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

PRT311

SPONSORED BY GRANULATE

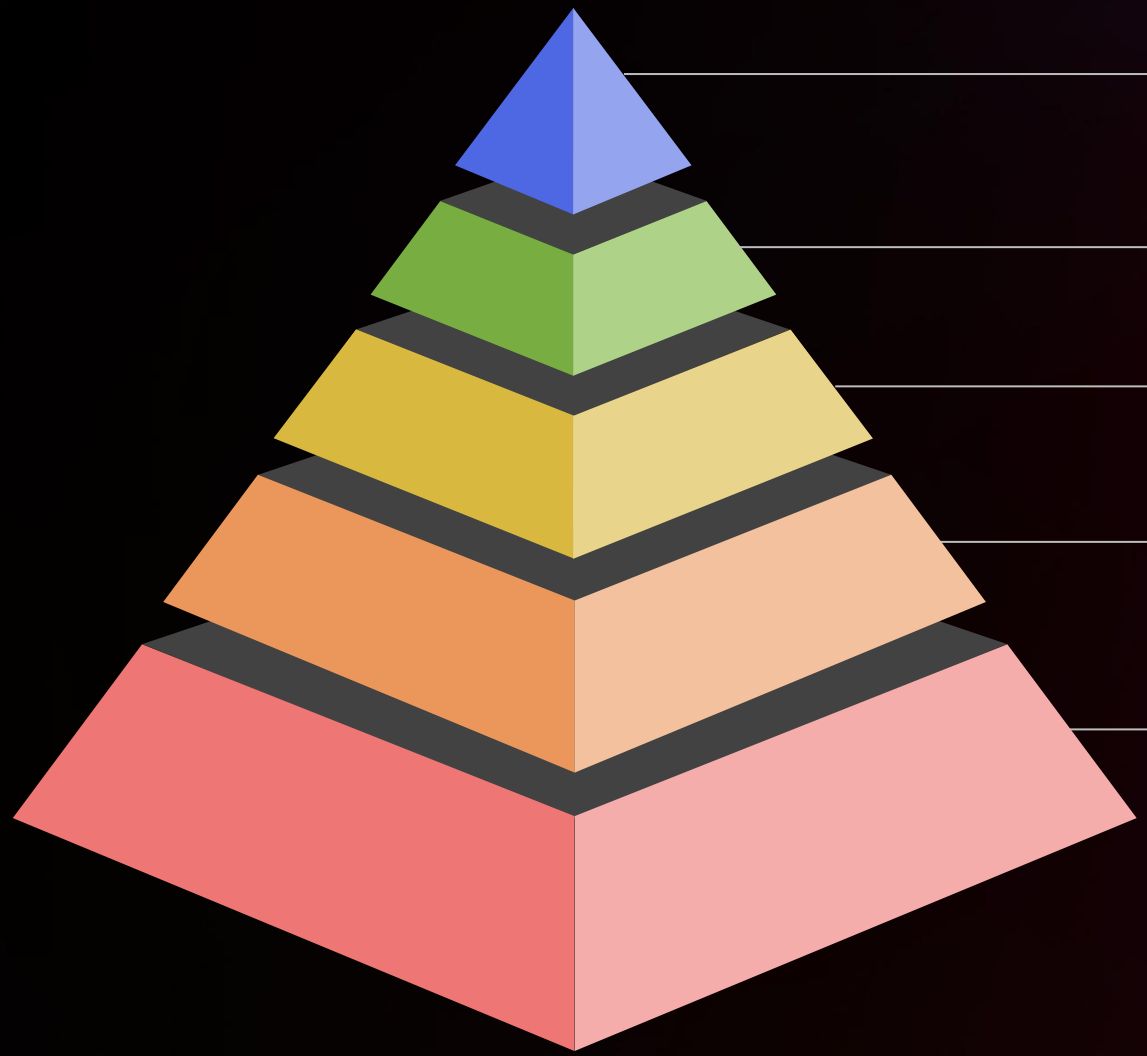
Integrating optimization into your cloud journey

Omer Mesika

Director of Solution Engineering
Granulate

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

Cloud App Hierarchy of Needs



5 **Self-actualization**
Optimizations - code, resources, tuning
Internal Training

4 **Self-esteem**
Reliability, quality maintenance, cost optimization

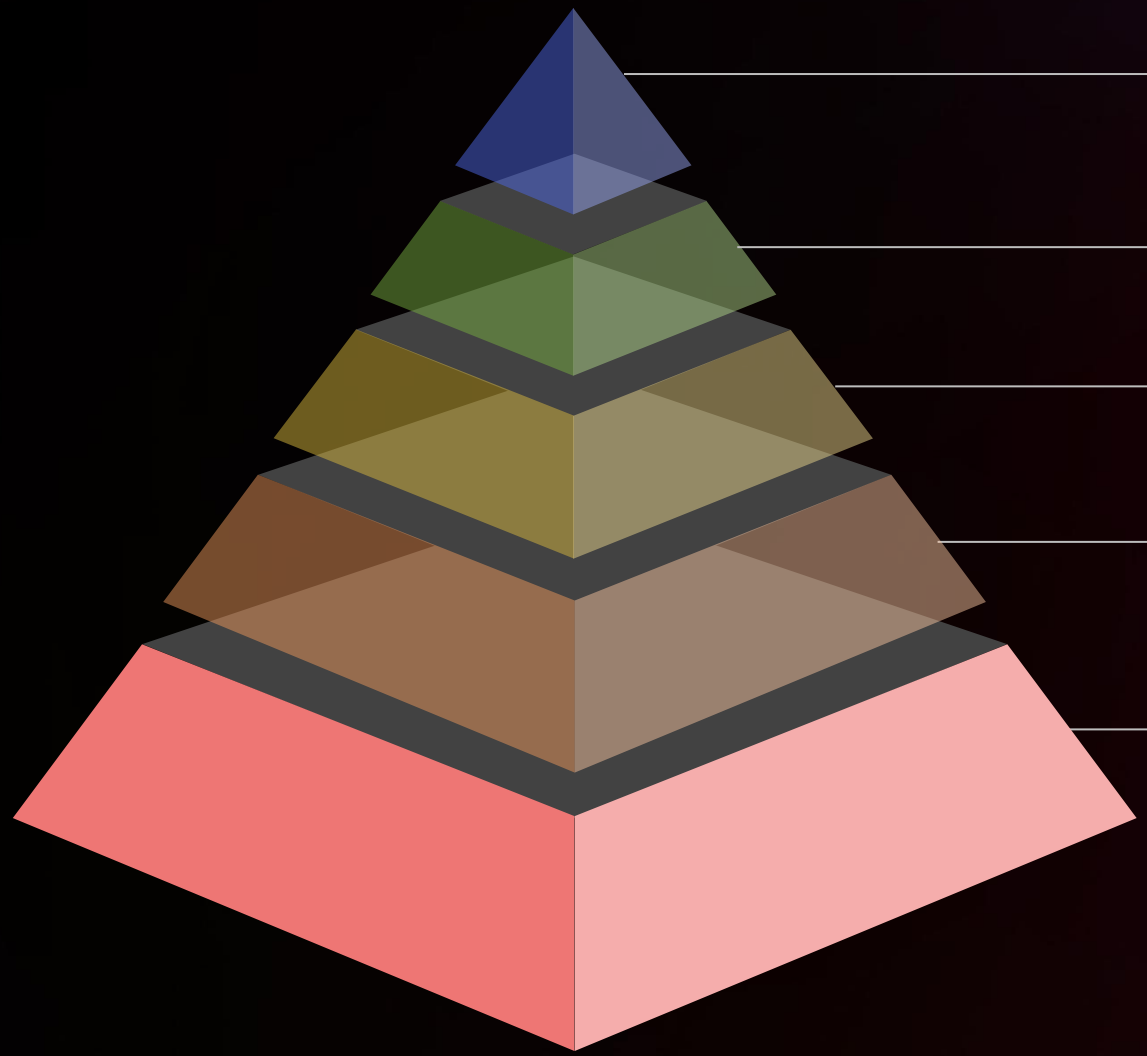
3 **Love and Belonging**
Product adoption, scaling up, rightsizing

