# ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ МОНИТОРИНГА ЗА РАБОТОЙ ВЫЧИСЛИТЕЛЬНЫХ КОМПЛЕКСОВ

**Нигматов Х., Умаров У.А., Турсунбаев Т.Б.**

https://doi.org/10.5281/zenodo.7856477

***Аннотация.*** *Для построения интеллектуальных систем в статье произведен анализ различных систем мониторинга вычислительных комплексов в распределенных компьютерных сетях с иерархической структурой, позволяющие отслеживать статистику и историю вычислений в реальном времени для каждого из наблюдаемых узлов.*

***Ключевые слова:*** *Информационная система, инциденты информационной безопасности, критерии полезности, программное обеспечение, компьютерная сеть.*

Задачей обеспечения надежности и бесперебойности работы вычислительных комплексов в распределенных компьютерных сетях является создание централизованного интеллектуального мониторинга, а также оперативного реагирования на инциденты информационной безопасности, возникающие в работе тех или иных узлов информационно-коммуникационной вычислительной инфраструктуры. Результирующие аналитические сведения, получаемые в рамках функционирования данной мониторинг-системы также являются предпосылкой к проведению компьютерно-технической экспертизы, направленной на выявление источников инцидента информационной безопасности, а также общего проектного управления (включая маркетинговый анализ).

Решение и реализация поставленной задачи осуществляется за счёт:

- внедрения механизма само исцеления сервисов и работоспособности основных программных, системных компонентов, обеспечивающих функционирование программных продуктов (проектов) на исследуемом серверном (и ином сетевом коммуникационном) оборудовании;
- внедрения системы мгновенного оповещений операторов системы мониторинга через мессенджеры и СМС (включая лиц, ответственных за обеспечение технической поддержки информационно-коммуникационной инфраструктуры и обеспечение информационной безопасности в предприятии) об аномальной активности;
- внедрения системы отображения актуальной информации о состоянии жизнедеятельности основных узлов и систем в структуре обеспечения информационно-коммуникационного и вычислительного аппарата.

Основными входными данными являются метрики, на основе которых формируется заключении о состоянии функционирования того или иного модуля (компонента) исследуемого узла в компьютерной сети [1].

Критериями полезности данного решения по разным аспектам являются:

- возможность получать оперативные уведомления об инцидентах информационной безопасности с целью своевременного реагирования и предотвращения возможных последствий инцидента информационной безопасности;
- возможность следить за жизнедеятельностью наиболее важных узлов, обеспечивающих работоспособность автоматизированной системы с целью обеспечения непрерывности производства и принимать решения на основе аналитических сведений;

- возможность проводить общую оценку рентабельности проектов (коммерческих программных продуктов) с учётом расходов на содержание отдельных компонентов программно-аппаратной вычислительной инфраструктуры и менеджмента с целью экономии затрат финансовых средств на содержание неперспективных проектов;

- возможность экономить расходы человеко-часов (в частности, системных администраторов) на устранение типовых сбоев в системной части серверной платформы за счёт механизма-само исцеления;

- возможность получать первичные сведения для проведения расследования инцидента информационной безопасности и анализа состояния информационной безопасности вычислительных платформ при помощи интуитивно понятного интерфейса.

Система не вторгается в целостность установленных программных средств и продуктов, систем, а также не выполняет на серверной платформе никаких операций удалённого управления. Статистическая информации собирается в коллектор – удалённый хост в сети TCP/IP путём передачи сведений по зашифрованному каналу, протоколу SSH.

Рассмотрим некоторые современные системы мониторинга кластеров параллельных и распределенных вычислительных комплексов.

Ganglia — эта масштабируемая распределённая система мониторинга кластеров параллельных и распределённых вычислений и облачных систем с иерархической структурой. Позволяет отслеживать статистику и историю (загруженность процессоров, сети) вычислений в реальном времени для каждого из наблюдаемых узлов.

Система построена по иерархическому принципу для интеграции кластеров. Для мониторинга состояния кластеров и их объединения используется древовидная система, основанная на P2P-соединениях и широковещательных протоколах. Использует такие технологии, как XML для представления данных, XDR для сжатия данных, RRDtool для хранения и визуализации данных. Для отображения страниц статистики используется шаблон затор TemplatePower.

Система пор тирована на широкий спектр операционных систем и процессорных архитектур, известно об её использовании более чем 500 кластерах по всему миру. Существуют сборки для следующих операционных систем: Linux (i386, x86-64, SPARC, DEC Alpha, powerpc, m68k, MIPS, ARM, PA-RISC, S390), FreeBSD, NetBSD, OpenBSD, Dragonfly BSD, Mac OS X, Solaris (SPARC), AIX, IRIX, Tru64, HP-UX и Windows NT/XP/2000/2003/2008. Используется для связи кластеров в университетских кампусах по всему миру и может масштабироваться для обработки кластеров, имеющих до 2000 узлов в своем составе.

Необходимые пакеты для установки Ganglia присутствуют в большинстве репозиториев современных дистрибутивов Linux.

