

# **HTCondor Workshop Autumn 2023**

Monday 18 September 2023 - Friday 22 September 2023

IJCLab

## **Book of Abstracts**



# Contents

News from the DESY Clusters . . . . .	1
The Virgo Data Quality Reports: an HTCondor automated framework to vet gravitational-wave candidates . . . . .	1
The journey to modern storage. . . . .	2
Just-in-time matching of workflows for the DUNE experiment . . . . .	2
Reaching new scales in the CMS Global pool . . . . .	3
Multi-tenancy HTCondor for the European Weather Cloud . . . . .	4
Energy savings by power modulation of a HTC pool . . . . .	4
CERN Update . . . . .	5
What's New, What's Coming Up . . . . .	5
HTCSS Vocabulary, Architecture, and User View . . . . .	5
IGWN Data Story . . . . .	5
Getting Your Data with HTCondor . . . . .	6
Pelican: Data Federations and HTCSS . . . . .	6
Whirlwind Tour: Job Placement Options . . . . .	6
Linux tuning for large scale APs . . . . .	6
HTCondor Services for Non-Humans . . . . .	6
What Are My Jobs Doing? . . . . .	7
Microarchitecture . . . . .	7
Advanced debugging with eBPF and Linux perf tools . . . . .	7
What are glideins using my cluster for? Monitoring via the CE . . . . .	7
Collector Overviews . . . . .	8
Energy Savings: Roosters and Condors, oh my! . . . . .	8
Token Taxonomy . . . . .	8

(Sci)Tokens authentication in HTCondor-CE (incl new features) . . . . .	8
Welcome . . . . .	8
Welcome and housekeeping . . . . .	9
Wrap Up and Goodbye . . . . .	9
Lightning Talks . . . . .	9
Show your Toolbox . . . . .	9
Office Hours . . . . .	10
Deployment of HTCondor at GRIF . . . . .	10
Talk of all the Tools . . . . .	10
Data and Security Wrap Up . . . . .	10
Pools and Execution Points Wrap Up . . . . .	11
Jobs and Access Points Wrap Up . . . . .	11
Welcome drink . . . . .	11
Social Diner . . . . .	11
Thoughts on EDIA in dHTC Cyber-Infrastructure . . . . .	11
Data and Security intro . . . . .	12
Jobs and Access Points Introduction . . . . .	12
Pools and Execution Points Intro . . . . .	12
IJCLab in a nutshell . . . . .	12
Visit of ThomX and SUPRATECH facilities . . . . .	12
Visit of ThomX and SUPRATECH facilities . . . . .	12
Discussion on EDIA . . . . .	13

**HTCondor User Presentations / 1****News from the DESY Clusters****Author:** Thomas Hartmann<sup>1</sup>**Co-author:** Christoph Beyer<sup>1</sup> *Deutsches Elektronen-Synchrotron (DE)***Corresponding Authors:** christoph.beyer@desy.de, thomas.hartmann@cern.ch

With the ever changing IT and political landscapes also the HTC clusters at DESY are evolving with these changes. We will give an overview and status, what over the past year is new and where we plan to go with our compute clusters. Efficient utilization of resources has become even more pressing following recent geopolitical turmoil and with climate change becoming even more pressing, for which we are optimizing the energy profile of our user Condor cluster. Also a Cobald/Tardis Condor meta cluster is under development with the idea to use it for backfilling of untapped idle resources.

**Desired slot length:**

15

**Speaker release:**

Yes

**HTCondor User Presentations / 2****The Virgo Data Quality Reports: an HTCondor automated framework to vet gravitational-wave candidates****Author:** Nicolas Arnaud<sup>1</sup><sup>1</sup> *IJCLab (Université Paris-Saclay and CNRS/IN2P3)***Corresponding Author:** nicolas.arnaud@ijclab.in2p3.fr

Transient gravitational-wave (GW) signals have been discovered since 2015 by the LVK global network of giant, ground-based, interferometric detectors. It currently includes four instruments: the LIGO Hanford and LIGO Livingston detectors located in the USA, the Virgo detector in Italy –hosted by the European Gravitational Observatory (EGO) –, and the KAGRA detector in Japan. A key component of the search for GWs is the broadcast of a low-latency public alert to the astronomical community, each time a significant GW candidate is identified by the pipelines which jointly analyze in real time the data from all the running detectors.

