Setting up a VPC MarkLogic Cluster using AWS Cloud Formation

The purpose of this document is to detail how AWS Cloud Formation can be used to set up a MarkLogic Cluster in an AWS[1] VPC[2] environment. It should be read in conjunction with two supplied files – MarkLogic-Sample-VPC-1.json and ThreePlusClusterForVPC-1.json. If reading this via GitHub, these files are in the same directory as this document.

1 Background

MarkLogic includes an in-built managed cluster capability allowing it to be both deployed easily to AWS and to take advantage of the high availability features that AWS offers. For full coverage of MarkLogic's AWS capabilities see the online EC2 guide - http://docs.marklogic.com/guide/ec2.pdf.

The recommended mechanism for deploying MarkLogic is using Cloud Formation Templates (CFTs). These allow deterministic repeatable deployment and also simplify the process of expanding or upgrading the cluster.

A standard CFT for MarkLogic deployment is available[3] but for deployment to an AWS Virtual Private Cloud (VPC) a customized template is required as it is not possibly to craft a generic template that would allow for all customer requirements and configurations.

The purpose of this document and associated assets is to provide a sample CFT for MarkLogic setup in a VPC and to show how to use it. It is hoped that this will be instructive.

2 Sample VPC

In order to demonstrate deployment into a VPC, it is necessary to have a VPC to deploy into. The accompanying template, MarkLogic-Sample-VPC-1.json can be used to create such a VPC.

This template will create

- Externally (to the VPC) accessible sub-nets in three availability zones. Use of three availability zones is required to make full use of MarkLogic's resiliency features
- Three internal sub-nets – hosts placed in these will not be accessible by external resources
- Security Groups
    - rMgmtSecurityGroup - allows ssh and mstsc access from a named IP Range
    - rExternalSecurityGroup – allows http and https access from external addresses
    - rWebSecurityGroup – allows ssh access from hosts in rMgmtSecurityGroup and http/https access from hosts in rExternalSecurityGroup
    - rInternalSecurityGroup – allows ssh access from rMgmtSecurityGroup and full access from rWebSecurityGroup

[1] Amazon Web Services

[2] Virtual Private Cloud

[3] https://s3.amazonaws.com/marklogic-releases/8.0-5.4/ThreePlusCluster-BYOL.template currently,or see http://developer.marklogic.com/products/aws if this URL changes

- A linux 'jump box' placed in rMgmtSecurityGroup allowing **ssh** accesss to rInternalSecurityGroup
- A windows 'jump box' placed in rWebSecurityGroup allowing **web** access to rInternalSecurityGroup.
- Access control lists (ACLs). These are placeholders as they permit all traffic – the security groups control access for us. These are required however. In production these settings would normally be reviewed.
- NAT Gateways[4] allowing hosts in the private sub-net to access the internet

3 Pre-requisites

In order to deploy the cluster two pre-requisites need to be in place. There should be no difficulty with these, but for the record they are

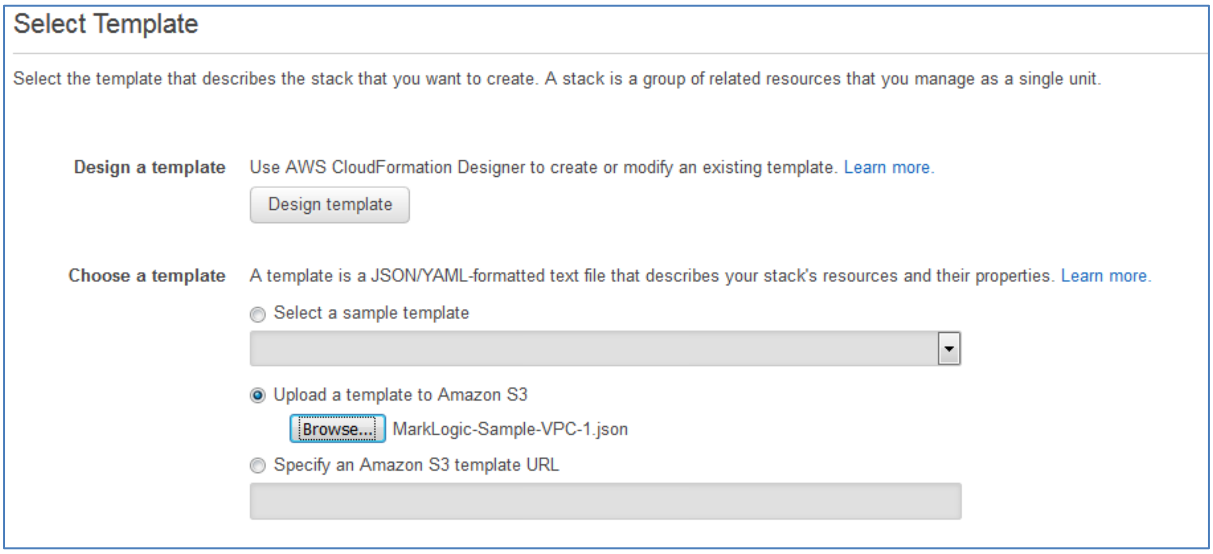
1) An AWS account – see http://docs.marklogic.com/guide/ec2/GettingStarted#id_52961

2) A SSH 'key pair' – see http://docs.marklogic.com/guide/ec2/GettingStarted#id_24571. This is used to secure access to instances.

4 Using the Sample VPC Template

To create the VPC we will use Amazon's management console – http://console.aws.amazon.com, although tools such as the AWS Command Line Interface[5] could also be used.

Access the Cloud Formation console via http://console.aws.amazon.com/cloudformation. From there, choose 'Create Stack'. In the next page, choose 'upload template to S3' then browse to a location where the Sample VPC template has been saved as per the screenshot below. Select the required template then click the 'next' button.



Select Template

Select the template that describes the stack that you want to create. A stack is a group of related resources that you manage as a single unit.

Design a template Use AWS CloudFormation Designer to create or modify an existing template. Learn more.

Choose a template A template is a JSON/YAML-formatted text file that describes your stack's resources and their properties. Learn more.

