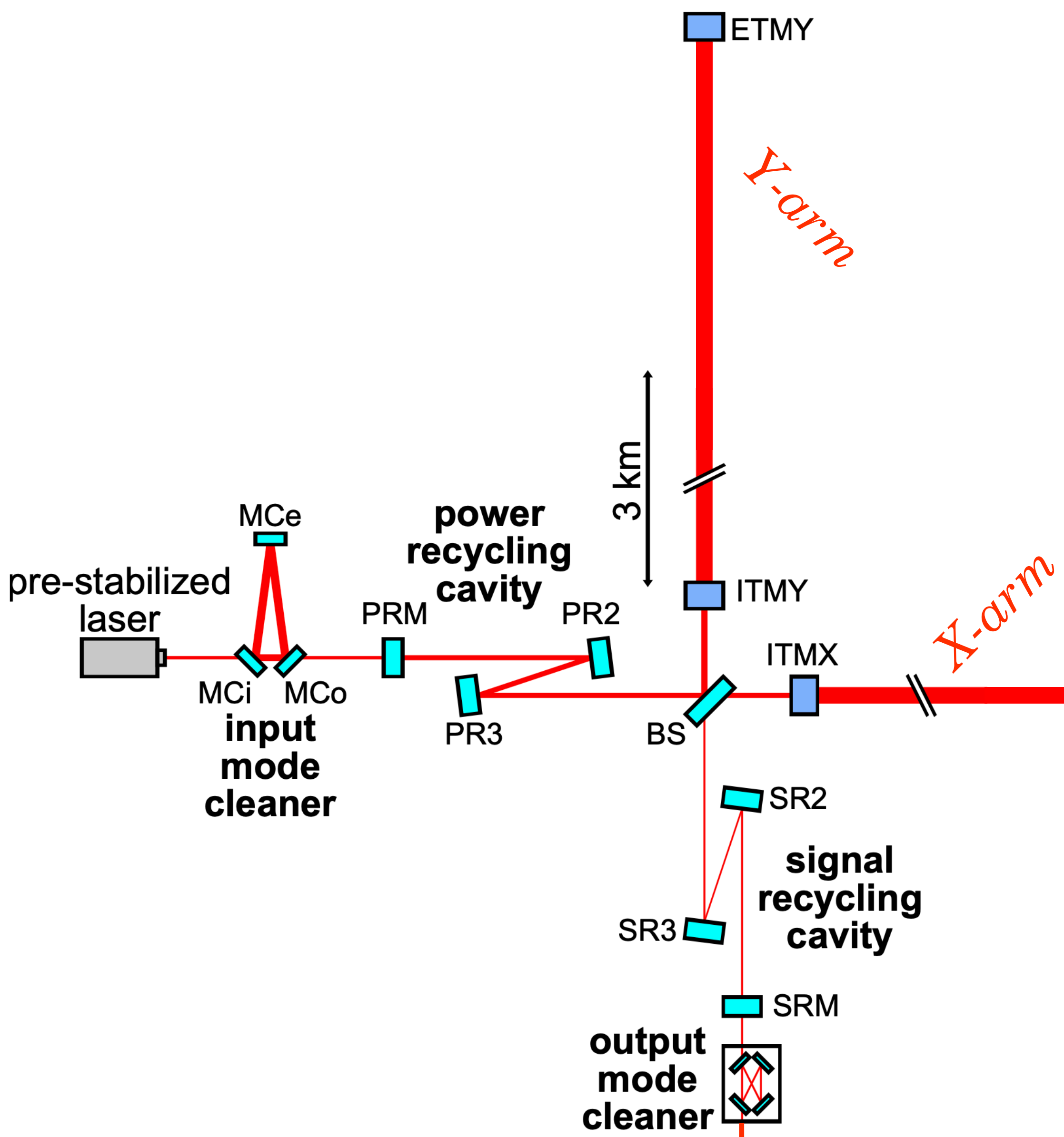
cryogenic suspension system for sapphire test masses

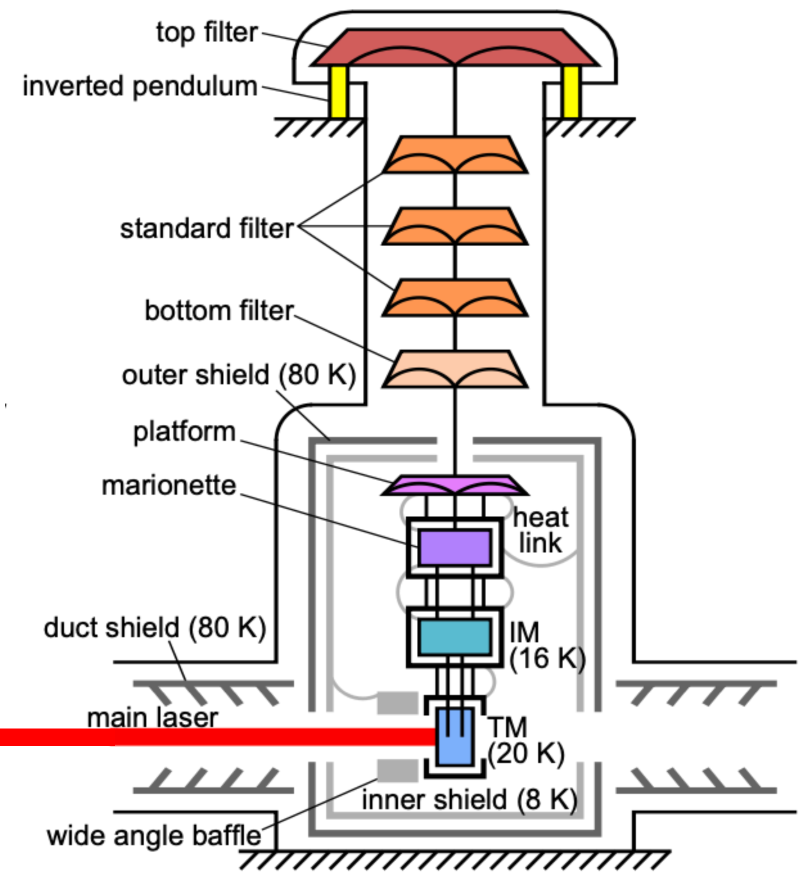
- Fabry-Perot Interferometer (power recycling)
  — Laser 180W continuous source at $\lambda = 1064$ nm
  — XY-arms, 20K-cryogenic cooled suspended mirrors
- On-site infrastructure actions
  — Data acquisition, pre-processing, environmental monitor.
  — Feedback control to actuators on ETMX,ETMY

Field Racks (Sensors) → µs → Electronics Room (ADC converter) → ms → →

T. Akutsu *et al* 2020 *J. Phys.: Conf. Ser.* **1342** 012014
https://www.ligo.org/science/Publication-S6PE/index.php

ON-SITE DATA ACQUISITION (AT KAGRA)



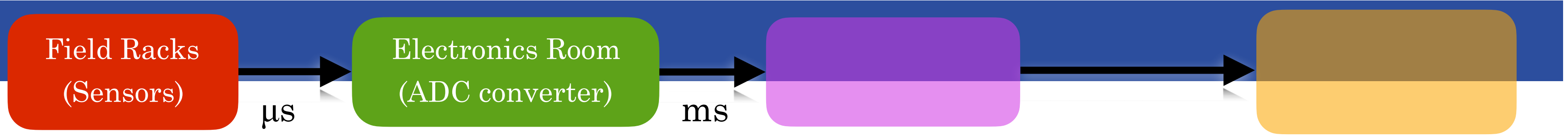


cryogenic suspension system for sapphire test masses

Raw data XY-arm → calibration pipeline → DARM Channel Signal f_s = 16kHz max.