These alerts require a quick, yet accurate, vetting of the quality of the corresponding data. The main input for such a decision –to confirm the alert or to retract it if the candidate is found not to be of astronomical origin –comes from the Data Quality Report (DQR), a set of predefined checks which is triggered automatically when a new candidate has been identified, and which runs on an HTCondor farm. In this contribution, we focus on the Virgo DQR framework which has been developed jointly by the Virgo Collaboration and the EGO IT department, following standards defined at the LIGO-Virgo Collaborations level in 2018-2019. It allows vetting the data acquired by the Virgo detector. After a short description of the EGO HTCondor farm and of the various software which run on it during a joint data-taking period of the LVK network, we will describe the Virgo DQR; the way a significant GW candidate leads to the generation of a global HTCondor DAG (running about 40 different checks in parallel, for a total of roughly 120 jobs); its main inputs and outputs; its performance; and finally the live monitoring system which has been developed to parse every minute

the dag.dagman.out DAG log file. Similar but independent DQR frameworks are running at various LIGO computing centers to vet the LIGO Hanford and LIGO Livingston data.

**Desired slot length:**

12+3 minutes should be enough, 20 minutes if there is space in the agenda

**Speaker release:**

Yes

**HTCondor User Presentations / 3**

## The journey to modern storage.

**Author:** David Handelman<sup>None</sup>

**Corresponding Author:** dudu@final.co.il

About a year ago we had a project that needs 20PB of low performance storage 20GBs throughout. We ended up buying 40PB with 2.7TBs throughout. I would like to share with you the journey from open source storage to commercial storage.

Choosing the technology, creating network to support this throughout, benchmarking, understanding the needs and at the end real world vs the benchmark.

I hope 20 minutes will be enough

**Desired slot length:**

30

**Speaker release:**

Yes

**HTCondor User Presentations / 4**

## Just-in-time matching of workflows for the DUNE experiment

**Author:** Andrew McNab<sup>1</sup>

**Co-authors:** Chris Brew<sup>2</sup>; Raja Nandakumar<sup>2</sup>

<sup>1</sup> *University of Manchester*

<sup>2</sup> *Science and Technology Facilities Council STFC (GB)*

**Corresponding Author:** andrew.mcnab@cern.ch

The DUNE experiment is a large international particle physics project which is currently under construction at Fermilab in Illinois and SURF in South Dakota, with prototypes at CERN. The experiment relies on Fermilab's investment in HTCondor and GlideInWMS, and on the LArSoft ecosystem of applications software. Initially data management was done with Fermilab's SAM system but this is gradually being replaced by other components.

MetaCat and Rucio are now in use as DUNE's file metadata and replica catalogues, and DUNE has developed a just-in-time workflow management system, justIN, to replace the SAM workflow functionality and provide higher level management of processing requests which are carried out in GlideInWMS/HTCondor jobs. The new system's philosophy of matching tasks to resources as they become available will be described. justIN provides a workflow submission interface and then submits suitable jobs to the DUNE HTCondor pool. Jobs call back to justIN when they eventually start at sites, and a decision is made at that point about what workflows to carry out on that machine and which files to process. These decisions are based on the available memory, processors, maximum local job duration, and the availability of nearby files which are still to be processed as part of the current workflows. This just-in-time approach is able to take unplanned downtimes at sites and storages into account immediately, as well as higher level changes such as fluctuations in the demand from other user communities. This system was validated during the DUNE Data Challenge 4 in late 2022 and has been used in the simulation campaigns of 2023. justIN uses token information obtained from CILogon with users authenticating with the Fermilab Identity Provider service. This in turn allows users to authenticate to the justIN web dashboard or to use the justIN command line tool to launch and manage workflows. To enforce DUNE policies on the use of Rucio-managed storage, justIN jobs carry out data write operations on behalf of user supplied scripts and code, which are isolated from higher level credentials by justIN's use of Singularity/Apptainer containers. Further work to increase the integration of justIN and the new dedicated DUNE HTCondor pool will be described.

**Desired slot length:**

20

**Speaker release:**

Yes

**HTCondor User Presentations / 5**

## Reaching new scales in the CMS Global pool

**Authors:** Antonio Perez-Calero Yzquierdo<sup>1</sup>; Edita Kizinevic<sup>2</sup>; Farrukh Aftab Khan<sup>3</sup>; Hyunwoo Kim<sup>3</sup>; Marco Mascheroni<sup>4</sup>; Nikos Tsipinakis<sup>2</sup>

<sup>1</sup> *Centro de Investigaciones Energéticas Medioambientales y Tecnológicas*

<sup>2</sup> *CERN*

<sup>3</sup> *Fermi National Accelerator Lab. (US)*

<sup>4</sup> *Univ. of California San Diego (US)*

**Corresponding Author:** marco.mascheroni@cern.ch

The computing resource needs of LHC experiments, such as CMS, are expected to continue growing significantly over the next decade, during the Run 3 and especially the HL-LHC era. The SI team manages a set of federated HTCondor pools, currently aggregating around 400k CPU cores distributed worldwide, supporting the simultaneous execution of over 200k CMS computing tasks. In order to detect and overcome performance degradation driven by scalability barriers, the SI team

regularly runs tests to explore the scalability reach of our infrastructure. In this contribution, we will report on the test results for potential scalability limitations of our infrastructure.

