**Autoscaling**

Provide on demand infrastructure

Vertical scaling -: Nothing but server resource / system resource if you increase it is called vertical scaling

OR

If you increase the system resource on same server resource its called as Vertical resource

Horizontal Scaling -: rather than increasing the same server capacity will add number of server according to demand it’s called as horizontal scaling.

In Aws autoscaling is nothing but horizontal scaling.

Scenarios -:

* We have very less time you need to provide resources to run your application continuously.
* Every day particular time my system takes more CPU utilization for that particular time you can increase your system capacity.

How does Autoscaling work ?

* Cloud watch will monitor our server, if CPU matrix britches threshold limit
* Then automatically trigger alarm
* In alarm we will define the scale out / scale in alarm
* This trigger alarm will execute policy.
* Scale out for adding server
* Scale in for removing server

Steps-:

Firstly Select Launch Configuration

* Create launch configuration
* Select AMI
* Select instance type
* Configure details 🡪 You have to add details 🡪 next
* Add storage -> add how many storages you want
* Configure security group-🡪 here you can assign security groups
* Review 🡪 create launch configuration
* Assign existing key or create new one -🡪 Create launch configuration
* Launch configuration successfully created

Then Create Auto Scaling Group

* Click on Auto scaling group
* Create autoscaling group
* Select launch configuration
* You are configuration started now

1.Configure autoscaling group details

-> Give Group name

-> select VPC under Network tab

-> Group size -: you can select number of instance

-> subnet -: select subnets

2.Configure Scaling Policy

-> You will show 2 options there

a. keep this group its initial size

b. Use scaling policy to adjust the capacity of this group

-> Select b option

-> you will show more options

-> Scale between and instances. These will be the minimum and maximum size of your group. (select how many min and max size ,here I given 1 and 3)

-> Select Scale the Auto Scaling group using step or simple scaling policies

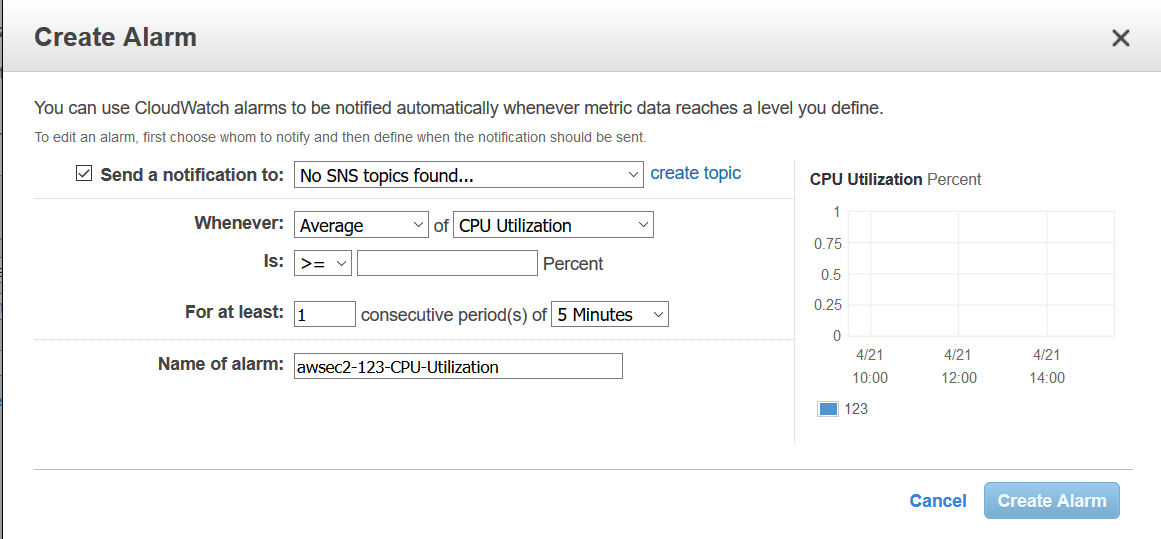
-> will show Increase Group Size and Decrease Group Size

