Computing contributions to Virgo

from UCLouvain

Andres Tanasijczuk

Workshop: Essential Technologies for the Einstein Telescope

Brussels, 27 November 2023

Outline

- Virgo computing domains: online, low-latency, offline

- LVK offline computing – International Gravitational-Wave observatories Network (IGWN)

- IGWN data management for offline computing

  › Open Science Data Federation (OSDF) from Open Science Grid (OSG)

  › Contribution from UCLouvain – OSDF origin (storage) server for Virgo aggregated h(t)

- IGWN resource usage and contributions for offline data processing

  › Local batch processing vs grid computing

  › Contribution from UCLouvain – computing cluster

Virgo computing domains

Online

Low-latency

Offline

Virgo computing domains

Online

- Happens in quasi-real time, in dedicated computing infrastructures close to the detectors
- Concerns detector control and monitoring, data acquisition and preparation, aggregation in manageable chunks and production of reduced sets of data for different uses

Virgo computing domains

Low-latency

- Common to the whole time-domain astronomical community

- Essential for multi-messenger astronomy - timely exchange of alerts between observatories

- LIGO, Virgo and KAGRA agreed to exchange their data also in low-latency and search for event candidates together

  › Data are transferred from the active detector sites to dedicated computing clusters at LIGO CIT and EGO, where low-latency pipelines search for candidate events

  › Whenever a candidate event is detected, its summary data are stored in a publicly accessible dedicated database (GraceDB) and an alert is distributed

  › Updates to the (super)event are made as more computational expensive tasks (e.g. precise sky localization and parameter estimation) are performed at the LIGO CIT, EGO and other clusters (e.g. Penn State, Wisconsin, CNAF, CIEMAT)

  › If the event is confirmed, a public alert is sent (including through NASA's GCN)

- Virgo low-latency computing needs at EGO for O4: ~2K virtual cores and 200 TB of storage

Virgo computing domains

Offline

- Asynchronous data analysis on pre-processed data, and simulations

    E.g. deep searches, parameter estimation, some detector characterization-related activities

Virgo computing domains

Offline

- Asynchronous data analysis on pre-processed data, and simulations

  E.g. deep searches, parameter estimation, some detector characterization-related activities

- During O1 and O2, offline computing activities were taking place on the proprietary uniform LIGO Data Grid platform and in the two main Virgo computing centres (CNAF and CC-IN2P3)
  › GW pipelines where designed to run on these specific environments (e.g. shared filesystem)
  › All these clusters used HTCondor batch system and some pipelines are now dependent on it (e.g. DAGs)

    High Throughput Computing (not High Performance Computing, i.e. no parallelism)

Virgo computing domains

Offline

- Asynchronous data analysis on pre-processed data, and simulations

    E.g. deep searches, parameter estimation, some detector characterization-related activities

- During O1 and O2, offline computing activities were taking place on the proprietary uniform LIGO Data Grid platform and in the two main Virgo computing centres (CNAF and CC-IN2P3)
    › GW pipelines where designed to run on these specific environments (e.g. shared filesystem)
    › All these clusters used HTCondor batch system and some pipelines are now dependent on it (e.g. DAGs)
        High Throughput Computing (not High Performance Computing, i.e. no parallelism)

- O3 increased computing requirements prompted a shift towards a more flexible infrastructure
    › A (WLCG-inspired) distributed infrastructure was implemented to serve the International GW observatories Network (IGWN)
        Strong support from Open Science Grid (OSG) that operates the system
    › Computing centres from collaborating institutions (e.g. UCLouvain) can share their computing resources via "the grid"
    › Ensure maximum interoperability (between LIGO, Virgo and KAGRA) with a common infrastructure
    › Workload management with HTCondor
        Allows for gradual transition from local running to fully grid-like distributed computing

Virgo computing domains

Offline

- Asynchronous data analysis on pre-processed data, and simulations

    E.g. deep searches, parameter estimation, some detector characterization-related activities

- During O1 and O2, offline computing activities were taking place on the proprietary uniform LIGO Data Grid platform and in the two main Virgo computing centres (CNAF and CC-IN2P3)
    - › GW pipelines where designed to run on these specific environments (e.g. shared filesystem)
    - › All these clusters used HTCondor batch system and some pipelines are now dependent on it (e.g. DAGs)
        High Throughput Computing (not High Performance Computing, i.e. no parallelism)