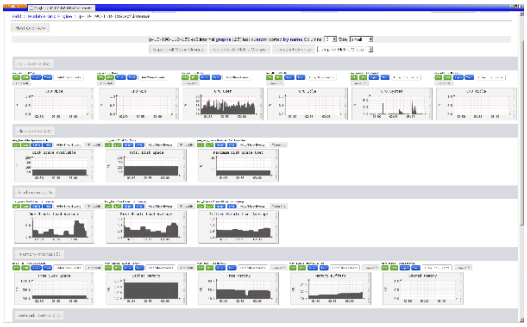

Рис 1. Интерфейс Ganglia

Collectd — это легковесный демон, который собирает данные (каждые 10 сек) об использовании системных ресурсов. Он позволяет собирать данные с нескольких хостов и отправлять их на централизованный сервер для дальнейшего использования (Например, можно использовать influxDB и потом с данной БД, строить графики в grafana). Главное отличие данного ПО, — это в том, что работает как push (а не poll/pull). Это означает что он находится в ожидании и ждет пока другие хосты пришлют ему данные по серверу [2].

Все данные для сбора прописываются вручную. Имеется библиотека плагинов, подключаемых для расширения функционала, разрабатываемых открытым сообществом разработчиков.

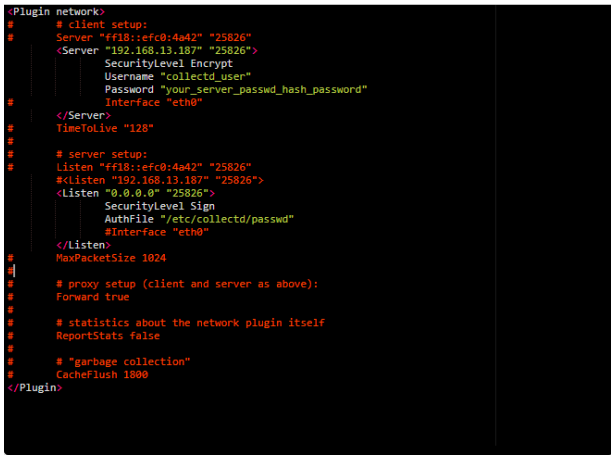


Рис. 2. Библиотека плагинов

Graphite - это бесплатное программное обеспечение с открытым исходным кодом (FOSS), которое отслеживает и графически отображает числовые данные временных рядов, такие как производительность компьютерных систем. Graphite был разработан Orbitz Worldwide, Inc и выпущен как программное обеспечение с открытым исходным кодом в 2008 году.

Graphite собирает, хранит и отображает данные временных рядов в реальном времени.

Инструмент состоит из трех основных компонентов:

Carbon - демон Twisted , который прослушивает данные временных рядов.

Whisper - простая библиотека базы данных для хранения данных временных рядов (по дизайну аналогична RRD ) [2].

Graphite webapp Graphite - веб-приложение Django, которое отображает графики по запросу с использованием библиотеки Cairo.

Графит используется в производстве такими компаниями, как Ford Motor Company, Booking.com, GitHub, Etsy, The Washington Post и Electronic Arts.

Рис 3. Интерфейс Netdata

Netdata – инструмент отслеживания большого количества показателей: статистику использования процессора, потребления памяти, операций ввода-вывода, сети и многого другого, в частности оснащён плагинами отслеживания различных служб, таких как Postfix, Squid, PHP-FPM и другие [3]. В частности:

- Ядро CPU – прерывания, частоты и т.д.
- Память – Общий объем памяти, ОЗУ, своп-файл и использование ядра.
- Дисковый ввод-вывод на диск: пропускная способность, операции, невыполненная работа, использование и т. д.
- Сети – пропускная способность, пакеты, ошибки, падение и т. д.
- Брандмауэр – мониторинг с netfilter/iptables в Linux подключение брандмауэра, событий, ошибок и т. д.
- Процессы – запущенные, заблокированные, активные и др.
- Системные приложения – с деревом процессов для процессора, памяти, подкачки, чтение/запись на диск, threads и т. д.
- Статус Apache и Nginx.
- База данных MySQL – запросы, обновления, блокировки, вопросы и т. д.
- Очередь Сообщений почтового сервера Postfix.
- Мониторинг пропускной способности прокси-сервера Squid и запросов.
- Аппаратные датчики-температуры, напряжение, вентиляторы, мощность, влажность и др.
- SNMP-устройство.

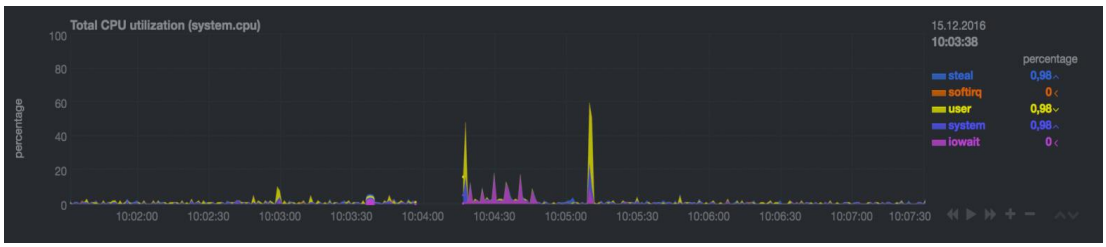


Рис 4. Интерфейс нагрузки на сантальной процессор

InfluxData – комплекс нескольких компонентов позволяющий собирать определённые данные с серверов:

Telegraf - это агент сбора данных, который собирает данные из растущего списка источников и переводит их в формат линейного протокола InfluxDB для хранения в InfluxDB. Расширяемая архитектура Telegraf позволяет создавать плагины, которые

извлекают данные (плагины ввода) и отправляют данные (плагины вывода) в разные источники и конечные точки и из них.

