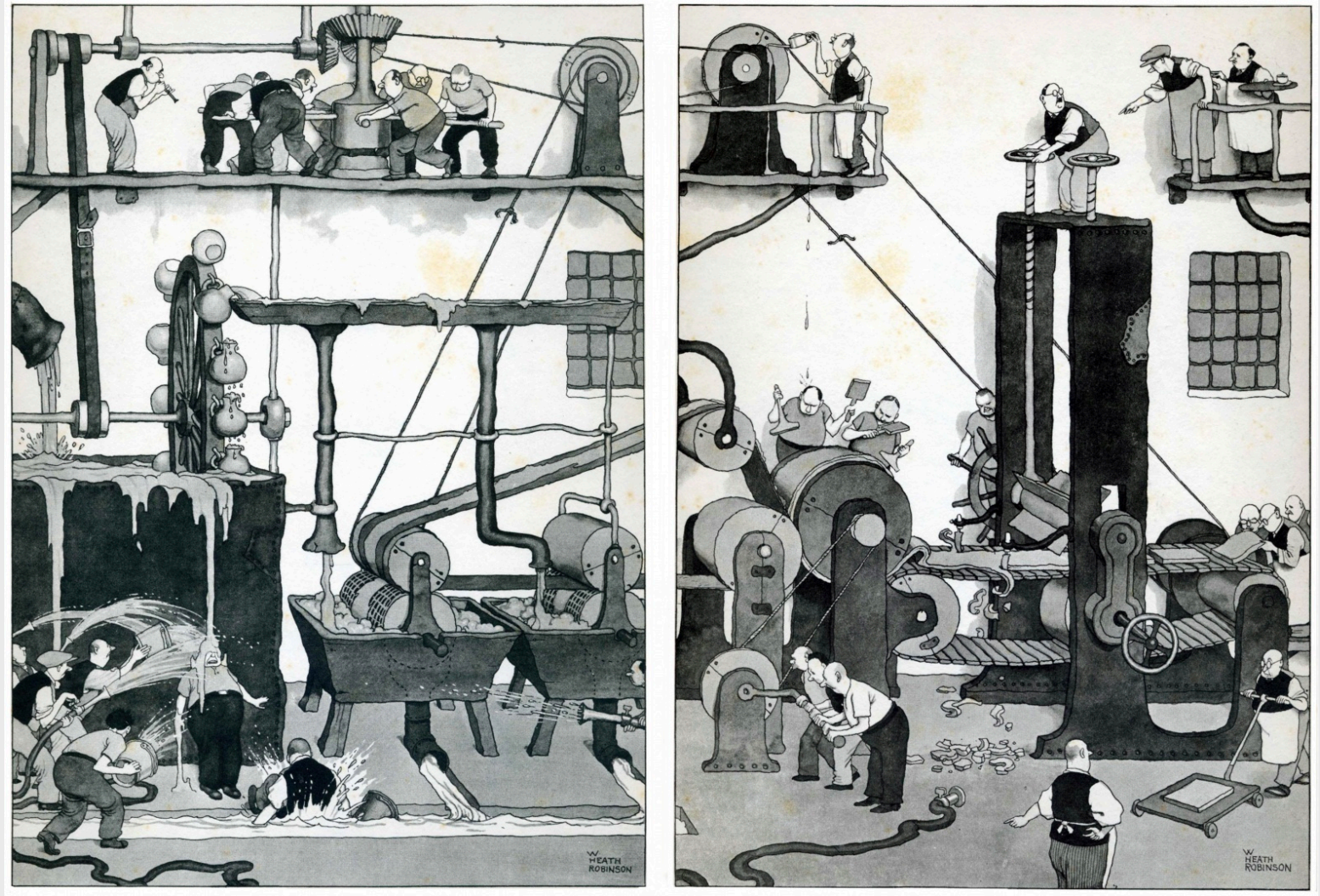
Computing infrastructure & activities



Looking for...

Burst sources:

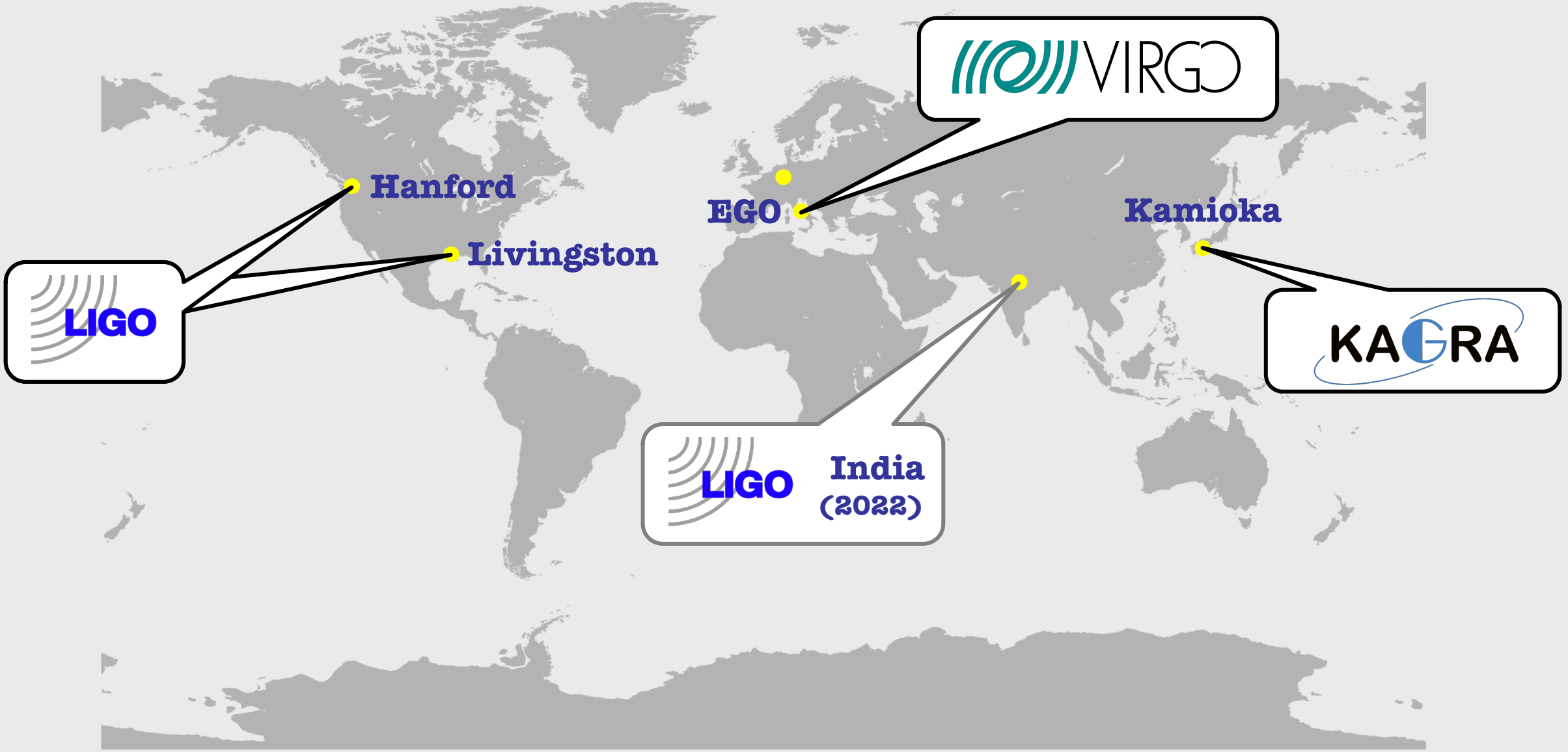
- Compact Binary Coalescence
  - Coalescing Compact Binary Systems (Neutron Star- NS, Black Hole-NS, BH-BH): Strong emitters, well modelled
- Unmodeled transient bursts
  - Asymmetric Core Collapse Supernovae: weak emitters, not well-modelled ('bursts'), transient
  - Cosmic strings, soft gamma repeaters, pulsar glitches

Low latency searches

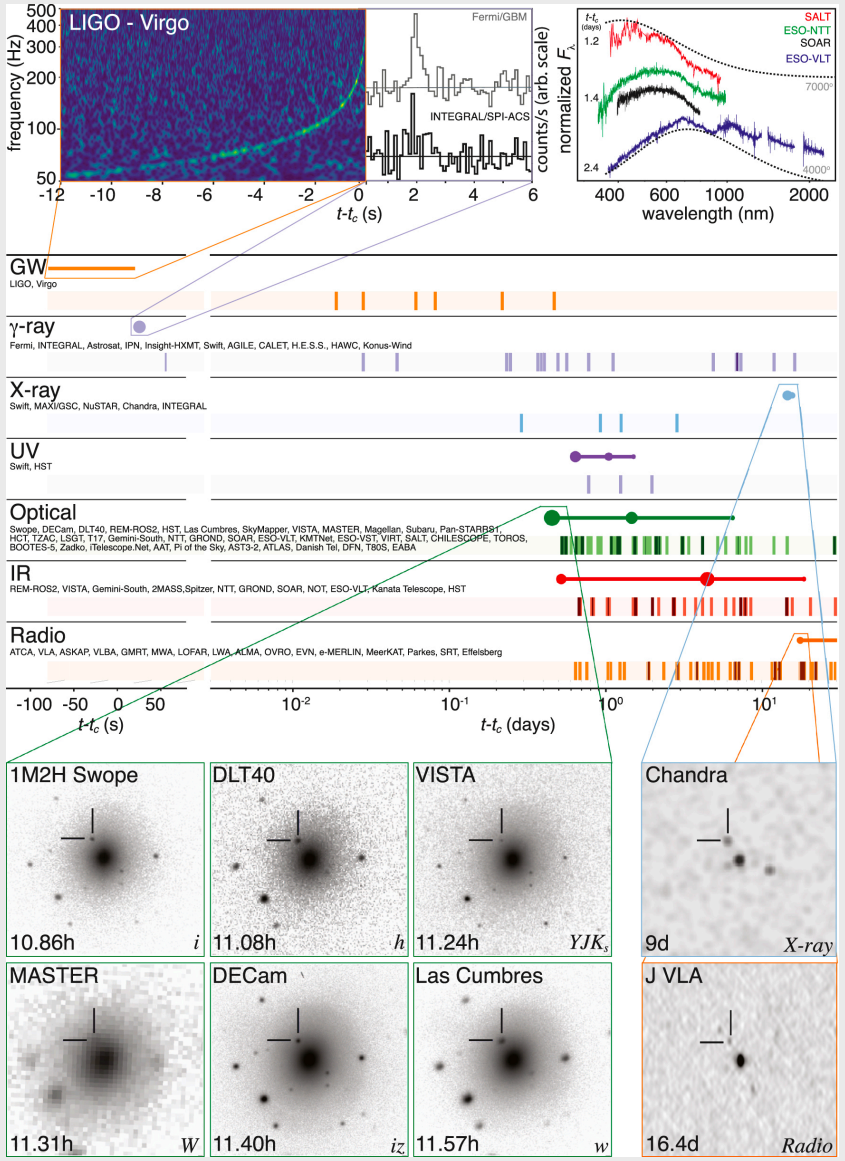
Continuous sources:

