

*А. И. Аникина, Д. В. Беляков, Т. Ж. Бежанян, М. Х. Киракосян,
А. А. Кокорев, М. А. Любимова, М. А. Матвеев, Д. В. Подгайный,
А. Р. Рахмонова, О. И. Стрельцова, Ш. Г. Торосян, М. Валя, М. И. Зувев*

Возможности программно-информационной среды гетерогенной вычислительной платформы HybriLIT для задач ОИЯИ

Гетерогенная вычислительная платформа HybriLIT [1] является частью Многофункционального информационно-вычислительного комплекса ЛИТ им. М. Г. Мещерякова ОИЯИ. Платформа представляет собой многокомпонентную систему, при этом основным вычислительным ресурсом является суперкомпьютер «Говорун» с пиковой производительностью 1,7 Пфлопс для вычислений с двойной точностью и 26 Пфлопс с половинной точностью для задач искусственного интеллекта. Не менее важным компонентом платформы является иерархическая система обработки и хранения данных, реализованная на базе параллельной файловой системы Lustre общей емкостью 8,6 ПБ. Для распро-

странения программного обеспечения на платформе используются сетевая файловая система CernVM-FS и менеджеры лицензий FlexLM/MathLM. Доступ к ресурсам платформы в различных режимах реализован через пул виртуальных машин (пользовательских интерфейсов). На платформе развернуты информационные сервисы, предназначенные для поддержки работы пользователей. Программно-аппаратная структура и сервисы платформы представлены на рис. 1.

Все компоненты платформы объединены единой программно-информационной средой (рис. 2), позволяющей пользователям применять доступные пакеты прикладных программ, разрабатывать собственные

*A. I. Anikina, D. V. Belyakov, T. Zh. Bezhanyan, M. Kh. Kirakosyan,
A. A. Kokorev, M. A. Lyubimova, M. A. Matveev, D. V. Podgainy,
A. R. Rakhmonova, O. I. Streltsova, Sh. G. Torosyan, M. Vala, M. I. Zuev*

Capabilities of the Software and Information Environment of the HybriLIT Heterogeneous Computing Platform for JINR Tasks

The HybriLIT Heterogeneous Computing Platform [1] is part of the Multifunctional Information and Computing Complex of MLIT JINR. The Platform represents a multicomponent system, with the major computing resource being the Govorun supercomputer with a peak performance of 1.7 PFlops for double-precision computations and 26 PFlops with half-precision for artificial intelligence tasks. An equally significant component of the Platform is the hierarchical data processing and storage system, which is implemented on the basis of the Lustre parallel file system with a total capacity of 8.6 PB. The CernVM-FS

network file system and FlexLM/MathLM license managers are employed to distribute software on the Platform. Access to the Platform's resources in various modes is implemented via a pool of virtual machines (user interfaces). Information services designed to support user work are deployed on the Platform. The software-hardware structure and services of the Platform are shown in Fig. 1.

All components of the Platform are combined by a unified software and information environment (Fig. 2), which enables to use available application software packages, develop applications and perform computations us-

приложения и проводить расчеты с использованием различных типов вычислительных архитектур (CPU и GPU). Программно-информационная среда платформы представлена тремя уровнями: системным, программным и информационным.

На информационном уровне размещены информационные сервисы, помогающие пользователям в работе на платформе HybriLIT. На веб-сайте [1] представлено подробное описание всей платформы: программно-аппаратная структура, характеристики вычислительных ресурсов, приведены примеры работы с установленным прикладным программным обеспечением. Для взаимодействия пользователей предоставлен сервис совместной разработки GitLab [2]. Сервис HybriLIT User Support в среде JINR Project Management

Service [3] используется для консультирования и решения возникающих вопросов пользователей по работе на платформе. Оперативное информирование пользователей ведется с помощью Telegram-канала HybriLIT User Support [4].

