HTCondor Workshop Autumn 2023

Report of Contributions

Contribution ID: **1**   Type: **not specified**

News from the DESY Clusters

Wednesday 20 September 2023 14:50 (20 minutes)

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

Author: HARTMANN, Thomas (Deutsches Elektronen-Synchrotron (DE))

Co-author: BEYER, Christoph

Presenter: HARTMANN, Thomas (Deutsches Elektronen-Synchrotron (DE))

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **2**   Type: **not specified**

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

Thursday 21 September 2023 11:50 (20 minutes)

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

Author: Dr ARNAUD, Nicolas (IJCLab (Université Paris-Saclay and CNRS/IN2P3))

Presenter: Dr ARNAUD, Nicolas (IJCLab (Université Paris-Saclay and CNRS/IN2P3))

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **3**

Type: **not specified**

The journey to modern storage.

Thursday 21 September 2023 14:35 (20 minutes)

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

Choosing the technology, creating network to support this throughout, benchamrking, understand 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

Author: HANDELMAN, David

Presenter: HANDELMAN, David

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **4**   Type: **not specified**

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

Thursday 21 September 2023 14:10 (20 minutes)

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

Author: MCNAB, Andrew (University of Manchester)

Co-authors: BREW, Chris (Science and Technology Facilities Council STFC (GB)); NANDAKUMAR, Raja (Science and Technology Facilities Council STFC (GB))

Presenter: MCNAB, Andrew (University of Manchester)

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **5**

Type: **not specified**

Reaching new scales in the CMS Global pool

Thursday 21 September 2023 11:25 (20 minutes)

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

Authors: PEREZ-CALERO YZQUIERDO, Antonio (Centro de Investigaciones Energéticas Medioambientales y Tecnológicas); KIZINEVIC, Edita (CERN); KHAN, Farrukh Aftab (Fermi National Accelerator Lab. (US)); KIM, Hyunwoo (Fermi National Accelerator Lab. (US)); MASCHERONI, Marco (Univ. of California San Diego (US)); TSIPINAKIS, Nikos (CERN)

Presenter: MASCHERONI, Marco (Univ. of California San Diego (US))

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **6**   Type: **not specified**

Multi-tenancy HTCondor for the European Weather Cloud

Thursday 21 September 2023 13:45 (20 minutes)

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

Authors: MURDACA, Francesco (EUMETSAT); GRANT, Michael (EUMETSAT)

Presenter: MURDACA, Francesco (EUMETSAT)

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **7**   Type: **not specified**

Energy savings by power modulation of a HTC pool

Thursday 21 September 2023 12:10 (20 minutes)

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

Authors: BEYER, Christoph; HARTMANN, Thomas (Deutsches Elektronen-Synchrotron (DE))

Presenter: BEYER, Christoph

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **8** Type: **not specified**

CERN Update

Wednesday 20 September 2023 16:10 (25 minutes)

Speaker release

Desired slot length

Presenter: JONES, Ben (CERN)

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **10**   Type: **not specified**

What's New, What's Coming Up

Thursday 21 September 2023 09:00 (40 minutes)

Speaker release

Desired slot length

Presenter: TANNENBAUM, Todd (University of Wisconsin Madison (US))

Session Classification: Workshop Session

Track Classification: Miscellaneous

Contribution ID: **11**   Type: **not specified**

HTCSS Vocabulary, Architecture, and User View

Tuesday 19 September 2023 09:30 (50 minutes)

Speaker release

Desired slot length

Presenter: TANNENBAUM, Todd (University of Wisconsin Madison (US))

Session Classification: Workshop Session

Track Classification: Miscellaneous

Contribution ID: **12**   Type: **not specified**

IGWN Data Story

Tuesday 19 September 2023 11:35 (30 minutes)

Speaker release

Desired slot length

Presenters: COUVARES, Peter; Mr COUVARES, Peter (California Institute of Technology)

Session Classification: Data and Security

Track Classification: The Data Story

Contribution ID: **13**

Type: **not specified**

Getting Your Data with HTCondor

Tuesday 19 September 2023 14:35 (30 minutes)

Speaker release

Desired slot length

Presenter: BOCKELMAN, Brian Paul (University of Wisconsin Madison (US))

Session Classification: Data and Security

Track Classification: The Data Story

Contribution ID: **14**   Type: **not specified**

Pelican: Data Federations and HTCSS

Tuesday 19 September 2023 11:10 (20 minutes)

