

The Virgo Data Quality Reports (DQRs): an HTCondor automated framework to vet gravitational-wave candidates

HTCondor Workshop Autumn 2023, September 21st, 2023

Nicolas Arnaud (nicolas.arnaud@ijclab.in2p3.fr)

IJCLab – Laboratoire de Physique des 2 Infinis Irène Joliot-Curie
(Université Paris-Saclay & CNRS/IN2P3)

EGO – European Gravitational Observatory
(Consortium: CNRS, INFN & NIKHEF)

Data Quality Report for S200311bg
Event: GPS = 1267963151.397788 (2020-03-11 11:58:53.397788+00:00 UTC), duration = 2.70s
DQR generation starting at 2020-03-11 12:06:31+00:00 UTC

Clickable buttons: GraceDB event, GraceDB joint LIGO-Virgo DQR, Condor monitoring

Color caption: pass [Data OK], fail [Data not OK], human_input_needed [No automated decision], error or bad_state [Code crash or processing problem], missing [Check still running], higher latency tier

Virgo DQR documentation: Introduction, Checks, FAQ, Instructions for shifters and RRT, For experts, LIGO DQR documentation: Introduction

THE MOST IMPORTANT CHECKS, TO BE CHECKED FIRST

- Brute-force coherence reports (bruce)
- Is the candidate GPS time in Virgo suspicious?
- Status of the Virgo systems
- UPV on last 24 hours
- Virgo data quality antifiags
- Virgo data quality flags
- Virgo glitch characterization

What was the LIGO noise stationarity while the candidate signal was observed?

What was the Virgo noise stationarity while the candidate signal was observed? What was the Virgo status while the candidate signal was observed?

What was the status of the environment around Virgo at the time of the candidate?

```
09/19/23 04:34:24 *****
09/19/23 04:34:24 ** condor_scheduniv_exec.6122842.0 (CONDOR_DAGMAN) STARTING UP
09/19/23 04:34:24 ** /usr/bin/condor_dagman
09/19/23 04:34:24 ** SubsystemInfo: name=DAGMAN type=DAGMAN(9) class=CLIENT(2)
09/19/23 04:34:24 ** Configuration: subsystem=DAGMAN local:<NONE> class=CLIENT
09/19/23 04:34:24 ** $CondorVersion: 10.0.7 2023-07-25 BuildID: 664317 PackageID: 10.0.7-1 $
09/19/23 04:34:24 ** $CondorPlatform: x86_64_CentOS7 $
09/19/23 04:34:24 ** PID = 4051368
09/19/23 04:34:24 ** Log last touched time unavailable (No such file or directory)
09/19/23 04:34:24 *****
09/19/23 04:34:24 Using config source: /etc/condor/condor_config
09/19/23 04:34:24 Using local config sources:
09/19/23 04:34:24 /etc/condor/config.d/00-htcondor-9.0.config
09/19/23 04:34:24 /etc/condor/config.d/10-stash-plugin.conf
09/19/23 04:34:24 /etc/condor/condor_config.local
09/19/23 04:34:24 /olusers/virgorun/.condor/user_config
09/19/23 04:34:24 config Macros = 236, Sorted = 236, StringBytes = 7444, TablesBytes = 8568
09/19/23 04:34:24 CLASSAD_CACHING is ENABLED
09/19/23 04:34:24 Daemon Log is logging: D_ALWAYS D_ERROR D_STATUS
09/19/23 04:34:24 DaemonCore: No command port requested.
09/19/23 04:34:24 Using DAGMan config file: /data/dev/web/detchar/dqr/pre04/202309/S230919bf/dag/dag.config
09/19/23 04:34:24 DAGMAN_USE_STRICT setting: 1
09/19/23 04:34:24 DAGMAN_VERBOSITY setting: 3
09/19/23 04:34:24 DAGMAN_DEBUG_CACHE_SIZE setting: 5242880
09/19/23 04:34:24 DAGMAN_DEBUG_CACHE_ENABLE setting: False
09/19/23 04:34:24 DAGMAN_SUBMIT_DELAY setting: 0
09/19/23 04:34:24 DAGMAN_MAX_SUBMIT_ATTEMPTS setting: 6
09/19/23 04:34:24 DAGMAN_STARTUP_CYCLE_DETECT setting: False
09/19/23 04:34:24 DAGMAN_MAX_SUBMITS_PER_INTERVAL setting: 100
```

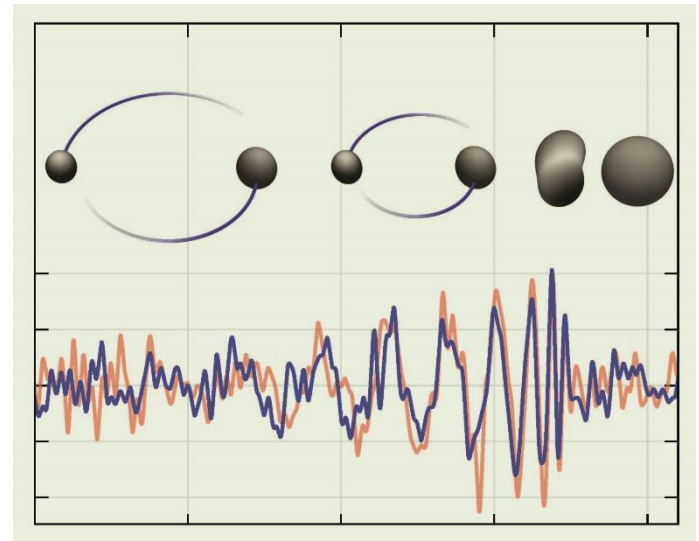
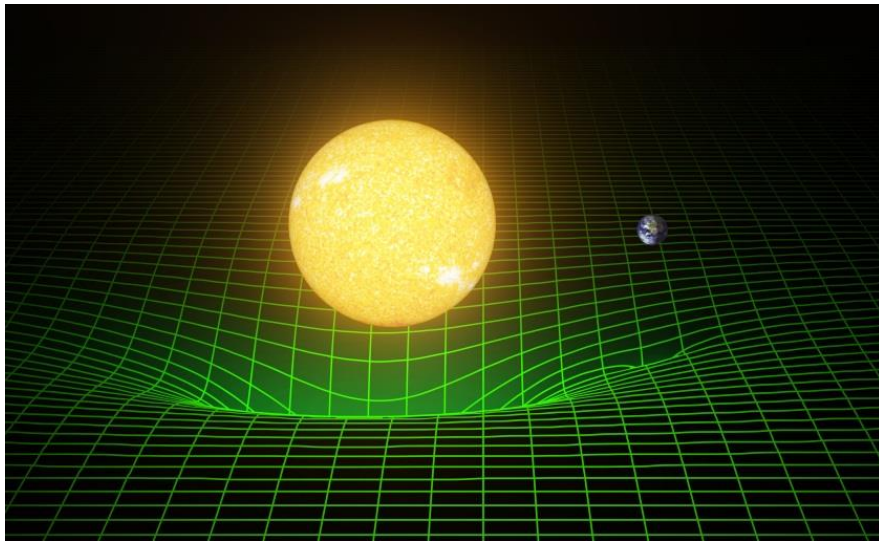


Introduction

- I am a member of the **Virgo collaboration**
 - **IJCLab** is my home lab, but I'm currently seconded to **EGO** in Italy
- Also member of the “meta-collaboration” **LIGO-Virgo-KAGRA** – like **Peter Couvares**
- I am a **physicist**
- Although I do **mostly software work**, I would **not** consider me as a **good developer**
- The **Data Quality Report (DQR)** was a **new project** in **2018**
 - I did like the idea of **developing from scratch** something that would be(come) important to assess the quality of our data, in particular in low-latency
- **I knew nothing about Condor** at the time
 - **I learned the basics on the job**
 - ◆ And kind of stopped when my framework started working fine...
 - **Limited personpower in Virgo:**
 - no computing professional available to tackle (or supervise) the project**
- The **DQR** has been my **main software project** these past five years
 - **Help from colleagues for data quality checks**
 - **Continuous support from the EGO computing department**

Gravitational waves

- “Spacetime tells matter how to move; matter tells spacetime how to curve”
John Archibald Wheeler (1990)