2 **Safety and Performance Security**
Making sure the app runs smoothly with all its core functions, Discounts

1 **Physiological**
Setting the infrastructure, getting the app to run

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

Cloud App Hierarchy of Needs



5 **Self-actualization**
Optimizations - code, resources, tuning
Internal Training

4 **Self-esteem**
Reliability, quality maintenance, cost optimization

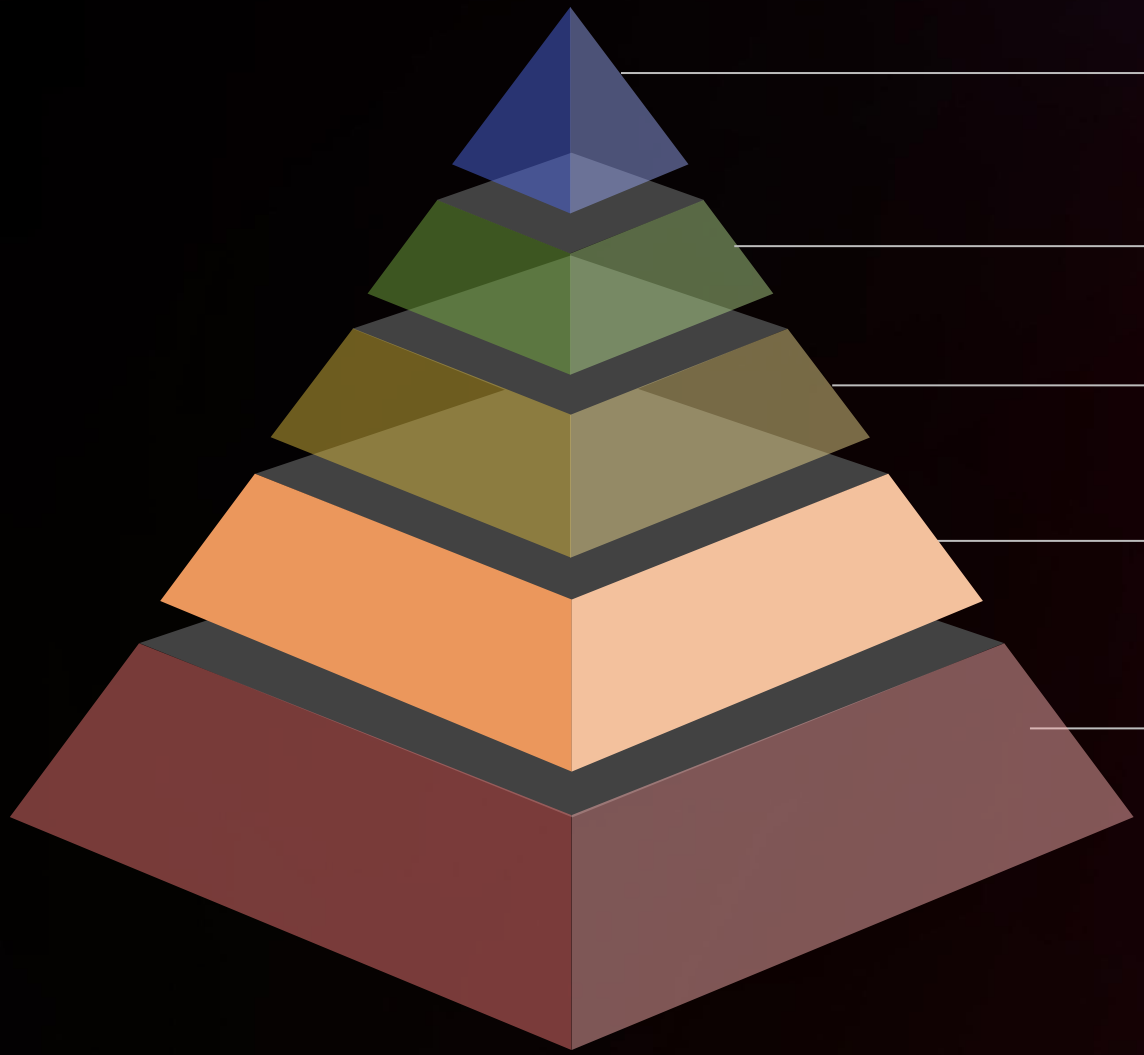
3 **Love and Belonging**
Product adoption, scaling up

2 **Safety and Security**
Making sure the app runs smoothly with all its core functions

1 **Physiological**
Setting the infrastructure, getting the app to run

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

Cloud App Hierarchy of Needs



5 **Self-actualization**
Optimizations - code, resources, tuning
Internal Training

