Backup and disaster recovery

This page covers the basics of backups and disaster recovery for storage providers. A backup strategy is only as good as the last successful restore.

It is crucial to have a backup of any production system. It is even more crucial to be able to restore from that backup. These concepts are vital to a Filecoin storage provider because not only are you storing customer data for which you have (on-chain) contracts, you have also pledged a large amount of collateral for that data.

If you are unable to restore your Lotus miner and start proving your storage on-chain, you risk losing a lot of money. If you are unable to come back online in 6 weeks, you are losingall of your collateral, which will most likely lead to bankruptcy.

As such it matters less what kind of backup you have, as long as you are able to restore from it fast.

High availability (HA) versus Disaster recovery (DR)

It is a common misconception to assume you are covered against any type of failure by implementing a highly available (HA) setup. HA will protect against unplanned unavailability in many cases, such as a system failure. It willnot protect you against data corruption, data loss, ransomware, or a complete disaster at the datacenter level.

Backups and (tested) restores are the basis for a DR (disaster recovery) plan and should be a major point of attention for any Filecoin storage provider, regardless of your size of operation.

Recovery Time Objective (RTO) and Recovery Point Objective

When planning for backup and recovery, the terms RPO and RTO are important concepts to know about.

- Recovery Time Objective (RTO)
- is the time taken to recover a certain application or dataset in the event of a failure. Fast recovery means a shorter RTO (typically measured in hours/minutes/seconds). Enterprises plan for very short RTOs when downtime is not acceptable to their business. Application and file system snapshots typically provide the lowest possible RTO.
- Recovery Point Objective (RPO)
- is the last known working backup from which you can recover. A shorter RPO means the time between the last backup and the failure is short. Enterprises plan for very short RPOs for systems and data that changes very often (like databases). Synchronous replication of systems and data typically provides the lowest possible RPO.
-

RPO/RTO for storage providers

Although ‘RPO zero’ and ‘RTO zero’ are the ideal, in practice it is rarely economical. DR planning requires compromises and if you are a storage provider you need to consider cost versus RPO.

RTO is typically less concerning for storage providers. The most critical parts to recover are your sealed storage and your wallets. Wallet addresses typically do not change, so the only thing to worry about is your sealed storage. With storage level snapshots (such as ZFS snapshots), you can reduce your RTO to almost zero.

For RPO, although synchronous replication, together with snapshots, can reduce RPO to nearly zero, that is not a cost-efficient solution. Asynchronous replication of sealed storage is the most viable option if you are running at small-to-medium scale. Once you grow beyond 10PB of storage, even replicating the data will become an expensive solution.

In such cases you might want to look intostorage cluster solutions with built-in redundancy. Very large storage providers will operateCeph clusters or other solutions with built-inerasure coding . Although this does more become more like a HA setup than a DR setup, at scale, it becomes the only economically viable option.

Running a storage cluster comes with its own operational challenges though, which does not make this a good fit for small-to-medium setups.

RPO/RTO for customers

Both storage providers and data owners (customers) should look at RPO and RTO options. As a customer, you can achieve HA/DR by having multiple copies of your data stored (and proven) across multiple storage providers. In the event of data loss at one provider, other providers will hold a copy of your data from which you can retrieve. As a customer, you choose how much redundancy you need, by doing storage deals with more providers.

RTO for data owners is a matter of how fast the storage provider(s) can provide you the data.

- Do your storage providers offer “fast retrieval” of the data through unsealed copies? If not, the unsealing process (typically multiple hours) must be calculated into the RTO.
- Do your storage providers offer retrieval throughSaturn, (the Web3 CDN)
- for ultra-fast retrieval?

- Do your storage providers pin your data on IPFS, in addition to storing it on Filecoin?
-

RPO for data owners is less of a concern, especially once the data is sealed. The Filecoin blockchain will enforce availability and durability of the data being stored, once it is sealed. It is therefore important, as a data owner, to know how fast your storage provider can prove the data on-chain.

Backup techniques

- A first level of protection comes from ZFS (if you are using ZFS as the file system for your storage). Having ZFS snapshots available protects you against data loss caused by human error or tech failure, and potentially even against ransomware. Other file systems typically also have a way to make snapshots, albeit not as efficiently as ZFS.
- A second level of defense comes from a dedicated backup system. Not only should you have backup storage (on a different storage array than the original data), you also need to have a backup server that can at a minimum run the Lotus daemon, Lotus miner and 1 WindowPoSt worker (note: this requires a GPU). With that you can sync the chain, offer retrievals and prove your storage on-chain, from your backup system, whilst you bring your primary back online.
- An alternative technique to having a dedicated backup system and copy is to have a storage cluster. This still requires a backup system to run the Lotus daemon, Lotus miner and PoST worker on. Implementing a storage cluster is usually only done for large-scale deployments as it comes with additional operational tasks.
-

For maximum resilience, you could host your backup system (server + storage) in a different datacenter than your primary system.

DR failover techniques

One way to prepare for an easy failover of the software components in the event of a failure is to configure floating IP addresses. Instead of pinning lotus daemon and lotus-miner to the host IP address of the server they are running on, you can configure a secondary IP address and pin the daemon to its own IP, and lotus-miner to yet another IP.

This helps to reduce the amount of manual tasks for a failover drastically. If the recovered daemon or miner instance changes IP address it requires quite a lot of reconfiguration in various places.

Having the services on a floating IP allows to assign this IP to another machine and start the service on it.

Previous Network Next Reference architectures

Last updated6 months ago