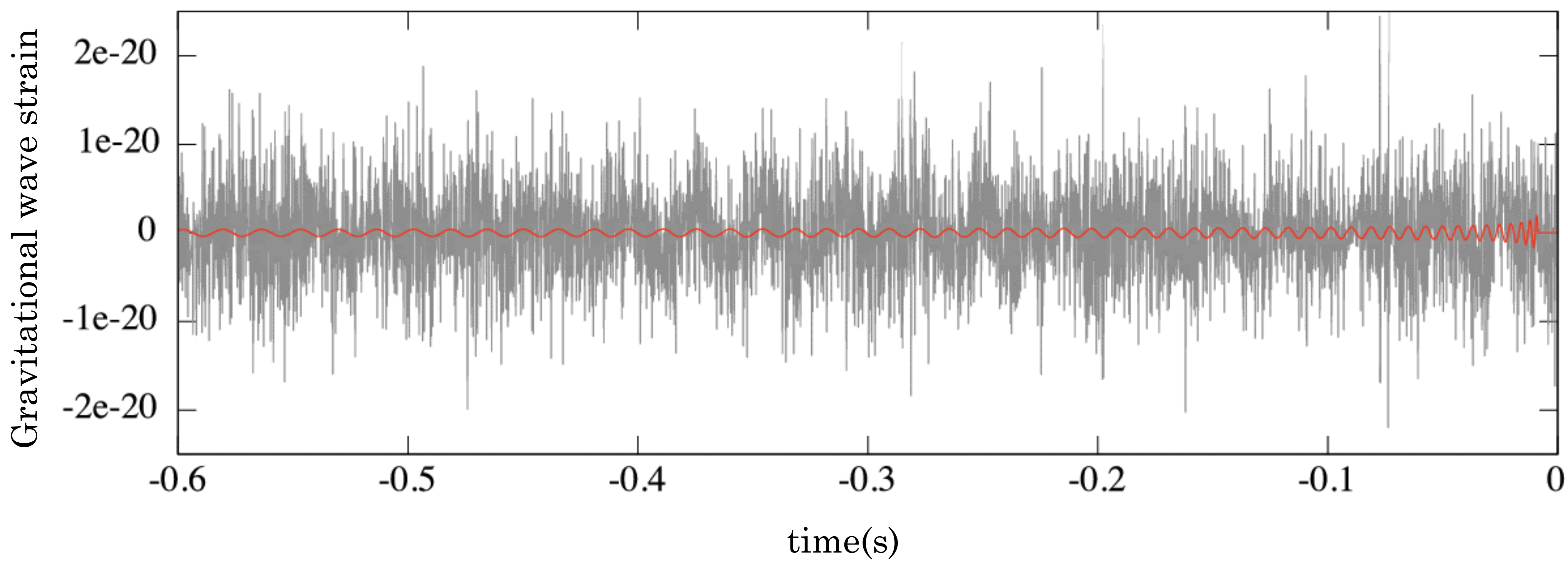
→ High-latency offline analysis

→ Low-latency online analysis!



T. Akutsu *et al* 2020 *J. Phys.: Conf. Ser.* **1342** 012014
https://www.ligo.org/science/Publication-S6PE/index.php

ON-SITE DATA ACQUISITION (AT KAGRA)

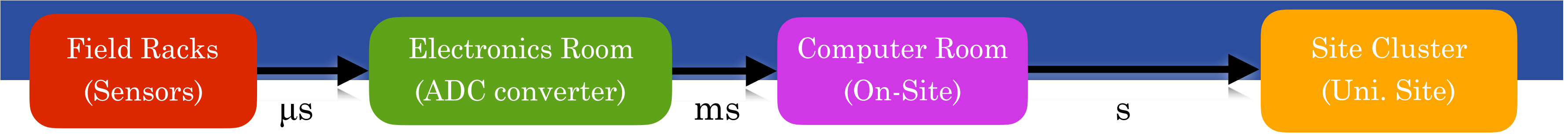


- Gravitation wave signal (here is LIGO experiment); typical signal injection in red
- Data transfer via Apache Kafka (message broker): low-latency reliable sustained data-stream
- Typical latencies at Kashiwa Data center for “1 second” files:

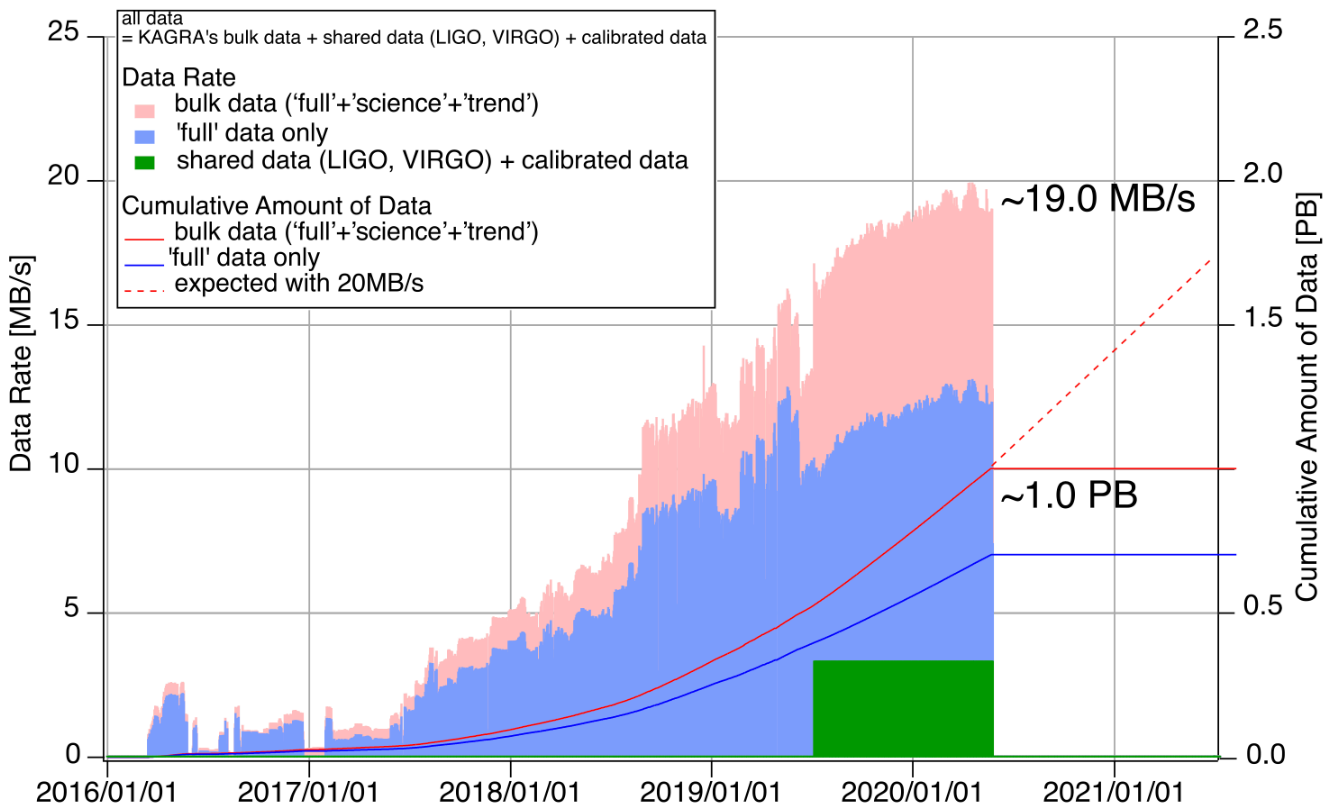
LIGO to Kashiwa	6-8s latency
VIRGO to Kashiwa	10s latency
KAGRA to Kashiwa	2.5s latency