4 **Self-esteem**
Reliability, quality maintenance, cost optimization

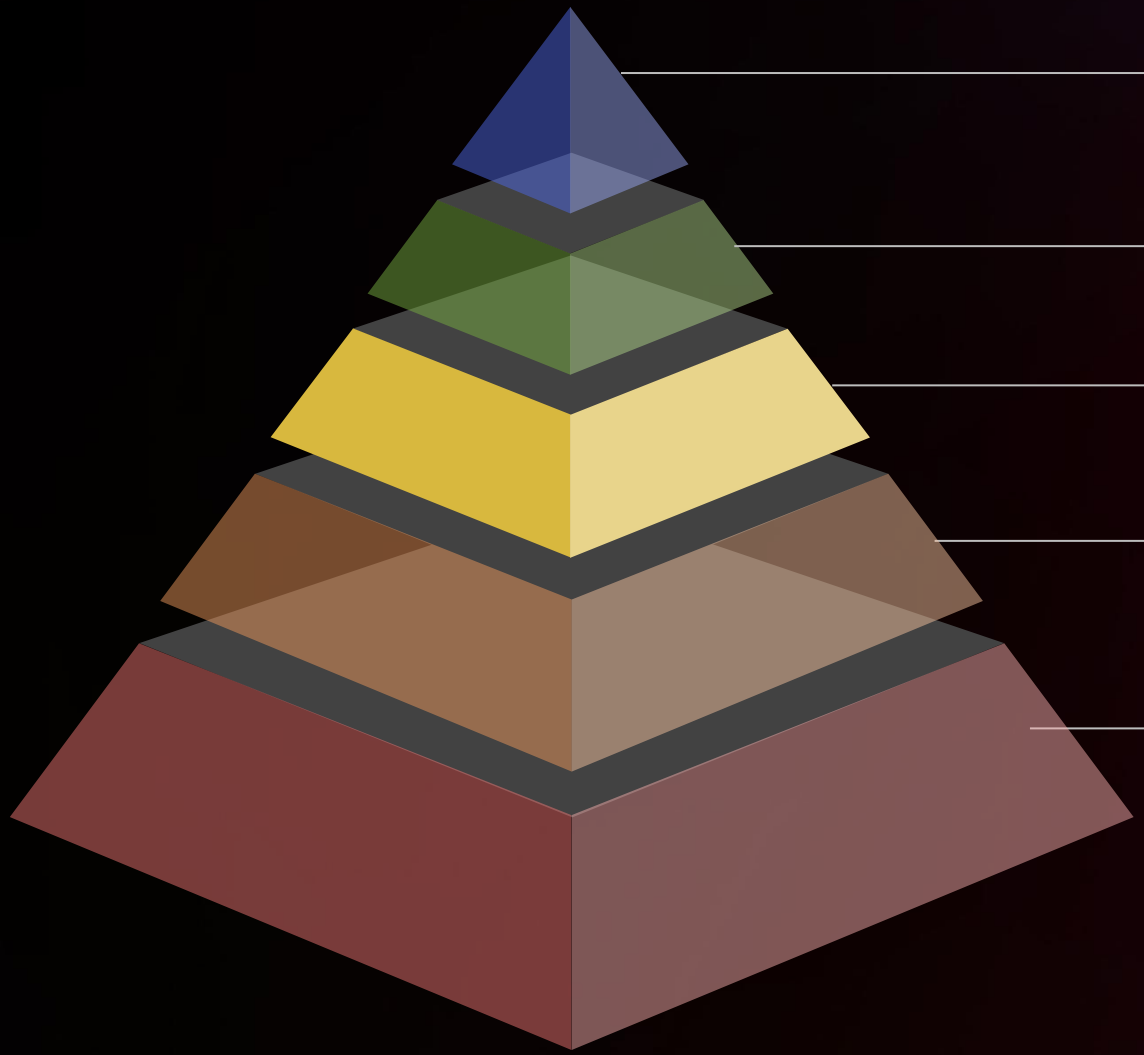
3 **Love and Belonging**
Product adoption, scaling up

2 **Safety and Security**
Making sure the app runs smoothly with all its core functions

1 **Physiological**
Setting the infrastructure, getting the app to run

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

Cloud App Hierarchy of Needs



5 **Self-actualization**
Optimizations - code, resources, tuning
Internal Training

4 **Self-esteem**
Reliability, quality maintenance, cost optimization

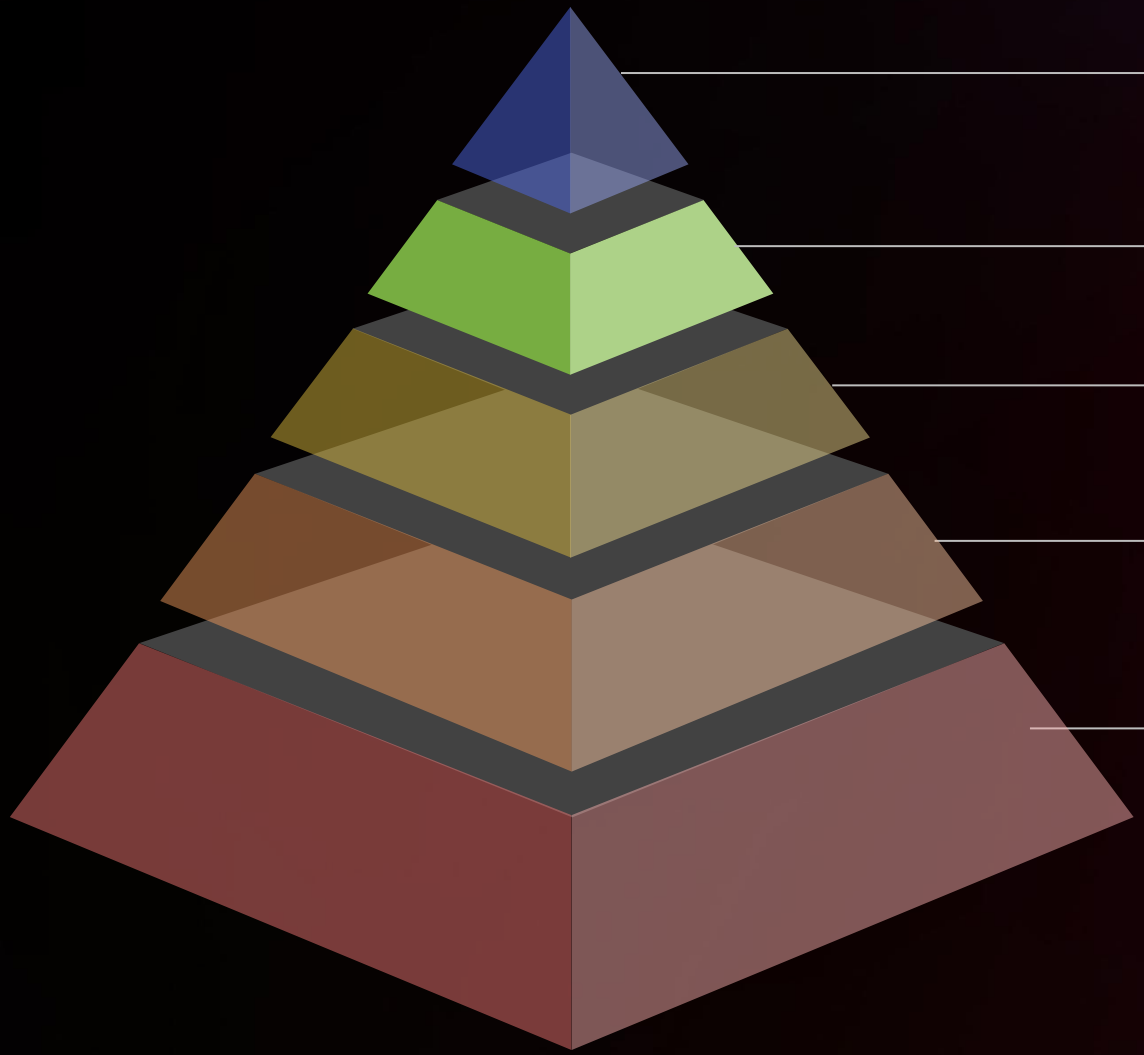
3 **Love and Belonging**
Product adoption, scaling up

