
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

CMP213

Building a budget-conscious culture at Standard Chartered Bank

Daniel Kazarinow (he/him)

Head of Cloud, Containers & Grid SRE
Standard Chartered Bank

Carlos Manzanedo Rueda (he/him)

WW SA Lead Flexible Compute
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

CMP213

Building a budget-conscious culture at Standard Chartered Bank

How to use cloud efficiently and effectively

Daniel Kazarinow (he/him)

Head of Cloud, Containers & Grid SRE
Standard Chartered Bank

Carlos Manzanedo Rueda (he/him)

WW SA Lead Flexible Compute
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

- Introduction
- Cloud journey at Standard Chartered Bank
- Cost and compute optimization
- Consumption model, instance selection, elasticity, tooling

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Daniel Kazarinow

Standard Chartered

Singapore

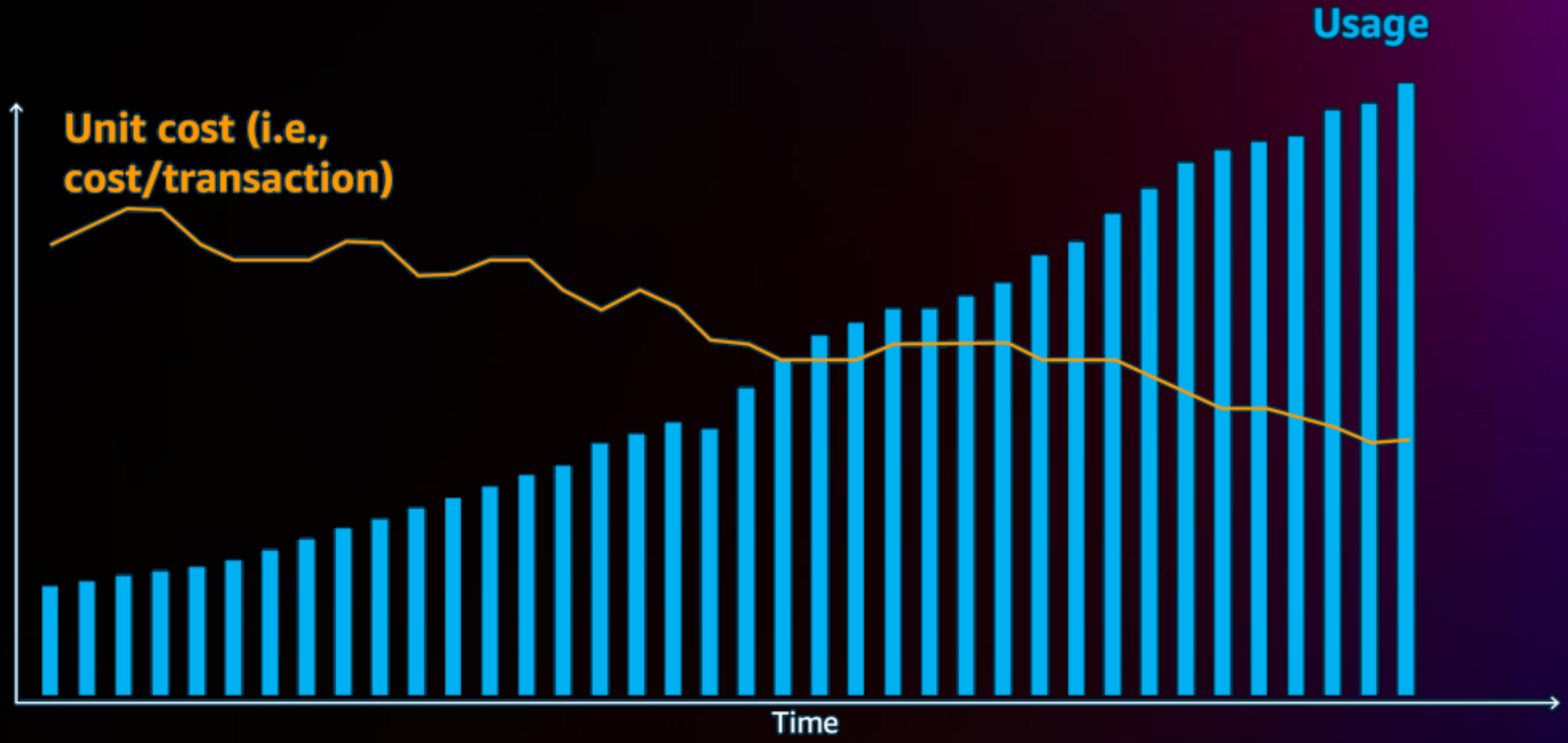
Head of Cloud, Containers & Grid SRE
- Head of Cloud & DevOps SRE
  - Head of Grid Computing
    - Grid Manager

London, UK

- Grid (HPC) Technical Design & Delivery
  - Support Analyst
    - Summer Intern

aws © 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cloud adoption journey – what good looks like



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

FinOps at Standard Chartered Bank

- Collaboration
- Ownership
- Centralized Team
- Reporting
- Business Value
- Variable Cost Model

