Новый выпуск

Прыгать на дно

Функция расширения RAIDZ #12225

Закрыто

Архренс Хочет слить 26 Обязуется в `Открытые: мастер` от `ahrens : rowz-expand`

Эта функция позволяет добавлять диски по одному в группу RAID-Z,
постепенное расширение его потенциала. Эта особенность полезна для
небольшие бассейны (обычно только с одной группой RAID-Z), где нет
достаточного оборудования для добавления мощности путем добавления совершенно
новой RAID-Z группы
(обычно удваивает количество дисков).

Для получения дополнительного контекста, а также обзора дизайна, см. мой доклад на саммите FreeBSD Developer Summit (видео) 2021 года (слайды) и новостную статью от Ars Technica.

Описание

Инициирование расширения

Новое устройство (диск) может быть прикреплено к существующему RAIDZ vdev, работая `zpool attach POOL raidzP-N NEW_DEVICE`, например. `zpool attach tank raidz2-0 sda`.
Новое устройство станет частью группы RAIDZ. «Расширение рейда» будет
быть инициированным, и новое устройство внесет дополнительное пространство в RAIDZ
Группа после завершения расширения.

The `feature@raidz_expansion` Флаг на диске должен быть `enabled` к
Инициировать расширение, и оно остается `active` Для жизни бассейна. В
другие слова, пулы с расширенными RAIDZ vdevs нельзя импортировать старыми
Выпуск программного обеспечения ZFS.

Во время расширения

Расширение предполагает чтение всего выделенного пространства с существующих дисков в
RAIDZ группа, и переписывание его на новые диски в группе RAIDZ (в том числе
Недавно добавленное устройство).

Прогресс расширения можно контролировать с помощью `zpool status`.

Избыточность данных поддерживается во время (и после) расширения. Если диск
терпит неудачу, пока расширение продолжается, расширение останавливается до тех пор, пока здоровье не
RAIDZ vdev восстанавливается (например, путем замены неисправного диска и ожидания
для завершения реконструкции).

Бассейн остается доступным во время расширения. После перезагрузки или
экспорт/импорт, расширение возобновляется там, где оно остановилось.

После расширения

Когда расширение завершается, дополнительное пространство доступно для

Расширение не меняет количество сбоев, которые можно терпеть без
Потеря данных (например, RAIDZ2 по-прежнему является RAIDZ2 даже после расширения).

RAIDZ vdev можно расширить несколько раз.

После завершения расширения старые блоки остаются со своими старыми данными по паритету.
соотношение (например, 5-широкий RAIDZ2 имеет 3 данных до 2 паритета), но распределено между
Больше набор дисков. Новые блоки будут написаны с новыми данными-паритет
соотношение (например, 5-широкий RAIDZ2, который был расширен один раз до 6-й, имеет 4 данных
до 2 паритета). Однако "предполагаемый коэффициент паритета" RAIDZ vdev не изменение, поэтому немного меньше места, чем ожидается, может быть сообщено для
Вновь написанные блоки, согласно `zfs list` , `df` , `ls -s` и подобные инструменты.

Изменения Manpage

zpool-прикрепление.8:

```
NAME
     zpool-attach — attach new device to existing ZFS vdev

SYNOPSIS
     zpool attach [-fsw] [-o property=value] pool device new_device

DESCRIPTION
     Attaches new_device to the existing device.  The behavior differs depend-
     ing on if the existing device is a RAIDZ device, or a mirror/plain
     device.

     If the existing device is a mirror or plain device ...

     If the existing device is a RAIDZ device (e.g. specified as "raidz2-0"),
     the new device will become part of that RAIDZ group.  A "raidz expansion"
     will be initiated, and the new device will contribute additional space to
     the RAIDZ group once the expansion completes.  The expansion entails
     reading all allocated space from existing disks in the RAIDZ group, and
     rewriting it to the new disks in the RAIDZ group (including the newly
     added device).  Its progress can be monitored with zpool status.

     Data redundancy is maintained during and after the expansion.  If a disk
     fails while the expansion is in progress, the expansion pauses until the
     health of the RAIDZ vdev is restored (e.g. by replacing the failed disk
     and waiting for reconstruction to complete).  Expansion does not change
     the number of failures that can be tolerated without data loss (e.g. a
     RAIDZ2 is still a RAIDZ2 even after expansion).  A RAIDZ vdev can be
     expanded multiple times.

     After the expansion completes, old blocks remain with their old data-to-
     parity ratio (e.g. 5-wide RAIDZ2, has 3 data to 2 parity), but distrib-
```

```
"assumed parity ratio" does not change, so slightly less space than is
expected may be reported for newly-written blocks, according to zfs list,
df, ls -s, and similar tools.
```

Статус

Считается, что эта функция полноценная. Однако, как и все PR, это предмет
измениться в рамках процесса рассмотрения кода. Так как этот PR включает в себя бортовой
изменения, он не должен использоваться на производственных системах, прежде чем он будет интегрирован в
Кодовая база OpenZFS. Задачи, которые еще предстоит выполнить перед интеграцией:

- [ ] Проверка теста очистки кода
- [ ] Дополнительная очистка кода (адрес всех комментариев XXX)
- [ ] Документируйте дизайн высокого уровня в комментарии «большое теоретическое заявление»
- [ ] Удалить/отключить многословную журналирование
- [ ] Несколько последних тестовых неудач
- [ ] удалить первый бросок (необходимо для получения более чистых тестовых запусков)

Благодарности

Спасибо Фонду FreeBSD за
заказать эту работу в 2017 году и продолжать спонсировать ее далеко за пределами нашей
Оригинальные оценки времени!

Спасибо также авторам **@FedorUporovVstack, @stuartmaybee, @thorsteneb** и **@Fmstrat** за порции
О реализации.

Спонсор: Фонд FreeBSD
Вклады-by: Стюарт Мэйбью stuart.maybee@comcast.net
Вклады: Федор Упоров фупоров.vstack@gmail.com
Вклады: Thorsten Behrens tbehrens@outlook.com
Вклады: Fmstrat nospam@nowsci.com

Как Это Проверялось?

Испытания добавлены в ZFS Test Suite, в дополнение к ручному тестированию.

Типы изменений

- [ ] Исправление ошибки (неразбивающая смена, которая решает проблему)

эффективность)

- [ ] Очистка кода (неразбивка изменений, которые делают код меньше или более читаемым)
- [ ] Слоеное изменение (исправление или функция, которая приведет к изменению существующей функциональности)
- [ ] Документация (изменение страниц человека или другой документации)

Контрольный список:

- [x] Мой код следует ZFS на Linux требованиям к стилю кода.
- [x] Соответствующим образом я обновил документацию.
- [x] Я прочитал **Вносящий вклад** в документ.
- [x] Я добавил тесты Чтобы покрыть мои изменения.
- [ ] Все новые и существующие тесты прошли.
- [ ] Все сообщения о совершении правильно отформатированы и содержат `Signed-off-by` .

👍 272 😄 1 🎉 189 ❤️ 138 🚀 86 👀 67

Архренс Упомянули эту просьбу о вытягива on 11 июн. 2021 г.

[WIP] рейзовое расширение, alfa preview 1 #8853

Закрыто

12 задач

Архренс назначенный **Мейби** on 11 июн. 2021 г.

Архренс Добавлено Статус: Необходим Обзор Кода Тип: Особенность этикетки on 11 июн. 2021 г.

Эверов Комментировать on 11 июн. 2021 г.

Поздравляем с прогрессом!

👍 68 🎉 23 ❤️ 19

Корним Комментировать on 11 июн. 2021 г.

Также поздравляю с тем, что вывело это из альфы.

У меня есть вопрос относительно

данных к паритету.

Раздел и это может помочь сделать небольшой пример здесь:

Если у меня есть RAIDZ 8x1TB и загрузить в него 2 ТБ, он будет заполнен на 33%.

Для сравнения, если я начну с 4x1TB RAIDZ2, который заполнен, и расширим его до 8x1TB RAIDZ2, он будет заполнен на 50%.

Правильно ли это понимание? Если это так, это будет означать, что всегда нужно начинать расширение с как можно более пустым vdev. Поскольку это не всегда будет вариантом, есть ли возможность (планируется) перезаписать старые данные в новый паритет? Будет ли перемещение данных с Vdev, а затем обратно делать работу?

👍 1

ГурлиДжебис Комментировать on 11 июн. 2021 г.

Перемещение данных и обратно должно быть сделано, я думаю, поскольку он переписывает данные (снимки могут помешать им).
Если дедупликация отключена, копирование файлов и удаление старой версии может сделать трюк (или я что-то упускаю?)

Архренс Силовой толчок the `рейнц-расширение` филиал от **`63c41fc`** к **`9f49205`** Сравнить
4 года назад

Архренс Комментировать on 11 июн. 2021 г. Член Автор

@cornim Я думаю, что математика правильная - это один из худших случаев (вам было бы лучше начать с зеркал, чем с 4-широкого RAIDZ2; и если вы удваиваете свое количество дисков, вам может быть лучше добавить новую группу RAIDZ). Возьмем другой пример, если у вас есть 5-широкий RAIDZ1 и добавить диск, вы все равно будете использовать 1/5 место (20%) пространства в качестве паритета, тогда как недавно написанные блоки будут использовать 1/6-е (17%) пространства в качестве паритета - разница в 3%.

@GurliGebis Rewriting блоков приведет к тому, что они будут обновлены до нового соотношения данных: паритет (например, экономия 3% места в 5-м экземпляре). Предполагая, что нет никаких моментальных снимков, копирование каждого файла или загрязнение каждого блока (например, чтение первого байта, а затем запись того же байт обратно) сделало бы трюк. Если есть снимки (и вы хотите сохранить их обмен блоками), вы можете использовать `zfs send -R` Для копирования данных. Должна быть возможность добавить некоторую автоматизацию, чтобы было легче перезаписывать ваши блоки в обычных случаях.

Может быть полезно прямо указать, становится ли дополнительное пространство доступным во время расширения или только после завершения расширения.

Архренс Комментировать on 11 июн. 2021 г. • • отредактировано Член Автор

@stuartthebruce Хороший пункт. Об этом говорится в сообщении о совершенствовании и PR-письме:

> Когда расширение завершается, дополнительное пространство доступно для использования

~~Но я добавлю его и к странице.~~ О, это также указано в манифесте:

> Новое устройство будет способствовать
> дополнительное пространство для группы RAIDZ после завершения расширения.

Стюартабрюса Комментировать on 11 июн. 2021 г.

> ~~Но я добавлю его и к странице.~~ О, это также указано в манифесте:
>
> > Новое устройство будет способствовать
> > дополнительное пространство для группы RAIDZ после завершения расширения.

Если это не педантично, как насчет «дополнительного пространства ... только после завершения расширения». Нынешняя формулировка оставляет открытой возможность того, что пространство может стать доступным постепенно во время расширения.

👍 3 ❤️ 2

фелисукойби Комментировать on 12 июн. 2021 г. • • отредактировано

Один вопрос, если я добавлю к пулу 10 драйверов еще 5 с этой системой, то 5 новых имеют новый паритет, и после этого старые 10 заменяются шаг за шагом с заменяющей командой, в конце, когда 10 старых заменят все рейд будут иметь новый паритет? при замене старых паритет используется или используется новый паритет, поэтому мы можем потреблять дополнительное пространство.

Архренс Комментировать on 12 июн. 2021 г. Член Автор

@felisucoibi Я не уверен, что полностью понимаю ваш вопрос, но вот пример, который может быть связан с вашим: если вы начинаете с 10-широкого RAIDZ1 vdev, а затем делаете 5 `zpool attach` операции по добавлению еще 5 дисков к нему, у вас будет 15-широкий RAIDZ1 vdev. Если 5 новых дисков были больше, чем 10 старых дисков, и вы тогда

а вновь написанные блоки будут использовать новый коэффициент данных: паритет 14:1 (6,7%). Так что разница в пространстве, используемом по паритету, составляет всего 3,3%.

фелисукойби Комментировать on 12 июн. 2021 г.

> **@felisucoibi Я** не уверен, что полностью понимаю ваш вопрос, но вот пример, который может быть связан с вашим: если вы начинаете с 10-широкого RAIDZ1 vdev, а затем делаете 5 `zpool attach` операции по добавлению еще 5 дисков к нему, у вас будет 15-широкий RAIDZ1 vdev. Если 5 новых дисков были больше, чем 10 старых дисков, и вы тогда `zpool replace` каждый из 10 старых дисков с новым большим диском, тогда vdev сможет расширить пространство 15 новых больших дисков. В любом случае, старые блоки продолжат использовать существующий коэффициент данных: паритета 9:1 (10% паритет), а вновь написанные блоки будут использовать новый коэффициент данных: паритет 14:1 (6,7%). Так что разница в пространстве, используемом по паритету, составляет всего 3,3%.

Спасибо за ответ. Таким образом, единственный способ пересчитать старые блоки - переписать данные, как вы предложили.

Лоуэнриус Комментировать on 12 июн. 2021 г. • • отредактировано Участник

> После завершения расширения старые блоки остаются со старым соотношением данных к паритету.

