

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОИЗВОДИТЕЛЬНОСТИ И ЭНЕРГОЭФФЕКТИВНОСТИ ПРОФИЛЕЙ КОНФИГУРАЦИИ ВЫЧИСЛИТЕЛЬНЫХ УЗЛОВ В СОСТАВЕ ГРИД-САЙТА JINR Tier-1

А. О. Голунов^{1,*}, *А. О. Голунов*^{1,**}, *С. С. Паржицкий*^{1,***}

¹ Объединенный институт ядерных исследований, Дубна, Россия

Предложен комплексный подход к настройке низкоуровневого программного обеспечения серверов для ресурсоемких задач обработки данных физических экспериментов с целью достижения баланса между максимальной производительностью и минимальным энергопотреблением. Для реализации подхода разработана система мониторинга, осуществляющая сбор метрик энергопотребления и температуры ключевых компонентов сервера в реальном времени. Создан специализированный инструмент анализа, позволяющий коррелировать данные мониторинга с результатами нагрузочного тестирования. Представлены расчетные коэффициенты производительности и энергоэффективности для различных настроек.

The paper proposes an integrated approach to configuring low-level server software for resource-intensive tasks of processing physical experiment data, in order to achieve a balance between maximum performance and minimum energy consumption. To implement the approach, a monitoring system that collects energy consumption and temperature metrics of key server components in real time has been implemented. A specialized analysis tool that allows one to correlate monitoring data with the results of load testing has been created. Calculated performance and energy efficiency coefficients for various settings are presented.

PACS: 89.20.Ff

ВВЕДЕНИЕ

Основными целями и задачами Лаборатории информационных технологий являются развитие и модернизация основных аппаратно-программных компонентов Многофункционального информационно-вычислительного комплекса (МИВК) [1] в соответствии с Семилетним планом развития ОИЯИ. Одним из таких компонентов является грид-сайт JINR

* E-mail: golunov@jinr.ru

** E-mail: golao@jinr.ru

*** E-mail: pss@jinr.ru

Tier-1 [2] для эксперимента CMS, серверные ресурсы которого сосредоточены преимущественно в выделенном аппаратном модуле МИВК. Все вычислительные узлы оснащены процессорами, оптимизированными под запуск 16-ядерных пилотов для задач по обработке данных эксперимента CMS на Большом адронном коллайдере (LHC). Согласно плану и для выполнения текущих обязательств перед коллаборацией в 2025 г. необходимо увеличить вычислительные ресурсы JINR Tier-1 до 23 000 процессорных ядер и до 400 000 HEPScore23 [3]. Для сохранения отказоустойчивости аппаратного модуля энергопотребление и тепловыделение новых серверов на данном этапе модернизации компонента не должно превышать пиковое потребление 25 кВт·ч.

ОБЪЕКТ ТЕСТИРОВАНИЯ

Для тестирования производительности процессора AMD EPYC9654 в 2025 г. у компании «Нерпа» приобретен двухпроцессорный сервер NерпаNORD5015 в следующей конфигурации:

- серверное шасси: 1U ASUS RS700A-E12-RS4U;
- процессоры: $2 \times$ AMD EPYC 9654 (96C/192T, 2,4 ГГц–3,7 ГГц);
- оперативная память: $24 \times$ 96 ГБ DDR5 ECC REG 4800 МГц;
- блоки питания: $2 \times$ 2600 Вт (N+1);
- диски: $2 \times$ 3,84 ТБ NVMe SSD;
- сеть: $2 \times$ 10 Гбит/с Ethernet;
- сервисный процессор (BMC): Aspeed AST2600.

В качестве нагрузочного тестирования использовался бенчмарк HEPScore23 под установленной ОС AlmaLinux 9.6 (5.14.0-570.24.1.el9_6.x86_64), включающий набор тестов, обеспечивающих нагрузку, максимально приближенную к обработке данных с крупных физических экспериментов на LHC.

Основными объектами исследования являлись:

- результаты бенчмарка в зависимости от примененных настроек BMC/BIOS;
- стабильность работы системы при выполнении нагрузочного тестирования;
- способность серверной платформы к эффективному охлаждению процессоров с сохранением максимальной доступной тактовой частоты при условии размещения в воздушной среде эксплуатации аппаратных модулей МИВК;
- суммарное накопленное потребление электроэнергии сервера, необходимое для выполнения бенчмарка.