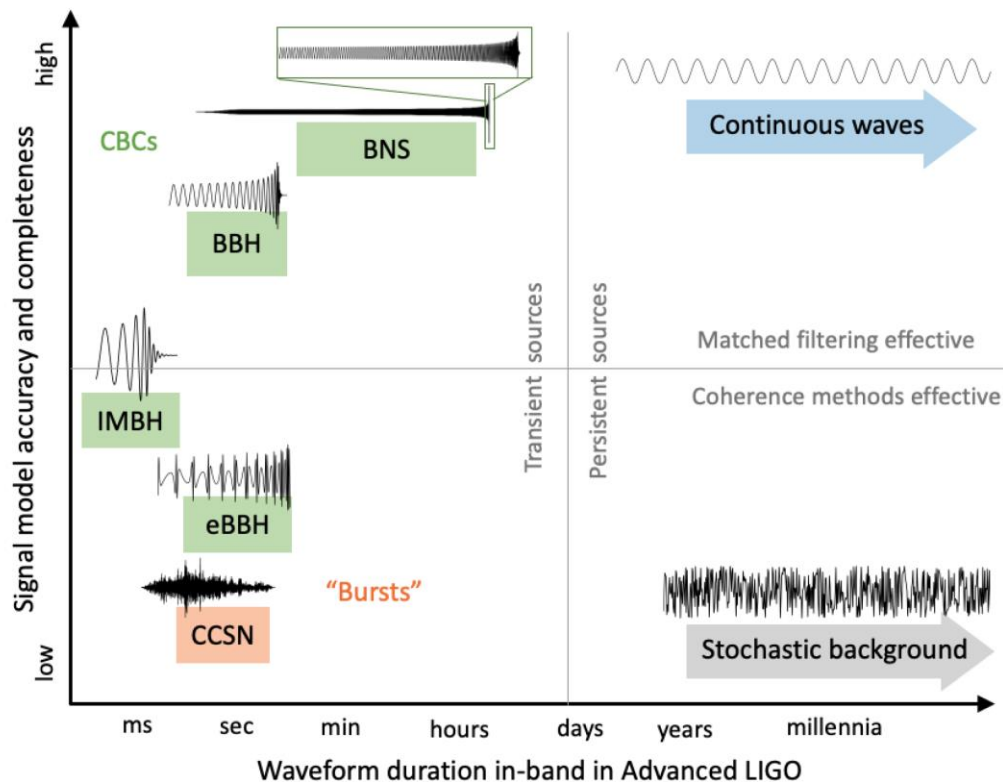
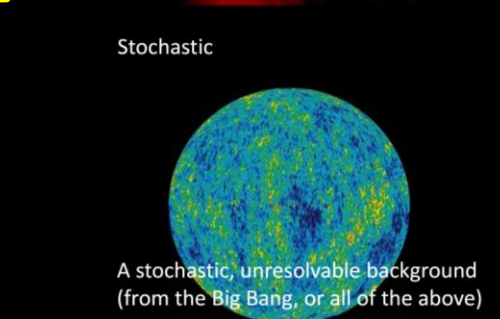
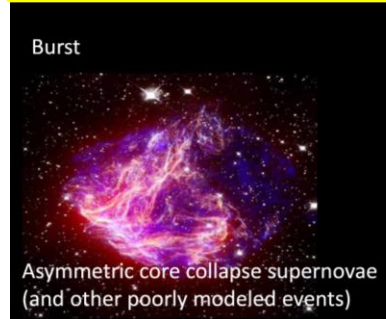
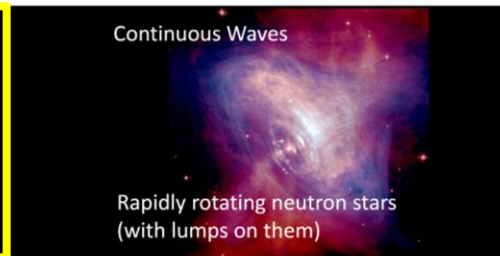
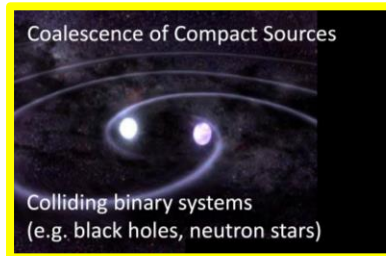
- **Gravitational Waves (GWs)** are ripples in the fabric of the spacetime
 - Propagation at the speed of light
 - Amplitude scales like $1 / \text{distance}$
 - Gravitation very weak \leftrightarrow Spacetime extremely rigid
 - Although pretty much all accelerated body emit GW,
nothing on Earth can produce enough GW to be detectable

GW sources

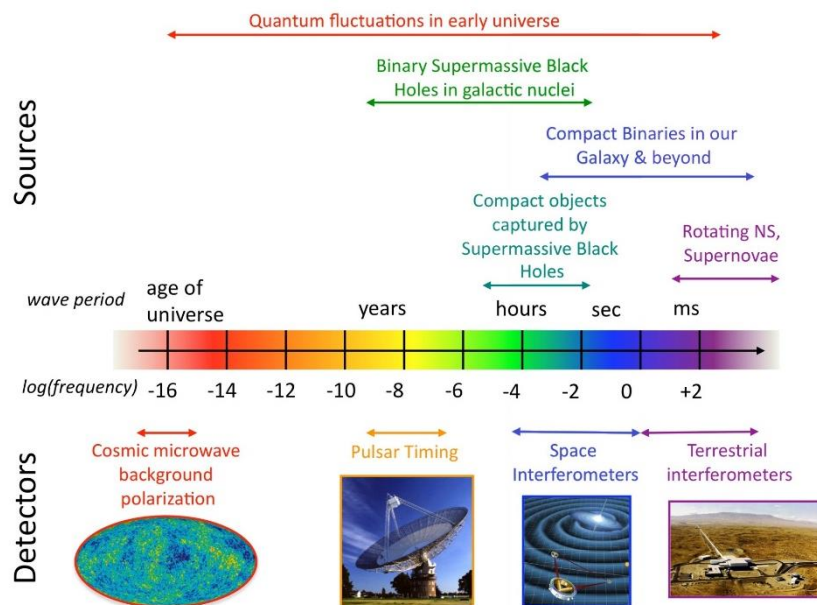
- **Classification**

- **Transient** / **Continuous**
- **Modeled** / **Unmodeled**

→ **Drives** the choice of the **data analysis** methods

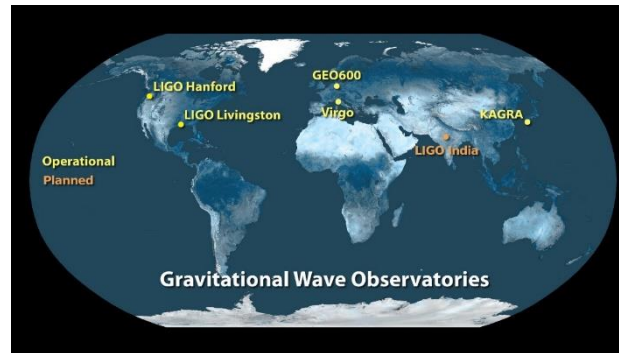


- **GW frequency contents / evolution**
- **Detector bandwidth**



Detectors

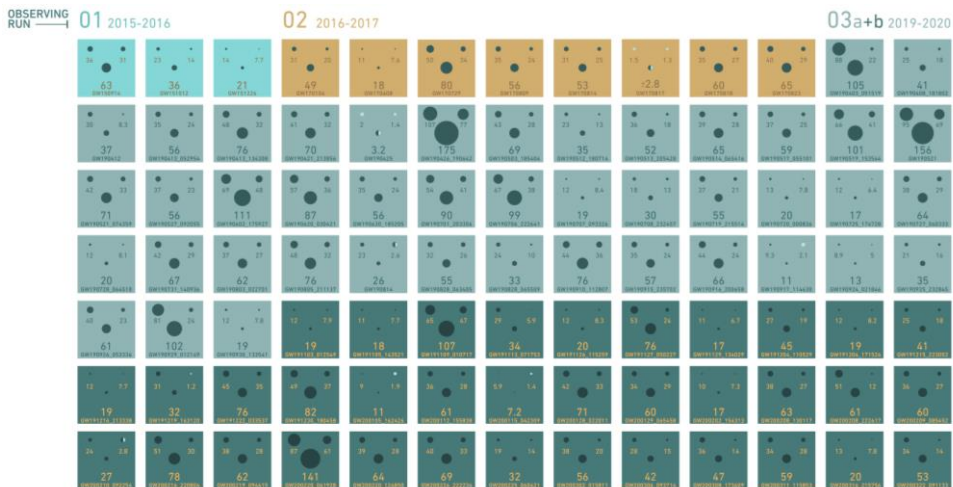
- Only **GW sources from the cosmos** can be powerful enough to be detected
 - **Very far away**: **GW signals** received on Earth are **tiny**
- **Extremely sensitive detectors** required
 - **Giant ground-based interferometric detectors**
 - ◆ **A passing GW distorts spacetime locally, thus changes the interference pattern**
 - Signal at the interferometer output port
 - Detectors see more or less the whole sky – unlike telescopes
 - **Network of instruments analyzing data jointly** are **much more powerful**



- **LIGO-Virgo-KAGRA (LVK) global network**

Detections

- 1915-2015: **one century** between General Relativity and **first GW detection**
 - **GW150914**: 1st GW detection – Binary black hole merger
 - **GW170817**: 1st binary black hole merger – Multi-messenger astronomy with GW
- **90 events** in the latest catalog issue: **GWTC-3**



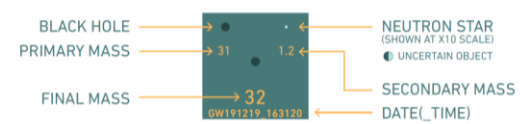
Note that the mass estimates shown here do not include uncertainties, which is why the final mass is sometimes larger than the sum of the primary and secondary masses. In actuality, the final mass is smaller than the primary plus the secondary mass.