Прежде всего, это замечательная особенность, спасибо. Если я могу спросить: почему старые блоки не переписываются, чтобы вернуть себе дополнительное пространство? Я могу себе представить, что перераспределение данных влияет только на меньшую часть всех данных и, таким образом, быстрее, но тогда пользователю все еще приходится переписывать данные, чтобы восстановить пространство для хранения. Было бы неплохо, если бы это можно было сделать в рамках процесса расширения в качестве дополнительного варианта, если люди готовы принять необходимое дополнительное время? За что бы это ни стоило.

👍 17 👎 1

йорикдаун Комментировать on 12 июн. 2021 г.

ZFS имеет философию «не связывайтесь с тем, что уже на диске, если вы можете избежать этого. Если нужно идти на крайности, чтобы не связываться с тем, что было написано (например, отображение памяти удаляет диски в пулах зеркал). Кто-то, кто хочет переписать старые данные, может сделать этот выбор и отправить / рецензирование, что является простой операцией. Мне нравится, как это работает сейчас.

командabvd Комментировать on 12 июн. 2021 г. • • отредактировано

@louwrentius Я тот же ум - хотя, исходя из другого направления, и, учитывая сложность кода, я задавался вопросом, возможно, будет ли компонент перераспределения данных, возможно, будет указан как последующий PR ...? Я искал существующий в том случае, если он уже был там, но это может быть что-то, что уже продумано / спланировано, и я просто не смог его найти.

Учитывая количество задействованных компонентов и сложность операций, которые были бы необходимы, особенно в том, что касается памяти и моментальных снимков, я мог видеть, что имеет смысл разделить задачи. Я представляю себе что-то вроде...

- Расширение завершено
- Существующая полоса чтения - как мы обеспечиваем, чтобы считываемые данные были записаны обратно в тот же vdev? Если в журнале намерений, нам понадобится метод, чтобы направить эти записи обратно в конкретный vdev и обойти текущий метод распределения vdev
- Stripe rewritten на «новую» полосу - Предполагая, что есть сфотографированные данные о vdev, как эти снимки метаданные обновляются, чтобы отразить новые местоположения блока? Метаданные о других vdevs могут (вероятно,) указывают на bbas, размещенные на этом vdev, и обновление снимка может быть ... проблематичным. Обязываем ли мы это сделать требование о том, что в бассейне не может существовать снимков для этой операции?
- Существующая полоса освобождена

По крайней мере, для меня, чем больше я думаю об этом, тем больше смысла будет иметь перебалансировку / перераспределение «уровня пула», поскольку все существующие данные в ветхих пула обычно зависят друг от друга. Это, безусловно, упростило бы вещи по сравнению с тем, что я описываю выше, я бы подумал. Это также помогает решить другие проблемы, которые были давними, особенно в том, что касается производительности долгоживущих бассейнов, которые, возможно, со временем добавили несколько vdev, поскольку существующий пул стал полным.

В любом случае, я не хочу слишком много болтать - я мог просто увидеть, как было бы разумно, по крайней мере, на каком-то уровне, чтобы перераспределение данных было еще одним пиаром.

👍 4

Муфуньо Комментировать on 12 июн. 2021 г.

> Предполагая, что нет никаких моментальных снимков, копирование каждого файла или загрязнение каждого блока (например, чтение первого байта, а затем запись того же байт обратно) сделало бы трюк. Если есть снимки (и вы хотите сохранить их обмен блоками), вы можете использовать `zfs send -R` Для копирования данных. Должна

Было бы очень признательно наличие легкодоступной команды для переписывания всех старых блоков или предпочтительно возможности сделать это в рамках процесса расширения.

👍 11

Архренс Комментировать on 12 июн. 2021 г. Член Автор

@louwrentius **@teambvd** **@yorickdowne** **@mufunyo** Я думаю, что вы все попадают на несколько разных вопросов:

1. Что будет вовлечено в перераспределение существующих блоков, чтобы они потребляли меньше места? Делать это должным образом - онлайн, работа с другими существующими функциями (снимками, клонами, dedup), не требуя тонн дополнительного пространства - потребует постепенного изменения моментальных снимков, что является проектом аналогичного масштаба RAIDZ Expansion. Существуют обходные пути, которые достигают одного и того же конечного результата в случаях ограниченного использования (без снимков? коснуться всех блоков. много места? `zfs send -R`)
2. Сколько это будет пользы? Я привел несколько примеров выше, но обычно это несколько процентов. (например, 5-широкий -> 6-широкий, вы получаете по крайней мере 5/6-е (83%) привода дополнительного полезного пространства, и если вы перераспределите существующие блоки, вы можете получить целый привод полезного пространства, то есть дополнительные 17% диска или ~ 3% всего бассейна. Более широкий RAIDZ будет видеть меньше воздействия (9-широкий -> 10-широкий, вы получаете по крайней мере 90% привода дополнительного полезного пространства; вы упускаете 1% всего бассейна).
3. Почему я еще не сделал этого? Потому что это невероятное количество работы для небольшой выгоды, и я считаю, что RAIDZ Expansion, как разработано и реализовано, полезен для многих людей.

Все это говорит о том, что я был бы счастлив, если бы оказался неправ в отношении сложности этого! Такая установка может быть использована для многих других задач, например, рекомпрессия существующих данных, чтобы сэкономить больше места (lz4 -> zstd). Если у кого-то есть идеи о том, как это может быть реализовано, возможно, мы сможем обсудить их на дискуссионном форуме или в новом запросе. Еще одна область, с которой люди могут помочь, - это автоматизация переписывания в случаях ограниченного использования (при прикосновении ко всем блокам или `zfs send -R`)

👍 25 ❤️ 11

Джеркисан Комментировать on 12 июн. 2021 г.

> **@louwrentius @teambvd @yorickdowne @mufunyo** Я думаю, что вы все попадают на

потребляли меньше места? Делать это должным образом - онлайн, работа с другими существующими функциями (снимками, клонами, dedup), не требуя тонн дополнительного пространства - потребует постепенного изменения моментальных снимков, что является проектом аналогичного масштаба RAIDZ Expansion. Существуют обходные пути, которые достигают одного и того же конечного результата в случаях ограниченного использования (без снимков? коснуться всех блоков. много места? `zfs send -R`)

2. Сколько это будет пользы? Я привел несколько примеров выше, но обычно это несколько процентов. (например, 5-широкий -> 6-широкий, вы получаете по крайней мере 5/6-е (83%) привода дополнительного полезного пространства, и если вы перераспределите существующие блоки, вы можете получить целый привод полезного пространства, то есть дополнительные 17% диска или ~ 3% всего бассейна. Более широкий RAIDZ будет видеть меньше воздействия (9-широкий -> 10-широкий, вы получаете по крайней мере 90% привода дополнительного полезного пространства; вы упускаете 1% всего бассейна).
3. Почему я еще не сделал этого? Потому что это невероятное количество работы для небольшой выгоды, и я считаю, что RAIDZ Expansion, как разработано и реализовано, полезен для многих людей.

Все это говорит о том, что я был бы счастлив, если бы оказался неправ в отношении сложности этого! Такая установка может быть использована для многих других задач, например, рекомпрессия существующих данных, чтобы сэкономить больше места (lz4 -> zstd). Если у кого-то есть идеи о том, как это может быть реализовано, возможно, мы сможем обсудить их на дискуссионном форуме или в новом запросе. Еще одна область, с которой люди могут помочь, - это автоматизация переписывания в случаях ограниченного использования (при прикосновении ко всем блокам или `zfs send -R`)

Я просто хотел бы заявить, что я очень ценю работу, которую вы делаете. Честно говоря, это одна из вещей, удерживающих многих людей от использования ZFS, и возможность выращивать пул zfs без необходимости добавлять vdevs будет буквально волшебной. Я смогу легко переключиться обратно на ZFS после того, как это будет завершено. ВНОВЬ СПАСИБО БОЛЬШОЕ!!

👍 39 ❤️ 7

Келлекиндт Комментировать on 12 июн. 2021 г. • • отредактировано

@ahrens Как кто-то, молчаливо следящий за прогрессом со времен оригинального пиара, я также хочу отметить, что я действительно ценю все усилия и обязательства, которые вы вкладываете и вкладываете в эту функцию! Я верю, что как только это приземлится, это будет действительно ценным дополнением к zfs :)

Спасибо вам

👍 21 ❤️ 7

@kellerkindt **@Jerkysan** **@Evernow** Спасибо за добрые слова! Это действительно делает мой день, чтобы знать, что эта работа будет полезной и оцененной! ❤️ 😄

👍 17 ❤️ 63

ДэйБлур Комментировать on 13 июн. 2021 г.

Спасибо за работу над этой функцией. Захватывающе, наконец, увидеть некоторый прогресс в этой области, и это будет полезно для многих людей после освобождения.

Закладывают ли эти изменения какую-либо основу для будущей поддержки добавления паритетного диска (вместо диска данных - т.е. повышения уровня RAID-Z)? Значительное увеличение количества дисков в существующем массиве, вероятно, вызовет желание повысить уровень отказоустойчивости.

Поскольку существующие данные просто перераспределяются, я понимаю, что старые данные не будут иметь повышенной избыточности, если не будут переписываться. Но мне все еще любопытно, может ли ваша работа, которая позволяет поддерживать старые / новые макеты данных + паритет одновременно в пуле, также может применяться к увеличению количества паритетных дисков (и алгоритма) для будущих письменных записей.

👍 5

рикатнойт11 Комментировать on 14 июн. 2021 г.

Я также невероятно взволнован этой проблемой и понимаю решение разделить перераспределение существующих данных на другой FR. Большое спасибо за всю тяжелую работу, которая вложилась в это. Я не могу дождаться, чтобы воспользоваться этим.

После выполнения расширения рейза существует, по крайней мере, точный механизм для определения *того, какие* объекты (файлы, снимки, наборы данных... Я признаю, что я не полностью обернул свою голову вокруг того, как это влияет на вещи, поэтому приносим извинения за то, что не использовали правильные термины) карта блоков с «старым» соотношением данных к паритету и, возможно, вычислим увеличение пространства? Я предполагаю, что многие администраторы хотят баланса между максимизацией пространства, извлекаением из всех преимуществ ZFS (счета, дедупликация и т. Д.) И гибкостью расширения хранения (да, мы хотим съесть все торты) и, естественно, сравнивают эту функцию с другими технологиями, такими как md рейд, где выращивание рейд-массива вызывает пересчет всех паритетов. Таким образом, эти администраторы захотят иметь возможность планировать, как сделать то же самое на группе с расширенным реймисловием, не просто слепо переписывая все файлы.

@DayBlur

> Закладывают ли эти изменения какие-либо основы для будущей поддержки добавления паритетного диска (вместо диска данных - т.е. повышения уровня RAID-Z)? ... Мне все еще любопытно, может ли ваша работа, которая позволяет поддерживать старые / новые макеты данных + паритет одновременно в пуле, также может применяться к увеличению количества паритетных дисков (и алгоритма) для будущих записей.

Да, это так. Эта работа позволяет добавлять диск, и будущая работа может заключаться в увеличении паритета. Как вы упомянули, переменная, основанная на времени схема геометрии RAIDZ Expansion может быть использована, чтобы знать, что старые блоки имеют старое количество паритета, а новые блоки имеют новое количество паритета. Эта работа была бы довольно простой.

Однако тот факт, что старые блоки остаются со старым количеством отказоустойчивости, означает, что в целом вы не сможете терпеть все большее количество сбоев, пока все старые блоки не будут освобождены. Поэтому я думаю, что практически важно иметь механизм, чтобы хотя бы наблюдать количество старых блоков, и, вероятно, также перераспределять старые блоки. В противном случае вы не сможете терпеть больше сбоев, не теряя [старых] данных. Это было бы сложно реализовать в общем случае, но, как упоминалось выше, есть некоторые OK решения для особых случаев.

👍 7

Архренс Комментировать on 15 июн. 2021 г. Член Автор

@rickatnight11

> После выполнения расширения рейза существует, по крайней мере, точный механизм для определения того, какие объекты (файлы, снимки, наборы данных... Я признаю, что я не полностью обернул свою голову вокруг того, как это влияет на вещи, поэтому приносим извинения за то, что не использовали правильные термины) карта блоков с «старым» соотношением данных к паритету и, возможно, вычислим увеличение пространства?

Нет простого способа сделать это, но вы можете получить эту информацию из `zdb`. В основном вы ищете блоки, время рождения которых до времени завершения расширения. Я думаю, что было бы относительно просто сделать некоторые инструменты, чтобы помочь с этой проблемой. Отправной точкой может быть сообщение о том, какие снимки были созданы до завершения паритетного расширения, и поэтому их необходимо будет уничтожить, чтобы освободить это пространство.

👍 2

@ahrens Я также хочу поблагодарить вас за потрясающую работу!

У меня есть несколько вопросов:

- Нужно ли DVA менять после того, как диск прикреплен к RAIDZ vdev?
- Если да, то как DVA переписывается? Вы используете «заполнитель», указывающий на новый DVA?
- Если DVA не изменяется, как на самом деле работает переезд?

Спасибо.

👎 1

Бгира Комментировать on 16 июн. 2021 г.

> Честно говоря, это одна из вещей, которая удерживает многих людей от использования ZFS.

Любопытно, с какой файловой системой или объединением / RAID настройками шли эти воображаемые люди; предположительно, он также проверяет большинство полей, которые делает ZFS? менеджер томов, поставщик шифрования, сжатие файловой системы, снимками и т.д.

