

# **ИНТЕГРАЦИЯ КЛАСТЕРНОЙ СИСТЕМЫ МОНИТОРИНГА НА БАЗЕ Icinga2 В МИВК ЛИТ ОИЯИ**

*И. А. Кашунин, В. В. Мицын, В. В. Трофимов, А. Г. Долбилов*

Объединенный институт ядерных исследований, Дубна

Оборудование Многофункционального информационно-вычислительного комплекса (МИВК) ЛИТ ОИЯИ обеспечивает необходимые условия для организации процесса обработки данных. Диагностику этого оборудования проводит система мониторинга. Развитие вычислительного комплекса предполагает наращивание вычислительной мощности, что увеличивает количество обрабатываемых параметров. В данной статье описана разработка и реализация новой концепции кластерной системы мониторинга, позволяющей осуществлять оперативное наращивание дополнительных узлов системы при увеличении подключаемого оборудования МИВК.

Equipment of the Multifunctional Information and Computing Complex (MICC) at JINR LIT provides the necessary conditions for organizing the data processing. Diagnostics of this equipment is carried out by the monitoring system. The development of the computing complex involves the increase in computing power, which increases the number of processed parameters. This article describes the development and implementation of a new concept of a cluster monitoring system, which allows the operational expansion of additional system nodes with an increase in the connected MICC equipment.

PACS: 07.05.Kf; 29.50.+v; 29.85.Fj

## **1. ОБЩИЕ СВЕДЕНИЯ**

Современная физика напрямую связана с моделированием и хранением данных больших объемов. Все это требует наличия обширных вычислительных мощностей, которые смогут обеспечить оперативную обработку и хранение данных. Для увеличения отказоустойчивости вычислительного оборудования требуется своевременно реагировать на различные сбои и оперативно их устранять. Наличие систем мониторинга позволяет сократить время реагирования на сбой и таким образом обеспечить надежность оборудования, что позволяет увеличить надежность вычислительного комплекса в целом.

## **2. СИСТЕМА МОНИТОРИНГА ПРОШЛОГО ПОКОЛЕНИЯ**

В ходе эксплуатации оборудования вычислительного комплекса необходимо иметь различные средства диагностики. С данной проблемой столкнулся и МИВК ЛИТ ОИЯИ [1].

В 2014 г. была начата работа над созданием единой системы мониторинга компонентов МИВК [2]. Предполагалась следующая концепция: все элементы вычислительного комплекса можно отследить в единой точке доступа. В качестве основы была взята система мониторинга Nagios 3.5 [3]. Данная система позволила считывать стандартные метрики с большого количества оборудования:

- дисковых и вычислительных серверов;
- системы питания;
- системы охлаждения;
- ленточного робота.

В процессе эксплуатации выяснилось, что мониторинг стандартных метрик, доступных по умолчанию в Nagios, не позволяет получать данные обо всех сбоях различных устройств, поэтому были созданы свои плагины сбора данных с оборудования. Это позволило поднять уровень диагностики оборудования. Однако с ростом количества метрик информация, выдаваемая на экран, стала вновь несодержательной. Виной тому стал тот факт, что многие отказы оборудования, такие как рост температуры, инженеры способны распознать заранее, на основании графиков, однако по умолчанию система мониторинга их не поддерживала. Внедрение системы графиков позволило расширить уровень диагностики, но не обеспечило улучшения визуализации в целом, так как операторам все так же приходилось переключаться между различными вкладками в браузере, в которых содержались уже целые группы устройств. Решением данной проблемы стала разработка концепции информационных дисплеев. Это решение предполагало вывод на экран наиболее важных объектов слежения. Изначально данное решение было внедрено для системных администраторов, однако вскоре стало понятно, что такой подход предоставляет очень гибкие возможности для мониторинга. В стандартный интерфейс Nagios, как правило, заходили только системные администраторы для выполнения команд. Операторам же для функций слежения хватало системы визуализации (рис. 1).