На системном уровне размещены базовые программные компоненты, обеспечивающие функционирование платформы в качестве вычислительной системы. Системное программное обеспечение включает в себя инструменты для развертывания и управления операционной системой, систему аутентификации и авторизации пользователей, менеджера ресурсов и планировщика задач, сетевые файловые системы и систему распространения прикладного программного обеспечения. Важным компонентом системного уровня

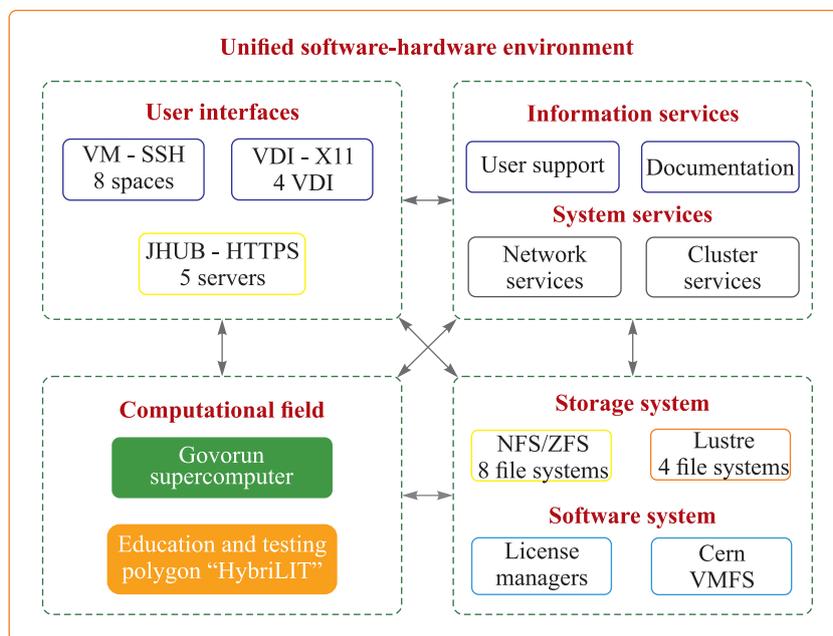


Рис. 1. Программно-аппаратная структура и сервисы платформы

Fig. 1. Software-hardware structure and services of the Platform

ing various types of computing architectures (CPU and GPU). The software and information environment of the Platform is represented by three levels: system, software, and information.

The information level embraces information services that help users work on the HybriLIT Platform. The website [1] provides a detailed description of the entire Platform: software-hardware structure, characteristics of computing resources, examples of working with the installed application software. The GitLab collaborative development service [2] is employed for user interaction. The HybriLIT User Support service in the JINR Project Management Service [3] environment is used to consult

and resolve issues related to the work process on the Platform. Users are promptly informed via the HybriLIT User Support Telegram channel [4].

The system level integrates the basic software components that ensure the functioning of the Platform as a computing system. The system software comprises tools for deploying and managing the operating system, a user authentication and authorization system, a resource manager and a task scheduler, network file systems, an application software distribution system. An important component of the system level is monitoring services that allow monitoring the operability and workload of the Platform.

The software level contains application software packages and services for user interaction with the resources of

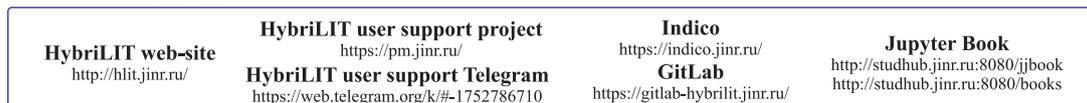
ня являются сервисы мониторинга, позволяющие следить за работоспособностью и загруженностью платформы.