2 **Safety and Security**
Making sure the app runs smoothly with all its core functions

1 **Physiological**
Setting the infrastructure, getting the app to run

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

Cloud App Hierarchy of Needs



5 **Self-actualization**
Optimizations - code, resources, tuning
Internal Training

4 **Self-esteem**
Reliability, quality maintenance, cost optimization

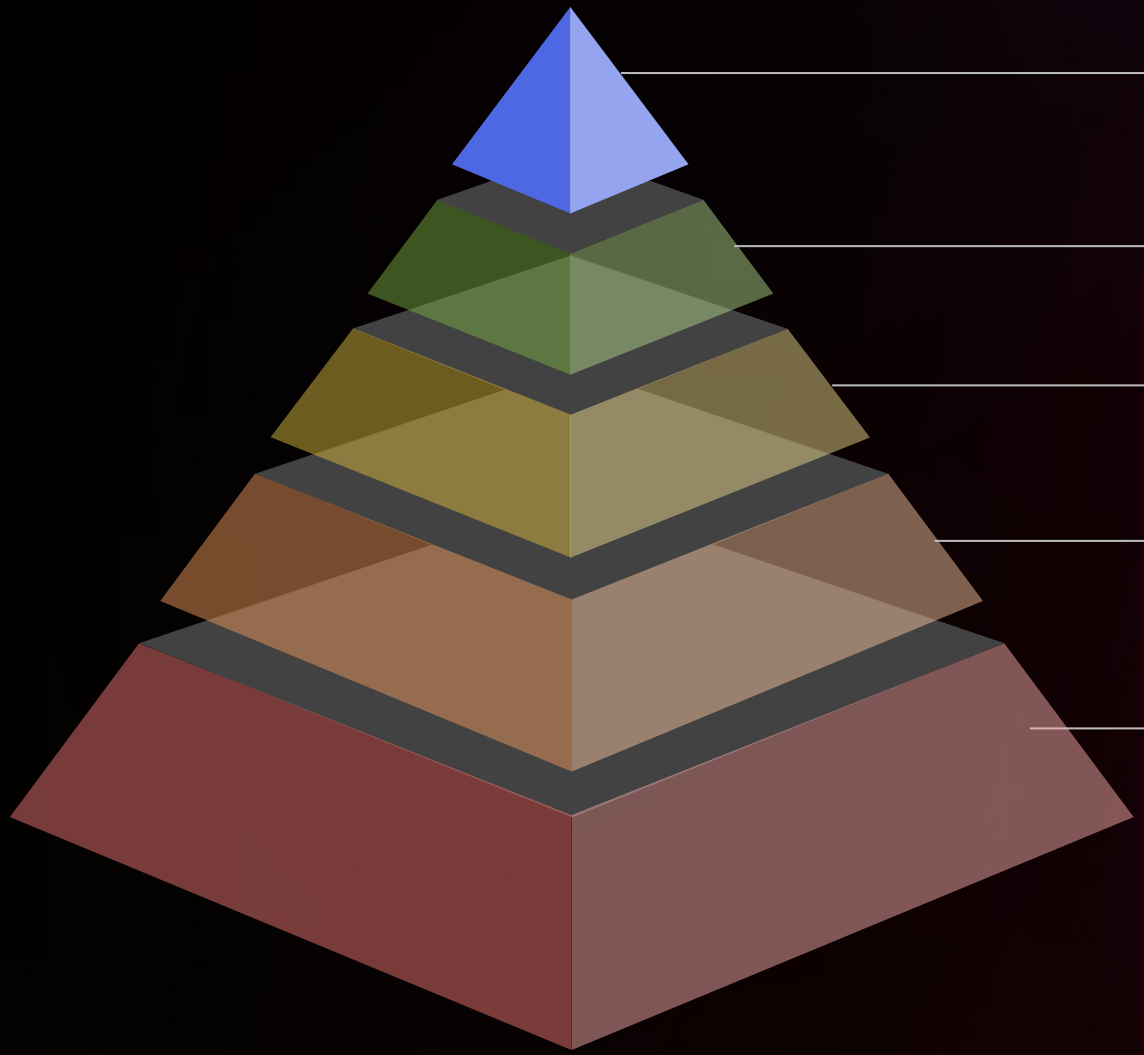
3 **Love and Belonging**
Product adoption, scaling up

2 **Safety and Security**
Making sure the app runs smoothly with all its core functions

1 **Physiological**
Setting the infrastructure, getting the app to run

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

Cloud App Hierarchy of Needs



5 **Self-actualization**
Autonomous, continuous optimization

4 **Self-esteem**
Reliability, quality maintenance, cost optimization

3 **Love and Belonging**
Product adoption, scaling up

2 **Safety and Security**
Making sure the app runs smoothly with all its core functions

1 **Physiological**
Setting the infrastructure, getting the app to run

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

Optimization Process, Methodology, and Practices

PHASE 1

Visibility

- Monitoring/APM
- Cloud cost tracking
- Actionable Alerts

PHASE 2

Awareness

- Savings plan
- Reserved instances
- Spot instances
- AIOps
- Best Practices

PHASE 3

Management

- Rightsizing
- Orchestration
- Schedule downtime
- Cheaper machine type
- Build a FinOps team/ finance analyst with DevOps

PHASE 4

Optimization

- Contract negotiations
- Workload migration
- Configuration Tuning
- Cose rework

