Basics - Nondisruptive Upgrades

In 'Nutanix Software Upgrade' and 'Hypervisor Upgrade' within the Prism section, we highlighted the steps used to perform an upgrade of AOS and hypervisor versions. This section will cover the techniques allowing us to perform different types of upgrades in a non-disruptive manner.

AOS Upgrades

For an AOS upgrade there are a few core steps that are performed:

1 - Pre-upgrade Checks

During the pre-upgrade checks, the following items are verified. NOTE: This must complete successfully before an upgrade can continue.

- Check version compatibility between AOS, hypervisor versions
- Check cluster health (cluster status, free space, and component checks (e.g. Medusa, Stargate, Zookeeper, etc.)
- Check network connectivity between all CVMs and Hypervisors

2 - Upload upgrade software to 2 nodes

Once the pre-upgrade checks have been completed, the system will upload the upgrade software binaries to two nodes in the cluster. This is done for fault-tolerance and to ensure if one CVM is rebooting the other is available for others to pull the software from.

3 - Stage Upgrade Software

Once the software has been uploaded to two CVMs, all CVMs will stage the upgrade in parallel.

The CVMs have two partitions for AOS versions:

- Active partition (the currently running version)
- Passive partition (where upgrades are staged)

When an AOS upgrade occurs, we perform the upgrade on the non-active partition. When the upgrade token is received it will mark the upgraded partition as the active partition and reboot the CVM into the upgraded version. This is similar to a bootbank / altbootbank.

NOTE: the upgrade token is passed between nodes iteratively. This ensures only one CVM reboots at a time. Once the CVM reboots and is stable (check service status and communication) the token can be passed to the next CVM until all CVMs have been upgraded.

Upgrade Error Handling

A common question is what happens if the upgrade is unsuccessful or has an issue partially through the process?

In the event some upgrade issue occurs we will stall the upgrade and not progress. NOTE: this is a very infrequent occurrence as pre-upgrade checks will find most issues before the upgrade actually begins. However, in the event the pre-upgrade checks succeed and some issue occurs during the actual upgrade, there will be **no impact to workloads and user I/O running on the cluster.**

The Nutanix software is designed to work indefinitely in a mixed mode between supported upgrade versions. For example, if the cluster is running x.y.foo and is upgrading to x.y.bar the system can run indefinitely with CVMs on both versions. This is actually what occurs during the upgrade process.

For example, if you have a 4 node cluster on x.y.foo and start the upgrade to x.y.bar, when the first node upgrades it will be running x.y.bar while the others are on x.y.foo. This process will continue and CVMs will reboot into x.y.bar as they receive the upgrade token.