Джеркисан Комментировать on 16 июн. 2021 г.

> Честно говоря, это одна из вещей, которая удерживает многих людей от использования ZFS.
>
> Любопытно, с какой файловой системой или объединением / RAID настройками шли эти воображаемые люди; предположительно, он также проверяет большинство полей, которые делает ZFS? менеджер томов, поставщик шифрования, сжатие файловой системы, снимками и т.д.

Ответ: «Мы договорились»... Я остановился на ненападении, хотя до этого момента я управлял ZFS в течение многих лет. Я больше не мог позволить себе покупать достаточно дисков сразу, чтобы построить коробку ZFS напрямую. У меня были жизненные обязательства, которые я должен был выполнить, желая продолжать свое хобби, которое, казалось бы, никогда не перестает расширяться в цене. Мне нужно было что-то, что могло бы расти вместе со мной. Это просто не дает мне сжимающую файловую систему, снимки и яда-яда.

В основном, мне пришлось пойти на жертвы, чтобы продолжить свое хобби, не буквально не нарушая банк. Эта функциональность позволит мне вернуться к ZFS. Мне все равно, если мне нужно переместить все файлы, чтобы заставить их использовать новые диски, которые я добавляю и тому подобное. Это не «критически важно для производства», но я

пытаюсь выяснить, как я собираюсь скатиться в это.

👍 25

Кьелегордон Комментировать on 16 июн. 2021 г.

> Честно говоря, это одна из вещей, которая удерживает многих людей от использования ZFS.
>
> Любопытно, с какой файловой системой или объединением / RAID настройками шли эти воображаемые люди; предположительно, он также проверяет большинство полей, которые делает ZFS? менеджер томов, поставщик шифрования, сжатие файловой системы, снимками и т.д.

Как один из этих воображаемых людей, один массив пошел использовать BTRFS, а другой массив использует mdadm, lvm2 и ext4. Я доволен ни тем, ни другим.

Основным драйвером для выбора были как возможности расширения, так и отказоустойчивость диска.

👍 5

йорикдаун Комментировать on 16 июн. 2021 г.

Чтобы быть реалистом, уже слишком поздно, чтобы это попало в OpenZFS 2.1 в августе, это означает, что, возможно, OpenZFS 2.2 в августе 2022 года, тогда ему нужно сделать его в «OS выбора». TrueNAS *может* пойти на это рано, как это не так, но это не то, на что можно положиться.

В то же время, «sorta расширение» с send/recv в el cheapo SAS от eBay и обратно определенно работает. Я сделал это от 5x8TB до 8x8TB. Не сломал банк и теперь имеет больше места, чем я, вероятно, буду использовать.

justinclift Комментировать on 16 июн. 2021 г.

> это означает, что, возможно, OpenZFS 2.2 в августе 2022 года

С одной стороны, хотя это кажется надолго... это кажется существенным изменением возможностей. Так что дополнительный ~ год (непроизводство!) тестирование, чтобы выявить странные, краевые ошибки могут оказаться полезными. 😄

👍 6

Хе Хе-хе. Есть статья Ars Technica об этом тоже:

https://arstechnica.com/gadgets/2021/06/raidz-expansion-code-lands-in-openzfs-master/

Это хорошая акция прямо здесь. 😉

👍 2

Соковысли Комментировать on 17 июн. 2021 г. • • отредактировано

> **@louwrentius** **@teambvd** **@yorickdowne** **@mufunyo** Я думаю, что вы все попадают на несколько разных вопросов:
>
> 1. ... Существуют обходные пути, которые достигают одного и того же конечного результата в случаях ограниченного использования (без снимков? коснуться всех блоков. много места? `zfs send -R`)

Можете ли вы указать на хорошую ссылку на любую из этих вещей? Единственный способ, которым я могу придумать, чтобы «затронуть все блоки», — это скопировать каждый файл через какой-то тупой gzip|gzip трубопровод.

👍 1

СтарманУК Комментировать on 18 июн. 2021 г.

> В то же время, «sorta расширение» с send/recv в el cheapo SAS от eBay и обратно определенно работает. Я сделал это от 5x8TB до 8x8TB. Не сломал банк и теперь имеет больше места, чем я, вероятно, буду использовать.

Не могли бы вы расширить это как метод?

👍 3

йорикдаун Комментировать on 18 июн. 2021 г. • • отредактировано

Конечно. Это просто осознание того, что используемые диски SAS дешевы на eBay. Я построил приемный пул примерно за 220 долларов США, сжег его, отправил туда свои данные, удалил оригинальный пул, переделал его 8-широко (с сожженными дисками) и отправил данные в другую сторону. В целом, это заняло бы 2 недели; вероятно, было бы быстрее, если бы я раньше понял, что диски SAS отключили кэш записи.

Этот метод подходит для расширения рейдов по варианту использования для: Hobbyists. Я в порядке, если мой дом NAS не работает в течение недели, пока его бассейн будет заселен. Это также единственная среда, о которой я могу думать, где планирование

and-test-system.86434/

👍 2

Лоуэнриус Комментировать on 19 июн. 2021 г. • • отредактировано Участник

@ahrens

Потеря эффективности функции расширения RAIDZ VDEV может стать более значительной в ситуации домашнего пользователя. Если этот сценарий актуален, полностью зависит от того, с чем люди лично согласны.

Домашний пользователь хотел бы начать с приличной избыточности, но небольшого количества дисков. Скажем, 4-приводный RAIDZ2. Пользователь купил шасси, которое может вместить до 10 дисков. Затем пользователь добавляет диск каждый раз, когда мы нажимаем на 80% использования емкости.

Таким образом, мы расширяем с 4 дисков с одним приводом за раз, до максимум 10 дисков. Если мой расчет верный, к тому времени, когда этот пользователь заполняет шасси 10 дисками, почти емкость диска (87%) теряется из-за этого накладного, если вы расширяетесь, когда емкость заполнена на 80%. (*Редактируйте, моя математика была неправильной, это гораздо больше, в этом сценарии теряется около 1,35 приводов.)*

Честно говоря, эти накладные расходы намного дешевле или намного меньше, чем текущие накладные расходы на добавление дополнительных VDEVS для добавления расширения. Таким образом, ясно, что это действительно гораздо лучшее решение. Но есть еще цена, которую нужно заплатить, если вы не переписываете данные. Если расширение ZFS + RAIDZ стоит этой цены, это личное, я думаю.

ZFS предназначен для корпоративного хранения, где бюджет не является таким препятствием, но кажется, что расширение RAIDZ VDEV по-прежнему рассматривается как достаточно важная проблема для решения. Может быть, в том же духе, этот вопрос все еще может представлять интерес, даже если потери значительно меньше.

Я думаю, что домашние пользователи более терпимы к решениям, которые связаны с временной потерей производительности или даже другими неудобствами (может быть, с существующими данными только для чтения? (просто чтобы проиллюстрировать способ мышления)) в обмен на очень простой процесс восстановления «потерянной» способности путем переписывания данных. Я думаю, что покупка дешевых накопителей SAS для временного хранения данных не является действительно удобным и жизнеспособным решением для большинства людей.

Это может быть двухэтапный процесс. Первый этап - это перераспределение всех данных по всем дискам, как это работает в настоящее время. Второй этап является необязательным, когда пользователь передает дополнительный параметр cmd-line, который, если установлен, ZFS будет переписывать данные в фоновом режиме. Было бы неплохо, если бы этот второй процесс мог быть начат в более позднее время, если это

некоторое время, я предполагаю, что для большинства домашних пользователей это не проблема.

Недавно добавленный накопитель + первый пропуск, где данные перераспределяются по всем дискам, обеспечивает свободное пространство, необходимое для переписывания существующих данных на месте, я бы предположил. Таким образом, это позволило бы этому процессу работать на месте, без необходимости в некотором «внешнем» хранилище для переписывания данных.

Особенно, если вы дойдете до точки, где пул состоит из более чем одного RAIDZ VDEV, возможность переписывания данных (удобным для пользователя способом) становится гораздо более важной. Редактировать: особенно если несколько VDEVS расширены, потери накапливаются.

Я надеюсь, что в этом есть смысл, и я не ошибаюсь. Если ничего из этого не осуществимо, это нормально, я просто хочу поделиться своим мнением об этой функции.

👍 11 ❤️ 6

Фмстрат Комментировать on 19 июн. 2021 г. • • отредактировано

@ahrens

То, что описывает **@louwrentius,** является именно вариантом использования для меня, и, честно говоря, для всех, кого я знаю, кто использует ZFS за пределами бизнес-среды.

Не так вероятно, что у пользователя Homelab будут дополнительные диски для временной отправки всех данных, поэтому возможность переписать на месте в качестве отдельного PR кажется полезной функцией (для меня, во всяком случае).

Черт возьми, даже если бы был процесс использования одного дополнительного диска для копирования файла, освободить старые блоки из файла, скопировать обратно, повторить для всех файлов в качестве скрипта было бы полезно.

👍 5

justinclift Комментировать on 19 июн. 2021 г. • • отредактировано

> ... для использования одного дополнительного диска для скопирования файла, освободить старые блоки из файла, скопировать обратно на ...

Имейте в виду, что независимо от этого временного места хранения, наличие некоторой избыточности данных (хотя и временно) довольно важно.

Мы, вероятно, все сталкивались со случаями, когда «совершенно новый блестящий» большой диск, который временно подключается ... начинает развивать проблемы 2 часа

dnelson-1901 Комментировать on 19 июн. 2021 г. • • отредактировано

По причинам производительности я предпочитаю наборы зеркальных vdev вместо одного рейза для моего основного пула, но я бы порекомендовал всем, кто использует большие пулы, следовать следующим общим правилам:

1. Иметь два диска за пределами основного бассейна, которые можно зеркалировать, чтобы сделать промежуточный бассейн, и
2. Никогда не позволяйте какой-либо одной файловой системе zfs расти больше, чем может удерживаться постановка.

Например, если у вас есть рейд 5x 6TB, у вас всегда должен быть один на полке в качестве запасного в любом случае, поэтому покупка другого позволит вам иметь промежуточную зону 6TB. Если у вас есть меньший диск, оставшийся от чего-то другого, используйте свой запасной и этот диск вместо того, чтобы покупать другой. Если у вас есть куча маленьких оставшихся дисков, налетайте на них в бассейн, если это дает вам больше места, чем зеркала.

Создайте столько файловых систем в своем основном пуле, сколько необходимо для сохранения правила 2. Если вы храните резервные копии других серверов, создайте «резервное копирование/хост1», «резервное копирование/хост2» и т.д. Если вы делаете CI и храните артефакты сборки или запускаете несколько серверов сборки параллельно, создайте файловую систему для каждого из них. Файловые системы почти бесплатны; они только добавляют строку к вашему выходу «df». :) Тогда, если ваш основной пул должен быть перебалансирован, это просто вопрос прохода через каждую файловую систему, выполняя последовательность «отправить | получать», «уничтожить», «отправить | получать» по очереди.

Если у вас нет свободных дисков, но у вас есть свободное место в вашем основном пуле, вы можете «отправить | получения» на временное имя в основном пуле, а затем перемешать переименовать имена, чтобы поменять имена и удалить исходную файловую систему. Обратите внимание, что если у вас есть только один большой рейд vdev, он будет бить ваши диски во время копии, так что промежуточный пул - это действительно путь.

👎 1

Соковысли Комментировать on 20 июн. 2021 г. • • отредактировано

> Я думаю, что домашние пользователи более терпимы к решениям, которые связаны с временной потерей производительности или даже другими неудобствами (может быть, с существующими данными только для чтения? (просто чтобы проиллюстрировать способ мышления)) в обмен на очень простой процесс восстановления «потерянной» способности путем переписывания данных. Я думаю,

> Это может быть двухэтапный процесс. Первый этап - это перераспределение всех данных по всем накопителям, как это работает в настоящее время. Второй этап является необязательным, когда пользователь передает дополнительный параметр cmd-line, который, если установлен, ZFS будет переписывать данные в фоновом режиме. Было бы неплохо, если бы этот второй процесс мог быть начат в более позднее время, если это будет более удобно.
>
> Хотя это требует в некотором смысле двух проходов к данным, и это может занять некоторое время, я предполагаю, что для большинства домашних пользователей это не проблема.

Это в значительной степени описывает меня точно. Я действительно не хочу покупать «промежуточную площадку», которая достаточно велика, чтобы хранить все данные моего основного массива [1], так как это в основном удваивает общую стоимость.

Я думаю, что это точная оценка, что люди, работающие в меньшем масштабе (например, малый бизнес, дом), гораздо более терпимы к временно ухудшенной производительности. До тех пор, пока данные безопасны (и все в конечном итоге вернется к норме), есть МНОГО, с чем я бы смирился. Бонусные баллы за оценку времени завершения.

Даже приступы только для чтения или необходимость кэша писать на внешнее зеркало во время реформы (для последующего применения) могут быть приемлемыми. (можно ли злоупотреблять зеркалами SLOG для этого, интересно?)

[1] Да, если бы основной массив был разделен на более мелкие массивы, постановка могла бы быть меньше = max (marray1, marray2,...). Тем не менее, это менее практично с меньшим количеством общих приводов (скажем, 6-8). Также значительно полезнее иметь один большой пул (который может быть повторно разделен по мере необходимости на наборы данных / zvols). Размахивание нескольких RAIDZ (я думаю) сделает невозможным вытащить один и реформировать данные.