**Desired slot length:**

**Speaker release:**

Yes

**HTCondor User Presentations / 6**

## Multi-tenancy HTCondor for the European Weather Cloud

**Authors:** Francesco Murdaca<sup>1</sup>; Michael Grant<sup>1</sup>

<sup>1</sup> *EUMETSAT*

**Corresponding Author:** francesco.murdaca@eumetsat.int

The European Weather Cloud (EWC) is the cloud-based collaboration platform for meteorological application development and operations in Europe and enables the digital transformation of the European Meteorological Infrastructure. Among the services to be provided to the meteorological community, batch processing has been requested for several type of analysis using satellite data. HTCondor has been selected as the batch system for data processing and it will be provided to the users through the EWC. In this talk you will discover what is EWC and how HTCondor was customized to work in the EWC in order to facilitate utilization/share of resources across multi-tenancies and to run jobs through HTCondor. In particular, you will see how we allow users to easily join the common pool of resources that can be used by the community, how nodes are automatically deployed, how we preserve the ability to access internal networks of different tenancies via VPN, how users can rely on multi tenancies resources, running jobs securely and with close to the data.

**Desired slot length:**

**Speaker release:**

Yes

**HTCondor User Presentations / 7**

## Energy savings by power modulation of a HTC pool

**Authors:** Christoph Beyer<sup>None</sup>; Thomas Hartmann<sup>1</sup>

<sup>1</sup> *Deutsches Elektronen-Synchrotron (DE)*

**Corresponding Author:** christoph.beyer@desy.de

Energy prices are reaching an all time high in europe and we are speculating at the same time to get a more elaborated billing model that calculates a real time electricity price depending on the availability of electricity on the market. As the price is predictable roughly 30h in forehand it seems to be desirable to not only know more about the energy consumption of a condor pool but also to drive the energy consumption of the pool along a timeline that is given by said prices ...

**Desired slot length:**



**Speaker release:**

Yes

**HTCondor User Presentations / 8**

## **CERN Update**

**Corresponding Author:** ben.dylan.jones@cern.ch

**Speaker release:**

**Desired slot length:**

**Workshop Session / 10**

## **What's New, What's Coming Up**

**Corresponding Author:** tannenba@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Workshop Session / 11**

## **HTCSS Vocabulary, Architecture, and User View**

**Corresponding Author:** tannenba@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Data and Security / 12**

## **IGWN Data Story**

**Corresponding Authors:** pcouvare@caltech.edu, peter.couvares@ligo.org

**Speaker release:**

**Desired slot length:**

**Data and Security / 13**

## **Getting Your Data with HTCondor**

**Corresponding Author:** brian.bockelman@cern.ch

**Speaker release:**

**Desired slot length:**

**Data and Security / 14**

## **Pelican: Data Federations and HTCSS**

**Corresponding Author:** brian.bockelman@cern.ch

**Speaker release:**

**Desired slot length:**

**Jobs and Access Points / 15**

## **Whirlwind Tour: Job Placement Options**

**Corresponding Author:** gthain@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Jobs and Access Points / 16**

## **Linux tuning for large scale APs**

**Corresponding Author:** gthain@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Jobs and Access Points / 17**

## **HTCondor Services for Non-Humans**

**Corresponding Author:** jfrey@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Jobs and Access Points / 18**

## **What Are My Jobs Doing?**

**Corresponding Author:** jfrey@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Pools and Execution Points / 19**

## **Microarchitecture**

**Corresponding Author:** gthain@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Pools and Execution Points / 20**

## **Advanced debugging with eBPF and Linux perf tools**

**Corresponding Author:** gthain@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Pools and Execution Points / 21**

## **What are glideins using my cluster for? Monitoring via the CE**

**Corresponding Author:** tannenba@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Pools and Execution Points / 22**

## **Collector Overviews**

**Corresponding Author:** jfrey@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Workshop Session / 23**

## **Energy Savings: Roosters and Condors, oh my!**

**Corresponding Author:** tannenba@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Data and Security / 24**

## **Token Taxonomy**

**Corresponding Author:** brian.bockelman@cern.ch

mixing them up... IDTOKENS for daemons, others for jobs, issued from OAuth vs Vault vs self signed  
vs ...

