Creating a Batch Processing Cluster

2015

Yu-Tsuen Lin

Learning object

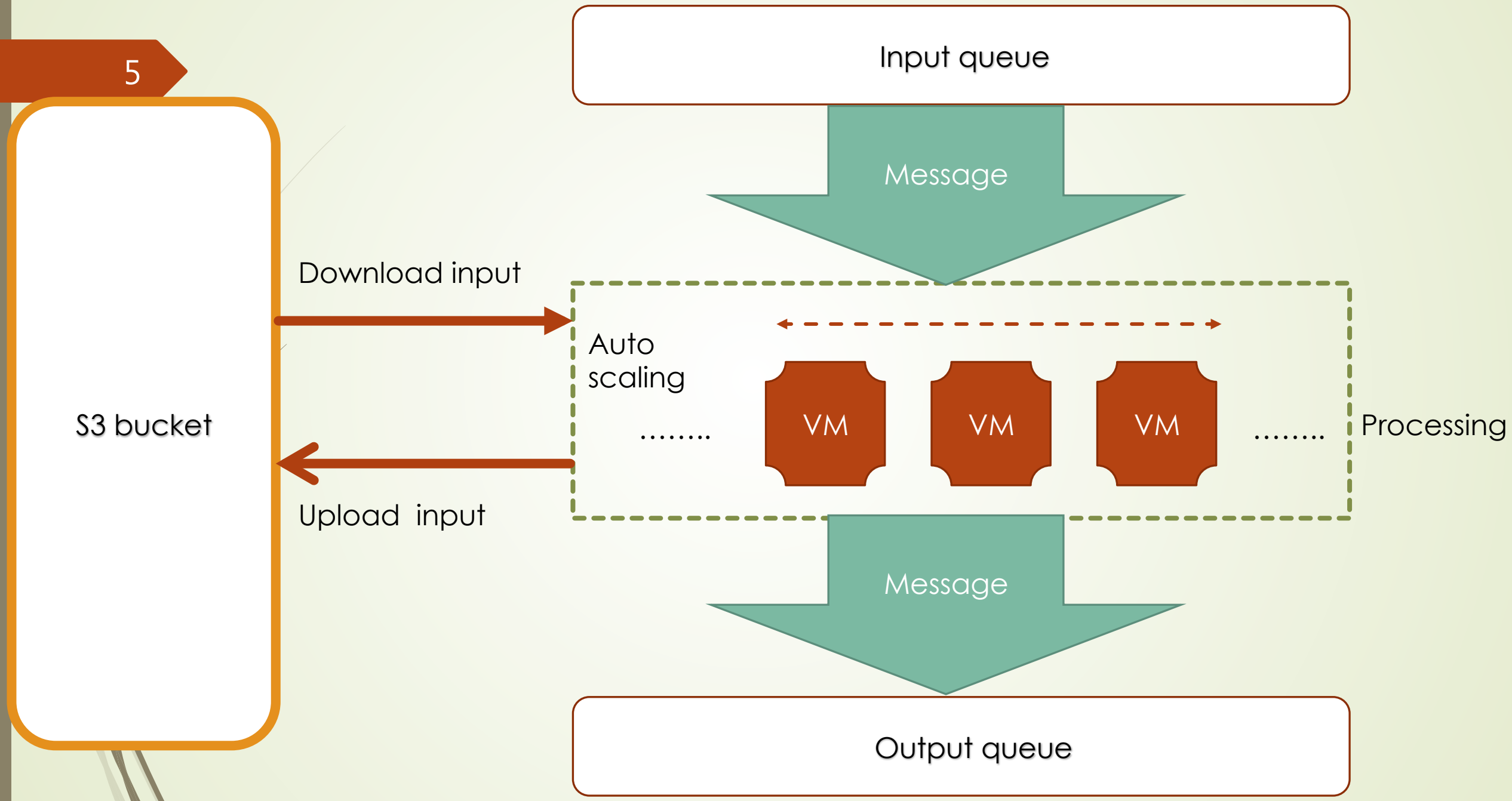
1. Use the AWS Management Console to create an SQS queue(6-7)
2. Bootstrap an EC2 instance using User Data (9-13)
3. Create an AMI from a running instance (17)
4. Create an Auto Scaling Group with Scaling Policies based on an SQS queue (18-29)

What service to use ?