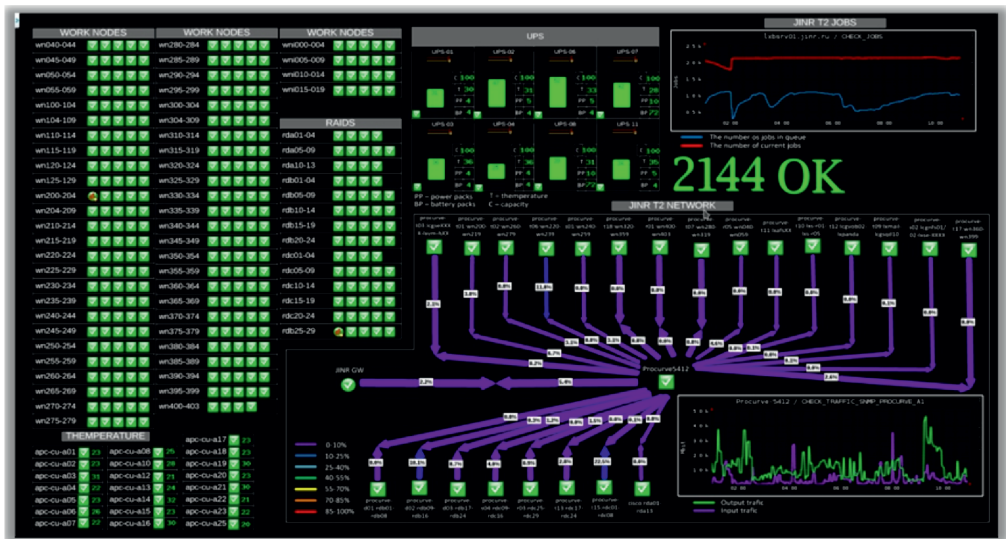


Рис. 1. Информационный дисплей актуальной системы мониторинга

### 3. ВОЗМОЖНОСТИ СИСТЕМЫ МОНИТОРИНГА ПРЕДЫДУЩЕГО ПОКОЛЕНИЯ

С ростом объема подключаемого оборудования вычислительного комплекса стало понятно, что производительность центрального процессора (ЦП) системы мониторинга является узким местом на пути увеличения масштабируемости Nagios 3.5.

На рис. 2 виден рост загрузки центрального процессора системы мониторинга.

Периодически в процессе работы сервера системы мониторинга появлялись всплески загрузки ЦП вплоть до 90%. Это говорило о том, что подключение нового оборудования вычислительного комплекса требует расширения системы мониторинга, но существующая система Nagios 3.5 не поддерживала добавление узлов распределения нагрузки, что позволило бы распределить задачи ЦП системы мониторинга на дополнительные узлы.

Кроме того, обнаружены и менее критичные проблемы, связанные с моральным устареванием системы мониторинга Nagios. К ним можно отнести:

- хранение данных в spool файлах, а не напрямую в базе данных;
- проблематичность создания различных переменных среды внутри конфигурации;
- отсутствие возможности создания карты сети с множественными родительскими связями;
- отсутствие возможности создания отказоустойчивого и распределенного кластера.

В поисках решения данных проблем были рассмотрены варианты миграции на различные программные пакеты, которые не будут иметь данных недостатков. Система визуализации Nagvis [4] и различные плагины сбора данных оказались совместимыми только с проектами, основанными на Icinga [5] и Nagios, поэтому рассмотрены возможности применения этих систем.

Выбор был сделан в пользу системы Icinga2, так как в ней осуществлялась поддержка переменных среды и создание различных кластерных решений.

