Настройка журналирования UFS для настольного компьютера

Аннотация

Журналируемая файловая система использует журнал для записи всех транзакций, происходящих в файловой системе, который также сохраняет ее целостность в случае краха системы или пропадания питания. Несмотря на то, что всё еще возможна потеря несохранённых изменений файлов, журналирование почти полностью исключает возможность повреждения структуры файловой системы, вызванное непредвиденным остановом работы. Журналирование также сокращает до минимума время, необходимое для проверки файловой системы после отказа. Несмотря на то, что в используемой FreeBSD файловой системе UFS нет поддержки журналирования, новый класс системы GEOM в FreeBSD 7.*X* может быть использован для ведения независимого от файловой системы журналирования. Эта статья объясняет, как реализовать журналирование UFS для типичного настольного компьютера.

Содержание

1. Введение 1
2. Реализация журналирования в FreeBSD 2
3. Действия, необходимые во время установки FreeBSD 3
4. Настройка журналирования 6
5. Устранение неполадок с журналированием 9
6. Для дальнейшего ознакомления 12

1. Введение

Серверное оборудование обычно хорошо защищено от потери питания. Настольный компьютер часто подвержен неожиданным пропаданиям питания, случайным нажатиям кнопки Reset и другим происшествиям (часто связанным с неосторожностью пользователей), которые могут привести к непредвиденным выключениям. Механизм Soft Updates, как правило, достаточно эффективно защищает файловую систему в таких случаях, однако в последствии требуется длительная фоновая проверка. В очень редких случаях повреждения файловой системы достигают того уровня, при котором становится необходимым вмешательство пользователя и данные могут быть утерянными.

Новая возможность журналирования, предоставленная системой GEOM, может существенно выручить в подобных случаях, исключая время, необходимое для проверки файловых систем и удостоверяясь, что файловая система быстро восстановлена в целостное состояние.

Эта статья описывает порядок действий, необходимых для конфигурирования журналирования UFS на типичном настольном компьютере, в котором один жесткий диск используется для размещения как операционной системы, так и данных. В статье подразумевается установка FreeBSD "с нуля". Шаги достаточно просты и не требуют чрезмерно сложных манипуляций с командной строкой.

После прочтения данной статьи вы будете знать:

- Как зарезервировать место для журнала во время новой установки FreeBSD.
- Как загрузить модуль `geom_journal` (или включить поддержку журналирования в специализированном ядре системы).
- Как преобразовать существующую файловую систему, в систему, использующую журналирование, и какие опции монтирования использовать в /etc/fstab.
- Как реализовать журналирование на новых (пустых) разделах.
- Как диагностировать неполадки, связанные с журналированием.

Перед прочтением этой статьи вам необходимо:

- Понимать базовые концепции таких операционных систем, как UNIX® и FreeBSD.
- Быть знакомым с процедурой установки FreeBSD, а также с программой Sysinstall.

Процедура, описанная здесь, подразумевает подготовку к новой установке, в которой на дисках еще нет пользовательских данных. Так как эту процедуру можно модифицировать и расширить на системы, которые уже используются, вам настоятельно рекомендуется сделать *резервную копию* всех ценных данных. Путаница в низкоуровневых операциях с дисками и разделами может привести к фатальным ошибкам и потере данных.

2. Реализация журналирования в FreeBSD

Журналирование, предоставляемое системой GEOM в FreeBSD 7.*X*, не является особенностью файловой системы (в отличие от, например, файловой системы ext3 в Linux®), оно функционирует на блочном уровне. А это значит, что оно может быть применено к разным типам файловых систем, однако для FreeBSD 7.0-RELEASE журналирование может быть применено только для UFS2.

Возможность журналирования обеспечивается загрузкой модуля geom_journal.ko в ядро (или сборкой собственного ядра с активированием соответствующих опций) и использованием команды `gjournal` для конфигурирования файловой системы. В общем, вы предпочтете журналировать файловые системы большого размера, к примеру - /usr. Однако, вам придется зарезервировать некоторое количество свободного места (см. следующий раздел).