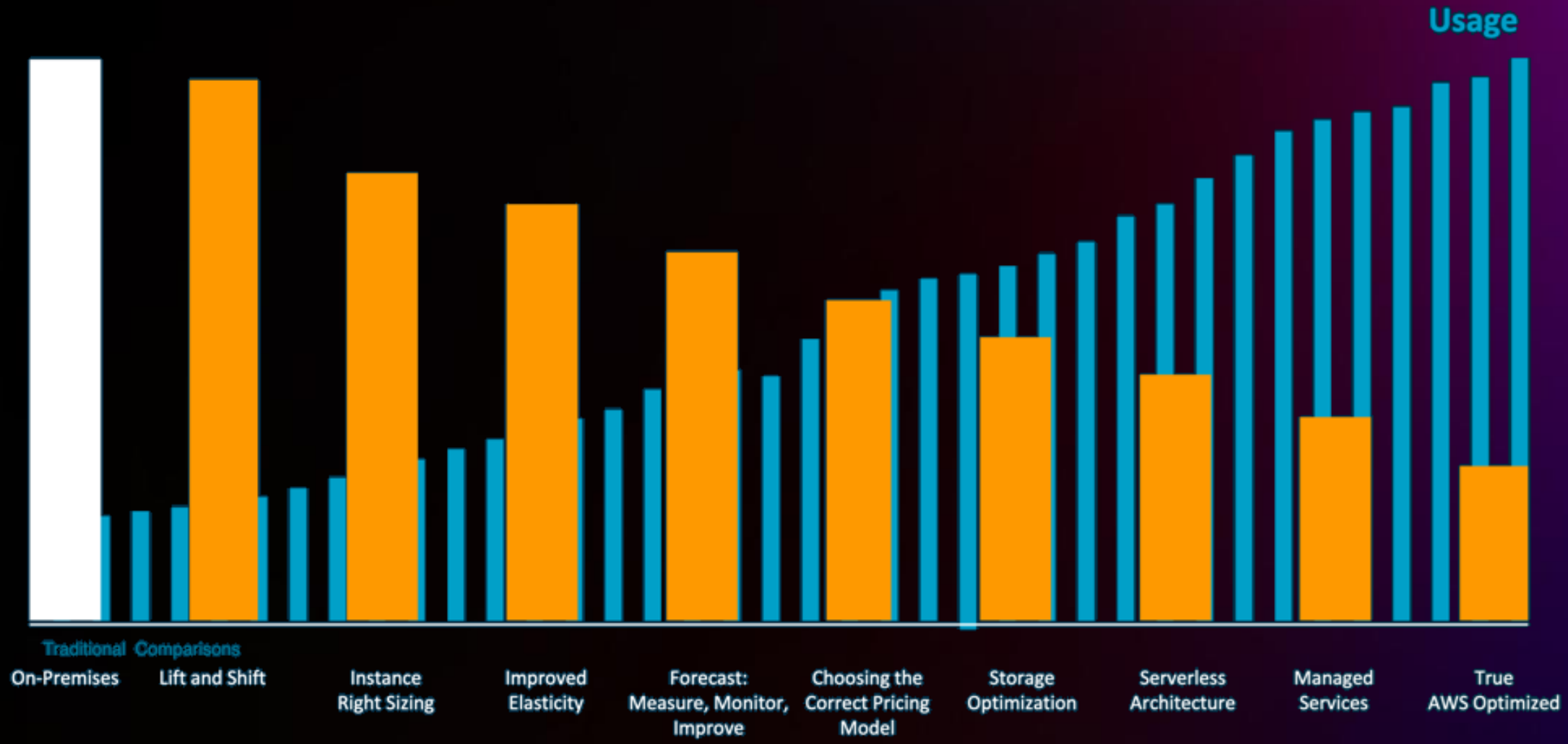


FinOps.org

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost optimization techniques and steps



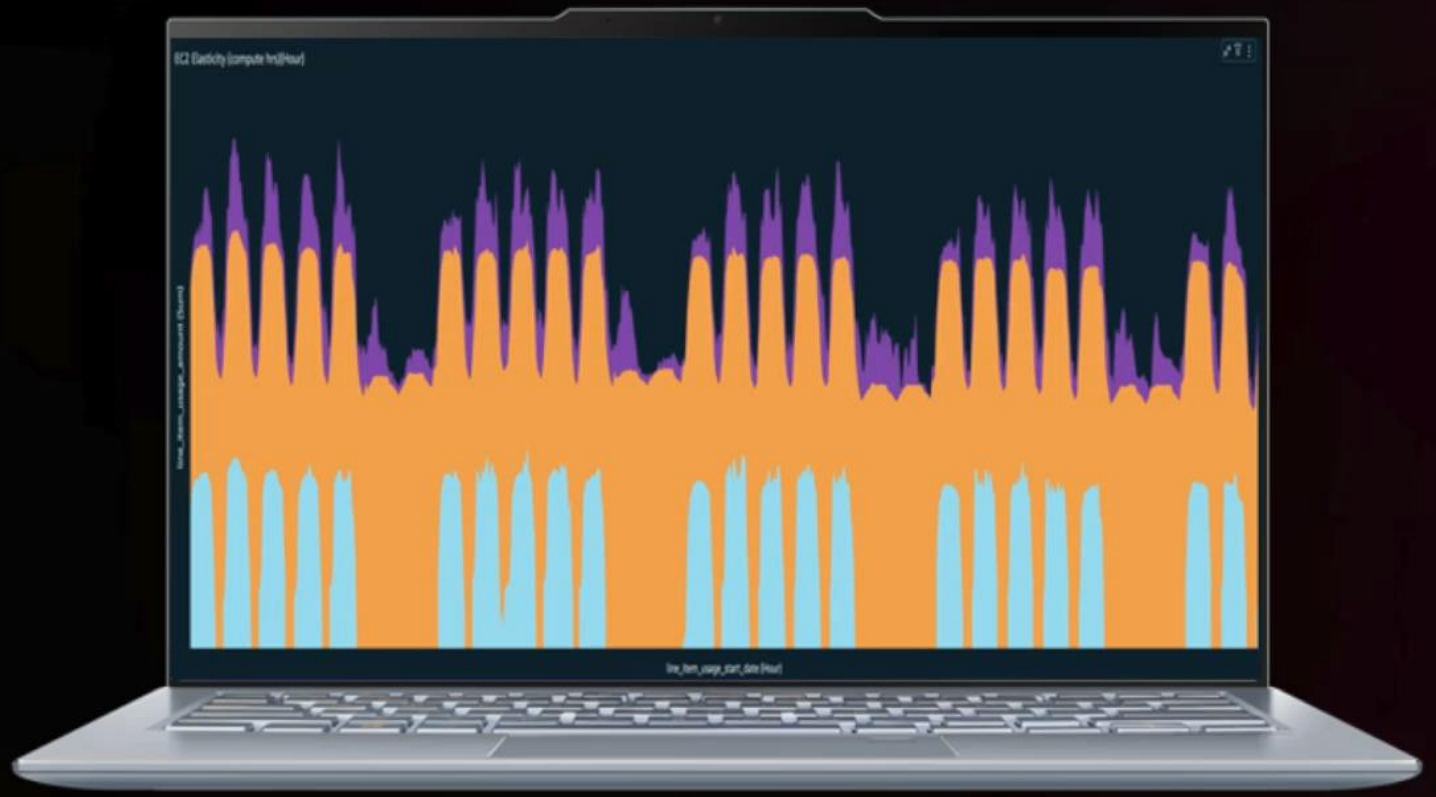
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Consumption model

