



Week 7  
18<sup>th</sup> June, 2022





# Agenda

1. Monitoring and Autoscaling
2. The Three Pillars of Observability  
- James
3. STAR Technique  
- Jamila



Become a Solutions Architect

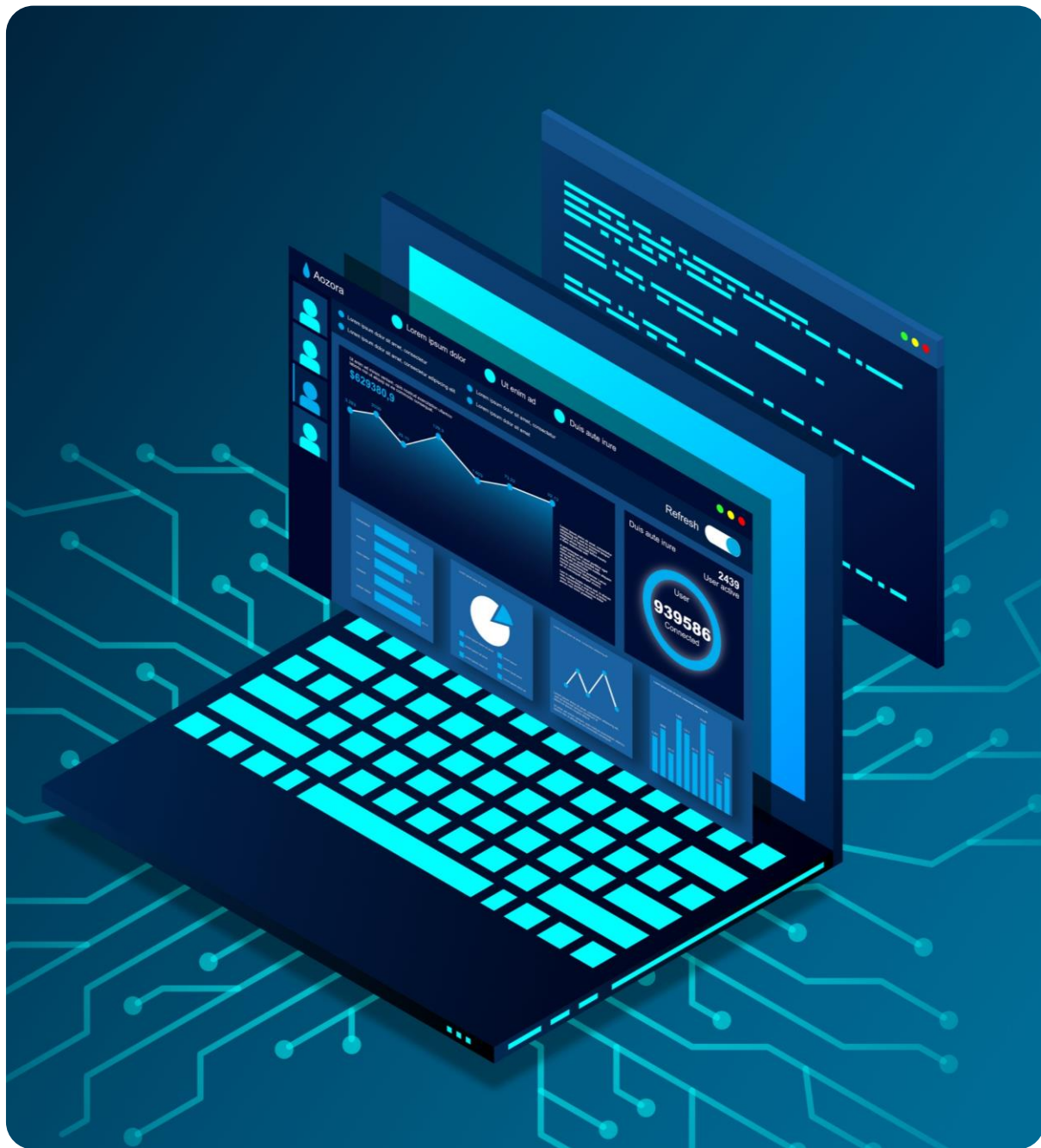
# Monitoring in AWS



Amazon  
CloudWatch







Why Monitor?