Когда файловая система журналируется, некоторая часть дискового пространства требуется для хранения самого журнала. Дисковое пространство, содержащее данные, называется *поставщиком данных (data provider)*, а часть пространства, содержащая журнал, называется

поставщиком журнала (journal provider). Поставщики данных и журнала должны быть на разных разделах, если журналирование достраивается к содержащему данные разделу. А если журналирование включается для нового раздела, у вас есть возможность использовать один поставщик для данных и журнала. В любом из двух вышеупомянутых случаев команда `gjournal` задействует поставщики и создаст конечную журналируемую файловую систему. Например:

- Вы намереваетесь журналировать файловую систему /usr, размещенную на /dev/ad0s1f, файловая система уже содержит данные.
- Вы зарезервировали часть дискового пространства на разделе /dev/ad0s1g.
- Используя команду `gjournal`, создаем новый файл устройства /dev/ad0s1f.journal, для которого /dev/ad0s1f является поставщиком данных, а /dev/ad0s1g - поставщик журнала. Это новое устройство необходимо использовать во всех последующих операциях.

Размер дискового пространства, отводимого под поставщик журнала, зависит от нагруженности файловой системы, а не от размера самого поставщика данных. Например, для типичного настольного компьютера достаточно отвести 1 Гб под поставщик журнала для файловой системы /usr, в то время как компьютеру, имеющему интенсивный дисковый ввод/вывод (например, редактирование видео) может потребоваться больше. Если свободное место на поставщике журнала заканчивается раньше, чем происходит сброс журнала на диск, - вы получите панику ядра.

Очень маловероятно то, что размеры журнала, предложенные здесь, станут причиной проблем с обычным настольным компьютером (на котором вы просматриваете веб-страницы, обрабатываете текст или проигрываете мультимедийные файлы). Если работа вашего компьютера подразумевает интенсивную дисковую активность, то для обеспечения стабильности следует придерживаться следующего правила: размер ОЗУ должен уместиться в 30% размера, отведенного под журнал. Например, если в вашем компьютере установлен 1 Гб ОЗУ, создайте под поставщик журнала раздел размером около 3.3 Гб. (Умножьте размер ОЗУ в 3.3 раза, чтоб получить размер журнала).

Для получения дополнительной информации о журналировании, пожалуйста, прочитайте страницу справочника, посвященную gjournal(8).

