

Система мониторинга многофункционального информационно-вычислительного комплекса

**И.А. Кашунин^а, А.Г. Долбилов, А. О. Голунов, В.В. Кореньков,
В.В. Мицын, Т.А. Стриж**

Объединенный институт ядерных исследований,
141980, г. Дубна Московской области, ул. Жолио-Кюри 6

E-mail: ^а miramir@jinr.ru

Система мониторинга Многофункционального Информационно-вычислительного Комплекса (МИВК) Лаборатории Информационных Технологий (ЛИТ) ОИЯИ направлена на оптимизацию его работы и увеличения уровня отказоустойчивости и надёжности вплоть до 100%.

Система мониторинга постоянно улучшается таким образом, чтобы включить в себя всё существующее в вычислительном комплексе программное и аппаратное обеспечение.

Наличие системы мониторинга позволяет предсказывать системные сбои на ранних этапах и оперативно решать их.

Удобный интерфейс обеспечивает пользователю своевременные оповещения, а так же включает в себя сохранение истории, на базе которой можно предсказывать отказы в будущем.

Всё это делает JINR-LCG2 Tier-2 и JINR-CMS Tier-1 сайты одними из самых эффективных в WLCG.

Ключевые слова: CMS, Tier, МІСС, система мониторинга.

© 2016 Иван.А. Кашунин, Андрей .Г. Долбилов, Алексей. О. Голунов, Владимир.В. Кореньков,
Валерий.В. Мицын, Татьяна.А. Стриж

1. Введение

Многофункциональный информационно-вычислительный комплекс (МИВК) в Лаборатории информационных технологий Объединённый институт ядерных исследований является многокомпонентной информационно-вычислительной инфраструктурой, непрерывное функционирование всех элементов которой на должном уровне является обязательным условием выполнения Институтом своих основных функций. Поддержка этой инфраструктуры в рабочем состоянии является одной из важнейших задач Лаборатории информационных технологий. Одними из базовых компонент комплекса являются грид сайт уровня Tier-1 (JINR-CMS Tier-1) [Astakhov, 2016] для эксперимента CMS (ЦЕРН) [CMS] и грид сайт уровня Tier-2 (JINR-LCG2 Tier-2), входящие в глобальную грид-инфраструктуру WLCG (Worldwide LHC Computing Grid) [WLCG]. Сайт JINR-LCG2 Tier-2 обеспечивает обработку данных всех четырех экспериментов на LHC (ATLAS, Alice, CMS, LHCb) и других масштабных экспериментов, использующих грид-среду и сотрудничающих с физическими группами в ОИЯИ и интегрирован с Центральным информационно-вычислительным кластером [Korenkov, 2015] для поддержки пользователей Лабораторий ОИЯИ и стран-участниц, не использующих грид-среду.

С инженерной точки зрения эти компоненты представляют собой различное оборудование и программное обеспечение, для штатной работы которого требуется наличие отказоустойчивой системы мониторинга. Среди основных групп этого оборудования можно выделить вычислительные сервера, дисковые массивы, коммутирующее сетевое оборудование, ленточный робот, а так же различные системы охлаждения и питания. Среди же программных продуктов это файловая система dCache [dCache], программное обеспечение для организации передачи данных PhEDEx [PhEDEx], система для организации вычислительных кластеров Torque [Torque].

Система мониторинга— это комплекс программного обеспечения, который должен позволять обслуживающему персоналу отслеживать состояние многофункционального информационно-вычислительного комплекса и своевременно реагировать на различные сбои.

Вычислительные компоненты МИВК имеют свои средства для получения сведений с датчиков состояния оборудования в реальном режиме времени. Для слежения за большим количеством датчиков обслуживающему персоналу требуется проверять их состояния на конкретных устройствах, что приводит к большим затратам времени и не позволяет оперативно реагировать на отказ работы оборудования.

При эксплуатации оборудования важно знать не только время когда произошёл сбой, но и отслеживать параметры оборудования за все время его работы, чтобы иметь возможность графического отображения процесса работы оборудования. Таким образом необходимо иметь систему накопления данных и системы построения графиков.