- Continuous stochastic background
  - Cosmological stochastic background (residue of the Big Bang, cosmic GW background, long duration)
  - Astrophysical stochastic background
- Continuous waves
  - Spinning neutron stars (known waveform, long/continuous duration)

Worldwide network

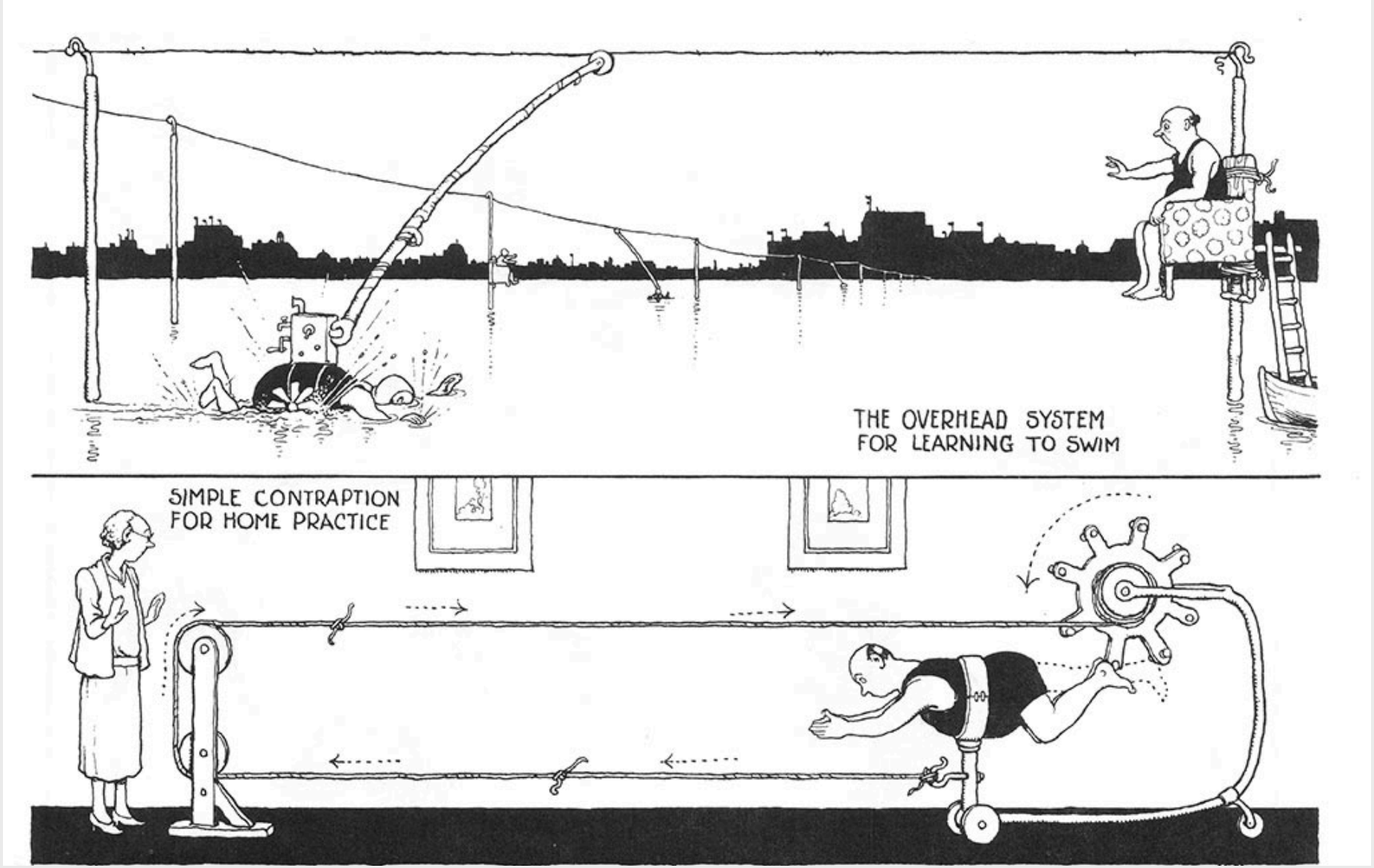


Multimessenger astronomy



"Multi-messenger Observations of a Binary Neutron Star Merger"
B. P. Abbott *et al.* 2017 *ApJL* 848 L12
doi:10.3847/2041-8213/aa91c9

The basics

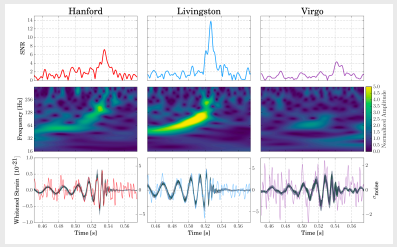

THE OVERHEAD SYSTEM
FOR LEARNING TO SWIM
SIMPLE CONTRAPTION
FOR HOME PRACTICE

Low Latency searches

Low-latency CBC searches

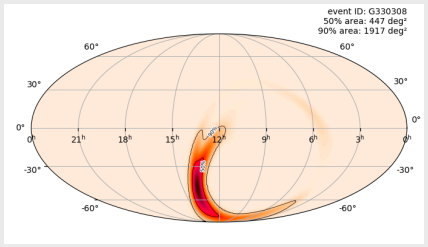
Detector sanity, Data Quality, localization,...

GW candidate

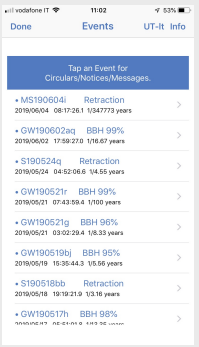


Event validation

Sky localization



Public Alert



On-site

A few minutes

½ hour

Off-site

- Parameter estimation
- GW Candidate Update

Hours, days

Raw Data, ~1PB/yr/detector:

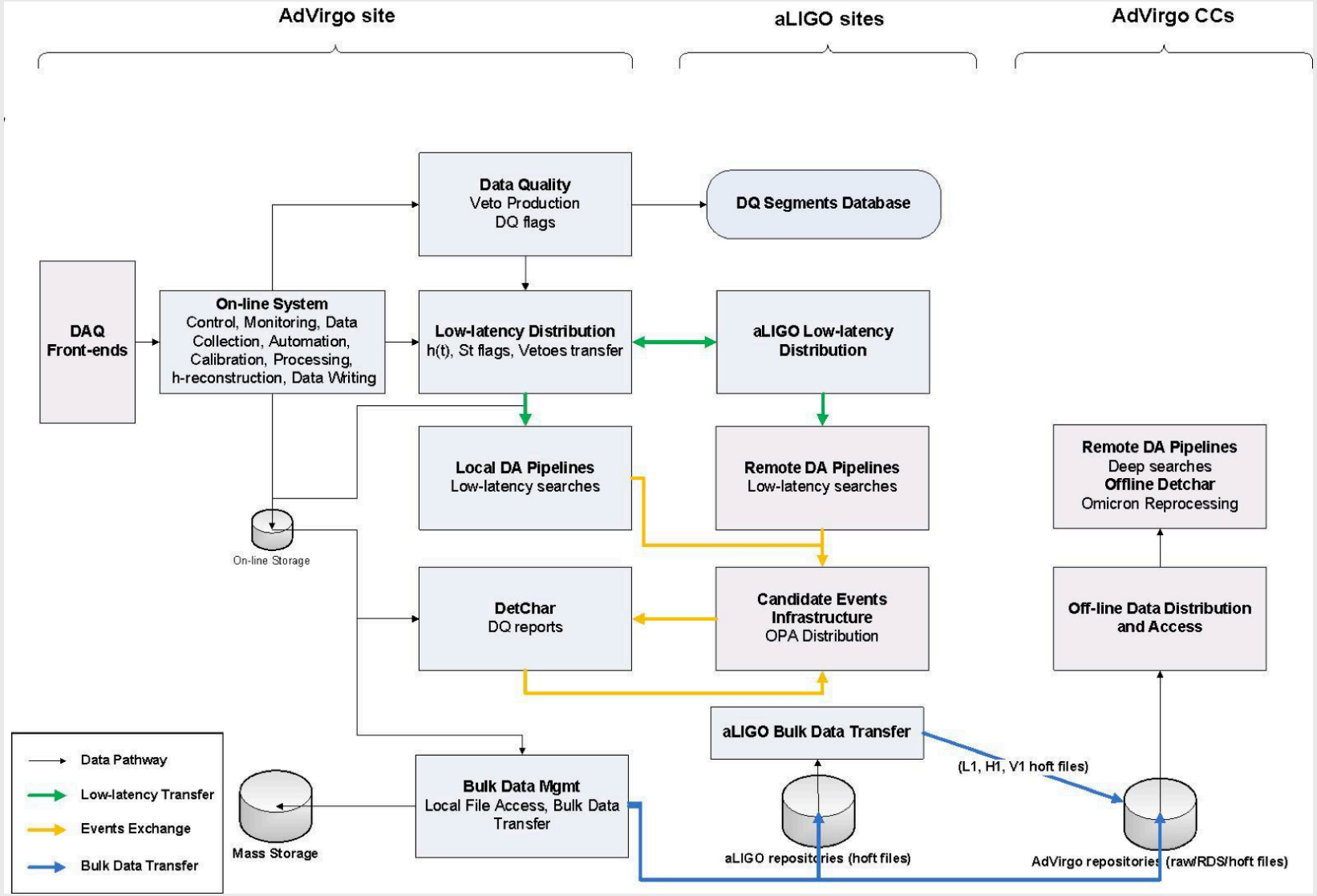
- **Full Bandwidth Raw**, not exported (7.5-10 TB/day)
- **Raw Data:** downsampled, calibrated and uncalibrated h(t) (3-4 TB/day)
- **A few levels of reduced data sets** for various uses

