CMS DATA PROCESSING AT NSC KIPT

O. Kurov, L. Levchuk, S. Lukyanenko, V. Popov, O. Pristavka, O. Bezpalov, L. Kovalyova
National Science Center "Kharkiv Institute of Physics and Technology", Kharkiv, Ukraine
E-mail: march@kipt.kharkov.ua

Present status of the CMS experiment at the Large Hadron Collider (LHC) is outlined. For many years, the NSC KIPT has been carrying out work on processing CMS experimental information. In particular, the institute operates a computing facility, which is a tier-2 (T2) center of the CMS grid infrastructure, T2_UA_KIPT, and the only Ukrainian site processing LHC experimental information practically in a continuous (24/7) mode. In 2025, a wide range of necessary changes were made in the T2_UA_KIPT configuration, and a high level of its reliability was provided being one of the best among CMS T2 sites. Last year, 1.74 PB of CMS experimental information have been transferred to the center for processing. Also in 2025, the full sample of 13 TeV proton-proton collisions recorded in the CMS experiment for the LHC Run 2 was analyzed in order to search for a supersymmetry (SUSY) signal, the direct pair production of the lightest chargions with their subsequent cascade decay to charged leptons (electrons/positrons and/or muons), neutrinos, and the lightest neutralinos. The signal regions were optimized and then used with statistical correctness to exclude a wide range of masses of the SUSY particles searched for.

PACS: 07.05.-t; 13.85.-t

In 2025, Run 3 of the Large Hadron Collider (LHC) was successfully continued with the record proton-proton collision energy for accelerators of 13.6 TeV and a peak luminosity of $\sim 2 \cdot 10^{34}$ $cm^{-2} \cdot s^{-1}$, which is 2 times greater than the LHC design peak luminosity value. The collider resumed operation (after a winter-spring break) in early May, and proton-proton collisions were supported until early November, and then, a run with heavy ion collisions (Pb-Pb) was fulfilled for another month. Runs with oxygen and neon ion beams were also successfully completed in July. In 2025, the collider has provided an integrated luminosity of ~126 fb^{-1} for proton-proton collisions, which is ~5 fb^{-1} greater than in the previous year 2024, thus making 2025 a new record-breaking year. In early December, the LHC was shut down for the scheduled winter break.

In the CMS experiment [1] during the LHC Run 3, the rate of event selection by the high-level trigger (HLT) [2, 3] for offline processing of proton-proton collision data reaches a few kHz, which is up to several times greater than over Run 2 (2015–2018) and by an order of magnitude higher than in Run 1 (2010–2013). As a result, a sample of events (proton-proton collisions at 13.6 TeV) has been recorded in the CMS experiment for 2025, which corresponds to an integrated luminosity of ~115 fb^{-1} thus setting a new record for data collection in a single year. In total, the CMS has already accumulated ~300 fb^{-1} of experimental information on 13.6 TeV proton-proton collisions for the LHC Run 3 (2022–2025). It should be also mentioned that, in addition to the conventional data-taking strategy, the CMS also intensively utilizes the data scouting and data parking techniques [4] to extend the data processing capabilities of the experiment.

One of the priority tasks in the CMS experiment is now continuation of the processing of the valuable experimental material obtained for LHC Runs 2 and 3. In 2025, a number of outstanding results were obtained, and, first of all, these are new data on the properties of the Higgs boson. In particular, based on the analysis of a 200 fb^{-1} sample of proton-proton collisions, information was obtained on spin correlations in Higgs decay into 4 leptons, which makes it possible to estimate possible contributions of anomalous couplings and mechanisms with violation of CP-invariance. In addition, the results were presented [6, 7] on processing of Run 3 data to estimate the Higgs pair production cross section. Events with four reconstructed hadron jets initiated by a b quark, as well as with two such jets accompanied by 2 photons, were selected in order to search for processes with pair production of Higgs bosons decaying as HH→4b and HH→bbγγ. For these processes, upper limits on the signal significance μ (the ratio of the cross-section obtained in the experiment to the one estimated within the Standard Model (SM)) of 4.4 [6] and 11 [7], respectively, were obtained at the confidence level of CL=95.

Also in 2025, the CMS continued an intensive search for signals of "new physics" beyond the SM – additional weak bosons, manifestations of the lepton flavor universality violation (LFUV), supersymmetry (SUSY) and dark matter signals, etc. However, as of the end of 2025, the CMS results turned out to be in a satisfactory agreement with the SM. In 2025, the CMS also obtained a number of important results on the study of the SM phenomena. In particular, using a sample of events recorded from 2016 to 2023 and advanced data analysis techniques based on the machine learning, the CMS experiment first observed the single top quark production in association with a W and a Z boson in proton-proton collisions [8] – an extremely rare process that occurs only once in a trillion proton-proton collisions.

