

УДК 004.4

## Программно-информационная среда Гетерогенной вычислительной платформы HybriLIT

М.А. Любимова<sup>1</sup>, А.И. Аникина<sup>1</sup>, Д. В. Беляков<sup>1</sup>, М. Валя<sup>2</sup>, Т. Н. Зайкина<sup>1</sup>, М. И. Зуев<sup>1</sup>, М. Х. Киракосян<sup>1</sup>,  
Д. В. Подгайный<sup>1</sup>, О. И. Стрельцова<sup>1</sup>, Ш. Г. Торосян<sup>1</sup>

<sup>1</sup> Лаборатория информационных технологий,  
Объединённый институт ядерных исследований,  
ул. Жолио-Кюри 6, Дубна, Московская область, Россия, 141980  
<sup>2</sup> Университет Павла Йозефа Шафарика, Кошице, Словакия

Email: lma@jinr.ru

Программно-информационная среда Гетерогенной вычислительной платформы Hybri-LIT (ЛИТ ОИЯИ), состоящей из суперкомпьютера «Говорун», учебно-тестового полигона, а также ряда специализированных экосистем, предназначена для обеспечения разработки приложений, проведения высокопроизводительных вычислений, обработки и хранения данных. Для повышения эффективности работы пользователей Платформы внедряются и развиваются информационные сервисы, сервисы по взаимодействию пользователей и группы HybriLIT, осуществляющей поддержку пользователей, консультирование пользователей по технологиям параллельного программирования, организации работы с данными, установку и поддержку программного обеспечения для решения прикладных задач по требованиям пользователей. Сама программно-информационная среда содержит три уровня: системный уровень, состоящий из системного программного обеспечения; программный уровень, предоставляющий пакеты прикладных программ для решения массивно-параллельных и ресурсоемких задач; информационный уровень, который включает в себя информационные сервисы, помогающие работе пользователей на Платформе.

В докладе представлена методология создания Программно-информационной среды платформы HybriLIT, которая может быть полезна при создании таких сред на других высокопроизводительных вычислительных системах.

**Ключевые слова:** высокопроизводительные системы, программное обеспечение, гетерогенная платформа, программно-информационные среды.

### 1. Введение

Гетерогенная вычислительная платформа HybriLIT [1] входит в состав Многофункционального информационно-вычислительного комплекса (МИВК) [2] Лаборатории информационных технологий имени М.Г. Мещерякова (ЛИТ) ОИЯИ. Платформа представляет собой многокомпонентную систему, состоящую из Учебно-тестового полигона и суперкомпьютера (СК) «Говорун», объединенных единой программно-информационной средой. Суперкомпьютер «Говорун» используется для проведения ресурсоемких и массивно-параллельных расчетов, позволяя решать широкий спектр задач, стоящих перед ОИЯИ, в том числе для мегапроекта NICA [3]. Использование различных ускорителей вычислений, делает его эффективным инструментом для проведения ресурсоемких вычислений. Учебно-тестовый полигон применяется для исследования новых вычислительных архитектур и IT-решений. Его ресурсы используются для проведения учебных курсов по параллельному программированию, разработке и отладке параллельных приложений пользователей.

Платформа HybriLIT создана в соответствии с основными принципами разработки информационных систем [4], такими как: гибкость — возможность быстрого формирования IT-среды для решения прикладных задач; масштабируемость — возможность быстрого увеличения/уменьшения вычислительного поля; открытость — оперативный отклик на запросы пользователей платформы.

Эффективное использование ресурсов Платформы невозможно без хорошо организованных средств и сервисов поддержки пользователей.

В данной статье приведен опыт создания программно-информационной среды платформы HybriLIT, который может быть полезен при разработке таких сред для других высокопроизводительных вычислительных систем.

## 2. Программно-информационная среда

Программно-информационная среда Платформы, призванная облегчить пользователям работу с вычислительными ресурсами и повысить эффективность их использования, может быть разделена три уровня: системный, программный и информационный (рис. 1).