На программном уровне размещены пакеты прикладных программ, сервисы для интерактивной работы пользователей с ресурсами платформы в различных режимах (рис. 3): для работы с помощью планировщика задач (в режиме очередей SLURM), для работы с программами с графическим интерфейсом (в режиме

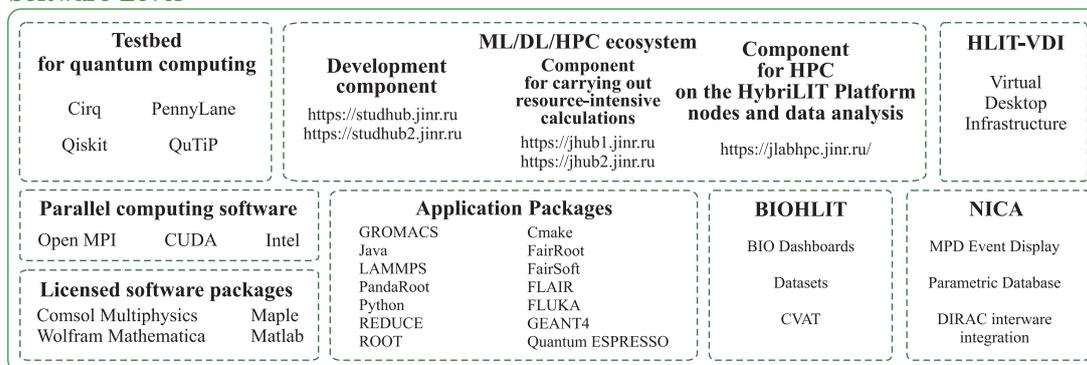
удаленного рабочего стола HLIT-VDI) и через веб-браузер для работы с экосистемой ML/DL/HPC и полигоном для квантовых вычислений.

Сервис HLIT-VDI [5] предназначен для работы с пакетами прикладных программ (Comsol, Wolfram Mathematica, Maple, Matlab и др.), использующими графический интерфейс, в режиме удаленного рабочего стола с помощью клиента X11 (TurboVNC Viewer)

Information Level



Software Level



System Level

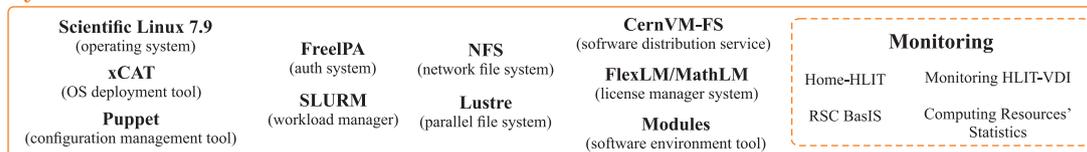


Рис. 2. Программно-информационная среда платформы

Fig. 2. Software and information environment of the Platform

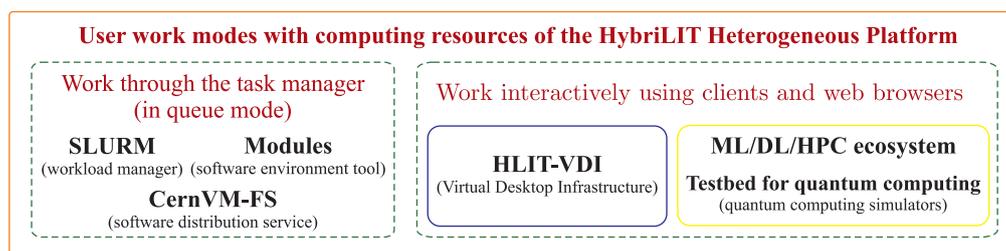


Рис. 3. Режимы работы пользователей с ресурсами платформы

Fig. 3. Modes of user interaction with the resources of the Platform

the Platform in various modes (Fig. 3): for work using the task scheduler (in the SLURM queue mode); for work with programs with a graphical interface (in the HLIT-VDI remote desktop mode) and via a web browser for work with the ML/DL/HPC ecosystem and the quantum computing polygon.

The HLIT-VDI service [5] is designed to work with application packages (Comsol, Wolfram Mathematica, Maple, Matlab, etc.), which use a graphical interface, in remote desktop mode via the X11 client (TurboVNC Viewer) on virtual machines hosted on a dedicated server.

на виртуальных машинах, размещенных на выделенном сервере.