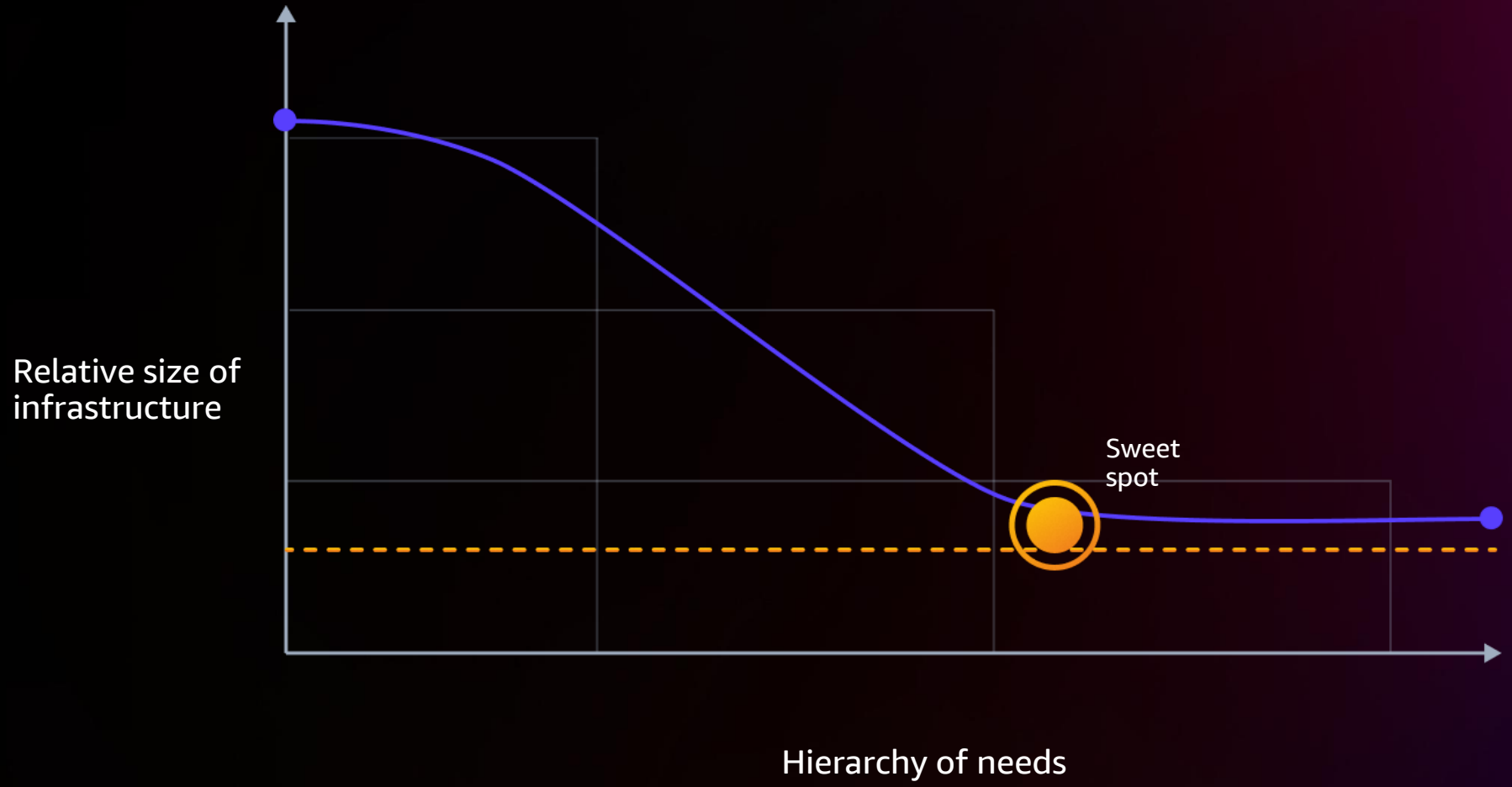
Low Effort - On the Cloud

High Effort - On & Off the cloud

aws

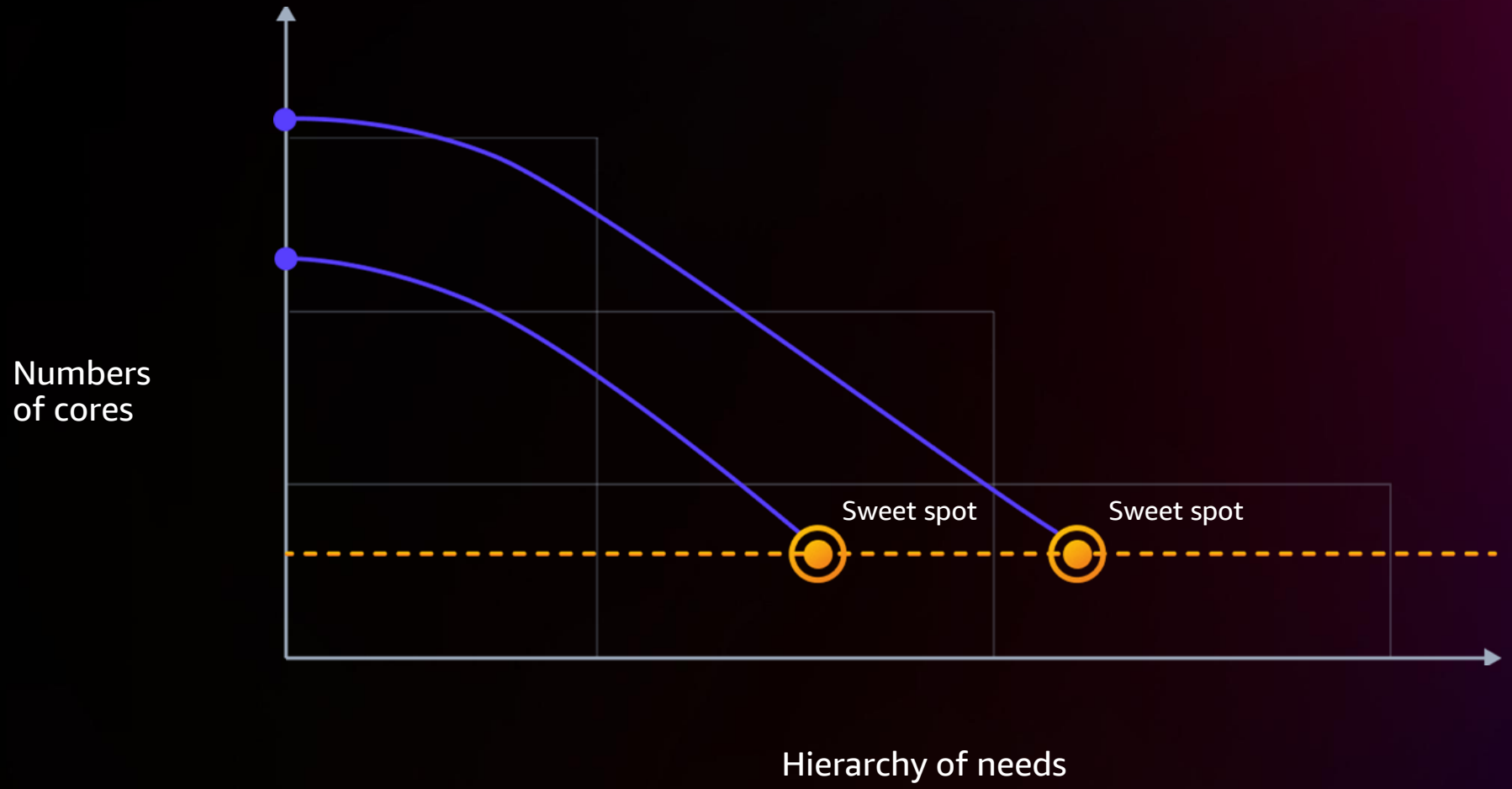
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Where Does Cost Fit into the Pyramid?