Speaker release

Desired slot length

Presenter: BOCKELMAN, Brian Paul (University of Wisconsin Madison (US))

Session Classification: Data and Security

Track Classification: The Data Story

Contribution ID: 15   Type: **not specified**

Whirlwind Tour: Job Placement Options

Wednesday 20 September 2023 12:05 (25 minutes)

Speaker release

Desired slot length

Presenter: THAIN, Gregory

Session Classification: Jobs and Access Points

Track Classification: The Access Point Story

Contribution ID: **16**　　　　　　　　　　　　　　　　　　　　　Type: **not specified**

Linux tuning for large scale APs

Wednesday 20 September 2023 11:00 (30 minutes)

Speaker release

Desired slot length

Presenter:　THAIN, Gregory

Session Classification:　Jobs and Access Points

Track Classification:　The Access Point Story

Contribution ID: **17**

Type: **not specified**

HTCondor Services for Non-Humans

Wednesday 20 September 2023 14:00 (20 minutes)

Speaker release

Desired slot length

Presenter: FREY, James (University of Wisconsin Madison (US))

Session Classification: Jobs and Access Points

Track Classification: The Access Point Story

Contribution ID: **18**   Type: **not specified**

What Are My Jobs Doing?

Wednesday 20 September 2023 11:35 (25 minutes)

Speaker release

Desired slot length

Presenter: FREY, James (University of Wisconsin Madison (US))

Session Classification: Jobs and Access Points

Track Classification: The Access Point Story

Contribution ID: **19**   Type: **not specified**

Microarchitecture

Friday 22 September 2023 09:40 (30 minutes)

Speaker release

Desired slot length

Presenter: THAIN, Gregory

Session Classification: Pools and Execution Points

Track Classification: The Execution Point Story

Contribution ID: **20**
Type: **not specified**

Advanced debugging with eBPF and Linux perf tools

Friday 22 September 2023 11:00 (50 minutes)

Speaker release

Desired slot length

Presenter: THAIN, Gregory

Session Classification: Pools and Execution Points

Track Classification: The Execution Point Story

Contribution ID: **21**   Type: **not specified**

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

Friday 22 September 2023 10:15 (15 minutes)

Speaker release

Desired slot length

Presenter: TANNENBAUM, Todd (University of Wisconsin Madison (US))

Session Classification: Pools and Execution Points

Track Classification: The Pool Story

Contribution ID: 22   Type: **not specified**

Collector Overviews

Friday 22 September 2023 09:05 (30 minutes)

Speaker release

Desired slot length

Presenter: FREY, James (University of Wisconsin Madison (US))

Session Classification: Pools and Execution Points

Track Classification: The Pool Story

Contribution ID: **23**   Type: **not specified**

Energy Savings: Roosters and Condors, oh my!

Thursday 21 September 2023 09:45 (15 minutes)

Speaker release

Desired slot length

Presenter: TANNENBAUM, Todd (University of Wisconsin Madison (US))

Session Classification: Workshop Session

Track Classification: The Pool Story

Contribution ID: 24   Type: **not specified**

Token Taxonomy

Tuesday 19 September 2023 12:10 (20 minutes)

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

Speaker release

Desired slot length

Presenter: BOCKELMAN, Brian Paul (University of Wisconsin Madison (US))

Session Classification: Data and Security

Track Classification: Security and Tokens

Contribution ID: 25   Type: **not specified**

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

Tuesday 19 September 2023 14:00 (30 minutes)

Speaker release

Desired slot length

Presenter: FREY, James (University of Wisconsin Madison (US))

Session Classification: Data and Security

Track Classification: Security and Tokens

Contribution ID: **26**

Type: **not specified**

Welcome

Tuesday 19 September 2023 09:00 (5 minutes)

Speaker release

Desired slot length

Author: BEYER, Christoph

Co-author: BREW, Chris (Science and Technology Facilities Council STFC (GB))

Presenter: BEYER, Christoph

Track Classification: Miscellaneous

Contribution ID: **27**   Type: **not specified**

Welcome and housekeeping

Tuesday 19 September 2023 09:05 (10 minutes)

Speaker release

Desired slot length

Presenter: JOUVIN, Michel (Université Paris-Saclay (FR))

Track Classification: Miscellaneous

Contribution ID: **28**   Type: **not specified**

Wrap Up and Goodbye

Friday 22 September 2023 12:10 (20 minutes)

Speaker release

Desired slot length

Author: BREW, Chris (Science and Technology Facilities Council STFC (GB))