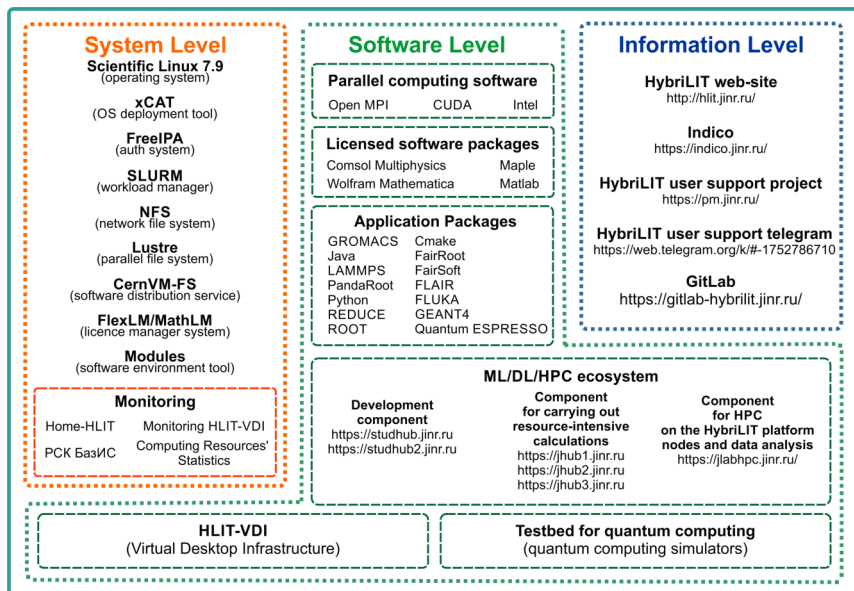


Рис. 1. Программно-информационная среда платформы HybriLIT

На системном уровне программно-информационной среды находятся основные программные компоненты, обеспечивающие функционирование всей системы в целом. Системное программное обеспечение (ПО) включает в себя операционную систему, batch-систему, файловые системы и другие компоненты, необходимые для управления аппаратными и программными ресурсами. Важным компонентом системного уровня являются сервисы мониторинга, позволяющие следить за работоспособностью и загруженностью Платформы.

Программный уровень содержит пакеты прикладных программ, предназначенных для обеспечения научных расчетов, инженерных вычислений, обработки и анализа данных.

Информационный уровень включает в себя различные сервисы, которые способствуют активному и тесному взаимодействию пользователей, в том числе с группой

HybriLIT. Эти сервисы значительно облегчают работу пользователей, предоставляют доступ к необходимой информации и позволяют обмениваться данными.

### 3. Системное программное обеспечение

На системном уровне находятся базовые компоненты, которые обеспечивают работу всей Платформы. К ним относятся операционная система (ОС) Scientific Linux 7.9, а также система управления конфигурацией ОС — xCAT. Данные компоненты позволяют управлять ресурсами системы, поддерживать ее стабильность и безопасность.

Важным инструментом в области безопасности является система аутентификации и авторизации — FreeIPA. Она предоставляет возможность создания и управления учетными записями пользователей, определения политик безопасности и управления доступом к различным ресурсам.

Планировщик задач SLURM позволяет распределять задачи пользователей в соответствии с доступными вычислительными ресурсами.

Сетевая файловая система NFS предоставляет пользователям доступ к рабочим файлам и папкам на различных узлах Платформы. Для высокоскоростной обработки данных используется параллельная файловая система Lustre, которая обеспечивает высокую скорость операций чтения и записи данных.

Системы распространения прикладного программного обеспечения, такие как CernVM-FS и FlexLM/MathLM, обеспечивают загрузку ПО в окружении среды пользователя. Пакет Modules позволяет настраивать переменные среды и использовать установленные на Платформе пакеты программ.