(Latency = transfer from ADC to on-site servers + FIR filtering + transfer from On-Site (L1,H1,V1,K1) to Kashiwa)

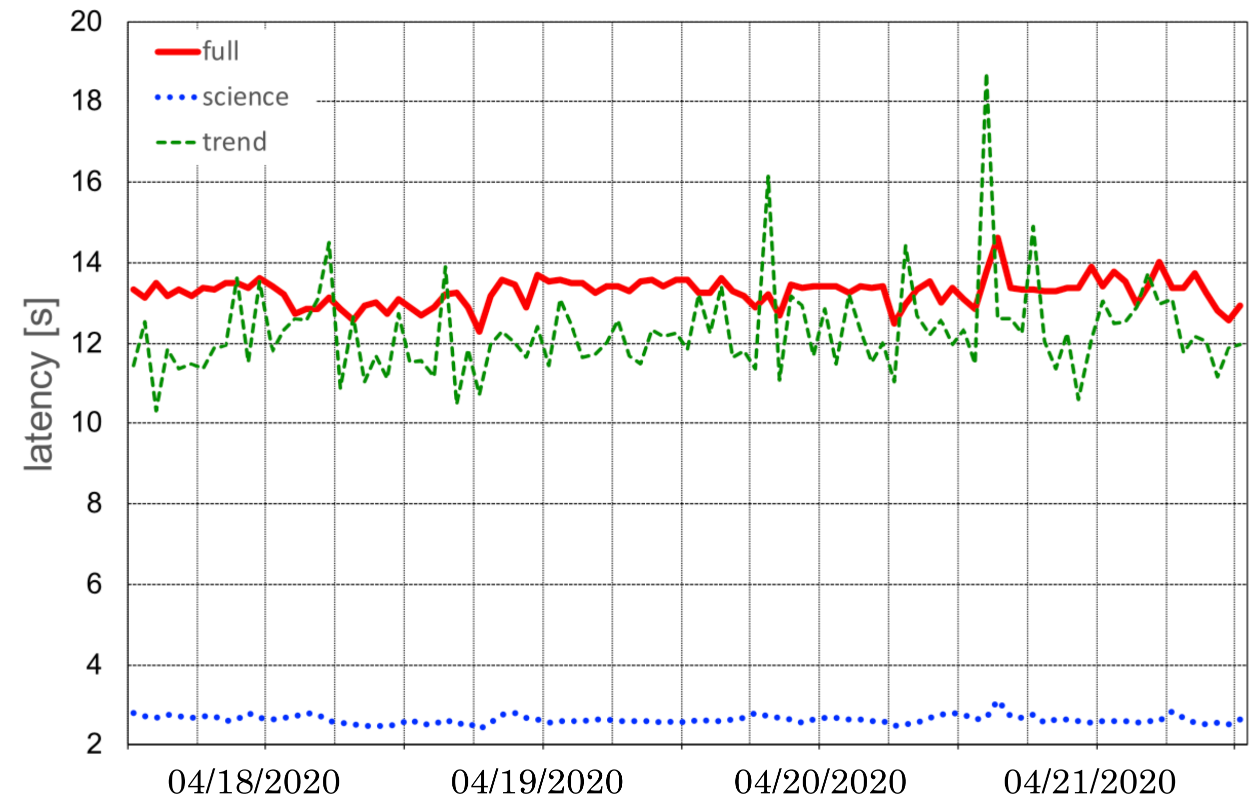
Apache kafka



FOCUS ON: KAGRA (KAMIOKA) TO KASHIWA MAIN CENTER (TOKYO)



(Data rate and quantity since beginning of KAGRA and projection for O4)

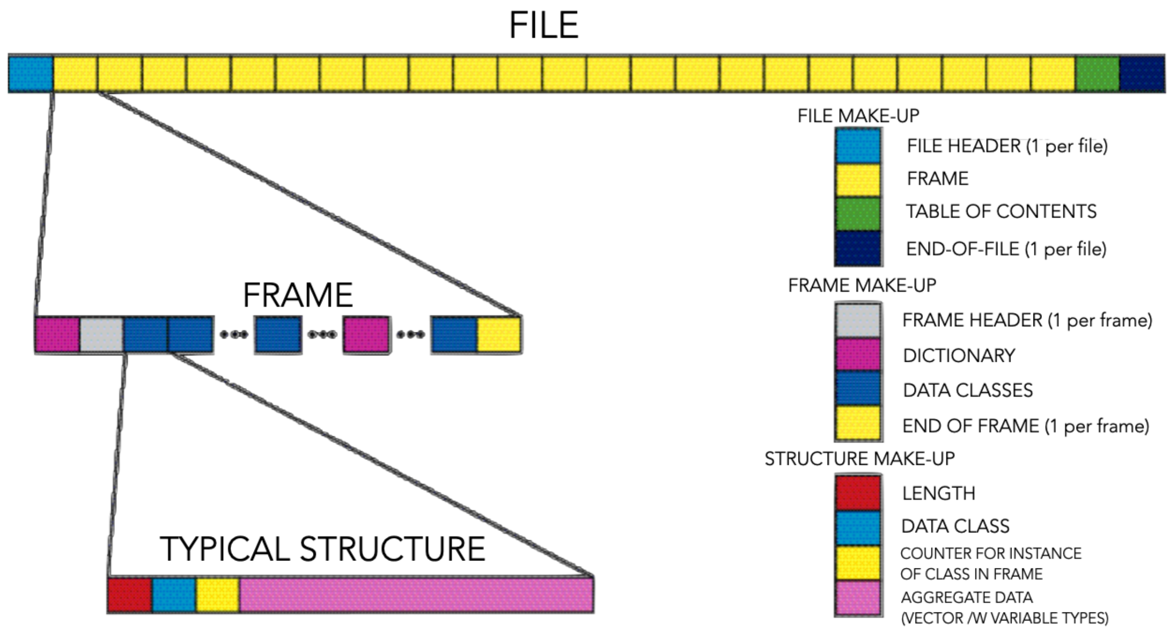


(Latency transfer of larger 32-seconds data files from KAGRA to Kashiwa)

