3- Tested Disaster Recovery

We introduce in this chapter key principles to implement a disaster recovery solution for Cloud Computing. The approach we suggest is inspired by quality assurance in software engineering. It combines the ability to test at any time the reconstruction of an application and the use of a decentralized infrastructure operated by independent suppliers.

Step 1: define a test case

The first step for disaster recovery plan is to define a test case based on which one can decide whether the disaster recovery plan was successful or not. This is just the same as defining a test case based on which one can decide whether a software application is working or not.

A typical test case for an accounting software could consist of producing the daily balances for the current month, producing monthly balances for the last year and yearly balances for the last decade, then comparing the base lines with daily balances printed and stored company sites located on 3 continents. The end of the test could consist in entering a new accounting transaction and making sure it is taken into account in the balance of current day.

This tests also involves the ability of skilled users to access the application through a network.

Step 2: define how to rebuild the environment

The second step consists of defining how one can rebuild the application environment from its operating system, source code, CDs, etc. and configure it. The steps can be made manually, following instructions of an installation manual, or automatically using some system build and configuration tool.

In the case of Virtual Machine images, ability to be able to reconstruct the system at any time is critical to cover the risks of filesytem corruption, system corruption and application corruption. A clean image of operating system should thus be provided or recreated if needed in the plan.

Step 3: define how to backup data

The third step consists of defining how the data of the application can be backed up using a historized backup technology. The notion of history is required to protect from the use of malware which could lead to mass destruction of data.

This step is often more complex than expected. Many applications store data in different places. Records are stored in a database whereas images are stored in the filesystem. It is quite easy to forget one data source in the plan.

Step 4: reserve recovery vaults

The fourth step consists of selecting recovery vaults on multiple continents from multiple providers. Recovery vaults can be baremetal servers or virtual machines. They can be hosted in certified data centers or in small offices. Their location can be well known or kept secret.

As long as strong encryption is used, we believe that the higher the diversity of data vaults, the lower the risk. Using specialized data centers brings in our opinion more risks in terms of data loss than using a grid of low cost data centers.

We also believe that relying on vaults which location is secret or at least requires multiple intermediaries to be known is the best protection against Disaster Case 1 "intentional simultaneous destruction".

Step 5: automate recovery tests

The last step consists of testing every day the possibility to reconstruct the environment, restore its data from backup and run the test cases. This should be done on at least one of the recovery vaults. In case data is encrypted in most disaster recovery vaults, testing that the SHA hash code of encrypted data and application is the same as the SHA hash code of data and application on the tested recovery vault is sufficient.