3. Действия, необходимые во время установки FreeBSD

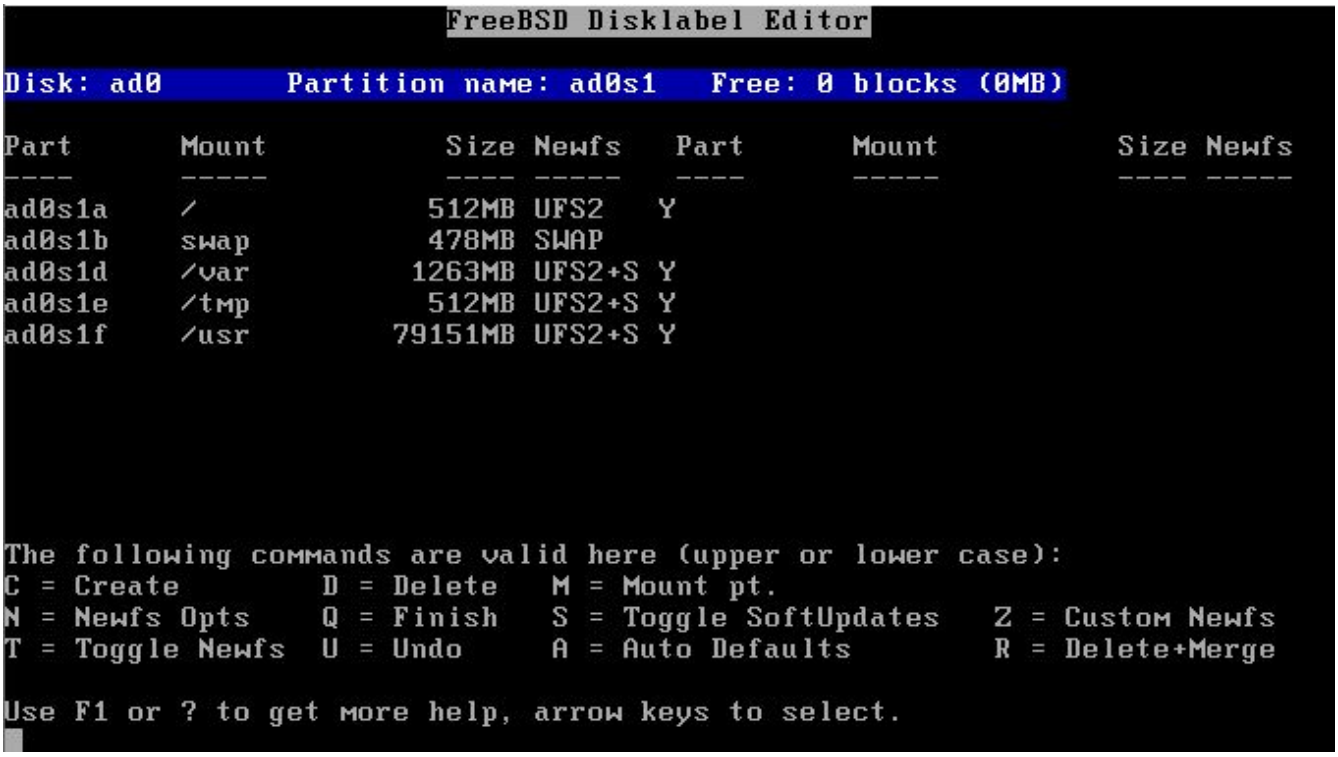
3.1. Выделение места под журналирование

Типичный настольный компьютер обычно имеет один жесткий диск, на котором хранится как операционная система, так и пользовательские данные. Вероятно, что схема разбития винчестера (по умолчанию), выбранная в меню Sysinstall, является более или менее подходящей: настольному компьютеру не требуется большой раздел /var, в то время, как

для раздела /usr выделяется значительный объем дискового пространства, ввиду того, что пользовательские данные и множество пакетов инсталлируются в подкаталоги /usr.

Разбиение по умолчанию (получаемое при нажатии `A` в редакторе разделов FreeBSD, называемом Disklabel) не оставляет свободного места. Каждый подлежащий журналированию раздел требует отдельного раздела для журнала. Ввиду того, что раздел /usr - наибольший, есть смысл немного уменьшить его размер, чтобы получить пространство, необходимое для журнала.

В нашем примере используется жесткий диск размером 80 Гб. Следующий скриншот показывает результаты разбиения по умолчанию, выполненного при помощи Disklabel в процессе установки операционной системы:



Если это разбиение более или менее вас устраивает, то его легко модифицировать для журналирования. Используйте клавиши со стрелками для того, чтобы выделить раздел, отведенный под /usr, потом нажмите `D` чтобы удалить его.

Теперь переведите подсвечивание к имени диска, находящемуся вверху экрана, и нажмите `C` - создайте новый раздел /usr. Новый раздел должен быть меньше на 1 Гб (если вы собираетесь журналировать только /usr) или на 2 Гб (если журналированию подлежат как /usr, так и /var). Во всплывающем окне выберите "создать файловую систему" и укажите /usr точкой монтирования.

Следует ли журналировать /var раздел? Обычно есть смысл журналировать большие разделы. Вы можете решить не журналировать /var, однако журналирование на обычном настольном компьютере не причинит вреда. Если файловая система не нагружена (что типично для настольной системы), то можно выделить меньше дискового пространства под журнал.