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost Optimization at Standard Chartered Bank

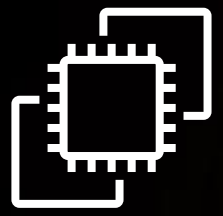
On-Demand Savings Plan Spot Instances



Consumption Model
Right Instance Choice
Elasticity
Tooling

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Different options and compute offerings



Amazon EC2

Virtual server instances in the cloud



Amazon ECS, Amazon EKS

Container management service for running Docker on a managed cluster of EC2



AWS Lambda

Serverless compute for stateless code execution in response to triggers

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Amazon EC2 purchase options

ON-DEMAND INSTANCES

Pay for compute capacity by **the second** with no long-term commitments



Spiky workloads to define needs

SAVINGS PLANS

Great discounts as Amazon EC2 up to 72% **more flexibility**



Flexible access to compute

SPOT INSTANCES

Spare Amazon EC2 capacity at **savings of up to 90%** on on-demand prices



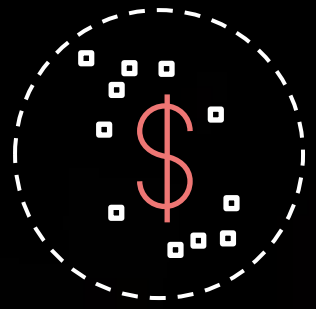
Fault-tolerant, flexible, stateless workloads

Focus on these for optimizing performance and cost

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Types of savings plans



Compute Savings Plans

Greatest flexibility

Discounts up to 66%

Flexible across

- ✓ Instance family & Instance Size e.g., move from C5.2xlarge to M5.4xlarge
- ✓ Region – e.g., change from US East (Ohio) to Europe (London)
- ✓ Operating System – Windows & Linux
- ✓ Tenancy – Default & Dedicated
- ✓ Compute options – e.g., move from Amazon EC2 to AWS Fargate



EC2 Instance Savings Plans

Provides Deepest savings

Discounts up to 72%

Flexible across

- ✓ Size – e.g., move from m5.xl to m5.4xl
- ✓ OS – e.g., change from m5.xl Windows to m5.xl Linux
- ✓ Tenancy – e.g., modify m5.xl Dedicated to m5.xl Default tenancy



SageMaker Savings Plans

Up to 64% off eligible SageMaker machine learning instances

Flexible across

- ✓ SageMaker ML usage
- ✓ Instance Family
- ✓ Size
- ✓ Region

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

On-demand capacity reservations

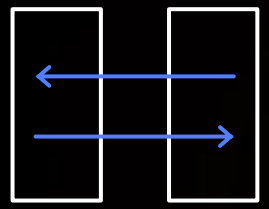
For steady-state workloads



- Manage capacity and discounts independently
- No commitment required; can be created and canceled as needed
- Reserve capacity by Availability Zone
- Capacity held whether or not you run instances
- Share reservations across accounts

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Spot instances for interruptible workloads



Same infrastructure

Spare EC2 capacity from the same infrastructure as On-Demand



Capacity

AWS can reclaim with 2-minute notice; interruptions happen when EC2 needs the spare capacity back

Spot is ideal for:



Fault-tolerant



Flexible



Loosely coupled



Stateless workloads

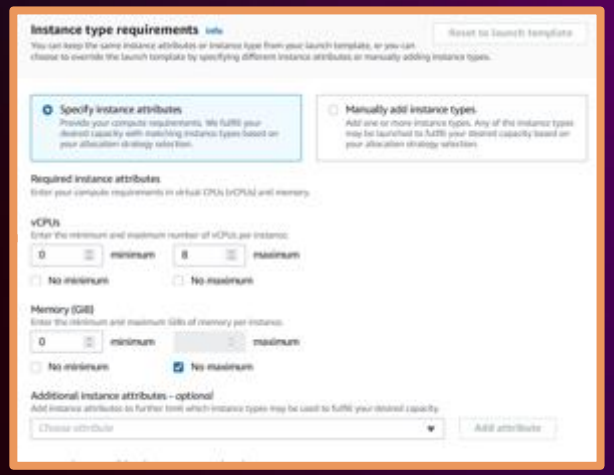
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Spot toolset

Make use of the Spot toolset to get started and get the most out of EC2 Spot:

