

## Объединение томов через aufs для отказоустойчивости и моментального восстановления

Объединение томов с разных физических устройств с распределением файлов/слепков файлов на предмет отказоустойчивости и моментального восстановления в случае нечисти по типу "шифровальщиков". Объединённый ресурс отдаётся через Samba и NFS.

Нужно было быстро решить проблему с восстановлением после попадания пользователей под "шифровальщиков". Тем не менее, решение уже показало себя весьма эффективным и буквально спасло. Дополнительно продумывались меры по повышению надёжности, так как если не по бедности, то по понятным соображениям, организовать правильно организованный бэкап сложно: инкрементальный бэкап периодически (сильно реже разумного) сливается на отдельный, не включенный постоянно, диск и (ещё реже), делается копия на блюрэй.

Общая идея: архив разбивается на две (или больше) частей. 1-ая всегда монтируется r/o и содержит холодные и редко изменяемые данные. Эта часть лежит на одной группе томов (рейд сконфигурирован так, что несколько групп томов лежат на своих физических дисках) и с неё достаточно редко делается резервная копия на блюрэй.

Для изменяемых данных выделен том, который лежит на другой группе томов (читай, другом физическом диске и в ещё одном случае, на другой ноде). Для хранения состояния файлов, сделано весьма идиотское решение, но опять-таки, практика показала его полезность. Раз в 30 минут ресурс попросту сканируется и если есть изменения, коммитится в Git. База гита лежит перекрёстно, на другом томе (= > другом физическом устройстве). В основном, люди работают с доковскими и автокадовскими файлами, но как оказалось, несмотря на очевидную неатомарность, файлы в гите не бьются. За счёт сжатия, база гита пухнет умеренно.

Раз в три месяца, изменяемые части ресурса сканируются на предмет не изменявшихся за это время файлов и таковые переносятся в r/o хранилище. База гита архивируется и отправляется на хранение.

За счёт того, что изменяемая часть ресурса относительно маленькая, всё работает весьма шустро и не жрёт ресурсы.

Изменяемые и холодные файлы "слиты" в один ресурс через aufs. А поскольку факт "стирания"/изменения файла в r/o хранилище aufs имитирует созданием спецфайла / копии изменяемого файла, работа с таким комбинированным хранилищем прозрачна. В то же время, при работе "шифровальщика" или (не)преднамеренном удалении, страдает весьма незначительная часть хранилища, инцидент легко разобрать руками и файлы восстанавливаются моментально. Перекрёстное хранение на разных физических устройствах r/o, r/w и слепков истории изменений файлов, делает сон значительно спокойнее: потерять сразу ВСЁ можно, но значительно менее вероятно, чем при задействовании механизмов перераспределения LVM/снапшотов btrfs. С последними тоже экспериментирую, но более осторожно.

Вот как-то так.

Ну и неоднократно приходилось доставать копии файлов с совсем небольшим временным лагом и/или сравнивать с "историческими". В обычной схеме со сжатым бэкапом можно было-бы рехнуться.

PS: грабли, на которые можно наступить: если нужно поменять ACL/права доступа, это нельзя делать на объединённом ресурсе: aufs тупо скопирует на r/w раздел файлы из всех слоёв. Но если последовательно применить setacl ко всем слоям, изменения подхватятся и будут нормально работать.

PPS: overlayfs в такой схеме работает скверно.

PPPS: права (любые) в гите не сохраняются. В данном случае это проблема, но схема всё равно оказывается спасительной.