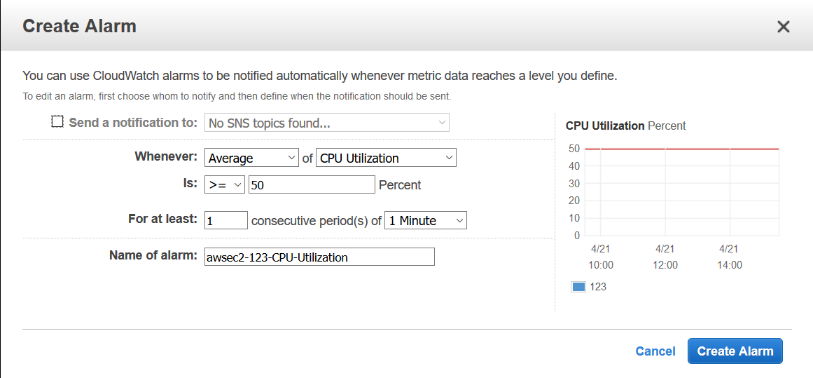
-> Firstly select Increase group size

. Select on add new alarm

. uncheck send a notification to option

. select conditions

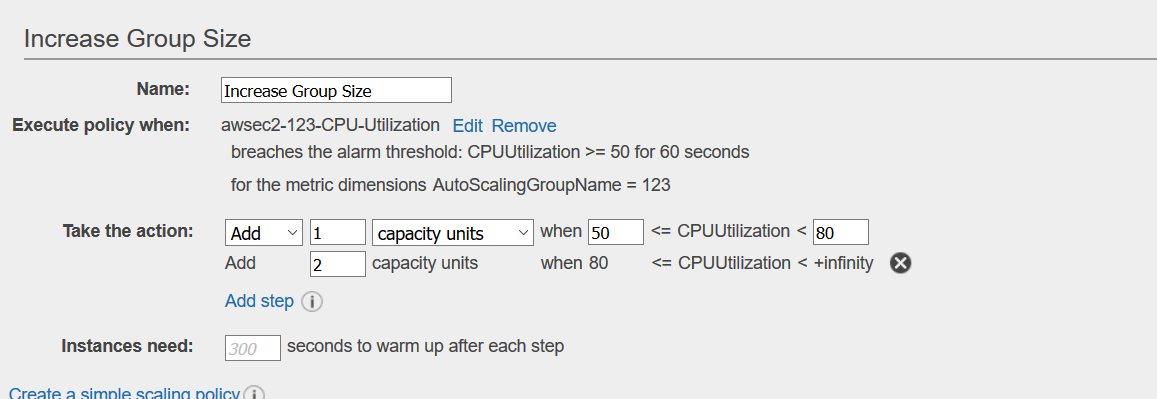


. create conditions 

Here ,I created for when cpu utilization will upto 50 % then It will alarm

. select on create alarm

. Then set Take The Action as belows



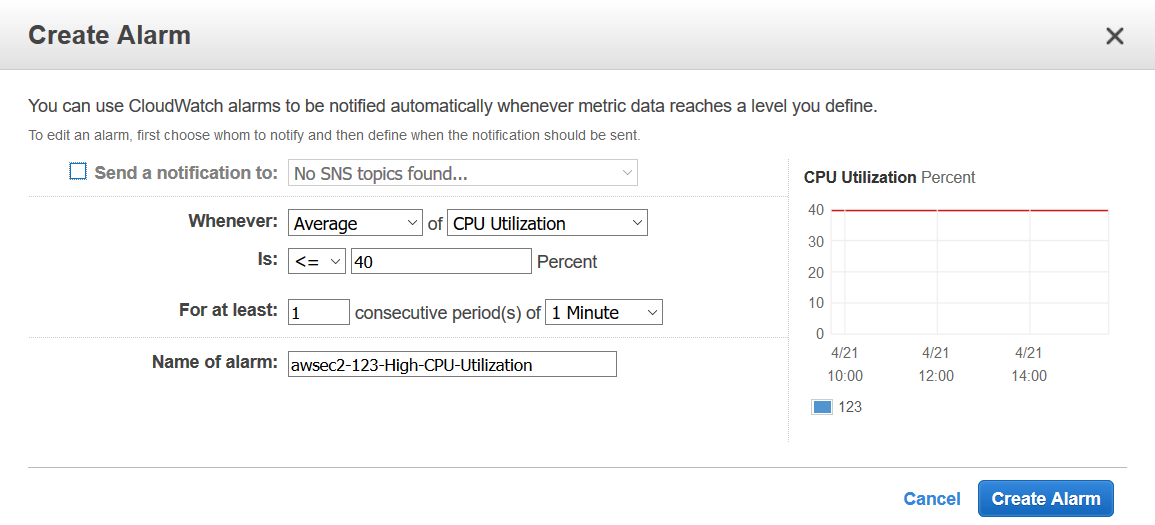
Here , when cpu utilization will more than 50 % then create one more instance