Data for physics, ~ 5TB/yr/detector:

- Virgo h(t): calibrated «strain» data
  - sampled at 10 KHz, as ~1kSec frame files
  - Includes state vector (data quality flags, vetoes,...)
- LIGO h(t)
  - Copied online to EGO for low-latency searches and exported to CCs for offline analysis
- KAGRA h(t)
  - Coming soon...

$$h(t) = \frac{1}{R}\frac{2G}{c^4}\ddot{I}(t)$$

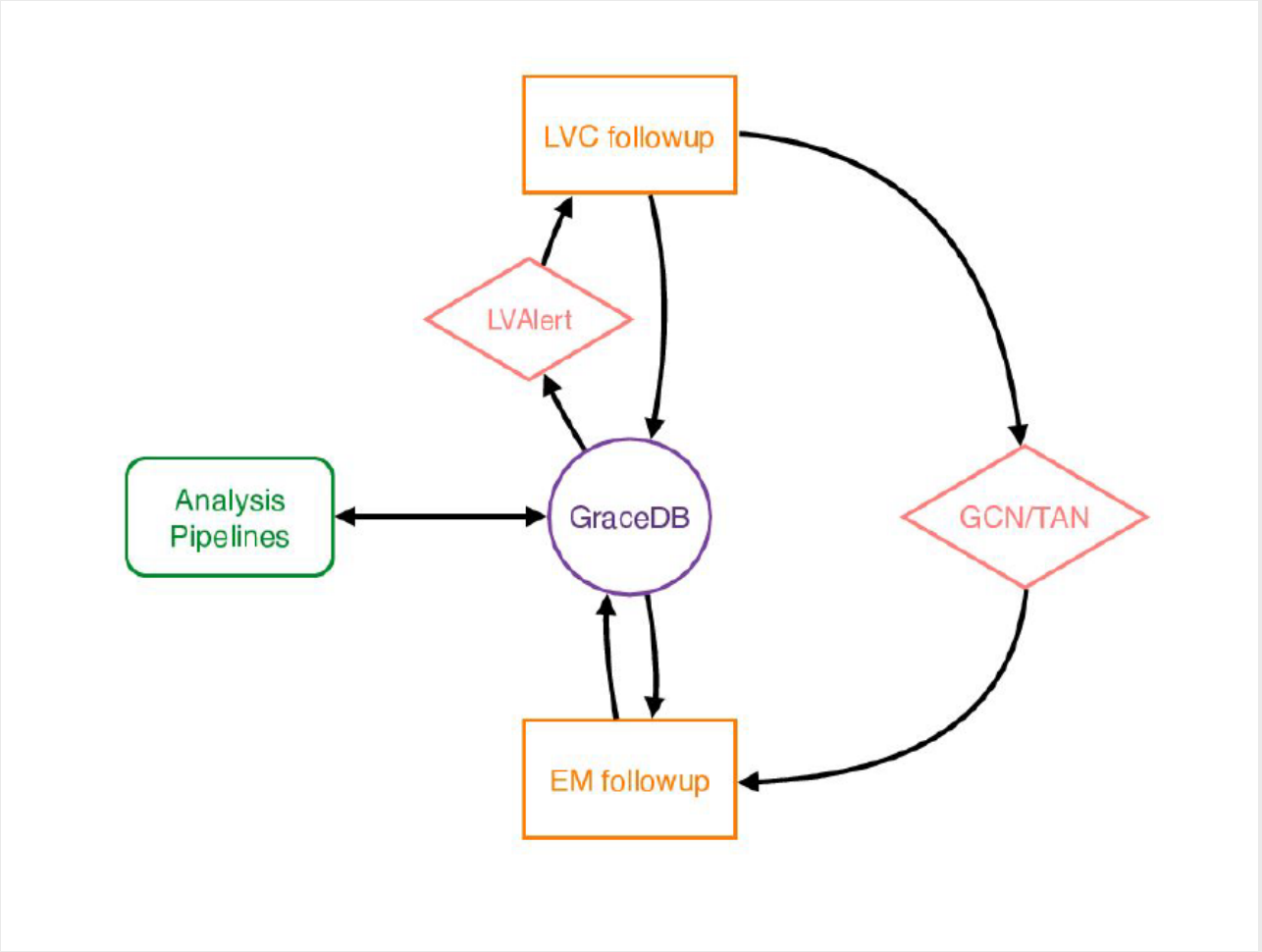
Online and low-latency dataflow schema



Franco.Carbognani@ego-gw.it

Low-Latency Cycle

- LVC analysis pipeline identifies a candidate and uploads the information to GraceDB.
- GraceDB sends out LVAlert messages, which trigger LVC follow-up processes
- These processes then add annotations to the candidate page.
- Once an event is deemed ready for release to astronomers, another process issues a GCN/TAN notice with the candidate information.



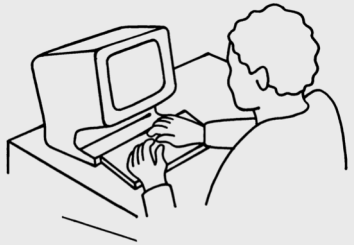
Offline activities

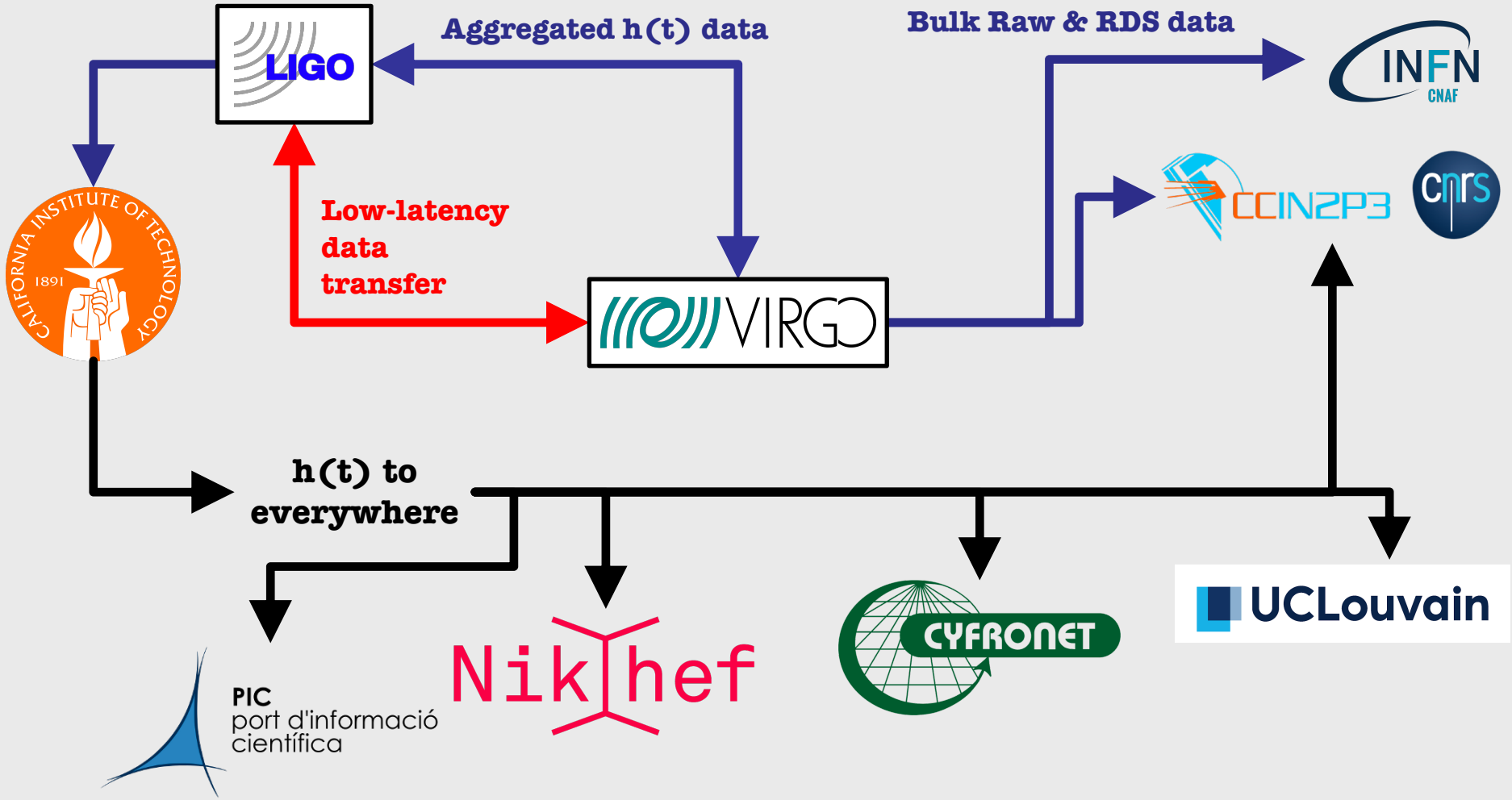
- The Virgo Offline computing infrastructure is growing
  - CNAF and CC-IN2P3 (“Tier-1”) plus Nikhef, Louvain, PIC, Cyfronet,... (“Tier-2”)
- The coordination with the LIGO counterpart is becoming ever more tight
  - Infrastructure interoperability becoming integration in a common computing infrastructure
  - And KAGRA is joining the party
- The overall architecture is essentially ready
  - And the components are being deployed
- Enter **IGWN** – the International Gravitational Waves observatories Network
  - A coordination effort aimed at jointly discussing the computing policy, management, and architecture issues of LIGO, Virgo, and KAGRA.