- Low-latency Online Analysis O4a ~ O(1min)
  - — Data transfer of 1s-files via "Apache Kafka"
  - — Run O4a ~ 1.5PB/y of raw data
  - — Research pipeline and rapid sky localization
  - — Goal: Alert generation and flagging
- High-latency Offline Analysis
  - — Data management /w "Rucio" (GridFTP/CERN)
  - — Offline event search, detector characterization
  - — Make use of for sharing data:
    - * CVMFS (CERN) for "data" and "software"
    - * OSDF Data origin (OSG) for "data"

Prog. Theor. Exp. Phys. 2023 10A102(18 pages) DOI: 10.1093/ptep/ptad112
https://www.ligo.org/science/Publication-S6PE/index.php

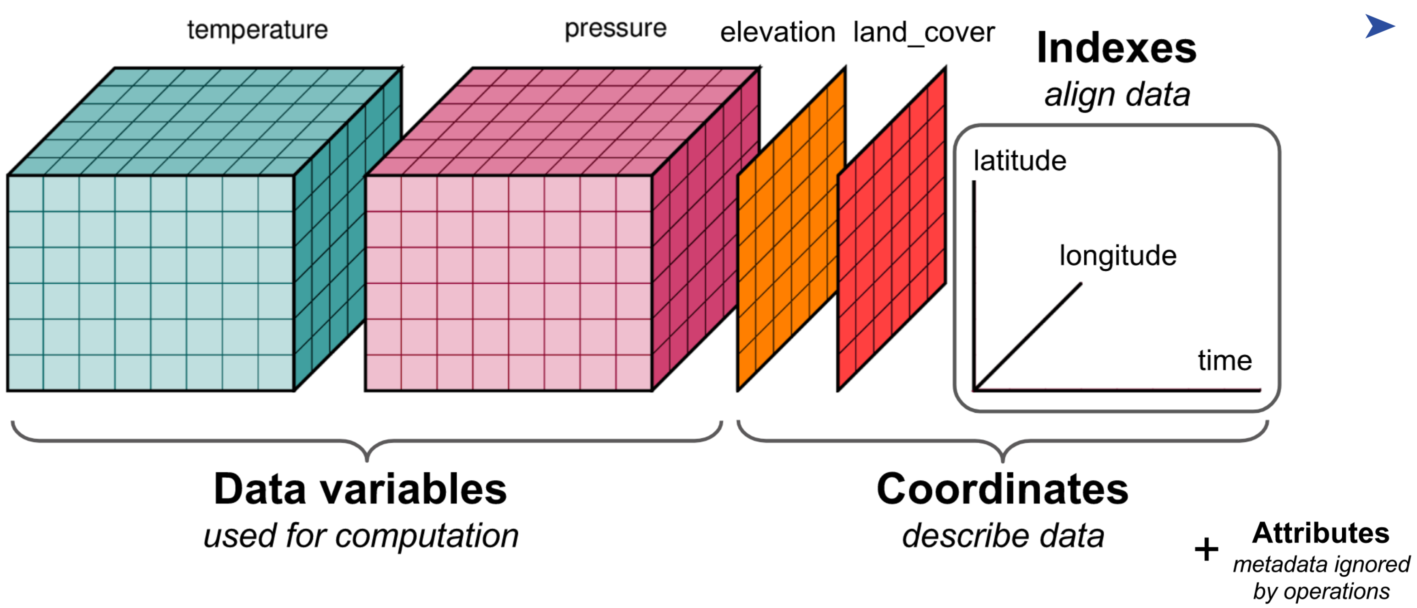
GRAVITATIONAL WAVE DATA FORMATS



— Picture of the Frame Format data structure for low latency —

- Low-latency data format: *.gwf
  - — "Frame Format" standard by LVK collaboration
  - — Since 2019, strain data h(t) are 1s frame size.
  - — Optimized format to fast reading/writing