Системы мониторинга и системы сбора статистики применяются для контроля за использованием ресурсов платформы, что помогает системным администраторам отслеживать техническое состояние всей Платформы. Некоторые сервисы являются собственной разработкой, в то время как другие основаны на готовых решениях, настроенных с учетом специфики Платформы. Сервис Home-HLIT, разработанный группой системных администраторов, позволяет отслеживать загрузку вычислительных ресурсов в режиме реального времени. Сервис Monitoring HLIT-VDI собирает и визуализирует информацию о состоянии удаленных рабочих столов и загрузке ресурсов на каждом сервере, входящем в сервис HLIT-VDI. Программное обеспечение РСК БазИС [5] – применяется для оркестрации вычислительных ресурсов и систем хранения данных суперкомпьютера «Говорун», а также включает в себя функции мониторинга, сбора и анализа статистики использования ресурсов и управления системой аварийного отключения в случае критических сбоев.

Разработанное в рамках развития Платформы web-приложение Computing Resources' Statistics формирует и визуализирует статистику по использованию вычислительных ресурсов в виде круговых диаграмм (рис. 2) и сводных таблиц, что позволяет оценить вклад каждой группы пользователей в процентном отношении по использованию ресурсов — по количеству задач и количеству ядро-часов.

Представленная структура системного уровня обеспечивает эффективное администрирование, что позволяет динамически увеличивать количество вычислительных узлов, обновлять и изменять программное обеспечение на любых узлах без прерывания работы всей Платформы, оперативно реагировать на сбои в работе Учебно-тестового полигона и СК «Говорун».

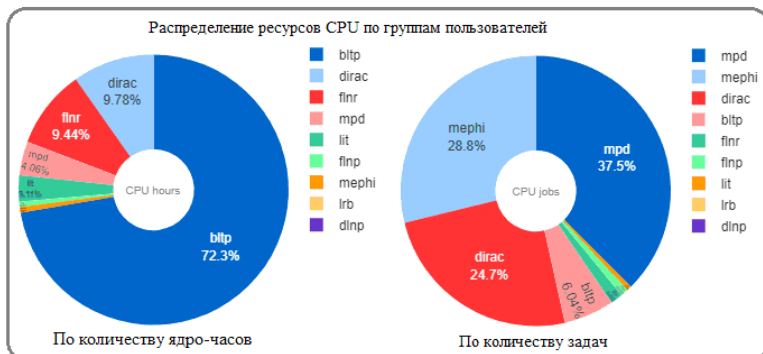


Рис. 2. Web-приложение Computing Resources' Statistics. Распределение ресурсов CPU по группам пользователей.

#### 4. Прикладное программное обеспечение

Второй уровень программно-информационной среды предоставляет пользователям возможность как разрабатывать собственные приложения с использованием компиляторов, библиотек и технологий для массивно-параллельных вычислений (Intel, CUDA, OpenMPI), так и применять установленное прикладное ПО. В настоящее время на Платформе установлено как специализированное лицензионное ПО Wolfram Mathematica для инженерных и математических расчетов, Comsol Multiphysics для моделирования физических процессов и решения задач в различных областях науки и техники, Matlab для решения задач технических вычислений и анализа данных и Maple для символьных и числовых вычислений, численного анализа и обработки данных, так и свободно распространяемое ПО для решения широкого спектра задач.

Для работы с прикладным программным обеспечением, использующим развитые графические интерфейсы, создан сервис HLLIT-VDI. Сервис позволяет через удаленный доступ на виртуальные машины работать с такими пакетами прикладных программ как Wolfram Mathematica, Maple, Matlab и др., обеспечивает как проведение небольших расчетов внутри виртуальных машин, так и выполнение ресурсоемких расчетов на вычислительных узлах Платформы.

Экосистема ML/DL/HPC разработана командой HybriLIT на основе JupyterHub — многопользовательской среды для работы с Jupyter Notebook. Данная среда используется для решения задач машинного и глубокого обучения с использованием фреймворков TensorFlow, PyTorch, Keras, поддерживающих обучение моделей на графических ускорителях.