РЕДАКТИРОВАТЬ: После перечитывания комментария выше я теперь понимаю, что он может работать с одним RAIDZ, если каждый набор данных / zvol не больше, чем постановка. Это более работоспособно, хотя это все еще неудачный произвольный предел. Я бы, вероятно, подумал о том, чтобы остановить все и запустить сценарий, чтобы скопировать каждый файл туда-сюда в гораздо меньшую постановку.

👍 3

Лоуэнриус Комментировать on 20 июн. 2021 г. • • отредактировано ▾ Участник

> Это в значительной степени описывает меня точно. Я действительно не хочу покупать «промежуточную площадку», которая достаточно велика, чтобы хранить все данные моего основного массива [1], так как это в основном удваивает общую стоимость.

вручную перемещать данные.

Идея использования SLOG или какого-либо другого устройства (идеально SSD?) для временного хранения данных, чтобы ускорить процесс - если это имеет смысл - звучит интересно для меня (как noob).

> Я думаю, что это точная оценка, что люди, работающие в меньшем масштабе (например, малый бизнес, дом), гораздо более терпимы к временно ухудшенной производительности. До тех пор, пока данные безопасны (и все в конечном итоге вернется к норме), есть МНОГО, с чем я бы смирился. Бонусные баллы за оценку времени завершения.

В том же духе временные функции отключения, такие как снимки, также могут быть приемлемы для этой группы пользователей. Я могу понять, что это может идти вразрез с фундаментальными принципами проектирования ZFS: возможно, это требование, чтобы все операции были возможны, пока все функции остаются в сети. Но не у всех пользователей есть это требование, я думаю.

Кроме того, я подозреваю (хотя у меня нет доказательств), что такая функция, как дедупликация, вероятно, не очень часто используется домашними пользователями и средами малого и среднего бизнеса из-за крутых требований к оборудованию. Может быть, это неправда, а просто мысль. Может быть, это сделает вещи проще.

Моя цель + надежда - найти способ, чтобы мы не создали совершенно новый «проект» с теми же усилиями, которые связаны с расширением RAIDZ VDEV. Для достижения этой цели, как домашние и малые и средние пользователи, мы можем принять определенные (временные) ограничения или ограничения, которые в обмен позволяют нам восстановить дополнительное пространство простым способом. Я признаю, что я не могу контролировать последствия этого запроса, но все, что я могу сказать, это то, что, если сказанные (временные) ограничения не против основных философий ZFS, я думаю, что есть место для «переговоров», где все могут быть еще счастливее, чем мы уже с расширением RAIDZ VDEV. 😃

👍 1

Соковысли Комментировать on 20 июн. 2021 г.

> Кроме того, я подозреваю (хотя у меня нет доказательств), что такая функция, как дедупликация, вероятно, не очень часто используется домашними пользователями и средами малого и среднего бизнеса из-за крутых требований к оборудованию. Может быть, это неправда, а просто мысль. Может быть, это сделает вещи проще.

По моему опыту, дедупликация не слишком дорогая, если вы не используете ZFS на очень ограниченной платформе, такой как распи. Эмпирическое правило 1 ГБ на ТБ обычно может быть поддержано, если у вас, скажем, 16 ГБ оперативной памяти в машине с 10-14 ТБ памяти.

Лоуэнриус Комментировать on 21 июн. 2021 г. • • отредактировано Участник

Я хотел бы уделять больше внимания в школе, когда преподается базовая математика. Я думаю, что мой накладной расчет был очень неправильным.

Я думаю, что эта электронная таблица показывает, что накладные расходы ~~очень малы,~~ хотя я был бы признателен, если бы кто-то проверил математику. *(редакция: моя математика была очень неправильной, накладная голова гораздо более значительна, благодаря ЙорикДауну и Дэйблуру)*

~~Если моя коррекция действительна, для меня означает, что функция расширения ZFS RAIDZ VDEV практически не имеет недостатков.~~
~~На мой взгляд, очень небольшая потеря емкости - это просто (низкая) цена, которую нужно заплатить за ZFS, и вы получаете так много обратно.~~ 😃

Редактировать: потеря емкости значительна для сценария домашнего пользователя, где мы начинаем с малого. Это очень мало, когда вы начинаете с гораздо большего vdev, где соотношение данных к паритету намного эффективнее. В любом случае, потерянная способность может быть восстановлена в конце концов, и это то, что имеет наибольшее значение, я думаю.

Я также хочу извиниться за путаницу.

ДэйБлур Комментировать on 21 июн. 2021 г.

> Я хотел бы уделять больше внимания в школе, когда преподается базовая математика. Я думаю, что мой накладной расчет был очень неправильным.
>
> Я думаю, что эта электронная таблица показывает, что накладные расходы очень малы, хотя я был бы признателен, если бы кто-то проверил математику.

Я не мог получить доступ к вашей электронной таблице, но ваш первоначальный расчет выглядел так, как будто он был на поле для меня, основываясь на моем понимании. Я подсчитал, что 96% одной емкости диска теряется только от 80% полного 4-широкого RAID-Z2 непосредственно до 10-широкого RAID-Z2 без пересчета паритета. Я думаю, что постепенно расширяясь, как вы изначально описали, я думаю, что потеряет ~136% от одной емкости диска к тому времени, когда вы достигнете 10-й ширины. Это не является незначительным, поэтому я также был бы признателен за разъяснение, если это не так.

👍 1

Лоуэнриус Комментировать on 21 июн. 2021 г. • • отредактировано Участник

> Я хотел бы уделять больше внимания в школе, когда преподается базовая математика. Я думаю, что мой накладной расчет был очень неправильным.
> Я думаю, что эта электронная таблица показывает, что накладные расходы очень

> расчет выглядел так, как будто он был на поле для меня, основываясь на моем понимании. Я подсчитал, что 96% одной емкости диска теряется только от 80% полного 4-широкого RAID-Z2 непосредственно до 10-широкого RAID-Z2 без пересчета паритета. Я думаю, что постепенно расширяясь, как вы изначально описали, я думаю, что потеряет ~136% от одной емкости диска к тому времени, когда вы достигнете 10-й ширины. Это не маловажно, поэтому я также был бы признателен за разъяснение, если это не так.

Я думаю, вы абсолютно правы. Моя голова была в неправильном месте, и потеря действительно может быть незначительной. Но я думаю, что в расчете есть нюанс, который я не принял во внимание. Когда вы добавляете диски, новые данные хранятся «более эффективно», и я думаю, что я пропустил это в расчете.

~~Я обновил ссылки на лист, но лист в настоящее время неправильный, мне нужно исправить математику.~~

Я не уверен, что понимаю, как вы получили эти 96% и ~136%. (извините).

Редактировать:

~~Согласно листу, мой сценарий, в котором вы постепенно обновляетесь с 4-приводного RAIDZ2 до 10-приводного RAIDZ2, приведет к потере 1,09 дисков на емкость. Это предполагает, что вы обновляетесь на 100% мощности, что не так реалистично.~~

~~Если мы обновим мощность на 80%, потеря составит около 0,7 приводов. Это не незначительно, но это также не так уж и экстремально, что многие люди не смогли бы с этим жить, я бы сказал.~~

Редактировать 2: Накладные расходы достаточно значительны, чтобы люди, по крайней мере, знали об этом и понимали, насколько могут быть потери. Я думаю.

йорикдаун Комментировать on 21 июн. 2021 г.

Я скопировал вашу ссылку и сделал новую электронную таблицу: https://docs.google.com/spreadsheets/d/1zEyVHW3Fs84W4ogbDY5LziPvtAdJPOYU6wMBlaLgGOsA/edit? usp=шеринг

йорикдаун Комментировать on 21 июн. 2021 г. • • отредактировано

Я думаю, что у меня сейчас математика, и не стесняйтесь проверить по вашей оригинальной электронной таблице **@louwrentius.** Увеличение емкости от 9 до 11% в зависимости от входных значений выглядит правильно. «Из 4» — худший сценарий для рейза из-за низкой первоначальной эффективности. Если я делаю «с 6», то прирост мощности больше в диапазоне от 5 до 6%.

Примечание: мы используем PR в качестве дискуссионного форума. Мэтт был снисходителен, и мы должны взять это на их форумы?

Лоуэнриус Комментировать on 21 июн. 2021 г. • • отредактировано Участник

> Я думаю, что у меня сейчас математика, и не стесняйтесь проверить по вашей оригинальной электронной таблице **@louwrentius.** Увеличение емкости от 9 до 11% в зависимости от входных значений выглядит правильно. «Из 4» — худший сценарий для рейза из-за низкой первоначальной эффективности. Если я делаю «с 6», то прирост мощности больше в диапазоне от 5 до 6%.

~~Я не уверен, что математика правильная. Сценарий заключается в том, что вы добавляете один диск за раз. Если мы возьмем оригинальный пример и перейдем от 4 дисков RAIDZ2 к 10 приводу RAIDZ2, то вы в основном будете наблюдать, что при добавлении диска будут храниться данные на различных уровнях эффективности. Это не учитывается в ваших расчетах, и это заставляет вещи казаться хуже, чем они есть, я думаю, но, возможно, я ошибаюсь.~~

Математика правильная.

> Примечание: мы используем PR в качестве дискуссионного форума. Мэтт был снисходителен, и мы должны взять это на ix форумы?

Если эта дискуссия не будет проводиться в этом PR, я бы предложил продолжить обсуждение здесь, на Github, в «Обсуждении».

👍 1

Фмстрат Комментировать on 21 июн. 2021 г.

> Увеличение емкости от 9 до 11% в зависимости от входных значений выглядит правильно. «Из 4» — худший сценарий для рейза из-за низкой первоначальной эффективности. Если я делаю «с 6», то прирост мощности больше в диапазоне от 5 до 6%.

@yorickdowne Я не понимаю. ;) Не должна ли увеличение мощности быть больше в ситуации «от 6 до 8»? Означает ли «прирост способности» что-то странное в этом контексте? Т.е. полная потеря мощности?

> Вот почему я надеюсь, что технически возможно использовать существующее свободное пространство на VDEV, которое было расширено для переписывания данных. Это гарантирует, что пользователям не нужно добавлять (временные) устройства хранения и вручную перемещать данные.
>
> Идея использования SLOG или какого-либо другого устройства (идеально SSD?) для временного хранения данных, чтобы ускорить процесс - если это имеет смысл - звучит интересно для меня (как noob).

Начнем с этого:

```
# The initial array is built
zpool create -o ashift=12 -O atime=off -O mountpoint=none s raidz2 <device> <device> <d
zfs create -o encryption=on -o keyformat=hex -o keylocation=file:///etc/zfskey -o compr
zfs create -o mountpoint=/storage/1 s/storage/1
# Over time, user creates a bunch of files until 90% of s/storage is taken up.
for n in {1..1000}; do
    dd if=/dev/urandom of=file$( printf %03d "$n" ).bin bs=1 count=$(( RANDOM + 1024 ))
done
```

Вариант 1: Отдельный vdev
Примечание: `find` команда полностью не проверена и, вероятно, не работает, но вы получите идею: перемещение 1 файла за раз, поэтому «свободное» пространство включено `s/storage` остается неизменным во время миграции минус размер одного файла.

```
zfs create -o mountpoint=/storage/2  s/storage/2
cd /storage/1
find . -exec mv {} /storage/2/{} \;
zfs destroy s/storage/1
```

Будет ли это «переписано» блоков?

Вариант 2: Кэш-устройство, по одному файлу за раз
Похожий на выше, но вместо того, чтобы переходить от одного vdev к другому, вы перемещаете файл на другое устройство, а затем перемещаете его обратно. Конечно, для этого потребуется создание каталогов/и т.д. Это просто для того, чтобы показать концепцию.

```
cd /storage/1
for F in $(find .); do
    mv "${F}" "/mnt/tmp/${F}"
    mv "/mnt/tmp/${F}" "${F}"
done
```

Как насчет этого?

👍 1

Гипфер Комментировать on 21 июн. 2021 г.

> Должны ли мы взять это на ix форумы?

ДА

👍 14

йорикдаун Комментировать on 21 июн. 2021 г.

Ссылка на форум TrueNAS: https://www.truenas.com/community/threads/raidz-expansion-its-happening.58575/page-10#post-648933

👍 7

Алкерин Комментировать on 23 июн. 2021 г. • • отредактировано

Позволит ли это также уменьшить количество приводов в рейзах?
также, как насчет «изменения» уровня рейза ex revestigz2 -> raignz3 с новым приводом или рейдz2 -> рейвz1

Тем не менее, даже если он не добавляет эти функции, спасибо за работу, так как это одна особенность, которую многие люди ждали.

👎 2 😕 1

ГурлиДжебис Комментировать on 23 июн. 2021 г.

> Позволит ли это также уменьшить количество приводов в рейзах?
> также, как насчет «изменения» уровня рейза ex revestigz2 -> raignz3 с новым приводом или рейдz2 -> рейвz1
>
> Тем не менее, даже если он не добавляет эти функции, спасибо за работу, так как это одна особенность, которую многие люди ждали.

