



Министерство науки и высшего образования Российской
Федерации
Федеральное государственное бюджетное
образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА КОМПЬЮТЕРНЫЕ СИСТЕМЫ И СЕТИ (ИУ6)

НАПРАВЛЕНИЕ ПОДГОТОВКИ 09.03.01 Информатика и вычислительная техника

Р Е Ф Е Р А Т

Название: Файловая система ReiserFS

Дисциплина: Операционные системы

Студент

ИУ6-52Б

(Группа)

(Подпись, дата)

С.В. Астахов

(И.О. Фамилия)

Преподаватель

(Подпись, дата)

(И.О. Фамилия)

Москва, 2021

Содержание

Введение.....	3
1. Устройство файловой системы.....	4
1.1 Структура раздела.....	4
1.2 Структура файловой системы.....	4
1.2.1 В-деревья.....	4
1.2.2 Узлы файловой системы и файлы.....	5
1.3 Журнал файловой системы.....	6
1.3.1 Общие сведения.....	6
1.3.2 Структура журнала.....	7
1.4 Блочное перераспределение.....	8
2. Сравнение производительности ReiserFS с другими файловыми системами	9
Заключение.....	11
Список использованных источников.....	12

Введение

ReiserFS - это журналируемая файловая система общего назначения, впервые представленная в 2001 году, изначально разработанная и реализованная командой Namesys во главе с Хансом Райзером. ReiserFS в настоящее время поддерживается в Linux (без поддержки квот) под лицензией GPLv2. Представленная в версии ядра Linux 2.4.1, это была первая файловая система с журналированием, включенная в стандартное ядро.

Namesys вышла из бизнеса в 2008 году после осуждения Райзера за убийство. В настоящее время продукт поддерживается добровольцами как открытый исходный код.

На момент своего появления ReiserFS предлагала функции, которые не были доступны в существующих файловых системах Linux. Одним из примеров является «упаковка хвоста» - схема уменьшения внутренней фрагментации.

Кроме того, отличительными чертами ReiserFS являются: использование B-деревьев, быстрое журналирование, эффективная работа с файлами малого объема и собственная модификация структуры стандартных для UNIX узлов inode файловой системы.

1. Устройство файловой системы

1.1 Структура раздела

Раздел ReiserFS представляет собой набор блоков фиксированного размера. Блоки нумеруются последовательно, начиная с нулевого. Раздел начинается с 64-х килобайт неиспользуемого пространства, оставленного под загрузчики, дисковые метки и прочие служебные надобности. Далее следует суперблок. Суперблок содержит важную информацию о разделе, например размер блока и местоположение корневого узла и узел журнала.

Сразу за суперблоком следует блок, содержащий битовую карту свободного места. Количество блоков, отслеживаемых картой, напрямую зависит от размера блока. Большой раздел может иметь несколько блоков, хранящих битовую карту.

За первым блоком битовой карты должен быть журнал, однако суперблок содержит более точную (для нестандартных случаев) информацию о его местоположении.

1.2 Структура файловой системы

1.2.1 В-деревья

В то время, как большая часть файловых систем, существовавших на момент разработки ReiserFS использовали для реализации директорий линейные списки, Ханс Рейзер принял решение основывать реализацию директорий в ReiserFS на В-деревьях^[1]. Прежде всего это решение было вдохновлено широким применением В-деревьев в СУБД. Пример В-дерева представлен на рисунке 1.

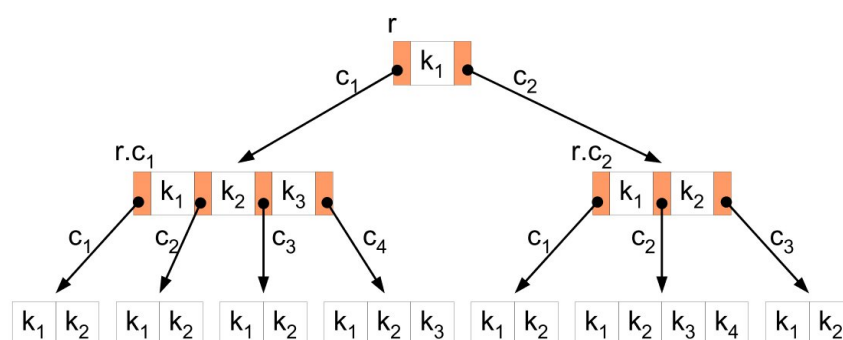


Рисунок 1 - пример В-дерева со степенью 3

В-дерево - сбалансированное, сильно ветвистое дерево. В-дерево должно удовлетворять следующим свойствам:

- Ключи в каждом узле обычно упорядочены для быстрого доступа к ним. Корень содержит от 1 до $2t-1$ ключей. Любой другой узел содержит от $t-1$ до $2t-1$ ключей. Здесь t — параметр дерева, не меньший 2 .
- Каждый узел В-дерева, кроме листьев, можно рассматривать как упорядоченный список, в котором чередуются ключи и указатели на потомков.
- Глубина всех листьев одинакова.