Data Format specific to LVK collaboration

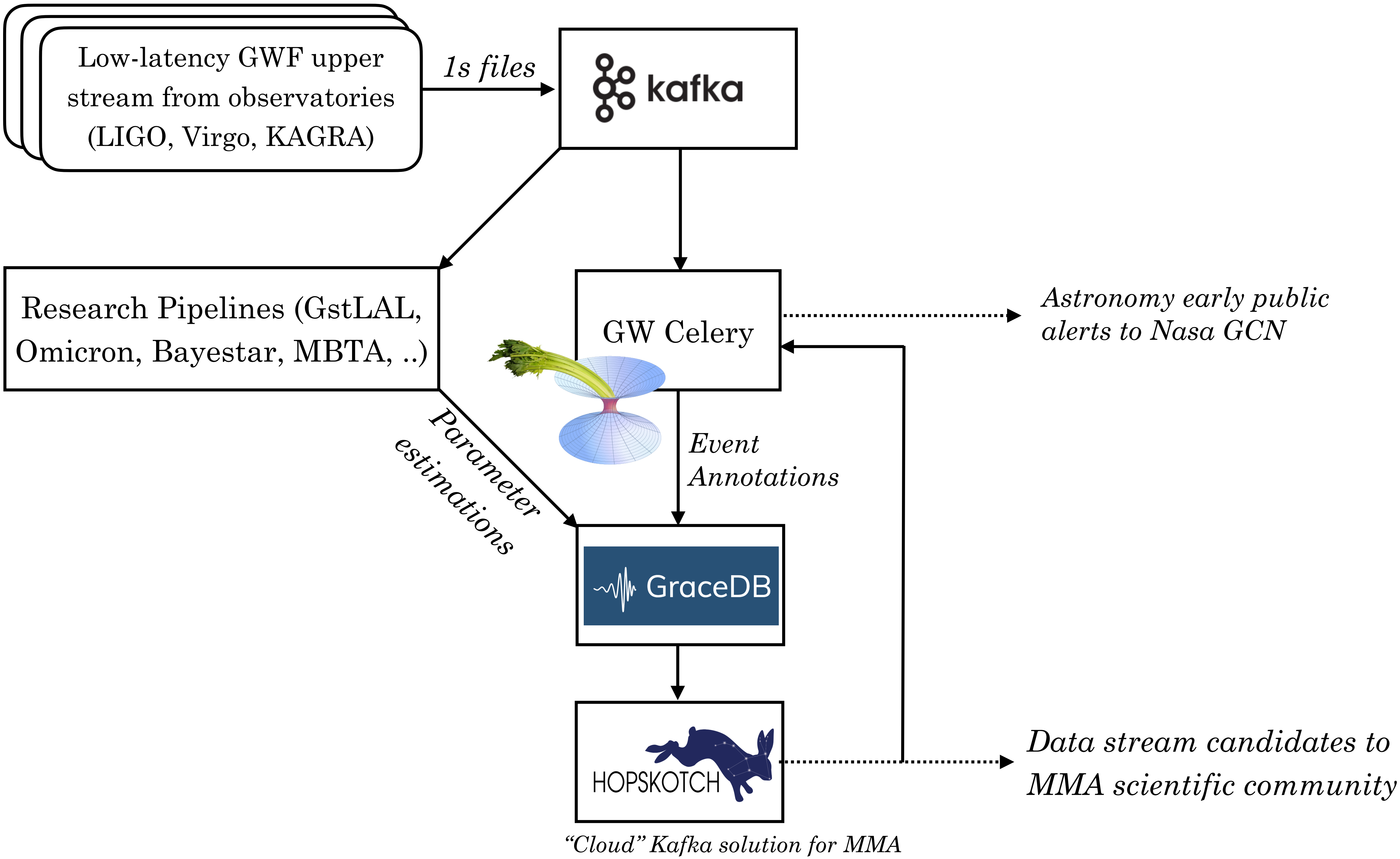


— Diagram of the HDF5 data format for high latency —

- High-latency data formats: *.gwf *.hdf5
  - — Usually HDF5 is used for Open Science Data
  - — More flexible when using with Python

(https://gitlab.cern.ch/escalade/standalone/hdf2root)

LOW-LATENCY ORCHESTRATION OVERVIEW

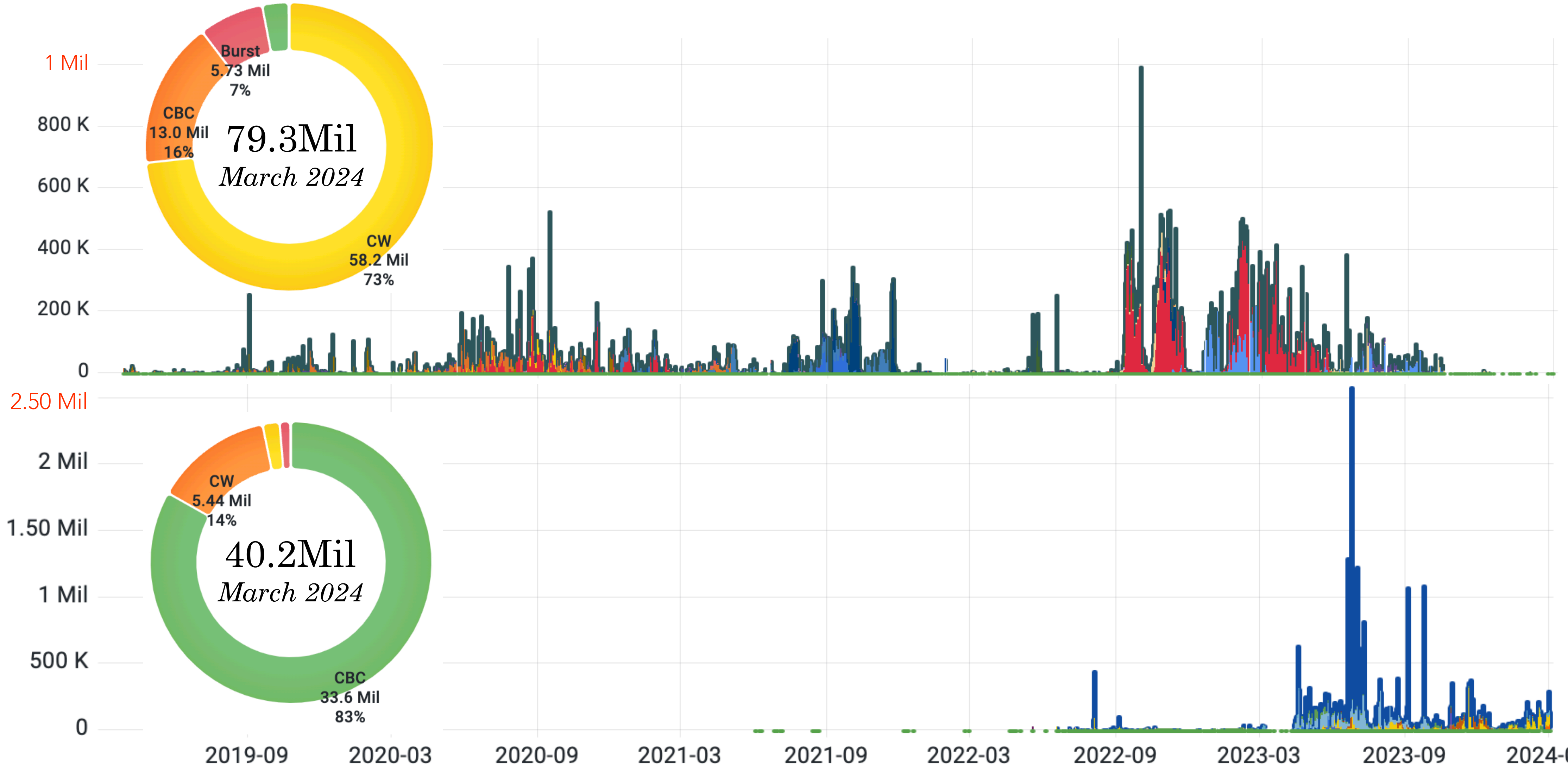


IGWN O3/O4 COMPUTING RESOURCES

- Observing run periods:
  - — O3a: April 1, 2019 to October 1, 2019
  - — O3b: November 1, 2019 to March 27, 2020
  - — O4a: May 24, 2023 to January 16, 2024
  - — O4b: April 3, 2024 to the end of year
- IGWN Computing Resources
  - — Computing policy and common coordination
  - — Code management: Gitlab, Conda & Docker
  - — Resource monitoring: Grafana
  - — No common framework for development.

(O3)

(O4)



COMPUTING PIPELINES AND COMMON FRAMEWORK

- IGWN Collaboration Goal: to reduce the amount of tools
  - — Interferometer sensitivity ↗ ; Detection range ↗ ; Smaller noise ↘ ; Signal-to-Noise-Ratio (SNR) ↗
  - — Raw data is not increasing much over years (main computing resources is about template generation)
  - — New technics developed in addition to the classic technics. (Matched Filtering, Wavelet, etc.)

- Computing pipelines are heterogeneous and complex (no common framework)
  - — Many languages: Matlab, C/C++, Python, ..
  - — Main scientific package: LVK Algorithm Library Suite (LALSuite; C/C++)
    [Noise simulation, template generation (GR modeling for matching data)]
  - — Machine learning increasing usage: First machine learning pipeline MLy is online.