○ Select a sample template

● Upload a template to Amazon S3

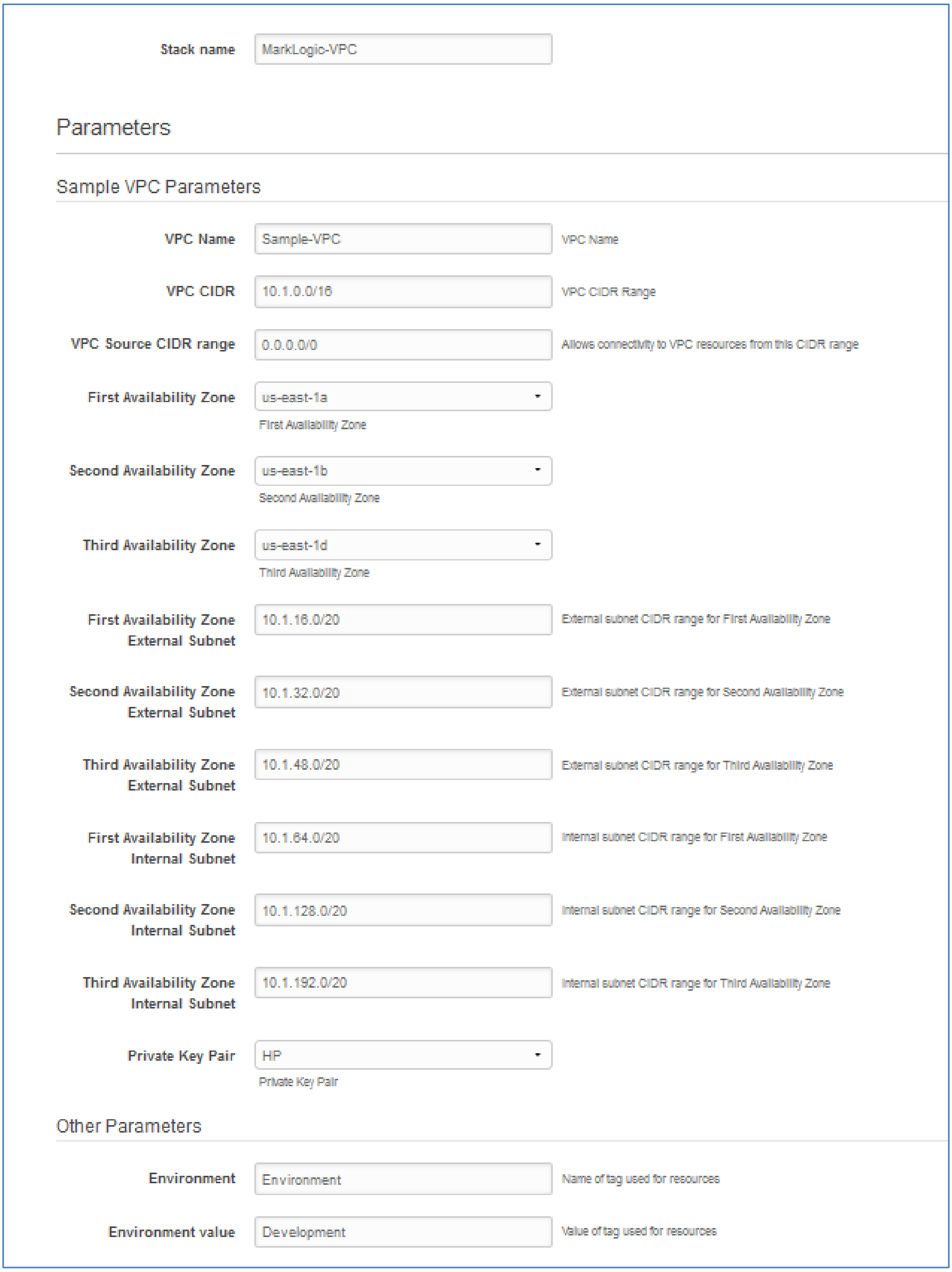
MarkLogic-Sample-VPC-1.json

○ Specify an Amazon S3 template URL

The next screen requires specification of cluster parameters.

[4] http://docs.aws.amazon.com/AmazonVPC/latest/UserGuide/vpc-nat-gateway.html

[5] See http://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/cfn-using-cli.html



Stack name	MarkLogic-VPC	

Parameters

Sample VPC Parameters

VPC Name	Sample-VPC	VPC Name
VPC CIDR	10.1.0.0/16	VPC CIDR Range
VPC Source CIDR range	0.0.0.0/0	Allows connectivity to VPC resources from this CIDR range
First Availability Zone	us-east-1a First Availability Zone	
Second Availability Zone	us-east-1b Second Availability Zone	
Third Availability Zone	us-east-1d Third Availability Zone	
First Availability Zone External Subnet	10.1.16.0/20	External subnet CIDR range for First Availability Zone
Second Availability Zone External Subnet	10.1.32.0/20	External subnet CIDR range for Second Availability Zone
Third Availability Zone External Subnet	10.1.48.0/20	External subnet CIDR range for Third Availability Zone
First Availability Zone Internal Subnet	10.1.64.0/20	Internal subnet CIDR range for First Availability Zone
Second Availability Zone Internal Subnet	10.1.128.0/20	Internal subnet CIDR range for Second Availability Zone
Third Availability Zone Internal Subnet	10.1.192.0/20	Internal subnet CIDR range for Third Availability Zone
Private Key Pair	HP Private Key Pair	

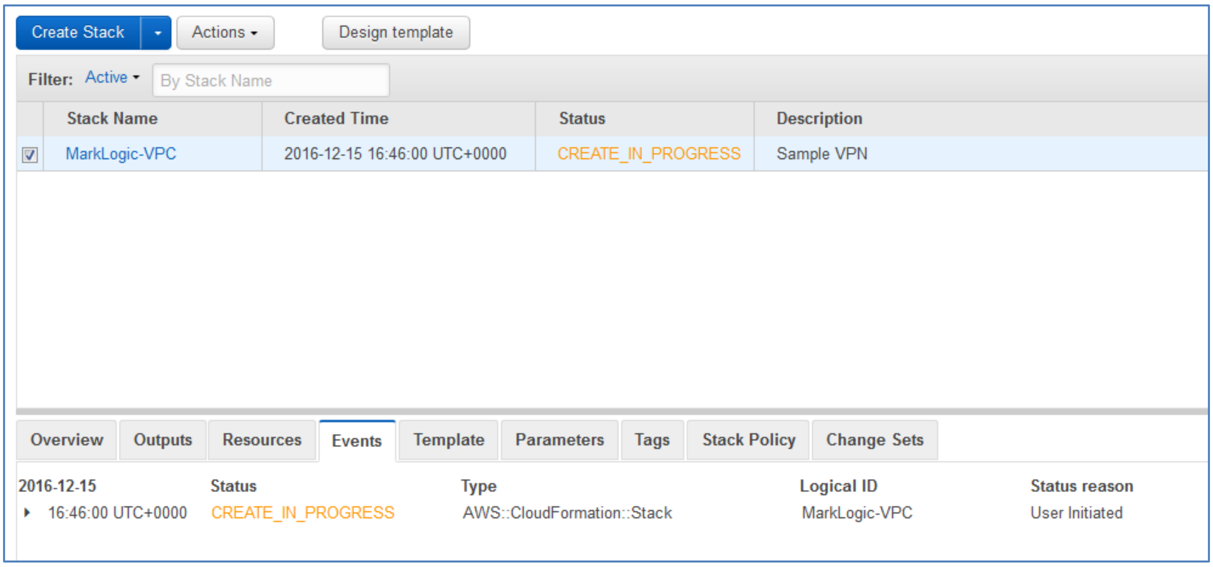
Other Parameters

Environment	Environment	Name of tag used for resources
Environment value	Development	Value of tag used for resources

Full detail of parameters is given below but for instruction purposes the default values may be accepted, with the exception of the private key pair, which will be specific to you.

Parameter Name	**Description**	**Comment**
pVpcName	VPC Name. Default : Sample-VPC	
pVpcCidr	Internal IP address range for VPC	Defaults to 10.1.0.0/16
pVpcSourceCidr	IP address range allowed management access	Can be set to 0.0.0.0/0 for testing purposes, or your own IP /32
pAvailabilityZone1	Availability zone to be used for subnet 1	Default : us-east-1a
pAvailabilityZone2	Availability zone to be used for subnet 2	Default : us-east-1b
pAvailabilityZone3	Availability zone to be used for subnet 3	Default : us-east-1d
pExternalSubnetAZ1Cidr	IP range for external subnet #1	Default : 10.1.16.0/20
pExternalSubnetAZ2Cidr	IP range for external subnet #2	Default : 10.1.32.0/20
pExternalSubnetAZ3Cidr	IP range for external subnet #3	Default : 10.1.48.0/20
pInternalSubnetAZ1Cidr	IP range for internal subnet #1	Default : 10.1.64.0/20
pInternalSubnetAZ2Cidr	IP range for internal subnet #2	Default : 10.1.128.0/20
pInternalSubnetAZ3Cidr	IP range for internal subnet #3	Default : 10.1.192.0/20
pPrivateKeyPair	SSH key pair for SSH access. See section 3	**Required**
pResourceTagName	Name of any additional tag required	Can leave blank
pResourceTagValue	Value of additional tag	Can leave blank

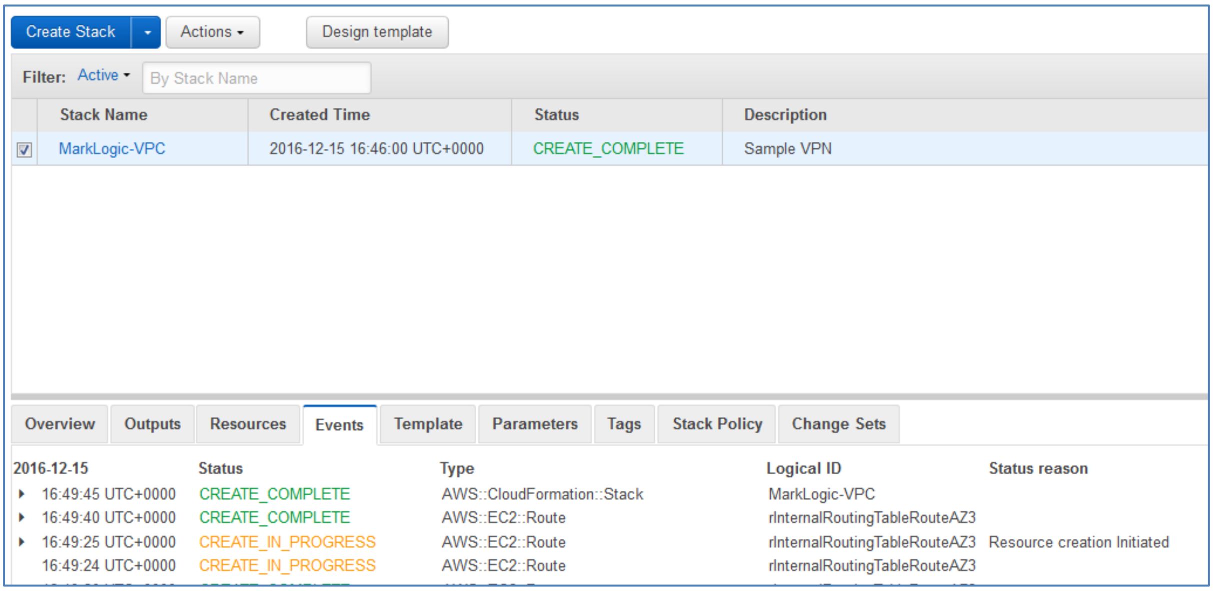
After configuring, click the ‘Next’ button twice followed by ‘Create’. You will see a screen similar to what is below.