Virgo-side computing infrastructure

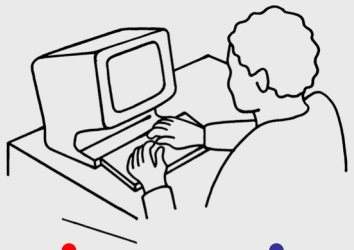
- EGO @ Cascina
  - Collaboration-wide services (AAI, Wiki, document repository,...)
  - DAQ and online data processing
  - DetChar and low-latency searches
  - Bulk data management
- CCs: CNAF, CCIN2P3 ("Tier-1")
  - Custodial data storage
  - Offline processing (reprocessing, parameter estimation, CW searches,...)
  - Support services (e.g. CVMFS Stratum-1)
- More and upcoming CCs ("Tier-2")
  - Nikhef (Amsterdam), PIC (Barcelona), Cyfronet (Krakow), Louvain (you guess)
  - Offline processing
  - ML pipelines also on "private" clusters

L'usine à gaz




LIGO
Aggregated h(t) data
Bulk Raw & RDS data
INFN
CNAF
Low-latency
data
transfer
VIRGO
CCIN2P3
cnrs
CALIFORNIA INSTITUTE OF TECHNOLOGY
1891
h(t) to
everywhere
UCLouvain
CYFRONET
Nikhef
PIC
port d'informació
científica

L'usine à gaz



Read data

Submit jobs

?

?

LIGO

Aggregated h(t) data

Bulk Raw & RDS data

INFN CNAF

Low-latency data transfer

CCIN2P3

cnrs

VIRGO

h(t) to everywhere

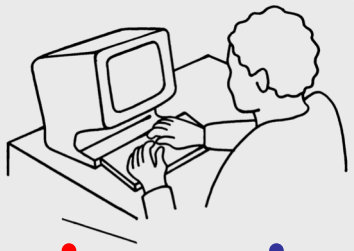
PIC port d'informació científica

Nikhef

CYFRONET

UCLouvain

What we want in one slide



Read data

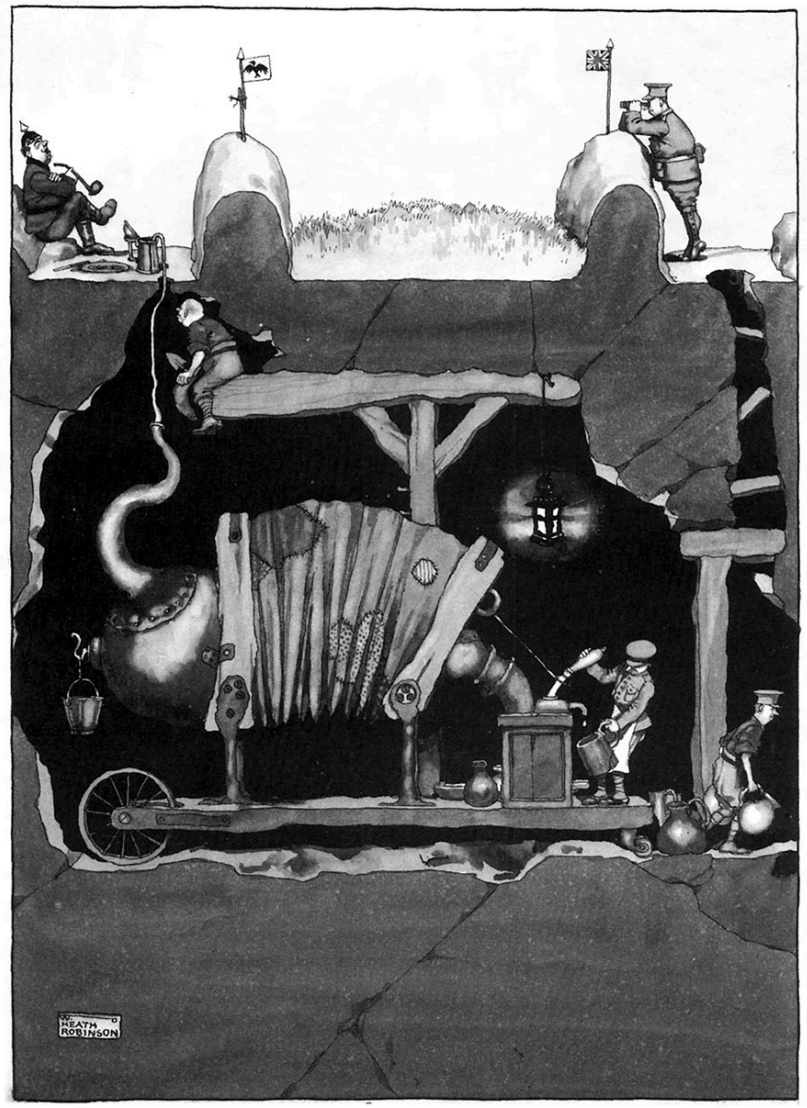
Submit jobs

HTCondor

/cvmfs

The IGWN Computing Infrastructure

The hidden machinery



The Pilsener Pump

Functionality shopping list

Bulk data transfer

Safely transfer all data to custodial storage in CCs

Software packaging and distribution

Make pipeline software available ubiquitously

Data distribution

Make h(t) data discoverable and available to worker nodes anywhere

Data cataloguing and bookkeeping

Organize all data and metadata and provide querying capabilities

Workload management

Provide a uniform job submission and runtime environment

High-level workload orchestration

Keep a database of all jobs and allow the enforcement of priorities

Monitoring and accounting

Monitor distributed computing and provide reliable accounting

Guiding ideas

- Provide a uniform runtime environment for offline pipelines
- And then help pipelines adopt common submission and data access tools
- Full interoperability with LIGO, aiming at fully common infrastructure
- Use (a small number of) mainstream, widely used tools to reduce support burden
- Define an interface now (i.e. what the pipelines need to use), all changes in the backend need to be completely transparent
- In perspective, try to transparently run on heterogeneous resources (OSG, wLCG-like, HPC, dedicated, opportunistic, you name it)
- Build something that will stand the test of time…

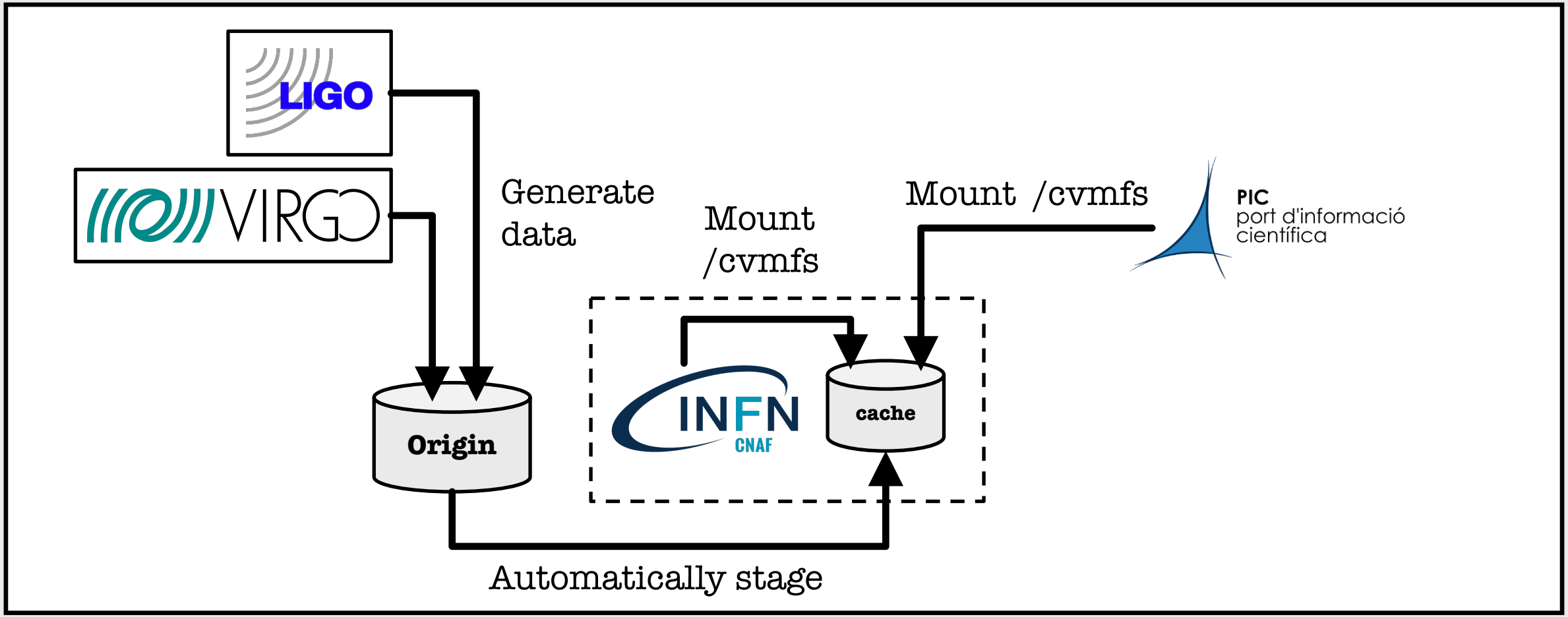
Data Distribution and Access

Make h(t) data discoverable and available to worker nodes (and workstations!) anywhere

