

Выбор системы обработки данных для мониторинга сервера в реальном времени

О.А. Мазуренко, И.О. Шамренко

Національний університет харчових технологій

В коммерческих предприятиях, которые используют и/или предоставляют услуги хостинга необходимо проводить мониторинг своих серверов в реальном времени. Чтобы произвести такой мониторинг, сетевым администраторам необходимо иметь дело с данными в форме временных рядов (time series): статистика скачиваний/установок, количество пакетов, статистика запросов и прочее. Для хранения и обработки такого количества данных необходимо иметь подходящий и оптимизированный инструмент.

Основной проблемой мониторинга сервера в реальном времени является огромное количество и объем данных, которые необходимо обработать и вывести с минимальной задержкой. Использование таких методов как MySQL или YAWNDB может привести к частым зависаниям и падениям самого сервера. Кроме того, эти решения не смогут обработать такое количество данных за очень короткое время, информация для мониторинга будет выводиться с большими задержками и уже не будут актуальными на момент взвода [2]. Возможным решением для данной проблемы может быть использование баз данных InfluxDB для минимизации времени обработки поступающих данных от сервера, тем самым ускоряя их вывод на монитор.

База данных InfluxDB относительно новый продукт, написанный на языке GO. Первые версии этого продукта были выпущены в октябре 2013 года и с тех пор он начал совершенствоваться [1]. Последняя версия на данный момент 0.9 представила много новинок, которые ожидали пользователи.

В числе преимуществ InfluxDB в первую очередь нужно выделить следующие:

- Вследствие того, что продукт написан на языке Go в нем отсутствуют зависимости.
- Имеет возможность работы в кластерном режиме;
- Имеет в наличии библиотеки для большого числа языков программирования (Python, JavaScript, PHP, Haskell и других);
- SQL-подобный язык запросов, с помощью которого можно производить различные операции с временными рядами (объединение, слияние, разбиение на части);
- удобный графический интерфейс для работы с БД.

InfluxDB позиционируется как база данных для хранения временных рядов, метрик и информации о событиях. Этот инструмент будет прекрасно обрабатывать и быстро выводить большие объемы данных [3].

Минусом для использования этого инструмента, является то, что он может работать только с целочисленными выражениями, вследствие чего некоторая информация не может быть обработана этой базой. Поскольку этот инструмент

решает лишь часть проблемы, необходимо придумать способ доставки необрабатываемых данных до InfluxDB в измененной форме.

Решением данной проблемы может послужить установка на сервер скрипта, который бы смог конвертировать получаемые данные от сервера и округлять значения с точностью до единицы, и посылал эти данных в InfluxDB на обработку.

- При использовании инструмента InfluxDB и дополнения его скриптом для конвертации данных мы получаем следующие:
- Большая скорость обработки данных
- Возможность обработки больших объемов данных.
- Предоставление этих данных в реальном времени для системы мониторинга.
- Возможность обработки и вывода ранее не доступных значений для InfluxDB.

При этом, существует несколько недостатков связанных с вводом дополнительного функционала. Скорость работы InfluxDB может быть уменьшена, в зависимости от скорости работы скрипта. Так же, значения, которые скрипт конвертирует в целые, показывают информацию с точностью до 1.

После ввода дополнительного функционала, можно увидеть полный набор данных, обработанный с помощью InfluxDB. Данные смогут обрабатываться и выводиться на экран с минимальной задержкой, тем самым выводя администратору актуальную информацию, на основе которой он сможет обеспечить более качественный мониторинг сервера для предоставления и\или использования услуг хостинга. Выявленные недостатки в предложенном методе, не являются критичными при использовании инструмента в коммерческих целях для мониторинга в реальном времени, поскольку точность и скорость обработки остаются в пределах нормы.

Литература

1. Документация по GitHub. [Электронный ресурс]. Режим доступа: <https://github.com/influxdb/influxdb> – Назва з екрану.
2. Time Series, метрики и статистика: знакомство с InfluxDB. [Электронный ресурс]. Режим доступа: <http://habrahabr.ru/company/selectel/blog/245515/> – Назва з екрану.
3. My Impressions of InfluxDB. [Электронный ресурс]. Режим доступа: <http://obfuscurity.com/2013/11/My-Impressions-of-InfluxDB> – Назва з екрану.
4. Документация по InfluxDB Docs v0.9. [Электронный ресурс]. Режим доступа: <https://influxdb.com/docs/v0.9/introduction/overview.html> – Назва з екрану.