

## ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ НА КУРСОВУЮ РАБОТУ

«Поиск оптимального решения задачи расчета синдромов в RAID N+M»  
студентки 444 группы кафедры системного программирования СПбГУ  
Быковой Юлии Сергеевны

В последнее время растет спрос на системы хранения данных (СХД), способные поддерживать высокий уровень производительности самых современных дисковых устройств. Так как объемы данных, которые требуется хранить с каждым годом возрастают, в СХД используется технология RAID. Она позволяет увеличить объем хранимой информации и, за счет распараллеливания чтения и записи данных, ускоряет процесс работы с ними. Кроме увеличения уровня производительности немаловажным фактором является необходимость увеличения быстродействия алгоритмов работы с RAID массивами.

Целью курсовой работы являлось проведение исследования и нахождение оптимального решения задачи расчета синдромов в поле Галуа размера 28. С помощью кода Рида-Соломона реализуется вычисление синдромов различными методами и также сравнивается их производительность.

Для достижения поставленной цели в ходе работы студенткой были успешно выполнены следующие задачи:

1. проанализированы и реализованы алгоритмы расчета;
2. подготовлена среда испытаний, программа тестов производительности;
3. осуществлена проверка корректности получившихся функций, протестирована их производительность;
4. выполнен анализ полученных результатов, сделан соответствующий вывод.

Разработка алгоритмов велась на языке программирования Си, что привело к следующим положительным результатам:

- ускорение разработки;
- возможность применения оптимизаций компилятора;
- переносимость программ;

Результат измерений выявил наилучший способ расчета синдромов на данный момент. Результаты работы представляют практическую ценность для компании RAIDIX. Курсовая работа Быковой Юлии заслуживает оценки «отлично».

Разработчик исследовательской лаборатории RAIDIX  Маров А. В.

В последнее время растет спрос на высокопроизводительные системы хранения данных.

Например, для высокопроизводительных вычислений используются СХД суммарной производительностью более 1ТБ/с и объемом более 10ПБ.

Так как объёмы данных, которые необходимо хранить, постоянно растут, в СХД используется технология RAID, а для распределенных облачных СХД применяются RAID-подобные технологии. Эти технологии позволяют увеличить возможный объем хранилища и обеспечивают высокую скорость доступа к данным во время процесса чтения-записи благодаря механизму распараллеливания данных.

Однако с увеличением числа накопителей и узлов хранения растет вероятность критичного для доступности данных отказа нескольких элементов, а значит, повышается необходимость в увеличении надежности системы.

В современных компаниях данные являются одним из основных и самых дорогих активов, а в таких сферах, как медицина и финансы, существуют законодательные акты, определяющие требования к хранению цифровых данных.

Одна из технологий, позволяющая высокого уровня доступности данных и надежности их хранения, это технология RAID(n+m), использующая n устройств для хранения информации и m для контрольных сумм.

Задача увеличения производительности восстановления данных является одной из наиболее важных, так как информационные системы должны удовлетворять SLA, и часто для критически важных для бизнеса процессов недопустимо снижение производительности. Также в распределенных системах хранения

применяется чтение неполных страйпов для снижения передаваемого между узлами трафика.

Целью данной курсовой работы являлась реализация восстановления данных на основе рассчитанных контрольных сумм (синдромов) в RAID(n+m) с использованием систематических кодов Рида-Соломона в арифметике поля  $GF(2^8)$ .

Для достижения цели студенткой были выполнены следующие задачи:

1. Рассмотрены различные алгоритмы восстановления данных, проведен их сравнительный анализ и реализован рабочий прототип некоторых из этих алгоритмов.
2. Произведена векторизация вычислений с использованием инструкций SSE для достижения максимальной производительности кода.
3. Осуществлена проверка производительности и правильности работы этих функций.

Разработка модуля велась студенткой на языке C. Разработанный модуль позволит увеличить число контрольных сумм с 3 до 64, что позволит повысить надежность системы, существенно не влияя на избыточность.

Результаты работы представляют большую практическую ценность для компании RAIDIX, а также были успешно представлены студенткой в рамках выступления с докладом на всероссийской научной конференции по проблемам информатики СПИСОК-2014.

Курсовая работа Зайберт Валерии заслуживает оценки «отлично».

Руководитель Исследовательской лаборатории



Платонов Сергей Михайлович