# Why monitoring is important?



Performance



Utilization



Health



Security

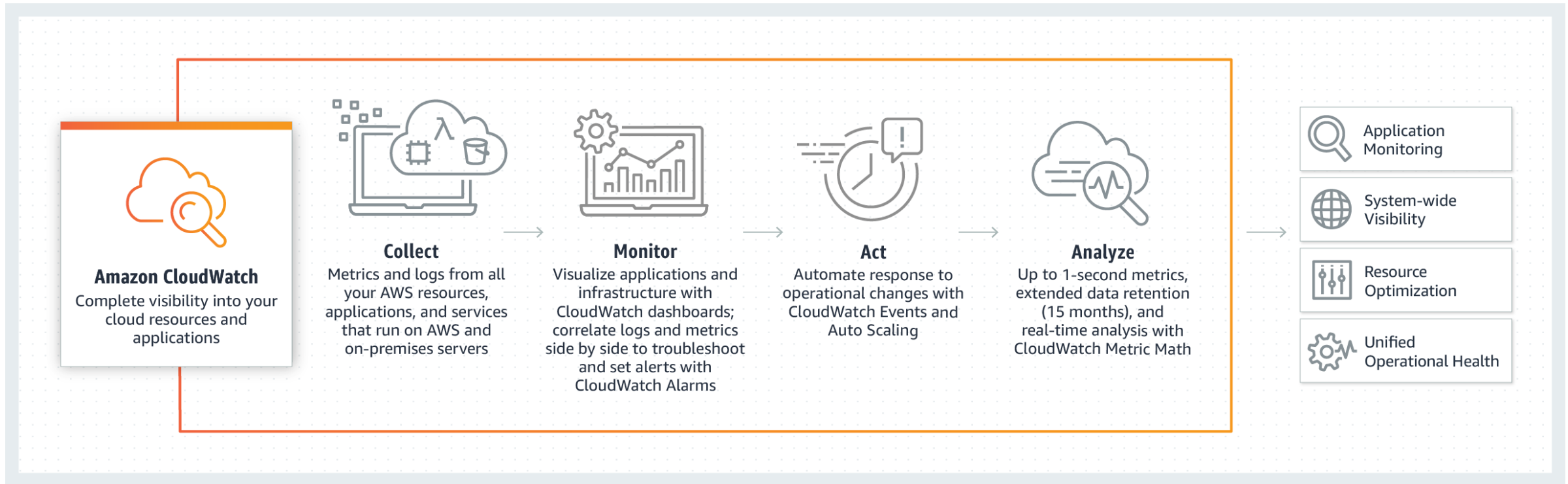




Amazon CloudWatch

# Amazon CloudWatch

- Amazon CloudWatch is a monitoring and observability service.



- CloudWatch collects monitoring and operational data in the form of logs, metrics, and events, and visualizes it using automated dashboards so you can get a unified view of your AWS resources, applications, and services that run on AWS and on premises.



# Amazon CloudWatch – Terminology

- Metrics

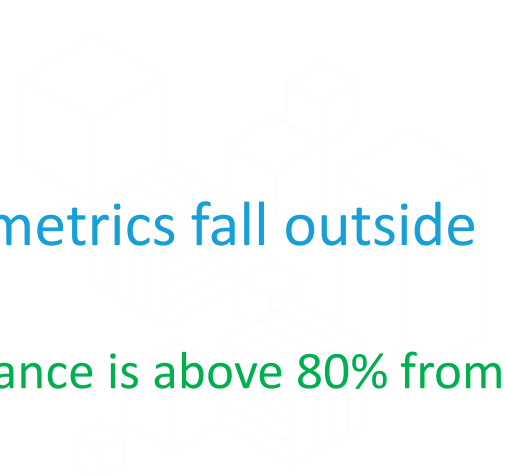
- Metrics are data about the performance of your system
  - Example – Amazon EC2 Instance: CPU Utilization

- Events

- Amazon CloudWatch Events (now EventBridge) delivers a near real-time stream of system events that describe changes in Amazon Web Services (AWS) resources.
  - Example – Amazon EC2 instance changes from *running* to *stopping*

- Alarm

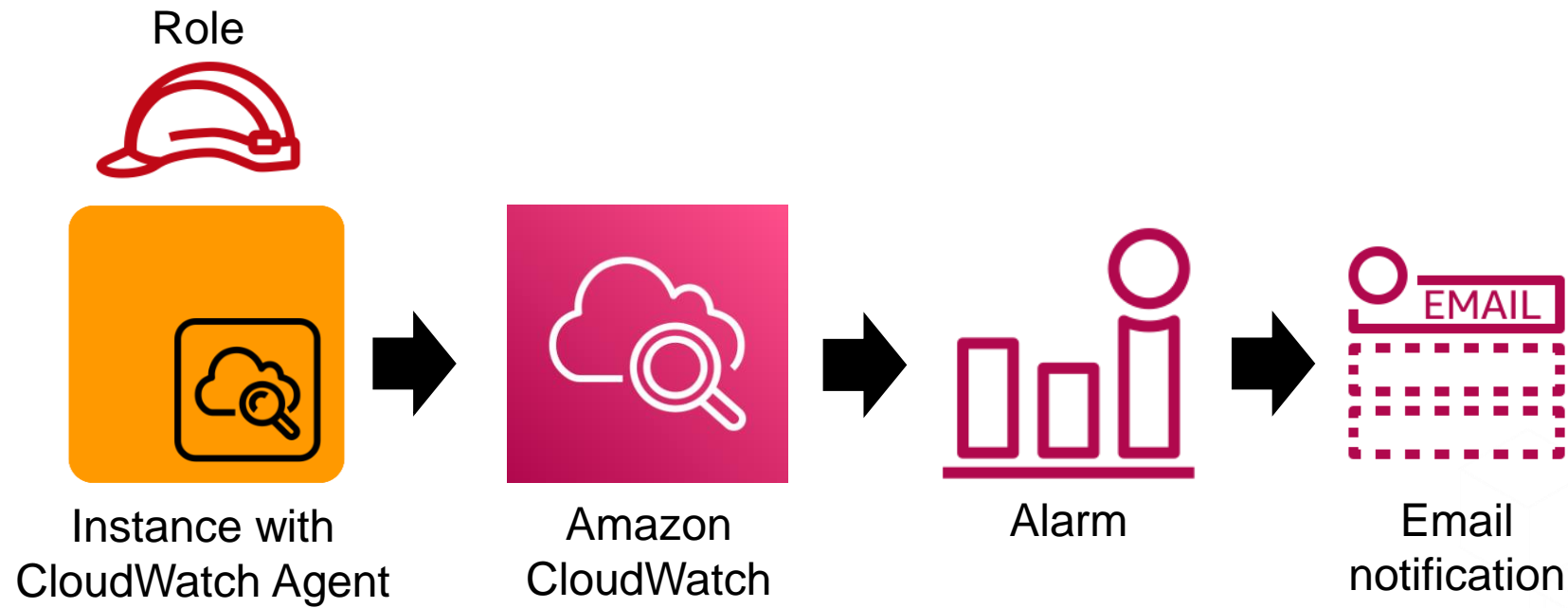
- Allows you to receive notifications or trigger an action when the metrics fall outside of the levels (high or low thresholds) that you configure
  - Example – Email an administrator if CPU utilization of an Amazon EC2 instance is above 80% from last 10 minutes





