How do I backup CouchDB? What data recovery strategies exist?

Overview

While CouchDB is a very reliable database, a careful engineer will always ask "What happens when something goes wrong?". Let's say your server has an unrecoverable crash and you lose all data... or maybe a hacker finds your top secret credentials and deletes your data... or maybe an undiscovered bug causes data corruption after an event... or maybe there is a logic error in your application code that accesses your database. Ideally we try to avoid these situations by preparing for the worst and hoping they never occur, but bad things do happen and we should be ready to react when they do.

There are a few traditional data backup strategies for CouchDB:

1. Replication
2. Database file backup
3. Filesystem snapshots

Replication Based Backup

CouchDB is well known for its push and pull replication functionality. Any CouchDB database can replicate to any other if it has HTTP access and the proper credentials.

Pro's & Con's

Pros:

- good
- good
- good

Cons:

- bad
- bad
- bad

Database File Backup

Under the hood, CouchDB stores databases and indexes as files in the underlying filesystem. Using a common command line back up tool, like rsync, we can perform incremental backups triggered by cron.

Pro's & Con's

Pros:

- good
- good
- good

Cons:

- bad
- bad
- bad

Filesystem/VM Snapshots

Most VM's and newer filesystems have snapshot capabilities to allow roll backs to preserve data.

Pro's & Con's

Pros:

- good
- good
- good

Cons:

- bad
- bad

- bad

Summary

STUB