ARC Centre of Excellence for Gravitational Wave Discovery

The events listed here pass one of two thresholds for detection. They either have a probability of being astrophysical of at least 50%, or they pass a false alarm rate threshold of less than 1 per 3 years.

KEY



UNITS ARE SOLAR MASSES
1 SOLAR MASS = 1.989×10^{30} kg

- New data taking period ongoing since May 2023: **O4**
 - Only with the two LIGO detectors so far
- **New alerts regularly found**

LIGO/Virgo/KAGRA Public Alerts

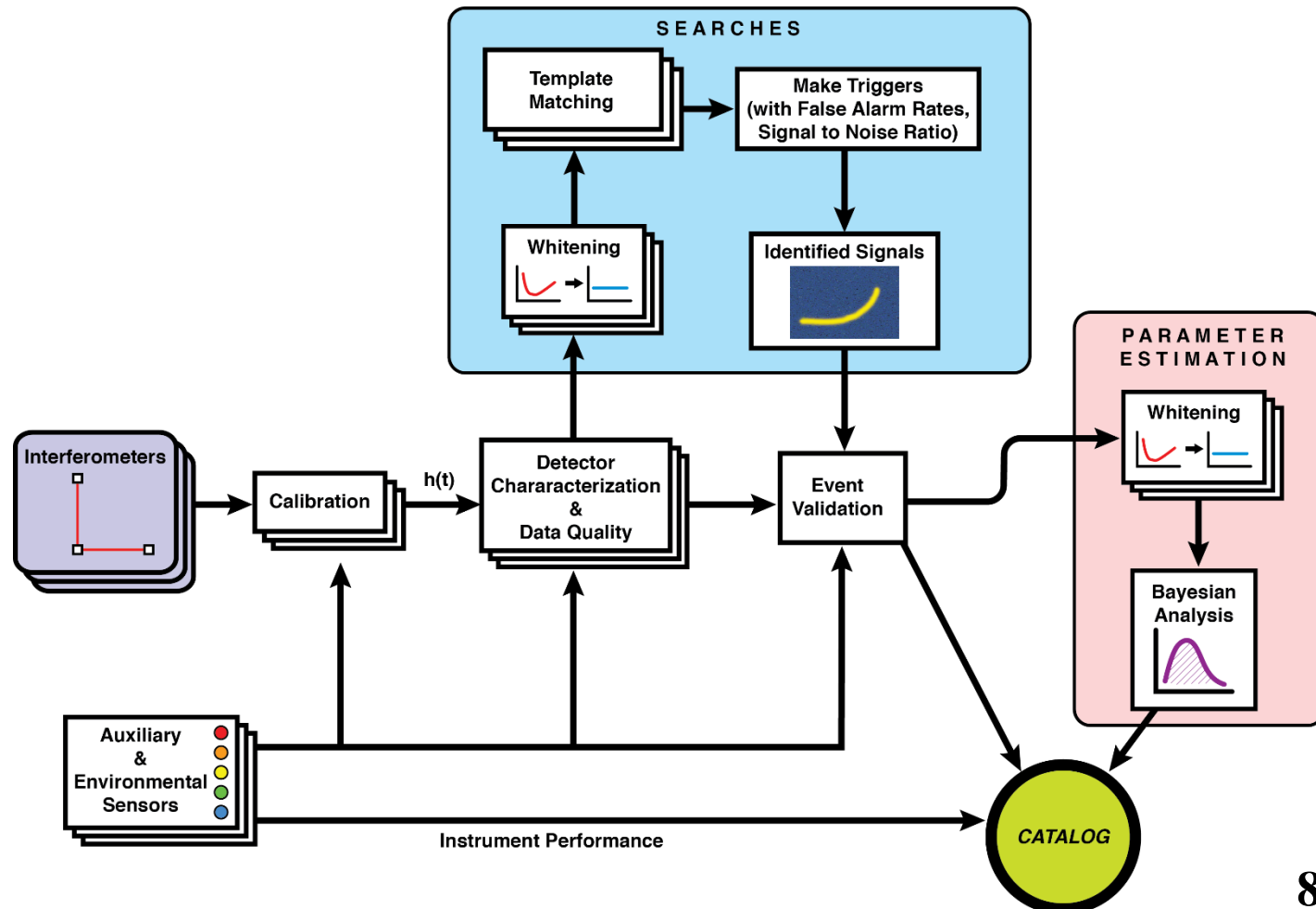
- More details about public alerts are provided in the [LIGO/Virgo/KAGRA Alerts User Guide](#).
- Retractions are marked in **red**. Retraction means that the candidate was manually vetted and is no longer considered a candidate of interest.
- Less-significant events are marked in **grey**, and are not manually vetted. Consult the [LVK Alerts User Guide](#) for more information on significance in O4
- Less-significant events are not shown by default. Press "Show All Public Events" to show significant and less-significant events.

O4 Significant Detection Candidates: **42** (51 Total - 9 Retracted)

O4 Low Significance Detection Candidates: **836** (Total)

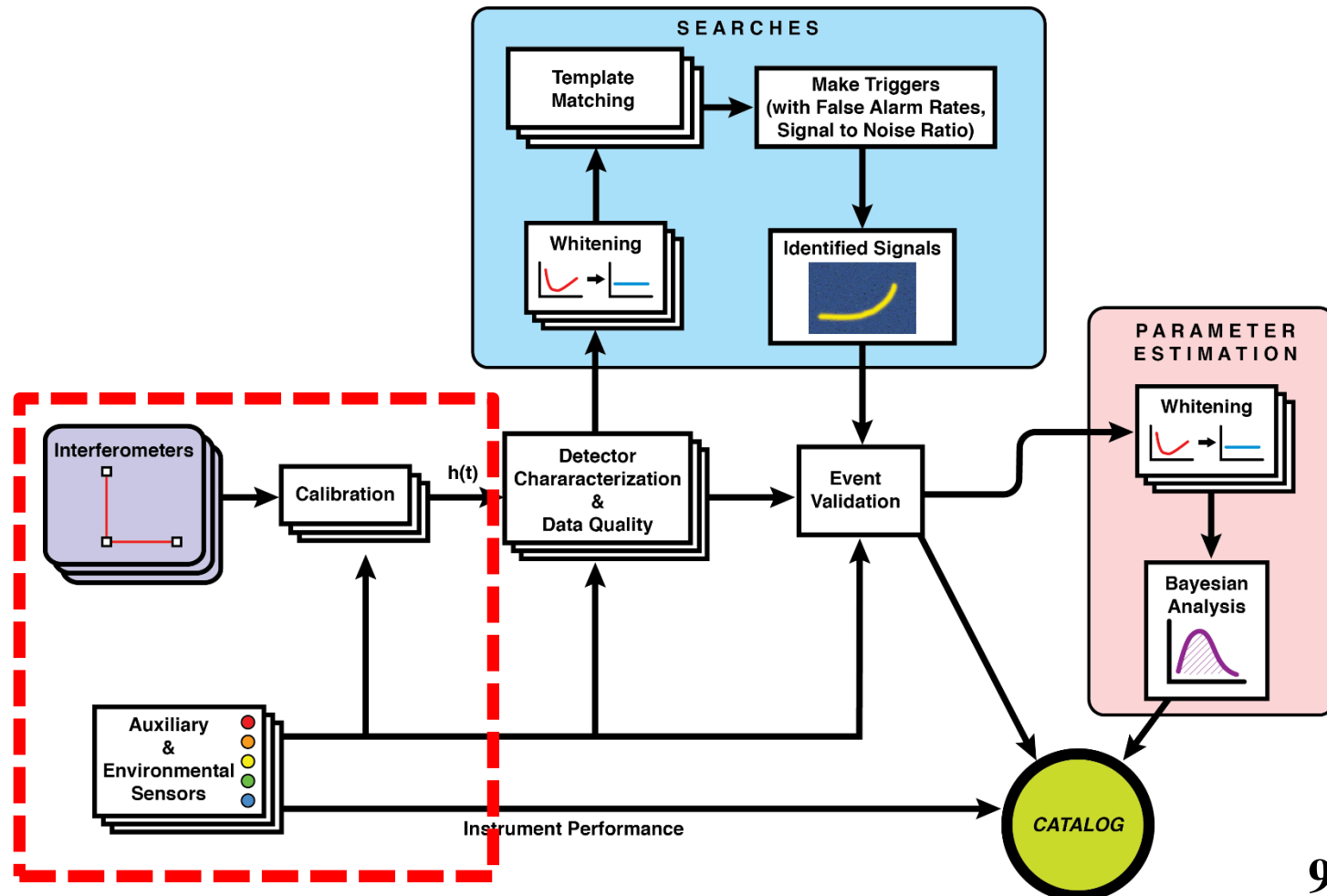
LVK dataflow

- From: A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals
 - [B. P. Abbott et al., 2020 *Class. Quantum Grav.* 37 055002](#)