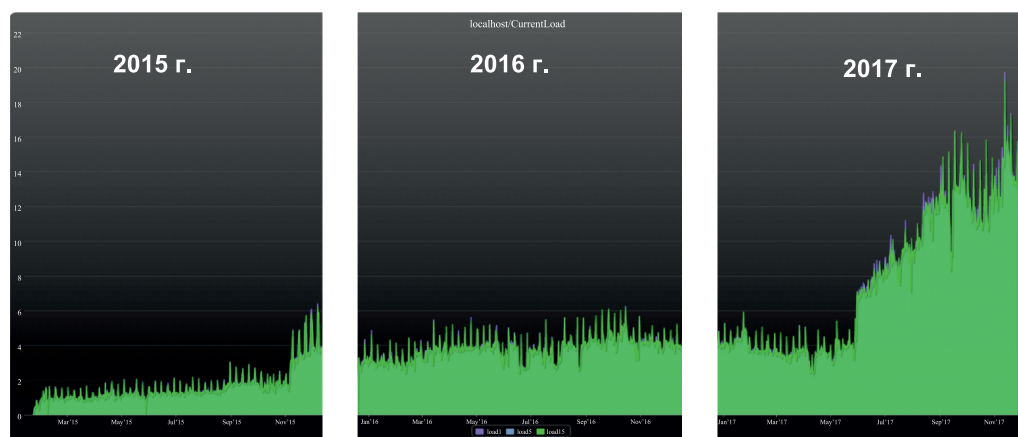


Рис. 2. Исходная загрузка ЦП сервера системы мониторинга

#### 4. ПРОТОТИП СИСТЕМЫ МОНИТОРИНГА

Icinga2 является дальнейшим развитием проекта Icinga. В нем реализована поддержка создания распределенного кластера. Данный подход позволяет значительно увеличить производительность системы мониторинга за счет наращивания дополнительных узлов распределения нагрузки.

На основе проекта Icinga2 проведена разработка прототипа кластерной системы мониторинга параллельно с работами по добавлению новых плагинов сбора данных в функционирующую систему мониторинга.

При разработке прототипа были поставлены задачи:

- создания рабочей конфигурации кластерной системы мониторинга и скриптов миграции конфигурации;
- тестирования совместимости дополнений (Nagvis, rnr4nagios [6], ngrpe) с Icinga2;
- доработки конфигурации Icingaweb2 для нужд МИВК;
- подготовки начальной версии кластерной системы мониторинга для установки ее на основной сервер.

Основная концепция построения кластера — распределение различных функций системы мониторинга между серверами. Построение кластера на базе Icinga2 предполагает наличие управляющего сервера (сервер хранения и обработки данных) и спутников (узлов распределения нагрузки) (рис. 3).

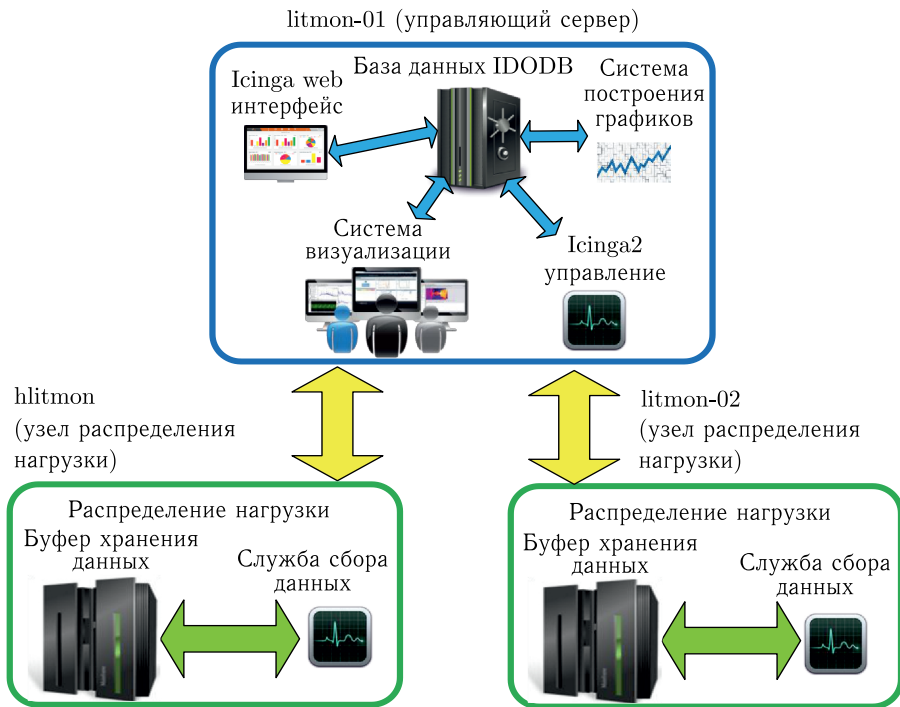


Рис. 3. Общая схема построения кластера системы мониторинга

Задачи управляющего сервера:

- хранение полученных данных;
- визуализация данных;
- построение графиков;
- управление системой мониторинга;
- сбор данных (опционально).

В задачи же сателлитов входит исключительно сбор данных.

