Написать данную заметку меня побудила ситуация вокруг хостера ihor.ru, с которым я сотрудничаю уже лет 5. Еще летом у них начались какие-то проблемы, а сейчас они вылились в открытое противостояние собственников. В чем там конкретно проблемы, я не знаю. Есть 2 противоборствующие стороны, у каждой из которых свой канал в телеграме — @ihor_hosting и @marosnet. Там есть подробности. Они делят между собой компанию. А на пользователях это сказывается тем, что постоянно что-то не работает. То связи нет, то сервера выключаются, тех. поддержка не работает вообще, платежи не принимают. В общем, жуть.

Содержание:

Введение

За столько лет я оброс всевозможными сервисами и услугами этого провайдера. Многим его советовал. Я туда десятки (может сотни) клиентов привел, в том числе через этот сайт. Долгое время у меня висела его реклама, так как я сам пользовался его услугами.

Последние дни торопливо переносил оттуда все критично важное (в основном в selectel). Оставил только то, что было оплачено заранее сильно вперед (например, дедики с оплатой на год). В связи с этой ситуацией, хотел еще раз обратить внимание на важность бэкапов и мои подходы к их созданию и обслуживанию.

Бэкапы нужно обязательно разворачивать и проверять

Мало просто настраивать мониторинг бэкапов и оповещения. Я регулярно вручную проверяю все бэкапы. Да, это очень хлопотно, но другого выхода я не вижу. Если я этого не делаю, то перестаю спокойно спать.

Вот свежий пример, когда большая продуктовая база бэкапилась, но бэкап не проверялся. Автору повезло, что потерял только несколько записей. А мог и всю таблицу. Там, где можно, я автоматизирую разворачивание на запасном сервере с помощью простых bash скриптов. У меня бывали всякие ситуации, когда бэкапы по каким-то причинам переставали делаться. Если они физически отсутствуют, я получаю уведомление от zabbix. Но этого мало.

Особенно важно проверять дампы баз. Физическое наличие файла с дампом не означает, что вы без проблем его развернете. Надо проверять данные, заливая дамп в базу. Это единственный способ убедиться, что данные реально есть и они актуальны.

Backup должен быть в другом дата центре у другого юр. лица

В описываемом выше случае, из-за проблем у конкретного юридического лица есть шанс потерять разом все — и прод и бэкапы от него. Я очень часто сталкиваюсь с ситуацией, когда заказчики предлагают купить сервис для бэкапов у того же хостера, где работает прод. Всегда от этого отговариваю, так как проблемы бывают разные.

Вас могут банально заблокировать за неуплату, из-за ошибки. Хостер может разорвать отношения в одностороннем порядке. Все это я видел и встречал подобные истории от других. Вот пример из одного чата.

С самого сервера с данными не должно быть доступа к бэкапам

Очень важное правило — бэкап сервер сам подключается к целевым машинам и забирает информацию. Это добавляет сложности к сервисам, с помощью которых вы будете делать бэкап. Не получится просто купить где-то место в s3, nfs, smb и класть туда бэкапы с серверов. Нужен активный агент, который

будет ходить по целевым серверам и собирать данные сам. Зачем такие сложности?

Представьте ситуацию, что на целевой сервер попадает злоумышленник или вирус. Он шифрует не только ваши данные, но и бэкапы, к которым есть доступ с машины. Иногда я делаю 2 разных бэкапа — в первое хранилище целевой сервер кладет данные самостоятельно. А с этого хранилища второй бэкап-сервер сам забирает данные. К нему нет доступа ни у кого. Для большей безопасности можно затирать в логах следы подключения, чтобы нельзя было получить информацию не только о местоположении второго сервера, но и в целом информации о том, что он существует.

Для подобного рода бэкапов хорошо подходят бюджетные vps с большими дисками. Такая услуга, к примеру, есть у ruvds. В списке услуг выбирайте **большой диск**. Доступен не на всех тарифах и датацентрах.

