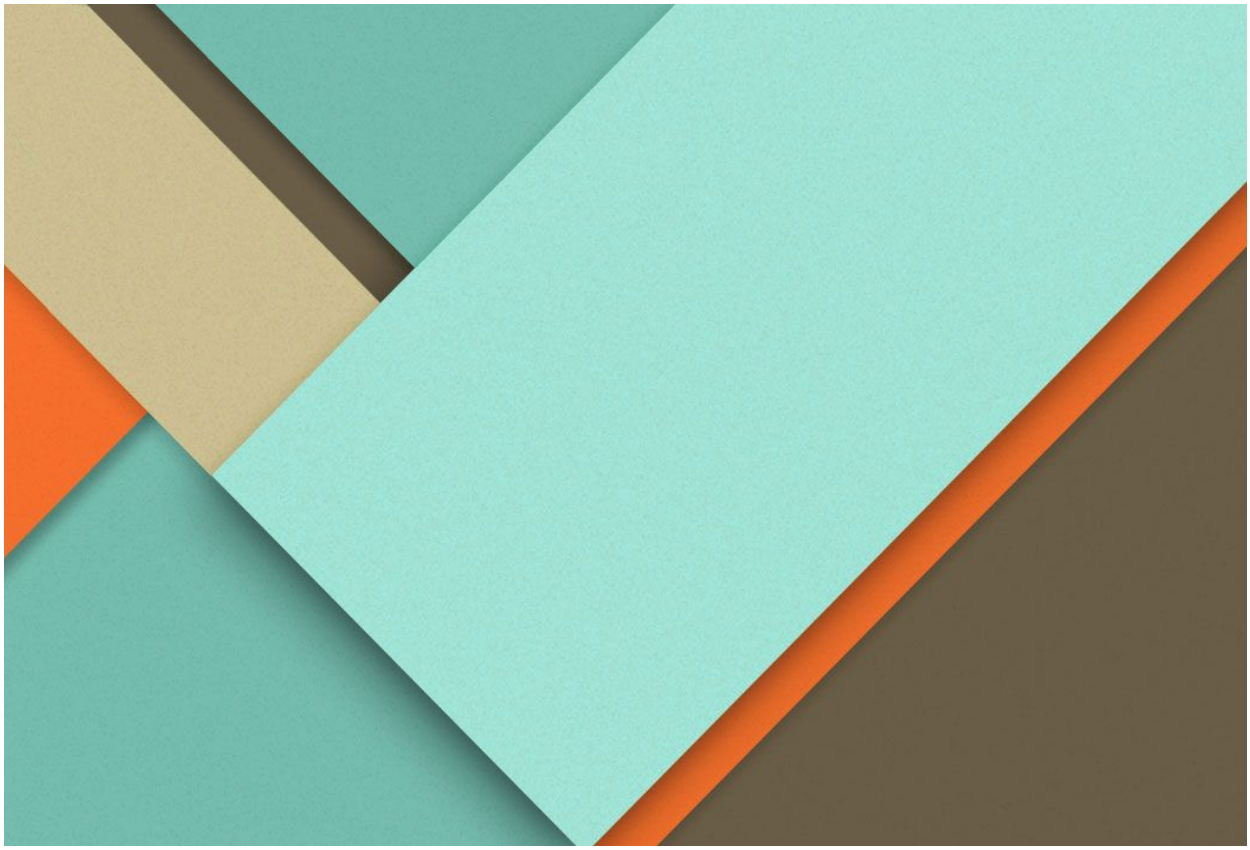


Data Protection

Prevent data loss from operational errors or system failures

Group 9
Zeyu Yao
Tao Yu
Morgan Greer
Jeremy Wong

Overview

Data protection is the process of safeguarding important information from corruption, loss or endangerment. Data protection needs to be taken very seriously and careful planning needs to be done before implementation. There are many different types of data that needs protecting and there are many different ways to go about protecting it.

What makes data protection so important?

Without protecting their data, organisations will lose important information required to run their business. Also, if a business is producing a service for their consumers, they will need to protect personal information. If this information were to get out it could compromise consumers and the businesses integrity. Data protection is one of the most important processes businesses must invest in.

Strategies

There are several concepts related to data protection, we have chosen to implement one out of the 3 listed below:

- Mirroring Storage
    - Is the replication of data onto separate physical hard disks in real time to ensure continuous availability.
- Storage Snapshots
    - Is a set of reference numbers that points to data that is saved at a certain time/date.
- Traditional Database Backup **(Implemented)**

Traditional Database Backup

This method makes a copy of data and stores it on another drive or server for safe keeping. This was the main system in use before the introduction of cloud computing, and is still in use today. As this is a manual process, administration time is required and schedules that are frequent enough to protect recently created data. An advantage of traditional database backup is the simplicity for implementation.