ROOT
Data Analysis Framework

ROOT has already been used in GW community (mainly people in Europe, originating from HEP).

- Requirements to make use of ROOT in GW community
  - — (1) Handling typical I/O files: GWF files and HDF5 files
  - — (2) Advanced digital signal processing and (3) complex digital filtering

- Personal interests in using CERN modern/high-level tools (RNTuple, RDataFrame and SOFIE)
  - — investigation and development planned for RDataFrame to include TimeSeries capability (1 event = 1s Δt)

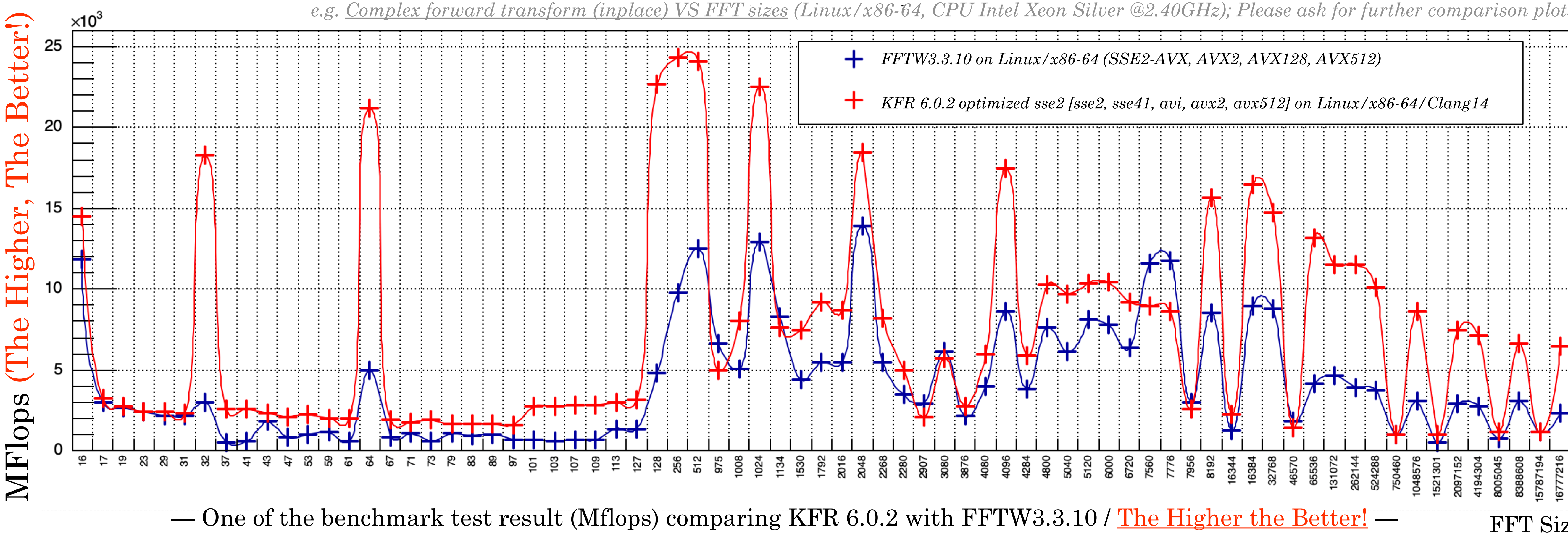
KFR https://www.kfrlib.com

FAST, MODERN C++ DSP FRAMEWORK

KFR is built for developers who need a powerful toolkit for audio and dsp applications

- **KFR library under GPL licensing.**
  — DFT & FFT at N-dimensions available
  — Advanced signal processing (IIR/FIR filtering)
- **Highly optimized modern C++ library**
  — optimal on x86-64 architectures using Clang
  — optimized perf. up to 40% in best cases vs. FFTW3
- **Compatible with both ARM64 and x86-64 archs.**
  — SIMD SSE, SSE2, AVX instructions, for x86_64
  — SIMD NEON64 instructions for ARM64.
- **Benchmark tests performed with last version.**
  — Comparison between KFR6.0.2 and FFTW3.3.10
  — https://git.ligo.org/kagra/containers/fft-benchmark

e.g. Complex forward transform (inplace) VS FFT sizes (Linux/x86-64, CPU Intel Xeon Silver @2.40GHz); Please ask for further comparison plots



— One of the benchmark test result (Mflops) comparing KFR 6.0.2 with FFTW3.3.10 / The Higher the Better! —

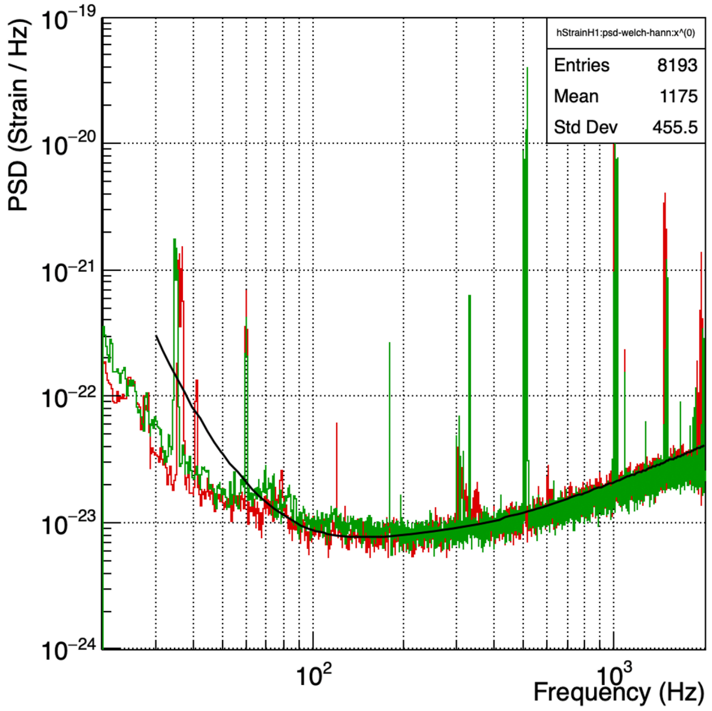
ROOT SIGNAL LIBRARY PROTOTYPE

- Advanced DFT features using KFR library.
  — ASD/PSD calculation, SNR calculation, Matched filtering,
  Convolution, Window, resampling, etc..