Я думаю, что перебалансировка потребуется, прежде чем расти избыточность. Поскольку это гарантирует, что все данные будут разделены по всем дискам

ИванВоложюк Упомянули эту просьбу о вытягива on 25 июн. 2021 г.

Запрос на функцию: zfs команда/собственность ребалансировки #12267

Открыть

Эндрю-Джадр Комментировать on 8 авг. 2021 г.

Работает ли это и на dRAID vdevs? #10102

Юрикдаун Комментировать on 8 авг. 2021 г.

Это не так. Учитывая размер и вариант использования драги, является ли драповое расширение приоритетной особенностью, это хороший вопрос, который будет обсуждаться, возможно, на Reddit ZFS или других социальных сетях ZFS.

сёданшок Комментировать on 8 авг. 2021 г. Участник

> Это не так. Учитывая размер и вариант использования драги, является ли драповое расширение приоритетной особенностью, это хороший вопрос, который будет обсуждаться, возможно, на Reddit ZFS или других социальных сетях ZFS.

Имея как список рассылки, так и проблемы/обсуждения GitHub, я бы предложил использовать эти каналы, а не фрагментировать информацию по различным социальным сетям.

Только мои 2 цента

Соковыслi Комментировать on 13 авг. 2021 г.

> Это не так. Учитывая размер и вариант использования драги, является ли драповое расширение приоритетной особенностью, это хороший вопрос, который будет обсуждаться, возможно, на Reddit ZFS или других социальных сетях ZFS.
>
> Имея как список рассылки, так и проблемы/обсуждения GitHub, я бы предложил использовать эти каналы, а не фрагментировать информацию по различным социальным сетям.
>
> Только мои 2 цента

Действительно. Все, что на Reddit, будет дезорганизовано и потеряно во времени в течение нескольких месяцев.

Эндрю-Джадр Комментировать on 14 авг. 2021 г.

Кажется ли разумным создать вопрос «расширения dRAID», чтобы где-то централизовать дискуссию?

Соковыслi Упомянули эту просьбу о вытягива on 27 авг. 2021 г.

Возможность запуска/триггерного сжатия/дедупликации пула/объема вручную #3013

Сомест Комментировать on 16 сент. 2021 г.

Похоже, что первое обязательство в этом запросе на вытягивание сделало его в (`069bf40`) и это должно быть переосмыслено

👍 5

кокомано1 Комментировать on 30 сент. 2021 г. • • Отредактировано ahrens

Я сделал первое расширение, и это было успешно, затем я попробовал другое, но это висит в конце с этим

```
pool: test
state: ONLINE
scan: resilvered 0B in 00:00:01 with 0 errors on Tue Sep 21 03:16:11 2021
raidz expand: Expansion of vdev 0 in progress since Mon Sep 27 14:46:40 2021
1.00T copied out of 1.00T at 5.23M/s, 100.00% done, 0h0m to go
config:

NAME        STATE     READ WRITE CKSUM
test        ONLINE       0     0     0
raidz2-0  ONLINE       0     0     0
sda     ONLINE       0     0     0
sdb     ONLINE       0     0     0
sdc     ONLINE       0     0     0
sdf     ONLINE       0     0     0
sde     ONLINE       0     0     0
sdg     ONLINE       0     0     0

errors: No known data errors
```

а затем висит, когда я импортирую его с этой ошибкой

```
VERIFY3(c < SPA_MAXBLOCKSIZE >> SPA_MINBLOCKSHIFT) failed (36028797018963967 <
32768)
[Wed Sep 29 22:13:28 2021] PANIC at zio.c:322:zio_buf_alloc()
[Wed Sep 29 22:13:28 2021] Showing stack for process 2437
[Wed Sep 29 22:13:28 2021] CPU: 1 PID: 2437 Comm: z_wr_int Tainted: P           OE
5.11.0-37-generic #41-Ubuntu
[Wed Sep 29 22:13:28 2021] Call Trace:
[Wed Sep 29 22:13:28 2021]  show_stack+0x52/0x58
[Wed Sep 29 22:13:28 2021]  dump_stack+0x70/0x8b
[Wed Sep 29 22:13:28 2021]  spl_dumpstack+0x29/0x2b [spl]
[Wed Sep 29 22:13:28 2021]  spl_panic+0xd4/0xfc [spl]
[Wed Sep 29 22:13:28 2021]  ? _cond_resched+0x1a/0x50
[Wed Sep 29 22:13:28 2021]  ? _cond_resched+0x1a/0x50
[Wed Sep 29 22:13:28 2021]  ? __kmalloc_node+0x144/0x2b0
[Wed Sep 29 22:13:28 2021]  ? spl_kvmalloc+0x82/0xb0 [spl]
[Wed Sep 29 22:13:28 2021]  zio_buf_alloc+0x5e/0x60 [zfs]
[Wed Sep 29 22:13:28 2021]  abd_borrow_buf_copy+0x6e/0xa0 [zfs]
```

```
[Wed Sep 29 22:13:28 2021]  zio_vsd_default_cksum_finish+0x147/0x20 [zfs]
[Wed Sep 29 22:13:28 2021]  zio_done+0x2fb/0x11b0 [zfs]
[Wed Sep 29 22:13:28 2021]  zio_execute+0x8b/0x130 [zfs]
[Wed Sep 29 22:13:28 2021]  taskq_thread+0x2b7/0x500 [spl]
[Wed Sep 29 22:13:28 2021]  ? wake_up_q+0xa0/0xa0
[Wed Sep 29 22:13:28 2021]  ? zio_gang_tree_free+0x70/0x70 [zfs]
[Wed Sep 29 22:13:28 2021]  kthread+0x11f/0x140
[Wed Sep 29 22:13:28 2021]  ? taskq_thread_spawn+0x60/0x60 [spl]
[Wed Sep 29 22:13:28 2021]  ? set_kthread_struct+0x50/0x50
[Wed Sep 29 22:13:28 2021]  ret_from_fork+0x22/0x30
```

Его можно установить с

```
zpool import -o readonly=on test
```

Есть идеи о том, что делать? thx

Я пытался отсчитывать данные, но они не работают

```
./zpool status -v
  pool: test
 state: ONLINE
status: One or more devices has experienced an error resulting in data
        corruption.  Applications may be affected.
action: Restore the file in question if possible.  Otherwise restore the
        entire pool from backup.
   see: https://openzfs.github.io/openzfs-docs/msg/ZFS-8000-8A
  scan: resilvered 0B in 00:00:01 with 0 errors on Tue Sep 21 03:16:11 2021
raidz expand: Expansion of vdev 0 in progress since Mon Sep 27 14:46:40 2021
    1.00T copied out of 1.00T at 4.00M/s, 100.00% done, 0h0m to go
config:

        NAME        STATE     READ WRITE CKSUM
        test        ONLINE       0     0     0
          raidz2-0  ONLINE       0     0     0
            sda     ONLINE       0     0   887
            sdb     ONLINE       0     0   825
            sdc     ONLINE       0     0   887
            sdf     ONLINE       0     0   879
            sde     ONLINE       0     0   876
            sdh     ONLINE       0     0     0

errors: Permanent errors have been detected in the following files:
list of files with errors
```

👀 1

Небес Волков Комментировать on 10 окт. 2021 г.

против: https://github.com/openzfs/zfs/tree/zfs-2.1.1

```
--- a/module/zfs/vdev_raidz.c   2021-10-09 23:46:44.565994410 -0300
+++ b.module/zfs/vdev_raidz.c   2021-10-09 23:47:28.783521961 -0300
@@ -4189,7 +4189,7 @@
 {
        ASSERT3P(spa->spa_raidz_expand_zthr, ==, NULL);
        spa->spa_raidz_expand_zthr = zthr_create("raidz_expand",
-           spa_raidz_expand_cb_check, spa_raidz_expand_cb, spa);
+           spa_raidz_expand_cb_check, spa_raidz_expand_cb, spa, defclsyspri);
 }

 void
```

👍 2

Архренс Комментировать on 14 окт. 2021 г. Член Автор

@kocoman1 Спасибо за сообщение об этой проблеме и сожалею о задержке в возвращении к вам. Можете ли вы попробовать воспроизвести с помощью отладки? Это должно поймать ошибку немного раньше, ближе к тому, где это происходит. Также, пожалуйста, соберите dbgmsg выход (от `/proc/spl/kstat/zfs/dbgmsg`) Если вы можете создать свалку, когда происходит сбой утверждения, это также будет полезно.

Архренс Силовой толчок the `рейнц-расширение` филиал от **`9f49205`** к **`23ba99c`** Сравнить
4 года назад

Архренс Комментировать on 24 окт. 2021 г. Член Автор

Я переосмыслил это, удалив одну из уже приземленных коммитов.

👍 5 🎉 1

Сомест Комментировать on 10 нояб. 2021 г.

Пара логистических вопросов - запрос на вытягивание в настоящее время имеет три патча,

`249f28b` "Отключить zstd mempool"
`7def1f9` "Функция расширения рейда"
`23ba99c` "слиться"

- "Отключить zstd mempool" - это "первая обязанность" удалить, упомянутая в TODO?

- "Немного последних тестовых провалов" - у вас случайно есть снимок текущего состояния тестов?

Архренс Комментировать on 10 нояб. 2021 г. Член Автор

@emaste

> "Отключить zstd mempool" - это "первая обязанность" удалить, упомянутая в TODO?

Правильно. Это обязательство делает так, что мы можем проверить как (AddressSaniter), не поражая несвязанных сбоев. Я использовал его, чтобы отследить ошибку без использования после использования, которая, я думаю, была решена.

> «Слияние» выглядит так, как будто оно исправляет типы, предупреждения и т.д., и должно быть выдавлено в основной коммит?

Правильно, это исправление некоторых ошибок, введенных, когда я перестраиваюсь. Я раздавлю его до того, как это будет интегрировано.

> "Немного последних тестовых провалов" - у вас случайно есть снимок текущего состояния тестов?

Не сейчас, но я планирую добраться до этого до конференции FreeBSD на следующей неделе.

👍 3

Соковысли Комментировать on 11 нояб. 2021 г. • • отредактировано

В общем, каково состояние этого патсета? Сколько грубых углов / тестов / ошибок необходимо исправить, прежде чем можно подумать о слиянии его с основной линией?

Я не хочу показаться нетерпеливым (действительно, с работой в файловой системе терпение является ключом к тому, чтобы не повредить файловые системы пользователей). Я просто пытаюсь понять, где обстоят дела.

👍 5

 `disable zstd mempool` `dda7bf2`

Архренс Комментировать on 16 нояб. 2021 г. Член Автор

@Matthew-Bradley Я не думаю, что есть слишком много грубых краев, все должно работать. Я немного очистлю код и надеюсь, что люди начнут просматривать и

Архренс Силовой толчок the `рейнц-расширение` филиал от **`23ba99c`** к **`fdcba17`** Сравнить
4 года назад

Архренс и другие добавлено 8 коммитов 4 года назад

`raidz expansion feature` ⋯	`2f2ae11`
`Increase the number of txgs added to last reflow complete sync txg` ⋯	`c64244b`
`panic in zthr_iscancelled`	`c32cc85`
`ztest: Skip ztest_vdev_LUN_growth() if raidz expansion is in-progress`	`8cb83b7`
`ztest: Make ztest_vdev_raidz_attach() error checking more closer to z…`	`50e8dae`
⋯	
`ztest: Restore scratch object testing`	`16a74ed`
`ztest: Add raidz expansion testing as CLI option`	`05a8136`
`ztest: Fix integer printing`	`2020f19`

Архренс Силовой толчок the `рейнц-расширение` филиал от **`fdcba17`** к **`2020f19`** Сравнить
4 года назад

Соковысли Комментировать on 19 нояб. 2021 г. • • отредактировано

Как обрабатывается отказ устройства во время расширения radz? Правильно ли себя ведет себя, если в качестве замены используется привод другого размера? (больше или меньше)

Я, по крайней мере, не вижу, чтобы «трюк изменения размера массива» происходил во время расширения, так как вам нужно было бы заменить все диски для этого. Либо все старые диски уже были бы нового размера (и, таким образом, уже расширены), либо все еще остались бы некоторые из старых размеров.

Архренс Комментировать on 19 нояб. 2021 г. Член Автор

@Matthew-Bradley

> Как обрабатывается отказ устройства во время расширения radz?

Расширение будет приостановлено до тех пор, пока бассейн снова не станет здоровым

> Правильно ли себя ведет себя, если в качестве замены используется привод другого размера? (больше или меньше)

Применяются обычные правила замены привода, независимо от того, происходит ли расширение рейза: новый диск должен быть такого же размера или больше.

> Я, по крайней мере, не вижу, чтобы «трюк изменения размера массива» происходил во время расширения, так как вам нужно было бы заменить все диски для этого. Либо все старые диски уже были бы нового размера (и, таким образом, уже расширены), либо все еще остались бы некоторые из старых размеров.

Я не понимаю, что такое "трюк с изменением размера мозга"?

Соковыслu Комментировать on 19 нояб. 2021 г. • • отредактировано

> Я не понимаю, что такое "трюк с изменением размера мозга"?

