MarkLogic

MarkLogic in the Cloud Workshop

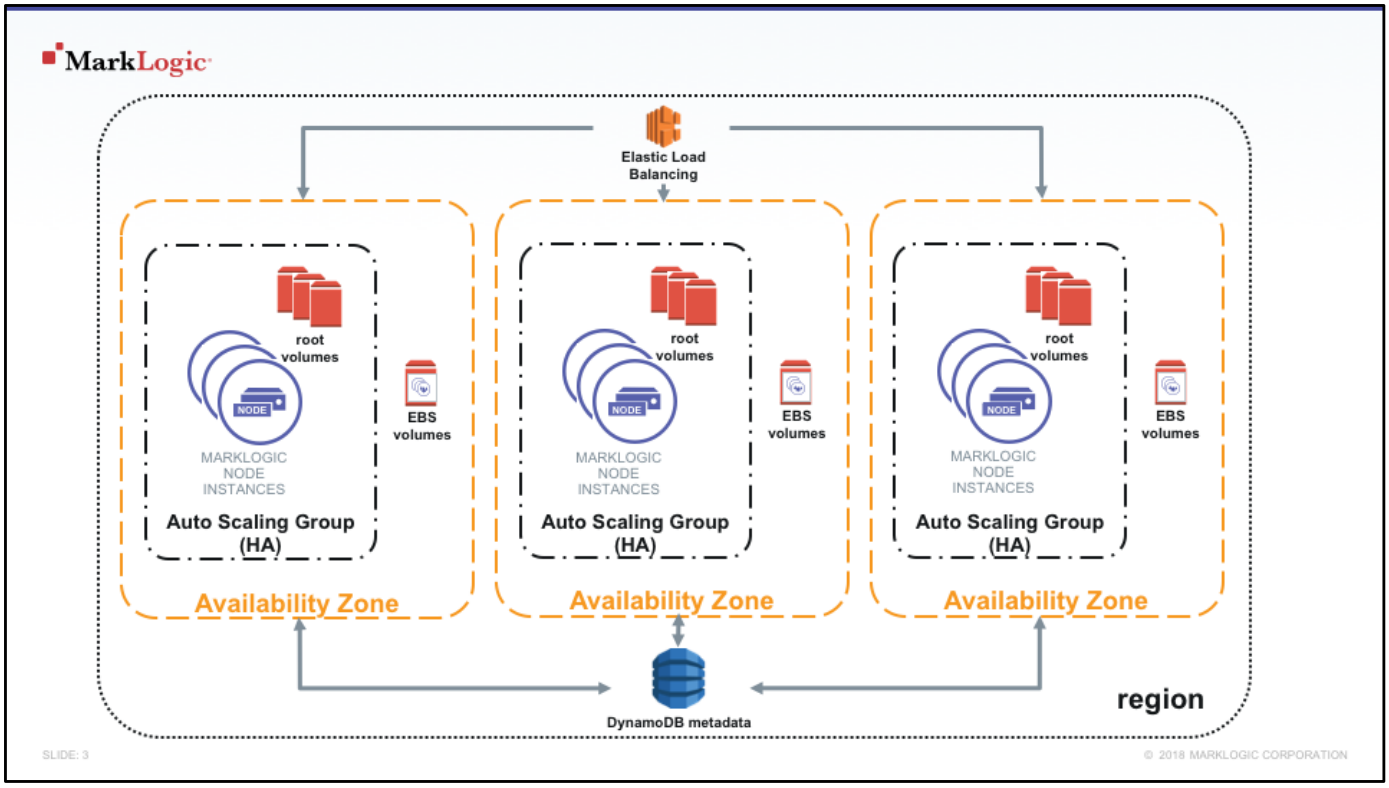
Unit 2 – Node Recovery

MarkLogic

Goal: Investigate Node Loss Recovery

- What happens when a node is unhealthy?
- How does recovery affect my data?

We will learn about what occurs should one or more MarkLogic server instances in AWS become unhealthy. We will look at how they recover and how that affects your data.



MarkLogic's Managed Clusters feature help mitigate challenges in the cloud that you'd perhaps not encounter with on-premise clusters.