- **What we need:**
  - h(t) files readable from worker nodes (without pre-staging)
  - A logical file structure for easy finding of relevant data
- **How we do it:**
  - A filesystem-like namespace with all available data is exported to worker nodes and even workstations
  - File metadata is encoded in filenames and directory structure, gw_data_find scans it and, you guessed, finds data
  - Actual bytes transparently read, posix-like, from a (local or remote) cache. No need to explicitly manage data transfers, the files are staged from upstream origin at first access.
- **How it works now:**
  - You already know about CVMFS!
  - StashCache is a hierarchy of xrootd caches designed to work like a content delivery network
  - Authentication provided by an OSG-developed plugin

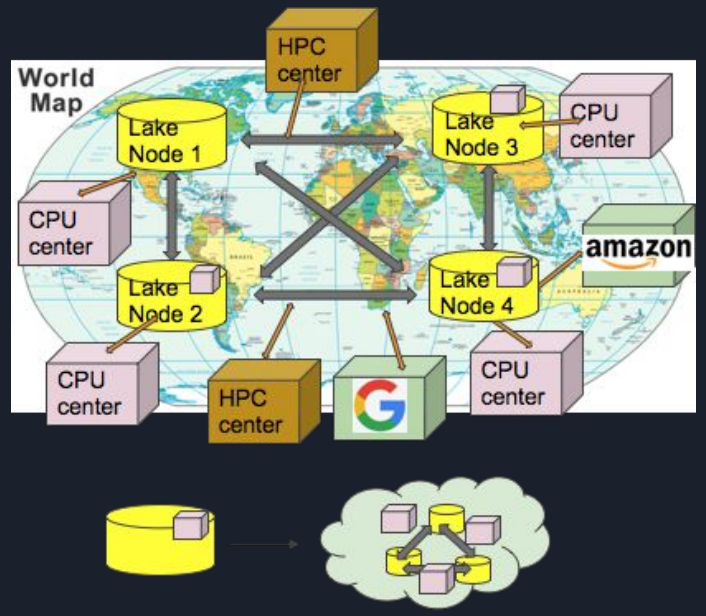
Data Distribution and Access

Make h(t) data discoverable and available to worker nodes (and workstations!) anywhere



The DataLake

- A DataLake is today's prefered R&D direction for LHC; but it has nothing which prevents it to be used by other sciences
- Idea:
  - Build a **small number of owned data centers**, which can keep the data safe
  - Make them appear as **a logical single entity** (no need for the experiments to know exactly where a file is)
  - .. which means you need to be able and **serve efficiently data** to remote sites, possibly transparently
- The gain
  - The experiment sees fewer sites (at the limit, 1 big logical storage system)
  - **A single copy is ok** (for performance; still want 2 copies of irreproducible data)
- What is needed?
  - A lot of bandwidth to fake remote sites are "as local"
    - The ability to shield a "CPU only site" with caches if the network is not good enough
  - The capability to switch on/off certain route paths on demand



Tommaso.Boccali@pi.infn.it

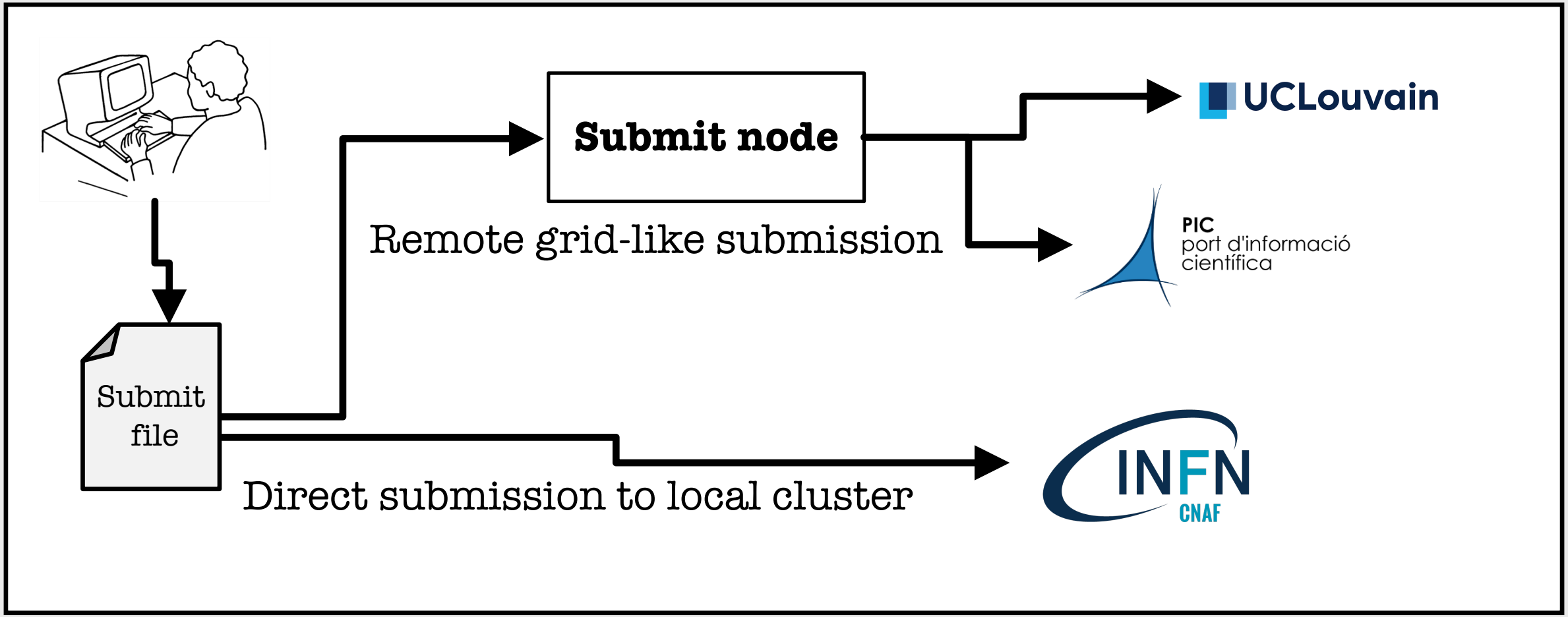
Workload Management

Provide a uniform job submission and runtime environment

- **What we need:**
  - Use the same semantics (and syntax!) for job submission and output management everywhere
  - Submit to different resources (local clusters, grid sites, etc.) using the same tools and workflow
- **How we do it:**
  - Use an HTCondor gateway for all resources
  - Same general submit file, only details change
  - Use HTCondor for job configuration and output file management wherever possible
- **How it works:**
  - Local submission to HTCondor clusters wherever available through local HTCondor schedd
  - Grid-like submission to upcoming IGWN HTCondor pool via GlideinWMS and dedicated submit nodes (one being deployed at Nikhef, more to come)
  - HTCondor can take care of output in many common use cases