- O3 increased computing requirements prompted a shift towards a more flexible infrastructure
    - › A (WLCG-inspired) distributed infrastructure was implemented to serve the International GW observatories Network (IGWN)
        Strong support from Open Science Grid (OSG) that operates the system
    - › Computing centres from collaborating institutions (e.g. UCLouvain) can share their computing resources via "the grid"
    - › Ensure maximum interoperability (between LIGO, Virgo and KAGRA) with a common infrastructure
    - › Workload management with HTCondor
        Allows for gradual transition from local running to fully grid-like distributed computing
    - › Unfortunately, only a few pipelines are currently able to run on the grid (e.g. pyCBC)

IGWN architecture for (offline) distributed computing

IGWN requirements are the same as in High Energy Physics experiments

- Bulk data transfer and storage

- Software versioning, packaging and distribution

- Data distribution

- Data cataloguing and bookkeeping

    Rucio ??

- Job and workload management

- Monitoring and accounting

- Authentication, Authorization and Identity management

(x509 certificates still supported, but probably not for a long time)

IGWN data distribution architecture for offline processing
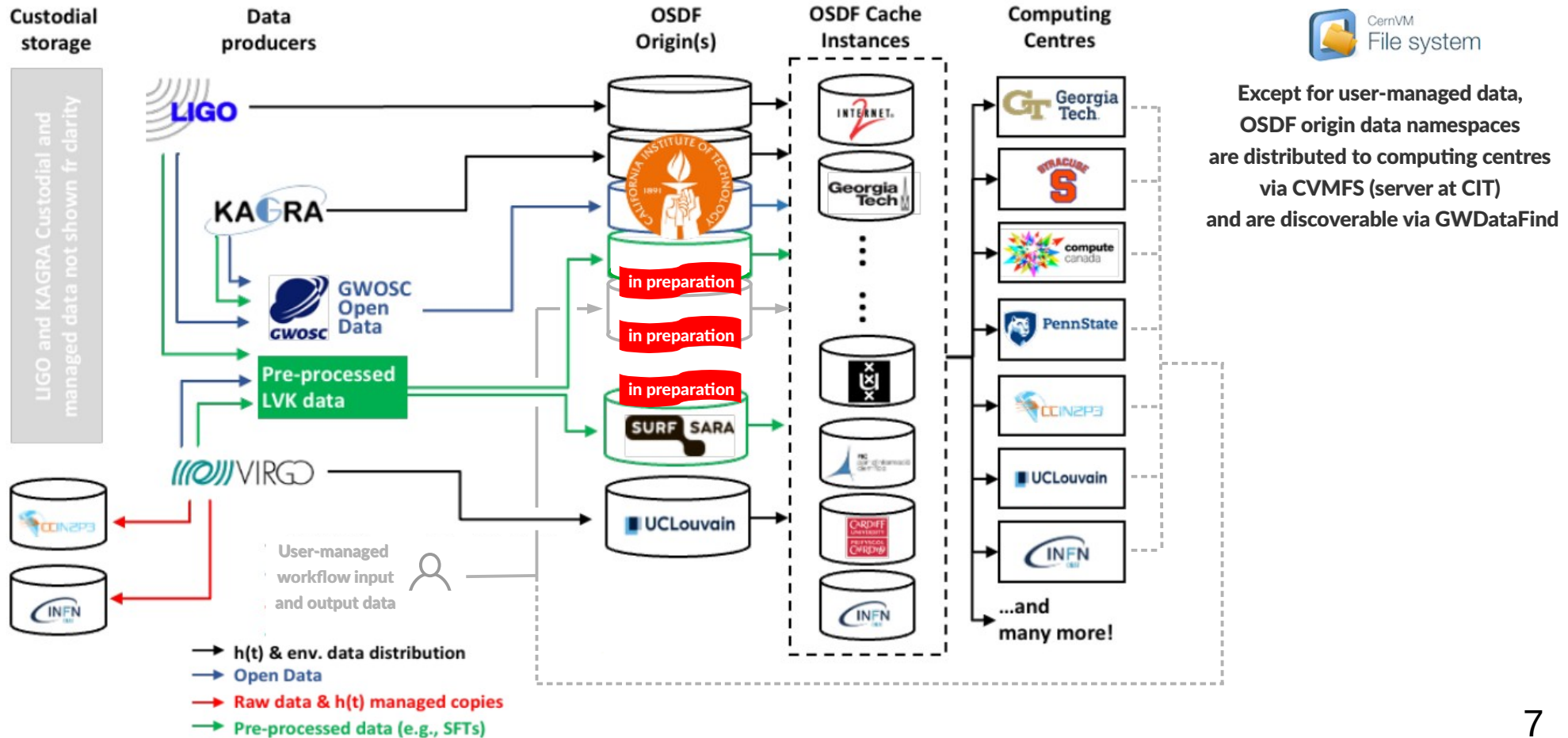
Open Science Data Federation