**Speaker release:**

**Desired slot length:**

**Data and Security / 25**

## **(Sci)Tokens authentication in HTCondor-CE (incl new features)**

**Corresponding Author:** jfrey@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

## Welcome

**Author:** Christoph Beyer<sup>None</sup>

**Co-author:** Chris Brew<sup>1</sup>

<sup>1</sup> *Science and Technology Facilities Council STFC (GB)*

**Corresponding Author:** christoph.beyer@desy.de

**Speaker release:**

**Desired slot length:**

27

## Welcome and housekeeping

**Corresponding Author:** jouvin@lal.in2p3.fr

**Speaker release:**

**Desired slot length:**

28

## Wrap Up and Goodbye

**Author:** Chris Brew<sup>1</sup>

**Co-author:** Christoph Beyer

<sup>1</sup> *Science and Technology Facilities Council STFC (GB)*

**Corresponding Author:** chris.brew@stfc.ac.uk

**Speaker release:**

**Desired slot length:**

**Lightning Talks / 29**

## Lightning Talks

**Speaker release:**

**Desired slot length:**

**Show your Toolbox / 30**

## Show your Toolbox

**Speaker release:**

**Desired slot length:**

31

## Office Hours

**HTCondor User Presentations / 32**

## Deployment of HTCondor at GRIF

**Author:** Emmanouil Vamvakopoulos<sup>1</sup>

<sup>1</sup> *Université Paris-Saclay (FR)*

**Corresponding Author:** emmanouil.vamvakopoulos@ijclab.in2p3.fr

In this communication, we are going to present the multiple HTCondor instances that have been deployed at GRIF for several years. GRIF is a distributed Tier-2 WLCG/EGI site made of four (4) different subsites (IJCLab, IRFU, LLR, LPNHE), in different locations of the Paris region. The worst network latency between the subsites is within 2-4 msec with three (3) of them connected with a 100Gbit/sec connection. In particular, a distributed HTCondor pool, with HTCondor-CE gateways, gives unified access to the IJCLab and LLR resources. IRFU and LPHNE are running independent condor pools based on an ARC-CE and HTCondor-CE gateways which are providing access to the computing resources of those sites. Future intentions and plans about the incorporation and usage of Cloud computing resources with or without Kurbenetes infrastructure will be discuss.

**Speaker release:**

Yes

**Desired slot length:**

30

**Jobs and Access Points / 33**

## Talk of all the Tools

**Corresponding Author:** gthain@cs.wisc.edu

**Speaker release:**

**Desired slot length:**

**Data and Security / 34**

## **Data and Security Wrap Up**

**Corresponding Author:** miron@cs.wisc.edu

**Pools and Execution Points / 35**

## **Pools and Execution Points Wrap Up**

**Corresponding Author:** miron@cs.wisc.edu

**Jobs and Access Points / 36**

## **Jobs and Access Points Wrap Up**

**Corresponding Author:** miron@cs.wisc.edu

37

## **Welcome drink**

La Part des Anges is an excellent and friendly Wine Bar.

Tapas/apetizers offered, drinks at your charge

Possibility to have a diner afterward on an individual basis

[https://osm.org/go/0BOQg93\\_M?node=6574870872](https://osm.org/go/0BOQg93_M?node=6574870872)

38

## **Social Diner**

**Equality, Diversity, Inclusion & Accessibility / 39**

## **Thoughts on EDIA in dHTC Cyber-Infrastructure**

**Author:** Matthew West<sup>None</sup>

**Corresponding Author:** m.t.west@exeter.ac.uk

Making space for new groups within our existing dHTC infrastructure community is essential for “democratizing computation.” This talk calls on participants to reflect on why so many research computing conferences have such homogeneous attendees. Some are long-term structural issues beyond our scope, but this should not absolve us from taking proactive measures to rectify this problem where possible.

**Desired slot length:**

5

**Speaker release:**

Yes

**Data and Security / 40**

## **Data and Security intro**

**Corresponding Author:** miron@cs.wisc.edu

**Jobs and Access Points / 41**

## **Jobs and Access Points Introduction**

**Corresponding Author:** miron@cs.wisc.edu

**Pools and Execution Points / 42**

## **Pools and Execution Points Intro**

**Corresponding Author:** miron@cs.wisc.edu

43

## **IJCLab in a nutshell**

**Corresponding Authors:** chambert@ipno.in2p3.fr, valerie.chambert@cern.ch

**Speaker release:**

**Desired slot length:**

44

## **Visit of ThomX and SUPRATECH facilities**

By groups of 15



45

## **Visit of ThomX and SUPRATECH facilities**

2 groups of 15

**Equality, Diversity, Inclusion & Accessibility / 46**

## **Discussion on EDIA**