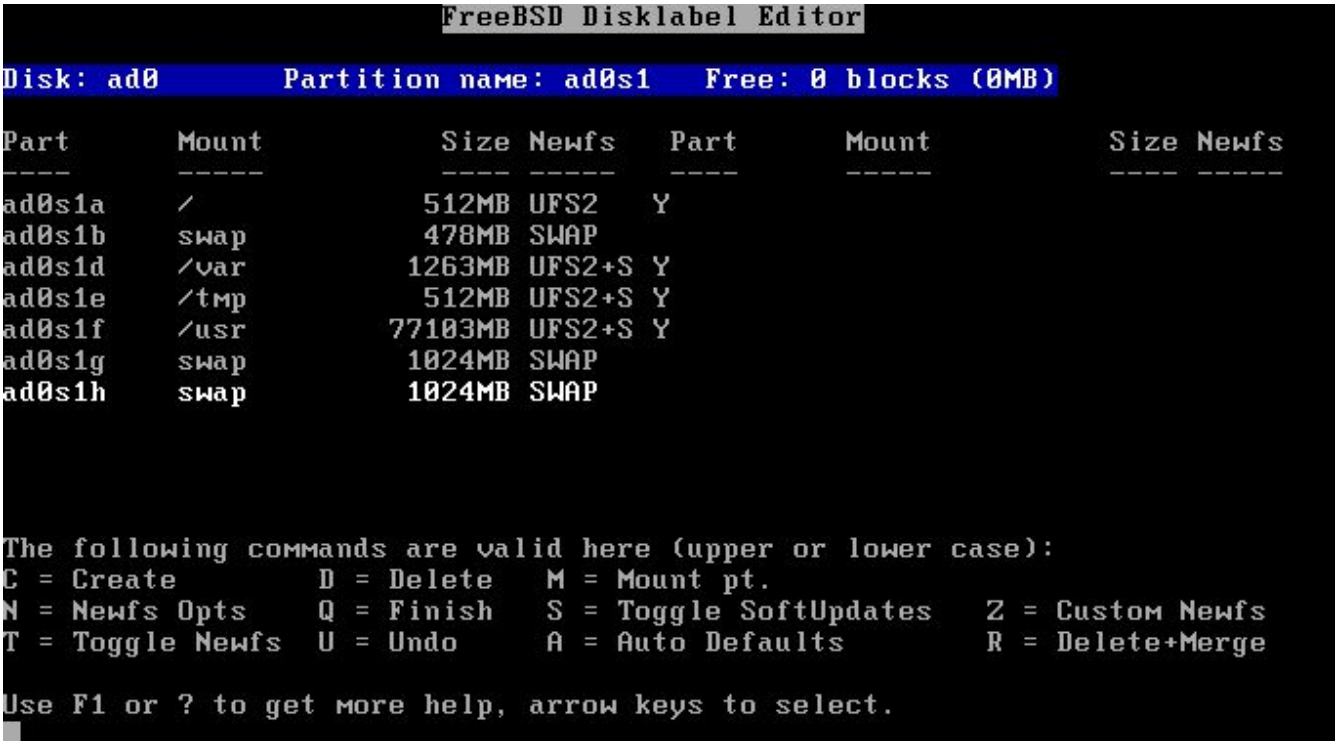
В этом примере подразумевается журналирование двух файловых систем:

> /usr и /var. Естественно, вы можете подкорректировать процедуру под свои задачи.

Чтобы не усложнять описываемую методику, для создания разделов, необходимых для размещения журналов, мы будем использовать утилиту Sysinstall. Однако, во время установки утилита Sysinstall требует указания точек монтирования для каждого созданного вами раздела. Но разделы, содержащие журналы, вам никогда и никуда монтировать не придется.

Чтобы избежать вопросов о точках монтирования, мы создадим разделы под журналы и установим их тип в swap. Раздел, предназначенный для свопа, никогда и никуда не монтируется, плюс к тому, утилита Sysinstall позволяет создавать столько разделов под своп, сколько необходимо. После первой перезагрузки необходимо подредактировать файл /etc/fstab, удалив в нём лишние записи о своп-разделах.