Экосистема ML/DL/HPC [6] разработана командой HybriLIT на основе многопользовательской среды JupyterLab для работы с Jupyter Notebook. Она способствует разработке методов и алгоритмов машинного и глубокого обучения, упрощению работы с большими данными, анализу данных и научной визуализации, а также проведению лекций, tutorиалов и мастер-классов для студентов, аспирантов и научных сотрудников ОИЯИ.

С использованием ресурсов экосистемы ML/DL/HPC для решения задач, связанных с разработкой квантовых алгоритмов с применением симуляторов квантовых вычислений, развивается полигон для квантовых вычислений [7]. Работа на полигоне организована в двух режимах.

1. С помощью планировщика задач (в режиме очередей SLURM).

В данном случае используются симуляторы квантовых вычислений и необходимые библиотеки, установленные в сетевой файловой системе CernVM-FS. Преимуществом данного режима работы является возможность использовать все вычислительные ресурсы

суперкомпьютера «Говорун» при использовании квантового симулятора.

2. В интерактивном режиме через веб-браузер.

В данном случае используются выделенные серверы с графическими ускорителями. Симуляторы квантовых вычислений и необходимые библиотеки установлены в локальную файловую систему сервера в независимые виртуальные окружения Python (virtualenv), доступные в виде ядер интерактивного Python (ipython) в среде JupyterLab. К преимуществам работы в данном режиме относятся возможности вести разработку и отладку квантовых алгоритмов и визуализировать квантовые схемы.

С использованием ресурсов экосистемы ML/DL/HPC в рамках совместного проекта между ЛРБ и ЛИТ разработана информационная система BIOHLIT [8] для автоматизации разметки и анализа данных экспериментов: фото- и видеоматериалов. Эти исследования направлены на изучение влияния ионизирующего излучения на мелких лабораторных животных (мышей, крыс) и состоят из двух этапов. На первом этапе осуществляется проведение поведенческих тестов на специализированных установках, таких как «Открытое поле» и «Водный лабиринт Морриса», имеющихся в ЛРБ. Для этих исследований приме-

The ML/DL/HPC ecosystem [6] was developed by the HybriLIT group on top of the JupyterLab multi-user environment for work with Jupyter Notebook. It contributes to the development of machine and deep learning methods and algorithms, the simplification of work with Big Data, data analysis and scientific visualization, as well as to the holding of lectures, tutorials and workshops for students, postgraduates, and JINR researchers.

Using the resources of the ML/DL/HPC ecosystem to solve tasks related to the elaboration of quantum algorithms with the application of quantum computing simulators, the quantum computing polygon is being developed [7]. The polygon operates in two modes.

1. Using the task scheduler (SLURM queue mode).

In this case, quantum computing simulators and the required libraries installed in the CernVM-FS network file system are used. The advantage of this operating mode is the ability to utilize all computing resources of the Gоворун supercomputer when using a quantum simulator.

2. Interactively via a web browser.

In this case, dedicated servers with graphics accelerators are used. Quantum computing simulators and the required libraries are installed in the local file system of the

server in independent Python virtual environments (virtualenv), available as interactive Python kernels (ipython) in the JupyterLab environment. The advantages of working in this mode include the ability to develop and debug quantum algorithms, and visualize quantum circuits.

Using the resources of the ML/DL/HPC ecosystem, the BIOHLIT information system [8] was developed within a joint project of LRB and MLIT to automate the marking and analysis of experimental data, namely, photo and video materials. This research is aimed at studying the effect of ionizing radiation on small laboratory animals (mice, rats) and comprises two stages. At the first stage, behavioral tests are conducted on specialized setups, such as the Open Field and the Morris Water Maze, available at LRB. Computer vision algorithms are used for these studies to record the orientation and exploratory reactions of laboratory animals. For this purpose, the number of sectors passed, exits to the center, as well as grooming, movement in place, freezing, etc., are recorded. At the second stage, based on machine and deep learning algorithms, a morphofunctional analysis of cerebral cortex sections is performed to detect damaged neurons.