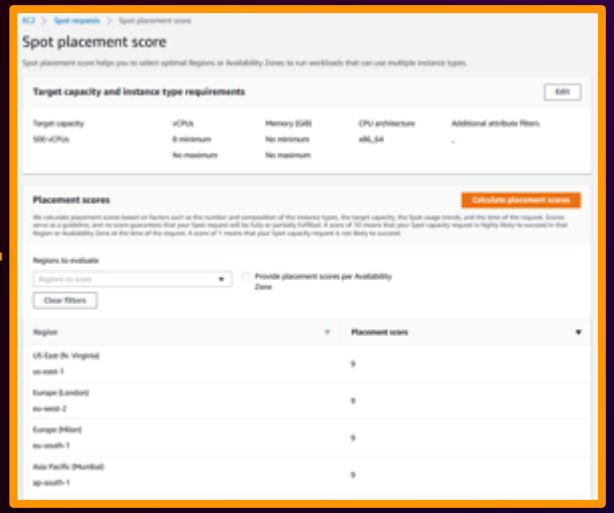
Attribute Based Instance Selection

Rather than selecting specific instance types, this tool allows you to provision capacity on instance attributes (such as CPU/Memory) depending on requirements, also supporting new instances as soon as they become available.



Spot Placement Score

Find the optimal location (Reg/AZ) for your workload. You provide the resource & capacity requirements and Spot Placement Score recommends locations (Reg/AZ) and the likelihood of getting the capacity you need.



Fault Injection Simulator

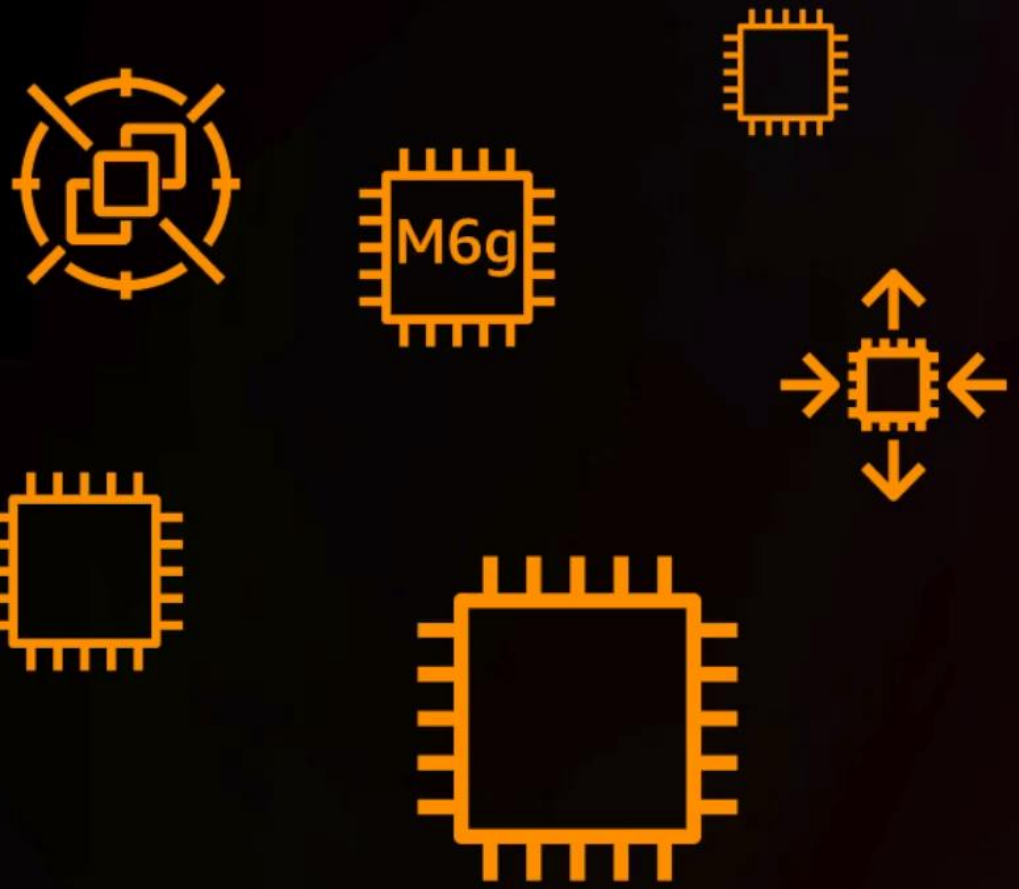
Enables chaos engineering and supports imitation of Spot interruptions, enabling you to prove the resilience of workloads running on Spot.

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Instance selection

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost Optimization at Standard Chartered Bank



Consumption Model
Right Instance Choice
Elasticity
Tooling

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Select the right compute options for your workload

550+

INSTANCE TYPES

for virtually every workload and business need

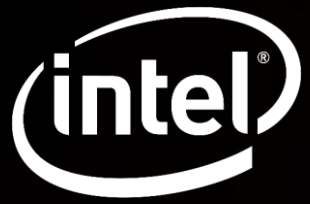
1. What processors can my workload use?
2. What are my workload's performance requirement?
3. What is my workload's consumption pattern?