	Stack Name	Created Time	Status	Description
☑	MarkLogic-VPC	2016-12-15 16:46:00 UTC+0000	CREATE_IN_PROGRESS	Sample VPN

2016-12-15	Status	Type	Logical ID	Status reason
16:46:00 UTC+0000	CREATE_IN_PROGRESS	AWS::CloudFormation::Stack	MarkLogic-VPC	User Initiated

After around five minutes the stack should finish creating. Your console view will resemble



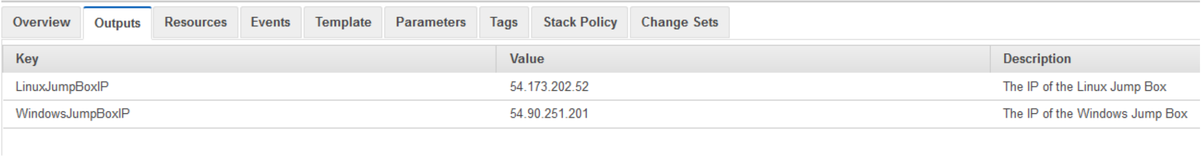
	Stack Name	Created Time	Status	Description
☑	MarkLogic-VPC	2016-12-15 16:46:00 UTC+0000	CREATE_COMPLETE	Sample VPN

2016-12-15	Status	Type	Logical ID	Status reason
16:49:45 UTC+0000	CREATE_COMPLETE	AWS::CloudFormation::Stack	MarkLogic-VPC	
16:49:40 UTC+0000	CREATE_COMPLETE	AWS::EC2::Route	rInternalRoutingTableRouteAZ3	
16:49:25 UTC+0000	CREATE_IN_PROGRESS	AWS::EC2::Route	rInternalRoutingTableRouteAZ3	Resource creation Initiated
16:49:24 UTC+0000	CREATE_IN_PROGRESS	AWS::EC2::Route	rInternalRoutingTableRouteAZ3	

5 Checking your Jump Boxes

We will be positioning the MarkLogic stack in the internal sub-nets, so we need jump boxes[6] to access them. As above, this template sets these up for you.

To check they are available, click the 'Outputs' tab in the 'Create Stack' view above. You should see something like the screenshot below.

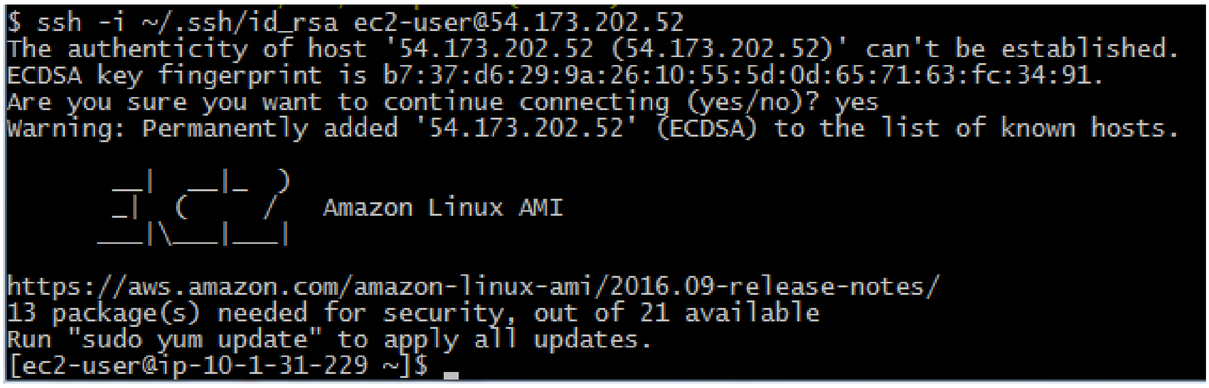


Key	Value	Description
LinuxJumpBoxIP	54.173.202.52	The IP of the Linux Jump Box
WindowsJumpBoxIP	54.90.251.201	The IP of the Windows Jump Box

5.1 Checking Linux Jump Box

Here, my Linux jump box has IP 54.173.202.52. The private half of my ssh key pair is in .ssh/id_rsa in my home directory. I verify my jump box is available by successfully executing

ssh -i ~/.ssh/id_rsa ec2-user@54.173.202.52



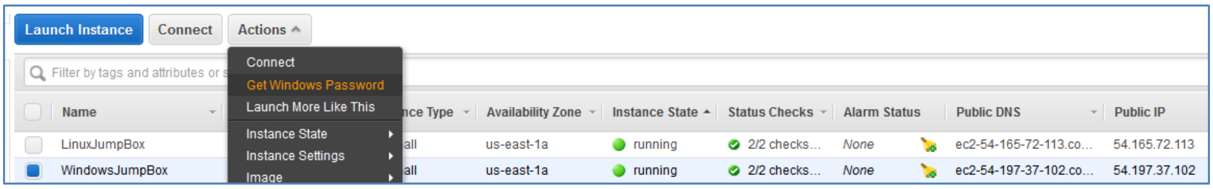
```
$ ssh -i ~/.ssh/id_rsa ec2-user@54.173.202.52
The authenticity of host '54.173.202.52 (54.173.202.52)' can't be established.
ECDSA key fingerprint is b7:37:d6:29:9a:26:10:55:5d:0d:65:71:63:fc:34:91.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '54.173.202.52' (ECDSA) to the list of known hosts.

       __|  __|_  )
       _|  (     /   Amazon Linux AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/
13 package(s) needed for security, out of 21 available
Run "sudo yum update" to apply all updates.
[ec2-user@ip-10-1-31-229 ~]$ _
```

[6] A jump box is an intermediate host allowing indirect access to private resources. Also referred to as a bastion host.

You should similarly verify for your IP/key combination. You should use user ec2-user.

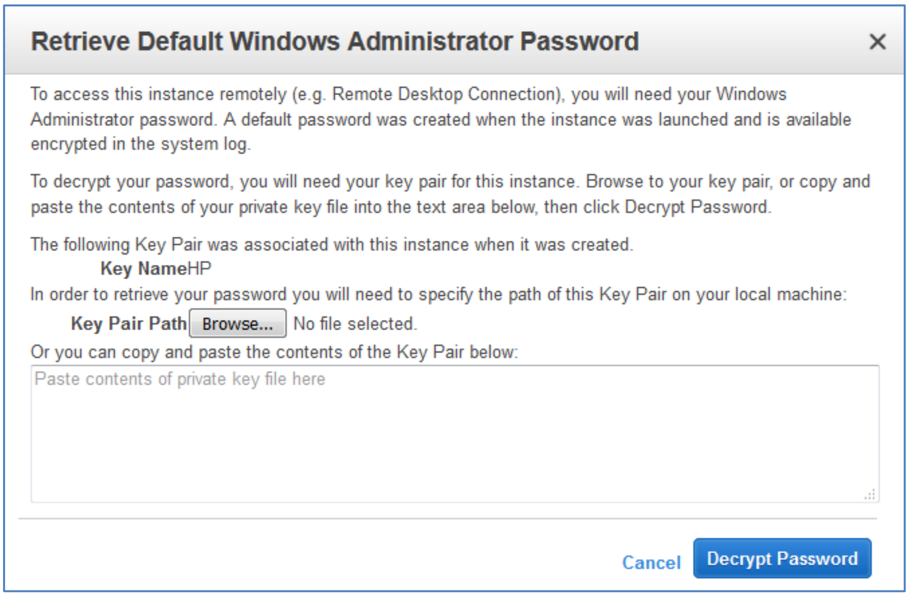
5.2 Checking Windows Jump Box

For the Windows jump box, it's slightly more complicated, but not much.



