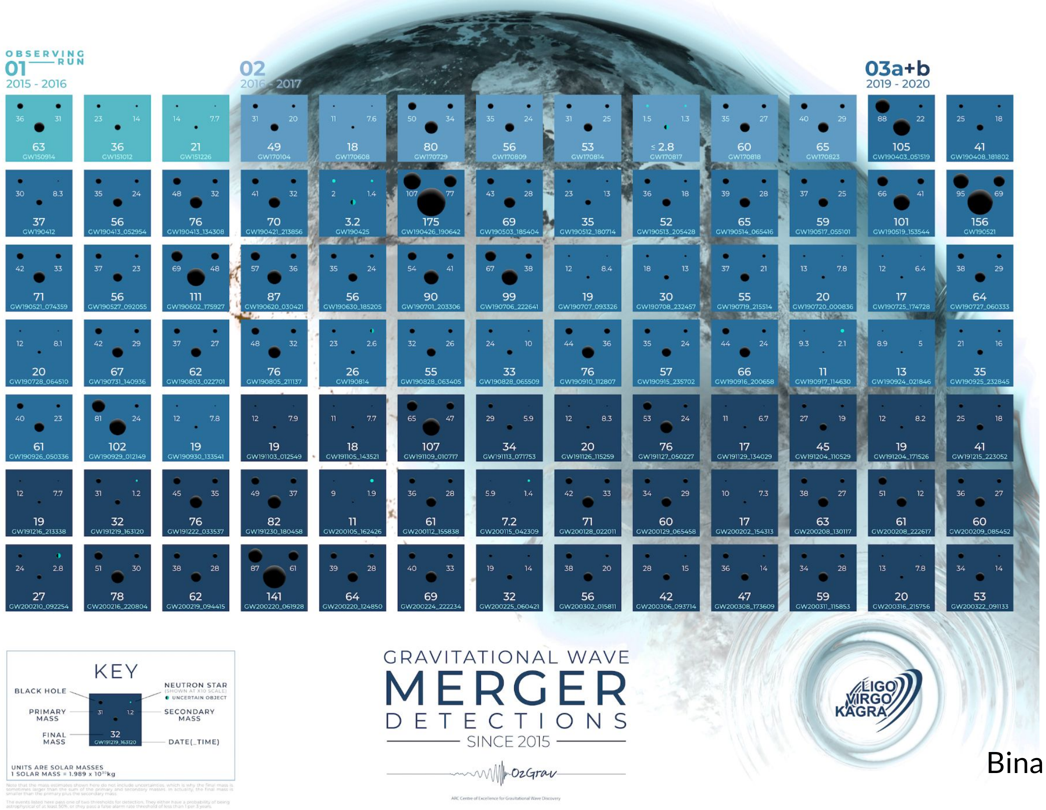
The Ligo-Virgo-KAGRA Computing Infrastructure for Gravitational-Wave Research

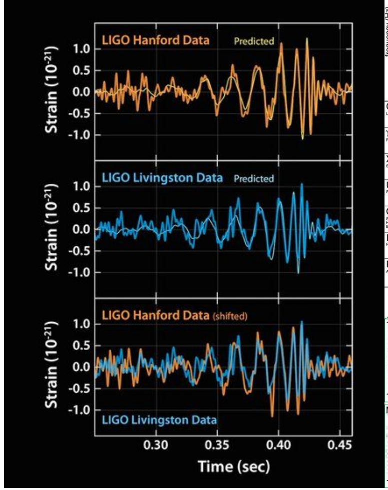
F. Legger,

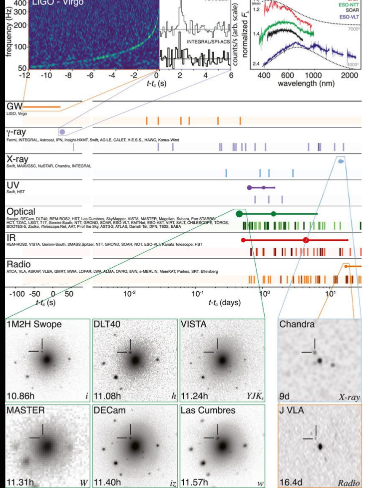
On behalf of the LVK Collaboration

CHEP23, Norfolk (VA), May 8th, 2023

What is this about?