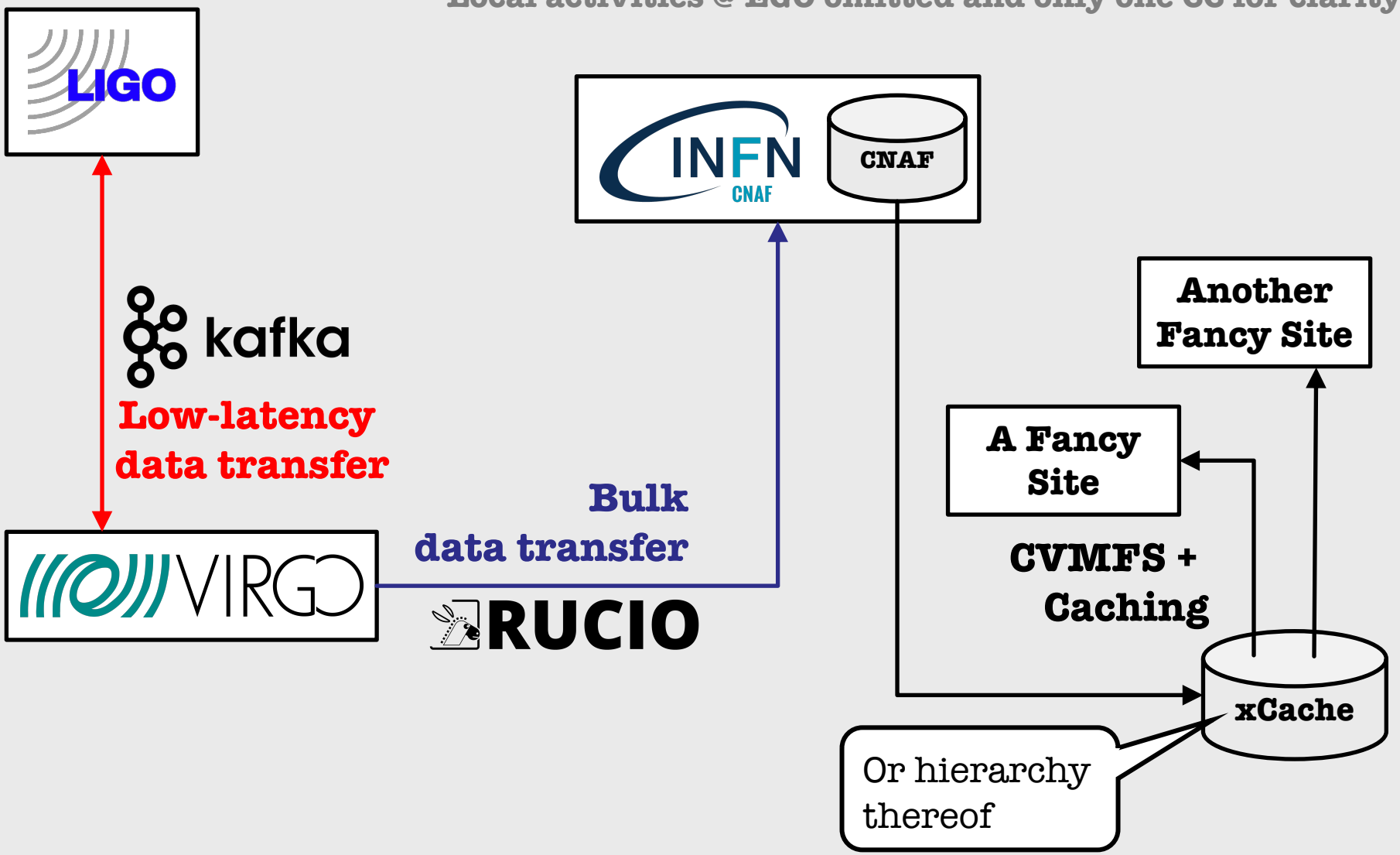
Workload Management

Provide a uniform job submission and runtime environment



A possible overall picture: data

Local activities @ EGO omitted and only one CC for clarity



Improving detector data access

```
[gfronze@my-worker-node]# tree /ruciofs

/ruciofs
|-- rucio-server-1
|   |-- user.gfronze
|   `-- user.root
|       |-- subfolder
|       |   `-- test-file-1.txt
|       `-- test-file-2.txt
`-- rucio-server-2
    |-- production
        |-- 01
        |-- 02
        |-- 03
    |-- simulation
    `-- user.gfronze
        |-- analysis-output
        |   `-- output-31102019.txt
        `-- test-file-3.txt
```

At the latest **Rucio Coding Week** the **Fuse-POSIX** module for Rucio has been **kickstarted**.

It is intended to allow to **mount multiple Rucio remote catalogs** in a common root path.

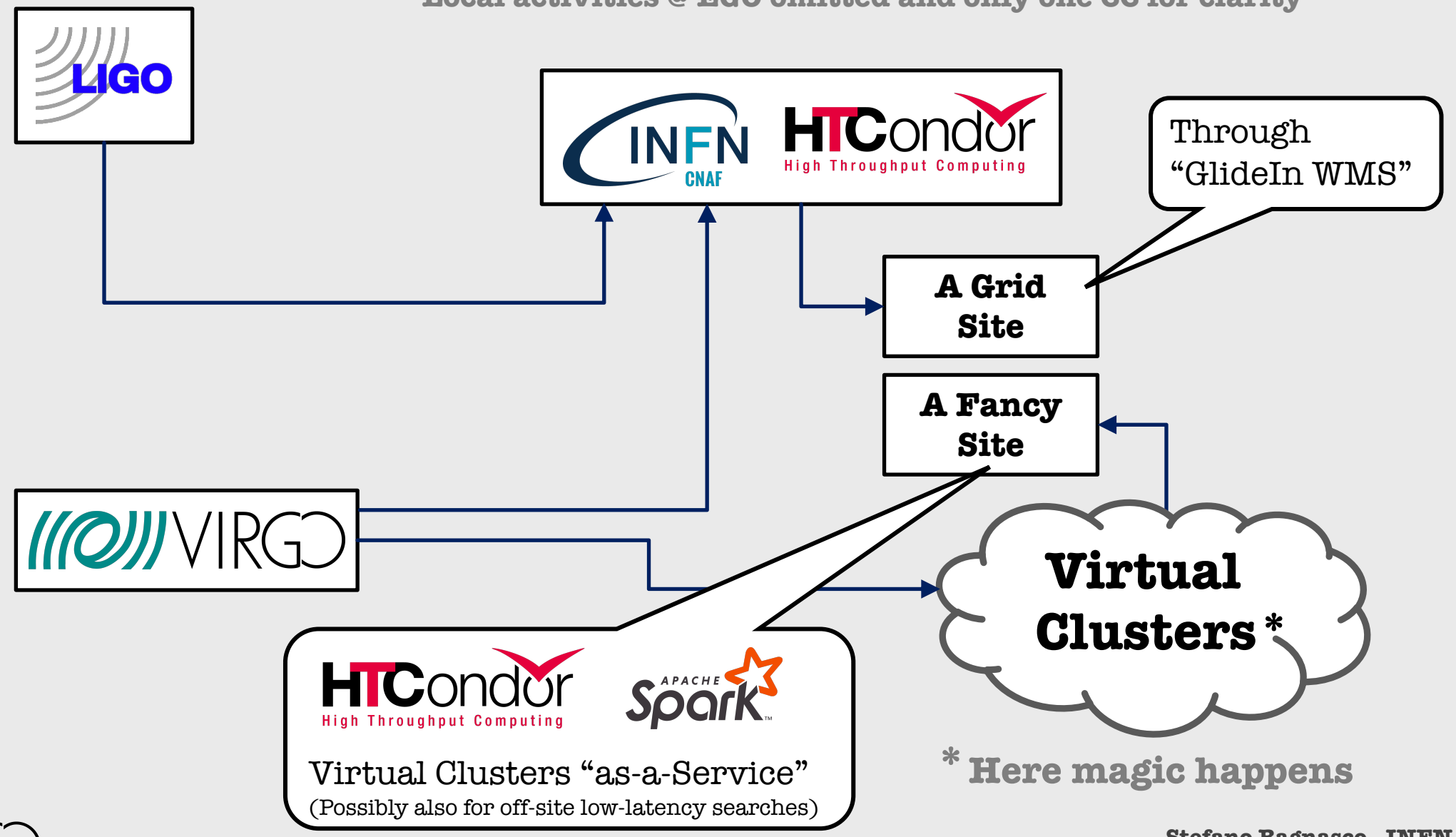
The Rucio **entities** are rendered as **directories** and **files**. They are accessible by local POSIX paths.