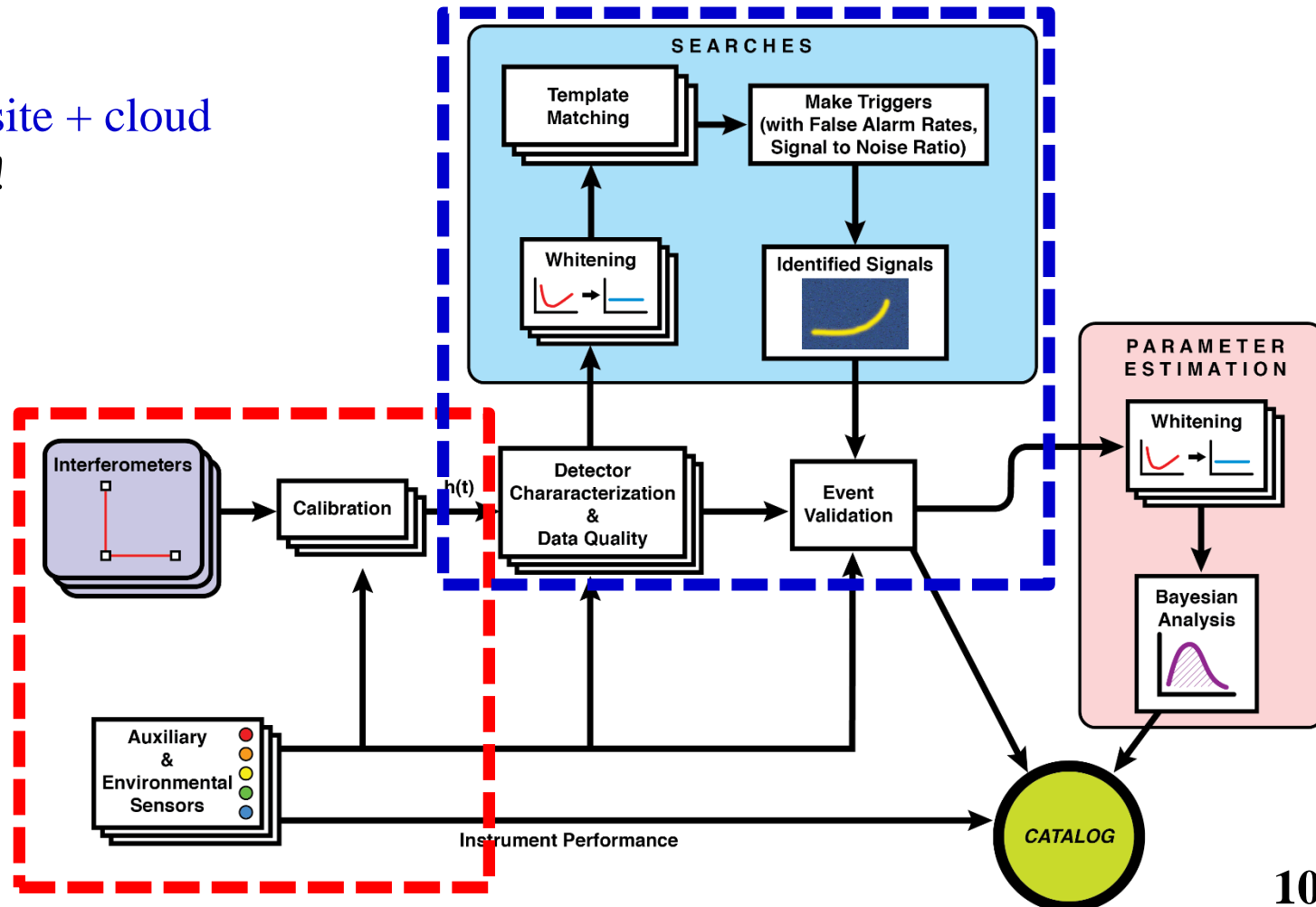
LVK dataflow

- From: **A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals**
 - [B. P. Abbott et al., 2020 *Class. Quantum Grav.* **37** 055002](#)
- Online: onsite



LVK dataflow

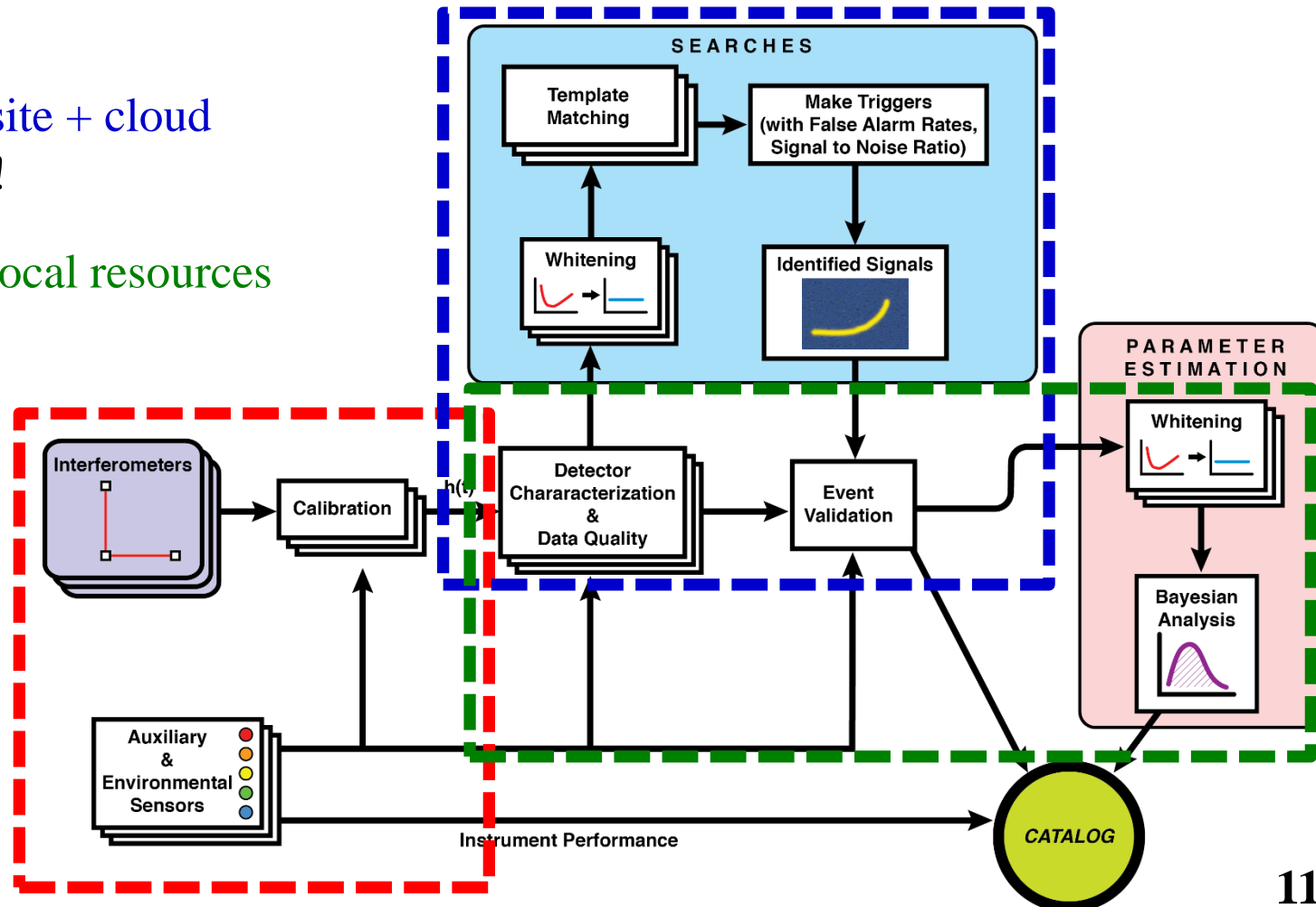
- From: **A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals**
 - [B. P. Abbott et al., 2020 *Class. Quantum Grav.* **37** 055002](#)
- **Online: onsite**
- **Low latency: onsite + cloud**
 - **DQR: this talk!**



LVK dataflow

- From: **A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals**
 - [B. P. Abbott et al., 2020 *Class. Quantum Grav.* **37** 055002](#)

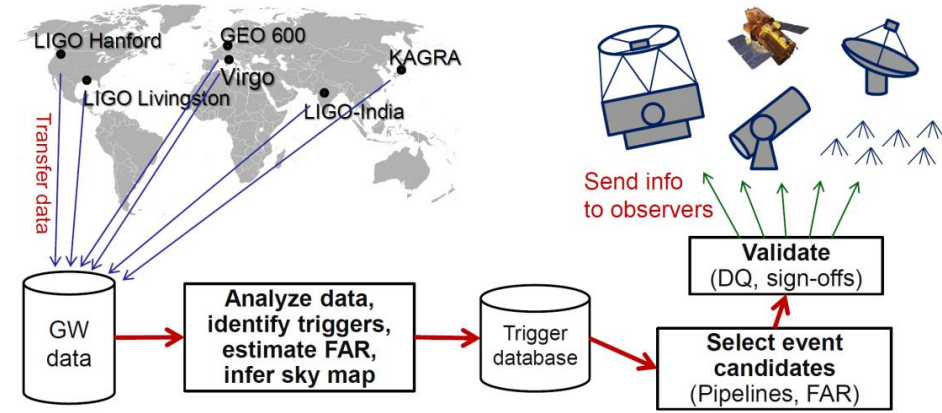
- **Online: onsite**
- **Low latency: onsite + cloud**
 - **DQR: this talk!**
- **Offline: HTC + local resources**