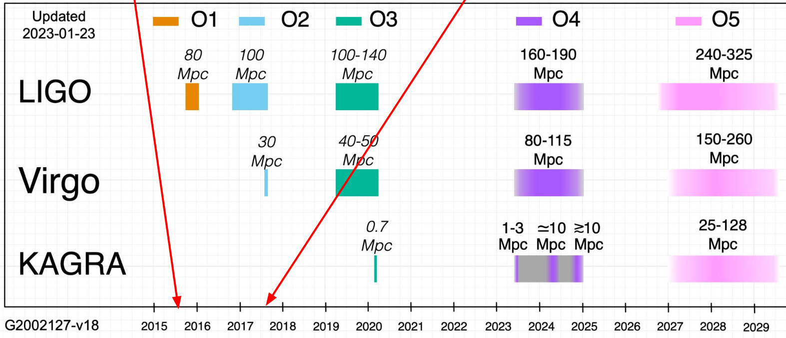






September 14, 2015

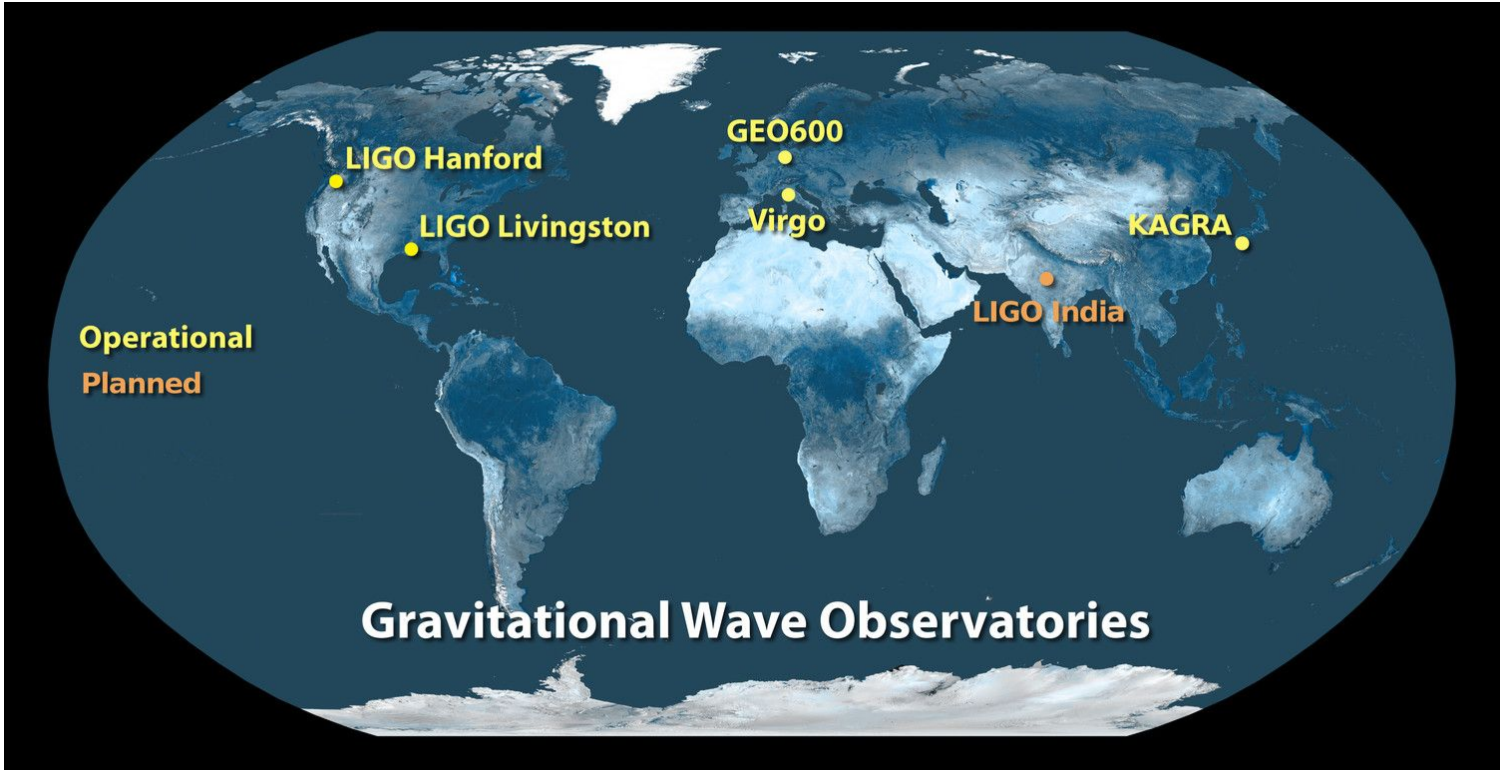
August 17, 2017



Binary Black Hole Merger

Binary Neutron Star Merger

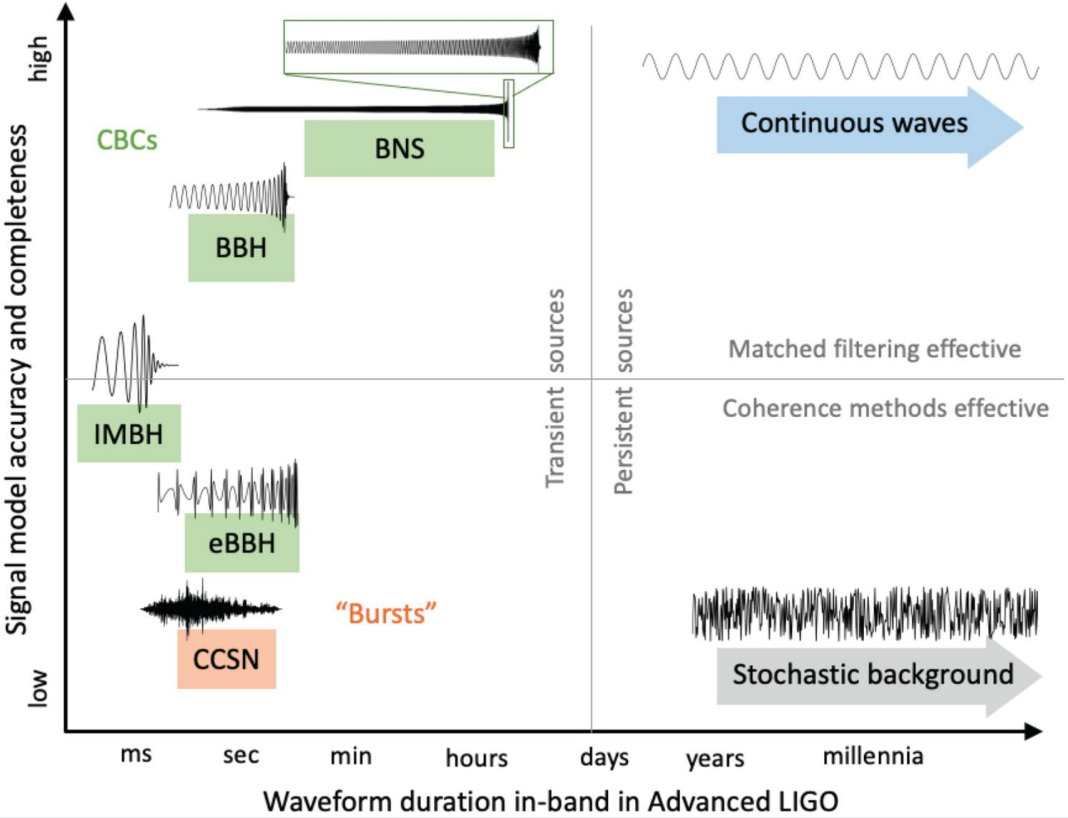
A distributed network of Observatories



Gravitational Wave (GW) Detection

CBC: Compact Binary Coalescence Coalescing Compact Binary Systems (Neutron Star-NS, Black Hole-NS, BH-BH): Strong emitters, well modelled for most parameter space

Burst: Unmodeled transient bursts Asymmetric Core Collapse Supernovae: weak emitters, not well-modelled (“bursts”), transient. Cosmic strings, soft gamma repeaters, pulsar glitches, ...