Многие люди, по-видимому, расширяют свой массив без добавления дисков, заменяя каждый из дисков один за другим более крупной моделью.
https://serverfault.com/questions/15290/how-to-upgrade-a-zfs-raid-z-array-to-larger-disks-on-opensolaris

Я, кажется, помню эмпирическое правило о том, чтобы не выбирать диски, которые примерно в 10x больше, чем первые диски, используемые для создания массива.

Архренс Комментировать on 19 нояб. 2021 г. Член Автор

@Matthew-Bradley Да, это работает. Эта техника является ортогональной для этого пиара.

Соковыслu Комментировать on 19 нояб. 2021 г.

Я упоминаю об этом только потому, что мне было интересно, могут ли такие махинации повлиять на расширение, если они попытаются.

кокомано1 Комментировать on 23 нояб. 2021 г. • • отредактировано

редактирование снова: см. обновленный в конце `Пчела78бб` (не знаю, как сделать новый ответ на мой пост)

> **@kocoman1** Спасибо за сообщение об этой проблеме и сожалею о задержке в возвращении к вам. Можете ли вы попробовать воспроизвести с помощью отладки? Это должно поймать ошибку немного раньше, ближе к тому, где это происходит.

Спасибо за ответ..

Я скачиваю ваш последний сбор 6 дней назад

в dbgmsg говорит

```
1637638370   ffff9b8a0225c500 vdev_raidz.c:3573:raidz_reflow_write_done(): comple
reflow offset=297071726592 size=4096 txg
1637638370   ffff9b8a0d830000 vdev_raidz.c:3717:raidz_reflow_impl(): initiating
reflow write offset=297081155584 length=4096
```

(продолжает идти)

```
zpool status
pool:
state: ONLINE
status: One or more devices has experienced an unrecoverable error.  An
        attempt was made to correct the error.  Applications are unaffected.
action: Determine if the device needs to be replaced, and clear the errors
        using 'zpool clear' or replace the device with 'zpool replace'.
   see: https://openzfs.github.io/openzfs-docs/msg/ZFS-8000-9P
  scan: resilvered 0B in 00:00:01 with 0 errors on Tue Sep 21 03:16:11 2021
raidz expand: Expansion of vdev 0 in progress since Mon Sep 27 14:46:40 2021
    1.00T copied out of 1.00T at 221K/s, 100.22% done, (copy is slow, no estimated
time)
config:

        NAME        STATE     READ WRITE CKSUM
        expa        ONLINE       0     0     0
          raidz2-0  ONLINE       0     0     0
            sde     ONLINE       0     0     8
            sdf     ONLINE       0     0     7
            sdb     ONLINE       0     0     7
            sda     ONLINE       0     0     7
            sdd     ONLINE       0     0     7
            sdg     ONLINE       0     0     0

errors: No known data errors
```

не знаю, что сейчас делает, прервочный тоже неправ
Просто оставлю его включенным и подожди...

Так что это все, что я знаю, как запечатлеть на данный момент

```
VERIFY3(0 == metaslab_load(msp)) failed (0 == 52)
[Tue Nov 23 01:06:42 2021] PANIC at vdev_raidz.c:4143:spa_raidz_expand_cb()
[Tue Nov 23 01:06:42 2021] Showing stack for process 2811
[Tue Nov 23 01:06:42 2021] CPU: 0 PID: 2811 Comm: raidz_expand Tainted: P
OE     5.11.0-38-generic #42-Ubuntu
[Tue Nov 23 01:06:42 2021] Call Trace:
[Tue Nov 23 01:06:42 2021]  show_stack+0x52/0x58
```

```
[Tue Nov 23 01:06:42 2021]  ? _cond_resched+0x1a/0x50
[Tue Nov 23 01:06:42 2021]  ? do_raw_spin_unlock+0x9/0x10 [zfs]
[Tue Nov 23 01:06:42 2021]  ? __raw_spin_unlock+0x9/0x10 [zfs]
[Tue Nov 23 01:06:42 2021]  ? metaslab_load_impl+0x2a3/0x4d0 [zfs]
[Tue Nov 23 01:06:42 2021]  ? kfree+0x3ce/0x400
[Tue Nov 23 01:06:42 2021]  ? totalram_pages+0x10/0x20 [zfs]
[Tue Nov 23 01:06:42 2021]  ? arc_all_memory+0xe/0x20 [zfs]
[Tue Nov 23 01:06:42 2021]  ? metaslab_potentially_evict+0x4b/0x270 [zfs]
[Tue Nov 23 01:06:42 2021]  spa_raidz_expand_cb+0x567/0x640 [zfs]
[Tue Nov 23 01:06:42 2021]  zthr_procedure+0x13b/0x150 [zfs]
[Tue Nov 23 01:06:42 2021]  ? spl_mutex_clear_owner+0x10/0x10 [zfs]
[Tue Nov 23 01:06:42 2021]  thread_generic_wrapper+0x79/0x90 [spl]
[Tue Nov 23 01:06:42 2021]  kthread+0x11f/0x140
[Tue Nov 23 01:06:42 2021]  ? __thread_exit+0x20/0x20 [spl]
[Tue Nov 23 01:06:42 2021]  ? set_kthread_struct+0x50/0x50
[Tue Nov 23 01:06:42 2021]  ret_from_fork+0x22/0x30
```

```
pool:
state: ONLINE
status: One or more devices has experienced an error resulting in data
        corruption.  Applications may be affected.
action: Restore the file in question if possible.  Otherwise restore the
        entire pool from backup.
   see: https://openzfs.github.io/openzfs-docs/msg/ZFS-8000-8A
  scan: resilvered 0B in 00:00:01 with 0 errors on Tue Sep 21 03:16:11 2021
raidz expand: Expansion of vdev 0 in progress since Mon Sep 27 14:46:40 2021
    1.10T copied out of 1.00T at 241K/s, 109.50% done, (copy is slow, no estimated
time)
config:

        NAME        STATE     READ WRITE CKSUM
        expa        ONLINE       0     0     0
          raidz2-0  ONLINE       0     0     0
            sde     ONLINE       0     0   642
            sdf     ONLINE       0     0   668
            sdb     ONLINE       0     0   640
            sda     ONLINE       0     0   628
            sdd     ONLINE       0     0   642
            sdg     ONLINE       0     0     0

errors: x data errors, use '-v' for a list
```

журнал /proc/spl/kstat/zfs/dbgmsg

паста.txt

модуль паника, так что я не знаю, как получить сброс краш-дана для этого

Я попробовал патч, который вы положили 1 час назад bee78ee

команда импорта пула не возвращается к снаряду
Паника, замеченная в dmesg

Другая паника

VERIFY3(0 == metaslab_load(msp)) потерпел неудачу (0 == 52)
[Ср Ноя 24 12:57:28 2021] ПАНИКЯ на vdev_raidz.c:4160:spa_raidz_expand_cb()
[Ср Ноя 24 12:57:28 2021] Показать стек для процесса 2459
[Ср Ноя 24 12:57:28 2021] ЦП: 3 PID: 2459 Комплект: rrawl_expand Испорченный: P OE
5.11.0-38-generic #42 -Ubuntu
[Ср 24 Ноя 12:57:28 2021] Трейс:
[Ср Ноя 24 12:57:28 2021] show_stack+0x52/0x58
[Ср Ноя 24 12:57:28 2021] dump_stack+0x70/0x8b
[Ср Ноя 24 12:57:28 2021] spl_dumpstack+0x29/0x2b [spl]
[Ср 24 ноя 12:57:28 2021] spl_panic+0xd4/0xfc [spl]
[Ср Ноя 24 12:57:28 2021] ? _cond_resched+0x1a/0x50
[Ср Ноя 24 12:57:28 2021] ? do_raw_spin_unlock+0x9/0x10 [zfs]
[Ср Ноя 24 12:57:28 2021] ? __raw_spin_unlock+0x9/0x10 [zfs]
[Ср Ноя 24 12:57:28 2021] ? metaslab_load_impl+0x2a3/0x4d0 [zfs]
[Ср Ноя 24 12:57:28 2021] ? обновление_load_load_avg+0x82/0x5d0
[Ср Ноя 24 12:57:28 2021] ? totalram_pages+0x10/0x20 [zfs]
[Ср Ноя 24 12:57:28 2021] ? arc_all_memory+0xe/0x20 [zfs]
[Ср Ноя 24 12:57:28 2021] ? metaslab_potentially_evict+0x4b/0x270 [zfs]
[Ср Ноя 24 12:57:28 2021] spa_raidz_expand_cb+0x567/0x640 [zfs]
[Ср Ноя 24 12:57:28 2021] zthr_procedure+0x13b/0x150 [zfs]
[Ср Ноя 24 12:57:28 2021] ? spl_mutex_clear_owner+0x10/0x10 [zfs]
[Ср Ноя 24 12:57:28 2021] thread_generic_wrapper+0x79/0x90 [spl]
[Ср Ноя 24 12:57:28 2021] Хтлест-свет+0x11ф/0x140
[Ср Ноя 24 12:57:28 2021] ? __thread_exit+0x20/0x20 [spl]
[Ср Ноя 24 12:57:28 2021] ? set_kthread_struct+0x50/0x50
[Ср Ноя 24 12:57:28 2021] ret_from_fork+0x22/0x30
[Ср Ноя 24 13:00:18 2021] ИНФОРМАЦИЯ: задание zpool:2217 заблокировано более 120
секунд.
[Ср Ноя 24 13:00:18 2021] Испорченный: P OE 5.11.0-38-generic #42 -Ubuntu
[Ср Ноя 24 13:00:18 2021] "Эхо 0 > /proc/sys/kernel/hung_task_timeout_secs" отключает это
сообщение.
[Ср Ноя 24 13:00:18 2021] задание:zpool состояние:D stack: 0 pid: 2217 pid: 2216 флаги:
0x00004006
[Ср Ноя 24 13:00:18 2021] Трейс Позвонить:
[Ср Ноя 24 13:00:18 2021] __schedule+0x23d/0x670
[Ср Ноя 24 13:00:18 2021] расписание+0x4f/0xc0
[Ср Ноя 24 13:00:18 2021] cv_wait_common+0xf4/0x130 [spl]
[Ср 24 Ноя 13:00:18 2021] ? wait_woken+0x80/0x80
[Ср Ноя 24 13:00:18 2021] __cv_wait+0x15/0x20 [spl]
[Ср Ноя 24 13:00:18 2021] spa_config_enter+0x72/0x100 [zfs]
[Ср Ноя 24 13:00:18 2021] spa_import+0x20f/0x620 [zfs]

[Ср 24 Ноя 13:00:18 2021] ? __check_object_size.part.0+0x4a/0x150
[Ср Ноя 24 13:00:18 2021] zfsdev_ioctl+0x57/0xe0 [zfs]
[Ср Ноя 24 13:00:18 2021] __x64_sys_ioctl+0x91/0xc0
[Ср Ноя 24 13:00:18 2021] do_syscall_64+0x38/0x90
[Ср Ноя 24 13:00:18 2021] запись_SYSCALL_64_after_hwframe+0x44/0xa9
[Ср Ноя 24 13:00:18 2021] RIP: 0033:0x7fa38b2c4ecb
[Ср Ноя 24 13:00:18 2021] РСП: 002b:0000709094955918 ФЛАГИ: 00000246 ORIG_RAX: 00000000000010
[Ср Ноя 24 13:00:18 2021] РАКС: ffffffffffffda RBX: 000000000000 RCX: 00007fa38b2c4ecb
[Ср Ноя 24 13:00:18 2021] RDX: 00007ffdc4955990 RSI: 00000000005a02 RDI: 00000000000003
[Ср Ноя 24 13:00:18 2021] РБП: 00007ффдк4959880 R08: 000000000000 R09: 000055f373115af0
[Ср Ноя 24 13:00:18 2021] R10: 000000000045c0 R11: 0000000000000246 R12: 000055f3730b3570
[Ср Ноя 24 13:00:18 2021] R13: 00007ffdc4955990 R14: 00007fa3840200 R15: 000000000000
[Ср 24 ноября 13:00:18 2021] ИНФОРМАЦИЯ: задание txg_sync:2457 заблокировано более чем на 120 секунд.
[Ср Ноя 24 13:00:18 2021] Испорченный: P OE 5.11.0-38-generic #42 -Ubuntu
[Ср Ноя 24 13:00:18 2021] "Эхо 0 > /proc/sys/kernel/hung_task_timeout_secs" отключает это сообщение.
[Ср Ноя 24 13:00:18 2021] задание:txg_sync состояние: D stack: 0pid: 2457 pid: 2 флага: 0x00004000
[Ср Ноя 24 13:00:18 2021] Трейс Позвонить:
[Ср Ноя 24 13:00:18 2021] __schedule+0x23d/0x670
[Ср Ноя 24 13:00:18 2021] расписание+0x4f/0xc0
[Ср Ноя 24 13:00:18 2021] cv_wait_common+0xf4/0x130 [spl]
[Ср 24 Ноя 13:00:18 2021] ? wait_woken+0x80/0x80
[Ср Ноя 24 13:00:18 2021] __cv_wait+0x15/0x20 [spl]
[Ср Ноя 24 13:00:18 2021] spa_config_enter+0xe8/0x100 [zfs]
[Ср Ноя 24 13:00:18 2021] spa_txg_history_init_io+0x5e/0xd0 [zfs]
[Ср Ноя 24 13:00:18 2021] txg_sync_thread+0x211/0x290 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? txg_dispatch_callbacks+0xf0/0xf0 [zfs]
[Ср Ноя 24 13:00:18 2021] thread_generic_wrapper+0x79/0x90 [spl]
[Ср Ноя 24 13:00:18 2021] Хтрейд+0x11ф/0x140
[Ср 24 Ноя 13:00:18 2021] ? __thread_exit+0x20/0x20 [spl]
[Ср 24 Ноя 13:00:18 2021] ? set_kthread_struct+0x50/0x50
[Ср Ноя 24 13:00:18 2021] ret_from_fork+0x22/0x30
[Ср 24 ноября 13:00:18 2021] ИНФОРМАЦИЯ: задание рейдz_expand:2459 заблокировано более чем на 120 секунд.
[Ср Ноя 24 13:00:18 2021] Испорченный: P OE 5.11.0-38-generic #42 -Ubuntu
[Ср Ноя 24 13:00:18 2021] "Эхо 0 > /proc/sys/kernel/hung_task_timeout_secs" отключает это сообщение.
[Ср 24 ноя 13:00:18 2021] задание:raidz_expand состояние:D stack: 0 pid: 2459 pid: 2 флага: 0x00004000
[Ср Ноя 24 13:00:18 2021] Трейс Позвонить:
[Ср Ноя 24 13:00:18 2021] __schedule+0x23d/0x670
[Ср 24 Ноя 13:00:18 2021] ? арк_локальный_irq_restore+0x6/0xd
[Ср Ноя 24 13:00:18 2021] расписание+0x4f/0xc0

