

# ИСПОЛЬЗОВАНИЕ ELASTIC STACK ДЛЯ СБОРА АНАЛИЗА ЖУРНАЛЬНЫХ ЗАПИСЕЙ В ECOMMERCE СИСТЕМЕ НА БАЗЕ MAGENTO

**Копылов Денис Валерьевич**

Аспирант, Институт компьютерных технологий  
и информационной безопасности Южного  
федерального университета, г. Таганрог  
dv.kopilov@gmail.com

## USING THE ELASTIC STACK TO COLLECT LOG DATA IN A MAGENTO BASED ECOMMERCE SYSTEMS

**D. Kopylov**

*Summary.* This article describes the main ways to organize the collection and storage of journal entries in online stores on the basis of Magento. Their characteristics are analyzed and drawbacks are revealed. The architecture of the system of collecting and storing log records on the basis of advanced software complex Elastic Stack is proposed. The developed software implementation of the proposed solution in the form of the Magento module and the main results achieved with the proposed solution are described.

*Keywords:* Magento, log files, journal entries, analyse, storage, elastic stack, kibana, elasticsearch.

*Аннотация.* В рамках данной статьи рассматриваются основные способы организации сбора и хранения журнальных записей в интернет-магазинах на базе Magento. Анализируются их характеристики и выявляются недостатки. Предлагается архитектура системы сбора и хранения журнальных записей на основе передового программного комплекса Elastic Stack. Описываются разработанная программная реализация предложенного решения в виде модуля Magento и основные результаты, достигнутые с помощью предложенного решения.

*Ключевые слова:* Magento, log файлы, журнальные записи, анализ, хранение, elastic stack, kibana, elasticsearch.

## Введение

**Т**ехнологический и технический прогресс не стоят на месте, и системы интернет торговли (eCommerce системы) развиваются и эволюционируют вместе с ними. Малые и крупные торговые бизнесы все меньше смотрят в сторону разработки собственных систем, и больше стараются использовать готовые платформенные решения. Не малую роль при выборе играет наличие поддержки, обширное сообщество разработчиков, широкий функционал доступный в системе. Одной из таких систем является eCommerce платформа Magento, которая имеет богатую историю и входит в тройку лидеров среди eCommerce систем, а также является лидером среди систем с открытым исходным кодом [1].

С учетом постоянно растущего количества продаж, которые, согласно данным аналитических исследований, в среднем увеличивались на 31% в период с 2011

по 2021 год [2], вырос и объем данных, которыми оперируют данные системы.

На основе вышесказанного можно сделать вывод, что мониторинг таких систем является важной составляющей для поддержания системы в работоспособном состоянии при постоянно растущих нагрузках и объемах информации [3].

Анализ существующих подходов к организации сбора журнальных записей в Magento

Для начала введем следующее определение. Журналирование или логирование — это процесс записи данных или событий происходящих с каким-либо объектом в хронологическом порядке в журналы событий, так же называемые log-файлами [4]. В информационных системах объектами могут считаться приложения. Регулярный анализ журнальных записей позволяет понять, что происходило с приложением в определен-

ный момент времени, выявить предпосылки и последовательность действий для воспроизведения и следования какого-либо инцидента, а также собирать различную статистику работы приложения. Можно представить следующие примеры журналируемых событий:

- ◆ пользовательские действия в системе;
- ◆ внешние запросы к серверу;
- ◆ запросы, формируемые приложением для отправки на сторонний сервер;
- ◆ ошибки, генерируемые приложением в процессе исполнения и т.д.

Основными подходами к организации сбора и хранения журнальной информации в платформе Magento являются:

1. сбор информации в log файлы. По умолчанию, Magento разделяет собираемые журнальные записи на 4 типа [5]:
  1. 1.1. debug log — предназначен для сбора журнальных записей отладочной информации. Стоит отметить, что данный тип логирования активен только в режиме разработки и выключен в production режиме;
  1. 1.2. system log — предназначен для сбора журнальных записей различных событий, генерируемых в процессе работы приложения;
  1. 1.3. exception log — предназначен для сбора журнальных записей о проходящих инцидентах в работе приложения;
  1. 1.4. cron log — предназначен для сбора журнальных записей во время работы встроенного в Magento планировщика задач;
  1. 1.5. db log — предназначен для сбора журнальных записей об использовании базы данных. Данный тип логов доступен только в коммерческой версии Magento;
  1. 1.6. reports — специальный тип журнальных записей, содержащий информацию о причине “падения” Magento и стек вызовов функций.
2. так же разработчики часто прибегают к кастомным решениям в виде сбора и хранения информации в базе данных, для последующего отображения в административной панели.