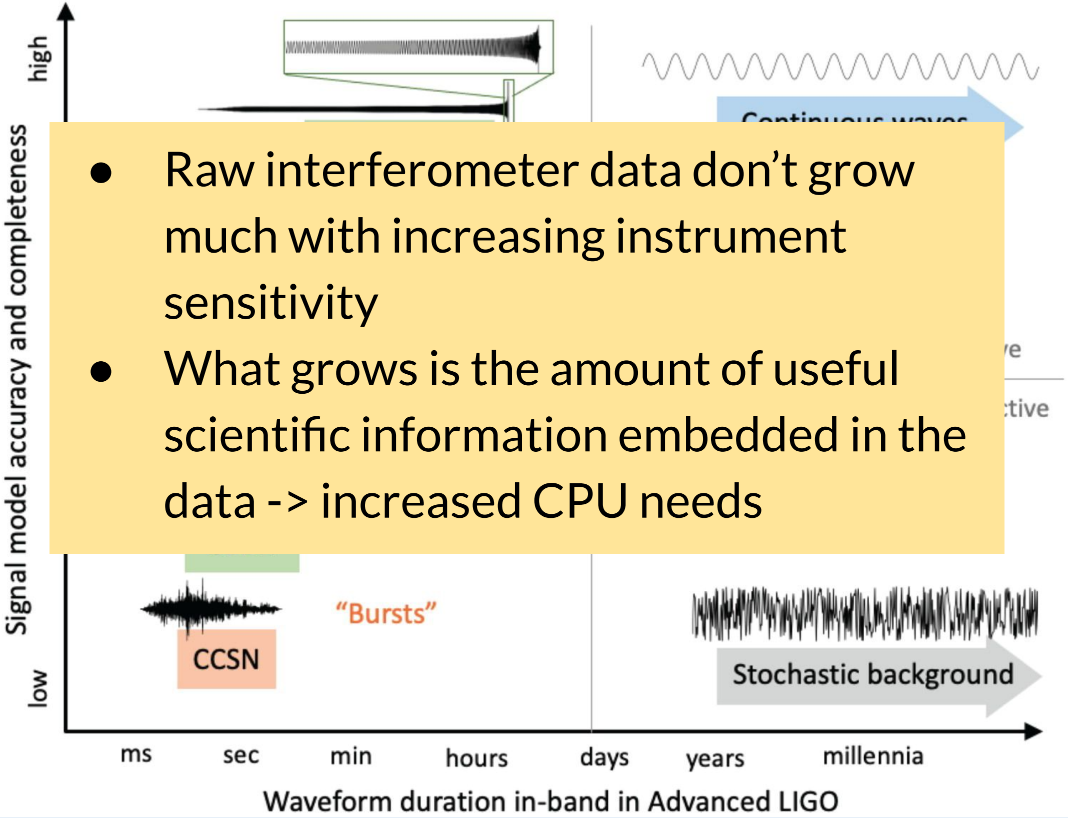
CW: Continuous waves Spinning neutron stars (known waveform, long/continuous duration). All-sky and targeted searches

SGWB: Continuous stochastic background Cosmological stochastic background (Big Bang residuals, cosmic GW background, long duration). Astrophysical stochastic background

Gravitational Wave (GW) Detection

CBC: Compact Binary Coalescence Coalescing Compact Binary Systems (Neutron Star-NS, Black Hole-NS, BH-BH): Strong emitters, well modelled for most parameter space

Burst: Unmodeled transient bursts Asymmetric Core Collapse Supernovae: weak emitters, not well-modelled (“bursts”), transient. Cosmic strings, soft gamma repeaters, pulsar glitches, ...