Co-author: BEYER, Christoph

Presenter: BREW, Chris (Science and Technology Facilities Council STFC (GB))

Track Classification: Miscellaneous

Contribution ID: **29**　　　　　　　　　　　　　　　　　　　Type: **not specified**

Lightning Talks

Speaker release

Desired slot length

Session Classification:　Lightning Talks

Track Classification:　Miscellaneous

Contribution ID: **30** Type: **not specified**

Show your Toolbox

Speaker release

Desired slot length

Session Classification: Show your Toolbox

Track Classification: Miscellaneous

Contribution ID: **31**　　　　　　　　　　　　　　　　Type: **not specified**

Office Hours

Wednesday 20 September 2023 16:40 (50 minutes)

Contribution ID: **32**   Type: **not specified**

Deployment of HTCondor at GRIF

Wednesday 20 September 2023 15:15 (25 minutes)

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

Author: Dr VAMVAKOPOULOS, Emmanouil (Université Paris-Saclay (FR))

Presenter: Dr VAMVAKOPOULOS, Emmanouil (Université Paris-Saclay (FR))

Session Classification: HTCondor User Presentations

Track Classification: HTCondor user presentations

Contribution ID: **33**   Type: **not specified**

Talk of all the Tools

Wednesday 20 September 2023 09:10 (45 minutes)

Speaker release

Desired slot length

Presenter: THAIN, Gregory

Session Classification: Jobs and Access Points

Track Classification: HTCondor presentations and tutorials

Contribution ID: **34** Type: **not specified**

Data and Security Wrap Up

Tuesday 19 September 2023 15:10 (15 minutes)

Presenter: LIVNY, MIRON

Session Classification: Data and Security

Contribution ID: **35**

Type: **not specified**

Pools and Execution Points Wrap Up

Friday 22 September 2023 11:50 (20 minutes)

Presenter: LIVNY, MIRON

Session Classification: Pools and Execution Points

Contribution ID: **36**　　　　　　　　　　　　　　　　　　Type: **not specified**

Jobs and Access Points Wrap Up

Wednesday 20 September 2023 14:25 (20 minutes)

Presenter:　LIVNY, MIRON

Session Classification:　Jobs and Access Points

Contribution ID: **37**   Type: **not specified**

Welcome drink

Tuesday 19 September 2023 18:30 (1h 30m)

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

Contribution ID: **38** · Type: **not specified**

Social Diner

Thursday 21 September 2023 19:15 (2 hours)

Contribution ID: **39**   Type: **not specified**

Thoughts on EDIA in dHTC Cyber-Infrastructure

Thursday 21 September 2023 15:00 (10 minutes)

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

Author: WEST, Matthew

Presenter: WEST, Matthew

Session Classification: Equality,Diversity, Inclusion & Accessibility

Track Classification: HTCondor user presentations

Contribution ID: **40**　　　　　　　　　　　　　　　　　　Type: **not specified**

Data and Security intro

Tuesday 19 September 2023 11:00 (10 minutes)

Presenter: LIVNY, MIRON

Session Classification: Data and Security

Contribution ID: 41   Type: **not specified**

Jobs and Access Points Introduction

Wednesday 20 September 2023 09:00 (10 minutes)

Presenter: LIVNY, MIRON

Session Classification: Jobs and Access Points

Contribution ID: **42**   Type: **not specified**

Pools and Execution Points Intro

Friday 22 September 2023 09:00 (5 minutes)

Presenter: LIVNY, MIRON

Session Classification: Pools and Execution Points

Contribution ID: 43 Type: **not specified**

IJCLab in a nutshell

Tuesday 19 September 2023 09:15 (15 minutes)

Speaker release

Desired slot length

Presenters: CHAMBERT, Valerie; CHAMBERT, Valerie (Universite de Paris-Sud 11 (FR))

Contribution ID: 44　　　　　　　　　　　　　　　　　　Type: **not specified**

Visit of ThomX and SUPRATECH facilities

Thursday 21 September 2023 10:05 (30 minutes)

By groups of 15

Contribution ID: **45** Type: **not specified**

Visit of ThomX and SUPRATECH facilities

Wednesday 20 September 2023 10:00 (30 minutes)

2 groups of 15

Contribution ID: 46　　　　　　　　　　　　　　　　　Type: **not specified**

Discussion on EDIA

Thursday 21 September 2023 15:10 (20 minutes)

Session Classification: Equality,Diversity, Inclusion & Accessibility