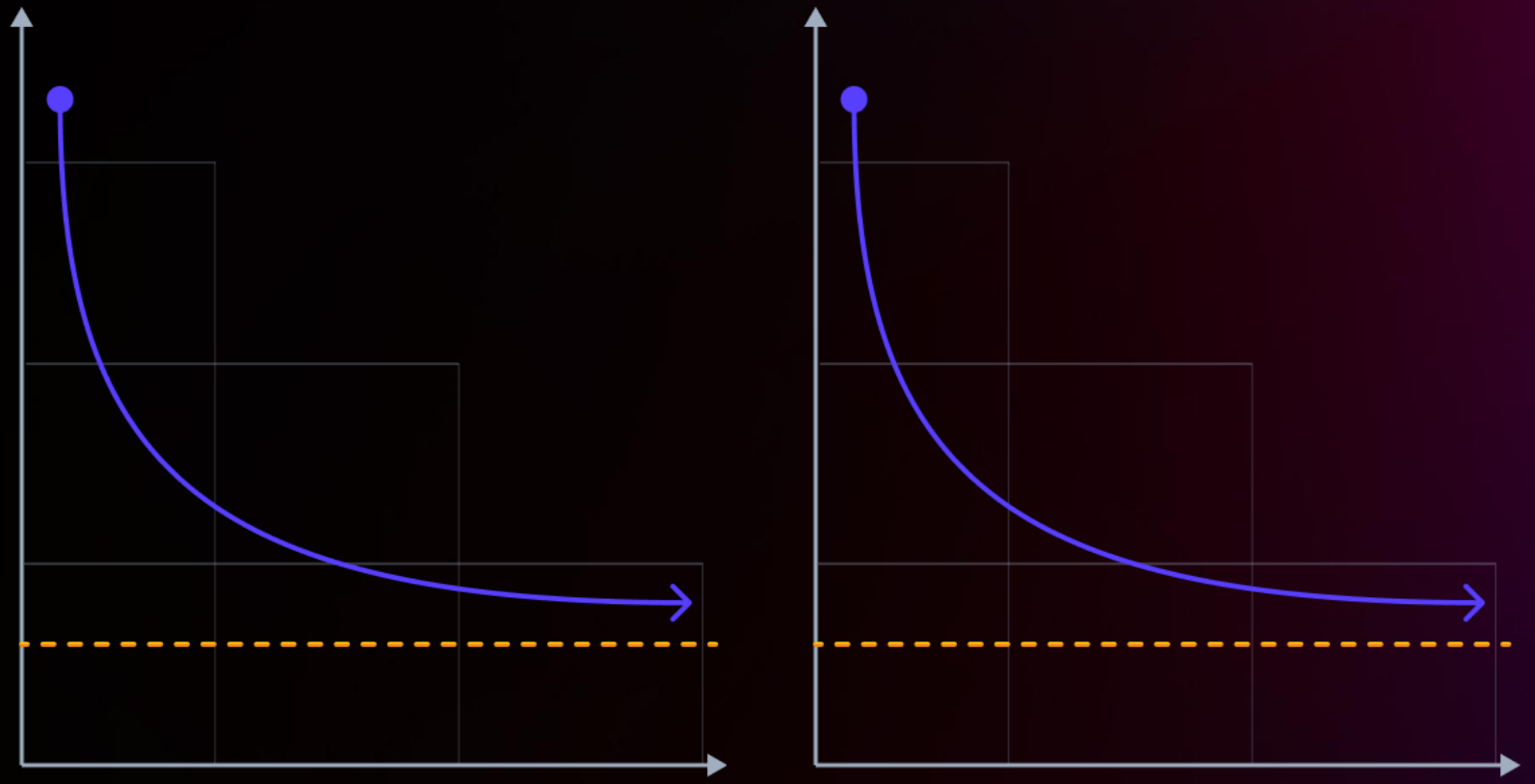
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Startups vs. Enterprises Optimization Optimum



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

Migration



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

You Are Here: Saturation



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

Challenges When Hitting Saturation



Further optimization



Resource exhaustion



Diminishing returns

GRANULATE

- Additive to all cost tools
- Little engineering time
- Starts with better performance

- Improves CPU utilization
- Reduces latency
- Up to 63% reduction in costs

aws

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

Autonomous Optimization

1. Application Profiling & Analysis
Granulate automatically learns your application specific resource usage patterns and data flow



3. Improve Performance
Granulate reduces response time and increases throughput of every machine

2. Continuously Optimize
Granulate tailors OS-level scheduling and prioritization decisions regarding CPU, network, locks, and memory accesses

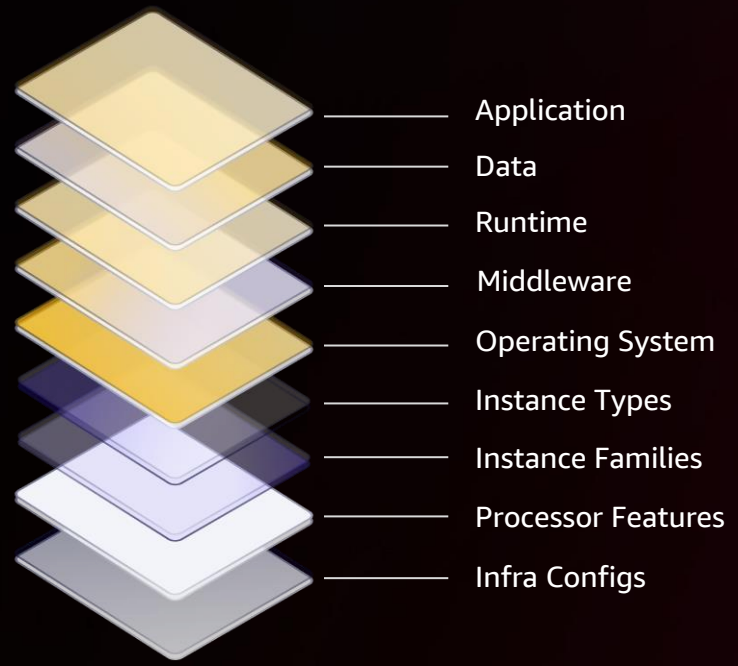
4. Automatic Capacity Reduction
Leverage improved performance to reduce cluster size and downsize compute resources