Учитывая особенности работы МИВК, были определены основные требования к системе мониторинга:

- универсальность;
- наличие простого и удобного интерфейса;
- наличие системы построения графиков и ведения истории;
- наличие системы оповещений;
- наличие системы визуализации данных;
- возможность включения в систему мониторинга нового оборудования;
- возможность разработки новых плагинов для сбора данных с датчиков;
- организация системы аутентификации;
- интеграция с Kerberos;
- модульная структура;

- настройка расширений.

Для создаваемой системы ключевой характеристикой является модульность. Это свойство позволяет более гибко настраивать систему и реализовать её расширение. Наиболее подходящими системами являются Nagios [Nagios], Icinga [Icinga], Icinga2 [Icinga2]. В таблице 1 приведены сравнительные характеристики некоторых распространенных систем мониторинга.

Разработка системы мониторинга проходила в 2014 году и на тот момент все плагины, которые необходимо было применить, в полной мере, поддерживал только Nagios. Среди них: NagVis [NagVis], check_mkLivestatus [mkLivestatus], pnp4nagios [pnp4nagios].

На данный момент проекты Icinga и Icinga2 так же получили их полную поддержку.

Таблица 1. Анализ систем мониторинга.

	Расширяемость за счёт плагинов	Интеграция с Kerberos	Модульность	Универсальность
Nagios [Nagios]	Да	Да	Да	Да
Ganglia [Ganglia]	Да	Нет	Нет	Система мониторинга для кластеров
Zabbix [Zabbix]	Да	Да	Нет	Да
Icinga [Icinga]	Да	Да	Да	Да
Icinga2 [Icinga2]	Да	Да	Да	Да

На базе Nagios основываются многие системы мониторинга. Это позволяет в кратчайшие сроки заменить ядро системы мониторинга оперативно без написания новых плагинов. Следует отметить, что в перспективе возможна миграция на проект Icinga2, который обеспечивает более лучшую производительность.

2. Особенности построения системы мониторинга

Основное предназначение системы мониторинга — это информативное представление данных системным администраторам и обслуживающему персоналу.

Система мониторинга собирает данные по состоянию оборудования опрашивая его специальными скриптами через определённые промежутки времени. Полученная информация сохраняется в файлах и базе данных. Далее с помощью специального интерфейса пользователя организуется визуализация данных в удобном для анализа виде. При определенных критических условиях срабатывает система оповещения. На рисунке 1 приведена логическая схема системы мониторинга.

Пакет Nagios, как модульная система, имеет множество расширений для увеличения своих возможностей. В базовом варианте она позволяет только отслеживать состояние узлов вычислительной инфраструктуры. Для расширения её возможностей была установлена система визуализации NagVis, система построения графиков pnp4nagios и система отображения графиков Nagios Highchart.



Рис. 1. Логическая схема системы мониторинга

Организацию системы мониторинга можно представить в виде 3-х уровней (рис. 2).

На первом уровне находятся датчики оборудования, передающие информацию по своим протоколам.

На втором уровне данные от оборудования преобразуются в формат, требуемый системой Nagios. Для этого были написаны специальные плагины. В нештатных ситуациях срабатывает система оповещения. Данные системы мониторинга хранятся в специальных файлах и базе данных.

На третьем уровне происходит визуализация данных в виде таблиц состояний, информационных панелей и графиков. По запросу пользователя система предоставляет определённый тип данных.

При настройке Nagios имеются возможности включить дополнительный функционал, необходимый для различных групп пользователей с определёнными правами доступа.

