

РАЗВИТИЕ СИСТЕМЫ МОНИТОРИНГА ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ ГЕТЕРОГЕННОЙ ПЛАТФОРМЫ HybriLIT

Д. В. Беляков^{1,}, Д. Г. Дереза², Г. А. Карпов²,
М. П. Лебедев², М. А. Сказкин², М. И. Зуев¹*

¹ Объединенный институт ядерных исследований, Дубна, Россия

² Дальневосточный федеральный университет, Владивосток, Россия

В работе представлен обзор систем мониторинга различных компонент гетерогенной платформы HybriLIT. Сформулированы цели и назначение применяемых систем, являющихся одним из важных инструментов системного администрирования платформы. Для контроля за состоянием вычислительных ресурсов применяется разработанная ранее система мониторинга, позволяющая в режиме реального времени отслеживать загрузку CPU- и GPU-компонент вычислительного узла, использование оперативной памяти, систем хранения данных, объем сетевого трафика и т.п. В работе представлена новая система мониторинга, являющаяся логическим развитием разработанной ранее системы, обеспечивающая ряд дополнительных функций для контроля за состоянием вычислительных ресурсов платформы.

This paper presents an overview of monitoring systems for various components of the HybriLIT heterogeneous platform. The goals and purposes of these systems, which serve as important tools for system administration of the platform, are outlined. Previously developed monitoring system is used for monitoring the state of computing resources of the platform. It allows real-time tracking of the load of CPU and GPU of the computing node, RAM usage, storage systems usage, network traffic, etc. This paper presents a new monitoring system which is a logical evolution of the previously developed system. It provides a number of additional functions for monitoring the state of computing resources of the platform.

PACS: 89.20.Ff

ВВЕДЕНИЕ

Гетерогенная платформа HybriLIT [1] представляет собой многокомпонентную вычислительную систему, предназначенную для выполнения расчетов в рамках решения научно-исследовательских задач и проектов Объединенного института ядерных исследований. В состав платформы входит суперкомпьютер «Говорун», учебно-тестовый полигон, экосистема ML/DL/HPC и ряд систем хранения данных, построенных на основе

* E-mail: dmitry@jinr.ru

файловых систем NFS/ZFS и Lustre. Для поддержки работоспособности такой сложной системы используется ряд систем мониторинга, позволяющих отслеживать состояния компонент платформы, суммарной нагрузки на вычислительные ресурсы, системы хранения данных, локальную сеть. В связи с тем, что гетерогенная платформа HybriLIT является многопользовательской вычислительной системой, одной из важных задач системного администрирования является выявление и мониторинг нагрузки, создаваемой счетными задачами отдельного пользователя. В работе представлена новая система мониторинга вычислительных ресурсов, предоставляющая функции для отслеживания нагрузки на компоненты вычислительного узла (CPU, RAM, disk, network), возникающей в ходе выполнения счетной задачи пользователя.

СИСТЕМЫ МОНИТОРИНГА ГЕТЕРОГЕННОЙ ПЛАТФОРМЫ HybriLIT

На гетерогенной платформе HybriLIT применяется несколько различных систем мониторинга, которые позволяют отслеживать работу компонент платформы и осуществлять мониторинг нагрузки на вычислительные ресурсы.

1) Мониторинг системы охлаждения суперкомпьютера «Говорун» (РСК) [2]. Система охлаждения суперкомпьютера «Говорун» организована на основе жидкостного охлаждения. Для обеспечения непрерывной и бесперебойной работы вычислительных узлов необходимо контролировать работу узлов регулирования подачи охлаждающей жидкости и выполнять мониторинг температуры «холодной» и «горячей» жидкости. Для управления системой охлаждения суперкомпьютера «Говорун» применяется ПО, разработанное компанией РСК.

2) Мониторинг МИБК суперкомпьютера «Говорун» (ЛИТ) [3]. Счетные задачи пользователей на суперкомпьютере «Говорун» используют различные вычислительные ресурсы, организованные в виде нескольких счетных очередей. Для контроля за текущим состоянием счетных очередей и количеством выполняемых задач внедрен мониторинг ресурсов МИБК, построенный на основе системы GRAFANA и использующий актуальные данные от менеджера ресурсов SLURM.