- Raw interferometer data don’t grow much with increasing instrument sensitivity
- What grows is the amount of useful scientific information embedded in the data -> increased CPU needs

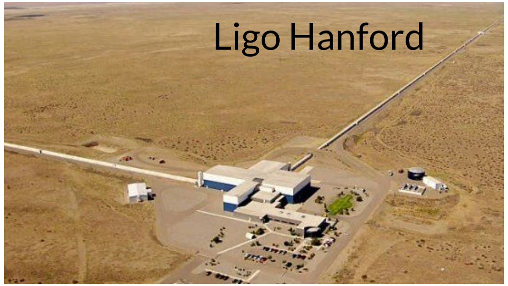
CW: Continuous waves Spinning neutron stars (known waveform, long/continuous duration). All-sky and targeted searches

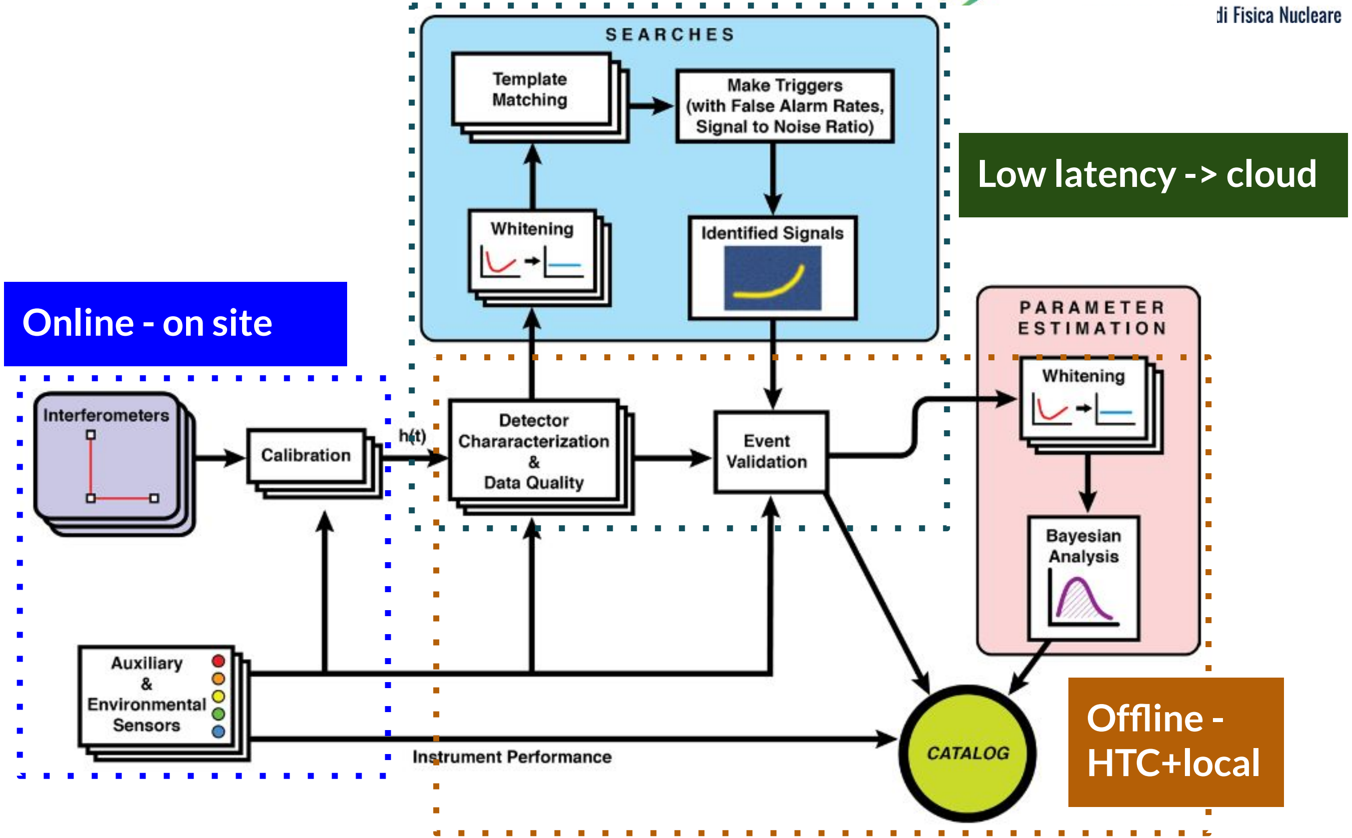
SGWB: Continuous stochastic background Cosmological stochastic background (Big Bang residuals, cosmic GW background, long duration). Astrophysical stochastic background