Проведено более 17 тестов с различными профайлами настроек BMC/BIOS. Для контроля за состоянием температуры и энергопотребления во время тестирования использовались встроенные сенсоры BMC. Температура воздушной среды эксплуатации в периоды тестирования составляла 17 °C.

СИСТЕМА МОНИТОРИНГА И ИНСТРУМЕНТЫ АНАЛИЗА

На базе существующего сервиса мониторинга инженерной инфраструктуры M4top в связке с СУБД MariaDB и платформой Grafana развернута система мониторинга сервера Nерра (рис. 1). Тип получаемых данных: температура процессоров и оперативной памяти, входящего и исходящего воздушных потоков, системной платы, входящее и исходящее энергопотребление блоков питания сервера, энергопотребление процессоров и оперативной памяти, скорость вращения вентиляторов шасси, напряжение основных каналов питания. Частота обновления данных 0,2 Гц.

Бенчмарк NEPScore23 протоколирует временные интервалы запуска тестов. Зная временной интервал запущенной операции, можно выполнить анализ энергопотребления данной операции как для всего сервера, так и отдельно для ЦПУ и ОЗУ сервера. Для удобства анализа получаемых данных был внедрен Telegram bot, который по команде пользователя, с указанием временного периода анализа, отправляет запрос на сервер мониторинга. На сервере выполняется Python-скрипт анализа суммарного и среднего энергопотребления, а результаты возвращают-



Рис. 1. Страница мониторинга сервера Nерра

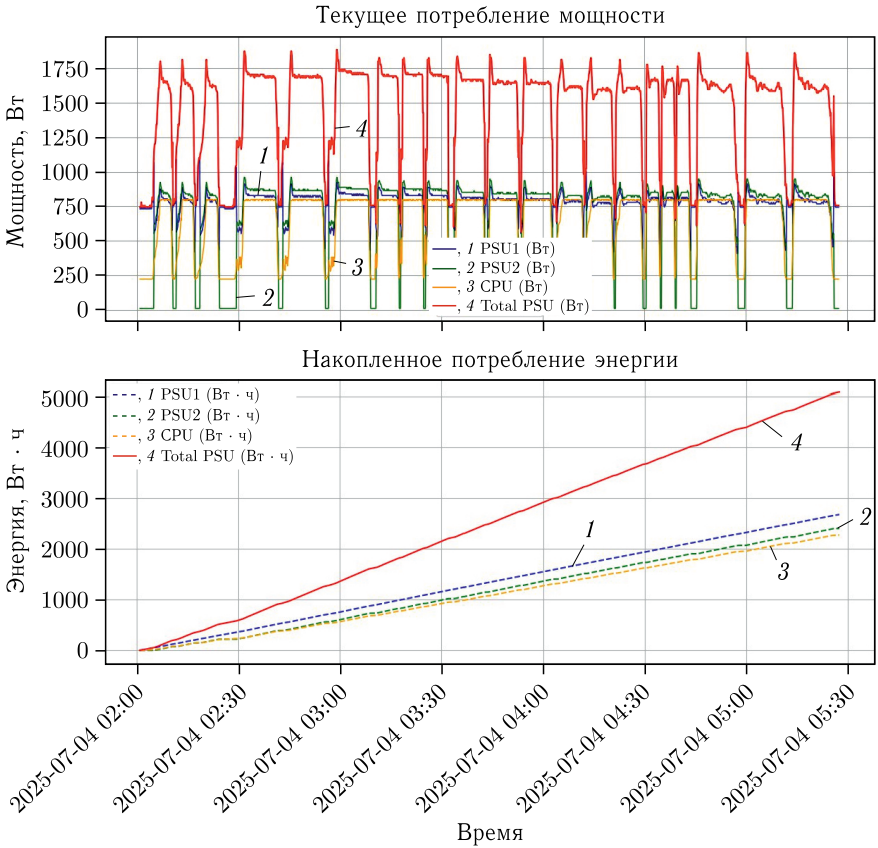


Рис. 2. Пример вывода в Telegram результатов анализа энергопотребления за период

ся в Telegram-messenger в текстовом виде. Опционально интегрирован функционал визуализации данных энергопотребления на базе пакета Matplotlib [4] (рис. 2).

ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

Основные результаты тестирования приведены в табл. 1. Данные по расчету коэффициентов производительности представлены в табл. 2.

Конфигурация сервера показала стабильную работу системы при размещении в аппаратном модуле МИВК во всех тестируемых режимах производительности. Оптимизация производительности осуществ-

ляется за счет комплексных настроек BIOS: Prefetcher, Simultaneous Multithreading (SMT), Engine boost, Configurable TDP.

Согласно полученным коэффициентам наиболее производительным режимом является HPC. Основными требованиями для корректного функционирования режима являются наличие дискретного контроллера дисковой подсистемы и температура среды эксплуатации менее 25 °С. Режим Energy efficiency оказывает минимальную нагрузку на инженерную инфраструктуру аппаратного модуля. Кроме режима производительности влияние на энергоэффективность сервера оказывает скорость работы вентиляторов шасси. В большинстве проведенных тестов максимальная скорость вентиляторов шасси оказалась избыточной и не оказывала существенного влияния на производительность системы. Ввиду разных временных интервалов, требуемых для выполнения сервером полного цикла бенчмарка, при размещении серверов в аппаратном модуле МИВК оптимальным режимом производительности BIOS является

Таблица 1. Результаты тестирования профилей

Профиль	FAN	SMT	HEPScore23	Время работы, ч	Потребление энергии, Вт · ч
Default	auto	disable	4976,89	2,63	2357,67
Default	full	disable	4979,38	2,62	2825,09
Energy efficiency	auto	disable	4329,14	3,27	2259,93
Energy efficiency	full	disable	4334,44	3,26	2998,67
HPC	auto	disable	6126,96	2,27	2636,03
HPC	full	disable	6214,93	2,25	3016,37
HPC	full	enable	7816,84	3,45	5081,41

Таблица 2. Расчет коэффициентов энергоэффективности и производительности

Профиль	FAN	SMT	Удельная энерго-эффективность, Вт/HEPScore23 [5]	Производительность на единицу энергии, HEPScore23/Вт · ч
Default	auto	disable	0,180123023	2,110935797
Default	full	disable	0,216548772	1,762556237
Energy efficiency	auto	disable	0,159641428	1,915608006
Energy efficiency	full	disable	0,212215978	1,445454151
HPC	auto	disable	0,189530648	2,324313456
HPC	full	disable	0,215707802	2,060400415
HPC	full	enable	0,188423003	1,538321057

профиль НРС с выключенной функцией SMT, а также настройка ВМС для автоматической регулировки скорости вентиляторов.

ЗАКЛЮЧЕНИЕ

Реализованы программные инструменты для хранения, определения и сравнения энергопотребления сервера при выполнении задач по обработке данных экспериментов в физике высоких энергий.

Протестировано серверное оборудование, и определены оптимальные настройки ВМС/BIOS, применимые для выполнения задачи по увеличению вычислительной мощности кластера грид-сайта JINR Tier-1 для эксперимента CMS в 2025 г.

СПИСОК ЛИТЕРАТУРЫ

1. *Korenkov V.* The JINR Multifunctional Information and Computing Complex // 2020 Intern. Scientific and Techn. Conf. "Modern Computer Network Technologies" (MoNeTeC). Moscow, 2020. P. 1–4.
2. *Baginyan A., Balandin A., Dolbilov A. et al.* JINR Grid Infrastructure: Status and Plans // Phys. Part. Nucl. 2024. V. 55. P. 355–359.
3. *Giordano D.* HEPsScore: A New CPU Benchmark for the WLCG // CERN Eur. Phys. J. Web Conf. 2024. V. 295; https://www.epj-conferences.org/articles/epjconf/pdf/2024/05/epjconf_chep2024_07024.pdf (accessed 17.09.2025).
4. *Hunter J. D.* Matplotlib: A 2D Graphics Environment // Comput. Science Engin. 2007. V. 9, No. 3. P. 90–95; DOI: 10.1109/MCSE.2007.55.
5. *George S.* HEPsScore and Power Measurements with AMD Bergamo // HEPiX Benchmarking Working Group Meeting. CERN, 2025; <https://indico.cern.ch/event/1508924/contributions/6349546/attachments/3008721/5304144/Power%20and%20HEPScore%20measurements%20talk.pptx> (accessed 17.09.2025).