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

Full Stack Optimization Across All Stages

Continuous profiler for production – across any environment

gProfiler



GRANULATE

Application Performance Optimization

gMaestro
Powered By Granulate

Cloud Infrastructure Optimization

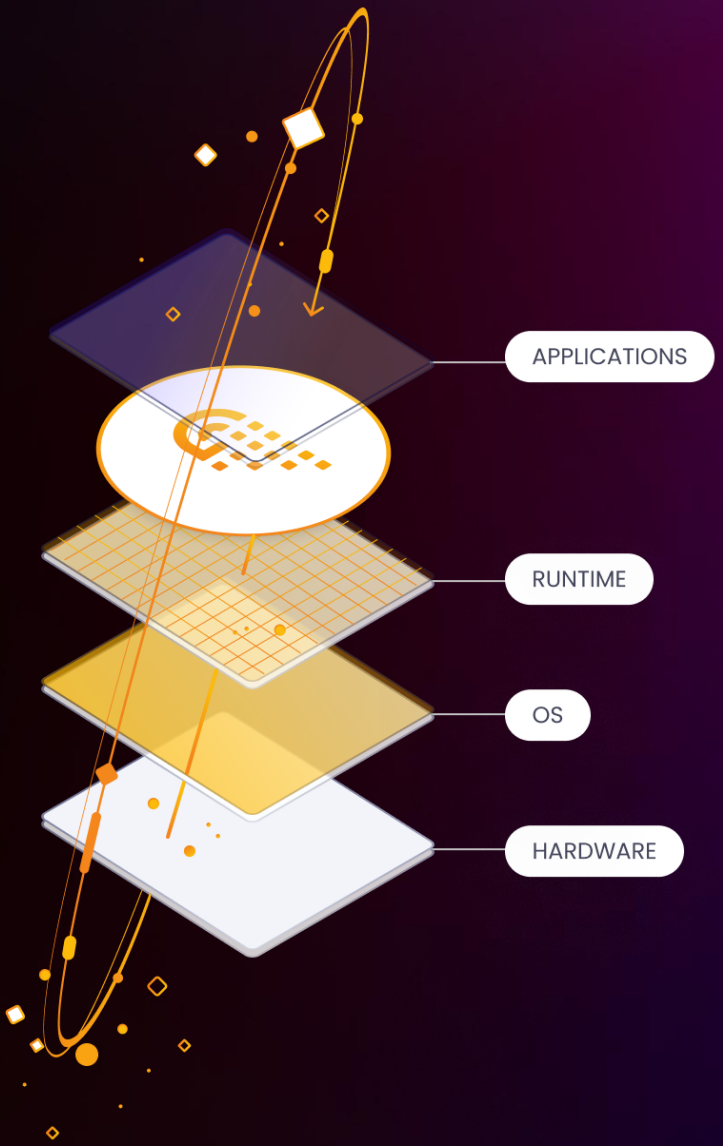
aws

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

The Technology

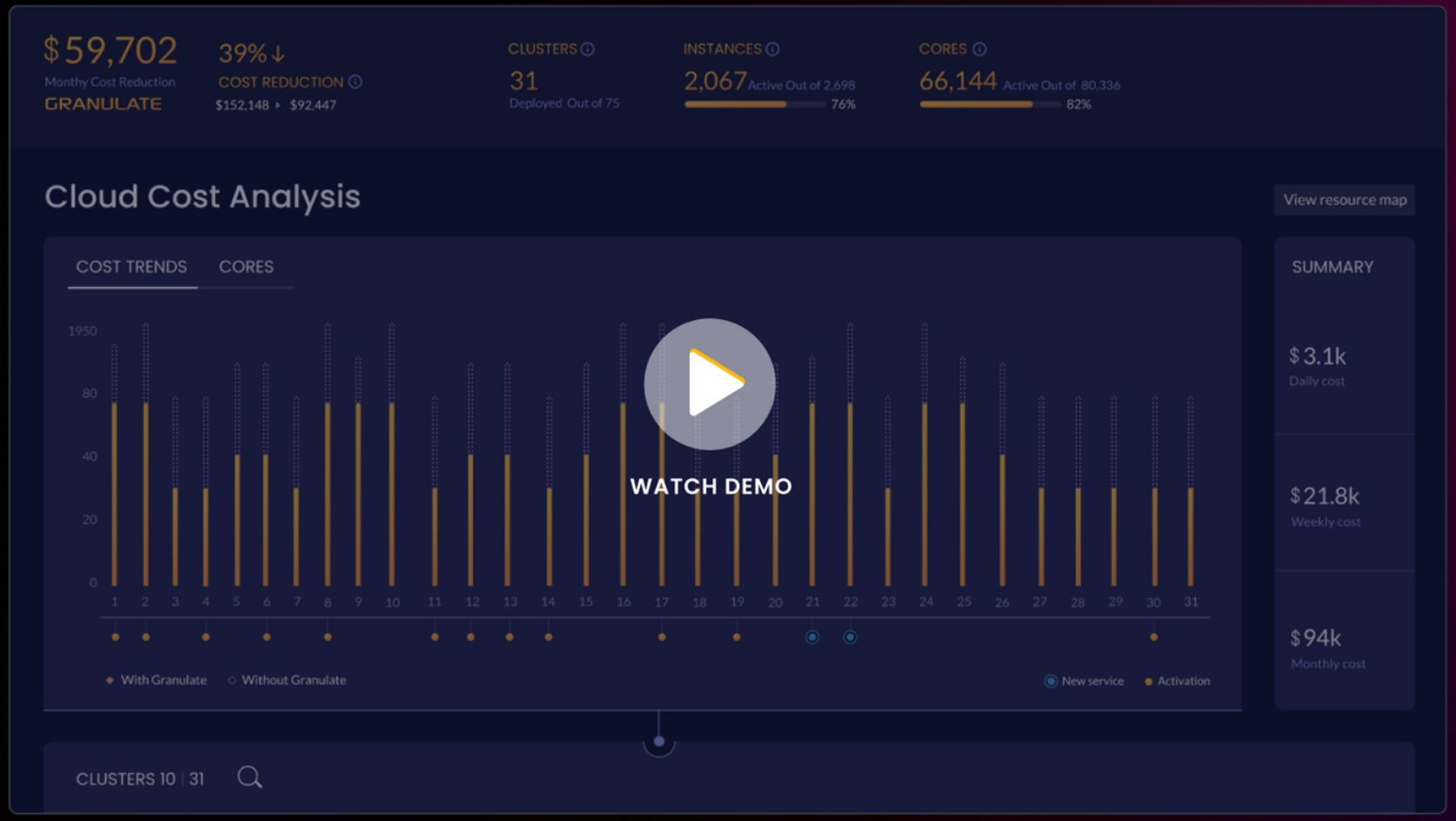
Granulate agents are optimizing performance continuously by adapting to the application workload patterns