Data Flow

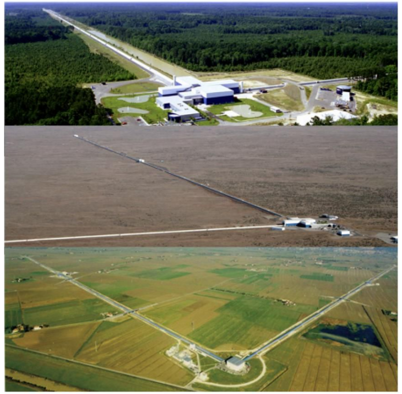

Virgo


Ligo Livingstone

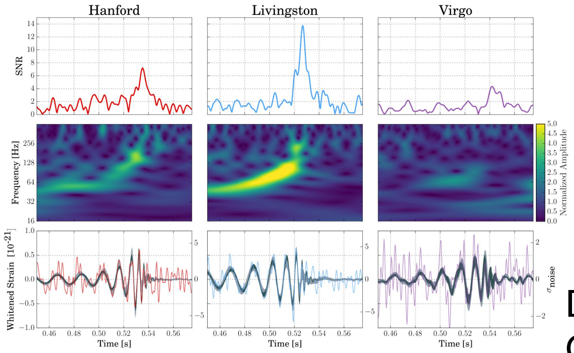

Ligo Hanford


SEARCHES
Template Matching
Make Triggers (with False Alarm Rates, Signal to Noise Ratio)
Whitening
Identified Signals
Low latency -> cloud
PARAMETER ESTIMATION
Online - on site
Whitening
Interferometers
Calibration
h(t)
Detector Chararacterization & Data Quality
Event Validation
Bayesian Analysis
Auxiliary & Environmental Sensors
CATALOG
Instrument Performance
Offline - HTC+local

Low Latency Searches in O3

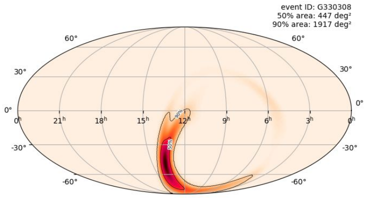


GW candidate



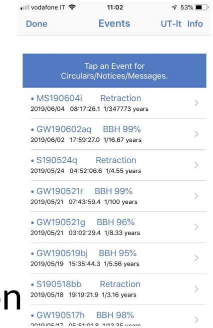
Low latency searches

Sky localisation



Detector sanity, Data Quality, localization

Public alert



Event validation

On site

A few minutes

Half an hour

Off site

- Parameter estimation
- GW Candidate Update

Hours, days

The International GW Network (IGWN)

- A coordination effort aimed at jointly discussing the computing policy, management, and architecture issues of LIGO, Virgo, and KAGRA