няются алгоритмы компьютерного зрения для регистрации ориентировочно-исследовательских реакций лабораторных животных. С этой целью фиксируется количество пройденных секторов, выходов в центр, а также груминг, движение на месте, замирание и др. На втором этапе на основе алгоритмов машинного и глубокого обучения осуществляется морфофункциональный анализ срезов коры головного мозга с целью нахождения поврежденных нейронов.

Ресурсы экосистемы ML/DL/HPC задействованы в реализации совместного проекта ЛТФ и ЛИТ по моделированию гибридных наноструктур сверхпроводник/магнетик. В рамках этого проекта был разработан пакет инструментов в виде Jupyter Notebook, который размещен в формате электронных публикаций Jupyter Book на ресурсах платформы [9]. Эти инструменты позволяют моделировать динамику φ_0 джозефсоновского перехода и явления переворота намагниченности в нем. Также реализованы алгоритмы моделирования динамики джозефсоновского перехода под воздействием внешнего электромагнитного излучения и сверхпроводникового квантового интерферометра. Разработаны параллельные алгоритмы, позволяющие ускорить вычисления различных физических

характеристик систем, содержащих джозефсоновские структуры.

Ресурсы платформы, особенно вычислительные ресурсы суперкомпьютера «Говорун» и иерархическая система обработки и хранения данных, широко задействованы в решении задач, связанных с реализацией мегасайнс-проекта NICA. На них выполняются ресурсоемкие расчеты в области решеточной квантовой хромодинамики для исследования свойств адронной материи при высоких плотностях энергии и в присутствии сверхсильных электромагнитных полей. Эти ресурсы были также интегрированы в распределенную гетерогенную среду на базе DIRAC Interware [10], что позволило решать задачи массовой генерации и реконструкции событий эксперимента MPD, а также привлечь эти ресурсы для обработки данных эксперимента BM@N.

Стоит отметить, что ресурсы HybriLIT используются как базовая платформа для изучения новых IT-технологий и подготовки IT-специалистов, что позволяет поддерживать высокий уровень компетенций пользователей и обеспечивать эффективное использование программно-информационной среды. Проводятся обучающие курсы и практикумы, которые посещают сотрудники ОИЯИ и участники научных

The resources of the ML/DL/HPC ecosystem are used in the implementation of a joint project of BLTP and MLIT on modeling hybrid superconductor/magnetic nanostructures. Within this project, a toolkit in the form of Jupyter Notebook was developed, it was posted in the Jupyter Book electronic publication format on the resources of the Platform [9]. These tools allow modeling the dynamics of the φ_0 Josephson junction and magnetization reversal phenomena in it. Algorithms for modeling the dynamics of the Josephson junction under the influence of external electromagnetic radiation and a superconducting quantum interferometer were also implemented. Parallel algorithms that enable the speedup of the computations of various physical characteristics of systems containing Josephson structures were elaborated.

The Platform's resources, especially the computing resources of the Govorun supercomputer and the hierarchical data processing and storage system, are widely used in solving tasks related to the implementation of the NICA megascience project. Resource-intensive computations in the field of lattice quantum chromodynamics to investigate the properties of hadronic matter at high energy densities and in the presence of ultrastrong electromagnetic fields

are performed on these resources. The resources were also integrated into a distributed heterogeneous environment based on the DIRAC Interware [10], which made it possible to solve the tasks of mass generation and reconstruction of MPD experiment events, as well as to involve these resources in processing BM@N experiment data.

It is noteworthy that HybriLIT resources are employed as a basic platform for studying novel IT technologies and training IT specialists, which allows maintaining a high level of user competence and ensuring the efficient use of the software and information environment. The educational activity embraces training courses and workshops attended by JINR staff members and scientific school participants. MLIT annually hosts the JINR International School of Information Technologies aimed to involve young specialists in solving tasks that face JINR using state-of-the-art information technologies. Within the IT School, the HybriLIT group conducts lectures and tutorials.