Open Science Data Federation (OSDF) managed by Open Science Grid (OSG)

Service designed to support the sharing of files staged in autonomous "origins", for efficient access to those files from anywhere in the world via a global namespace and network of caches.

US Map featuring the locations of current OSDF architectural components.

OSDF resources in EU, used by LVK

· From institutions affiliated to Virgo

  · Origin   **UCLouvain**   **SURF SARA**   *in preparation*

  · Caches   **INFN CNAF**   **PIC** port d'informació científica

  [logo]   **CCIN2P3**   *in preparation*

· From institutions affiliated to LIGO or other

  · Caches   **CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD**   **ESnet** London   **ESnet** Amsterdam

LVK is the main user of OSDF, but not the only one:

ET EINSTEIN TELESCOPE   **IceCube** SOUTH POLE NEUTRINO OBSERVATORY   **XENON**   **Fermilab**   **GlueX**

OSDF origin at UCLouvain

- 1 storage machine with 100TB disk space, 16 CPUs, 48GB RAM, 10Gbps connection
- Dedicated to host Virgo aggregated h(t) and detector environmental data
  - › Also used to distribute ET input data for current mock-data challenge
  - › Current used space: Virgo: 8.1TB, ET: 1.8TB
- OSDF origins are essentially XRootD servers with special configuration and additional supporting services
  - › At UCLouvain, we run 2 such instances on the same server, 1 for Virgo and 1 for ET, serving separate namespaces (filesystem directories) and listening on different ports
- OSDF origins (and caches) can be private or public
  - › IGWN origins are private, ET origin is public
- UCLouvain origin server is independent from the cluster storage machines
  - › 2.75PB storage mounted (readonly) on the cluster
  - › Used to host CMS and Na62 data

9

Offline CPU resource usage timeline

- Overall CPU resource usage timeline, by observation period tag
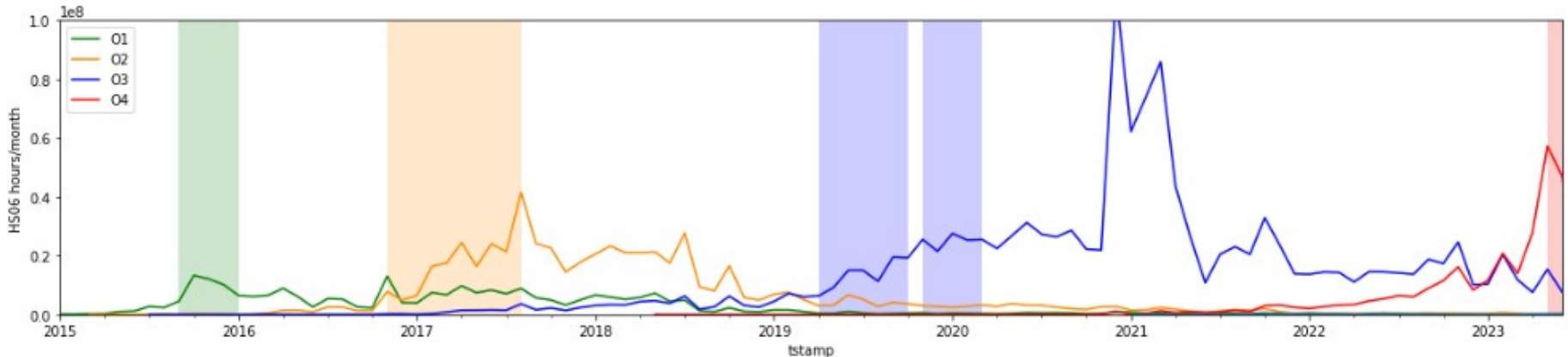
- The total amount of CPU used for O3 data processing was more than 7.5×10^9 HS06 hours

    Mhh, this number doesn't seem to agree with the blue graph integral...