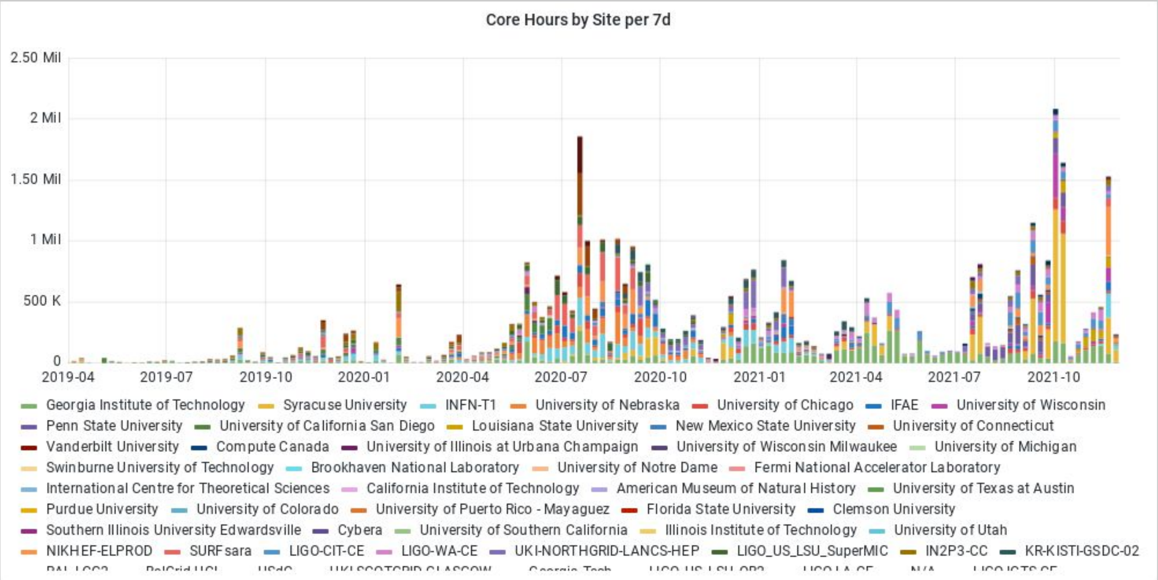
IGWN computing in a nutshell

- **Data handling:**
  - low-latency distribution with Kafka
  - higher-latency distribution via CVMFS
- **Software and computing environments:**
  - low-latency workflows -> dedicated resources
  - higher-latency workflows -> HTC platform
- **Resource provision:**
  - HTC platform allows easy contribution of resources (GPUs, HPCs, ...)

Requirements

Overall: O(10%) of an LHC experiment of today

- **Compute:**
  - **O3:** 700M CPU core hours/year
    - Astrophysical searches (total: about 80, 10 of these using 90% of CPU resources)
    - Low latency (10%)
    - Detector characterization
  - **O4**: 1.5 -2x O3