## 5. ПОСТРОЕНИЕ КЛАСТЕРНОЙ СИСТЕМЫ МОНИТОРИНГА

В соответствии с поставленными задачами был развернут сервер, на котором установлено программное обеспечение Icinga2. Формат опций конфигурационных файлов Icinga2 отличается от формата Nagios, поэтому были созданы специальные скрипты миграции конфигурации с Nagios на Icinga2. Этот механизм позволил решить две проблемы: перенести актуальную конфигурацию Nagios на сервер с Icinga2; упростить конфигурацию Icinga2 в целом.

После начальной настройки Icinga2 и переноса конфигурации была проведена настройка всех дополнений системы мониторинга:

- системы отображения графиков;
- системы визуализации;
- системы оповещения;
- плагинов сбора данных.

Кластерная система мониторинга предусматривает высокий уровень масштабируемости. На начальном этапе было принято решение о включении в систему одного узла управления и одного узла распределения нагрузки. В перспективе это позволит проводить наращивание до необходимого количества узлов распределения с учетом возрастающей нагрузки.

Сервер распределения нагрузки базируется на Icinga2 в режиме сателлита (см. рис. 3). С помощью плагина `check_pgrep` он запускает сбор активных метрик с различного оборудования. Данный плагин подключается к службе `pgrep`, которая установлена на различных узлах МИВК, и считывает с них определенные данные. Например, с дискового массива берутся данные о состоянии жестких дисков. Если узел не поддерживает установку `pgrep`, то система мониторинга считывает данные специально написанными скриптами, которые, в свою очередь, используют специальные протоколы, такие как `snmp` и `ipmi`.

## 6. СРАВНЕНИЕ ПРОИЗВОДИТЕЛЬНОСТИ

Вопрос сравнения производительности систем мониторинга достаточно неоднозначен. Это связано с тем, что в системах применяются различные способы сбора данных, что в итоге может предоставить недостоверную информацию о производительности. Для того чтобы оценить производительность системы мониторинга, необходимо провести сравнение всего комплекса программных пакетов.

Так как при переходе с одной системы мониторинга на другую поменялось фактически только ядро (рис. 4), но остались дополнения, производительность которых не изменялась, сравнение является наиболее корректным.