- EC2
- AMIs
- ELB
- Security Group
- Auto Scaling
- Cloud Watch

Amazon EC2 Key Pairs

- Use your .pem file


Input queue
Message
Download input
S3 bucket
Auto
scaling
.....
VM
VM
VM
.....
Processing
Upload input
Message
Output queue

Creating two SQS task queues

- In the AWS Management Console, select SQS from the Services menu.
- Click Create New Queue, then
  a. Queue Name: input
  b. Default Visibility Timeout: 90 seconds
  c. Click Create Queue
- Repeat the previous step to create another queue named: output

Creating two SQS task queues (continued…)

- Select your input queue.
- From the Queue Actions menu, select Send a Message.

==========

https://us-east-1-aws-training.s3.amazonaws.com/arch-static-assets/static/20120728-DSC01265-L.jpg

https://us-east-1-aws-training.s3.amazonaws.com/arch-static-assets/static/20120728-DSC01267-L.jpg

https://us-east-1-aws-training.s3.amazonaws.com/arch-static-assets/static/20120728-DSC01292-L.jpg

https://us-east-1-aws-training.s3.amazonaws.com/arch-static-assets/static/20120728-DSC01315-L.jpg

https://us-east-1-aws-training.s3.amazonaws.com/arch-static-assets/static/20120728-DSC01337-L.jpg

==========

- Click Send Message then click Close.

Create an S3 bucket

- In the AWS Management Console , Select S3 form the Services menu.
- Click Create Bucket, Then:

  a. Bucket Name: lab2-account

  b. Click Create

Creating a 'Master' EC2 Instance

- From the AWS Management Console, select EC2 from the Services menu.
- Click Launch Instance.
- Locate the Amazon Linux AMI and click Select.
- At the Choose an Instance Type panel, click Next: Configure Instance Details.

Creating a 'Master' EC2 Instance (continued…)

- At the Configure Instance Details panel:

  a. IAM Role: BatchProcessing

  b. Expand the Advanced Details section.

  c. User Data: Paste the text in next slide.

  d. Click Next: Add Storage.

Creating a 'Master' EC2 Instance (continued…)

```
#!/bin/bash
# Install ImageMagick, a Python library, and create a directory
yum install -y ImageMagick
easy_install argparse
mkdir /home/ec2-user/jobs
# Download and install the batch processing script
# The following command must be on a single line:
wget -O /home/ec2-user/image_processor.py https://us-west-2-aws-training.s3.amazonaws.com/awsu-ilt/architecting/lab-3-creating-a-batch-processing-cluster-3.2/static/image_processor.py
```

Creating a 'Master' EC2 Instance (continued…)

- There are no modifications needed in the Add Storage panel. Click Next: Tag Instance.
- At the Tag Instance panel, enter the Value: Master-Username
- Click Next: Configure Security Group.
- At the Configure Security Group panel:

  a. Security group name: BatchProcessing

  b. Description: Batch Processing

  c. Verify there is an existing rule for port 22 (SSH).

Creating a 'Master' EC2 Instance (continued…)

- Click Review and Launch.
- Click Launch. You are presented with the Select an existing key pair or create a new key pair dialog.
- Choose your key pair
- Check the acknowledgement box and click Launch Instances.
- Click View Instances (you might need to scroll down to see it).