Flexible list of instance types that fit

Your workload

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Broadest choice of processors and architectures



Intel Xeon Scalable processors



AMD EPYC processor

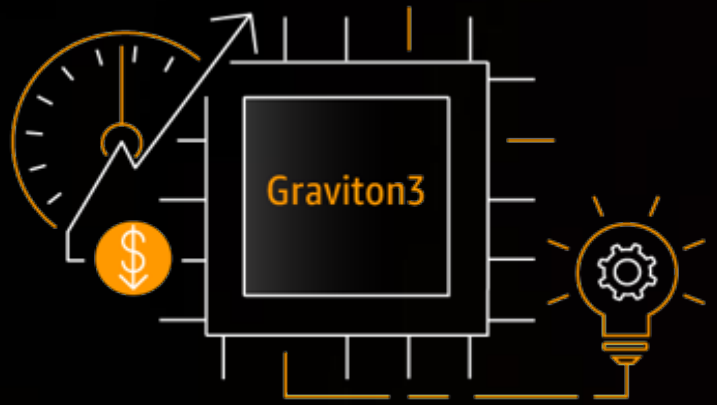


AWS Graviton Processor

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Graviton3 – C7g Instances

Enabling the best price performance for workloads in Amazon EC2



Up to 25% better performance compared to Graviton2

Up to 2x higher floating-point performance, up to 2x faster cryptographic workload performance, and up to 3x better machine learning performance compared to Graviton2

First in the cloud to feature DDR5 memory

60% more energy efficient over comparable EC2 instances

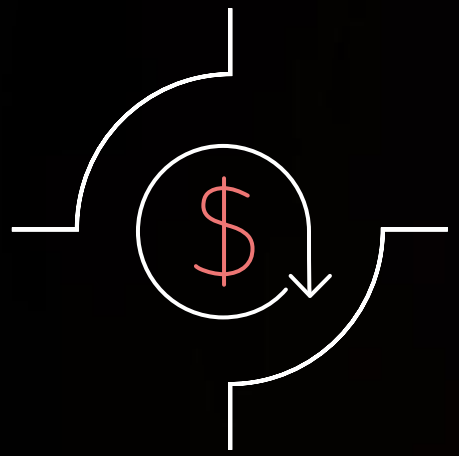
C7g instances, will provide the best price performance for compute-intensive workloads in Amazon EC2

aws

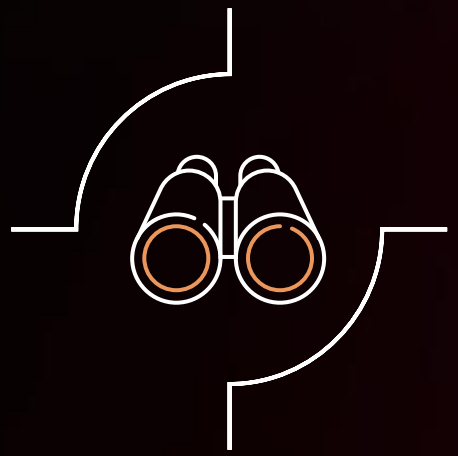
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Compute Optimizer

Recommends optimal instances for AWS Lambda, Amazon EC2, and Amazon EC2 Auto Scaling groups from 145+ instances from M, C, R, T, and X families



Lowers **costs** and improves workload **performance**



Applies insights from millions of workloads to make recommendations



Saves time comparing and selecting optimal resources for your workload

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

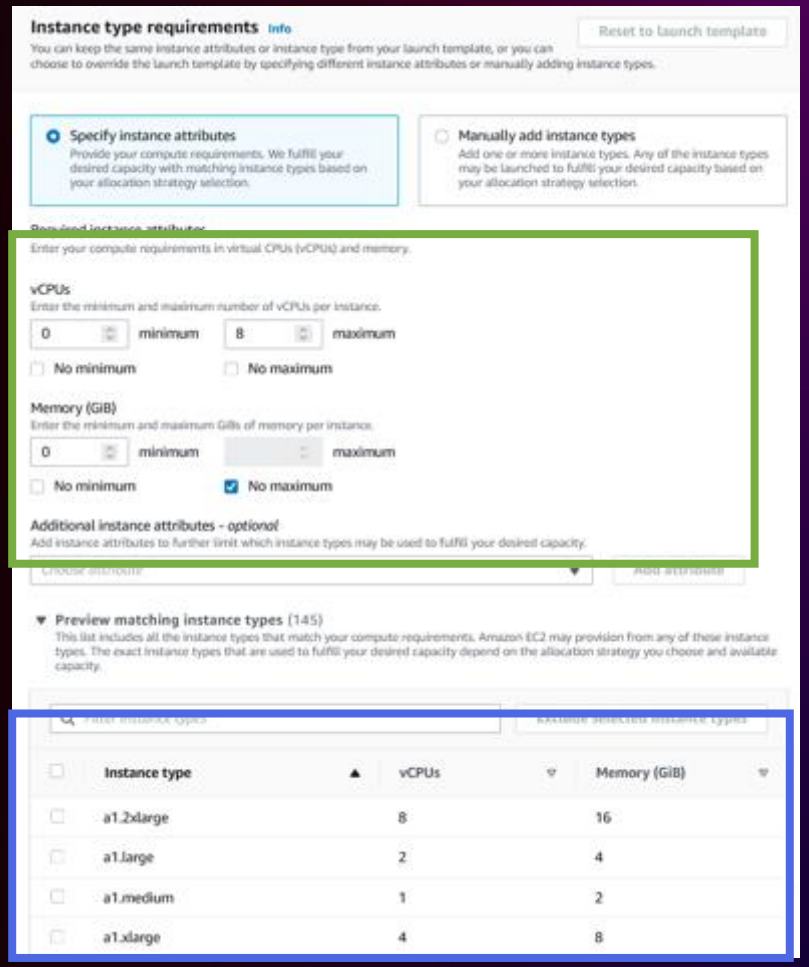
Diversify with attribute-based instance selection

PROVISION CAPACITY BASED ON INSTANCE ATTRIBUTES DEPENDING ON COMPUTE REQUIREMENTS, SUCH AS VCPU, RAM, STORAGE TYPE, AND MORE

Benefits

- Price protected
- Instance flexibility
- Simplifies provisioning
- Mixed instance groups
- Future proof
- Preview instances

- Attribute based instance selection requires the use of Launch templates
- On-Demand allocation strategy: **"lowest-price"**
- Spot Allocation Strategy: **"price-capacity-optimized"**