The high intensity of the experimental information flow over the LHC Run 3 tightens requirements set on the data processing infrastructure. The NSC KIPT specialized computing facility is an active element of that distributed system, taking CMS data for processing since the LHC startup in 2009. The facility operates within the Worldwide LHC Grid (WLCG) as a tier-2 (T2) center of CMS grid infrastructure (where it is registered as T2_UA_KIPT), and its resources are used

for physics analyses of CMS data, as well as for the Monte Carlo (MC) generation of CMS events.

In 2025, a wide range of work was carried out to support operation of the T2_UA_KIPT center, which is the only Ukrainian site where experimental information from the LHC is processed in a practically continuous (24/7) mode. It should be noted that 2025 (as well as the previous 3 years) was very hard to provide a due operation of the T2_UA_KIPT site. Numerous emergency situations were caused, in particular, by failures of nodes and systems that had reached their service life, as well as by continuation of the war in Ukraine. In particular, such situations arose as a result of strikes on the energy infrastructure of the Kharkiv region and Ukraine as a whole, which led to emergency power outages. In addition to solving problems related to these emergency situations, a large amount of other work was carried out during 2025 to provide the functionality of the T2_UA_KIPT facility as a CMS tier-2 center. In particular, detected hardware malfunctions of the facility's nodes were promptly corrected, and numerous necessary system updates and changes of its configuration were also performed in a timely manner.

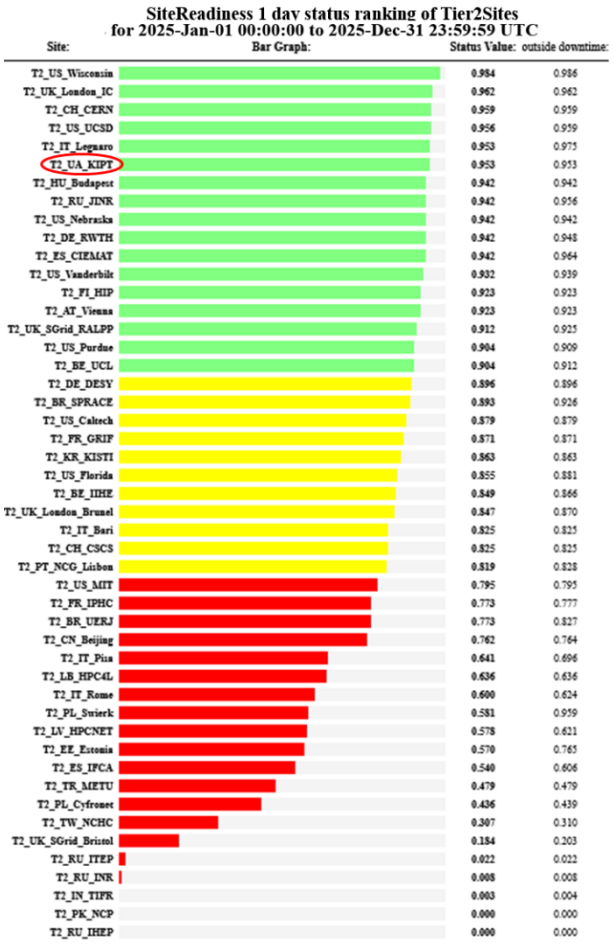

Fig. 1. CMS tier-2 site readiness ranking for the entire year 2025 (see Ref. [9])

As a result of the work completed, the T2_UA_KIPT center readiness for participation in CMS data processing in 2025 (following the metrics adopted in the CMS) was 95.3% (Fig. 1, which displays the corresponding CMS T2 site ranking table given by Ref. [9]), and the center turned out to be one of the six most reliable CMS T2 sites with the readiness indicator exceeding 95%. In addition, a high rate of CMS experimental information transfers to the T2_UA_KIPT center for processing was supported (Fig. 2), which in certain periods of time practically saturated the current facility's external link bandwidth of 10 Gb·ps. The total amount of the transferred information for the year reached 1.74 PB [10], which substantially exceeds the total T2_UA_KIPT mass disk storage capacity of 1.1 PB. At the same time, more than 40 thousand kHEPScore23 hours of T2_UA_KIPT computing resources were utilized for CMS data processing (see Ref. [11]), which was the only Ukrainian contribution to the WLCG computing resources usage for processing of LHC data for the year.