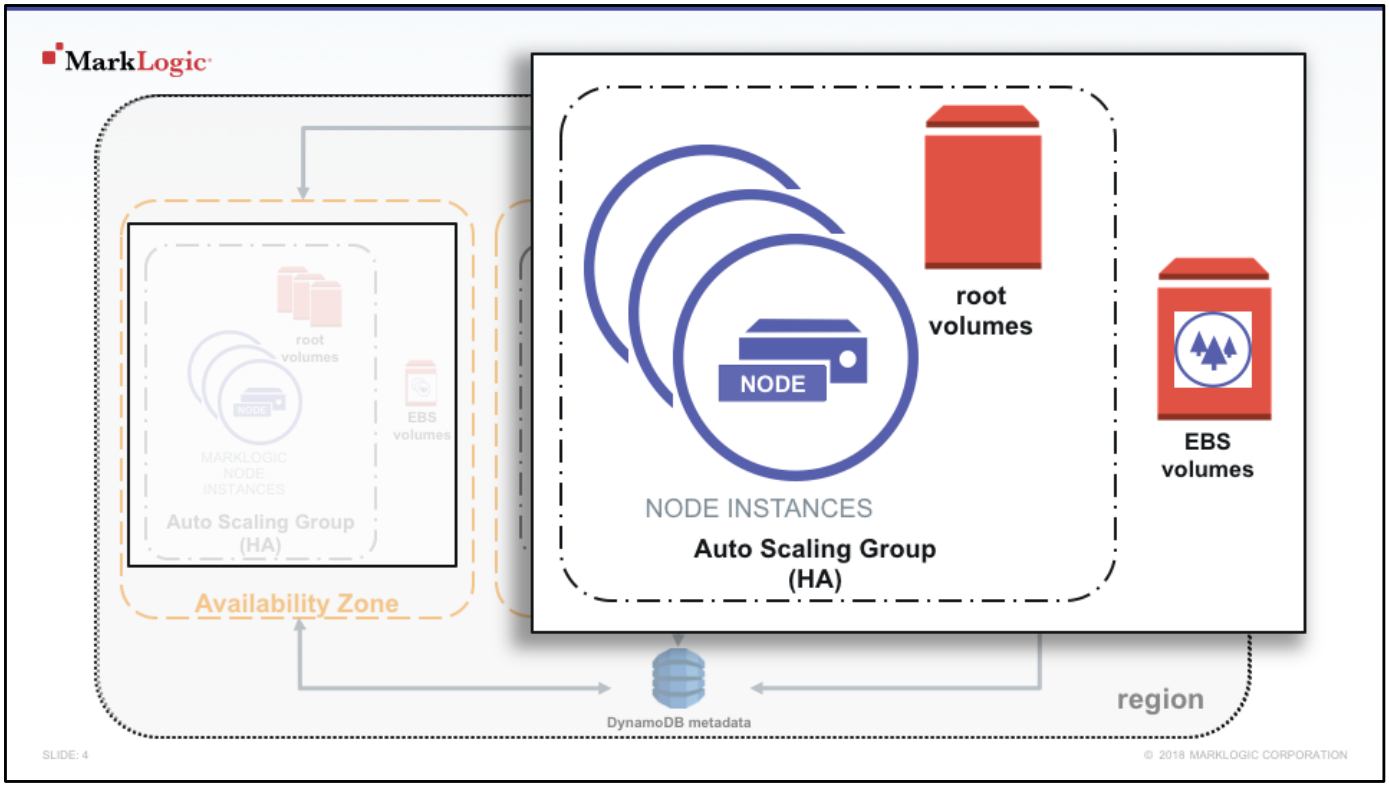
For example, the Managed Cluster feature automatically handles creating your Administrator credentials then configuring remaining nodes to connect to the cluster. The Managed Cluster feature also automatically applies your license keys to all of the nodes in the cluster.

When EC2 instances are terminated, the root volumes are released, deleting any installed applications and data. MarkLogic configuration data and your content are not put on root volumes, but rather on attached EBS volumes. MarkLogic's Default Data directory lives on an EBS volume, not a root volume.