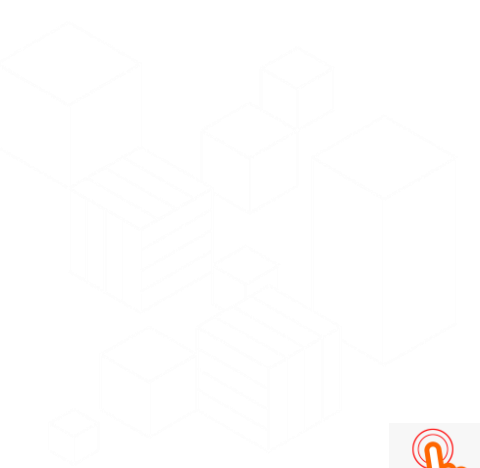
# Amazon CloudWatch Metrics for Amazon EC2

- Standard Metrics
- Custom Metrics



# Installing Amazon CloudWatch Agent using UserData Script

- `#!/bin/bash`
- `yum update -y`
- `sudo yum install -y perl-Switch perl-DateTime perl-Sys-Syslog perl-LWP-Protocol-https perl-Digest-SHA.x86_64`
- `cd /home/ec2-user/`
- `curl https://aws-cloudwatch.s3.amazonaws.com/downloads/CloudWatchMonitoringScripts-1.2.2.zip -O`
- `unzip CloudWatchMonitoringScripts-1.2.2.zip`
- `rm -rf CloudWatchMonitoringScripts-1.2.2.zip`



# Sending Custom EC2 Metrics to CloudWatch

- Verify the access to Amazon Cloudwatch

```
# /home/ec2-user/aws-scripts-mon/mon-put-instance-data.pl --mem-util --verify --verbose
```

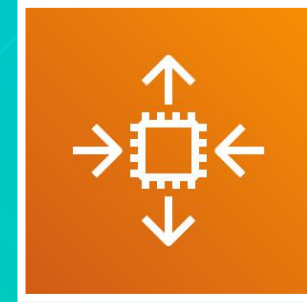
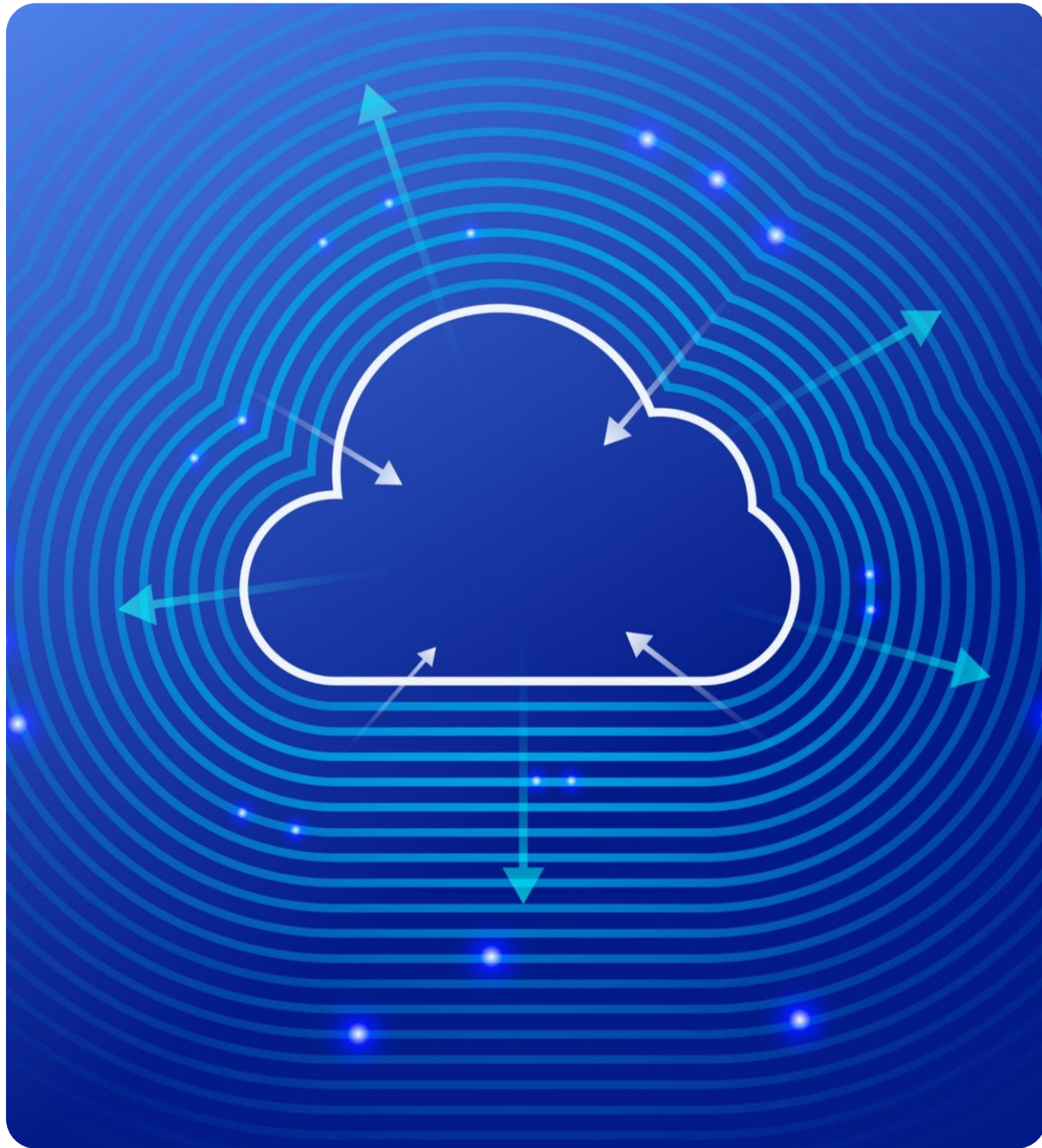
- Send Metrics

```
# /home/ec2-user/aws-scripts-mon/mon-put-instance-data.pl --mem-util --mem-used --mem-avail
```

