Инкрементное резервное копирование

Это небольшой tutorial о том, как правильно и экономно делать резервное копирование (backup) данных на сервере.

Всё началось с того, что файлы нашего сервера (tuffle.com) стали занимать > 300GB. Естественно, эти данные очень ценны для нас и мы должны позаботиться о их сохранности. Первым делом мы настроили автоматический мастер резервного копирования в ISP Manager. Он справлялся не так уж плохо, когда данных было немного. Потом в день “X” было обнаружено, что каждый архив бэкапа весит очень много и что процесс копирования длится около 10 часов. Естественно, это сказывается на производительности сервера.

Решение.

Главной ошибкой было то, что хранились дубликаты файлов. Именно эту проблему решает инкрементное резервное копирование. Оно сохраняет только уникальные файлы.

Как это работает.

Один раз в день (лучше ночью, когда посетителей минимальное количество) мы сохраняем только файлы, которые были загружены в этот день. Бэкапы хранятся за последние N дней. Есть общий архив, который хранит все файлы до первого бэкапа и отдельные ежедневные архивы с изменениями.

Какие инструменты использовались? Консоль и bash Создадим папку, где будут хранится наши архивы:

```
mkdir /root/backup
```

Создаем sh скрипт, который и будет заниматься резервированием:

```
touch /root/backup/backup-files.sh
touch /root/backup/exclude.list
```

Со следующим содержимым:

```
#!/bin/sh

cd /root/backup || exit

date=`date +%F`
days=10
host=`hostname`

rsync -aR --delete --link-dest=$PWD/$host.current --exclude-from=exclude.list \
 /etc/apache2 \
 /var/www \
 $host.$date && \
 ( rm -f $host.current; ln -s $host.$date $host.current ) && \
 find $host.* -maxdepth 0 -type d -ctime +$(( $days - 1 )) -exec rm -r {} +
```

Этот скрипт проверяет файлы, копирует новые, создавая папку с датой в названии. В exlude.list запишите файлы и папки, которые не стоит проверять. Осталось запустить этот скрипт, а лучше делать это автоматически. В cron-файле добавить следующую строку:

```
0 3 * * * /root/backup/backup-files.sh
```

Теперь этот скрипт будет запускаться каждый день в 3 часа ночи.

Что с базой данных?

Сейчас мы делаем бэкапы БД обычным mysqldump так как она занимает относительно немного места. Если мы выберем более оптимальное и интересное решение, я об этом напишу.