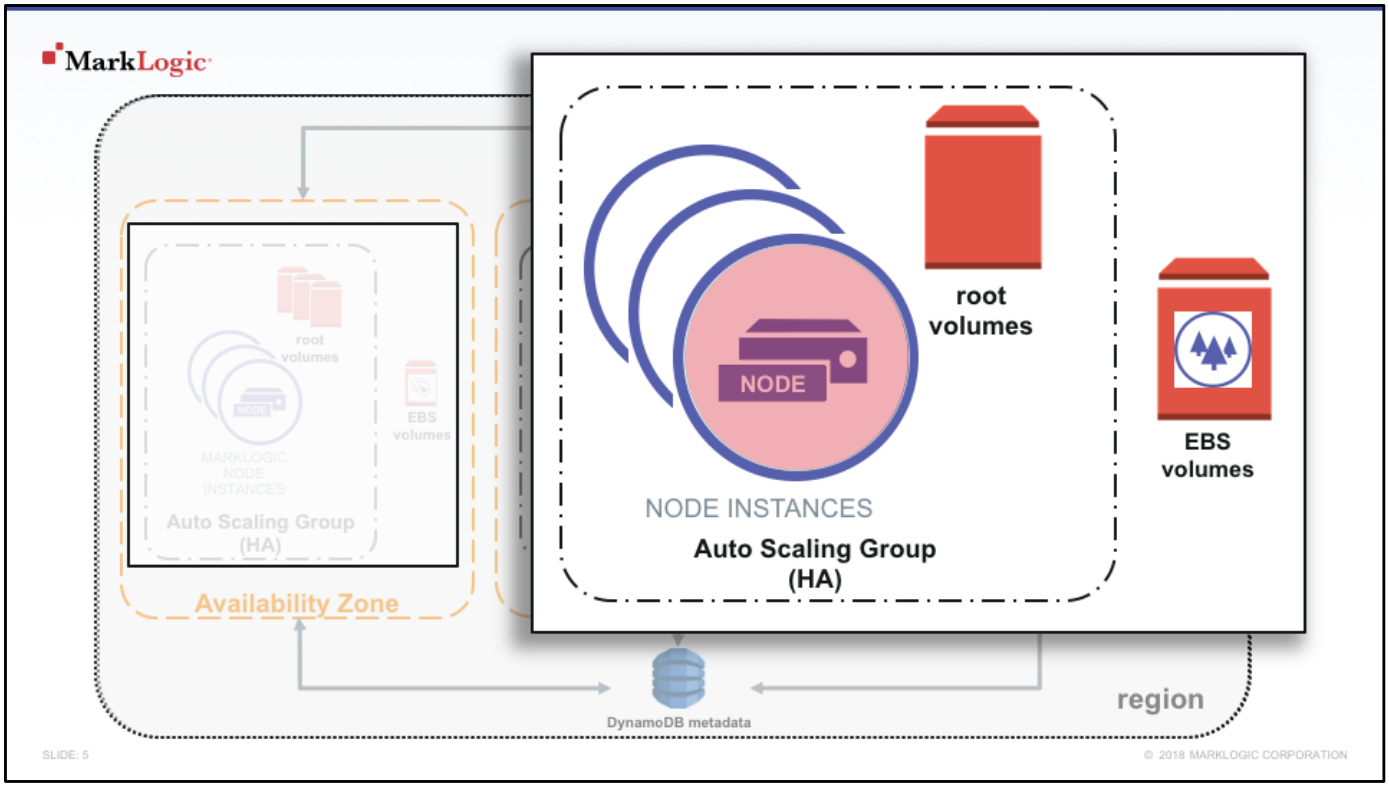
The Managed Cluster feature automatically keeps track of each EBS volume, along with its related EC2 instance and mount directory. When MarkLogic EC2 instances restart, the Managed Cluster feature automatically re-attaches and mounts your volumes to the appropriate locations to ensure that your Forests and Databases remain intact.

When MarkLogic node instances start or restart, they receive new, unique hostnames from AWS. The Managed Cluster feature automatically detects hostname changes and propagates the changes throughout the cluster ensuring these new node instance re-join the cluster properly.

The DynamoDB database that stores and indexes metadata such as all of the configuration data required to manage a cluster of one or more MarkLogic Servers. For details on AWS' DynamoDB, see http://docs.aws.amazon.com/amazondynamodb/latest/developerguide/GettingStartedDynamoDB.html.