- Automating through Cronjob

- # sudo su
- # vi /etc/crontab
- # \*/1 \* \* \* \* root /home/ec2-user/aws-scripts-mon/mon-put-instance-data.pl --mem-util --mem-used --mem-avail

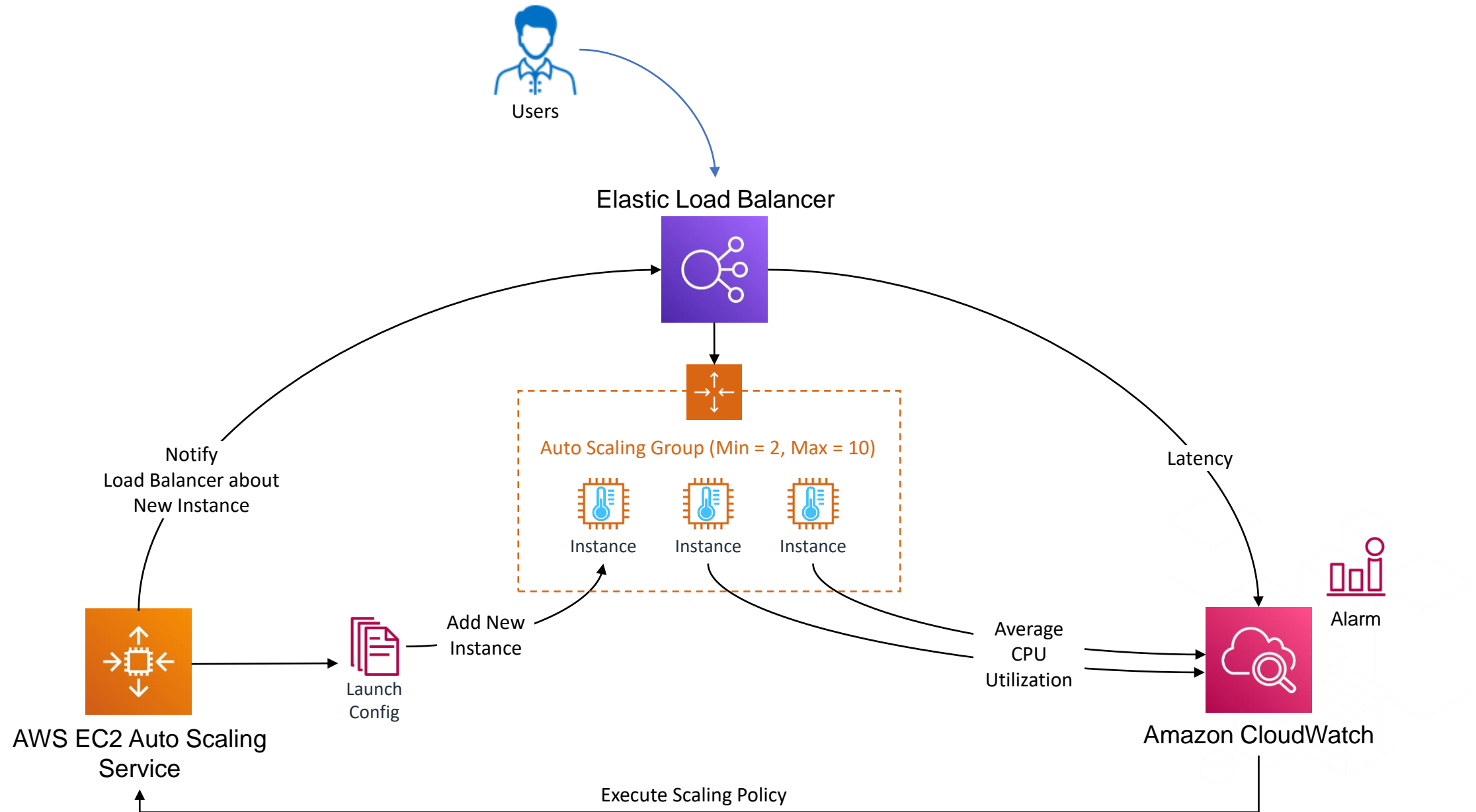




Amazon EC2 Auto Scaling

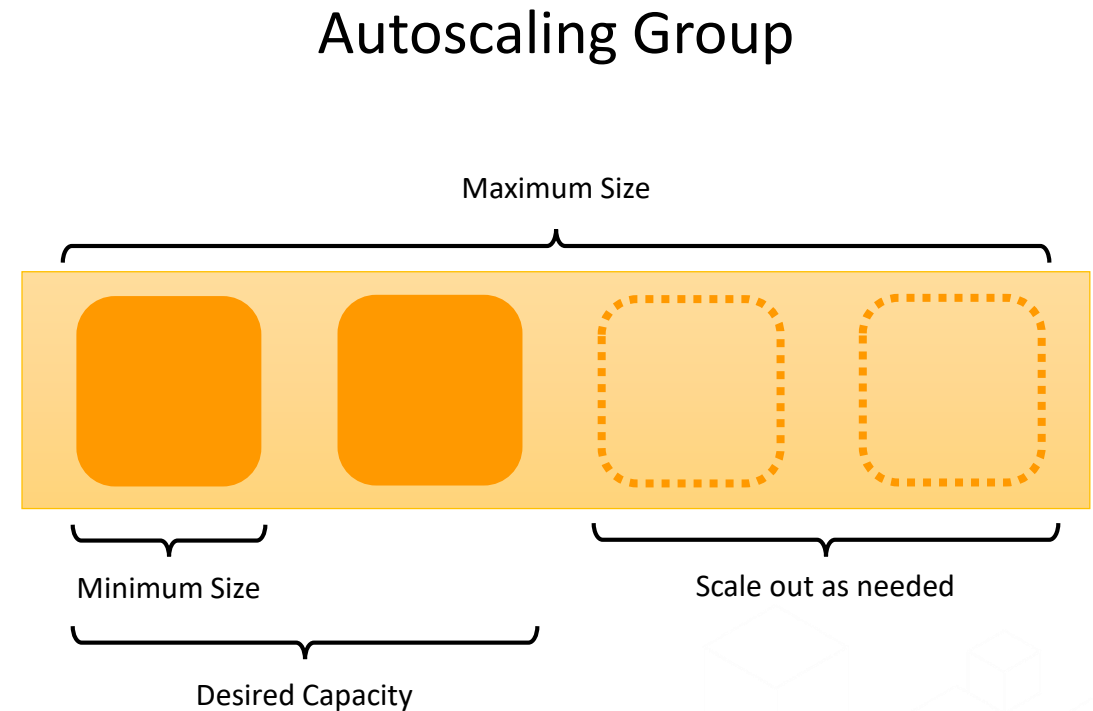


# Amazon EC2 Auto Scaling



# Autoscaling Group

- Min, Max and Desired
- Different Instance Types
- Different Billing Family
- Across Availability Zones

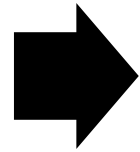


Generate Load for testing

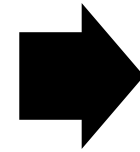
```
# sudo amazon-linux-extras install epel
# sudo yum -y install stress
# uptime
# sudo stress --cpu 8 -v --timeout 3000s
```

# Configuring Autoscaling

Create a  