Connecting to the EC2 instance (Windows)

- Download PuTTY and PuTTYGEN from:

http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html

- Run PuTTYGEN for .pem to translate to .ppk.

  a. Run PuTTYgen

  b. choose SSH-2 RSA in Type of key to generate panel

  c. click Load and choose All Files(*.*),find your.pem

  d. click Save private key and click yes

- Copy the Public DNS name of your EC2 instance from the EC2 Management Console and paste it into the Host Name field in PuTTY:

Connecting to the EC2 instance (Windows)

- Expand the SSH setting in the left panel and click Auth.
- Click the Browse button and locate the PPK file that you downloaded at the start of this lab.
- Click the Open button to initiate the connection. Accept any warning messages that appear.
- Login as: ec2-user

Change setting in python

```
def main(argv=None):
        # Handle command-line arguments for AWS credentials and resource names
        parser = argparse.ArgumentParser(description='Process AWS resources and credentials.')
        parser.add_argument('--input-queue', action='store', dest='input_queue', required=False, default="input-testaaa", help='SQS queue from which input jobs are retrieved')
        parser.add_argument('--output-queue', action='store', dest='output_queue', required=False, default="output-testaaa", help='SQS queue to which job results are placed')
        parser.add_argument('--s3-output-bucket', action='store', dest='s3_output_bucket', required=False, default="", help='S3 bucket where list of instances will be stored')
        parser.add_argument('--region', action='store', dest='region', required=False, default="", help='Region that the SQS queus are in')
        args = parser.parse_args()

        # Get region
        region_name = args.region

##############################################################################
# Verify S3 bucket, create it if required
##############################################################################
def create_s3_output_bucket(s3_output_bucket, s3_endpoint, region_name):

        # Connect to S3
        s3 = boto.connect_s3(host=s3_endpoint)

        # Find any existing buckets starting with 'image-bucket'
        buckets = [bucket.name for bucket in s3.get_all_buckets() if bucket.name.startswith('youraccount')]
        if len(buckets) > 0:
          return buckets[0]

        # No buckets, so create one for them
        name = 'youraccount' + str(uuid.uuid4())
        s3.create_bucket(name, location=region_name)
        return name

##############################################################################
```

Creating an AMI from your batch processing instance

- Select your instance in the EC2 Management Console.
- From the Actions menu, select Create Image, then:
  a. Image Name: Worker Image
  b. Image Description: Batch Processing worker
  c. Click Create Image, then click Close
- Click AMIs in the left panel to view your AMI.

Creating an auto scaling launch configuration

- In the AWS Management Console, select EC2 from the Services menu.
- Click Launch Configurations in the left panel (you may need to scroll down to see it).
- Click Create Auto Scaling group.
- Click Create launch configuration.
- Click My AMIs in the left panel.
- Select your Worker Image AMI:
- On the Choose Instance Type panel, click Next: Configure details.

Creating an auto scaling launch configuration(continued…)

- On the Configure details panel:
  a. Name: Workers-username
  b. IAM Role: BatchProcessing
  c. Expand the Advanced Details section.
  d. User Data: Paste following in the text.

```
#!/bin/sh
/usr/bin/python /home/ec2-user/image_processor.py &
```

  e. Click Next: Add Storage.

AWS | Services | Edit

1. Choose AMI | 2. Choose Instance Type | 3. Configure details | 4. Add Storage | 5. Configure Security Group | 6. Review

Create Launch Configuration

Name workers-testaaa

Purchasing option Request Spot Instances

IAM role BatchProcessing

Monitoring Enable CloudWatch detailed monitoring
Learn more

Advanced Details

Kernel ID Use default

RAM Disk ID Use default

User data As text | As file | Input is already base64 encoded

```
#!/bin/sh
/usr/bin/python /home/ec2-user/image_processor.py &
```

IP Address Type
- Only assign a public IP address to instances launched in the default VPC and subnet. (default)
- Assign a public IP address to every instance.
- Do not assign a public IP address to any instances.

Note: this option only affects instances launched into an Amazon VPC

Later, if you want to use a different launch configuration, you can create a new one and apply it to any Auto Scaling group. Existing launch configurations cannot be edited.