- **Many monitoring levels**
 - **Detector**
 - **Network**
 - **Analyses**

Multi-messenger astronomy with GW

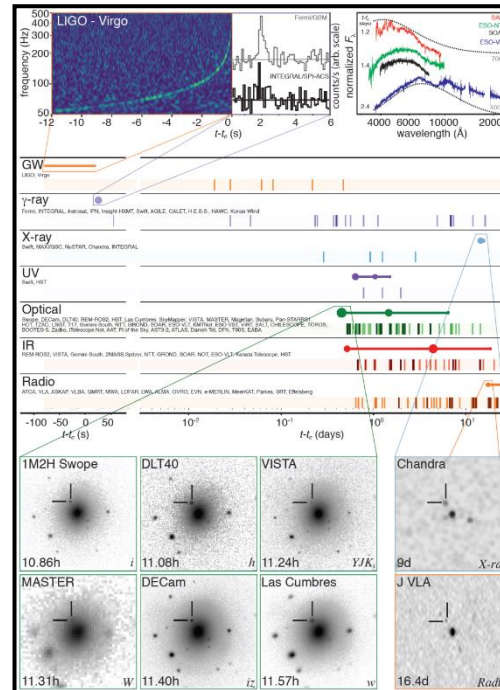
- **GW**: a **new messenger from the cosmos**
 - **Light** (full electromagnetic spectrum), cosmic rays, neutrinos



→ A source can emit different messengers

- **Complementary signatures**
 - ◆ Interest in **observing all of them**
 - ◆ From **as early as possible** after the event has “occurred”

- **Exceptional** (and **unique so far**) example: **GW170817**



GW170817

Binary neutron star merger
A LIGO / Virgo gravitational wave detection with associated electromagnetic events observed by over 70 observatories.

Distance
130 million light years

Discovered
17 August 2017

Type
Neutron star merger

12:41:04 UTC
A gravitational wave from a binary neutron star merger is detected.

gamma ray burst
A short gamma ray burst is an intense beam of gamma ray radiation which is produced just after the merger.

+2 seconds
A gamma ray burst is detected.

+10 hours 52 minutes
A new bright source of optical light is detected in a galaxy called NGC 4993, in the constellation of Hydra.

+11 hours 36 minutes
Infrared emission observed.

+15 hours
Bright ultraviolet emission detected.

+9 days
X-ray emission detected.

+16 days
Radio emission detected.

Au
The observation of a kilonova allowed us to show that neutron star mergers could be responsible for the production most of the heavy elements, like gold, in the universe.

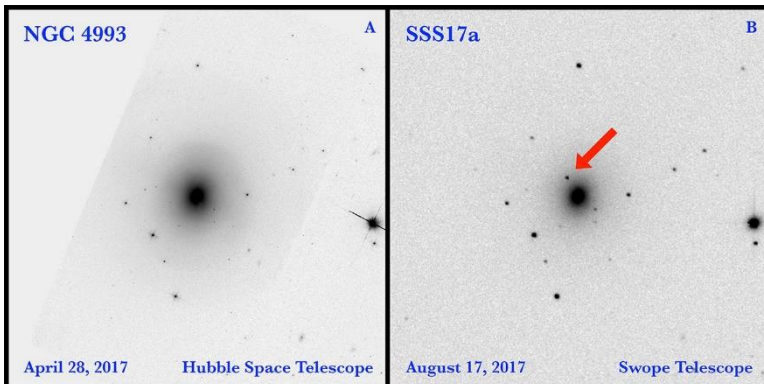
radio remnant
As material moves away from the merger it produces a shockwave in the interstellar medium - the turbulent material between stars. This produces emission which can last for years.

GW170817 allows us to measure the expansion rate of the universe directly using gravitational waves for the first time...

Detecting gravitational waves from a neutron star merger allows us to find out more about the structure of these unusual objects.

This multimessenger event provides confirmation that neutron star mergers can produce short gamma ray bursts.

Observing both electromagnetic and gravitational waves from the event provides compelling evidence that gravitational waves travel at the same speed as light.



Low-latency alerts

- **General Coordinates Network (GCN)**

- <https://gcn.nasa.gov>

- **Real-time processing of LVK data**

- Dedicated data analysis pipelines searching for **transient GW events**

- **Latency is the main challenge**

for the public alert

- The lower, the better

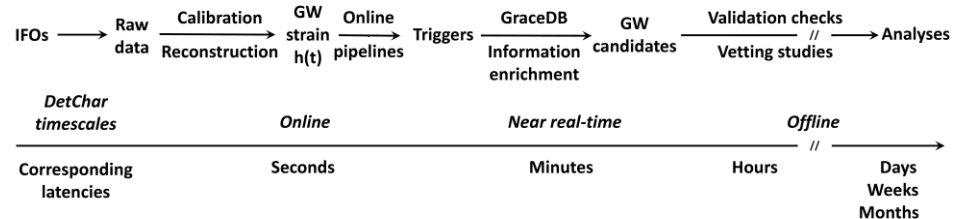
- An **alert** must be **informative** for the **astronomy community**

- Automated alerts later found not to originate from the cosmos are **retracted**

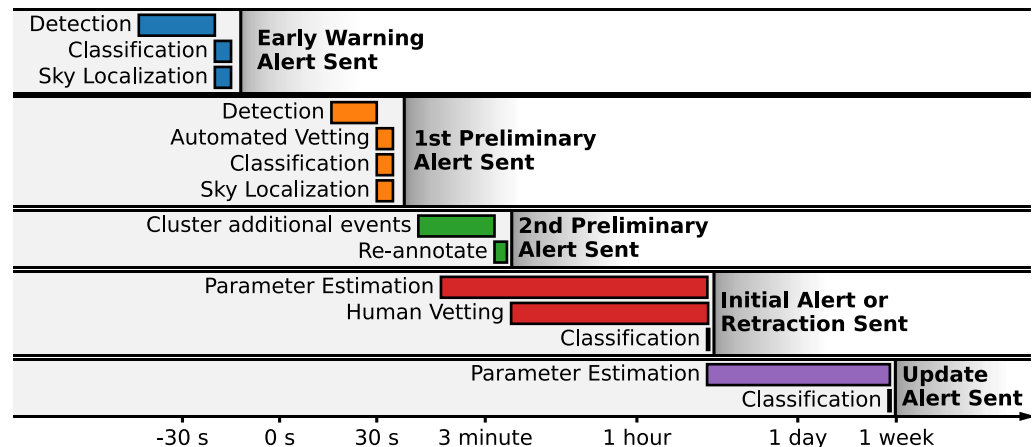
- Central database: **GraceDB**

- **Gravitational wave candidate event DataBase**

→ **Public portal:** <https://gracedb.ligo.org/superevents/public/O4>



Time relative to gravitational-wave merger





Virgo @ EGO

- **E**uropean **G**ravitational **O**bservatory (**EGO**):
the lab hosting the **Virgo** detector
- Recent snapshot: ~800 members / ~530 authors
- ~140 participating institutions
from 15 countries
 - Gathered in ~35 groups
from 9 countries



The EGO HTCondor farm (1/2)

- For more information: **Giuseppe Di Biase** dibiase@ego-gw.it
- **Architecture**
 - **HTCondor**
 - ◆ \$CondorVersion: 10.0.7 2023-07-25 BuildID: 664317 PackageID: 10.0.7-1 \$
 - ◆ \$CondorPlatform: x86_64_CentOS7 \$
 - **1 Central Manager Node** (Negotiator, Collector)
 - ◆ 2 VCPU
 - ◆ 4 GB RAM
 - **1 Submit Node** (Schedd)
 - ◆ 4 VCPU
 - ◆ 16 GB RAM
 - **352 Execute Nodes** (Startd)
 - ◆ 8 VCPU (Intel(R) Xeon(R) Gold 6148 CPU @ 2.40GHz)
 - ◆ 26GB RAM

The EGO HTCondor farm (2/2)