Основные достоинства В-дерева:

- Произвольный доступ к записи реализуется посредством малого количества подопераций (обращения к физическим блокам).
- В среднем достаточно эффективно реализуются операции включения и удаления записей; при этом сохраняется естественный порядок ключей с целью последовательной обработки, а также соответствующий баланс дерева для обеспечения быстрой произвольной выборки.
- Неизменная упорядоченность по ключу обеспечивает возможность эффективной пакетной обработки.

Основной недостаток В-деревьев состоит в отсутствии для них эффективных средств выборки данных (то есть метода обхода дерева), упорядоченных по отличному от выбранного ключа.

1.2.2 Узлы файловой системы и файлы

В ReiserFS вместо традиционных для UNIX структур inode используются собственные узлы трех типов: внутренние, форматированные и неформатированные^[2].

На рисунке 2 эти три типа узлов выделены соответственно красным, желтым и зеленым цветами.

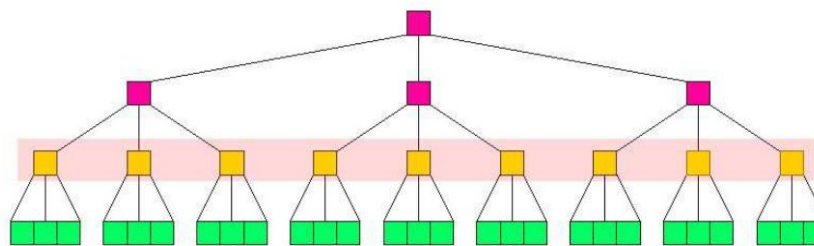


Рисунок 2 - узлы файловой системы

Блок внутреннего узла состоит из заголовка блока, ключей и указателей на узлы-потомки. В начале узла располагается заголовок блока, затем все ключи, отсортированные по значению, далее – указатели на узлы-потомки. Именно эти узлы используются для поиска элемента по ключу в дереве.

Форматированные узлы содержат четыре типа элементов (т.н. «items») с уникальными ключами. Типы элементов неформатированного узла:

- Прямые - содержат «хвосты» файлов
- Косвенные (непрямые) - указатели на неформатированные узлы
- Директории
- Статистические данные

Неформатированные узлы не содержат указателей, имеют размер блока и хранят тело файлов, размер которых больше размера блока.

Хранение файлов осуществляется за счет форматированных и неформатированных узлов. Форматированные узлы хранят «хвост» файла и указатели на неформатированные узлы. Неформатированные узлы хранят непосредственное содержимое файла.

1.3 Журнал файловой системы

1.3.1 Общие сведения

Журналируемая файловая система — файловая система, в которой осуществляется ведение журнала, хранящего список изменений и, в той или иной степени, помогающего сохранить целостность файловой системы при сбоях.

Журналируемая файловая система сохраняет список изменений, которые она будет проводить с файловой системой, перед фактическим их осуществлением. Как только изменения файловой системы внесены в журнал, она применяет эти изменения к файлам или метаданным, а затем удаляет эти записи из журнала. При перезагрузке компьютера программа монтирования может гарантировать целостность журналируемой файловой системы простой проверкой лог-файла на наличие ожидаемых, но не произведённых изменений и последующей записью их в файловую систему^[3].

ReiserFS использует журналирование, напоминающее упорядоченный режим журналирования в ext3, при котором запись данных в файл производится гарантированно до записи информации об изменении этого файла^[4, 5]. Такая схема журналирования является компромиссом между ненадежным журналированием в режиме обратной связи (журналируются только метаданные) и медленным полным журналированием (журналируются как метаданные, так и данные).

Использование журнала позволяет избежать длительной проверки целостности ФС (fsck) при, например, непредвиденном отключении питания.

1.3.2 Структура журнала

Журнал в ReiserFS имеет фиксированный размер: для реализации в Linux 2.4.x – это 8192 блоков + 1 блок для заголовка журнала. Журнал состоит из транзакций переменной длины и заголовка журнала. Он начинается списком транзакций, а заканчивается заголовком. Транзакция охватывает по крайней мере 3 дисковых блока и заголовок журнала, занимающий в точности один блок. Журнал реализован в виде кольцевого буфера^[4].

В отличие от обычного упорядоченного журналирования, когда журналируются только метаданные, ReiserFS журналирует и некоторые дисковые блоки, т.к. они должны появиться в файловой системе лишь после успешной модификации метаданных, и записывать их прямо на диск до завершения транзакции нельзя.