Base/Golden  
Template



Create a  
Launch  
Configuration



Configure  
AutoScaling  
Group



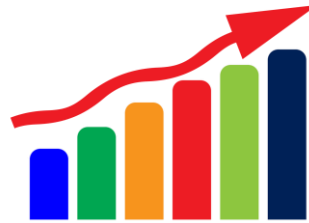
# Autoscaling Policies



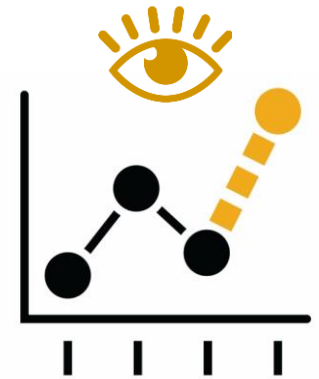
Manual  
Scaling



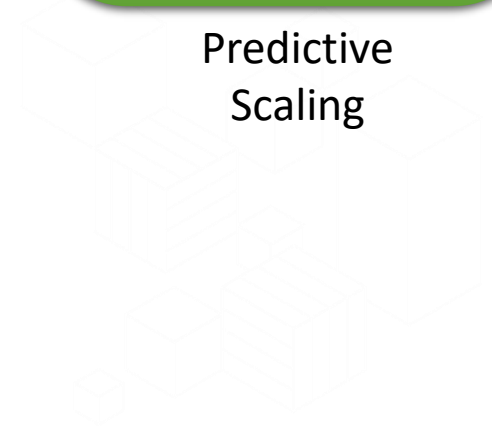
Scheduled  
Scaling



Dynamic  
Scaling



Predictive  
Scaling





# Manual Scaling

**Desired Capacity**



1

**Min**



1

**Max**



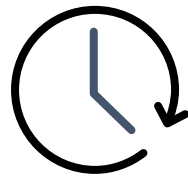
5



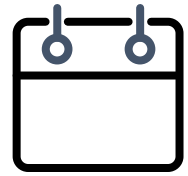
# Scheduled scaling

Name	Sample	
Auto Scaling Group	LCtoLT	
Provide at least one of Min, Max and Desired Capacity		
Min	5	
Max	10	
Desired Capacity	7	
Recurrence	Once	
Start Time	: 00 UTC Specify the start time in UTC	