If cpu utilization upto 80 % then will create 2’nd instance

* Secondly, create Decrease Group Size

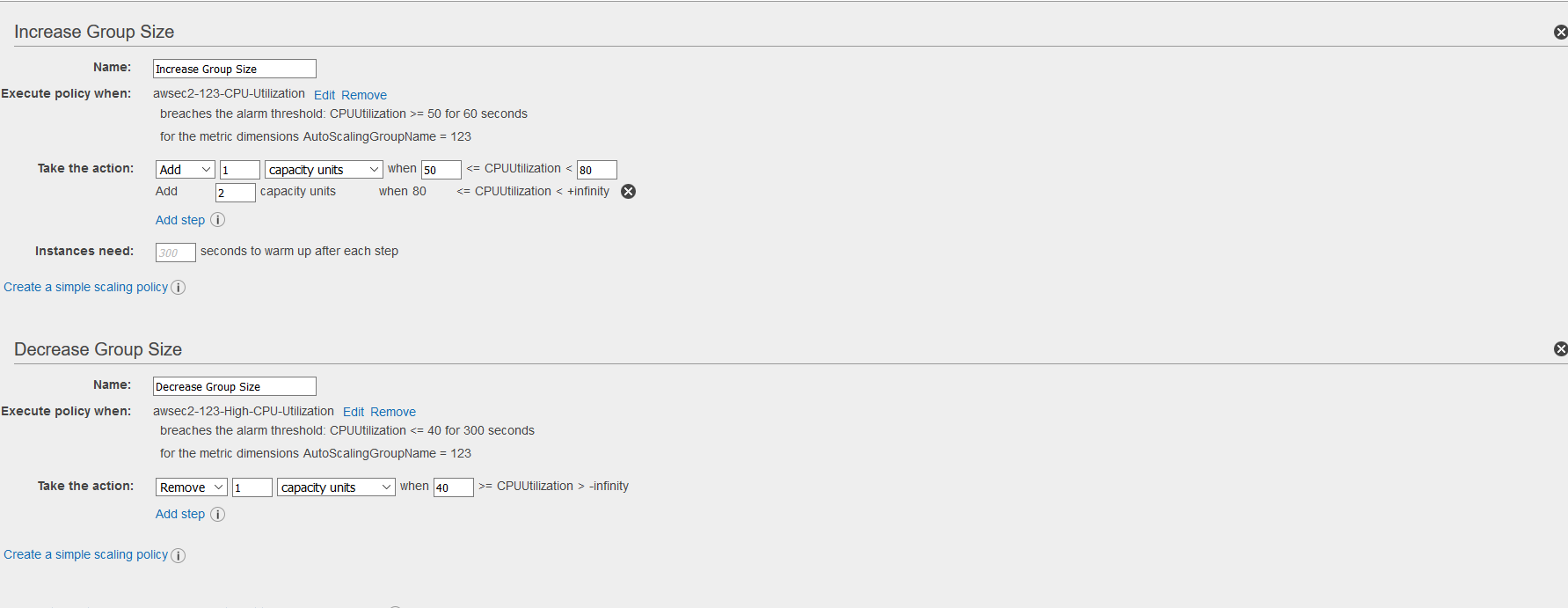
. Use same steps as above

. create alarm (set alarm if cpu utilization will below 40 % )



. select create alarm

. Set Take the action as remove



.select on next

3. Configure notification

-> Its notify you via sms when instance added or removed

-> It is using simple notification service (sns)

-> select on next

4.Configure Tags

-> Give Key – Name

Value – which you want

-> next

5.Review

-> Review it and select on Create Auto Scaling Group

Your group created