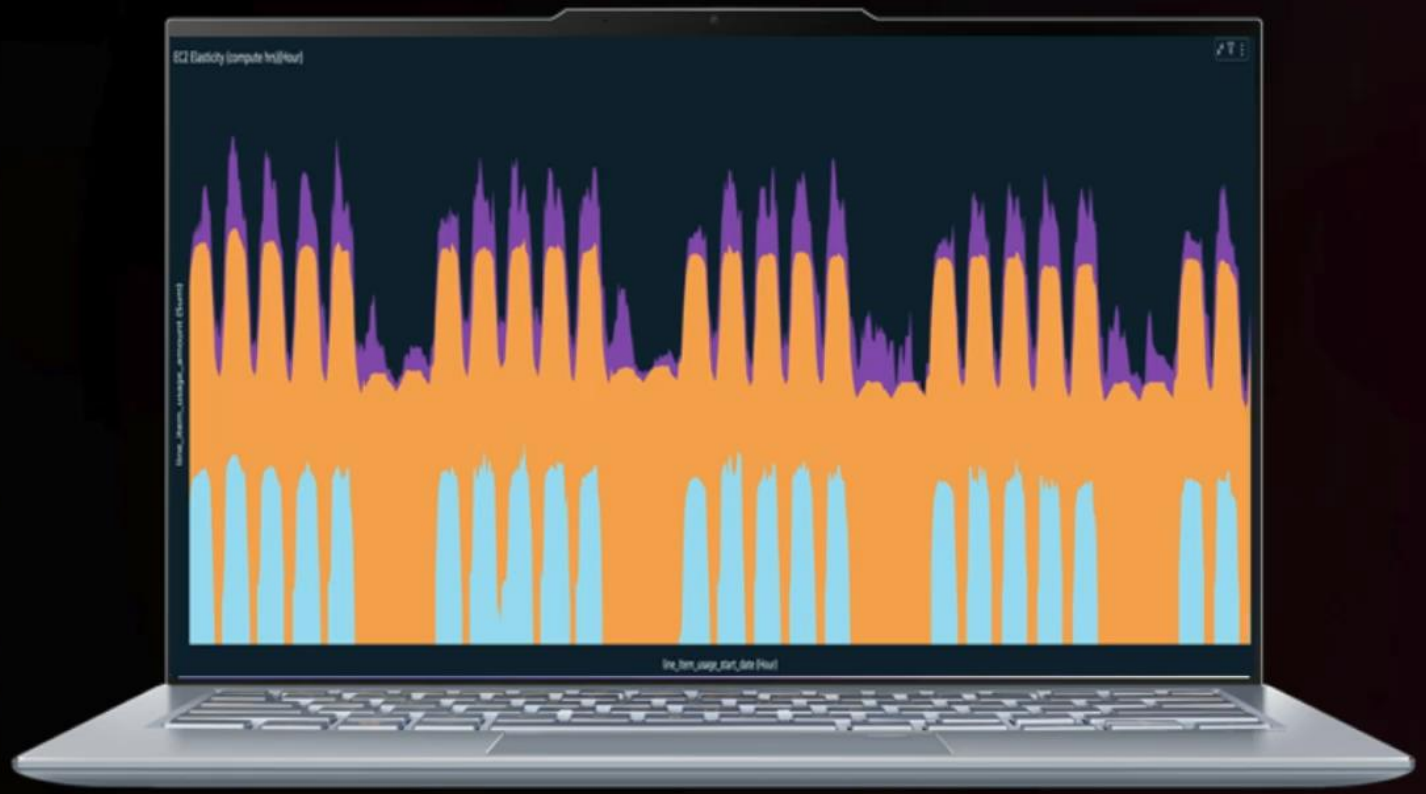
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Elasticity

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost Optimization at Standard Chartered Bank

On-Demand Savings Plan Spot Instances

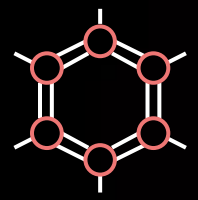


Consumption Model
Right Instance Choice
Elasticity
Tooling

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Auto Scaling groups

CONTAINS A COLLECTION OF AMAZON EC2 INSTANCES THAT ARE TREATED AS A LOGICAL GROUPING FOR THE PURPOSES OF AUTOMATIC SCALING AND MANAGEMENT

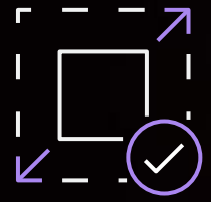


SIZE

Launch instances to meet desired capacity

Periodic health checks on instances

Maintain a fixed number of instances



SCALING POLICY

Manual, scheduled, predictive, on-demand

Dynamically change size of group to meet changing conditions

Size is determined by min/max capacity



PURCHASE/ARCH.

On-demand, Saving Plans and EC2 Spot

x86, arm64 or mixed architecture group



INTEGRATIONS

Lifecycle hooks

Instance health checks

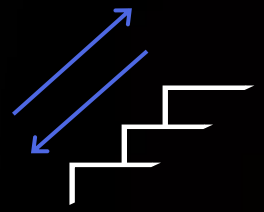
Load balancer integration

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Scaling policies with EC2 Auto Scaling

DYNAMIC SCALING

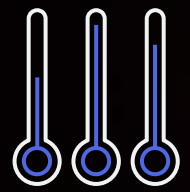


SIMPLE/STEP SCALING

Monitors metrics and adds/removes instances as per steps defined by the customers

Manually calculate capacity

Reactive in nature



TARGET TRACKING

Thermostat-like control mechanism that automatically adds or removes instances to maintain metrics at a customer defined target