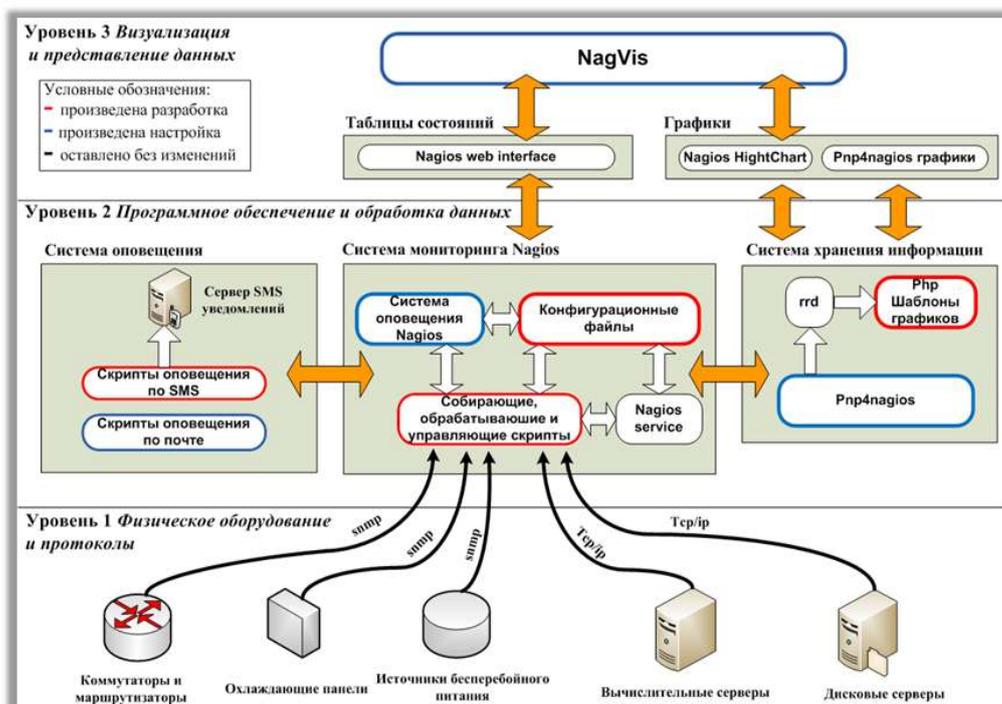


Рис. 2. Структурная схема системы мониторинга

При установке системы мониторинга по умолчанию организовано оповещение по электронной почте. Однако существует возможность добавлять свои механизмы оповещения, например по SMS. Для этого потребовалось написать собственный плагин, алгоритм работы которого представлен на рисунке 3.



Рис. 3. Алгоритм работы SMS оповещения

В процессе работы вычислительного комплекса возможны включения дополнительных функций подтверждения ситуаций, решение которых возможно в перспективе. Эти функции позволяют отключать оповещения для определённого узла и информировать пользователя о данной ситуации. Система мониторинга обеспечивает поддержку функций подтверждения проблемы и временного отключения узла.

3. Эксплуатация системы мониторинга

Система мониторинга позволяет в режиме реального времени наблюдать как за состоянием всего комплекса в целом, так и за состоянием каждой единицы оборудования в отдельности. При настройке системы в таблицу статусов были внесены основные группы оборудования. К ним относятся:

- вычислительные узлы;
- дисковые массивы;
- система климат контроля;
- система бесперебойного питания;
- ленточная библиотека.

Для получения детальной информации о работе группы в целом и конкретного оборудования необходимо активировать соответствующую ячейку таблицы (рис 4). Состояния различных элементов вычислительного комплекса отображаются соответствующими цветами:

- красный имеют проблемы;
- жёлтый предупреждение;
- зелёный проблем нет.

Map Index					
CICC	❌	Cooling_doors	✅	Tier-1_WH	✅
Tier-1_tape_library	✅	Tier-1_therometer	✅	Tier-2_Dashboard	🌐
Tier-2_network	✅	Tier-2_network_max	✅	Tier-2_raid	🌐
dCache	❌	other_servers	🟡	test	🌐

Рис. 4. Таблица статусов системы мониторинга

В системе мониторинга предусмотрен вывод информации по состоянию инфраструктур Tier-2 или Tier-1 компонент на информационные дисплеи. Это позволяет очень быстро оценить текущие проблемы и оперативно принять решение. Расширенную информацию по каждой из компонент можно получить, нажав соответствующую иконку в таблице статусов. На рисунке 5 приведен вид информационного дисплея компоненты Tier-2. Цветом выделены группы однотипного оборудования (вычислительные сервера и дисковые массивы, вентиляционные панели и т.д.)

Система мониторинга является программным продуктом, который используется персоналом, ответственным за определённые типы оборудования: операторами системы климат-контроля, дежурными операторами ЦИВК, операторами Tier-1, системными администраторами, инженерами, обслуживающими локальную сеть вычислительного комплекса.

4. Заключение

Неотъемлемой частью для обеспечения непрерывного и качественного функционирования любого вычислительного комплекса является эффективный мониторинг всех его составляющих.