Для создания своп-раздела, используя клавиши со стрелками, перемещайте подсвечивание к верхней части экрана в утилите Disklabel так, чтобы стало подсвеченным имя диска. Потом, нажмите `N`, введите необходимый размер раздела (*1024M*), а после - выберите во всплывшем окне "swap space". Повторите эти шаги для всех оставшихся журналов. В этом примере мы создаем два раздела, на которых будут размещаться журналы для /usr и /var. Конечный результат показан на следующем скриншоте:



По завершении создания разделов мы рекомендуем вам записать на бумагу названия разделов и их точек монтирования: с этой информацией вы будете сверяться во время конфигурирования. Это также поможет уменьшить количество ошибок, приводящих к повреждению установки. Следующая табличка отображает наши заметки, сделанные для данного примера:

Таблица 1. Разделы и журналы

Раздел	Точка монтирования	Журнал
ad0s1d	/var	ad0s1h
ad0s1f	/usr	ad0s1g

Дальше продолжайте обычную установку. Однако, мы рекомендуем вам отложить инсталляцию приложений сторонних разработчиков (пакетов) до полной настройки журналирования.

3.2. Первая загрузка

Ваша система загрузится нормально, однако вам необходимо будет подредактировать /etc/fstab и удалить те лишние своп-разделы, которые вы создавали под журналы. Как правило, в названии файла устройства, созданного автоматически утилитой Sysinstall, присутствует суффикс "b" (в нашем примере это ad0s1b). Удалите другие записи о своп-разделах и перезагрузите компьютер, после чего FreeBSD перестанет их использовать.

После второй перезагрузки, компьютер будет готов к конфигурированию журналирования.

4. Настройка журналирования

4.1. Работа с командой `gjournal`

Подготовив необходимые разделы, перейдем к конфигурированию журналирования. Нам будет необходимо загрузиться в однопользовательском режиме, для этого залогинимся пользователем `root` и напечатаем:

```
# shutdown now
```

Нажмите `Enter` для получения приглашения командного интерпретатора. Нам необходимо будет размонтировать разделы, которые подлежат журналированию, в нашем примере это /usr и /var:

```
# umount /usr /var
```

Загрузите модуль ядра, необходимый для журналирования:

```
# gjournal load
```

На данном этапе сверьтесь со своими записями и определите, какие разделы будут использоваться под какой журнал. В нашем примере /usr располагается на ad0s1f, а его журнал будет располагаться на ad0s1g, и, по аналогии, для /var: файловая система располагается на ad0s1d, а ее журнал - на ad0s1h. Наберите следующие команды:

```
# gjournal label ad0s1f ad0s1g
GEOM_JOURNAL: Journal 2948326772: ad0s1f contains data.
GEOM_JOURNAL: Journal 2948326772: ad0s1g contains journal.

# gjournal label ad0s1d ad0s1h
GEOM_JOURNAL: Journal 3193218002: ad0s1d contains data.
GEOM_JOURNAL: Journal 3193218002: ad0s1h contains journal.
```

Если последний сектор любого из двух разделов (поставщиков данных) используется, команда `gjournal` возвратит ошибку. Вам необходимо будет использовать флаг `-F` для принудительной перезаписи, например:

```
# gjournal label -f ad0s1d ad0s1h
```

Так как это - новая установка, очень маловероятен факт, что что-нибудь будет действительно переписано.