Automated

Reactive in nature

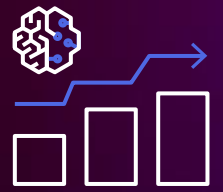


SCHEDULED SCALING

Launch/terminate instances as defined by customer on a schedule

Manually calculate capacity

Proactive in nature



PREDICTIVE SCALING

Proactively launch capacity based on historic trends

Automated

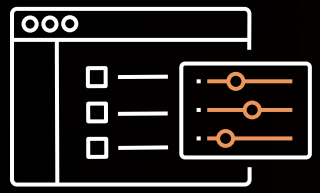
Proactive in nature

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

EC2 Auto Scaling: Warm Pools

A **pool** of pre-initialized EC2 instances that you can draw out of when your application need to scale out



Proactively initialize capacity



Reduces overprovisioning



Lower cost in stopped state

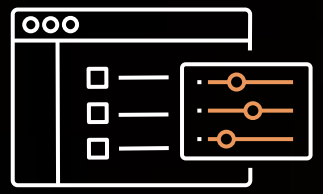


Adjusts automatically

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

EC2 Auto Scaling: Predictive scaling

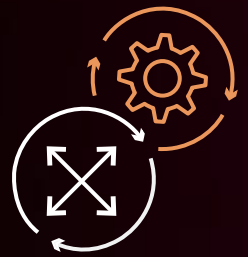
A **new** Amazon EC2 Auto Scaling policy that proactively launches Amazon EC2 instances in anticipation of predicted demand



Proactively manages capacity



Reduces overprovisioning



Works within defined boundaries



Adjusts automatically

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Tools and services that help you scale



Auto scaling groups





EKS Managed Node Groups

&

Kubernetes Cluster Autoscaler



Karpenter

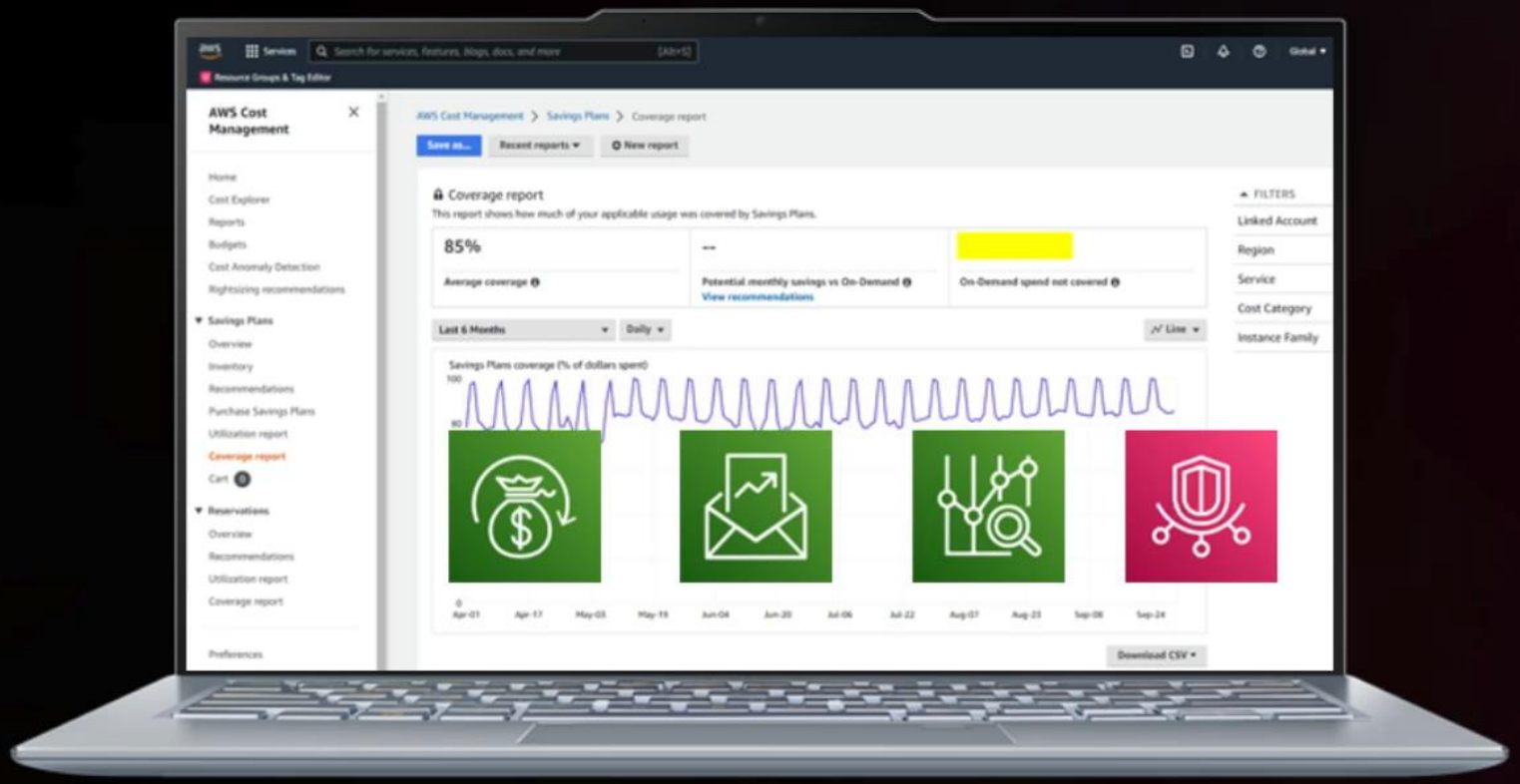
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Tooling

