

# MarkLogic in the Cloud Workshop

Unit 3 – Pausing and Resuming Clusters

© 2018 MARKLOGIC CORPORATION

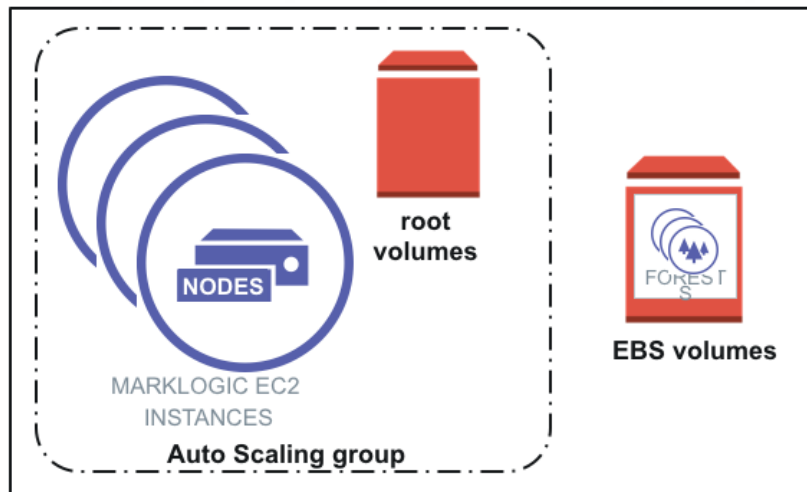
## Goal: Pause and Resume a Cluster

---

- Pausing an AWS MarkLogic cluster
- Resuming the cluster again.

We will examine why and how to pause and resume an AWS MarkLogic cluster.

# Auto Scaling Groups



SLIDE: 3

© 2018 MARKLOGIC CORPORATION

Amazon's Auto Scaling Group launches the desired number of MarkLogic server instances and ensures this number is maintained. Should an instance node in the cluster fail to respond, that instance is terminated and a new one replaces it. Remember that new MarkLogic node instances always re-attach to the EBS volumes and mount MarkLogic's Default Data directory as part of the Managed Cluster feature. New hostnames from AWS are also resolved in the cluster.

The number of instances is a parameter in our CloudFormation templates. As long as the number of nodes is set to 1 or greater, there will be MarkLogic servers to respond to our requests. The Auto Scaling Group responds to our template parameters to always maintain the desired number of MarkLogic nodes per Availability Zone.

The screenshot shows a configuration interface for MarkLogic. A blue callout box points to the 'Nodes per Zone' input field, stating '0 = "pause"'. The 'Nodes per Zone' field is currently set to '1'. Below it, the 'Availability Zone' section shows three selected zones: 'us-west-2a', 'us-west-2b', and 'us-west-2c'. The interface includes descriptive text for each field: 'Number of Zones' (Total number of Availability Zones. 1 or 3.), 'Nodes per Zone' (Total number of nodes per Zone. Set to 0 to shutdown/hibernate), and 'Availability Zone' (The Availability Zones for VPC subnets and instances. Accept either 1 zone or 3 zones. In the order of Subnet 1, Subnet 2 and Subnet 3 (if applicable)).

SLIDE: 4

© 2018 MARKLOGIC CORPORATION

A great benefit of Cloud Computing is that you only pay for what you use. If you are running a development server, or a site that doesn't need to be running 24/7, you can pause your entire cluster so you don't incur EC2 charges while it is not running.

The Auto Scaling Group and Managed Cluster features, along with CloudFormation, enable you to quickly pause your cluster. It is just as simple to resume the cluster and have your resources re-attach to all of your data, so that your cluster will be up and running with all of your data and configuration intact.

Amazon's Auto Scaling uses the "Node per Zone" value to determine how many MarkLogic server instances are in each Availability Zone. A value of "0" essentially pauses the cluster by terminating all of the MarkLogic instance nodes. When the value is reset to the original value, in this case a value of "1" per Availability Zone, new MarkLogic server instances are created.

Remember, your data and configuration are not discarded and will be available when the new MarkLogic server instances mount the existing EBS MarkLogic volumes again.



## Unit 3 exercise – Pause & Resume

- Display the CloudFormation stack.
  - Update the CloudFormation template.
    - Set the number of instances to 0.
  - Check the status of the Cluster instances.
  - Update the CloudFormation template again.
    - Set the number of instances to original value of 1.
- **DO**
    - Refresh to check the CloudFormation stack status.
    - Ensure each MarkLogic EC2 instance is available to Amazon's Elastic Load Balancer.

## Unit 3 QUIZ

1. AWS CloudFormation templates create:

- a) MarkLogic EC2 instances.
- b) Elastic Block Storage (EBS) for MarkLogic data and configuration.
- c) An AWS Elastic Load Balancer that distributes requests to MarkLogic node instances.
- d) All of the above.

1. d

## Unit 3 QUIZ

2. You can pause an EC2 instance but cannot pause an entire MarkLogic cluster.

- a) True.
- b) False.

2. b

## Unit 3 QUIZ

3. Healthy MarkLogic EC2 instances that are ready to be contacted display an Elastic Load Balancer status of:

- a) Ready.
- b) InRange.
- c) InService.

3. c



## Unit 3 QUIZ

4. When viewing a CloudFormation stack, you'd select the main stack to:

- a) Update the MarkLogic CloudFormation stack parameters.
- b) Get the URL for the MarkLogic Administrative Interface on port 8001.
- c) View the resources created such as the Elastic Load Balancer.
- d) All of the above.

4. d

# Questions?