Creating an auto scaling launch configuration(continued…)

- There are no modifications needed in the Add Storage panel. Click Next: Configure Security Group.
- At the Configure Security Group panel:
  a. Click Select an existing security group
  b. Select the BatchProcessing security group you created earlier
  c. Click Review (in the bottom-right)
- Click Create launch configuration. You are presented with the Select an existing key pair or create a new key pair dialog.
- Check the acknowledgement box and click Create launch configuration.

Creating an auto scaling group

- In the Configure Auto Scaling group details panel:

  a. Group Name: worker-group

  b. Subnet: Select at least one subnet by clicking in the Subnet box

  c. Click Next: Configure scaling policies.
- On the Configure scaling policies panel

  a. Click Use scaling policies to adjust the capacity of this group

  b. Scale between: 1 and 3 instances

  c. Under Increase Group Size: Take the Action: Add 1 instances

  d. Under Decrease Group Size: Take the Action: Remove 1 instances

Creating an auto scaling group (continued..)

- Click Review
- Click Create Auto Scaling Group.
- Click Close.
- Click Instances in the left panel to view your worker instances.

Your Auto Scaling group has only been configured to run a single instance at the moment, and no alarms have been attached to the Scaling Policies.

Once you have verified that the worker node is functioning correctly, you will create alarms to automatically adjust the number of worker nodes.

Dispatching work and viewing results

- In the AWS Management Console, select SQS from the Services menu.
- Select your input queue.
- Confirm that there is 1 Message Available output queue. It may take a few minutes after creating your worker instances for the message to move to the output queue from the input queue.

Follow these steps to view the output and open the resulting link in a browser:

Dispatching work and viewing results (continued..)

- Select your output queue.
- From the Queue Actions menu, select View/Delete Messages.
- Click Start Polling for Messages.
- Find your message and click More Details to view the message body. The message will contain a link to the output.
- To view the montage image, select the link, right-click, and choose Go to Address in New Tab.
- Send more messages to your input queue (20+) so that the queue is above zero for several minutes. You can use the Send Another Message button to send the same message multiple times.

Monitoring the cluster

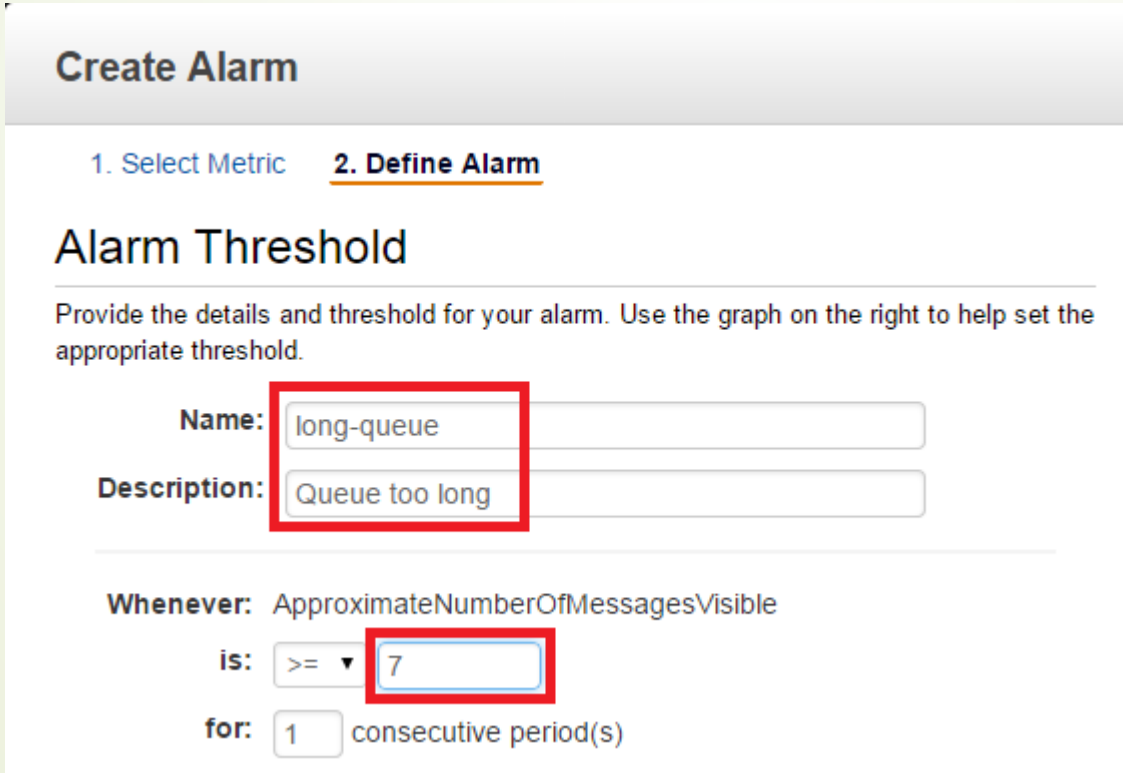
- In the AWS Management Console, select CloudWatch from the Services menu.
- Click Browse Metrics.
- Click the SQS Metrics header. If the header is not visible, return to your input queue and ensure that there are messages queued for processing. This will trigger metrics to be sent to CloudWatch after a few minutes.
- Select the line for:

  Queue Name: input

  Metric Name: ApproximateNumberOfMessagesVisible

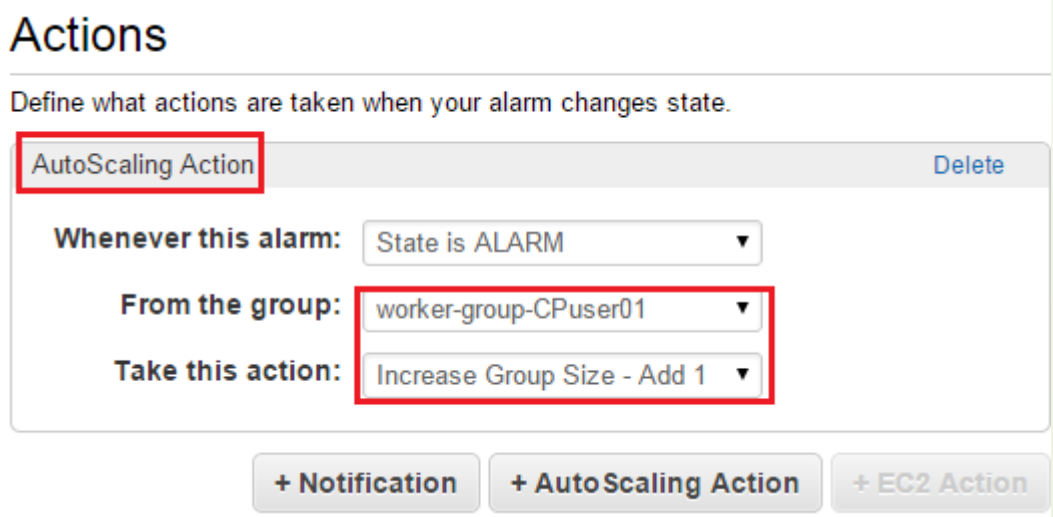
Monitoring the cluster (continued…)

- Click Create Alarm (in the bottom-right).
- In Alarm Threshold, use these values:



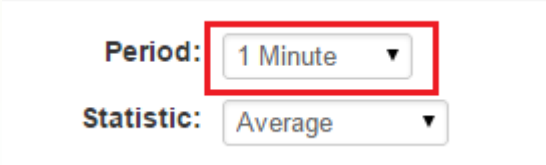
Monitoring the cluster (continued…)

- In Actions:
  - a. Delete the existing Notification action.
  - b. Add an AutoScaling Action, then use these values:



Monitoring the cluster (continued…)

- Set Period to 1 minute (bottom-right):



- Click Create Alarm.

- Try adding a CloudWatch Alarm to scale-in your worker nodes when the queue size is below 5.

LAB (2%)

- Bucket should have output image. (0.5%)
- The output queue should have output message (0.5%)
- You should show me auto scaling history. (1%)