At present, the Platform's resources, especially the resources of the Govorun supercomputer, are used by research groups from all of the Institute's laboratories, as well as by students and postgraduates within the JINR educational activity in the field of information technolo-

школ. ЛИТ ежегодно организует международную школу по информационным технологиям, направленную на привлечение молодых специалистов к решению задач ОИЯИ с использованием современных информационных технологий. В рамках IT-школы группа HybriLIT проводит лекции и практические занятия.

В настоящее время ресурсы платформы, и прежде всего суперкомпьютера «Говорун», используются научными группами из всех лабораторий ОИЯИ, а также студентами и аспирантами в рамках образовательной активности Института в области информационных технологий. Всего к ресурсам платформы имеют доступ 480 человек, 347 из них имеют доступ к ресурсам суперкомпьютера «Говорун», который предоставляется только пользователям, принимающим непосредственное участие в реализации Проблемно-тематического плана ОИЯИ. К настоящему времени пользователями платформы опубликовано 336 научных работ, результаты которых были получены с использованием ее ресурсов.

Список литературы

1. Гетерогенная платформа «HybriLIT». <http://hlit.jinr.ru/>.
2. HybriLIT GitLab. <https://gitlab-hybrilit.jinr.ru/>.

3. JINR Project Management Service. <https://pm.jinr.ru/>.
4. HybriLIT: User Support. <https://web.telegram.org/k/#-1752786710/>.
5. HLIT-VDI. <http://hlit.jinr.ru/hlit-vdi/>.
6. Экосистема для задач машинного обучения, глубокого обучения и анализа данных. http://hlit.jinr.ru/access-to-resources/ecosystem-for-ml_dl_bigdataanalysis-tasks/.
7. Полигон для квантовых вычислений. <http://hlit.jinr.ru/quantum-polygon/>.
8. Информационная система для радиобиологических исследований. <https://bio.jinr.ru/>.
9. Toolkit for Modeling Superconductor/Magnetic Hybrid Nanostructures. <http://studhub.jinr.ru:8080/books/>, <http://studhub.jinr.ru:8080/jjbook/>.
10. *Kutovskiy N., Mitsyn V., Moshkin A., Pelevanyuk I., Podgayny D., Rogachevsky O., Shchinov B., Trofimov V., Tsaregorodtsev A.* // Phys. Part. Nucl. 2021. V.52, No.4. P.835–841. doi:10.1134/S1063779621040419.

gies. In total, 480 people have access to the resources of the Platform, 347 of whom have access to the resources of the Govorun supercomputer, which is provided only to users directly involved in the implementation of the JINR Topical Plan. To date, 336 scientific papers, the results of which were obtained using the resources of the Platform, are published by the Platform's users.

References

1. HybriLIT Heterogeneous Platform. <http://hlit.jinr.ru/>.
2. HybriLIT GitLab. <https://gitlab-hybrilit.jinr.ru/>.
3. JINR Project Management Service. <https://pm.jinr.ru/>.
4. HybriLIT: User Support. <https://web.telegram.org/k/#-1752786710/>.
5. HLIT-VDI. <http://hlit.jinr.ru/hlit-vdi/>.
6. Ecosystem for the Tasks of Machine Learning, Deep Learning and Data Analysis. http://hlit.jinr.ru/access-to-resources/ecosystem-for-ml_dl_bigdataanalysis-tasks/.
7. Quantum Computing Polygon. <http://hlit.jinr.ru/quantum-polygon/>.

8. Information System for Radiobiological Research. <https://bio.jinr.ru/>.
9. Toolkit for Modeling Superconductor/Magnetic Hybrid Nanostructures. <http://studhub.jinr.ru:8080/books/>, <http://studhub.jinr.ru:8080/jjbook/>.
10. *Kutovskiy N., Mitsyn V., Moshkin A., Pelevanyuk I., Podgayny D., Rogachevsky O., Shchinov B., Trofimov V., Tsaregorodtsev A.* // Phys. Part. Nucl. 2021. V.52, No.4. P.835–841. doi:10.1134/S1063779621040419.