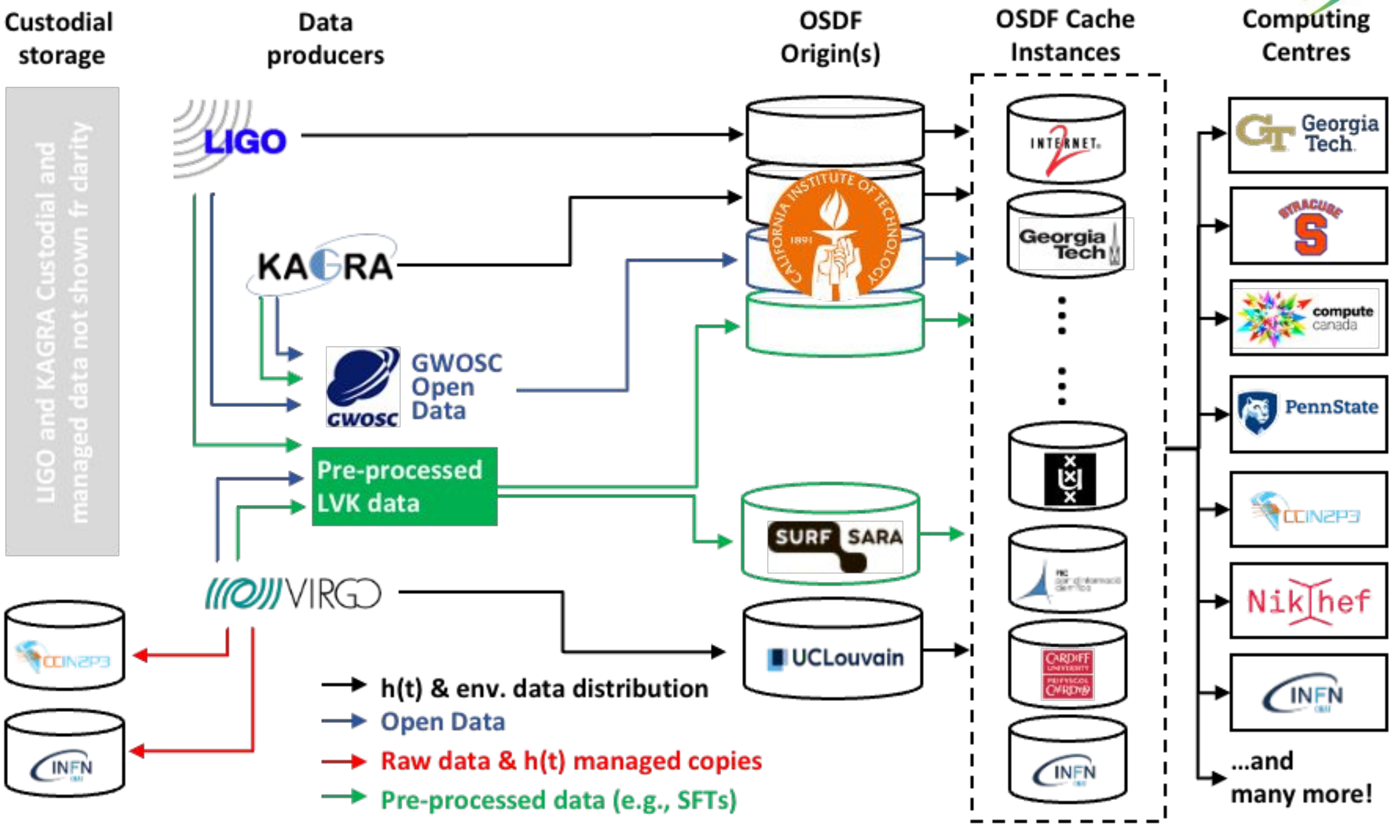


- **Data:**
  - Strain *h(t)* : 20 TB /year /experiment
  - Raw (auxiliary and control channels): 1.5 PB /year /experiment

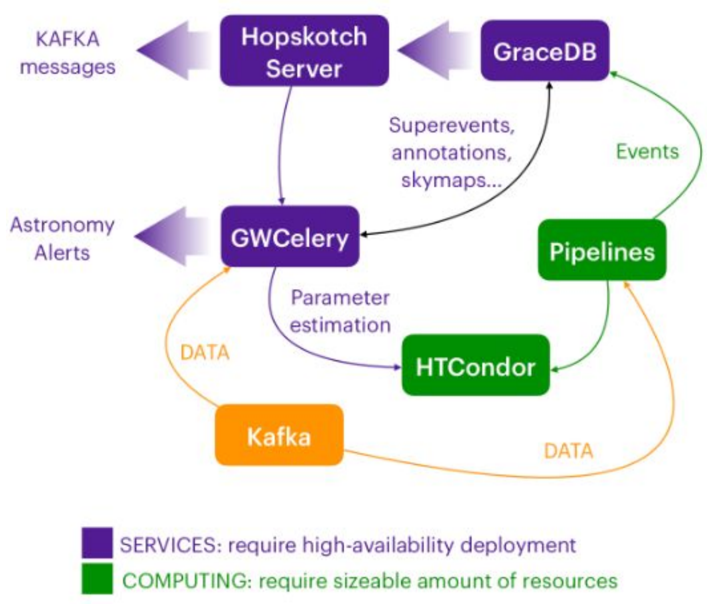
Technologies

- **HTCondor** for job management
- **Rucio** for (most) data management
- **CVMFS + StashCache (xrootd)** for data distribution
- **CVMFS** for software distribution
- **GitLab + Conda + cmake** for code management
- **Apache Kafka** for low-latency data exchange
- **Kubernetes** for service deployment on cloud resources
- Collaboration operations: **Federated identity (IAM)**

Higher-latency Data Distribution



Low Latency in O4



- Deployment of main services:
  - GraceDB -> AWS
  - GWCelery -> dedicated Hardware
  - Hopskotch -> AWS
- Additional deployment for development and testing of the main services:
  - on K8s in INFN-Cloud at CNAF
  - configuration option for self-contained installation on Minikube
  - GitLab CI for deployment and configuration
  - Monitoring via Prometheus/Grafana

See also Vallero's poster *GRAVITATIONAL WAVE ALERT GENERATION INFRASTRUCTURE ON YOUR LAPTOP*

Summary

- GW community entering a new computing era:
  - full interoperability between Virgo, LIGO (and KAGRA)
  - a common and sustainable computing environment
  - a uniform runtime environment for offline pipelines
  - scalability and the opportunity to exploit heterogeneous resources
  - adoption of mainstream, widely used tools

Keep calm and ready for O4!

Further Reading

- Discovering gravitational waves with Advanced LIGO, Jess McIverand D. H. Shoemaker, https://doi.org/10.1080/00107514.2021.1946264
- The IGWN Computing Grid
- Gravitational Wave Open Science Center (GWOSC)
- First Demonstration of Early Warning Gravitational-wave Alerts, Ryan Magee et al 2021 ApJL 910 L21
- Gravitational wave data analysis - Computing Challenges in the 3G era

Backup

Detailed Data Flow