Своевременный мониторинг и анализ журнальных записей, выявление и решение проблем на их основе являются неотъемлемой частью в работе любого интернет магазина. В то же время данные процессы являются критически важными для крупного бизнеса.

Существующие подходы к организации сбора, хранения и обработки журнальной информации в Magento являются малоэффективными, ввиду следующих факторов:

- ◆ затруднен доступ к log файлам, ввиду необходимости доступа к серверу, на котором запущено приложение;
- ◆ увеличивается время на поиск и анализ логов ввиду представления информации в простом текстовом виде;
- ◆ объем log файлов может достигать нескольких гигабайт, что накладывает определенные ограничения на ПК, на котором будет производиться анализ данных файлов;
- ◆ отсутствие какой-либо классификации журнальной информации, т.к. логирование производится в конец файла за последней существующей записью.

Таким образом можно сделать вывод о необходимости оптимизации и автоматизации данных процессов.

Использование Elastic Stack для организации хранения и анализа журнальных записей

Основываясь на результатах анализа имеющихся решений для сбора и анализа большого объема данных, можно выделить программный комплекс Elastic Stack [6]. Неоспоримыми преимуществами Elastic Stack является бесплатность, открытость исходного кода, а также высокий авторитет в индустрии как популярного решения, сумевшего доказать свою эффективность работы с большими объемами данных. Elastic Stack состоит из трех основных программных компонентов:

1. Elasticsearch — распределенное, масштабируемая поисковая система для полнотекстового поиска и анализа данных. Elasticsearch поддерживает REST API и передачу данных в формате JSON. Преимуществом данного программного комплекса является поиск в режиме реального времени по большим объемам данных — документов, а также поддержка многоуровневости.
2. Logstash — программный ковейер обработки данных, отвечающий за сбор данных из различных источников, обработку, преобразование и отправку данных в систему хранения Elasticsearch. Logstash позволяет нормализовывать информацию, поступающую из различных источников не зависимо от формата и структуры данных.
3. Kibana — программная панель визуализации данных из хранилища Elasticsearch [7]. Данные, полученные из Elasticsearch представляются в виде списков и диаграмм различных видов: линейные, столбчатные, точечные графики, круговые диаграммы и др.

Стоит отметить, что использование Elastic Stack так же позволяет проводить анализ данных из индек-

сов Elasticsearch на основе алгоритмов машинного обучения. Гибкая настройка работы алгоритмов машинного обучения доступна через интерфейс Kibana.

## Результаты

Решение для сбора и обработки большого объема журнальных записей, предложенное в данной статье, было реализовано в виде программного модуля Magento и позволяет достичь следующих результатов:

1. повысить качество хранения журнальных записей за счет использования централизованного хранилища вместо файлового;
2. снизить требования к дисковому пространству серверов;
3. повысить качество и скорость поиска и анализа журнальных записей;
4. повысить эффективность работы ИТ-отдела за счет предоставления доступа к журнальным записям посредством Kibana всему подразделению.

## ЛИТЕРАТУРА

1. Magento Open Source — software that delivers basic eCommerce capabilities to build a unique online store from the ground up. // Adobe URL: <https://business.adobe.com/products/magento/open-source.html> (дата обращения: 13.05.2022).
2. Российский eCommerce между весной и летом 2022 г. // Data Insight URL: [https://datainsight.ru/sites/default/files/Ecommerce\\_in\\_Russia\\_between\\_spring\\_and\\_summer\\_2022.pdf](https://datainsight.ru/sites/default/files/Ecommerce_in_Russia_between_spring_and_summer_2022.pdf) (дата обращения: 13.08.2022).
3. Мониторинг систем: управление на основе данных // URL: <https://www.computerworld.ru/cio/articles/210218-Monitoring-sistem-upravlenie-na-osnove-dannyh> (дата обращения: 14.08.2022).
4. Гадасин Д., Каледина А. Использование современных средств мониторинга для анализа состояния ИТ-системы. Сборник трудов XIV Международной отраслевой научно-технической конференции. 2020. 267 с.
5. View and manage logs // URL: <https://devdocs.magento.com/cloud/project/log-locations.html> (дата обращения: 14.08.2022)
6. Elastic Stack // URL: <https://www.elastic.co/elastic-stack/> (дата обращения: 14.08.2022).
7. Kibana // URL: <https://www.elastic.co/kibana/> (дата обращения: 14.08.2022).

© Копылов Денис Валерьевич ( dv.kopilov@gmail.com ).

Журнал «Современная наука: актуальные проблемы теории и практики»



г. Таганрог