1.4 Блочное перераспределение

На момент своего появления ReiserFS была единственной файловой системой, реализовавшей механизм «упаковки хвоста» (англ. «tail packing») - механизм, позволяющий бороться с внутренней фрагментацией^[6].

Суть внутренней фрагментации показана на рисунке 3 и заключается в том, что при записи файла на диск остается не полностью занятый блок, что приводит к неэффективному использованию дискового пространства. Особенно это заметно при большом размере блоков.

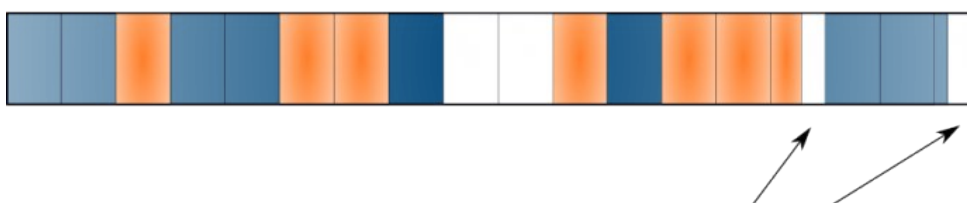


Рисунок 3 - внутренняя фрагментация данных

Предложенный ReiserFS механизм позволяет избежать этой проблемы, упаковывая «хвосты» нескольких файлов в общий хвостовой блок.

Однако, такой подход увеличивает внешнюю фрагментацию данных^[7], как это показано на рисунке 4. Внешняя фрагментация данных приводит к значительному замедлению чтения файлов из-за физической удаленности их частей на диске, что особенно заметно в случае HDD-накопителей.

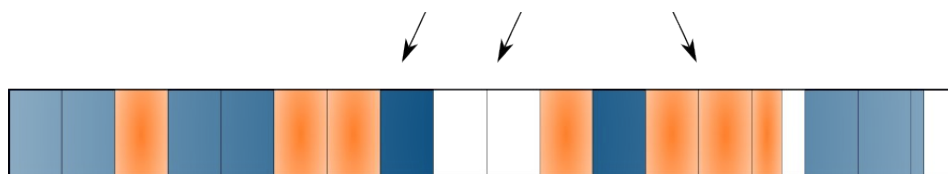


Рисунок 4 - внешняя фрагментация данных

Тем не менее, негативный эффект от внешней фрагментации может быть уменьшен, если «хвосты» хранятся близко к остальным блокам файла, а операционная система использует механизм опережающего чтения.

2. Сравнение производительности ReiserFS с другими файловыми системами

Опираясь на статью сотрудников университета MMU приведем и прокомментируем результаты тестов производительности ReiserFS в сравнении с EXT3, EXT4 и JFS^[8]. Для оценки производительности преимущественно была использована утилита IOZone, определяющая скорость операций над файлами различного размера.

Как видно из рисунка 5, ReiserFS обеспечивает самую высокую скорость чтения для файлов малого размера, но уступает JFS при среднем размере файла. При большом объеме файлов производительность рассматриваемых ФС примерно одинакова.

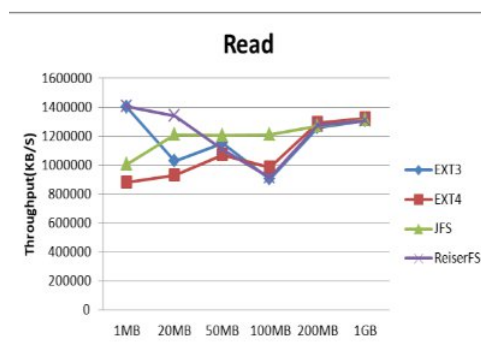


Рисунок 5 - скорость чтения файлов в различных ФС

Исходя из рисунка 6, можно заключить, что ReiserFS имеет самую низкую скорость записи из всех рассматриваемых файловых систем.

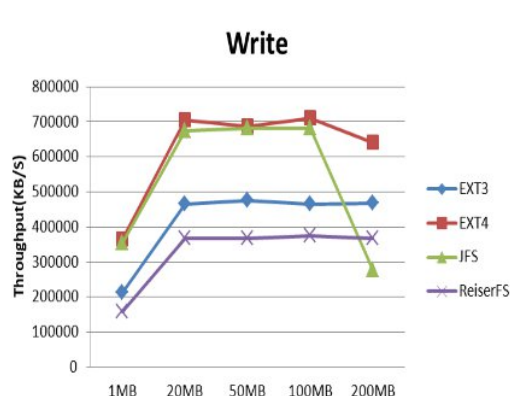


Рисунок 6 - скорость записи файлов в различных ФС

Рисунок 7 иллюстрирует, что EXT4 значительно уступает остальным рассматриваемым ФС по скорости повторного чтения. Скорость ReiserFS в данном случае значительно не отличается.