□	Name	nce Type	Availability Zone	Instance State	Status Checks	Alarm Status	Public DNS	Public IP
□	LinuxJumpBox	all	us-east-1a	● running	✓ 2/2 checks...	None	ec2-54-165-72-113.co...	54.165.72.113
■	WindowsJumpBox	all	us-east-1a	● running	✓ 2/2 checks...	None	ec2-54-197-37-102.co...	54.197.37.102

First, because you can't ssh into a Windows machine, you need to obtain the Windows admin password. Do this from the EC2 console[7]. Select the host with name 'WindowsJumpBox' and then 'Get Windows Password' as shown. You'll see



Retrieve Default Windows Administrator Password ✕

To access this instance remotely (e.g. Remote Desktop Connection), you will need your Windows Administrator password. A default password was created when the instance was launched and is available encrypted in the system log.

To decrypt your password, you will need your key pair for this instance. Browse to your key pair, or copy and paste the contents of your private key file into the text area below, then click Decrypt Password.

The following Key Pair was associated with this instance when it was created.

Key NameHP

In order to retrieve your password you will need to specify the path of this Key Pair on your local machine:

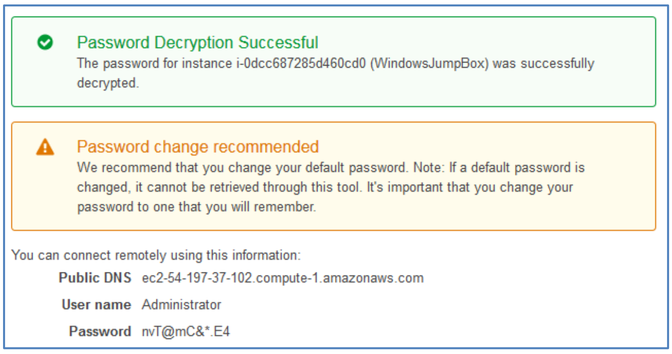
Key Pair Path Browse... No file selected.

Or you can copy and paste the contents of the Key Pair below:

Paste contents of private key file here

Cancel Decrypt Password

Click browse, and then select the private part of your ssh key pair[8], then 'Decrypt Password'. You'll see a screen similar to



Password Decryption Successful

The password for instance i-0dcc687285d460cd0 (WindowsJumpBox) was successfully decrypted.

Password change recommended

We recommend that you change your default password. Note: If a default password is changed, it cannot be retrieved through this tool. It's important that you change your password to one that you will remember.

You can connect remotely using this information:

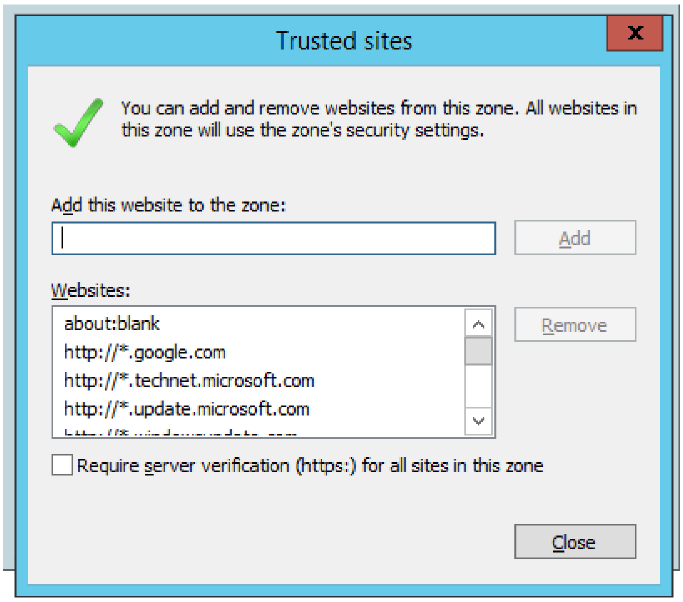
Public DNS ec2-54-197-37-102.compute-1.amazonaws.com
User name Administrator
Password nvT@mC&*.E4

[7] From console.aws.amazon.com select Services, then click EC2
[8] As per section 5.1

Use an RDP client (e.g. mstsc) to log into the above box using account ‘Administrator’ and the password given. You should be able to log in successfully. While you’re there, follow the recommendation above and change your password. It’s CTRL-ALT-END under Windows to do this.

To avoid difficulties with Internet Explorer, which will require explicit exemptions for any websites accessed, including the MarkLogic Elastic Load Balancer, installing Chrome[9] is recommended. Open Internet Explorer and access www.google.com. You’ll be told it’s untrusted. Add http://*.google.com and https://*.google.com as trusted sites then navigate to

https://www.google.com/chrome/browser/desktop/index.html



You can add and remove websites from this zone. All websites in this zone will use the zone's security settings.

Add this website to the zone:

Websites:

about:blank
http://*.google.com
http://*.technet.microsoft.com
http://*.update.microsoft.com

□ Require server verification (https:) for all sites in this zone

Download, install and make Chrome your default browser.

6 Setting up MarkLogic in the VPC

In this section we will use the ThreePlusClusterForVPC-1.json template to set up a MarkLogic cluster.

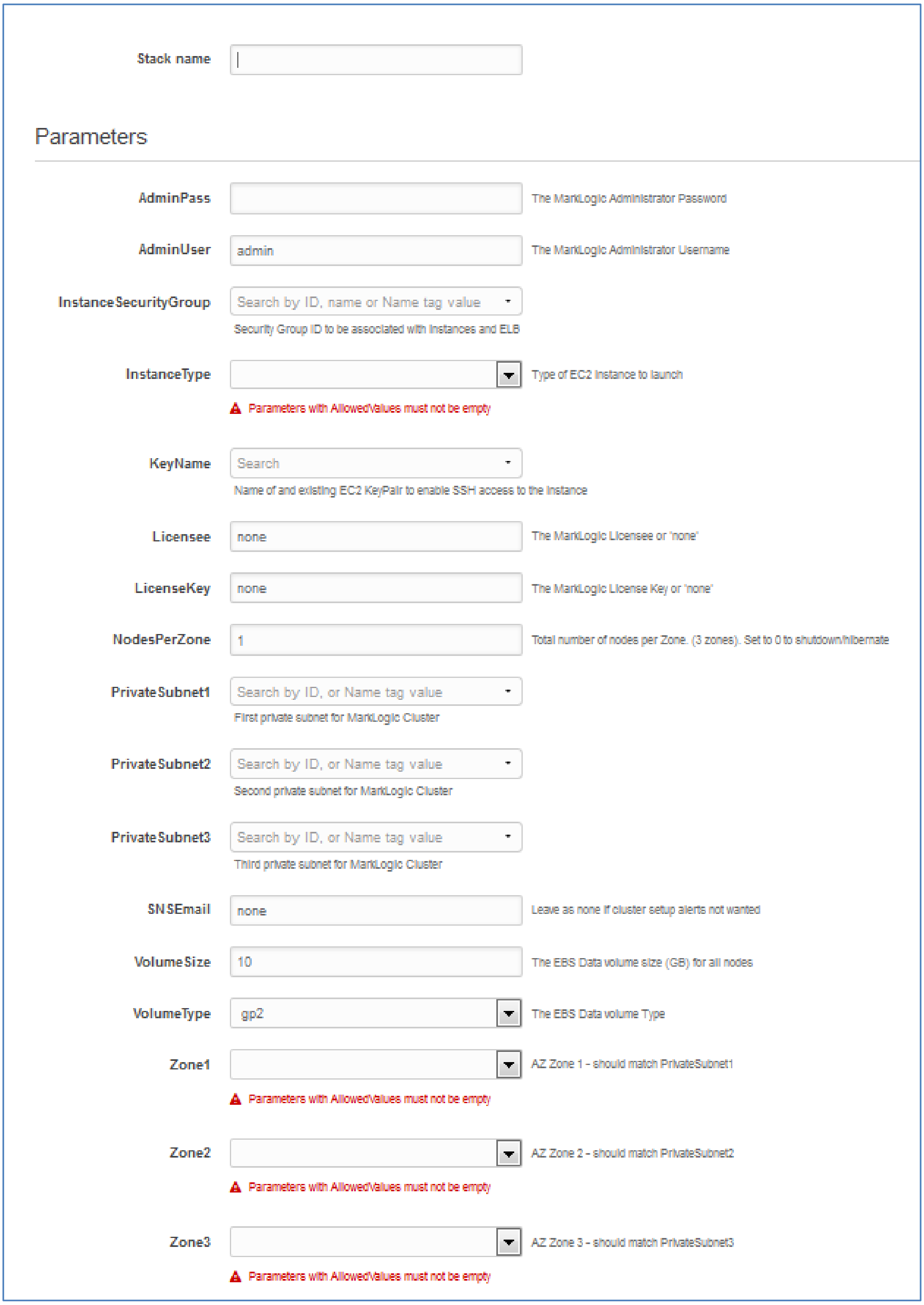
There is not a single fixed way in which you might do this. Here we deploy MarkLogic hosts into our private subnets. We will be assigning our hosts to the rInternalSecurityGroup created above, which means in turn that they can be seen by hosts in the rWebSecurityGroup. A typical configuration might be to place the middle tier (e.g. node.js/tomcat/apache) in rWebSecurityGroup and then add in a load balancer for these assets in the rExternalSecurityGroup, which would be visible externally. Later we see how we can demonstrate that the arrangement works using our Windows jump box.

