Considering Compute Performance Options



Mike Erickson
DEVELOPER, ARCHITECT, TRAINER, SPEAKER, AUTHOR
@mgerickson





Processing

Capacity

Scaling

Responsive

Economical





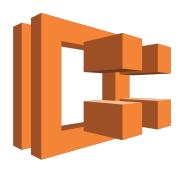
Understand your workload
Understand AWS compute
Gather/analyze data
Test, test, test



AWS Compute Options



EC2
Elastic Cloud Compute

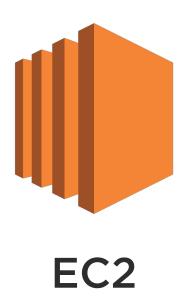


ECS
Elastic Container
Service



AWS Lambda





Original/default option

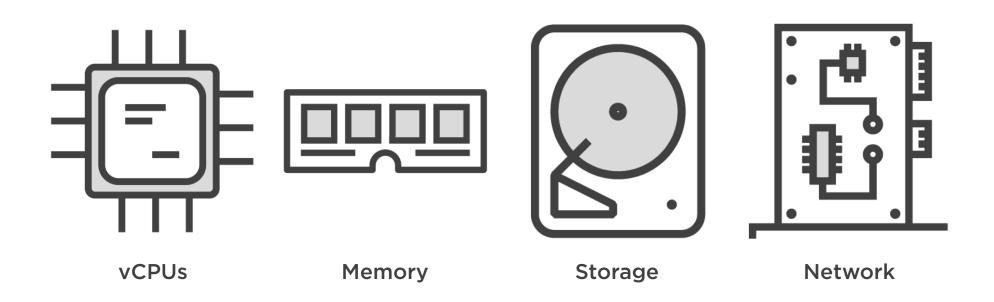
Virtualized servers

Choose resources

You own the OS



EC2 General Resources





Burstable

GPU

FPGA



Instance Types

General purpose

Compute optimized

Memory optimized

Accelerated computing

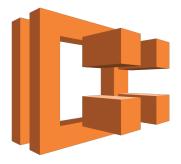
Storage optimized

Bare metal



EC2 Auto Scaling





ECS

Similar workloads as EC2

Migrate apps to the cloud

Long running apps

Batch processing

Microservices





Better utilize resources

ELB

Autoscaling





AWS Fargate





Lambda

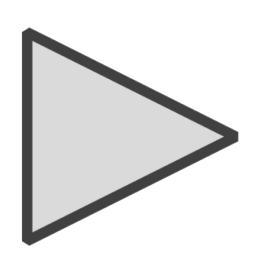
Backend processing

Event processing

Stream processing

Data processing





AWS resource triggers

API endpoints

Choose memory, get CPU

Beware of limits





Simply execute code

Automatic scaling

Fault tolerant

Pay for usage - really





Researching AWS compute options



Applying Our Knowledge



Migrating to AWS Compute



Portfolio of web applications

Global presence

Data collection from clinical trials



Considerations for Choice



First of many



Time to migrate



Predictable usage





ECS

Lift and shift

- Easy to containerize
- Able to scale

Size instances

Leverage for future applications



Built for the Cloud



New web application

Match medical devices to needs

Global user base

Considerations for Choice







Lambda

Services behind static site

Pay for runtime

Scales based on demand

Multiple regions

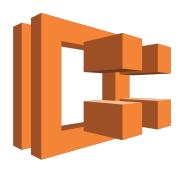
No servers to maintain



AWS Compute Options



EC2
Elastic Cloud Compute



ECS
Elastic Container
Service



AWS Lambda





Storage performance