На платформе HybriLIT развернут полигон для квантовых вычислений, позволяющий работать с квантовыми симуляторами (Cirq, Qiskit, PennyLane, QuTiP) в режиме удаленного доступа через терминал. Симуляторы подключаются путем добавления модулей из репозитория CernVM-FS. Для симуляторов, реализованных на языке Python, возможна работа через интерфейс JupyterLab, где квантовые симуляторы установлены в виде отдельных модулей со всеми необходимыми библиотеками. Примеры работы с симуляторами реализованы в JupyterNotebook, что позволяет моделировать и запускать квантовые схемы и сразу же видеть результаты выполнения кода.

## 5. Информационная поддержка пользователей

Внедрение современных IT-технологий и методов обработки информации, позволяющих эффективно использовать все возможности гетерогенной вычислительной платформы HybriLIT, требует наличия высококвалифицированной поддержки пользователей. Поддержка включает в себя разработку методических пособий, детальных инструкций и консультации специалистов технической поддержки. Для взаимодействия с пользователями используются электронные рассылки и оперативное оповещение пользователей через канал Телеграм.

Web-сайт HybriLIT [1] содержит описание вычислительной компоненты Платформы, методические пособия и инструкции для работы с различными технологиями программирования. На web-сайте описаны задачи, которые решают пользователи на ресурсах Платформы. Так, например, с 2023 по 2024 гг. пользователи опубликовали 77 статей, 21 из которых вышли в журналах с высоким уровнем цитирования (Q1/K1).

Система Indico применяется для организации конференций, семинаров и встреч, в том числе в гибридном формате. В период с 2023 по 2024 гг. с помощью Indico было проведено более 90 мероприятий. Сервис поддержки HybriLIT User Support, реализованный в системе Project Management Service, используется для взаимодействия пользователей с группой поддержки HybriLIT и предназначен для решения вопросов, возникающих у пользователей при работе на Платформе. За 2023 и 2024 годы с помощью этого сервиса было обработано 116 обращений. Для оперативного информирования пользователей используется Telegram-канал HybriLIT User Support. Сервис GitHub предназначен для совместной разработки ПО.

С начала работы Платформы команда поддержки HybriLIT собирает и анализирует предложения пользователей по улучшению Платформы. Также планируется подготовить раздел на web-сайте «Часто задаваемые вопросы», который будет содержать готовые решения распространённых проблем и ответы на вопросы пользователей.

Такая всесторонняя поддержка помогает пользователям осваивать новые IT-технологии и максимально использовать возможности гетерогенной вычислительной платформы HybriLIT.

## 6. Образовательная деятельность

Гетерогенная вычислительная платформа HybriLIT используется и 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.

Образовательная деятельность включает в себя проведение обучающих курсов и практикумов, которые посещают сотрудники ОИЯИ и участники научных школ.

В ЛИТ ежегодно проводится Международная школа по информационным технологиям ОИЯИ [6], направленная 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 2023 году в них приняли участие 60 студентов из разных вузов страны.

В Государственном университете «Дубна» на регулярной основе проводятся учебные курсы по таким дисциплинам как: «Архитектура и технологии высокопроизводительных систем», «Параллельные распределенные вычисления», «Языки и технологии анализа данных», «Технологии высокопроизводительных вычислений». За прошедший учебный год (2022-2023 гг.) эти курсы прослушали 230 студентов.

Также в Тверском государственном университете проведен курс «Программные средства математических вычислений», который посетили 40 студентов.

В 2023 году под руководством сотрудников HуbriLIT и с использованием ресурсов Платформы, было подготовлено пять бакалаврских работ и семь магистерских диссертаций.

## 7. Заключение

Гетерогенная вычислительная платформа HуbriLIT является активно развивающейся системой. В её состав добавляются новые вычислительные узлы и системы хранения данных, внедряются современные IT-решения. В соответствии с этим развивается программно-информационная среда Платформы, которая должна обеспечивать эффективное использование вычислительных ресурсов и сервисов на всех трех выделенных уровнях.

На развитие Платформы можно взглянуть и по-другому, а именно как на «горизонтальное» расширение и «вертикальный» рост.

На горизонтальном уровне Платформа развивается за счет добавления новых пакетов прикладного программного обеспечения, внедрения новых систем хранения данных и файловых систем. На информационном уровне — это внедрение новых систем коммуникации с пользователями.