6.1 Template Use

As previously, access the Cloud Formation console via http://console.aws.amazon.com/cloudformation. From there, choose ‘Create Stack’. In the next

[9] Other browsers are available.

page, choose ‘upload template to S3’ then browse to a location where the MarkLogic cluster template ThreePlusClusterForVPC-1.json has been saved. Select the required template then click the ‘next’ button. You will see something like the screen below.



Stack name	|	

Parameters

AdminPass		The MarkLogic Administrator Password
AdminUser	admin	The MarkLogic Administrator Username
InstanceSecurityGroup	Search by ID, name or Name tag value ▾ Security Group ID to be associated with Instances and ELB	
InstanceType	▾ ⚠ Parameters with AllowedValues must not be empty	Type of EC2 Instance to launch
KeyName	Search ▾ Name of and existing EC2 KeyPair to enable SSH access to the Instance	
Licensee	none	The MarkLogic Licensee or 'none'
LicenseKey	none	The MarkLogic License Key or 'none'
NodesPerZone	1	Total number of nodes per Zone. (3 zones). Set to 0 to shutdown/hibernate
PrivateSubnet1	Search by ID, or Name tag value ▾ First private subnet for MarkLogic Cluster	
PrivateSubnet2	Search by ID, or Name tag value ▾ Second private subnet for MarkLogic Cluster	
PrivateSubnet3	Search by ID, or Name tag value ▾ Third private subnet for MarkLogic Cluster	
SNSEmail	none	Leave as none if cluster setup alerts not wanted
VolumeSize	10	The EBS Data volume size (GB) for all nodes
VolumeType	gp2 ▾	The EBS Data volume Type
Zone1	▾ ⚠ Parameters with AllowedValues must not be empty	AZ Zone 1 - should match PrivateSubnet1
Zone2	▾ ⚠ Parameters with AllowedValues must not be empty	AZ Zone 2 - should match PrivateSubnet2
Zone3	▾ ⚠ Parameters with AllowedValues must not be empty	AZ Zone 3 - should match PrivateSubnet3

Fields are required as follows

Parameter Name	**Description**	**Value if using MarkLogic VPC**
Stack Name	Name of stack – free text field	
AdminPass	Configure admin user password	
AdminUser	Configure admin user name	
InstanceSecurityGroup	Security group for ML instances/ELB	rInternalSecurityGroup
InstanceType	EC2 instance type	m3.medium for instruction[10]
KeyName	SSH Key pair for ssh access. See section 3	
Licensee	MarkLogic Licensee	
LiceneKey	MarkLogic License Key	
NodesPerZone	Nodes per AZ zone	1 for instruction
PrivateSubnet1	Internal Subnet for AZ #1	rInternalSubnetAZ1
PrivateSubnet2	Internal Subnet for AZ #2	rInternalSubnetAZ2
PrivateSubnet3	Internal Subnet for AZ #3	rInternalSubnetAZ3
SNSEmail	Email address to send diagnostic setup messages to.	Your email address or none if not required.
VolumeSize	Size of EBS volume in Gb	10 for instruction
VolumeType	EBS volume type	gp2
Zone1	AZ corresponding to PrivateSubnet1[11]	us-east-1a
Zone2	AZ corresponding to PrivateSubnet2	us-east-1b
Zone3	AZ corresponding to PrivateSubnet3	us-east-1d

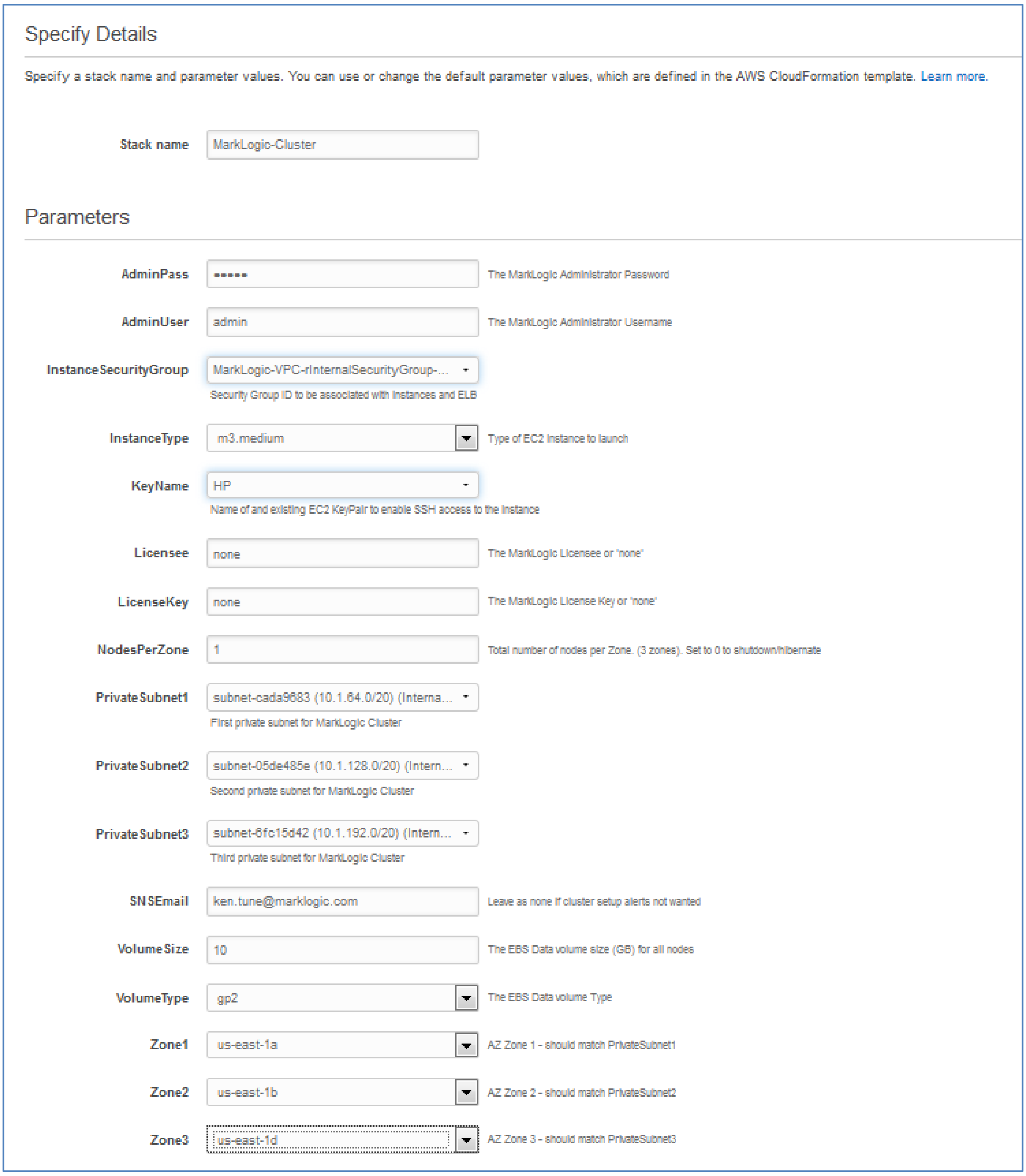
The SNSEmail parameter, if set, will be assigned to an SNS Topic[12] resulting in diagnostic cluster setup messages reaching the specified address. There will be a large number of these sent, but it is useful to enable this when setting up for the first time, or in case of failure to setup correctly. Leave as 'none' if you do not wish to receive these messages.

