A common example used in DR discussions is the failure of an entire datacenter, and will typically involve all the hardware in one geographic location simultaneously shut down. A good DR strategy will be able to recover from this by recreating the same infrastructure (compute, networking and storage) in another geographical location within a specific Recovery Time Objective (RTO).

These methods can be combined with the backup strategy to ensure you are backing up the platform effectively to allow you to recover from any given failure. When doing the system architecture it is important to consider the backup strategy and also if there is a need for high availability (HA) or disaster recovery (DR). The following specifications must be taken into account: Recovery Point Objective (RPO) and Recovery Time Objective (RTO).

The backup strategy will grant that the system is restored to a given time when the backup was taken and in a case of the disaster, the data created between the time that the backup was taken and the moment of the failure will be lost in most of cases. This should be considered when creating the backup policy

The high availability environment will grant that the servers still responsive including when one or more nodes fail. So other servers will remain online and perform the tasks. When thinking about HA it is frequently referred as the local system redundancy and each of HA environments should be able to handle the full load. The main difference between the HA and backup is that if data gets corrupted in one of the nodes, it will be propagated to the other nodes, with that both nodes will having the failure, and will need the backup to resume to its normal operation.

The disaster recovery (DR) environment is a copy of the production (or running) environment and in case of failure this environment will take over the requests and all operation, when the main system is ready to be put back online the data should be replicated from the DR side to the main system.

3.3 Backup options

When considering the backup strategy, it is recommended to verify the most suitable types of backup to meet the requirements. This section will briefly discuss each of the components within the IBM Cloud Private infrastructure and platform, and if/why each should be backed up.

3.3.1 Infrastructure backups

Infrastructure backups are the most common method to ensure a full copy of an IBM Cloud Private node exists in the event of a failure. Suitable tools for infrastructure backups are usually available with the hypervisor hosting IBM Cloud Private.

At the time of writing, there are 6 different types of IBM Cloud Private nodes; boot, master, proxy, management, Vulnerability Advisor and worker nodes. Each node plays a specific role in the cluster, and each will have a different impact on the backup and restore strategy.

- Boot nodes: A boot (or bootstrap) node is used for running installation, configuration, node scaling, and cluster updates. Only one boot node is required for any cluster, and a single boot node can cater for multiple installations of IBM Cloud Private. This node stores the cluster configuration data and is important to at least backup the filesystem so that the original configuration and cluster certificates are able to be reused if necessary. In small clusters, boot nodes are typically combined with the master nodes.