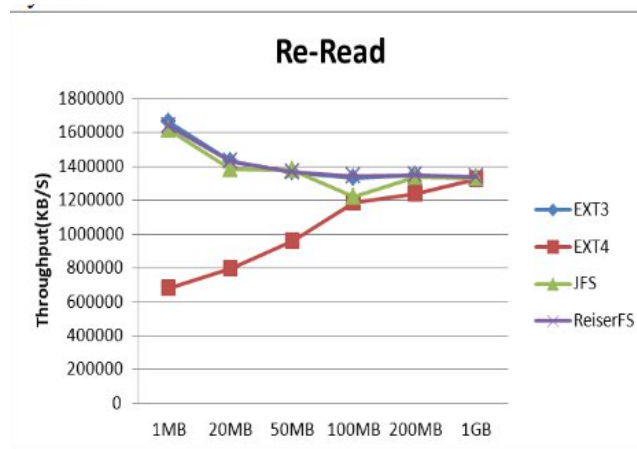


Рисунок 7 - скорость повторного чтения файлов в различных ФС

Согласно рисунку 8, ReiserFS имеет низкую скорость повторной записи при малом объеме файла, однако для средних и больших файлов, различия в скорости рассматриваемых ФС незначительны.

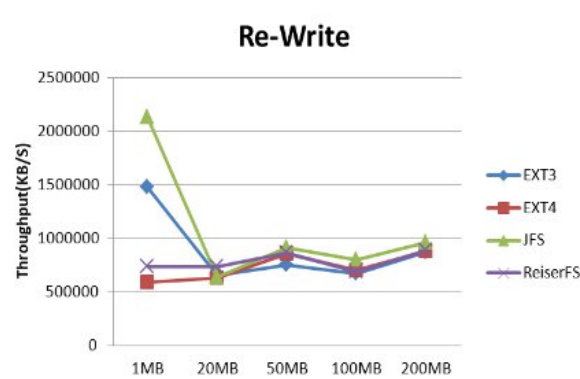


Рисунок 8 - скорость повторной записи файлов в различных ФС

Заключение

В ходе работы были рассмотрены устройство и специфика разделов, каталогов и файлов в файловой системе RaiserFS. Кроме того, была изучена система журналирования и схема борьбы с внутренней дефрагментацией и проанализированы показатели эффективности данной файловой системы в сравнении с распространенными аналогами.

Несмотря на то, что на момент выхода RaiserFS была весьма перспективной файловой системой, реализовавшей ряд новаторских идей, с уходом из проекта Ганса Рейзера его развитие сильно замедлилось и другие файловые системы значительно обошли ее в своем развитии.

Список использованных источников

1. Hans Reiser, Three reasons why ReiserFS is great for you / Hans Reiser // . – 2002. – С.
2. DCIC: ReiserFS [Электронный ресурс]. – Режим доступа: https://cs.uns.edu.ar/~jechaiz/sosd/clases/slides/07-FileSystemsExtra4_Reiser.pdf – Дата доступа: 27.10.2021.
3. Википедия: Журналируемая файловая система [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/%D0%96%D1%83%D1%80%D0%BD%D0%B0%D0%BB%D0%B8%D1%80%D1%83%D0%B5%D0%BC%D0%B0%D1%8F_%D1%84%D0%B0%D0%B9%D0%BB%D0%BE%D0%B2%D0%B0%D1%8F_%D1%81%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0 – Дата доступа: 27.10.2021.
4. Florian Buchholz, The structure of the Reiser file system / Florian Buchholz // . – 2006. – С. 21
5. Википедия: Ext3 [Электронный ресурс]. – Режим доступа: <https://ru.wikipedia.org/wiki/Ext3> – Дата доступа: 27.10.2021.
6. Wikipedia: ReiserFS [Электронный ресурс]. – Режим доступа: <https://en.wikipedia.org/wiki/ReiserFS> – Дата доступа: 27.10.2021.
7. Wikipedia: Block suballocation [Электронный ресурс]. – Режим доступа: https://en.wikipedia.org/wiki/Block_suballocation – Дата доступа: 27.10.2021.
8. Performance Evaluation of the EXT4 File System A Comparative Study Against EXT3, ReiserFS and JFS / Alireza Ghobadi, Amir Hesam Yaribakht, Sanam Maham, Mohammad Hossein Ghods [и др.]. – Cyberjaya : Faculty of Information Technology Multimedia University, 2011. – 6 с.