Do it in DevOps approach

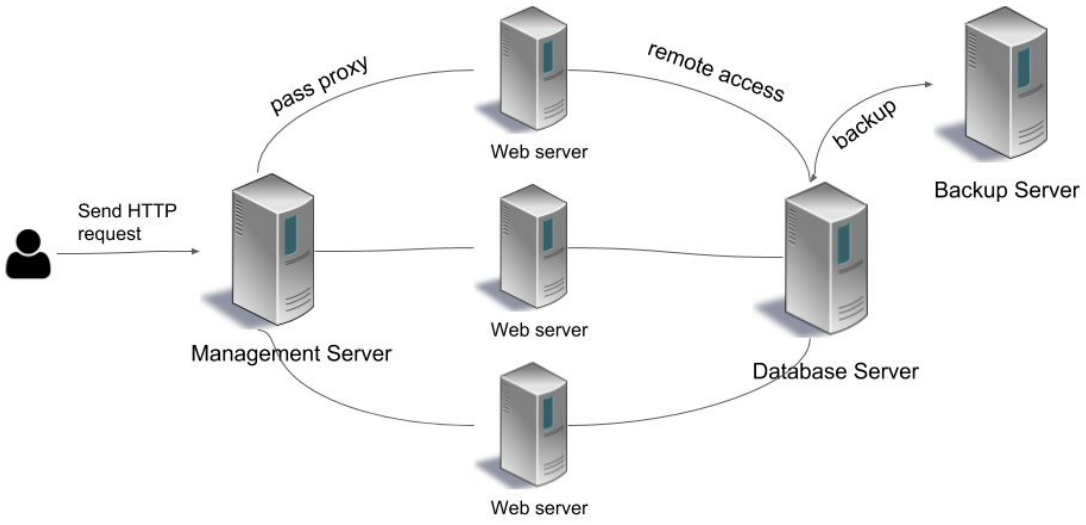
There are two Devops principles highly relevant in this case: continuous improvement and automation.

Continuous improvement: Organisations must continue to develop and change in the face of changing circumstances such as new technology and changing customer needs. Processes should be optimised for speed to improve efficiency and reduce cost.

Automation: Devops stipulates that everything should be automated where possible. This includes the process of database backups. Automation saves time and resources which can be used elsewhere in an organisation.

How did we implement?

Firstly, we still use a distributed system with a management server, three multiple web servers and a remote database server. The backup process has been executed periodically, the frequency can be specified (eg. 1 minutes/backup or 1 hours/backup). There is another backup server that keep accessing the remote database server and backup the database into a gzip file locally. When the remote database server was corrupted or collapsed, this backup server will be brought back online and restore the databases using the backup sql file.



Challenges and Comparison

While traditional data backup is simple, there are a number of possible drawbacks to consider. Firstly, this system backs up data periodically which means if the system was to suddenly fail all data after the most recent backup could be lost forever. Also, using storage media can incur high software and hardware costs depending on the size of the backup.

Cloud backup is an alternative way of backing up data that involves sending a copy of data to a server which is off-site.

Key advantages include:

- Improved scalability: Costs can be reduced if data volumes are low, and can also be increased easily.
- The service provider will take care of many tasks involved with the backup freeing up time and resources.
- More secure against ransomware attacks and other security problems due to the data being off-site on alternative servers and systems.

Key disadvantages:

- Speed: Depending on the bandwidth and active users the backup can be slow.
- Large amounts of data may escalate cost making traditional backups cheaper.
- An outside organisation has access and control.

Continuous data protection is becoming increasingly popular with organisations, and offers another alternative to data back-up.

Key advantages:

- Every transaction is backed up making it the most secure method of backup.
- If the system is infected with a virus or corrupted, a very recent copy of the affected file can be easily restored. Traditional backup may not be frequent enough to do this.
- Data recovery can occur in just seconds, and the installation of hardware is straightforward.

Key disadvantages:

- Performance may be impacted while confirmations between systems and data backups occur.
- Network bandwidth will affect backup restoration times significantly.
- Proprietary hardware is required making it more complex than cloud backup, and can be more costly than traditional backup.

[References]

6 Principles of DevOps by *DevOps Agile Skills Association*

Definition of Cloud backup from Whatls by *Margaret Rouse*

Pros and Cons of Continuous Data Protection by *Jim Lippie*