3) Мониторинг SALSA вычислительных ресурсов (HybriLIT) [4]. В состав программно-информационной среды гетерогенной платформы HybriLIT входит набор сервисов, размещенных на физических серверах и ряде виртуальных машин. Для контроля работы аппаратной составляющей данной среды разработана и применяется специализированная система мониторинга SALSA, позволяющая в режиме онлайн отслеживать нагрузку на различные компоненты физических серверов и виртуальных машин. Система мониторинга SALSA предоставляет системным администраторам различные данные о суммарной нагрузке (CPU/GPU, RAM, disk, network) на вычислительные ресурсы платформы.

По мере увеличения вычислительных мощностей суперкомпьютера «Говорун» (модернизации 2024–2025 гг.) пользователи гетерогенной платформы HybriLIT получили возможность запускать на счет все более массивно-параллельные задачи (пакеты задач), создавая конкурирующую нагрузку не только на вычислительные ресурсы, но и на локальную сеть платформы и системы хранения данных. В связи с этим мониторинг нагрузки, создаваемой ресурсоемкими задачами пользователей, является одной из важных задач системного администрирования платформы.

Разработанный ранее (2016–2017 гг.) мониторинг SALSА в первую очередь был ориентирован на отслеживание суммарной нагрузки на компоненты физических серверов и виртуальных машин, представляющих собой площадку для размещения сервисов платформы. Необходимость мониторинга нагрузки, создаваемой счетной задачей пользователя, потребовала разработки новой системы мониторинга, способной предоставлять детальную информацию об использовании ресурсов отдельной задачей пользователя на выбранном вычислительном узле.

В рамках научной темы «Разработка и внедрение новых систем сбора и анализа статистики использования вычислительных ресурсов и прикладного программного обеспечения гетерогенной платформы HybriLIT» была разработана новая система мониторинга StarLIT, являющаяся логическим развитием мониторинга SALSА.

Основными элементами системы мониторинга StarLIT являются: сенсор (программа, собирающая метрики), сервер (backend, часть системы, выполняющая функции сбора, агрегации и хранения данных) и клиент (frontend, часть системы, представляющая собой пользовательский интерфейс к функционалу системы). Взаимодействие элементов системы происходит на основе двухканальной асинхронной передачи данных (канал support, канал payload) с помощью TCP- (для сенсора и сервера) и веб-сокетов (для клиента и сервера). Каждый элемент построен на основе микросервисной архитектуры, что позволяет легко выполнять модернизацию функций, не затрагивая другие функции системы.

Серверная часть системы (рис. 1, 2), в которую входят сервер и сенсор, разработана на языке программирования Python 3 с применением фреймворка FastAPI, баз данных SQLite, MongoDB и библиотеки psutil. Клиентская часть системы (пользовательский интерфейс) разработана на языке программирования JavaScript с применением фреймворков Vue.js, Bootstrap, библиотек Charts.js и Axios.

Аутентификация пользователей в системе организована с помощью внешнего сервиса FreeIPA, отвечающего за аутентификацию и авторизацию пользователей гетерогенной платформы HybriLIT. При первичной аутентификации пользователя система, получив логин и пароль, обращается к сервису FreeIPA и в случае успеха формирует JWT-токен, который отдает веб-браузеру пользователя. Данный JWT-токен заносится в хранилище HTTP-only Cookies, что предотвращает обращение к данным из любого скрипта JavaScript и обеспечивает высокий уровень защиты