- Demand expected to grow by at least 50% for O4

10

Relative contribution from IGWN CCs to O3 data processing

- More than 50 computing centres contributed to the processing of O3 data
    - › Graph includes local + grid batch processing
    - › Accounting of local batch processing is currently done only on HTCondor clusters
    - › UCLouvain not seen in the graph, because it uses Slurm, plus it joined the IGWN grid at the end of O3
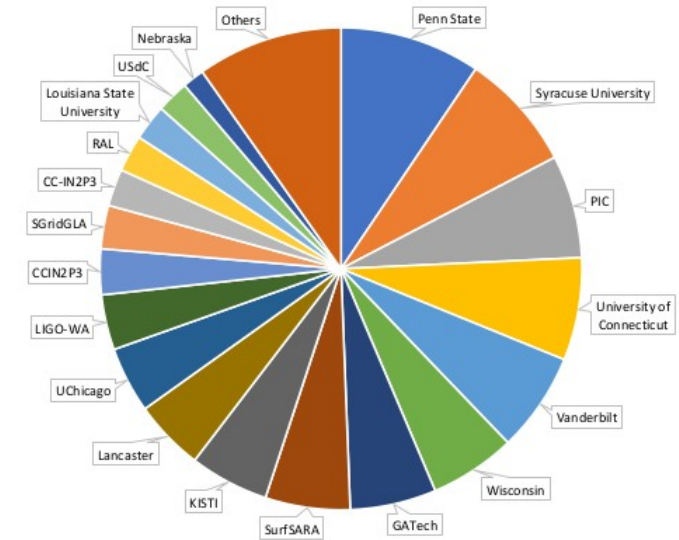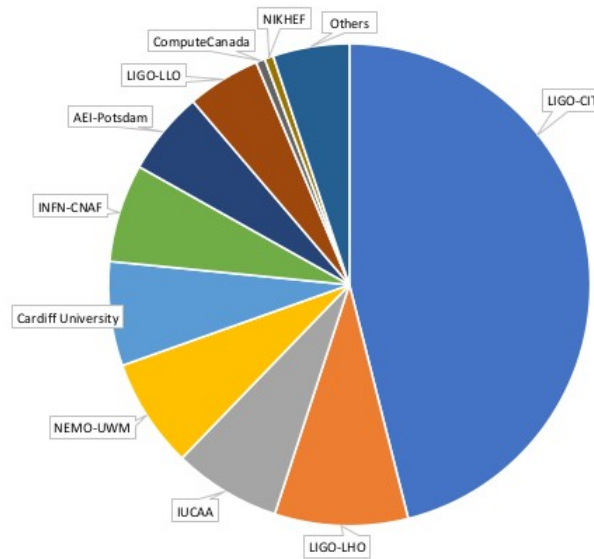
# Contribution from IGWN computing centres to O3 data processing

- During O1 and O2, contribution from Virgo-associated computing centres was much much smaller than that from LIGO-associated computing centres