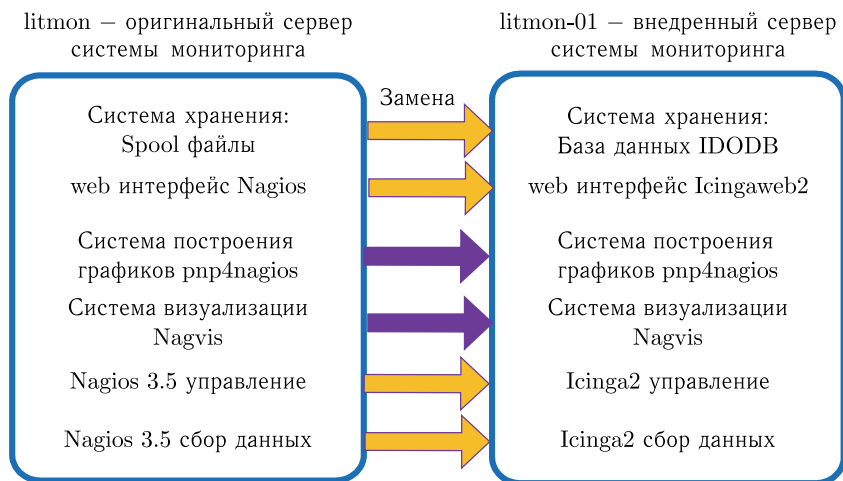


Рис. 4. Программные изменения при переходе с оригинальной системы мониторинга

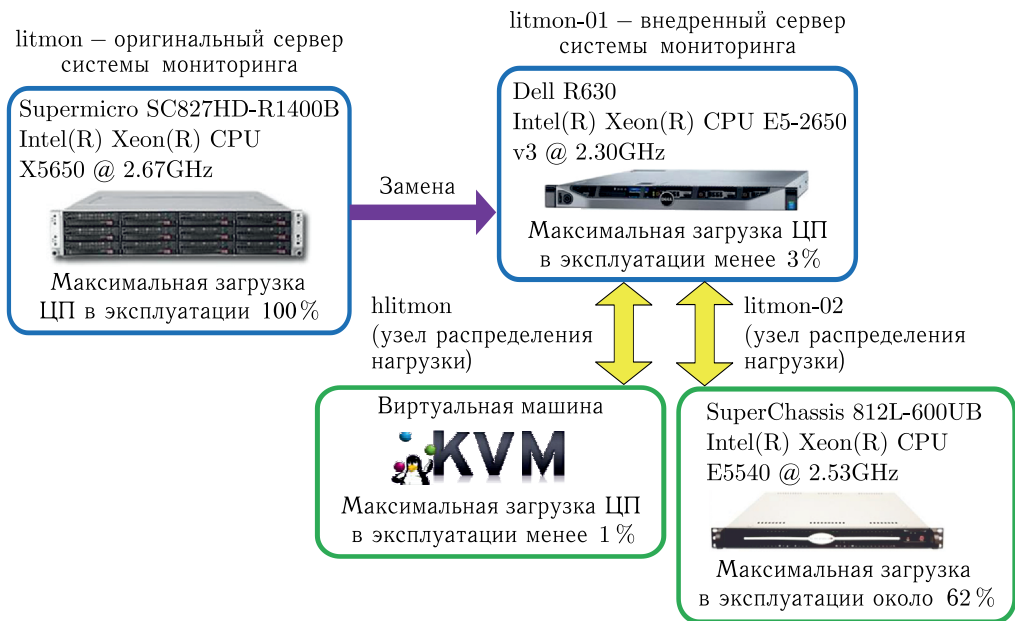


Рис. 5. Аппаратные изменения при переходе с оригинальной на новую систему мониторинга

Ранее все операции по сбору и обработке данных обеспечивал один сервер litmon, теперь в систему мониторинга МИВК включены три узла: litmon-01 — центральный сервер управления и два узла распределения нагрузки — litmon-02, hlitmon.

После конфигурирования системы серверов litmon-01 и litmon-02 в единую структуру был произведен перенос конфигурации с litmon на litmon-01 (рис. 5). В резуль-

тате этого загрузка на litmon значительно сократилась (рис. 6, а). Однако загрузка litmon-01 увеличилась незначительно.

В операционных системах загрузка каждого ядра измеряется в единицах, где полная загрузка соответствует 1 (100% загрузки). Соответственно, если сложить показатели загрузки всех ядер, можно получить среднюю загрузку сервера.

Был произведен переход с системы мониторинга, базирующейся на litmon, на систему, базирующуюся на сервере litmon-01. До перехода на сервер litmon-01 загрузка litmon была более чем 22 ядра (рис. 6, а). После изменения конфигурации достаточно только двух ядер ЦП (рис. 6, б). Это связано с тем, что загрузка с litmon перенесена