[Ср 24 Ноя 13:00:18 2021] ? do_raw_spin_unlock+0x9/0x10 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? __raw_spin_unlock+0x9/0x10 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? metaslab_load_impl+0x2a3/0x4d0 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? обновление_load_load_avg+0x82/0x5d0
[Ср 24 Ноя 13:00:18 2021] ? totalram_pages+0x10/0x20 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? arc_all_memory+0xe/0x20 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? metaslab_potentially_evict+0x4b/0x270 [zfs]
[Ср Ноя 24 13:00:18 2021] spa_raidz_expand_cb+0x567/0x640 [zfs]
[Ср Ноя 24 13:00:18 2021] zthr_procedure+0x13b/0x150 [zfs]
[Ср 24 Ноя 13:00:18 2021] ? spl_mutex_clear_owner+0x10/0x10 [zfs]
[Ср Ноя 24 13:00:18 2021] thread_generic_wrapper+0x79/0x90 [spl]
[Ср Ноя 24 13:00:18 2021] Хтрейд+0x11ф/0x140
[Ср 24 Ноя 13:00:18 2021] ? __thread_exit+0x20/0x20 [spl]
[Ср 24 Ноя 13:00:18 2021] ? set_kthread_struct+0x50/0x50
[Ср Ноя 24 13:00:18 2021] ret_from_fork+0x22/0x30

122 скрытых предмета

Загружать больше...

justinclift Комментировать on 4 мар. 2023 г. • • отредактировано

> Это вызывает потерю данных для меня (zfs паника)

@kocoman1 Это на Linux, или вы получили его для компиляции на macOS? 😄

👍 1

AT-Стивен Детомаси Комментировать on 4 мар. 2023 г.

> Эта функция была объединена или нет?!

Ответить на этот вопрос прямо - нет. Требуется больше работы и тестирования. Я бы сказал, что у нас как минимум 12 месяцев. Вернитесь в марте 2024 года и посмотрите, как идут дела.

👍 10 😕 9 👀 5

АльтерX76 Комментировать on 4 мар. 2023 г.

👍 6

кокомано1 Комментировать on 5 мар. 2023 г.

> Это вызывает потерю данных для меня (zfs паника)
>
> **@kocoman1** Это на Linux, или вы получили его для компиляции на macOS? 😄

Я больше не пытался в osx (это было в Linux) после того, как я даже не мог получить много из того, какие файлы я потерял. Sas2008 работает с внешним ксвоим на ventura ok. Я добавил extfs/apfs, но он перезагружается иногда во время нагрузки. не уверен, где баг, плюс unmount занимает так много времени, что он паникует, когда я выключаю машину в osx

плюс у меня есть куча STDM3000 и 2.5 4000 серии SMR, которые умирают.

Рейн Комментировать on 12 мар. 2023 г.

Так что я хотел бы расширение RaidZ для личного проекта, и я надеюсь попробовать его. Я вытащил пиар-отдел и попытался построить из источника, но филиал был слишком устаревшим, чтобы строить против моей текущей версии Ubuntu (самый последний публичный релиз).

Поэтому я рассматриваю следующее:

1. Установка старой версии Linux на запасной жесткий диск и создание ZFS из источника
2. Экспорт моего массива ZFS из моей текущей ОС и импорт его на новой
3. Расширение массива на новой ОС с помощью моей сборки из источника
4. Перенос массива после расширения обратно в мою текущую ОС Linux.

Но я упускаю контекст, чтобы узнать, хорошая ли это идея. Итак, несколько вопросов:

- На какой уровень риска я смотрю? Произошли ли какие-либо изменения в схеме экспортируемых массивов, о которых я должен беспокоиться? Есть ли причина, по которой я *не должен* этого делать?
- Какую версию Linux вы бы порекомендовали мне перейти / наиболее стабильно в этой отрасли для моего эксперимента?
- Есть ли какие-либо дополнительные флаги, которые я должен добавить или отладить функции, которые я должен включить при строительстве / запуске, чтобы помочь обеспечить полезный отчет об ошибке, если он потерпит неудачу?

У меня нет свободного времени, чтобы забрать новую кодовую базу, чтобы помочь с перебазированием ветки, поэтому я мог бы стать подопытным кроликом в следующий раз, когда у меня будет свободное время.

> - На какой уровень риска я смотрю? Произошли ли какие-либо изменения в схеме экспортируемых массивов, о которых я должен беспокоиться? Есть ли причина, по которой я *не должен* этого делать?

Уровень риска прямо пропорционален вашей способности восстановить данные из резервной копии.

👍 10 👎 2 😄 5 ❤️ 3 🚀 1

Фмульдер007 Комментировать on 27 мар. 2023 г.

Я начинающий пользователь FreeBSD и не так давно начал использовать ZFS для своих экспериментов. С нетерпением жду появления этого функционала! Большое спасибо за вашу работу!

👍 9 👎 1 ❤️ 5

кокомано1 Комментировать on 29 мар. 2023 г.

> Так что я хотел бы расширение RaidZ для личного проекта, и я надеюсь попробовать его. Я вытащил пиар-отдел и попытался построить из источника, но филиал был слишком устаревшим, чтобы строить против моей текущей версии Ubuntu (самый последний публичный релиз).
>
> Поэтому я рассматриваю следующее:
>
> 1. Установка старой версии Linux на запасной жесткий диск и создание ZFS из источника
> 2. Экспорт моего массива ZFS из моей текущей ОС и импорт его на новой
> 3. Расширение массива на новой ОС с помощью моей сборки из источника
> 4. Перенос массива после расширения обратно в мою текущую ОС Linux.
>
> Но я упускаю контекст, чтобы узнать, хорошая ли это идея. Итак, несколько вопросов:
>
> - На какой уровень риска я смотрю? Произошли ли какие-либо изменения в схеме экспортируемых массивов, о которых я должен беспокоиться? Есть ли причина, по которой я *не должен* этого делать?
> - Какую версию Linux вы бы порекомендовали мне перейти / наиболее стабильно в этой отрасли для моего эксперимента?
> - Есть ли какие-либо дополнительные флаги, которые я должен добавить или отладить функции, которые я должен включить при строительстве / запуске, чтобы помочь обеспечить полезный отчет об ошибке, если он потерпит неудачу?
>
> У меня нет свободного времени, чтобы забрать новую кодовую базу, чтобы помочь с

У меня была потеря данных после расширения его как 5 раз (расширение, копирование некоторых файлов из тех, чтобы стереть и расширить ... и т. Д.), Затем, когда я читаю данные, я получаю Zfs паника и любые ls и т. Д. Просто висит, я смог установить считывание только, но выполнение ls в некоторых каталогах приводит просто ошибка. копирование обратно получает контрольную сумму и ошибку прерывания ... так что это не готово, если они не исправили.

👍 1 👎 4

бахп Комментировать on 29 мар. 2023 г.

@kocoman1 Не могли бы вы предоставить более подробное описание ошибки? Это может помочь точно определить проблему.

Как минимум:

- Какие шаги вы сделали именно, это может помочь воспроизвести проблему.
- Каковы точные сообщения об ошибках, которые вы видите.

кокомано1 Комментировать on 29 мар. 2023 г. • • отредактировано

> **@kocoman1** Не могли бы вы предоставить более подробное описание ошибки? Это может помочь точно определить проблему.
>
> Как минимум:
>
> - Какие шаги вы сделали именно, это может помочь воспроизвести проблему.
> - Каковы точные сообщения об ошибках, которые вы видите.

Его спрятанный в "kocoman1 прокомментировал сентябрь 29, 2021"
#12225 (комментарий)

VERIFY3(c < SPA_MAXBLOCKSIZE >> SPA_MINBLOCKSHIFT) потерпел неудачу (36028797018963967 < 32768)
[Ср Сен 29 22:13:28 2021] ПАНИКЯ на zio.c:322:zio_buf_alloc()

[Ср Сен 29 22:13:28 2021] Следы звонков:

[Ср Сен 29 22:13:28 2021] show_stack+0x52/0x58
[Ср Сен 29 22:13:28 2021] dump_stack+0x70/0x8b
[Ср Сен 29 22:13:28 2021] spl_dumpstack+0x29/0x2b [spl]
[Ср Сен 29 22:13:28 2021] spl_panic+0xd4/0xfc [spl]
[Ср Сентябрь 29 22:13:28 2021] ? _cond_resched+0x1a/0x50
[Ср Сентябрь 29 22:13:28 2021] ? _cond_resched+0x1a/0x50
[Ср Сентябрь 29 22:13:28 2021] ? __kmalloc_node+0x144/0x2b0

[Ср Сен 29 22:13:28 2021] Аннотировать_эксум+0x85/0x4d0 [zfs]
[Ср Сентябрь 29 22:13:28 2021] ? abd_cmp_cb+0x10/0x10 [zfs]
[Ср Сен 29 22:13:28 2021] zfs_ereport_finish_finish_checksum+0x2c/0xb0 [zfs]
[Ср Сен 29 22:13:28 2021] zio_vsd_default_cksum_finish+0x14/0x20 [zfs]
[Ср Сен 29 22:13:28 2021] zio_done+0x2fb/0x11b0 [zfs]
[Ср Сен 29 22:13:28 2021] zio_execute+0x8b/0x130 [zfs]
[Ср Сен 29 22:13:28 2021] задачаq_thread+0x2b7/0x500 [spl]
[Ср Сентябрь 29 22:13:28 2021] ? win_up_q+0xa0/0xa0/0xa0
[Ср Сентябрь 29 22:13:28 2021] ? zio_gang_tree_free+0x70/0x70 [zfs]
[Ср Сен 29 22:13:28 2021] Хтлит+0x11ф/0x140
[Ср Сентябрь 29 22:13:28 2021] ? taskq_thread_pawn+0x60/0x60 [spl]
[Ср Сентябрь 29 22:13:28 2021] ? set_kthread_struct+0x50/0x50
[Ср Сен 29 22:13:28 2021] ret_from_fork+0x22/0x30

Майер Комментировать on 7 апр. 2023 г.

Есть ли что-то, что могут сделать конкретные добровольцы, чтобы помочь в тестировании этой функции? Что поможет больше всего?

Кроме того, можно ли переоснастить эту ветвь на мастера? Сейчас существует большое количество конфликтов слияний. Это делает его намного сложнее, чтобы проверить эту ветвь. Кроме того, я подозреваю, что это сделает любые тесты, которые проводятся, более бессмысленными, поскольку функции, уже присутствующие в upstream, не будут тестироваться вместе с расширением RAIDZ. Другими словами, чем более устаревшей становится эта ветвь, тем меньше смысла она использует или полагаться на нее.

👍 5

кокомано1 Комментировать on 8 апр. 2023 г.

> Есть ли что-то, что могут сделать конкретные добровольцы, чтобы помочь в тестировании этой функции? Что поможет больше всего?
>
> Кроме того, можно ли переоснастить эту ветвь на мастера? Сейчас существует большое количество конфликтов слияний. Это делает его намного сложнее, чтобы проверить эту ветвь. Кроме того, я подозреваю, что это сделает любые тесты, которые проводятся, более бессмысленными, поскольку функции, уже присутствующие в upstream, не будут тестироваться вместе с расширением RAIDZ. Другими словами, чем более устаревшей становится эта ветвь, тем меньше смысла она использует или полагаться на нее.

может попробовать, что я сделал, что вызывает ошибку

иметь кучу данных, которые могут быть потеряны
подготовьте как 10 hdd с некоторыми данными (может быть, небольшой размер так

случае) с мин 3 (?) drives, затем добавляйте другой диск, копируйте больше данных до полного (самого важного), промывайте и повторяйте до достижения 10 дисков, чтобы увидеть, может ли какая-либо ошибка во время процесса.. также может вымыть после каждого расширения

justinclift Комментировать on 8 апр. 2023 г.

