



Disaster Recovery

Databases are an integral component of any modern business, and their failure can often be quite devastating and disrupting to its operations. Planning for disasters and setting up contingencies for database related disasters is hence considered as an important task by many companies, particularly those related to database recovery.

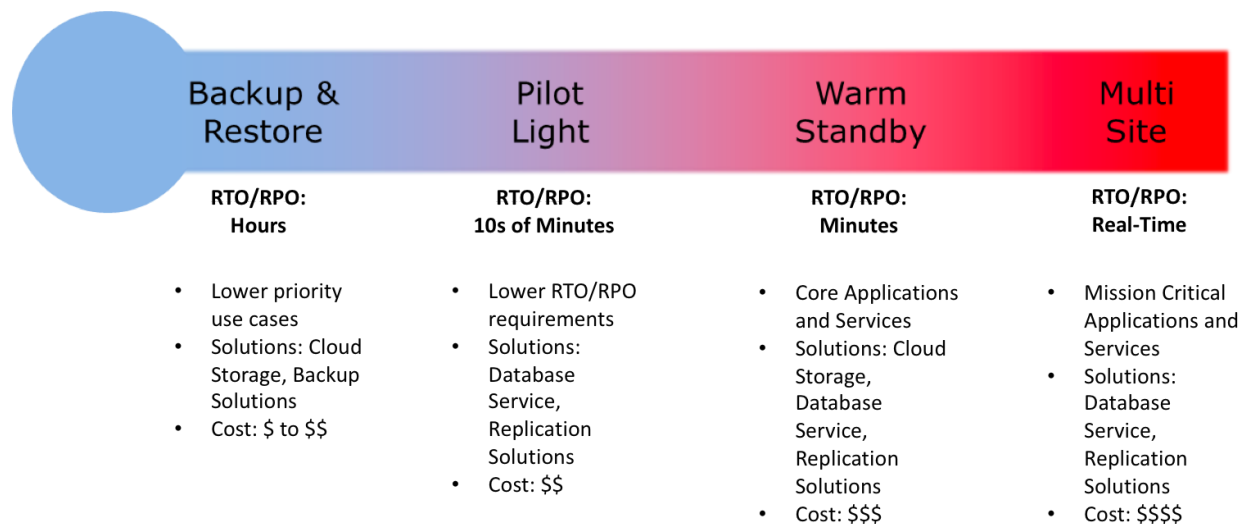
There exist however, a myriad of different approaches and practices for disaster recovery which are measured using the metrics **Recovery Point Objective (RPO)** and **Recovery Time Objective (RTO)**.

RPO defines the maximum allowable data loss in terms of time; for instance, if a business has an RPO of 30 minutes, it means that, in the event of a disaster, they can only tolerate losing data from the last 30 minutes.

RTO on the other hand, specifies the maximum time allowed for system recovery after a failure. For example, if a company has an RTO of 2 hours, it indicates that they must restore their database and resume normal operations within that timeframe following an outage.

As such, a low RTO/RPO method of disaster recovery is acceptable when downtime is not a dealbreaker, and the business can handle some amount of lost data whilst a high RTO/TPO method of disaster recovery is more suited for databases that must be maximally available and cannot handle the loss of data.

An in-depth discussion of all the disaster recovery methods is beyond the scope of this book, and hence I will be relying on a diagram made available by AWS. Consult the diagram for a brief overview of all the disaster recovery methods and their respective RTO/RPOs:



I encourage the readers to research the different disaster recovery methods listed above on their own, searching the internet for their accompanying architecture diagrams as well if desired.