A populated template is shown below.

[10] Currently ~$0.05 per hour

[11] Note it doesn't seem to be possible to infer this from PrivateSubnet1 – at least I can't make it happen. Fn::GetAtt operates on resources only, not parameters.

[12] See https://aws.amazon.com/sns/



Specify Details

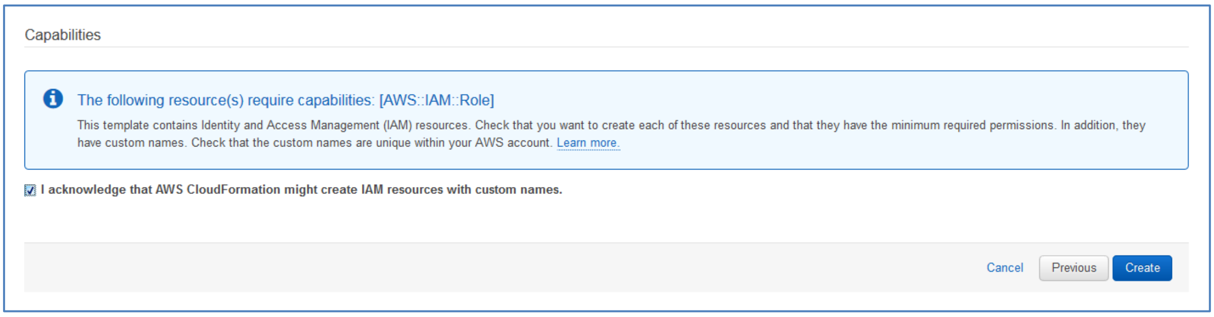
Specify a stack name and parameter values. You can use or change the default parameter values, which are defined in the AWS CloudFormation template. Learn more.

Stack name: MarkLogic-Cluster

Parameters

AdminPass	•••••	The MarkLogic Administrator Password
AdminUser	admin	The MarkLogic Administrator Username
InstanceSecurityGroup	MarkLogic-VPC-rInternalSecurityGroup-... Security Group ID to be associated with instances and ELB	
InstanceType	m3.medium	Type of EC2 instance to launch
KeyName	HP Name of and existing EC2 KeyPair to enable SSH access to the instance	
Licensee	none	The MarkLogic Licensee or 'none'
LicenseKey	none	The MarkLogic License Key or 'none'
NodesPerZone	1	Total number of nodes per Zone. (3 zones). Set to 0 to shutdown/hibernate
PrivateSubnet1	subnet-cada9683 (10.1.64.0/20) (Interna... First private subnet for MarkLogic Cluster	
PrivateSubnet2	subnet-05de485e (10.1.128.0/20) (Intern... Second private subnet for MarkLogic Cluster	
PrivateSubnet3	subnet-6fc15d42 (10.1.192.0/20) (Intern... Third private subnet for MarkLogic Cluster	
SNSEmail	ken.tune@marklogic.com	Leave as none if cluster setup alerts not wanted
VolumeSize	10	The EBS Data volume size (GB) for all nodes
VolumeType	gp2	The EBS Data volume Type
Zone1	us-east-1a	AZ Zone 1 - should match PrivateSubnet1
Zone2	us-east-1b	AZ Zone 2 - should match PrivateSubnet2
Zone3	us-east-1d	AZ Zone 3 - should match PrivateSubnet3

After populating the template, click ‘Next’ twice. At the bottom of the screen you will see



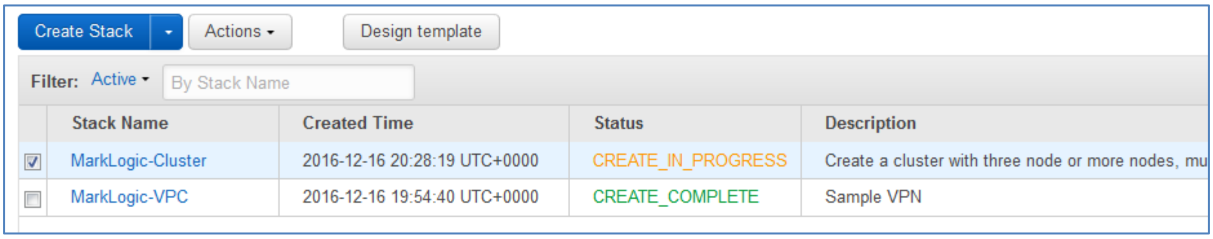
Capabilities

The following resource(s) require capabilities: [AWS::IAM::Role]

This template contains Identity and Access Management (IAM) resources. Check that you want to create each of these resources and that they have the minimum required permissions. In addition, they have custom names. Check that the custom names are unique within your AWS account. Learn more.

☑ I acknowledge that AWS CloudFormation might create IAM resources with custom names.

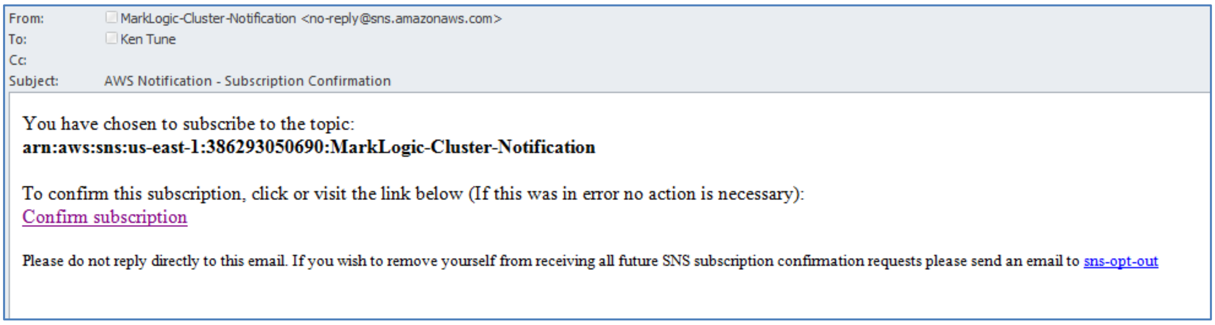
as the template creates an IAM role that governs permissions of the cluster nodes and this requires specific approval. Check the ‘I acknowledge’ box and click ‘Create’. You will see



Filter: Active By Stack Name

	Stack Name	Created Time	Status	Description
☑	MarkLogic-Cluster	2016-12-16 20:28:19 UTC+0000	CREATE_IN_PROGRESS	Create a cluster with three node or more nodes, mu
□	MarkLogic-VPC	2016-12-16 19:54:40 UTC+0000	CREATE_COMPLETE	Sample VPN

If you populated the SNSEmail field, shortly after, the specified address will receive a mail similar to



From: MarkLogic-Cluster-Notification <no-reply@sns.amazonaws.com>
To: Ken Tune
Cc:
Subject: AWS Notification - Subscription Confirmation

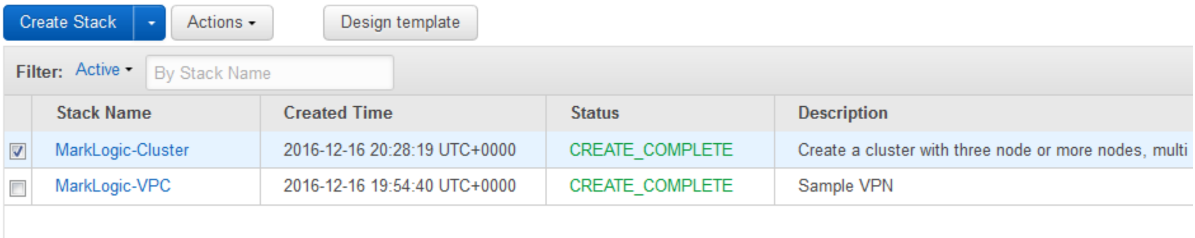
You have chosen to subscribe to the topic:
arn:aws:sns:us-east-1:386293050690:MarkLogic-Cluster-Notification

To confirm this subscription, click or visit the link below (If this was in error no action is necessary):
Confirm subscription

Please do not reply directly to this email. If you wish to remove yourself from receiving all future SNS subscription confirmation requests please send an email to sns-opt-out

Click ‘Confirm Subscription’. If you do not, you will not receive diagnostic messages.