Просто чтобы указать для всех, кто хотел бы протестировать это, но не имеет большого запасного оборудования ... тестирование этого в виртуальных машинах - это работоспособная идея. например, виртуальная машина с виртуальными дисками. Используйте их, чтобы имитировать добавление большего количества дисков, вытягивание вилки на некоторые, пока он делает вещи и т. Д.

В настоящее время моя проблема заключается в нехватке времени. ***

Соковысли Комментировать on 8 апр. 2023 г.

> Просто чтобы указать для всех, кто хотел бы протестировать это, но не имеет большого запасного оборудования ... тестирование этого в виртуальных машинах - это работоспособная идея. например, виртуальная машина с виртуальными дисками. Используйте их, чтобы имитировать добавление большего количества дисков, вытягивание вилки на некоторые, пока он делает вещи и т. Д.
>
> Моя проблема в нехватке времени в настоящее время. Ангел

Одной из великих сил ZFS была возможность поставлять поддельные файловые диски, используемые в тестировании. Мог бы все-таки сбить ядро, если бы не в виртуальный.

👍 2

Темная базировка Комментировать on 13 апр. 2023 г.

Реально я не думаю, что узким местом является отсутствие тестирования, а отсутствие обзоров кода (AFAIK, за исключением пары PR, нет никаких других отзывов для решения), и в основном сам **@ahrens** занят другими вещами (ранее упомянутые PR ожидаются более года). Поэтому я предлагаю либо подождать, пока его приоритеты совпадут с нашими, либо быть готовым внести какой-то значимый вклад. Эта функция была описана как, возможно, рассмотренная в следующей крупной версии zfs (я не помню, какая компания была заинтересована в этом), но, честно говоря, я начинаю сильно сомневаться в этом.

👍 6

> Просто чтобы указать для всех, кто хотел бы протестировать это, но не имеет большого запасного оборудования ... тестирование этого в виртуальных машинах - это работоспособная идея. например, виртуальная машина с виртуальными дисками. Используйте их, чтобы имитировать добавление большего количества дисков, вытягивание вилки на некоторые, пока он делает вещи и т. Д.
> Моя проблема в нехватке времени в настоящее время. Ангел
>
> Одной из великих сил ZFS была возможность поставлять поддельные файловые диски, используемые в тестировании. Мог бы все-таки сбить ядро, если бы не в виртуальный.

Есть ли для этого учебник? Я хочу объединить 1 и 2tb диск, чтобы сделать 3tb, чтобы заменить факультатив 3tb
thx

👎 18 😕 1

Соковысли Комментировать on 14 апр. 2023 г.

> > Просто чтобы указать для всех, кто хотел бы протестировать это, но не имеет большого запасного оборудования ... тестирование этого в виртуальных машинах - это работоспособная идея. например, виртуальная машина с виртуальными дисками. Используйте их, чтобы имитировать добавление большего количества дисков, вытягивание вилки на некоторые, пока он делает вещи и т. Д.
> > Моя проблема в нехватке времени в настоящее время. Ангел
> >
> > Одной из великих сил ZFS была возможность поставлять поддельные файловые диски, используемые в тестировании. Мог бы все-таки сбить ядро, если бы не в виртуальный.
>
> Есть ли для этого учебник? Я хочу объединить 1 и 2tb диск, чтобы сделать 3tb, чтобы заменить фавлирующий 3tb thx

Вы можете поставить два диска, поскольку видеоэлектроэнергии верхнего уровня и ZFS по существу будут разбираться по ним. Если любой привод (1TB или 2TB) не сработает, вы потеряете все данные, и любые ошибки будут неисправимы в этой конфигурации.

Пожалуйста, задавайте вопросы поддержки, не связанные с этой функцией, на публичном форуме, таком как https://www.reddit.com/r/openzfs/
Есть много людей, подписанных на эту проблему, и это пустая трата времени каждого, чтобы опубликовать о вещах, которые не помогают выполняемой работе.

👍 12 ❤️ 2

об этом, как и закончили (год назад)
https://freebsdfoundation.org/blog/raid-z-expansion-feature-for-zfs/
Но в то же время кажется, что здесь ппы устарели... что происходит?

👍 4 👎 4 😕 1 👀 2

Соковысли Комментировать on 20 апр. 2023 г. • • отредактировано

@felisucoibi

> Реально я не думаю, что узким местом является отсутствие тестирования, а скорее отсутствие обзоров кода (AFAIK, за исключением нескольких PR, нет никаких других отзывов для решения) и в основном @/Hrens сам занят другими вещами (ранее упомянутые PR ожидаются более года). Поэтому я предлагаю либо подождать, пока его приоритеты совпадут с нашими, либо быть готовым внести какой-то значимый вклад. Эта функция была описана как, возможно, рассмотренная в следующей крупной версии zfs (я не помню, какая компания была заинтересована в этом), но, честно говоря, я начинаю сильно сомневаться в этом.

Пожалуйста, прочитайте ветку, прежде чем ответить. Согласно комментарию, три комментария перед вашим:

> Реально я не думаю, что узким местом является отсутствие тестирования, а скорее отсутствие обзоров кода (AFAIK, за исключением нескольких PR, нет никаких других отзывов для решения) и в основном @/Hrens сам занят другими вещами (ранее упомянутые PR ожидаются более года). Поэтому я предлагаю либо подождать, пока его приоритеты совпадут с нашими, либо быть готовым внести какой-то значимый вклад. Эта функция была описана как, возможно, рассмотренная в следующей крупной версии zfs (я не помню, какая компания была заинтересована в этом), но, честно говоря, я начинаю сильно сомневаться в этом.

Эта работа является функциональной, но требует как очистки, так и обзоров кода. Другие люди также испытывали ошибки во время недавнего тестирования. Для того, чтобы это было сделано, должно быть время для @/ahrens и других, чтобы работать над этим, но они очень заняты другими приоритетами.

Есть много людей, подписанных на этот вопрос, и это пустая трата времени каждого, чтобы опубликовать о вещах, на которые уже был дан ответ.

👍 6

бочковой Комментировать on 20 апр. 2023 г. • • отредактировано

Я стараюсь верить в проекты с открытым исходным кодом со всей своей волей, но ситуации, подобные этому PR, создают большую демотивацию. Многие проекты и иногда

расстроенным и острым (может быть, не только я). В последний раз автор заключал код 25 февраля 2022 года, более года назад, затем мы получили некоторые изменения от **@fuporovvStack,** и последняя активность была 16 ноября 2023 года, почти 6 месяцев назад. Также сопровождающие и разработчики ничего не сообщают об интенсионалах или статусе. Мне кажется, нет никаких намерений продолжать его.

Мне грустно, что этот проект не кажется устойчивым, целеустремленным и уважающим, даже с сотнями заинтересованных людей в нем. Для меня эта функция устаревшая / заброшенная, я лучше потрачу свое время на поиск альтернативной FS, которая выполняет мои требования, если кто-то другой чувствует то же, что и я, мы можем попытаться найти вместе лучшее решение, поскольку у меня недостаточно знаний о внутренних источниках этого проекта, чтобы исправить код zfs (и кажется, что люди, у которых он не очень хочет).

👍 12 👎 33 😄 5 😕 1

Эверов Комментировать on 20 апр. 2023 г.

> Также сопровождающие и разработчики ничего не сообщают об интенсионалах или статусе. Мне кажется, нет никаких намерений продолжать его.

Могу ли я предложить добавить это в повестку дня, тогда на встрече руководства 25 апреля (5 дней спустя)? Подробности здесь: https://docs.google.com/document/d/1w2jv2XVYFmBVvG1EGf-9A5HBVsjAYOLIFZAnWHV-BM/edit

👍 10 👎 1 😄 1

Мыслить Чаос Комментировать on 20 апр. 2023 г.

Может ли кто-то попытаться найти решение, чтобы ограничить спам в этом пиаре? Большинство комментариев - это просто шум. (включая мой, извините!)
Может быть, блокировка его, чтобы только участники могли комментировать, и провести отдельное обсуждение для тестирования и обратной связи от третьих сторон?

👍 25 👎 19

Квинз Комментировать on 1 мая 2023 г.

> Также сопровождающие и разработчики ничего не сообщают об интенсионалах или статусе. Мне кажется, нет никаких намерений продолжать его.
>
> Могу ли я предложить добавить это в повестку дня, тогда на встрече руководства 25 апреля (5 дней спустя)? Подробности здесь: https://docs.google.com/document/d/1w2jv2XVYFmBVvG1EGf-9A5HBVsjAYOLIFZAnWHV-BM/edit

👍 19 😄 2

Хендрик42 Комментировать on 2 мая 2023 г.

Спасибо, что предоставили нам обновление на этой встрече **@behlendorf**

Вот стенограмма из youtube:

> Здесь в списке, и мне нечего сказать об этом, это расширение рейда Z.
> Я не совсем уверен, что происходит с этим - может быть, у других есть еще какое-то понимание - но я знаю, что это все еще работа, которую сделал Мэтт, которая нуждается в большем обзоре кода, нуждается в большем тестировании, ее необходимо переоснастить, поэтому там еще есть куча работы, чтобы сделать время, чтобы завершить его.
>
> Я знаю, что люди провели некоторые тесты с ним, но это все еще не совсем завершено, но было бы здорово продвинуть эту работу вперед и переосмыслить ее, потому что я знаю, что это функция, которой многие люди увлекаются. Это просто еще не совсем вместе, но это большой подъем.

👍 31 🎉 8

 **Крисдисмонсон** Упомянули эту просьбу о вытягива on 8 мая 2023 г.

Raidz расширяется #14840

Закрыто

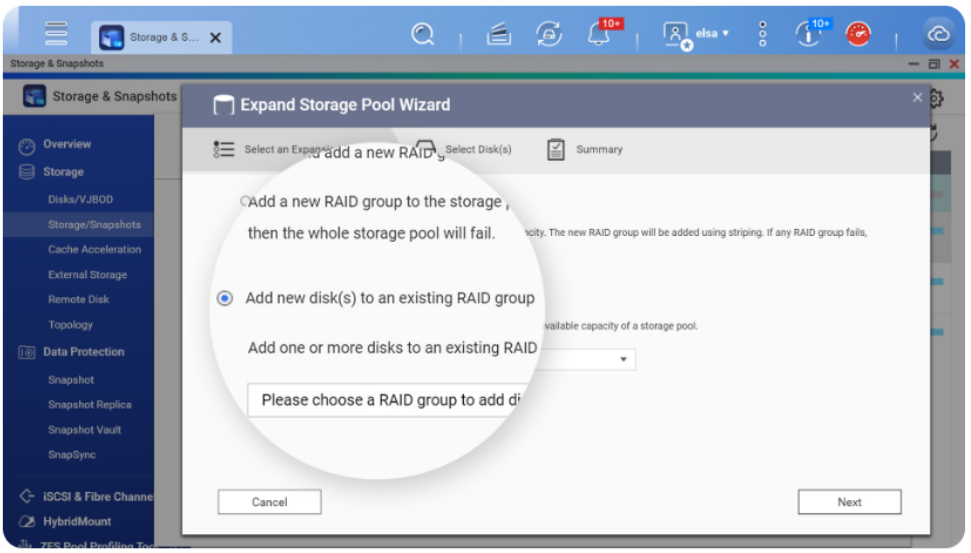
12 задач

Кесил Комментировать on 11 июн. 2023 г.

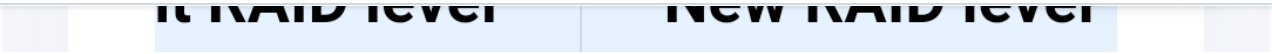
Это происходит?
Здесь:
https://www.qnap.com/en-us/operating-system/quts-hero/5.1.0

single disk

Expanding the capacity of a ZFS RAID-Z previously required adding another RAID group. Now, you can simply add a single disk to an existing RAID-Z for storage expansion, or conveniently add 2 to 3 disks for upgrading RAID level with Parity.



The minimum number of disks for expanding RAID capacity:



👍 13 👎 2 👀 26

kmore134 Комментировать on 27 июн. 2023 г.

Приятно сообщить, что iXsytems спонсирует усилия **@don-brady** по завершению и объединению этого. Спасибо **@don-brady** и **@ahrens** за обсуждение этого на встрече лидеров OpenZFS сегодня. С нетерпением ждем обновленного PR в ближайшее время.

https://www.youtube.com/watch? v=2p32m-7FNpM

👍 41 🎉 80 ❤️ 51 🚀 21 👀 3

Дон-Брэйди Упомянули эту просьбу о вытягива on 29 июн. 2023 г.

функция расширения рейза #15022

Слияние

19 задач

Архренс Комментировать on 29 июн. 2023 г. Член Автор

@don-brady взял на себя эту работу и открыл #15022. Спасибо Дону и спасибо iXsystems за спонсирование его работы! См. Совещание лидеров OpenZFS в июне 2023 года для краткого обсуждения переходного периода.

👍 27 🎉 36 ❤️ 22 🚀 10

Архренс Закрыл это on 29 июн. 2023 г.

Рецензенты

Набиджацлевелили

статический

Ассеенты

Мейби

Этикетки

Проекты

Пока нет

Развитие

Успешное слияние этого запроса может закрыть эти вопросы.

Пока нет

73 участника