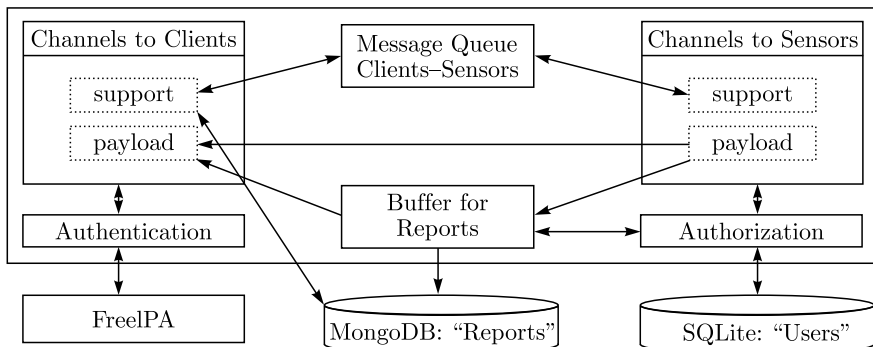


Рис. 1. Структура сервера

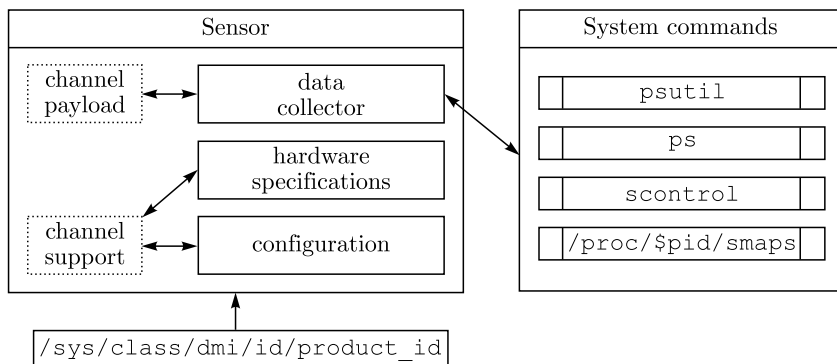


Рис. 2. Структура сенсора

учетной записи пользователя. При последующих обращениях к системе JWT-токен автоматически пересылается серверу в составе HTTP-запроса для подтверждения действий пользователя. В системе также реализована расширенная модель доступа к функциям системы на основе редактируемых ролей (ADMIN, USER), прав доступа к функциям системы и режимов мониторинга вычислительных ресурсов.

Мониторинг StarLIT поддерживает следующие режимы мониторинга нагрузки на компоненты вычислительных узлов:

- 1) режим «Standard» (суммарно по ресурсам) — аналог мониторинга SALSA;
- 2) режим «Extended» (детально по отдельным элементам компонентов);
- 3) режим «Intraday» (аналог режима «Extended» с историей до 24 ч);

4) режим «Analytic» (аналог режима «Extended» для накопления данных для внешней системы обработки, анализа и визуализации статистики);

5) режим «SLURM» (для мониторинга нагрузки от счетных задач SLURM).

Особенностью работы системы мониторинга StarLIT является динамическое поведение системы в зависимости от активности пользователя. В случае, когда в системе нет авторизованного пользователя, система выполняет лишь накопление данных для режимов «Intraday» и «Analytic». При наличии пользователя включаются режимы «Standard», «Extended» или «SLURM» в зависимости от выбора пользователя.

Пользовательский интерфейс системы организован в виде набора вкладок, на которых динамически отображаются различные мониторинговые данные в виде круговых диаграмм, таблиц и графиков. На вкладке «24-часового мониторинга» можно интерактивно менять диапазон «окна просмотра» от 1 мин до 24 ч. На вкладке «SLURM-мониторинг» отображается информация о нагрузке на вычислительный узел, создаваемой счетной задачей SLURM, включая все порождаемые процессы пользователя, запустившего счетную задачу.

В настоящее время мониторинг StarLIT развернут на ресурсах гетерогенной платформы HybriLIT и находится на этапе ввода в эксплуатацию.

СПИСОК ЛИТЕРАТУРЫ

1. Anikina A. I., Belyakov D. V., Bezhanyan T. G., Kirakosyan M. H., Koko-rev A. A., Lyubimova M. A., Matveev M. A., Podgainy D. V., Rahmonova A. R., Shadmehri S. A., Streltsova O. I., Torosyan Sh. G., Vala M., Zuev M. I. Structure and Features of the Software and Information Environment of the HybriLIT Heterogeneous Platform // Distributed Comput. and Commun. Networks (DCCN 2024). Lecture Notes in Comput. Sci. 2024. V. 15460. P. 444–457.
2. Жидкостное охлаждение ПСК. <https://rscgroup.ru/technology/liquidcooling/> (дата обращения 01.09.2025).
3. Baginyan A. S., Balashov N. A., Baranov A. V., Belov S. D., Belyakov D. V., Butenko Yu. A., Dolbilov A. G., Golunov A. O., Kadochnikov I. S., Kashunin I. A., Korenkov V. V., Kutovskiy N. A., Mayorov A. V., Mitsyn V. V., Pelevanyuk I. S., Semenov R. N., Strizh T. A., Trofimov V. V., Vala M. Multi-Level Monitoring System for Multifunctional and Computing Complex at JINR // Proc. 26th Intern. Symp. on Nuclear Electronics & Computing. 2017. P. 226–233.
4. Валя М., Майоров А. В., Бутенко Ю. А. Развитие сервиса Stat-ГибриЛит для мониторинга гетерогенного кластера HybriLIT // Тр. 7-й конф. (с междунар. участием) «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». 2017. С. 209–211.