GitHub https://github.com/rucio/fuse-posix

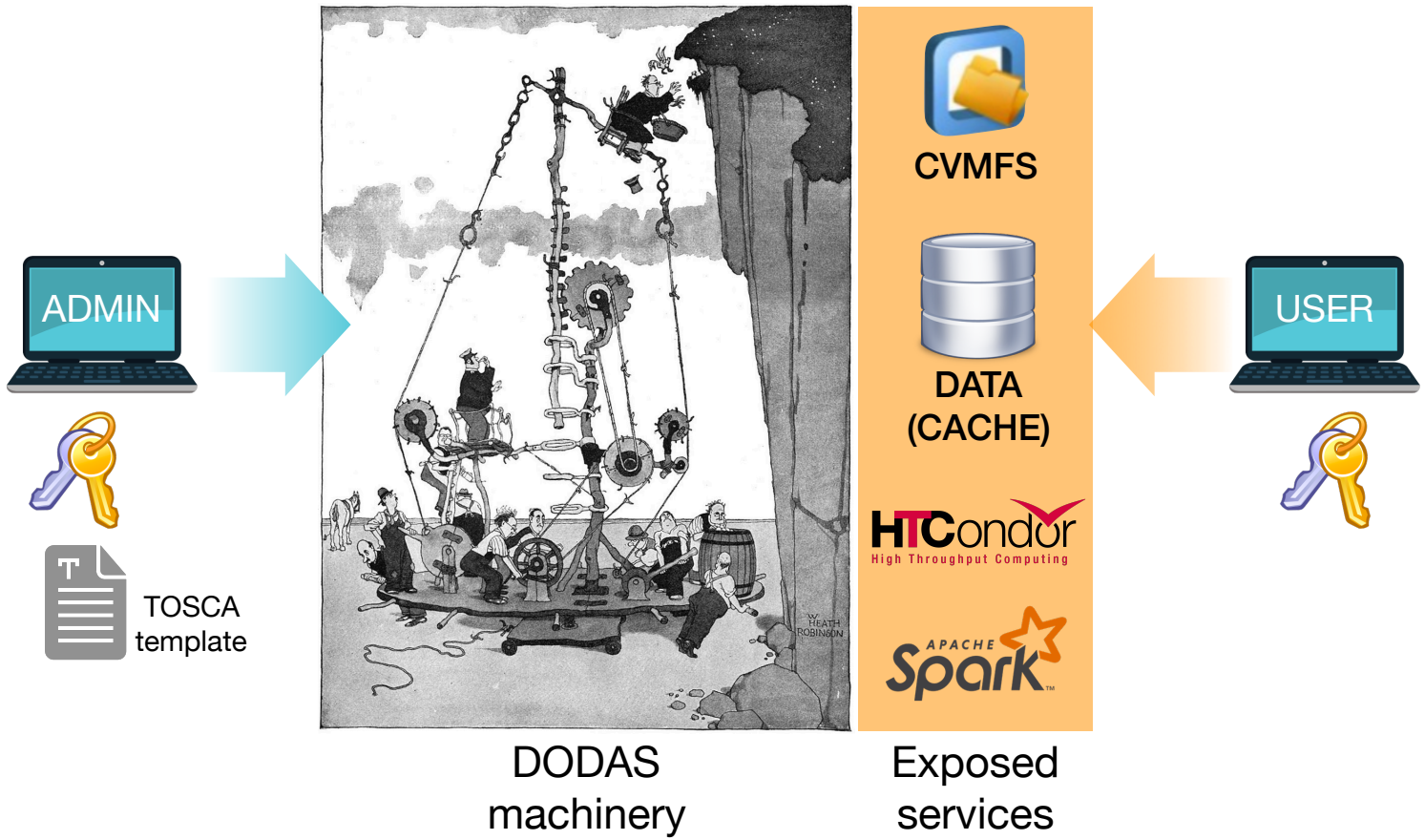
Gabriele.Fronze@to.infn.it

A possible overall picture: jobs

Local activities @ EGO omitted and only one CC for clarity

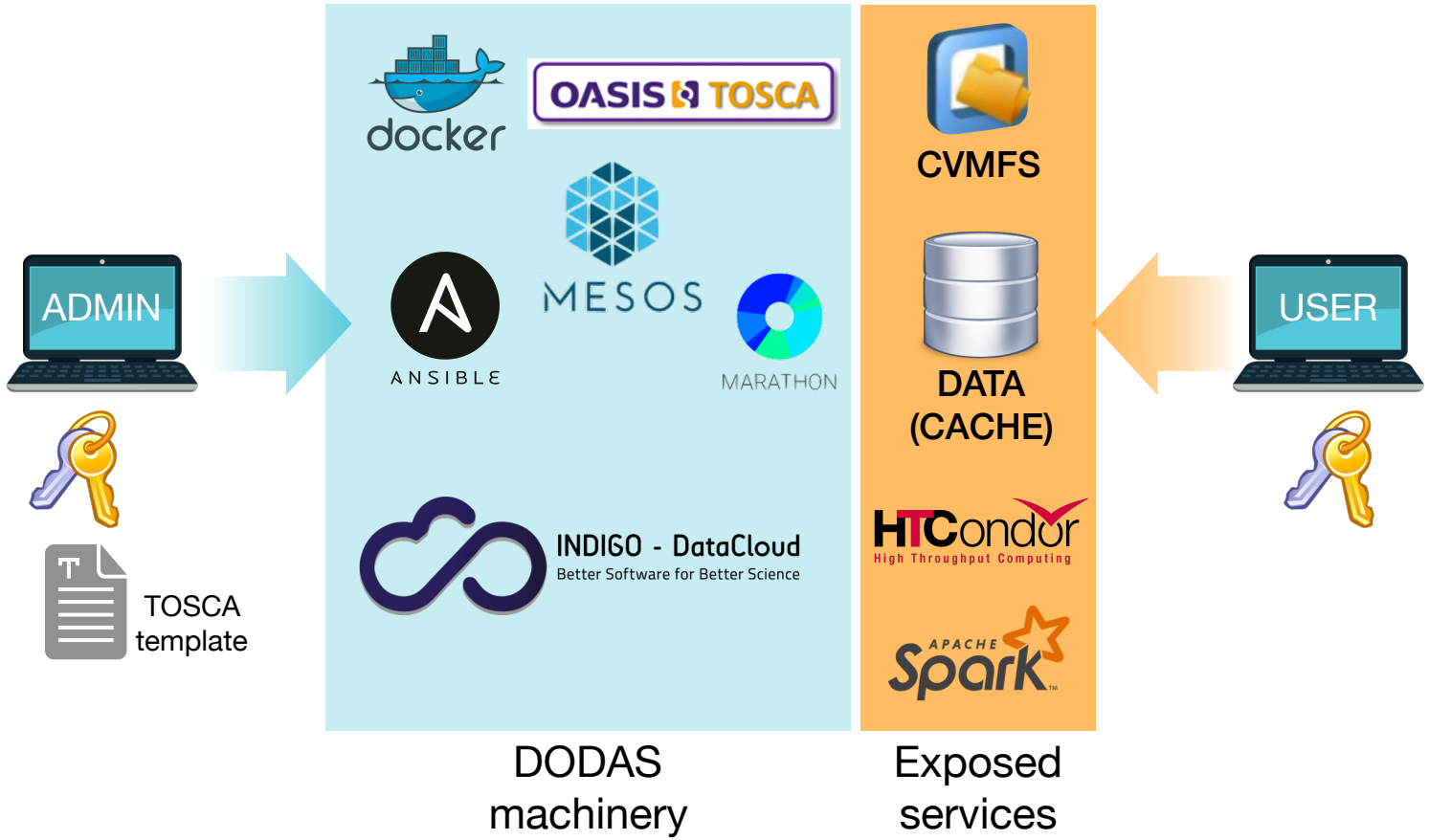


Dynamic On Demand Analysis Service



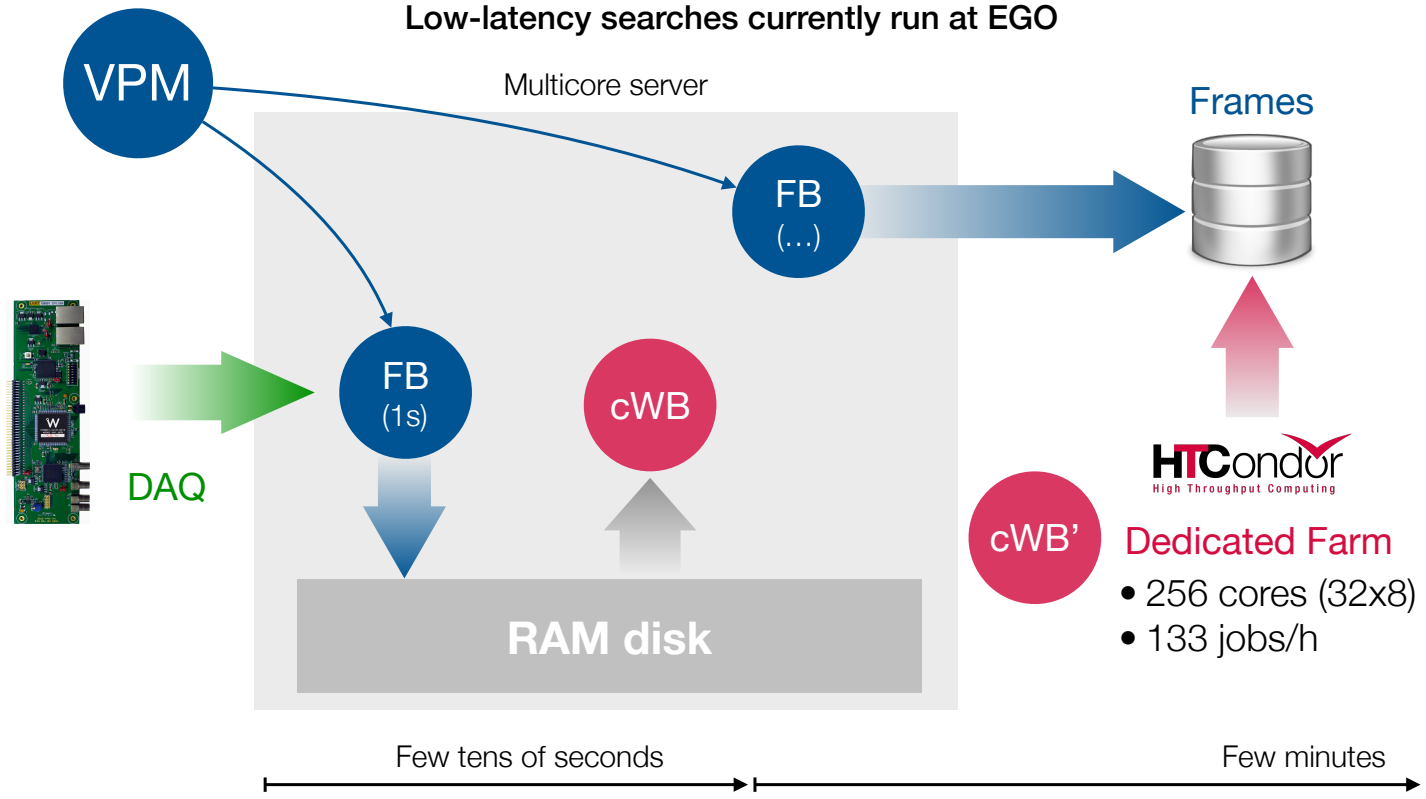
Sara.Vallero@to.infn.it

Dynamic On Demand Analysis Service



Sara.Vallero@to.infn.it

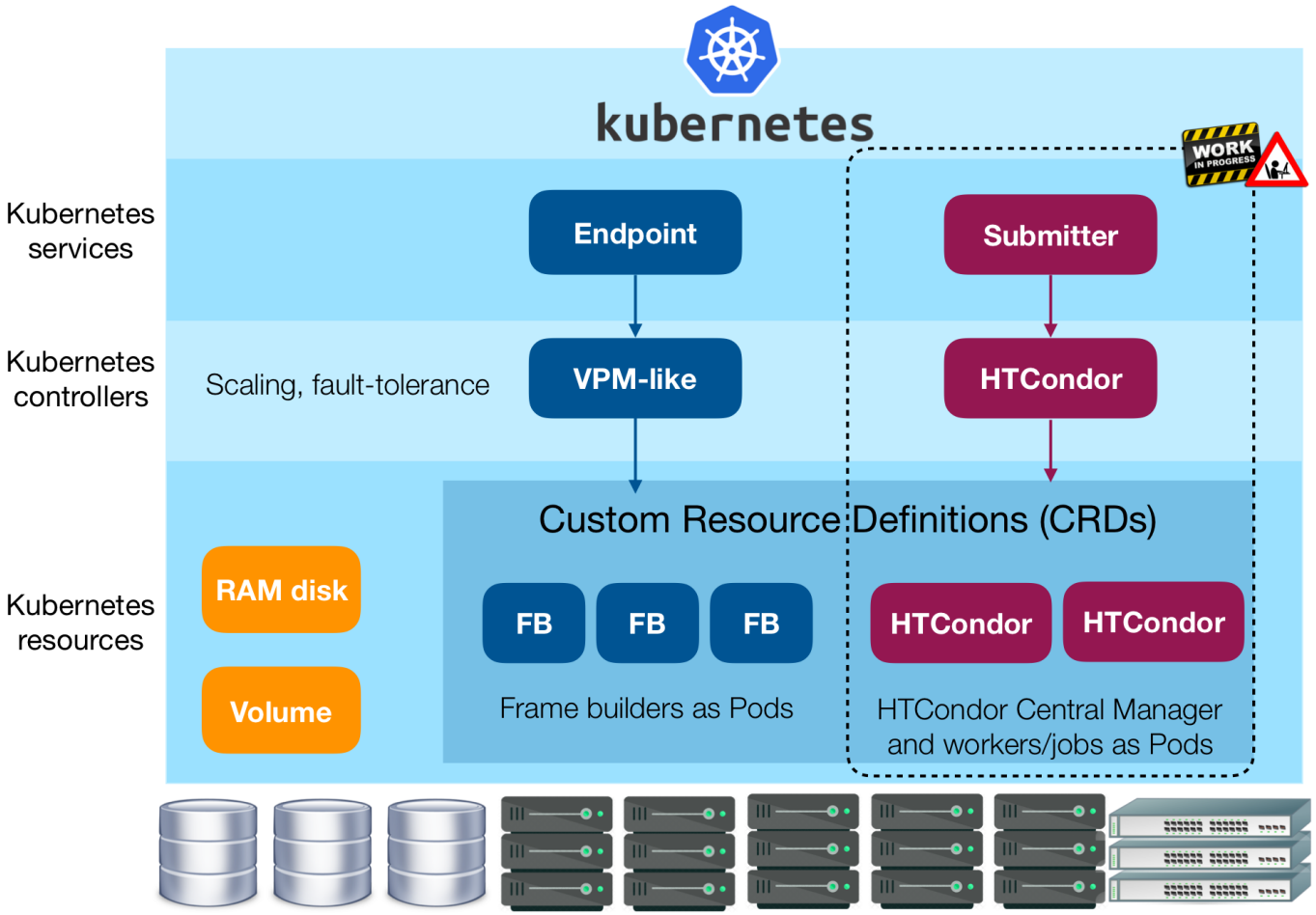
Low-latency workflow (simplified)



VPM = Virgo Process Monitor