В ЛИТ ОИЯИ введена в эксплуатацию система мониторинга, являющаяся расширением программного продукта Nagios и позволяющая следить за состоянием различного оборудования, входящего в состав вычислительного центра в реальном режиме времени. В системе мониторинга реализован модульный подход, что позволяет гибко настраивать различные компоненты. Система мониторинга состоит из следующих компонентов:

- **подсистема сбора и обработки данных** реализуется средствами программного обеспечения Nagios и обеспечивает запуск специальных программ плагинов, которые собирают и обрабатывают данные как от оборудования и программного обеспечения, и передают их в систему мониторинга;

подсистема хранения данных основана на различных базах данных (rrd, mysql) и на файлах, в которых хранятся данные системы мониторинга;

- **подсистема оповещения** – для обеспечения своевременного оповещения операторов используется система оповещения по почте и посредством SMS;

- **подсистема визуализации** – адаптирует систему мониторинга для вывода информации на видеостену в виде специальных информационных экранов.

Для системы мониторинга:

разработаны плагины для сбора и обработки данных с оборудования;

- написаны конфигурационные файлы, позволяющие собирать в единую систему данные об оборудовании;

написан плагин, позволяющий организовать оповещение через SMS сообщения;

разработаны шаблоны визуализации графиков;

организована система оперативного представления данных о Tier-1 и Tier-2 в режиме реального времени на экране.



Рис. 5. Tier-2 Dashboard

Список литературы

- Korenkov V., Strizh T., Adam Gh., Podgainy D. The development of distributed computing technologies and BigData in LIT-JINR Proceedings of the 8-th ROLCG 2015 Conference (“Grid, Cloud and High Performance Computing in Science”, 28th - 30th October, Cluj-Napoca, Romania, ISBN: 978-606-737-039-3
- Astakhov N.S., Baginyan A.S., Belov S.D. et al. JINR Tier-1 centre for the CMS Experiment at LHC. Particles and Nuclei, v.13,no 5, pp.1103-1107, 2016
- WLCG (The Worldwide LHC Computing Grid): <http://wlcg.web.cern.ch/LCG>
- CMS (Compact Muon Solenoid): <http://cms.web.cern.ch/>
- dCache: <https://www.dcache.org/>
- PhEDEx (Physics Experiment Data Export): <https://cmsweb.cern.ch/phedex/>
- Torque (Torque Resource Manager): <http://www.adaptivecomputing.com/products/open-source/torque/>
- Icinga (Icinga – Open Source Monitoring): <https://www.icinga.org/>
- Ganglia (Ganglia Monitoring System): ganglia.info/
- Zabbix (Zabbix Monitoring System): www.zabbix.com/ru/
- Icinga2 (Icinga Open Source Monitoring): <https://www.icinga.org/products/icinga-2/>
- Nagios (Nagios - Monitoring System): <https://www.nagios.org/>
- Check_mkLivestatus (Nagios plugin): https://mathias-kettner.de/checkmk_livestatus.html
- Pnp4nagios (Broker Module): <https://docs.pnp4nagios.org/>
- NagVis (visualization system): www.nagvis.org/

The monitoring system of Multifunctional Information and Computing Complex

**I.A. Kashunin^a, A.G. Dolbilov, A.O. Golunov,
V.V. Korenkov, V.V. Mitsyn, T.A. Strizh**

Joint Institute for Nuclear Research, 6 Joliot-Curie, Dubna, 141890, Russia

E-mail: ^a miramir@jinr.ru

The development of the monitoring system of the Multifunctional Information and Computing Complex (MICC) of JINR is aimed at optimizing its work and at increasing the availability and reliability toward the absolute 100% level limit.

The monitoring is continuously improved such as to encompass all the existing hardware and primary software service modules. Monitoring allows detecting predictable system failures at incipient stages and the provision of beforehand solutions.

Suitable interfaces provide timely notifications, enable history-based predictions concerning the future MICC functioning.

This makes the JINR-LCG2 Tier-2 and the JINR-CMS Tier-1 sites amongst the most efficient ones within the WLCG.

Keywords: CMS, Tier, MICC, monitoring system.

© 2016 Ivan.A. Kashunin, Andey.G. Dolbilov, Alexey. O. Golunov, Vladimir.V. Korenkov,
Valery.V. Mitsyn, Tatyana.A. Strizh