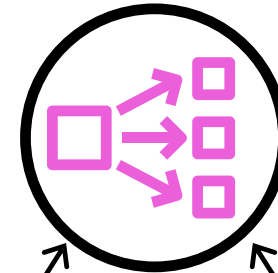
Recurring scaling events



Schedule individual events



Elastic Load Balancing



Amazon EC2 instances

Auto Scaling group

# Dynamic scaling with target tracking

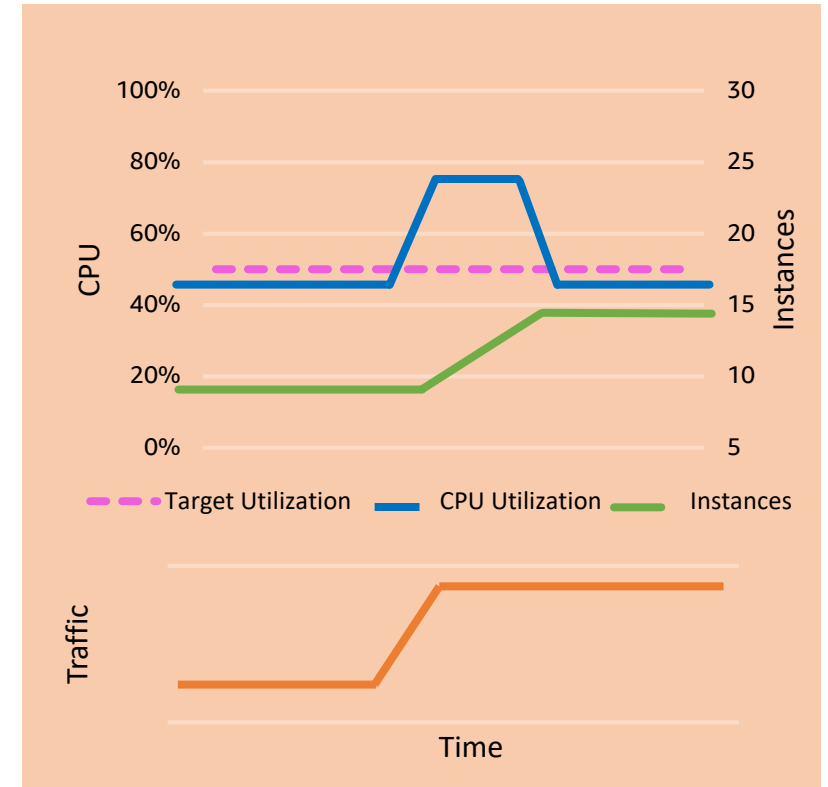
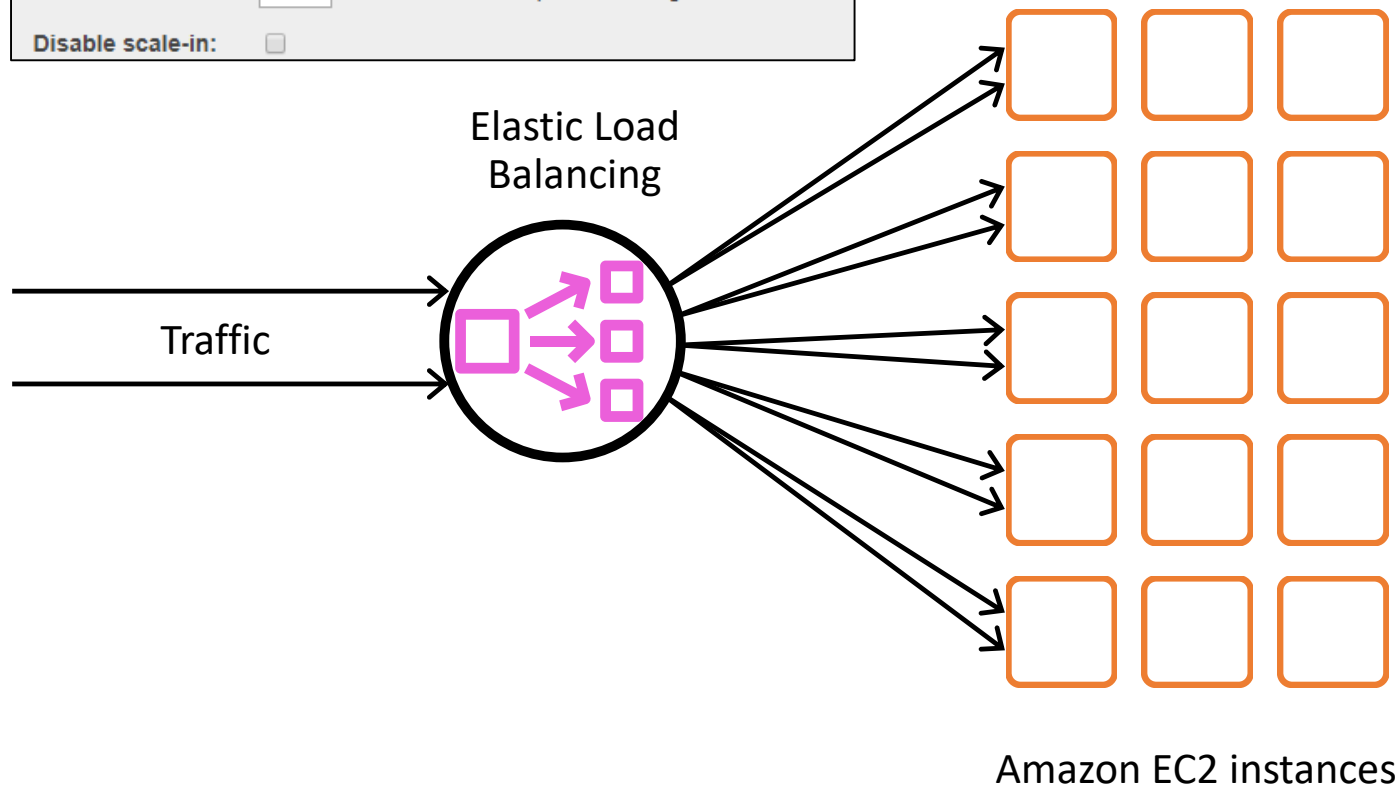
Name:

Metric type:

Target value:


Instances need:  seconds to warm up after scaling

Disable scale-in: ☐





# Dynamic scaling with step scaling


**Name:**

**Execute policy when:**   [Create new alarm](#)

breaches the alarm threshold: CPUUtilization  $\geq 50$  for 5 consecutive periods of 60 seconds  
for the metric dimensions AutoScalingGroupName = WebApp-WebServerGroup-13H20HQT519C


**Take the action:**

Add	<input type="text" value="2"/>	instances	when <input type="text" value="50"/>	$\leq$ CPUUtilization < <input type="text" value="60"/>	
Add	<input type="text" value="4"/>	instances	when 60	$\leq$ CPUUtilization < <input type="text" value="70"/>	
Add	<input type="text" value="6"/>	instances	when 70	$\leq$ CPUUtilization < +infinity	

[Add step](#) 


**Instances need:**  seconds to warm up after each step


**Name:**

**Execute policy when:**   [Create new alarm](#)

breaches the alarm threshold: CPUUtilization  $\leq 30$  for 10 consecutive periods of 60 seconds  
for the metric dimensions AutoScalingGroupName = WebApp-WebServerGroup-13H20HQT519C

**Take the action:**

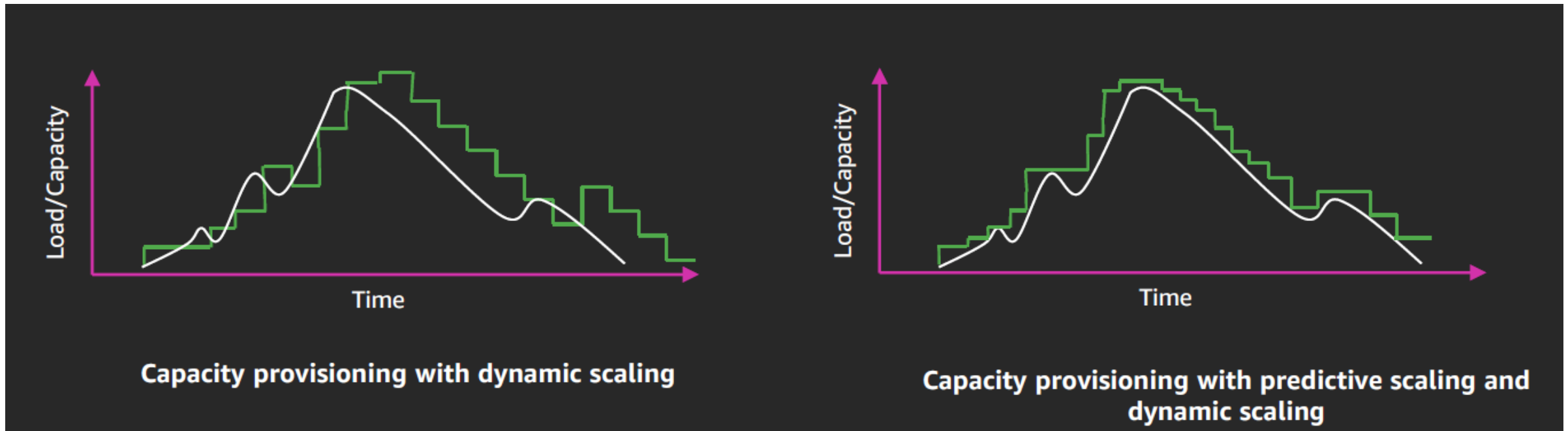
Remove	<input type="text" value="2"/>	instances	when <input type="text" value="30"/>	$\geq$ CPUUtilization > <input type="text" value="20"/>	
Remove	<input type="text" value="4"/>	instances	when 20	$\geq$ CPUUtilization > -infinity	

[Add step](#) 



# Predictive Scaling

- With predictive scaling, AWS Auto Scaling analyzes the history of the specified load metric from the past 14 days (minimum of 24 hours of data is required) to generate a forecast for two days ahead.
- It then schedules scaling actions to adjust the resource capacity to match the forecast for each hour in the forecast period.