FB = Frame Builder

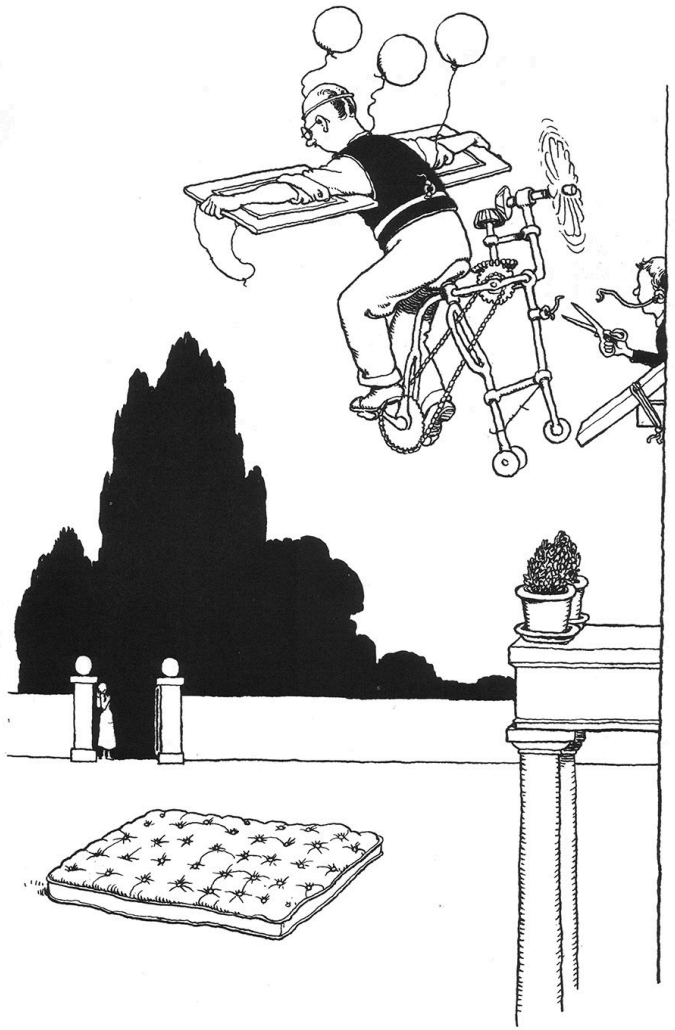
Sara.Vallero@to.infn.it

Example implementation



Sara.Vallero@to.infn.it

Wanna try?



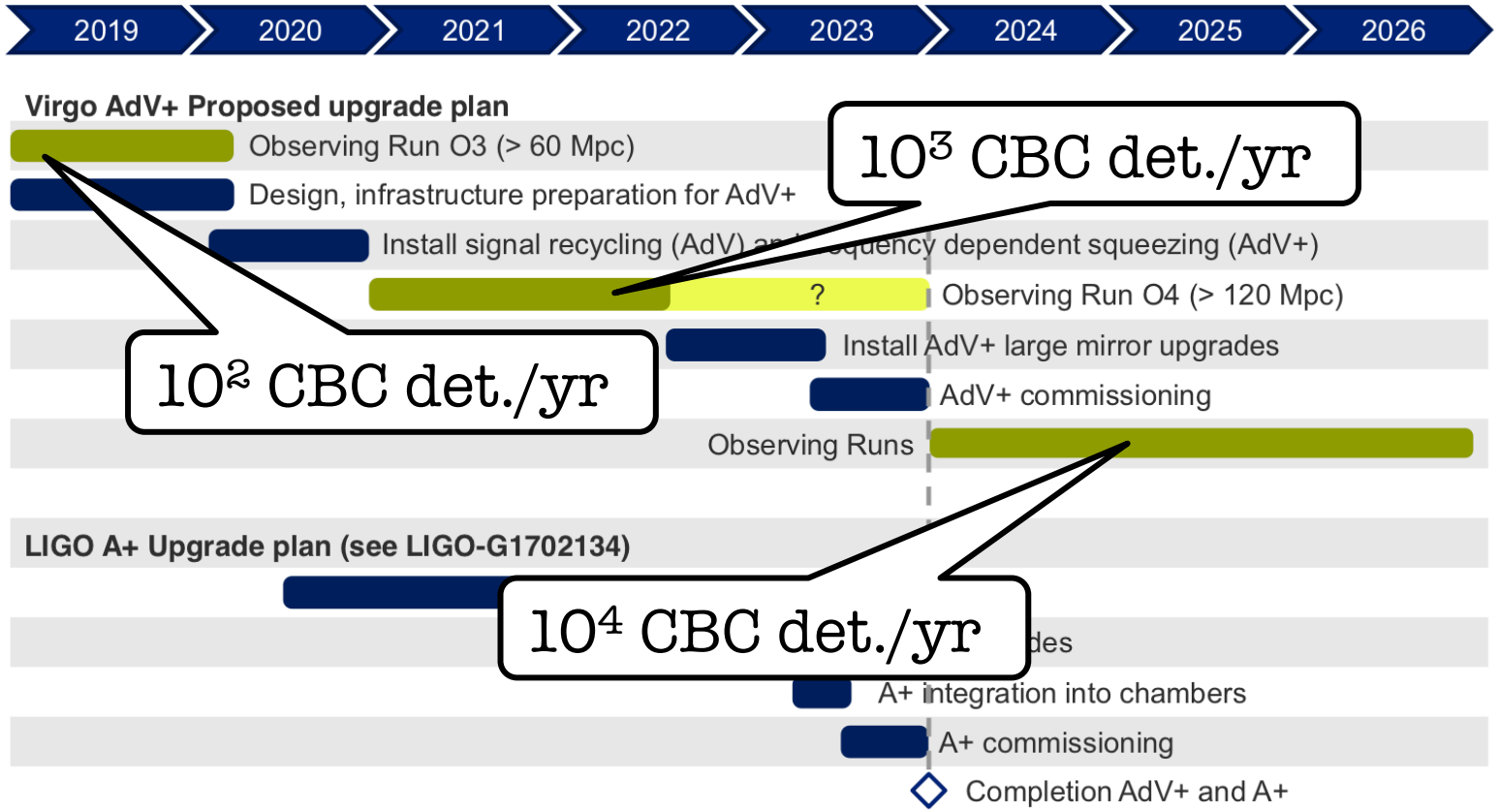
The Aero-Widow

Spare slides

LIGO – Virgo timeline

AdV+ to be carried out in parallel with LIGO's A+ upgrade

Five year plan for observational runs, commissioning and upgrades



Note: duration of O4 has not been decided at this moment
AdV+ is part of a strategy to go from 2nd generation to Einstein Telescope