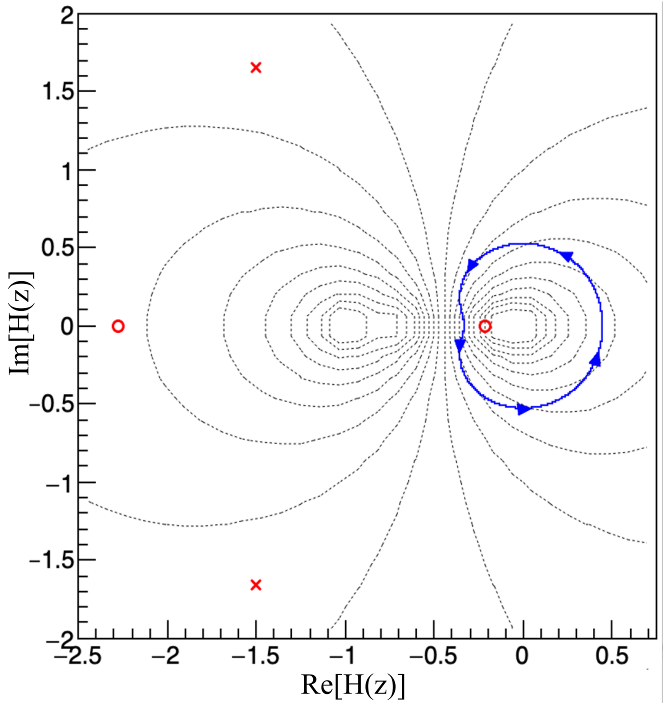
- Analog and digital filter H = B/A implementations. (ZPK, BA, SOS, SS representations)
  — Finite/Infinite Impulse Response filtering
  — Chebychev, Butterworth, Elliptical filter implemented, [..]

- TKFR Prototype class implemented with ROOT / Make use of the TVirtualFFT interface too
  ✅ ASD, PSD, Hilbert transform, Elliptical filters, Wiener & Kalman filter
  FIR/IRR, Windowing, Decimation, Waveform generation (sine, impulse, steps, ramps, etc.), *etc.*

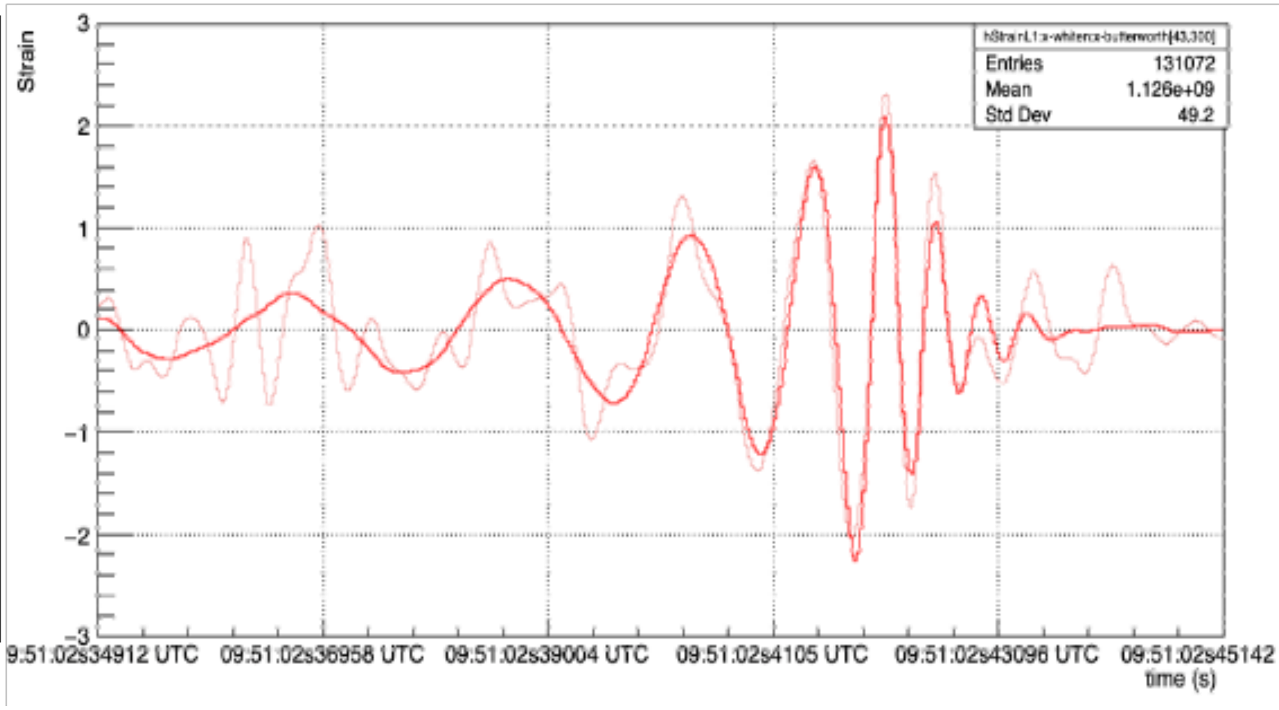
Power Spectrum Density calculation using GW150914 OpenScience data



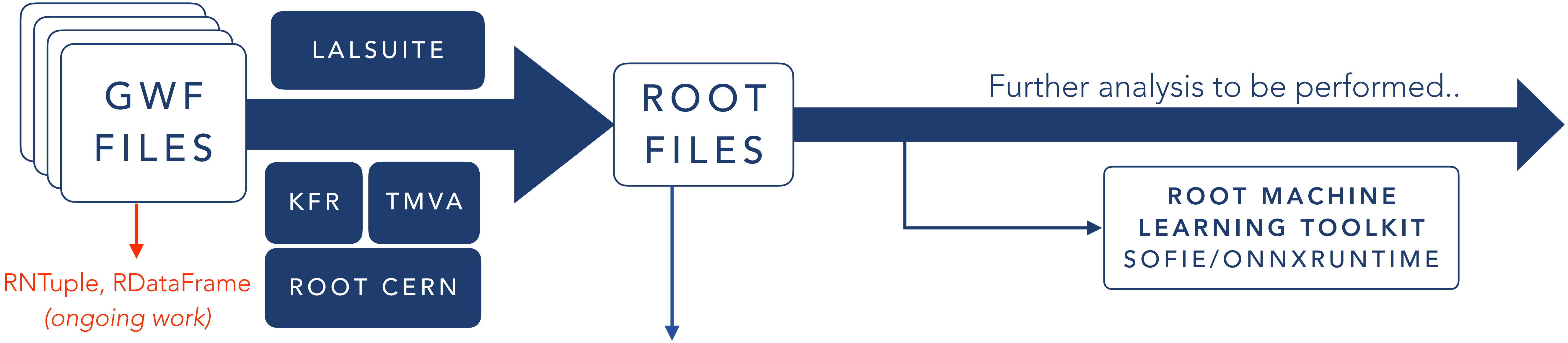
Example of digital filter implementation $H(z)=(5.0+3.0z+1.0z^2)/(1.0+5.0z+2.0z^2)$