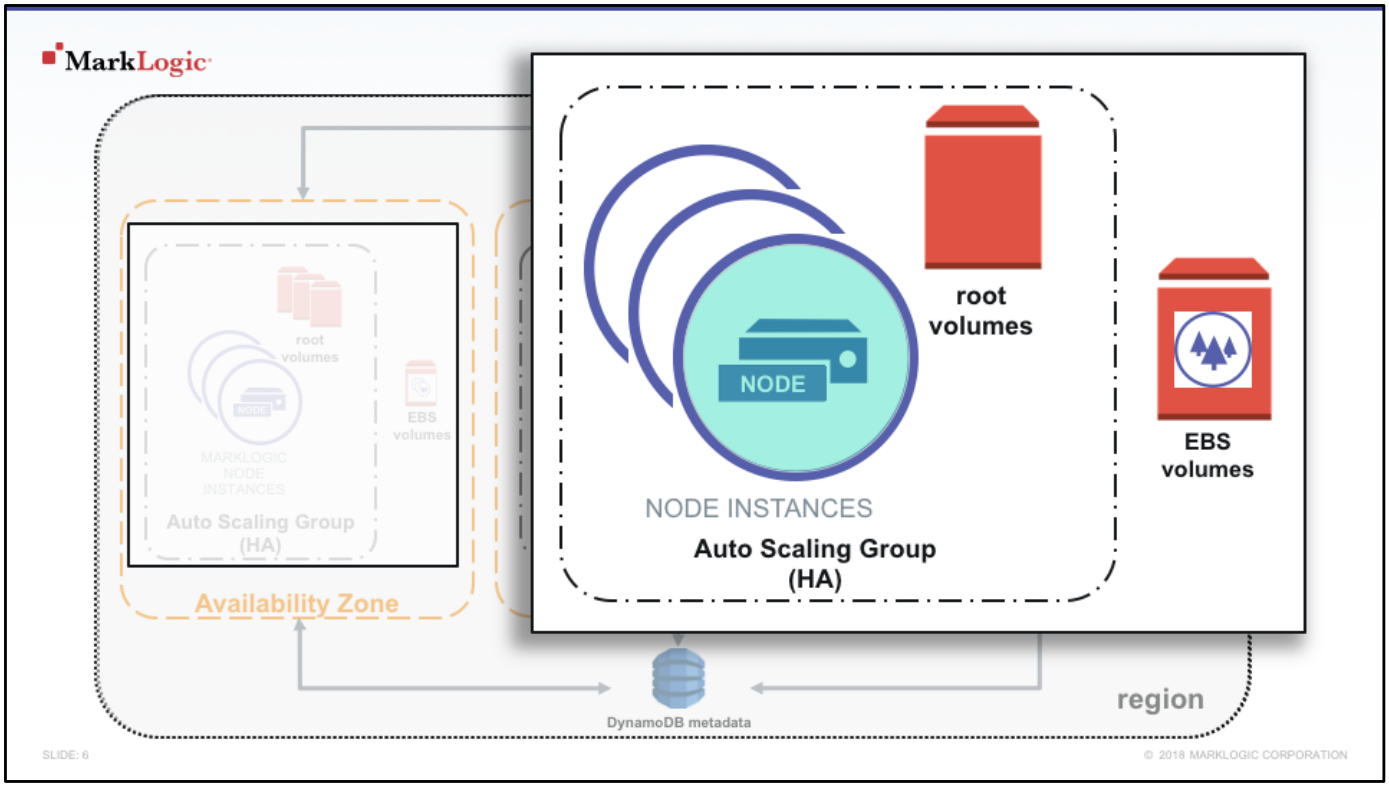


Data EBS volumes are maintained separately from the AMI instance root volumes. When creating a CloudFormation stack, the template asks you for the size of these EBS volumes. These store your content in the Forests. Other protections such as ensuring MarkLogic backup schedules to S3 buckets and replication for high availability should also be considered.



When a MarkLogic EC2 instance is unhealthy, a Managed Cluster feature, the HealthCheck application server, will either report the application server as not being contactable or, in a worst case scenario, will not respond at all. AWS will take action by terminating the instance.

When the instance terminates, its root volume also terminates. However, the EBS volumes containing your contents and configurations remain, protected but currently not mounted to a MarkLogic EC2 instance.



The Managed Cluster feature starts a new MarkLogic instance to replace the unhealthy one. The new EC2 instance gets a new root volume and the startup script for MarkLogic begins.

As part of the startup process, MarkLogic checks any previous saved environment information that was stored in an Amazon Dynamo database. Should it find that an EBS volume was previously mounted, the volume is automatically mounted again as the default data directory.

Your Forest contents are mounted again and life returns to normal.

MarkLogic

Unit 2 exercise – Terminate Instance

Lab time estimate:
- 15 minutes

- Display MarkLogic Cluster Instances.
- Terminate one of the instances.
  - Note the storage still remains.
- AWS starts a new instance, connects to the storage and recovers.

- **DO**
  - Note what is terminated versus what remains.
- **DON'T**
  - Start another instance if you think nothing is occurring. Be patient and you will be rewarded!

MarkLogic

Questions?