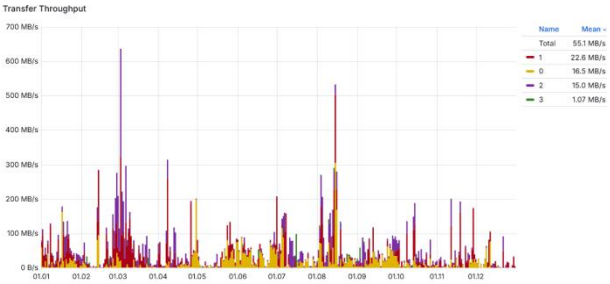

Fig. 2. Daily averaged rate of CMS experimental information transfers to T2_UA_KIPT site over 2025 from tiers 0 (CERN), 1, 2, and 3 of the CMS grid infrastructure (see Ref. [10])

In 2025, we also continued the physics analysis of 13 TeV proton-proton collision samples recorded in the CMS experiment for LHC Run 2 with the aim of searching for a SUSY signal, the direct pair production of the lightest charginos with their subsequent cascade decay into charged leptons (electrons/positrons and/or muons), neutrinos and neutralinos. For this analysis, we exploited the LatinoAnalysis software framework [12]. Compared to our previous consideration, a new optimization was fulfilled for signal regions (SR). For this purpose, the software code developed earlier to obtain optimal SR through maximizing the signal significance (ratio $signal/[signal+background]^{1/2}$) was upgraded. As a result, 6 SR were determined based on optimization over the missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ and the number of hadronic jets (either b-tagged ones or jets without identification of their origin). The analysis on the full Run 2 event sample was carried out, with providing statistical correctness and excluding a wide range of masses of the SUSY particles searched for. For a greater correctness, all the results were recalculated to eliminate discrepancies between versions of the data and generated MC background and signal event samples using a single UltraLegacy (UL) version in all the cases.

REFERENCES

1. CMS Collaboration, The CMS Experiment at the CERN LHC // *Journal of Instrumentation.* 2008, v. 3, S08004. https://doi.org/10.1088/1748-0221/3/08/S08004

2. CMS Collaboration, The CMS trigger system // *Journal of Instrumentation.* 2017, v. 12, P01020. https://doi.org/10.1088/1748-0221/12/01/P01020

3. CMS Collaboration, Strategy and performance of the CMS long-lived particle trigger program in proton-proton collisions at √s = 13.6 TeV. *e-Print: arXiv:2601.17544 [hep-ex]* (2026); https://doi.org/10.48550/arXiv.2601.17544

4. CMS Collaboration, Enriching the physics program of the CMS experiment via data scouting and data parking // *Physics Reports.* 2025, v. 1115, p. 678; https://doi.org/10.1016/j.physrep.2024.09.006

5. CMS Collaboration, Study of spin correlations in the Higgs boson decays to four leptons at CMS // *CMS Physics Analysis Summary.* 2025, CMS-PAS-HIG-25-011

6. CMS Collaboration, Improved measurement of di-Higgs production in the 4b final state with Run-3 data from the CMS Detector // *CMS Physics Analysis Summary.* 2025, CMS-PAS-HIG-24-010

7. CMS Collaboration, A search for HH→bbγγ production in proton-proton collisions at √s= 13.6 TeV with a partial CMS Run 3 dataset // *CMS Physics Analysis Summary.* 2025, CMS-PAS-HIG-25-007

8. CMS Collaboration, Observation of tWZ production at the CMS experiment // *Physical Review Letters.* 2026, v. 136, p. 081802; https://doi.org/10.1103/rk6w-1pcl

9. CMS Site Ranking Selection, https://cmssst.web.cern.ch/ranking

10. CMS FTS Metrics, https://monit-grafana.cern.ch/d/mtQFDScGk/cms-fts-metrics

11. EGI Production Accounting, https://accounting.egi.eu/

12. LatinoAnalysis, https://github.com/latinos/LatinoAnalysis

ОБРОБКА ДАНИХ ЕКСПЕРИМЕНТУ CMS У ННЦ ХФТІ

О. Куров, Л. Левчук, С. Лук'яненко, В. Попов, О. Приставка, О. Безпалов, Л. Ковальова

Представлено поточний стан експерименту CMS на Великому адронному колайдері (LHC). Протягом багатьох років у ННЦ ХФТІ виконуються роботи з обробки експериментальної інформації, яка отримується в експерименті CMS. Зокрема, функціонує спеціалізований інформаційно-обчислювальний комплекс, який є центром 2-го (T2) ярусу грід-інфраструктури CMS, T2_UA_KIPT, та єдиним обчислювальним центром України, в якому практично в безперервному (24/7) режимі виконується обробка експериментальної інформації з LHC. У 2025 р. виконано широкий спектр необхідних змін у конфігурації центру T2_UA_KIPT, і забезпечено високий рівень його надійності, який став одним із найкращих серед T2-сайтів CMS. За цей час обсяг переданої до центру експериментальної інформації CMS для обробки склав 1,74 ПБ. Також у 2025 р. було виконано аналіз повної вибірки протон-протонних зіткнень при 13 ТеВ, зареєстрованих в експерименті CMS у 2-му сеансі LHC, із метою пошуку сигналу суперсиметрії (SUSY) – прямого народження пар найлегших чарджіно з їх подальшим каскадним розпадом на заряджені лептони (електрони/позитрони та/або мюони), нейтрино та найлегші нейтраліно. Проведено оптимізацію сигнальних областей, на основі яких здійснено аналіз даних із забезпеченням статистичної коректності та виключенням широкого діапазону мас шуканих SUSY-частинок.