Помимо горизонтального расширения, вертикальный рост обеспечивается включением в состав Платформы экосистем и полигонов, основанных на новых IT-решениях и парадигм вычислений, например, на Платформе созданы и активно развиваются Экосистема ML/DL/HPC [7] и Полигон для квантовых вычислений [8]. Все новые компоненты органично вписываются в существующую программно-информационную среду.

В заключение следует отметить, что для освоения новых технологий разрабатываются учебные материалы [9] и проводятся учебные курсы для пользователей, сотрудников ОИЯИ и студентов — будущих пользователей, в соответствии с развитием Платформы.

## Литература

1. Гетерогенная платформа «HуbriLIT». — URL: <http://hlit.jinr.ru/> (Дата обращения: 05.03.2024).
2. Многофункциональный информационно-вычислительный комплекс ОИЯИ. — URL: <https://misc.jinr.ru/> (Дата обращения: 05.03.2024).
3. Мегaproект NICA. — URL: <https://nica.jinr.ru/> (Дата обращения: 05.03.2024)
4. Юрченко, Т. В. Информационные системы в экономике и управлении: учебное пособие / Т. В. Юрченко; Нижегород. гос. архит.-строит. ун-т – Н.Новгород: НН-ГАСУ, 2013. – 114 с.
5. Платформа оркестрации РСК БазИС. — URL: <https://rscgroup.ru/product/rsc-basis-platform/> (Дата обращения: 05.03.2024).
6. Школа по информационным технологиям ОИЯИ <https://indico.jinr.ru/event/4274/> (Дата обращения: 05.03.2024)
7. Экосистема для задач машинного обучения, глубокого обучения и анализа данных — URL: [http://hlit.jinr.ru/access-to-resources/ecosystem-for-ml\\_dl\\_bigdataanalysis-tasks/](http://hlit.jinr.ru/access-to-resources/ecosystem-for-ml_dl_bigdataanalysis-tasks/) (Дата обращения: 05.03.2024)
8. Полигон для квантовых вычислений — URL: <http://hlit.jinr.ru/quantum-polygon/> (Дата обращения: 05.03.2024)
9. Python-инструментарий для моделирования динамики джозефсоновского перехода под воздействием внешнего излучения — URL: <http://studhub.jinr.ru:8080/jjbook/intro.html>

UDC 004.4

## Software and information environment of the HybriLIT Heterogeneous Computing Platform

M. A. Lyubimova<sup>1</sup>, A. I. Anikina<sup>1</sup>, D. V. Belyakov<sup>1</sup>, M. Vala<sup>2</sup>, M. Kh. Kirakosyan<sup>1</sup>, D. V. Podgainy<sup>1</sup>, O. I. Streltsova<sup>1</sup>,  
Sh. G. Torosyan<sup>1</sup>, T. N. Zaikina<sup>1</sup>, M. I. Zuev<sup>1</sup>

<sup>1</sup> *Laboratory of Information Technologies*

*Joint Institute for Nuclear Research*

*Joliot-Curie 6, Dubna, Moscow region, 141980, Russia*

<sup>2</sup> *Pavol Jozef Šafárik University in Košice, Slovak Republic*

Email: lma@jinr.ru

The software and information environment of the HybriLIT Heterogeneous Computing Platform (LIT JINR), consisting of the Govorun supercomputer, a testing polygon and a number of specialized ecosystems, is designed to support application development, high-performance computing, data processing and storage. To improve the efficiency of the platform, new services aimed at both information support and interaction between users and the members of the HybriLIT team have been developed. The software and information environment itself contains three levels: the system level which includes system software; software level, which provides application software packages for solving massively parallel and resource-intensive tasks; information level, which includes information services that help users work on the platform.

The paper presents methodology for developing a Software Information Environment of the HybriLIT platform, which can be useful in developing such environments on other high-performance computing systems.

**Key words and phrases:** high-performance systems, software, heterogeneous platform, software and information environments.