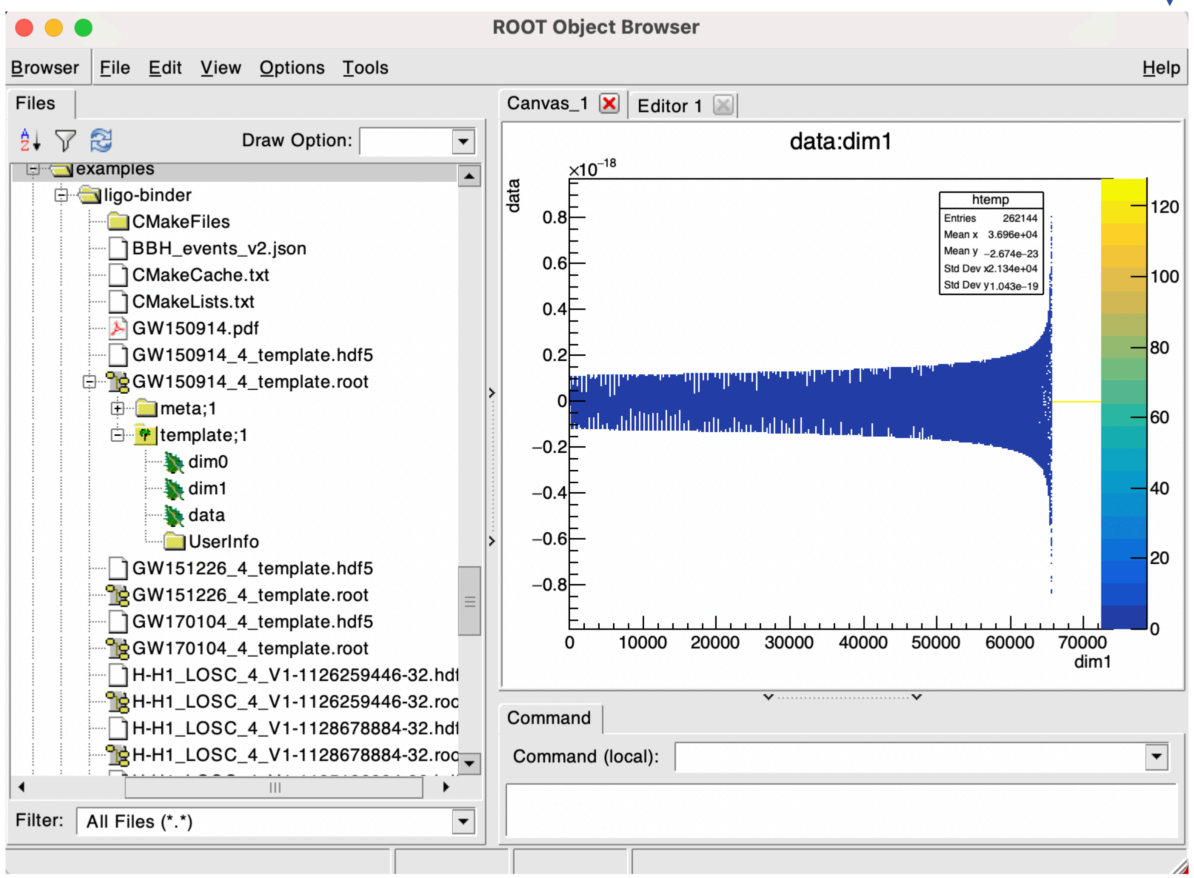


Matched filtering technics with "best template" using GW150914 data



SUMMARY: PROPOSED GW WORKFLOW





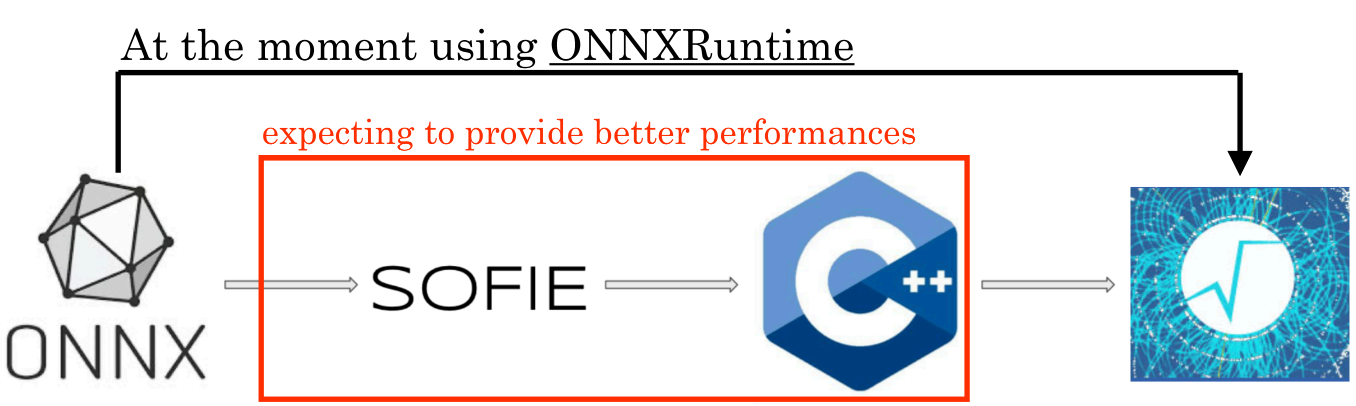
- ROOT is very much compatible with GW analysis:
  - ✅ KFR Library, TTimeSeries/TFrequencySeries prototype
  - ✅ TKafka consumer/broker client prototype
  - ✅ TConfigParser (YAML,JSON,TOML,XML,.. run-time files) prototype
  - ✅ TCurlFile / TArchive prototype (download/cache files and archive handling)
  - ✅ THDF5 parser prototype
- LALSuite Package linkable for direct use.
  - ✅ TStrain, TStrainNoise, TStrainTemplate
    (for direct strain, noise and template generation)
  - ✅ TLigoFormat for configuration file parsing
  - ✅ TFrameChain GWF proprietary format. (MT)

ULTIMATE GOAL: MACHINE LEARNING INFERENCE WITH ROOT

- Starformer project, a transformer-based model for early detection of CBC mergers
  — Collaboration with ISLab *(experts in transformer architecture development, Osaka University)*
  — *Explanable AI is of highest importance, flexibility for production models, parallel performances.*

- Machine learning universal format for models: “Open Neural Network Exchange” (ONNX) files. (Training: pyTorch / Inference: ROOT)

- Inference using ROOT + ONNX Runtime:
  — ONNX Runtime: developed by Microsoft and Facebook. (Please visit https://huggingface.co/ too!)
  Interoperability in many languages. (Python, C++, ..)
  — Performance enhanced expected using SOFIE library in the future (maybe)



CONCLUSION

- Gravitational wave community is building a new computing paradigm:
  — Computing power requirement and increasing detector sensitivity.
  — Ongoing work (since many years), now resulting in the IGWN computing collaboration.
- Computing power doesn't grow that much in GW compared to LHC experiments
  — Still very challenging to communicate quickly with MMA institutes and astronomers.
  — Kafka integration prototype in an external ROOT library.
- Goal: provide mature tools as a bedrock for analysis and pipeline development.
  — Common framework, performant tools + keep flexibility towards new IT technologies
  — Machine learning models is (since many years) intensively used
- Use of ROOT as main data framework in the future
  — provide tools to run optimized DSP analysis
  — communicate with cloud computing services.
- Development of Starformer project (transformer-based ML model)
  — collaboration with ISLab (Osaka University) for innovative architecture developments.