- **Configuration**

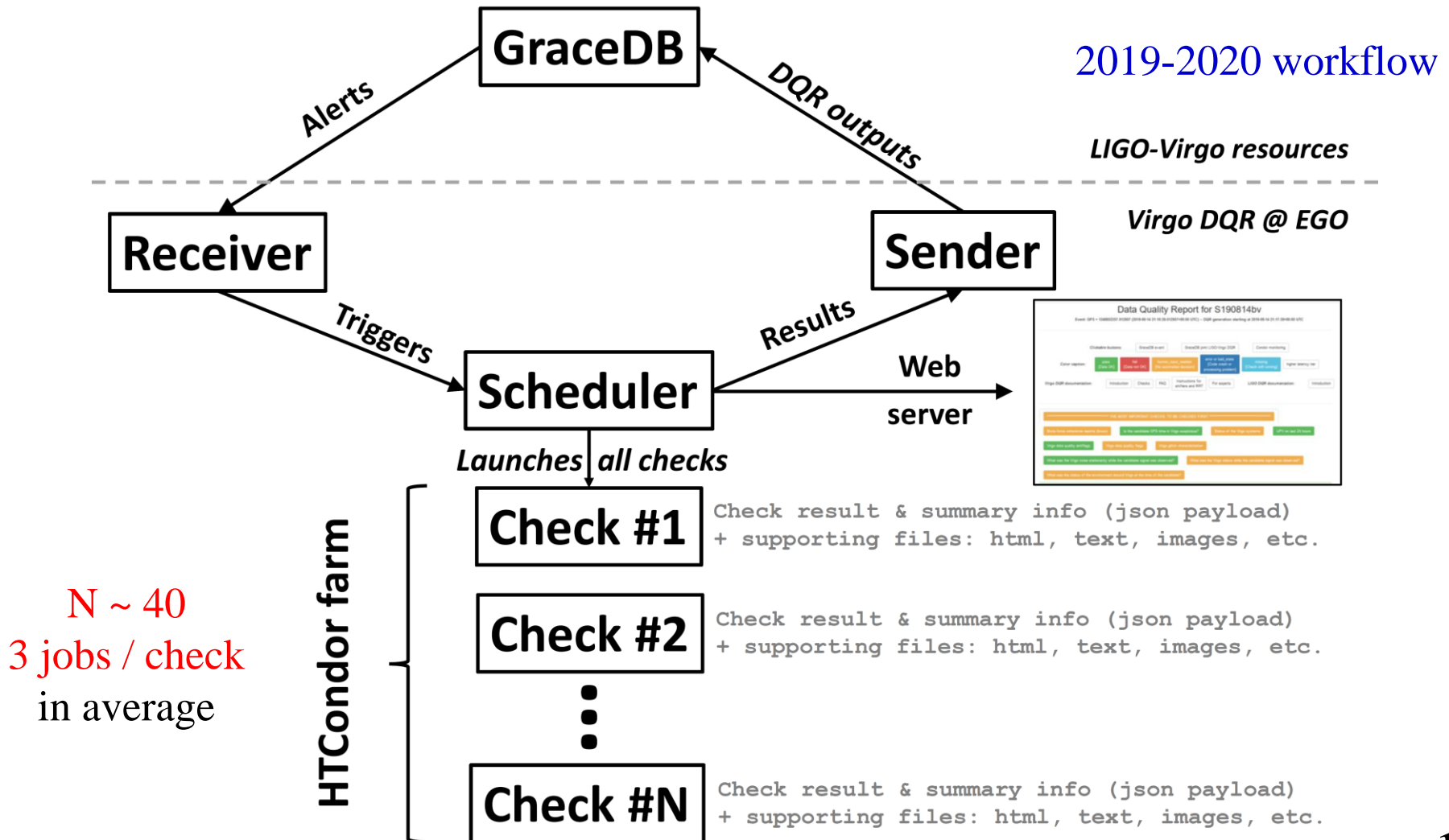
- All nodes are equal and mount the usual filesystems with r/w access to the storage farm volumes – no need to enable the HTCondor internal file transfer mechanism
- Jobs submitted only in the (default) "vanilla" universe: no need to relink
- Each machine has 8 Condor "slots" with 1 cpu 100% and 8GB of memory
- Currently no different "queues" defined
- The configuration of queues with specific characteristics for different applications requirements is the target of the tests
- To specify the pathnames for the "log" , "output" , "error" Condor files, one must put the absolute path – otherwise the NFS automounted path does not work
 - ♦ Put "initialdir = _working_directory_path_" in submit file
 - ♦ The "executable path" in the submit file must always be absolute
- DAGMan: only works if dag submitted from directory where the dag file is located

- **Policy**

- GROUP with Dynamic quota
- Slots assignment by group names
- Accept surplus
- RANK assignment by Experiment custom ClassAd

The Virgo Data Quality Report framework

- Set of automated checks triggered upon receiving an alert from GraceDB



$N \sim 40$
3 jobs / check
in average

HTCondor DAG + parent/child relationship

- DAG generation
 - Search for input data
 - Generate checks in parallel to minimize latency

```
[upv] Generate UPV on the last 24 hours
[upv]   - Do not run that (long) check further if Virgo was clearly not running at the time of the event
[upv] Using CMT Conda executable /users/narnaud/Software/RRT/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/worth_running_or_not-conda
[upv] Event is long: increasing nsecondsbackward from 10 s to 12 s
[upv]   - generate parameter file
[upv]   - Delay the job by 0 seconds to make sure that Omicron triggers are available on disk
[upv] Using CMT Conda executable /users/narnaud/Software/RRT/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/sleep-conda
[upv]   - generate upv .sub file
[upv]   - generate json .sub file
[upv] Using CMT Conda executable /users/narnaud/Software/RRT/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/read_segment_json-conda
[upv] TEST: ifo_checks=V1
[upv] Results upload location: CIT
[upv] Upload results
[upv] Adding line "UPV on last 24 hours | upv | 1 | high latency | Is there a high probability that a glitch was present based on
statistical inference of auxiliary information? | V1 | pass human_input_needed irrelevant | UPV on last 24 hours" to file /data/dev
/web/detchar/dqr/pre04/202309/S230919bf/config.txt
[upv] TIME to setup check upv = .550823775 s [UNIX time range: 1695090857.836443711 -> 1695090858.387267486]
```

HTCondor DAG + parent/child relationship

- **DAG generation**
 - Search for input data
 - Generate checks in parallel to **minimize latency**

```
[upv] Generate UPV on the last 24 hours
[upv]   - Do not run that (long) check further if Virgo was clearly not running at the time of the event
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/worth_running_or_not-conda
[upv] Event is long: increasing nsecondsbackward from 10 s to 12 s
[upv]   - generate parameter file
[upv]   - Delay the job by 0 seconds to make sure that Omicron triggers are available on disk
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/sleep-conda
[upv]   - generate upv .sub file
[upv]   - generate json .sub file
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/read_segment_json-conda
[upv] TEST: ifo_checks=V1
[upv] Results upload location: CIT
[upv] Upload results
[upv] Adding line "UPV on last 24 hours | upv | 1 | high latency | Is there a high probability that a glitch was present based on
statistical inference of auxiliary information? | V1 | pass human_input_needed irrelevant | UPV on last 24 hours" to file /data/dev
/web/detchar/dqr/pre04/202309/S230919bf/config.txt
[upv] TIME to setup check upv = .550823775 s [UNIX time range: 1695090857.836443711 -> 1695090858.387267486]
```

- **Example of DAG snippet**

```
JOB S200311bg_upv upv.sub
VARS S200311bg_upv initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv 1
JOB S200311bg_upv_exe upv_exe.sub
VARS S200311bg_upv_exe initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_exe 2
PARENT S200311bg_upv CHILD S200311bg_upv_exe
JOB S200311bg_upv_json upv_json.sub
VARS S200311bg_upv_json initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_json 1
PARENT S200311bg_upv_exe CHILD S200311bg_upv_json
JOB S200311bg_upv_upload upv_upload.sub
VARS S200311bg_upv_upload initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_upload 5
PARENT S200311bg_upv_json CHILD S200311bg_upv_upload
```

HTCondor DAG + parent/child relationship

- **DAG generation**
 - Search for input data
 - Generate checks in parallel to **minimize latency**

```
[upv] Generate UPV on the last 24 hours
[upv]   - Do not run that (long) check further if Virgo was clearly not running at the time of the event
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/worth_running_or_not-conda
[upv] Event is long: increasing nsecondsbackward from 10 s to 12 s
[upv]   - generate parameter file
[upv]   - Delay the job by 0 seconds to make sure that Omicron triggers are available on disk
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/sleep-conda
[upv]   - generate upv .sub file
[upv]   - generate json .sub file
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/read_segment_json-conda
[upv] TEST: ifo_checks=V1
[upv] Results upload location: CIT
[upv] Upload results
[upv] Adding line "UPV on last 24 hours | upv | 1 | high latency | Is there a high probability that a glitch was present based on
statistical inference of auxiliary information? | V1 | pass human_input_needed irrelevant | UPV on last 24 hours" to file /data/dev
/web/detchar/dqr/pre04/202309/S230919bf/config.txt
[upv] TIME to setup check upv = .550823775 s [UNIX time range: 1695090857.836443711 -> 1695090858.387267486]
```