Бэкапы должны быть максимально полные и подробные

Что я имею ввиду? Зачастую бэкапы делаются с расчетом на то, что что-то изменится на проде и нам надо будет развернуть бэкап и посмотреть на него. Бэкапятся, к примеру, исходники сайта и база к нему. Когда все в порядке и ты просто проверяешь бэкап, сложностей не возникает, потому что если ты даже что-то забыл, то залез на prod и посмотрел.

Например, у сайта может быть сложный и насыщенный конфиг nginx с кучей location и редиректов. В процессе эксплуатации вы могли тюнить и вносить много изменений в настройки mysql сервера. Если конфиги не забэкапить вместе с сайтом, то разворачивать и запускать его в работу может быть очень трудозатратно и хлопотно, особенно если оптимизация и настройки выполнялись давно и все уже забыто.

Отсюда правило — бэкапьте не только данные, но и все окружение из расчета на то, что вы разом можете потерять prod навсегда. Я много раз об это спотыкался и по второму разу настраивал все то, что у же было настроено ранее. Сейчас все конфиги проектов храню в своем gitlab. Из того, что часто забывают — ssl/tls сертификаты. А потом их приходится торопливо искать.

Backup виртуальных машин

Я знаю, что многие бэкапят целиком виртуалки. Я редко делаю такие бэкапы, так как с ними трудно работать. Получаются файлы огромных размеров. Если гоняете их через интернет, то они могут биться по дороге, либо в момент создания. У меня были ситуации, когда нужно было развернуть бэкап виртуальной машины на 200 Гб. Я несколько часов его заливал через интернет, а потом он не разворачивался из-за ошибки чтения. Пробовал несколько раз и каждый раз неудачно. Очень повезло, что заметил это в момент тестового разворачивания, а не тогда, когда была бы реальная авария. Это был бы

полный провал, так как это был единственный экземпляр архивной копии.

Если нужно делать бэкап виртуальной машины, то я делаю небольшой системный диск (30-50 Гб) и бэкаплю только его. Данные кладу на отдельный виртуальный диск и забираю их в сыром в виде с помощью rsync или аналогов. С его помощью легко и быстро делать инкрементные бэкапы, а потом их разворачивать. Нет проблем с докачкой и битыми файлами. Даже если что-то и повреждено, то 1 файл из десятков тысяч погоды не сделает. Можно пережить.

Проверяйте скорость восстановления данных

В завершение своего списка еще одна рекомендация — **проверяйте скорость доступа к бэкапам**. У меня были ситуации, когда надо срочно восстанавливать данные. Ты начинаешь их загружать с сервера бэкапа и понимаешь, что скорость катастрофически низкая.

В обычное время, во время проверок, этого можно не заметить, потому что спешить некуда. А вот если случилась авария и нужно как можно быстрее восстановить данные, скорость доступа становится очень критична.

Есть дешевые хранилища, где явно не указано, что скорость доступа будет низкой. Но нужно это понимать, если цена за хранение действительно ниже, чем в среднем по рынку. Многие хостеры на хранилищах с безлимитным трафиком на самом деле этот лимит имеют и включат вам его после того, как вы превысите определенный порог в скачанных данных. Вам просто сделают не 100 мегабит полосу, а 5-10.

У меня были ситуации, когда сервер с бэкапами живет на гигабитном порту с «безлимитным трафиком». А когда ты скачаешь 10-15 гигабайт, включается ограничение полосы до 100 мегабит. Возможно будет и дальнейшее урезание. С такими лимитами нет смысла платить за гигабит. Это просто маркетинговая уловка.

В общем, не доверяйте данным из описания тарифа, проверяйте их сами, чтобы потом не было сюрприза.

Заключение

> Не понравилась статья и хочешь научить меня администрировать? Пожалуйста, я люблю учиться. Комментарии в твоем распоряжении. Расскажи, как сделать правильно!

Это мои основные правила создания бэкапов. А как вы делаете бэкапы? Поделитесь своим опытом. Возможно, я что-то забываю или упускаю.