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

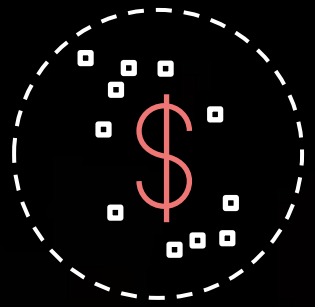
Cost Optimization at Standard Chartered Bank



Consumption Model
Right Instance Choice
Elasticity
Tooling

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost management and optimization



Cost Management

Ability to give visibility of cost to keep them under control and to enable cost optimization

- ✓ Consolidated Billing
- ✓ Cost Allocation Tags
- ✓ Cost Explorer & AWS Budget
- ✓ Cost Anomaly Detection
- ✓ Cloud Intelligence Dashboards
- ✓ Great 3rd Party Partner Ecosystem: CloudHealth, CludCheckr, Apptio, etc



Cost Optimization

Ability to run systems to deliver business value at the lowest possible price point

- ✓ Savings Plans & EC2 Spot Instances
- ✓ Auto Scaling
- ✓ Compute Optimizer & Trusted Advisor
- ✓ Automation: Instance Scheduler, Janitor Monkey
- ✓ Great 3rd Party Partner Ecosystem: GorillaStack, Skeddly, ParkMyCloud, etc

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Tagging Strategy

- Identify a set of tags using cross functional teams
- Define a tagging naming standard & dictionary
    - *env = dev | test | stg | prod*
- Publish the tagging dictionary
    - *Feedback + Announce + Refer*
- Define “rules of the game”
    - *Deployment automation*
- Enforce the rules
    - *Re-active enforcement*
- *Start small and evolve*

Monthly Activity

- *Tag Coverage %*
- *Un-allocatable spend/usage*
- *Tag dictionary review*

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Reporting and enforcement

	Proactive enforcement	Reactive enforcement	Reporting
Methods	Stakeholder approved tagging policy/dictionary	Resources with non-compliant tags are stopped and/or terminated	Awareness and gamification
AWS	AWS Organizations (Service Control Policy) AWS CloudFormation (tags as code) AWS Service Catalog (TagOptions) Amazon IAM (tag read/write controls)	AWS Tag Policies (supported services) AWS Config (required-tags rule) AWS auto-tagging solutions (link) Marketplace AMI cost allocation and tagging (link)	AWS Cost Categories AWS Cost Explorer AWS Budgets Amazon Quicksight (+ AWS CUR) AWS Resource Groups Amazon CloudWatch (tag alerts) AWS Config Dashboard
AWS Partners	Terraform (tags as code) Configuration Management (Puppet, Chef) AWS Cloud Management Competency Partners	AWS Cloud Management Competency Partners	AWS Cloud Management Competency Partners
Open source	GorillaStack (auto-tag)	Cloud Custodian GorillaStack (retro-tag) Graffiti Monkey (tag propagation) AWS Tagger (by Washington Post)	
DIY	Custom solutions via fill-in form integrated with AWS APIs	Custom enforcement solutions	Custom dashboards

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cloud Intelligence Dashboards

A series of in-depth and customizable dashboards that provide the most comprehensive cost and usage details to help you optimize cost, track usage goals, and achieve operational excellence.

Easy to use

- Self service
- Auto-generated recommendations
- Actionable insights for entire organization in a single place
- Multi-payer customer support

Secure

- No agents running
- Data stays in your organization
- All AWS native services

In Depth

- Resource level granularity
- Fully customizable
- ML-driven insights

Cost Efficient

- Serverless: Pay as you go and scale on demand
- No additional charges – pay only for underlying services Amazon S3, Athena, QuickSight
- Estimated cost in a large AWS deployment less than $100 per month



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What dashboard do I need?

Explore Cloud Intelligence Dashboards using the Public Demo

Deploy today using the Cloud Intelligence Well Architect Labs

Trends Dashboard



Executive summary

Cost Intelligence Dashboard



- Track success
- High level

CUDOS Dashboard



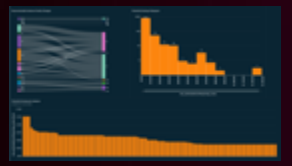
- In-depth insights
- Targeted recommendations

KPI Dashboard



- Define and track KPIs
- Take action

Compute Optimizer Dashboard



- Track wasted spend
- Find right-sizing opportunities

Trusted Advisor Organizational (TAO) Dashboard



Track security, performance, cost opt, and fault tolerance flags

Executives

Finance/Procurement

Business/Product Owners

FinOps Practitioners

Engineering and Operations

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Putting it all together

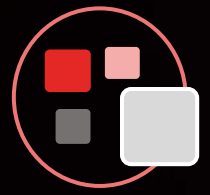


CONSUMPTION MODEL.

On-demand

Saving Plans & On-Demand Capacity Reservations

EC2 Spot

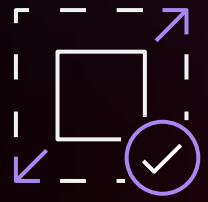


INSTANCE SELECTION

Intel, AMD, Graviton

Trusted Advisor & Compute Optimizer

Instance Flexibility & Attribute Based Instance Selection

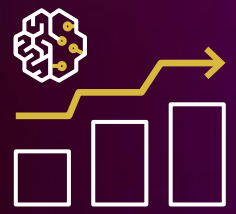


ELASTICITY

Auto Scaling group

Dynamic, Scheduled Predictive Scaling

Warm Pools & Instance Refresh



TOOLING

Cost Management & Cost Optimization

Tagging Strategy

Cloud Intelligence Dashboard



https://FinOps.org

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Daniel Kazarinow
Daniel.Kazarinow@sc.com

LinkedIn :

Carlos Manzanedo Rueda
ruecarlo@amazon.com

LinkedIn :



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.