# How Netflix uses Autoscaling?

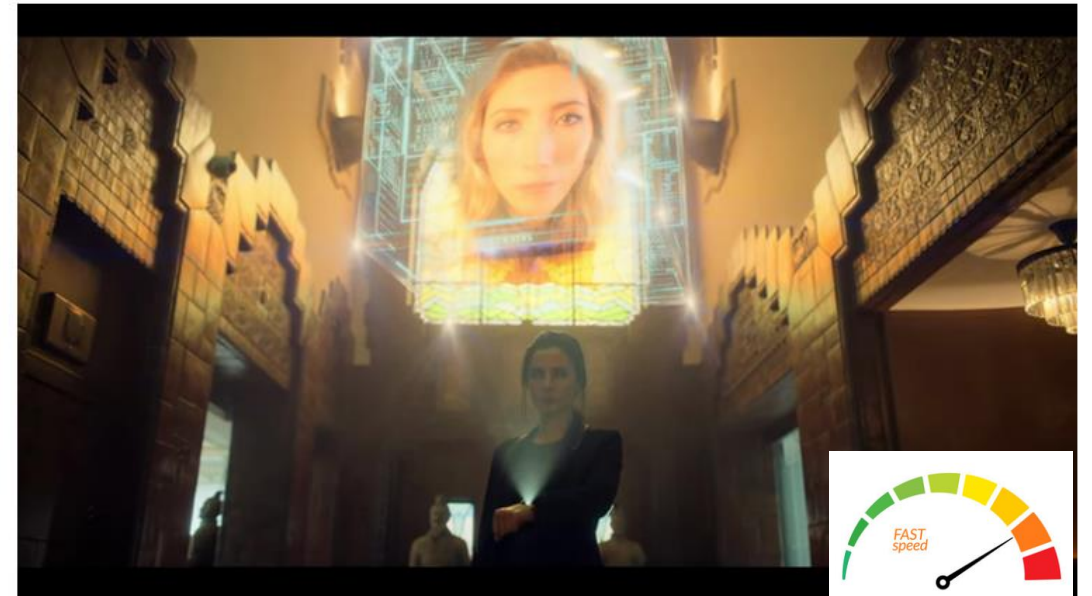
# NETFLIX

Data from 2018

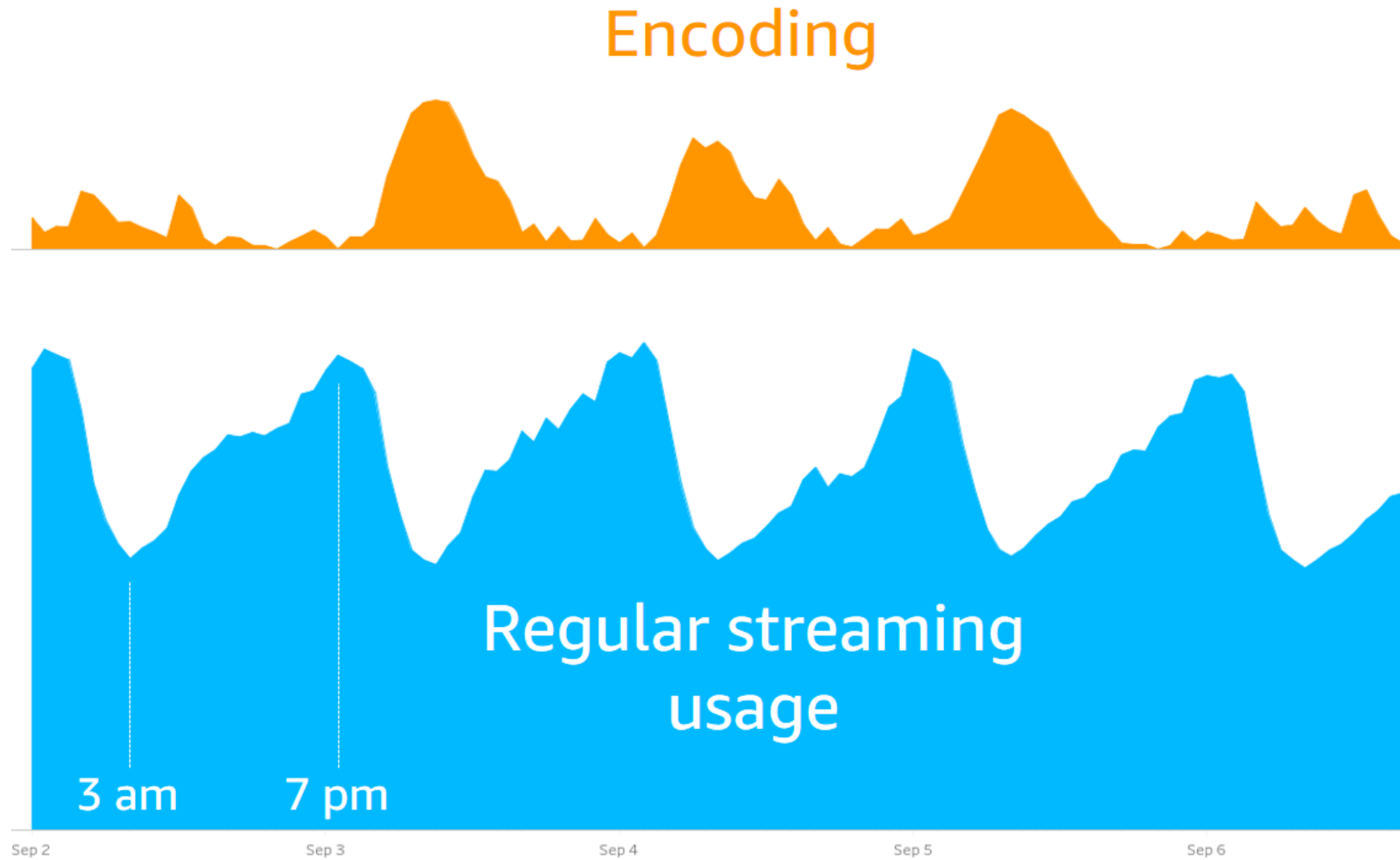
- 137+ million subscribers
- 190+ countries
- 1700+ supported device types



- Three regions
- 250,000+ reserved instances
- 20,000 Auto Scaling groups



# How Netflix uses Autoscaling?





# The Three Pillars of Observability

- James





‘STAR’ Technique

- Jamila



**‘STAR’ Technique to  
Answer Behavioural  
Interview Questions**



S



T



A



R

## SITUATION

## TASK

## ACTION

## RESULT

- Detail the background.
- Provide a context.
- Where? When?

- Describe the challenge and expectations.
- What needed to be done?
- Why?

- Elaborate specific action.
- What did you do? How?
- What tools did you use?

- Explain the results:
  - Accomplishments
  - Recognition
  - Savings, etc.
- Quantify.

Thank you for attending. See you next Saturday (25-Jun-2022)



Become a Solutions Architect

For content check **Resources Link** on BeSA Home Page