- **Example of DAG snippet**

```
JOB S200311bg_upv upv.sub
VARS S200311bg_upv initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv 1
JOB S200311bg_upv_exe upv_exe.sub
VARS S200311bg_upv_exe initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_exe 2
PARENT S200311bg_upv CHILD S200311bg_upv_exe
JOB S200311bg_upv_json upv_json.sub
VARS S200311bg_upv_json initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_json 1
PARENT S200311bg_upv_exe CHILD S200311bg_upv_json
JOB S200311bg_upv_upload upv_upload.sub
VARS S200311bg_upv_upload initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_upload 5
PARENT S200311bg_upv_json CHILD S200311bg_upv_upload
```

Preprocessing

HTCondor DAG + parent/child relationship

- **DAG generation**
 - Search for input data
 - Generate checks in parallel to **minimize latency**

```
[upv] Generate UPV on the last 24 hours
[upv]   - Do not run that (long) check further if Virgo was clearly not running at the time of the event
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/worth_running_or_not-conda
[upv] Event is long: increasing nsecondsbackward from 10 s to 12 s
[upv]   - generate parameter file
[upv]   - Delay the job by 0 seconds to make sure that Omicron triggers are available on disk
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/sleep-conda
[upv]   - generate upv .sub file
[upv]   - generate json .sub file
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/read_segment_json-conda
[upv] TEST: ifo_checks=V1
[upv] Results upload location: CIT
[upv] Upload results
[upv] Adding line "UPV on last 24 hours | upv | 1 | high latency | Is there a high probability that a glitch was present based on
statistical inference of auxiliary information? | V1 | pass human_input_needed irrelevant | UPV on last 24 hours" to file /data/dev
/web/detchar/dqr/pre04/202309/S230919bf/config.txt
[upv] TIME to setup check upv = .550823775 s [UNIX time range: 1695090857.836443711 -> 1695090858.387267486]
```

- Example of DAG snippet

```
JOB S200311bg_upv upv.sub
VARS S200311bg_upv initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv 1
JOB S200311bg_upv_exe upv_exe.sub
VARS S200311bg_upv_exe initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_exe 2
PARENT S200311bg_upv CHILD S200311bg_upv_exe
JOB S200311bg_upv_json upv_json.sub
VARS S200311bg_upv_json initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_json 1
PARENT S200311bg_upv_exe CHILD S200311bg_upv_json
JOB S200311bg_upv_upload upv_upload.sub
VARS S200311bg_upv_upload initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_upload 5
PARENT S200311bg_upv_json CHILD S200311bg_upv_upload
```

Processing

HTCondor DAG + parent/child relationship

- DAG generation
 - Search for input data
 - Generate checks in parallel to minimize latency

```
[upv] Generate UPV on the last 24 hours
[upv]   - Do not run that (long) check further if Virgo was clearly not running at the time of the event
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/worth_running_or_not-conda
[upv] Event is long: increasing nsecondsbackward from 10 s to 12 s
[upv]   - generate parameter file
[upv]   - Delay the job by 0 seconds to make sure that Omicron triggers are available on disk
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/sleep-conda
[upv]   - generate upv .sub file
[upv]   - generate json .sub file
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/read_segment_json-conda
[upv] TEST: ifo_checks=V1
[upv] Results upload location: CIT
[upv] Upload results
[upv] Adding line "UPV on last 24 hours | upv | 1 | high latency | Is there a high probability that a glitch was present based on
statistical inference of auxiliary information? | V1 | pass human_input_needed irrelevant | UPV on last 24 hours" to file /data/dev
/web/detchar/dqr/pre04/202309/S230919bf/config.txt
[upv] TIME to setup check upv = .550823775 s [UNIX time range: 1695090857.836443711 -> 1695090858.387267486]
```

- Example of DAG snippet

```
JOB S200311bg_upv upv.sub
VARS S200311bg_upv initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv 1
JOB S200311bg_upv_exe upv_exe.sub
VARS S200311bg_upv_exe initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_exe 2
PARENT S200311bg_upv CHILD S200311bg_upv_exe
JOB S200311bg_upv_json upv_json.sub
VARS S200311bg_upv_json initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_json 1
PARENT S200311bg_upv_exe CHILD S200311bg_upv_json
JOB S200311bg_upv_upload upv_upload.sub
VARS S200311bg_upv_upload initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_upload 5
PARENT S200311bg_upv_json CHILD S200311bg_upv_upload
```

Postprocessing

HTCondor DAG + parent/child relationship

- DAG generation
 - Search for input data
 - Generate checks in parallel to minimize latency

```
[upv] Generate UPV on the last 24 hours
[upv]   - Do not run that (long) check further if Virgo was clearly not running at the time of the event
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/worth_running_or_not-conda
[upv] Event is long: increasing nsecondsbackward from 10 s to 12 s
[upv]   - generate parameter file
[upv]   - Delay the job by 0 seconds to make sure that Omicron triggers are available on disk
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/sleep-conda
[upv]   - generate upv .sub file
[upv]   - generate json .sub file
[upv] Using CMT Conda executable /users/narnaud/Software/RRR/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/read_segment_json-conda
[upv] TEST: ifo_checks=V1
[upv] Results upload location: CIT
[upv] Upload results
[upv] Adding line "UPV on last 24 hours | upv | 1 | high latency | Is there a high probability that a glitch was present based on
statistical inference of auxiliary information? | V1 | pass human_input_needed irrelevant | UPV on last 24 hours" to file /data/dev
/web/detchar/dqr/pre04/202309/S230919bf/config.txt
[upv] TIME to setup check upv = .550823775 s [UNIX time range: 1695090857.836443711 -> 1695090858.387267486]
```

- Example of DAG snippet

```
JOB S200311bg_upv upv.sub
VARS S200311bg_upv initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv 1
JOB S200311bg_upv_exe upv_exe.sub
VARS S200311bg_upv_exe initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_exe 2
PARENT S200311bg_upv CHILD S200311bg_upv_exe
JOB S200311bg_upv_json upv_json.sub
VARS S200311bg_upv_json initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_json 1
PARENT S200311bg_upv_exe CHILD S200311bg_upv_json
JOB S200311bg_upv_upload upv_upload.sub
VARS S200311bg_upv_upload initialdir="/data/procdata/web/dqr/03b/202003/S200311bg/dag"
RETRY S200311bg_upv_upload 5
PARENT S200311bg_upv_json CHILD S200311bg_upv_upload
```

Upload of
the results

Example of .sub file

```
universe = vanilla
executable = /users/narnaud/Software/RRT/Virgo/Git/VirgoDQR/trunk/Linux-x86_64-CL7/bin/virgo_status-conda
arguments = "--event_gps 1263541772.015625 --event_id S230919bf \  
--data_stream /data/dev/web/detchar/dqr/preO4/202309/S230919bf/dqr_raw.ff1 \  
--output_dir /data/dev/web/detchar/dqr/preO4/202309/S230919bf --n_seconds_backward 12 --n_seconds_forward 10"
priority = 10
getenv = True
log_path = \  
/data/dev/web/detchar/dqr/preO4/202309/S230919bf/virgo_status/logs/$(cluster)-$(process)-virgo_status-$$ (Name)
error = $(log_path).err
output = $(log_path).out
notification = never

+Experiment = "DetChar"
+AccountingGroup= "virgo.prod.o4.detchar.transient.dqr"

# Remove jobs after 1800 seconds
periodic_remove = (JobStatus == 2) && (time() - EnteredCurrentStatus) > 1800

queue 1
```


Parsing the .dag.dagman.out file

- Python module **launched every minute** while the DAG is running, and **one last time after it is over**
→ Producing a **logfile updated at each scan** – example snippet below

```
*****
***** Scan 0064 *****
*** 2023-09-19 05:38:35 ***
*****

Parsing the DAG logfile...

Let's stop the monitoring of the Condor DAG now!

DAG dqr_S230919bf
  PID = 4051368
  Batch(es) = [('6122842', '2023-09-19 04:34:24')]
  Start date = 2023-09-19 04:34:24 LT
  End date = 2023-09-19 05:36:52 LT
  => duration = 3748 s

  Successful checks = 39 / 39

  Done checks = 39 / 39
  => The DAG is done!

* Check 0
- Info: "BP skymap"
- Question: "Was the detector in a nominal state?"
- Tier: 0 (low latency)
- 2 jobs: ['S230919bf_BP_skymap', 'S230919bf_BP_skymap_upload']
- Started at 2023-09-19 04:34:31 LT
- Status = Successful at 2023-09-19 04:36:02 LT
  => duration = 91 s
- Done = True
  => Check genealogy: job = S230919bf_BP_skymap [status: Successful] -> job = S230919bf_BP_skymap_upload [status: Successful]

* Check 1
- Info: "Glitch distribution in LIGO Hanford around the event -- print"
- Question: "Are known sources of noise without auxiliary witnesses active?"
- Tier: 1 (high latency)
- 4 jobs: ['S230919bf_omicronprinth1', 'S230919bf_omicronprinth1_chirp', 'S230919bf_omicronprinth1_exe_json', 'S230919bf_omicronprinth1_upload']
- Started at 2023-09-19 04:34:31 LT
- Status = Successful at 2023-09-19 04:36:23 LT
  => duration = 112 s
- Done = True
  => Check genealogy: job = S230919bf_omicronprinth1 [status: Successful] -> job = S230919bf_omicronprinth1_chirp [status: Successful] -> job = S230919bf_omicronprinth1_exe_json [status: Successful] -> job = S230919bf_omicronprinth1_upload [status: Successful]
```