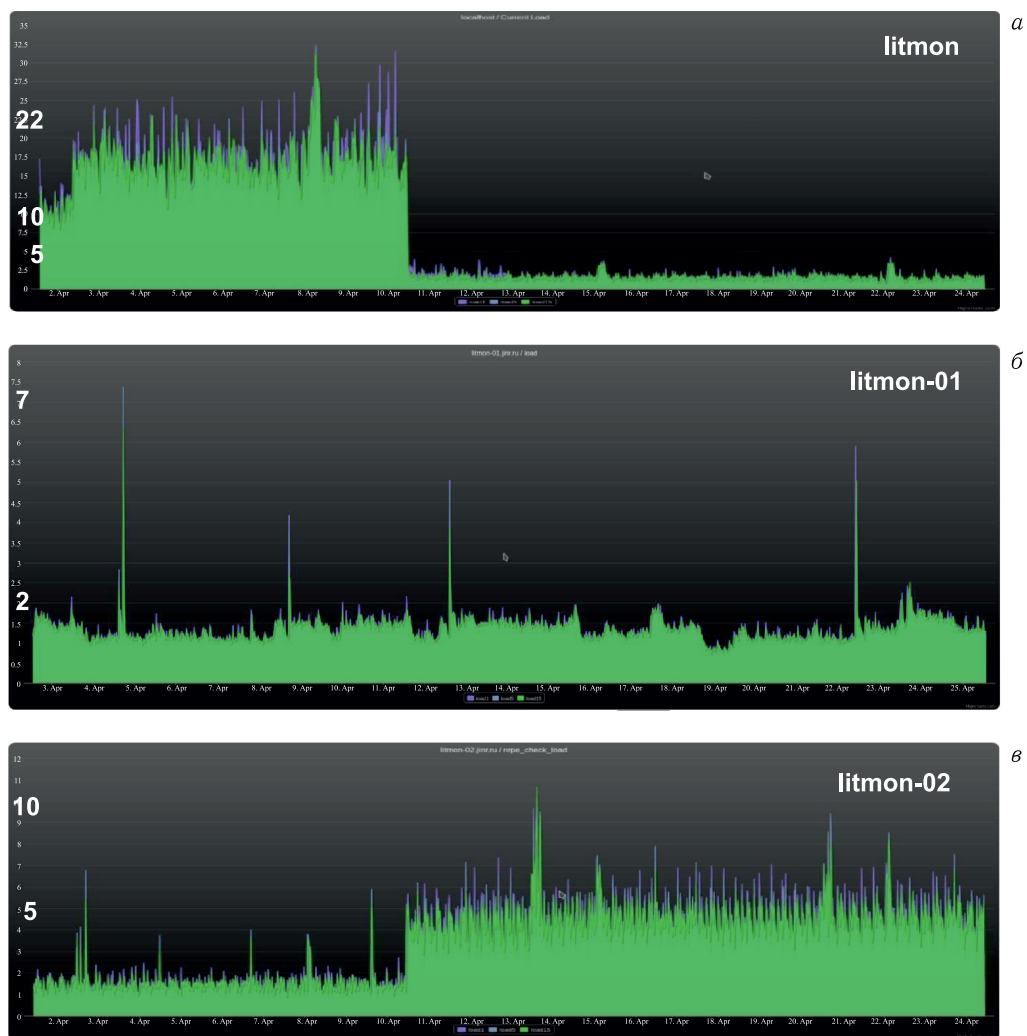


Рис. 6. Загрузка ЦП оригинальной системы мониторинга (а), новой системы мониторинга (б), узла распределения нагрузки новой системы мониторинга (в) (ось абсцисс — производительность ЦП в ядрах, ось ординат — временной промежуток с 02.04.2019 по 24.04.2019)

на сервер litmon-02 (рис. 6, в). Благодаря этому бэкэнд системы мониторинга хранится в базе данных mysql, а не в spool-файлах. Запросы к базе данных кэшируются, что уменьшает загрузку центрального процессора (см. рис. 5).

## 7. ВЫВОДЫ

Зависимость между загрузками litmon-02 и litmon-01 тщательно отслеживается и отображается в системе визуализации Nagvis. Можно с уверенностью сказать, что задачи, которые выполнял Nagios 3.5 с уровнем загрузки более чем 22 ядра, может выполнять более слабый сервер с загрузкой около 10 ядер.

При достижении максимальной загрузки узла распределения производится подключение следующего узла. Он принимает на себя дополнительную нагрузку. Таким образом, это позволяет осуществить наращивание узлов системы мониторинга при увеличении подключаемого оборудования МИВК.

Сравнивая загрузку центрального узла litmon-01 и litmon, можно говорить о выигрыше производительности.

Создана и введена в эксплуатацию кластерная система мониторинга на базе Icinga2 в МИВК.

## СПИСОК ЛИТЕРАТУРЫ

1. *Baginyan A., Balashov N., Baranov A., Belov S., Belyakov D., Butenko Yu., Dolbilov A., Golunov A., Kadochnikov I., Kashunin I., Korenkov V., Kutovskiy N., Mayorov A., Mitsyn V., Pelevenyuk I., Semenov R., Strizh T., Trofimov V., Vala M.* Multi-Level Monitoring System for Multifunctional Information and Computing Complex at JINR // CEUR Workshop Proc. (CEUR-WS.org). Aachen Univ.; RWTH, 2016.
2. *Кашунин И. А. и др.* Система мониторинга многофункционального информационно-вычислительного комплекса // CEUR Workshop Proc. (CEUR-WS.org). Aachen Univ.; RWTH, 2017. V. 1787. P. 256.
3. Nagios. <https://www.nagios.org>
4. NagVis. [www.nagvis.org](http://www.nagvis.org)
5. Icinga. <https://www.icinga.org>
6. pnp4nagios. <http://docs.pnp4nagios.org/start>

Получено 26 ноября 2019 г.