- IGWN helped to reduce the share imbalance during O3, but Virgo is still in debt
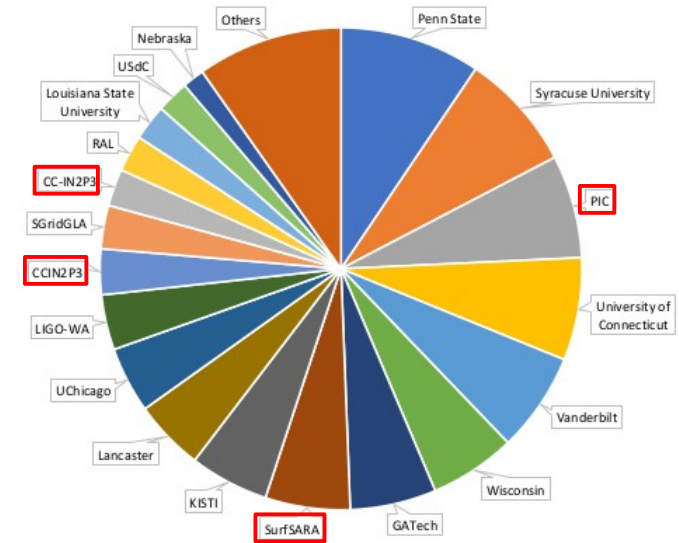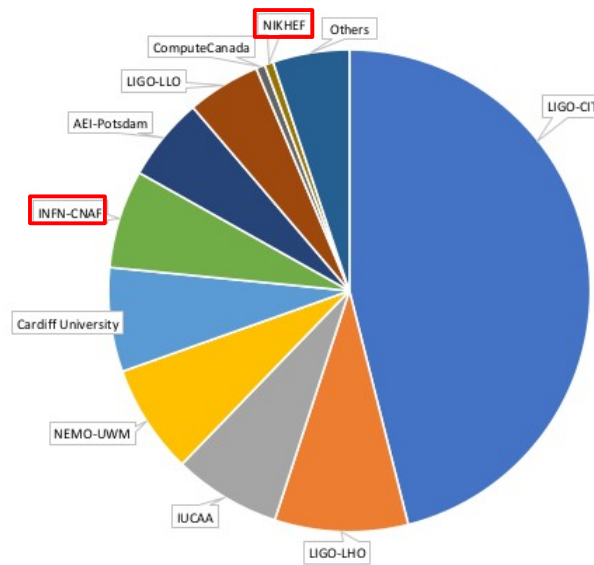
Contribution from IGWN computing centres to O3 data processing

- During O1 and O2, contribution from Virgo-associated computing centres was much much smaller than that from LIGO-associated computing centres

- IGWN helped to reduce the share imbalance during O3, but Virgo is still in debt
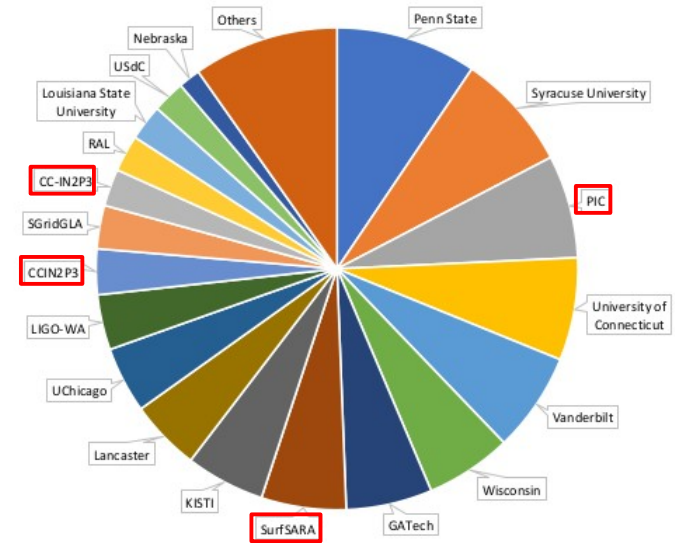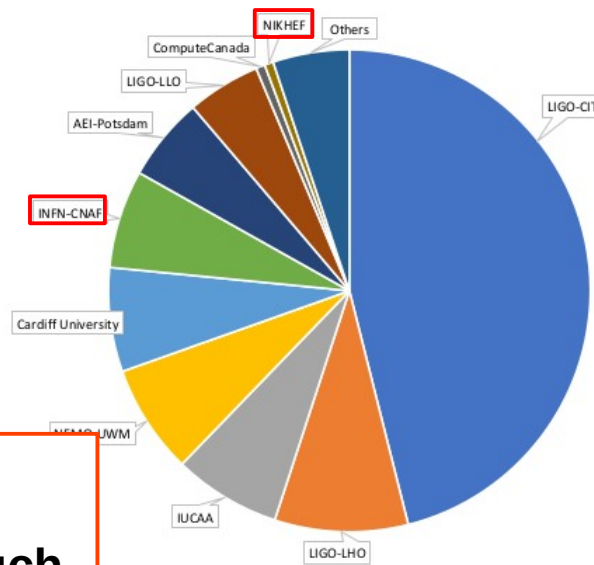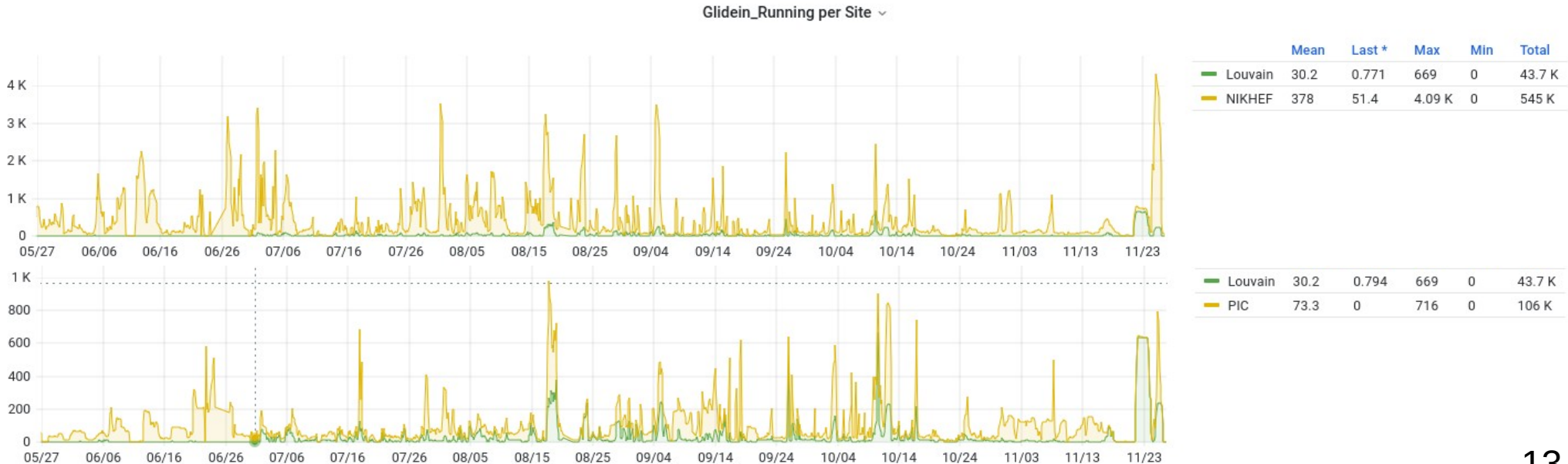
New computing resources for Virgo would be very much welcomed

