

УДК 004.428.4

РАЗРАБОТКА СИСТЕМЫ ЛОГИРОВАНИЯ ДЛЯ РАСПРЕДЕЛЕННЫХ ПРИЛОЖЕНИЙ

Гуляев Б. С. (Университет ИТМО)

**Научный руководитель – к.т.н., Косяков М.С.
(Университет ИТМО)**

В сложных высокопроизводительных распределённых приложениях система логирования событий играет важную роль. Выбор правильной архитектуры и ее корректная реализация позволяют организовать логирование событий с минимальным влиянием на производительность и масштабируемость приложений и высокой скоростью работы с собранной информацией. В данной работе сформирован ряд требований к системам логирования, рассмотрены примеры существующих систем и основные подходы к их реализации, проанализированы их недостатки в области распределенных вычислений, и описана архитектура системы, разработанной для исправления этих недостатков.

Наиболее распространенным хранилищем для логируемых сообщений в существующих системах являются текстовые файлы или системный журналы. Такой подход имеет ряд недостатков. Во-первых, в текстовом виде метаданные сообщения занимают значительно больше места, чем в двоичном представлении. Во-вторых, при форматировании данных в текст теряется их внутренняя структура, и при необходимости нетривиальной обработки преобразованных данных требуется их обратное распознавание.

Данный недостаток усугубляется необходимостью получения и управления сообщениями удаленных приложений, так как это либо потребует передачи неполностью отсортированных сообщений по медленным каналам связи, либо создаст нагрузку на работающее приложение, делая сбор логируемых сообщений процедурой, значительно влияющей на производительность приложения.

Для исправления этих недостатков необходимо разработать альтернативную архитектуру обработки, хранения и передачи логируемых сообщений, ориентированную на работу в распределенной среде.

В ходе данной работы был сформирован ряд требований к системе логирования, специфичных для распределенных приложений:

- Группирование логируемых сообщений по приложению
- Возможность запроса сообщений удаленного приложения
- Хранение метаданных логируемых сообщений (например, позиции в исходном коде, приложение, время создания) в дополнение к тексту сообщения
- Быстрая фильтрация сообщений по метаданным
- Регулирование уровня сохраняемых на диск сообщений во время работы приложения

Для реализации данных требований в первую очередь необходимо использовать табличный формат хранения логируемых сообщений вместо текстовых. Таким образом, появляется возможность эффективной фильтрации сообщений по метаданным, и, соответственно, повышается скорость обработки.

Для масштабируемости было решено реализовать систему логирования как сервис распределенных приложений. При запросе набора сообщений приложения будет передаваться набор критериев для метаданных сообщений. По этим критериям будет конструироваться запрос к таблице и производиться выборка сообщений для дальнейшей централизованной обработки.

Дополнительно, реализация системы логирования в качестве сервиса позволяет эффективный мониторинг событий по мере их возникновения – поступающие новые события можно дублировать для сохранения на диск и для отправки потоком клиентским утилитами.

Высокий уровень контроля над реализацией системы и наличие метаданных в сообщениях открывает возможность тонкого управления форматом сохраняемых сообщений и критериями сохранения сообщений на диск, в том числе и в момент выполнения, с минимальным влиянием на производительность самого приложения.

В ходе работы были изучены существующие реализации системы логирования, проанализированы их недостатки и сформирован ряд требований к системам логирования для распределенных приложений, разработана и реализована архитектура, удовлетворяющая этим требованиям.

Гуляев Б.С. (автор)

Подпись

Косяков М.С. (научный руководитель)

Подпись