After around five minutes your template should complete and your Cloud Formation view should resemble



Filter: Active By Stack Name

	Stack Name	Created Time	Status	Description
☑	MarkLogic-Cluster	2016-12-16 20:28:19 UTC+0000	CREATE_COMPLETE	Create a cluster with three node or more nodes, multi
□	MarkLogic-VPC	2016-12-16 19:54:40 UTC+0000	CREATE_COMPLETE	Sample VPN

After the template completes setup of the cluster starts. The included Auto Scaling Groups[13] will launch instances to match the ’Hosts per Zone’ parameter in the template and a number of

[13] https://aws.amazon.com/autoscaling/

operations will take place on each node in order to ready the cluster. If you provided an email address for the SNSEmail field, you will start to see diagnostic messages concerning the automated setup.

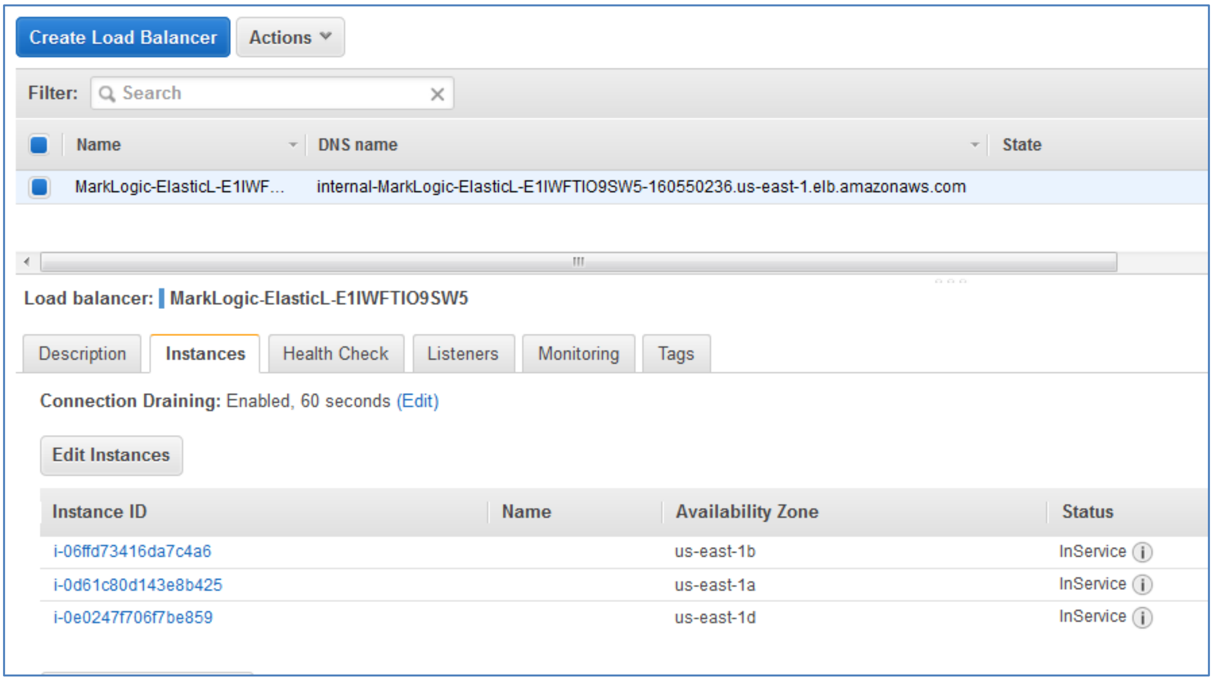
It will take 5 – 10 minutes from the point at which the cloud formation template completes before a fully available cluster is operational, or 10-15 minutes from the point at which the cloud formation stack creation was started.

7 Verifying your Cluster

To check setup has completed successfully, we verify two things.

7.1 All nodes in cluster available and connected to ELB[14]

The first test of whether the cluster has set up correctly is whether the full cluster is available via the included Elastic Load Balancer. Go to the EC2 console - https://console.aws.amazon.com/ec2 and select 'Load Balancers' from the left hand list. Select your load balancer followed by the 'Instances' tab. All three of our instances should be visible with status 'In Service', as below.



Name	DNS name	State
MarkLogic-ElasticL-E1IWF...	internal-MarkLogic-ElasticL-E1IWFTIO9SW5-160550236.us-east-1.elb.amazonaws.com	

Load balancer: MarkLogic-ElasticL-E1IWFTIO9SW5

Connection Draining: Enabled, 60 seconds (Edit)

Instance ID	Name	Availability Zone	Status
i-06ffd73416da7c4a6		us-east-1b	InService
i-0d61c80d143e8b425		us-east-1a	InService
i-0e0247f706f7be859		us-east-1d	InService

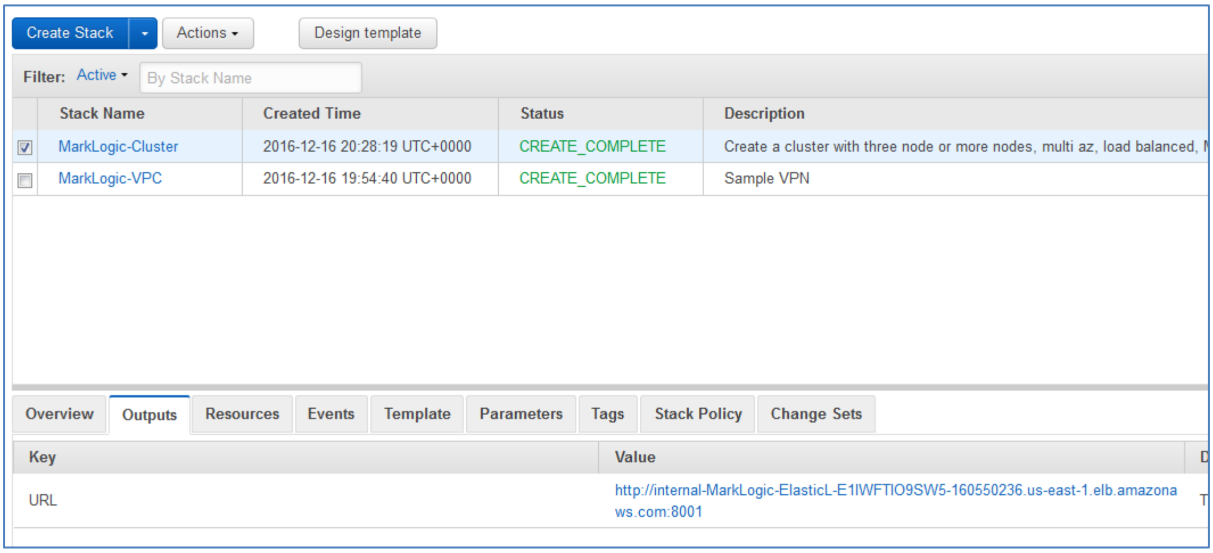
If any are out of service, or absent then cluster setup has not completed.

7.2 Cluster is reachable via the ELB from the Windows Jump Box

The cluster should be reachable via the ELB from the Windows Jump box. If this is the case, then it will be reachable by any host in the rWebSecurityGroup, which as per section 6 would typically be the security group used for the middle tier.