InfluxDB хранит данные для любого варианта использования, включающего большие объемы данных с отметками времени, включая мониторинг DevOps, данные журналов, метрики приложений, данные датчиков IoT и аналитику в реальном времени. Он предоставляет функциональные возможности, которые позволяют экономить место на вашем компьютере, сохраняя данные в течение определенного периода времени, а затем автоматически понижает дискретизацию или истекает и удаляет ненужные данные из системы.

Chronograf — это пользовательский интерфейс для стека TICK, который предоставляет настраиваемые информационные панели, визуализацию данных и исследование данных. Он также позволяет просматривать задачи Kapacitor и управлять ими.

Kapacitor — это фреймворк для обработки данных, который позволяет обрабатывать данные и действовать с ними по мере их записи в InfluxDB, который включает в себя обнаружение аномалий, создание предупреждений на основе пользовательской логики и выполнение заданий ETL.

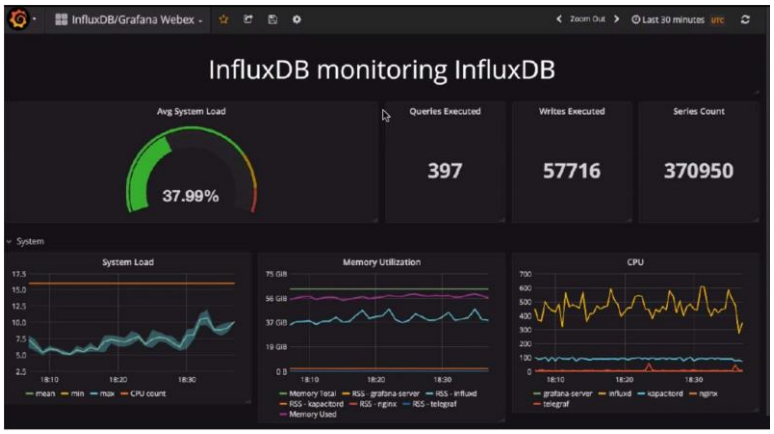


Рис 5. Интерфейс InfluxDB

Network Olympus. Программа работает как служба и имеет веб-интерфейс, что дает гораздо большую гибкость и удобство в работе. Главная особенность – конструктор сценариев, позволяющий отойти от выполнения примитивных проверок, которые не позволяют учитывать те или иные обстоятельства работы устройств. С его помощью можно организовывать схемы мониторинга любой сложности, чтобы точно выявлять проблемы и неполадки, а также автоматизировать процесс их устранения.

В основе сценария лежит сенсор, от которого можно выстраивать логические цепочки, которые в зависимости от успешности проверки будут генерировать разные оповещения и действия, направленные на решение ваших задач. Каждый элемент цепочки может быть отредактирован в любое время и сразу применится для всех устройств, за которыми закреплен сценарий. Вся сетевая активность будет отслеживаться при помощи журнала активности и специальных отчетов.

Кроме вышеприведенного существуют большое количество современных систем, позволяющие проводить мониторинг вычислительных комплексов не зависимо от места его

нахождения в глобальной компьюторной сети Internet, такие как Cacti, Nagios, Icinga, NeDi, ntop, Zabbix, OpenNMS и другие [2,4].

На основе анализа масштабируемых систем мониторинга кластеров параллельных и распределённых вычислений и облачных систем с иерархической структурой, позволяющие отслеживать статистику и историю вычислений в реальном времени для каждого из наблюдаемых узлов можно разработать новую интеллектуальную систему.

**REFERENCES**

1. Нигматов Х. Информационная безопасность и защита информации в сетях телекоммуникации. Учебное пособие. Казахстан. Изд.”ЖЕБЕ”. Чимкент. 2015. 188 стр.
2. Тарасов А. Г Трёхуровневая система мониторинга расширенной функциональности. Хабаровск: Изд-во ДВГУПС, ИПМ ДВО РАН, 2007,
3. Тарасов А. Г. Мониторинг вычислительного кластера с использованием java-технологий // XXX Дальневосточная математическая школа-семинар имени академика Е.В. Золотова: тезисы докладов. - Хабаровск: Изд-во ДВГУПС, ИПМ ДВО РАН, 2005, с. 201
4. Lambert M. Surhone Ganglia (software). — VDM Verlag Dr. Mueller AG & Co. Kg, 2000. — 120 с. — (Betascript). — ISBN 978-6-1319-6802-0.