UCLouvain's recent contribution to data processing

- UCLouvain contribution in terms of running payloads in the last 6 months
  - › 1/12 that from Nikhef, which seems to be not good
  - › 2/5 that from PIC

    PIC is a big computing centre in Spain, so UCLouvain is doing a relatively good contribution

Glidein_Running per Site

	Mean	Last *	Max	Min	Total
Louvain	30.2	0.771	669	0	43.7 K
NIKHEF	378	51.4	4.09 K	0	545 K

	Mean	Last *	Max	Min	Total
Louvain	30.2	0.794	669	0	43.7 K
PIC	73.3	0	716	0	106 K

13

Computing cluster at UCLouvain

- High Performance Computing cluster (named ingrid) with Slurm batch system

  › Logically divided in two parts: ~2150 Vcores for local jobs

    ~2000 Vcores for grid jobs

  › 4GB RAM / Vcore

  › Ingrid is part of the Manneback cluster at UCLouvain

    Additional ~1000 Vcores for local jobs

  › 1 Tesla K80 GPU card for local use

- No shared filesystem

- IGWN software and data available

  via CVMFS

- Other (scientific) softwares managed

  with easybuild and modules

Manneback cluster at UCLouvain

Computing cluster at UCLouvain

- High Performance Computing cluster (named ingrid) with Slurm batch system

  › Logically divided in two parts: ~2150 Vcores for local jobs

    ~2000 Vcores for grid jobs

  shared by:

  all local research groups (CMS, Na62, **Virgo**, ET, IceCube, KM3Net, pheno/theory)

  our supported grid VOs (CMS, **Virgo**, Na62)