Parsing the .dag.dagman.out file

- Python module **launched every minute** while the DAG is running and **one last time** after it is over
→ **Information stored in json file** as well

```
▼ dag:
  done: true
  duration: 3748
  end_date: "2023-09-19 05:36:52"
  end_gps: 1379129830
  id: "dag_S230919bf"
  n_checks: 39
  n_checks_done: 39
  n_initial_jobs: 0
  n_jobs: 123
  n_successful_checks: 39
  n_successful_jobs: 123
  pid: "4051368"
  start_date: "2023-09-19 04:34:24"
  start_gps: 1379126082
  evictions: {}
```

DAG info

```
▼ S230919bf_BP_skymap:
  ▼ ClassAd:
    CPUsUsage: "0.9713028963699825"
    DiskUsage: "1"
    MemoryUsage: "123"
    RemoteWallClockTime: "51.0"
    RequestCpus: "1"
    RequestDisk: "1"
    RequestMemory: "123"
  check: 0
  check_info: "BP skymap"
  cluster: "6122844"
  done: true
  duration: 61
  end_date: "2023-09-19 04:35:32"
  node: "o1node32"
  slot: "slot1_1"
  start_date: "2023-09-19 04:34:31"
  status: "Successful"
  sub: "BP_skymap.sub"
```

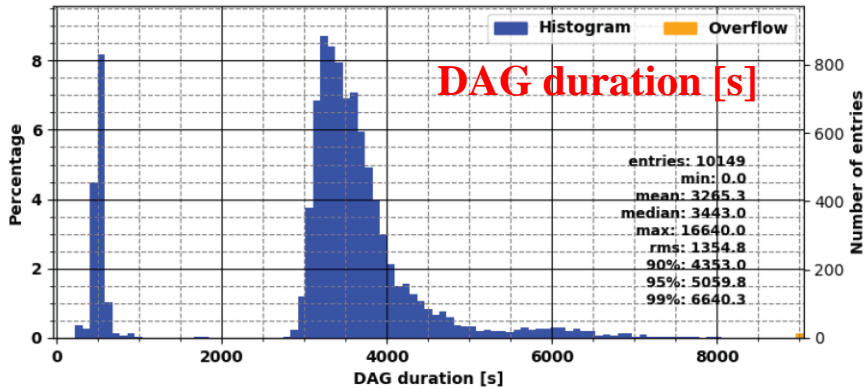
Job info

```
▼ 0:
  done: true
  duration: 91
  end_date: "2023-09-19 04:36:02"
  info: "BP skymap"
  ▼ n_jobs:
    0: "S230919bf_BP_skymap"
    1: "S230919bf_BP_skymap_upload"
  n_successful_jobs: 2
  question: "Was the detector in a nominal state?"
  start_date: "2023-09-19 04:34:31"
  status: "Successful"
  tier: "0"
  tier name: "low latency"
```

Check info

Long-term monitoring

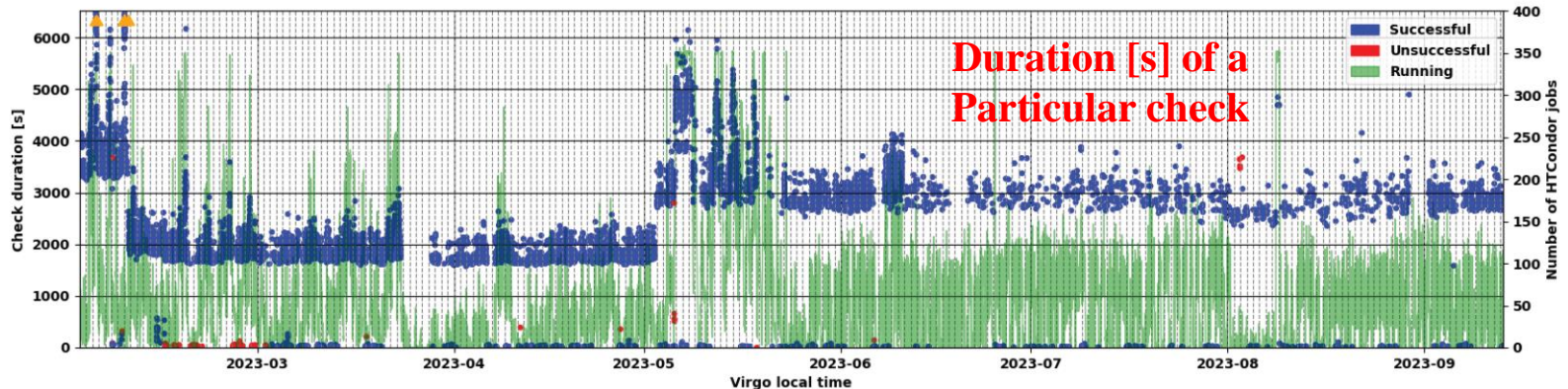
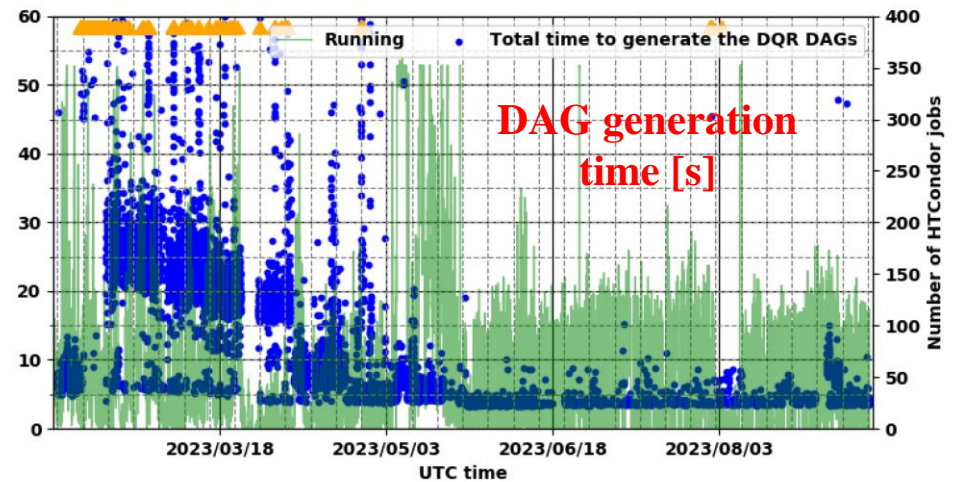
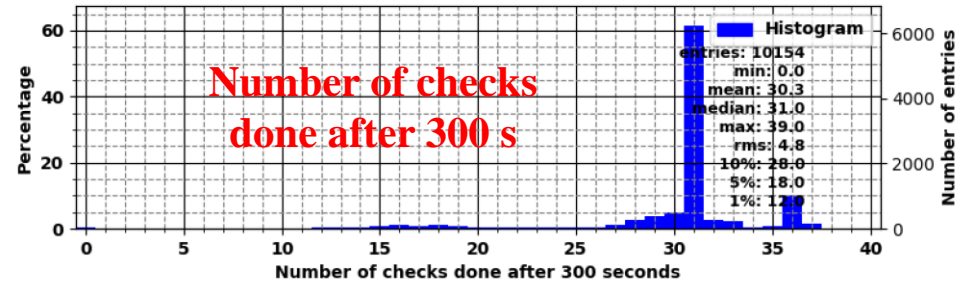
- Histograms



- Stripcharts

- Green: number of jobs running in the HTCondor farm

- ◆ Right y-axis



Conclusions

- **Virgo DQR** developed in 2018-2019
 - **Used extensively** during the **LIGO-Virgo O3 run** (2019/04 → 2020/03)
 - ◆ **Robust and fulfilling its requirements**
 - Now a **standard for data quality and event validation** in **LVK**
 - **Similar (but independent) DQR framework developed by LIGO, on the same timescale and for the same goals**
 - **Improved adiabatically** between 2020 and 2023 – **personpower limitation**
 - Thanks to **continuous Mock Data Challenges** in which **GW signal-enriched O3 data** are replayed and analyzed by the low-latency pipelines
 - **To be used extensively again, when Virgo joins the O4 run**
 - **HTCondor** easy and convenient to use, extremely reliable
 - DQR usage of the system quite **basic**
 - **All technical aspects handled** behind the (my) scene by **EGO computing dpt.**
- **Glad to have added this framework to my user knowledge!**

???

•

■ ■ ■ ■ ■ ■ ■ ■

•

■ ■ ■ ■ ■ ■ ■ ■

•

■ ■ ■ ■ ■ ■ ■ ■

•

■ ■ ■ ■ ■ ■ ■ ■

•

■ ■ ■ ■ ■ ■ ■ ■

•

■ ■ ■ ■ ■ ■ ■ ■