[14] Elastic Load Balancer

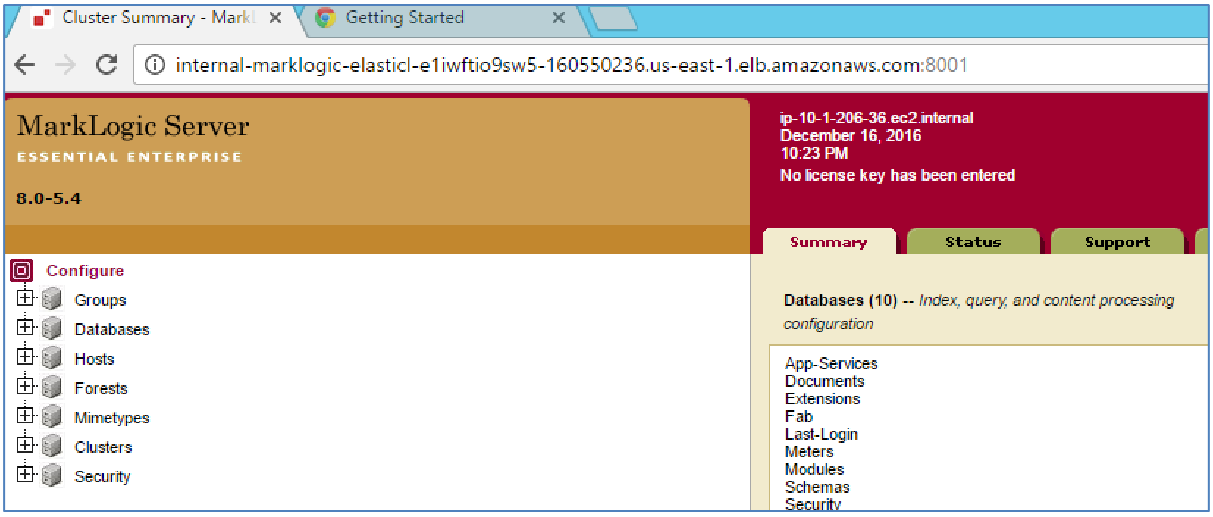
To verify this, first identify the ELB DNS name by going back to the Cloud Formation screen - https://console.aws.amazon.com/cloudformation, select the 'MarkLogic Cluster' stack and select the 'Outputs' tag.



	Stack Name	Created Time	Status	Description
☑	MarkLogic-Cluster	2016-12-16 20:28:19 UTC+0000	CREATE_COMPLETE	Create a cluster with three node or more nodes, multi az, load balanced, M
□	MarkLogic-VPC	2016-12-16 19:54:40 UTC+0000	CREATE_COMPLETE	Sample VPN

Key	Value	D
URL	http://internal-MarkLogic-ElasticL-E1IWFTIO9SW5-160550236.us-east-1.elb.amazonaws.com:8001	T

Copy and paste your ELB URL somewhere convenient. Now log into the Windows jump box as per section 5.2, open Chrome and paste your ELB URL into the address bar. You should see the admin console as below. Note that you're accessing it via your ELB URL.



internal-marklogic-elasticl-e1iwftio9sw5-160550236.us-east-1.elb.amazonaws.com:8001

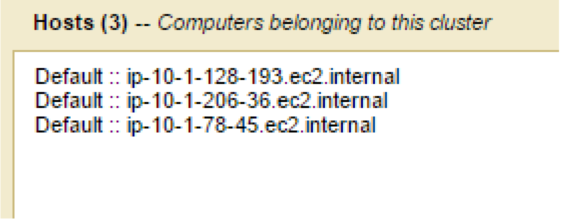
MarkLogic Server
ESSENTIAL ENTERPRISE
8.0-5.4

ip-10-1-206-36.ec2.internal
December 16, 2016
10:23 PM
No license key has been entered

Databases (10) -- *Index, query, and content processing configuration*

App-Services
Documents
Extensions
Fab
Last-Login
Meters
Modules
Schemas
Security

You should see your three connected hosts bottom right of this panel.

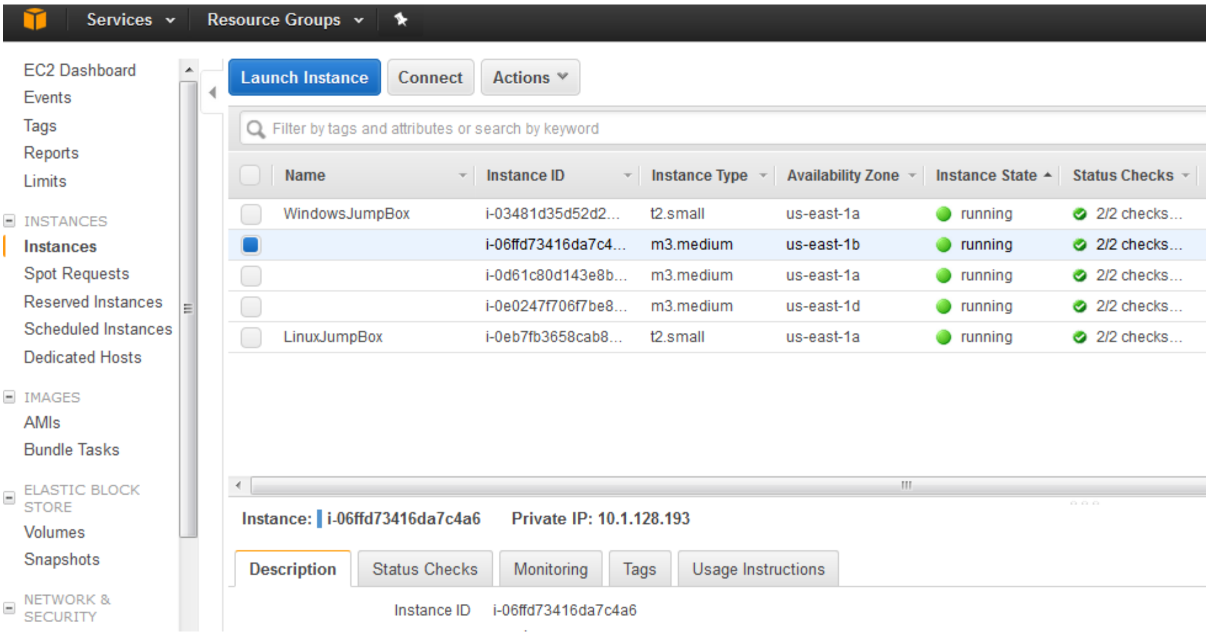


Hosts (3) -- *Computers belonging to this cluster*

Default :: ip-10-1-128-193.ec2.internal
Default :: ip-10-1-206-36.ec2.internal
Default :: ip-10-1-78-45.ec2.internal

7.3 Accessing hosts via ssh via Linux Jump Box

It’s also useful to verify that you can access your MarkLogic hosts at the command line. First obtain the IP of the hosts by visiting the EC2 console, https://console.aws.amazon.com/ec2, and clicking ‘Instances’. The IP addresses of the hosts may be found by selecting the hosts as shown. The first one below has IP 10.1.128.193.



	Name	Instance ID	Instance Type	Availability Zone	Instance State	Status Checks
□	WindowsJumpBox	i-03481d35d52d2...	t2.small	us-east-1a	● running	✓ 2/2 checks...
■		i-06ffd73416da7c4...	m3.medium	us-east-1b	● running	✓ 2/2 checks...
□		i-0d61c80d143e8b...	m3.medium	us-east-1a	● running	✓ 2/2 checks...
□		i-0e0247f706f7be8...	m3.medium	us-east-1d	● running	✓ 2/2 checks...
□	LinuxJumpBox	i-0eb7fb3658cab8...	t2.small	us-east-1a	● running	✓ 2/2 checks...

Instance: i-06ffd73416da7c4a6 Private IP: 10.1.128.193

Instance ID i-06ffd73416da7c4a6

To access this host, log into your Linux jump box as per section 5.1. Then ssh into the internal instance e.g. ssh 10.1.128.193. You should log in successfully, and also be able to do this for the other two boxes.

8 Troubleshooting

VPCs can be difficult to troubleshoot precisely because they hide their contents from the outside world. Any difficulties encountered may well be associated with restrictions due to security groups, access control lists, or lack of public IPs. Check the implications of these carefully if resources are inaccessible.

If your MarkLogic cluster does not start successfully, the first course of action is to log into the MarkLogic hosts, if you can, and look at /var/log/messages[15] for clues. Also look at the ‘Events’ tab associated with the Cloud Formation Stacks in the case of errors. The SNS messages are useful - look for ‘fail’. You will see that they are equivalent to the messages found in /var/log/messages.

9 Conclusion

If you’ve got this far, then you’ve successfully built your managed MarkLogic Cluster inside a Virtual Private Cloud.

[15] Sudo required

Note that there is more than one way to do this, and your own requirements may vary. Furthermore, for production use, you should review the detail thoroughly, especially the security aspects, to ensure configuration meets your requirements. Where applicable you should make changes – these examples are intended to help you get started rather than being rigid solutions.

Nevertheless, it is hoped that this document and the associated templates have been instructive.