  › 4GB RAM / Vcore

  › Ingrid is part of the Manneback cluster at UCLouvain

    Additional ~1000 Vcores for local jobs

  › 1 Tesla K80 GPU card for local use

- No shared filesystem

- IGWN software and data available via CVMFS

- Other (scientific) softwares managed with easybuild and modules

Manneback cluster at UCLouvain

Grid job submission overview

- The Glidein Factory submits a "pilot" job request to the cluster's grid computing entry-point (CE)

  The CE submits the pilot job to the worker nodes using the cluster's native batch system

    › Pilot jobs submitted to UCLouvain request 2 Vcores and 8GB RAM

- The pilot job starts a condor process that communicates directly with a HTCondor pool central manager advertising its resources

- HTCondor matches and submits user jobs (payloads) to these condor resources
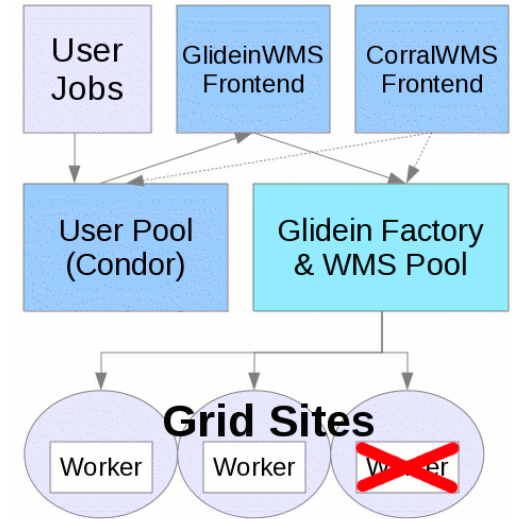
IGWN grid computing doesn't require HTCondor as native batch system

CÉCI

CÉCI

Consortium des Équipements de Calcul Intensif
6 clusters, 10k cores, 1 login, 1 home directory

- Access to Manneback requires a CÉCI account
    › Manneback is not part of CÉCI, but it is administered by the UCLouvain CISM team, which is also in charged of CÉCI
- Virgo users from a CÉCI university could also use the CÉCI clusters
- But... there is no CVMFS in the CÉCI clusters, so neither IGWN software nor LVK data are available
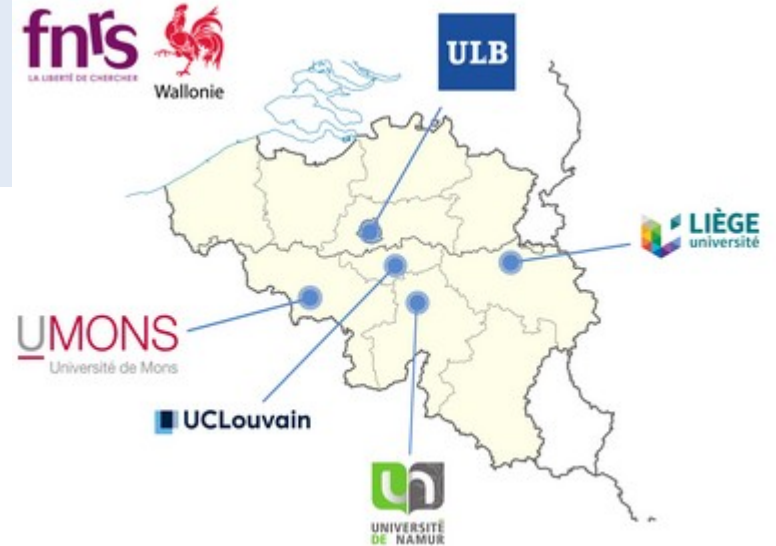
Thank you!

Backup slides

GPUs availability at IGWN

- ~400 GPUs available at LIGO CIT

- GPUs from Virgo associated institutions
  - › CNAF will have many GPUs in 1 year from now
  - › PIC has some GPUs
  - › 16 A10 GPUs available at SURFsara

Virtual Organizations in OSG/OSDF

Each supported experiment/collaboration is called a "virtual organization" (VO) in OSG

    Nothing to do with virtual organizations in grid computing

    LIGO is one of these VOs, Virgo is not (Virgo is "hiding" under the LIGO umbrella)

OSG accepted to add ET as a supported VO in OSDF

    "The entire global collaboration of GW observatories analyzes their data together"

Virtual organizations and OSDF resources are registered in OSG topology (yaml files in github)