На данном этапе созданы два устройства: ad0s1d.journal и ad0s1f.journal. Они представляют /var и /usr соответственно. Перед монтированием, нам необходимо установить флаг журналирования и снять флаг механизма Soft Updates:

```
# tunefs -J enable -n disable ad0s1d.journal
tunefs: gjournal set
tunefs: soft updates cleared

# tunefs -J enable -n disable ad0s1f.journal
tunefs: gjournal set
tunefs: soft updates cleared
```

Теперь, смонтируйте новые устройства в соответствующие места файловой системы (обратите внимание на то, что мы можем использовать опцию монтирования `async`):

```
# mount -o async /dev/ad0s1d.journal /var
# mount -o async /dev/ad0s1f.journal /usr
```

Откройте /etc/fstab и исправьте записи для следующих файловых систем: /usr и /var:

```
/dev/ad0s1f.journal /usr  ufs rw,async 2 2
/dev/ad0s1d.journal /var  ufs rw,async 2 2
```

Проверьте корректность всех указанных параметров - в противном случае после перезагрузки возможны проблемы со стандартным запуском системы!

Наконец, внесите изменения в /boot/loader.conf, добавив строку для автоматической загрузки модуля gjournal(8) при каждом старте системы:

```
geom_journal_load="YES"
```

Поздравляем! Ваша система теперь настроена для журналирования. Вы можете ввести команду `exit` для возврата в многопользовательский режим или перезагрузить систему для проверки конфигурации (рекомендуется). При загрузке вы увидите сообщения следующего вида:

```
ad0: 76293MB XEC XE800JD-00HBC0 08.02D08 at ata0-master SATA150
GEOM_JOURNAL: Journal 2948326772: ad0s1g contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1h contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1d contains data.
GEOM_JOURNAL: Journal ad0s1d clean.
GEOM_JOURNAL: Journal 2948326772: ad0s1f contains data.
GEOM_JOURNAL: Journal ad0s1f clean.
```

После некорректного завершения работы сообщения будут немного отличаться, например:

```
GEOM_JOURNAL: Journal ad0s1d consistent.
```

Обычно это означает, что gjournal(8) использовал информацию из журнала для приведения файловой системы в согласованное состояние.

4.2. Журналирование новых разделов

Процедура, описанная выше, необходима для подключения журналирования разделов, содержащих данные. Журналирование пустых разделов немного проще, ввиду того, что поставщик данных и поставщик журнала могут быть размещены на одном и том же разделе. Например, предположим, что был установлен новый жесткий диск и был создан новый раздел /dev/ad1s1d. Создание журнала не сложнее набора:

```
# gjournal label ad1s1d
```

Размер журнала - 1 Гб по умолчанию. Однако, вы можете изменить это значение используя ключ `-s`. Значение можно задавать в байтах, в килобайтах, мегабайтах или гигабайтах (используя суффикс `K`, `M` или `G`). Имейте ввиду, что команда `gjournal` не позволит вам создать журнал недопустимо малого размера.

К примеру, чтобы создать журнал размером в 2Гб, можно использовать следующую команду:

```
# gjournal label -s 2G ad1s1d
```

Далее, вы можете создать файловую систему на новом разделе, а также разрешить журналирование ключом `-J`:

```
# newfs -J /dev/ad1s1d.journal
```

4.3. Встраивание журналирования в специализированное ядро

Если вы не желаете загружать `geom_journal` как модуль, то можно встроить его функции прямо в ваше специализированное ядро. Редактируя конфигурационный файл ядра, убедитесь, что в нем находятся следующие две строки:

```
options UFS_GJOURNAL # Note: This is already in GENERIC

options GEOM_JOURNAL # You will have to add this one
```

Соберите и установите новое ядро следуя указаниям Руководства FreeBSD.

И не забудьте удалить соответствующую строку загрузки модуля ("load") из /boot/loader.conf (если на предыдущем этапе она была туда внесена).

5. Устранение неполадок с журналированием

Этот раздел содержит часто задаваемые вопросы касательно неполадок, связанных с журналированием.

5.1. Я получаю паники ядра во время высокой дисковой активности. Как это связано с журналированием?

Вероятно, что журнал заполняется раньше, чем происходит сброс его на диск. Помните, размер журнала зависит от загруженности диска, а не от размера поставщика данных. Если загрузка диска высокая, вам потребуется раздел большего размера для журнала. См. замечания в разделе Understanding Journaling in FreeBSD.

5.2. Я допустил некоторые ошибки во время конфигурирования, теперь система не загружается. Можно это как-нибудь исправить?

Вы либо забыли внести запись (опечатались) в /boot/loader.conf, либо есть ошибки в файле /etc/fstab. Это легко исправить. Нажмите `Enter`, чтобы получить приглашение командного интерпретатора в однопользовательском режиме. Потом, проверьте возможные варианты:

```
# cat /boot/loader.conf
```

Если отсутствует запись `geom_journal_load`, или она содержит ошибки, журналируемые устройства не создадутся. Загрузите модуль вручную, примонтируйте все разделы и переходите в многопользовательский режим (продолжайте загрузку):

```
# gjournal load

GEOM_JOURNAL: Journal 2948326772: ad0s1g contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1h contains journal.
GEOM_JOURNAL: Journal 3193218002: ad0s1d contains data.
GEOM_JOURNAL: Journal ad0s1d clean.
GEOM_JOURNAL: Journal 2948326772: ad0s1f contains data.
GEOM_JOURNAL: Journal ad0s1f clean.

# mount -a
# exit
(boot continues)
```

Если же запись о `geom_journal_load` верна, то проверьте /etc/fstab. Вероятней всего, что вы обнаружите опечатку или отсутствие необходимой записи. В этом случае смонтируйте вручную оставшиеся разделы и продолжите загрузку в многопользовательский режим.

5.3. Возможно ли отказаться от журналирования и вернуться к моей привычной файловой системе с механизмом Soft Updates?

Несомненно. Используйте приведенную ниже последовательность действий, которая обращает изменения. Разделы, созданные для поставщиков журналов, могут позже быть использованы для других целей.

Залогиньтесь `root` и переведите систему в однопользовательский режим:

```
# shutdown now
```

Размонтируйте журналируемые разделы:

```
# umount /usr /var
```

Синхронизируйте журналы:

```
# gjournal sync
```

Остановите поставщиков журналов:

```
# gjournal stop ad0s1d.journal
# gjournal stop ad0s1f.journal
```

Удалите метаданные журналирования со всех задействованных устройств:

```
# gjournal clear ad0s1d
# gjournal clear ad0s1f
# gjournal clear ad0s1g
# gjournal clear ad0s1h
```

Снимите флаг журналирования и установите флаг механизма Soft Updates:

```
# tunefs -J disable -n enable ad0s1d
tunefs: gjournal cleared
tunefs: soft updates set

# tunefs -J disable -n enable ad0s1f
tunefs: gjournal cleared
tunefs: soft updates set
```

Смонтируйте вручную старые (первоначальные) устройства:

```
# mount -o rw /dev/ad0s1d /var
# mount -o rw /dev/ad0s1f /usr
```

Откройте файл /etc/fstab и приведите его к изначальному виду:

```
/dev/ad0s1f /usr  ufs rw 2 2
/dev/ad0s1d /var  ufs rw 2 2
```

И напоследок, удалите строку, загружающую модуль `geom_journal`, из файла /boot/loader.conf и перезагрузите операционную систему.

6. Для дальнейшего ознакомления

Журналирование - относительно новая функциональная возможность FreeBSD, и как такова, она еще недостаточно документирована. Однако, вы можете сочти полезными следующие источники:

- Новый раздел Руководства FreeBSD, посвященный журналированию.
- Этот пост в списке рассылки Список рассылки, посвящённый обсуждению FreeBSD-CURRENT, написанный Paweł Jakub Dawidek <pjd@FreeBSD.org> - автором gjournal(8).
- Этот пост от Ivan Voras <ivoras@FreeBSD.org> в списке рассылки Список рассылки, посвящённый вопросам и ответам пользователей FreeBSD.
- Страницы справочника gjournal(8) и geom(8).