Each virtual organization exports unique namespaces in the global OSDF namespace

    Given the namespace, the VO can be unequivocally determined

    Namespaces can be private or public

    LIGO VO: `/igwn/[ligo,virgo,kagra,shared]` (private), `/gwdata` (public)

Each VO defines a list of allowed "origins" and caches for each of its namespaces

OSDF "origins"

OSDF "origins" can be configured to make data private (authenticated "origins") or public (unauthenticated "origins"), and can control the rules of sharing

- Authenticated or unauthenticated: must be consistent with exported namespace being private or public
- Authenticated (unauthenticated) "origins" give read access to VO allowed caches (everyone)
  - "Origin" admins can extend these default sharing rules, in particular for write access

OSDF "origins" are essentially XRootD servers with special configuration and additional supporting services

- Authorized clients can interact with an "origin" server using XRootD client commands
  - List directory contents: `xrdfs xroot://<origin-fqdn>:<port>/ ls <namespace>/<subdir>`
  - Download file: `xrdcp xroot://<origin-fqdn>:<port>/<namespace>/<subdir>/<file> .`
  - Download directory: `xrdcp -r xroot://<origin-fqdn>:<port>/<namespace>/<subdir> .`

OSG provides instructions for installing XRootD client from their software repositories

- Should already be installed in most computing centres (for sure in WLCG and IGWN)

Virgo & ET OSDF "origins"

Virgo "origin" was installed in a dedicated 100 TB storage server at UCLouvain

  Private "origin" (https) listening on port 1095 (IPv4 and IPv6)

   Uses Grid Security Infrastructure (GSI) protocol for authentication

   Only Virgo/LIGO OSDF caches and Virgo data transfer operators are authorized to read the "origin"

   Only Virgo data transfer operators are authorized to write into the "origin"

  FQDN: `ingrid-se09.cism.ucl.ac.be`

ET "origin" service running on same Virgo "origin" storage server at UCLouvain

  Runs in a docker container (provided by OSG)

   Isolated from Virgo "origin" service - there is always the risk of a human mistake (from myself) though

  Public "origin" (http) listening on port 1094 (IPv4 only)

   No authentication: everyone has read access to the "origin", from any host (except from writing host)

   Only I am authorized to write into the "origin" (and from a certain host from where authentication is enforced)

  FQDN: `et-origin.cism.ucl.ac.be` (DNS alias for `ingrid-se09.cism.ucl.ac.be`)

CVMFS

CernVM File System (CVMFS) is an HTTP-based file distribution service used to provide software and (small) data files in a fast, scalable, and reliable way

- Implemented as a POSIX read-only file system
- Files and directories are hosted on standard web servers (known as CVMFS Stratum-0) in so-called CVMFS repositories
- On client side, CVMFS repositories are mounted in the universal namespace `/cvmfs`
- Files are downloaded on demand and aggressively cached (squid proxies)
- Developed to deploy software for HEP collaborations on the Worldwide LHC Computing Grid (WLCG)
- Enhanced by OSG to distribute data files stored in OSDF
  - Supports authentication for private data (X509 proxies, tokens)

OSG provides instructions for installing CVMFS client from their software repositories

- Should already be installed in most computing centres (for sure in WLCG and IGWN)

CVMFS in GW community

CVMFS is used by LVK to distribute software and data

Managed by OSG, CVMFS server in the US

OSDF namespace `/igwn/ligo` is published in `/cvmfs/ligo.storage.igwn.org` (private)

OSDF namespace `/igwn/virgo` is published in `/cvmfs/virgo.storage.igwn.org` (private)

OSDF namespace `/igwn/kagra` is published in `/cvmfs/kagra.storage.igwn.org` (private)

OSDF namespace `/igwn/shared` is published in `/cvmfs/shared.storage.igwn.org` (private)

GW community is the main user of authenticated data in OSDF

OSDF namespace `/gwdata` is published in `/cvmfs/gwosc.osgstorage.org` (public)

Data from Gravitational Wave Open Science Center (GWOSC)

IGWN software is published in `/cvmfs/software.igwn.org`

OSG created a (public) CVMFS repository for ET data

OSDF namespace `/et-gw/PUBLIC` is published in `/cvmfs/et-gw.osgstorage.org` (public)