Runtime Module – OS resource management optimization from within runtime environments



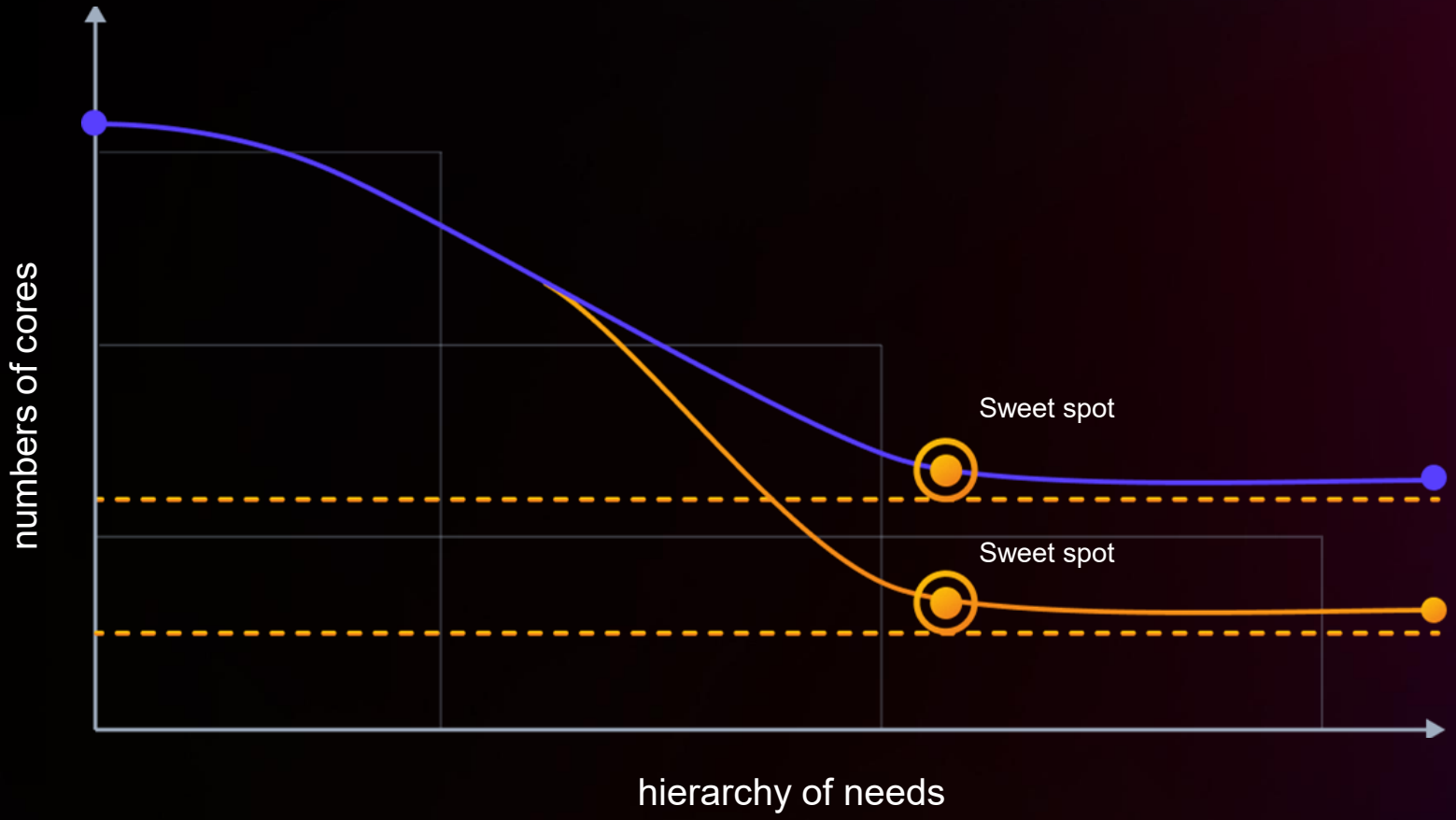
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Let’s See It in Action



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

Granulate Activates a New Optimum



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

Is This Relevant to Me?

- Infrastructure Agnostic - K8S, Big-Data, Dynamic applications, Hybrid
- Autoscaling by applicative metrics
- Different stages of the hierarchy
- Deployment process
- Spikey load

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

Retail and Ecommerce

About

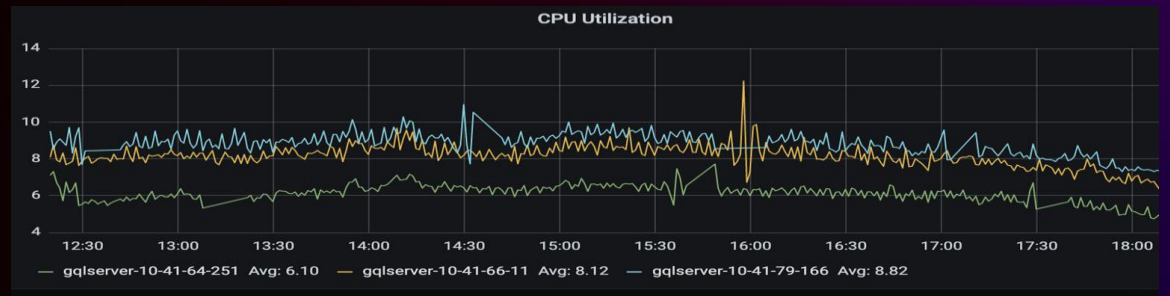
Industry: **Retail**
Employees: **5,000+**
Cloud: **AWS**
Use Case : **Golang**

Customer Leverages Intel Workload Optimizer To Achieve **Over 30% Cost Reduction**

30% Break of saturation

Reducing costs and achieving improved performance in less than 10 days with no code changes or efforts





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

Mobility and Autonomous Driving

About

Industry: **Mobility**
Employees: **1,500+**
Size: **150,000 Cores**
Cloud: **AWS**
Use Case: **Big Data**

Customer Sees **45% lower job completion time, to enable 45% cost reduction** with Granulate

Groundbreaking autonomous workload optimization yields value in days, leading to faster development and significant cost reduction

45% Cost reduction

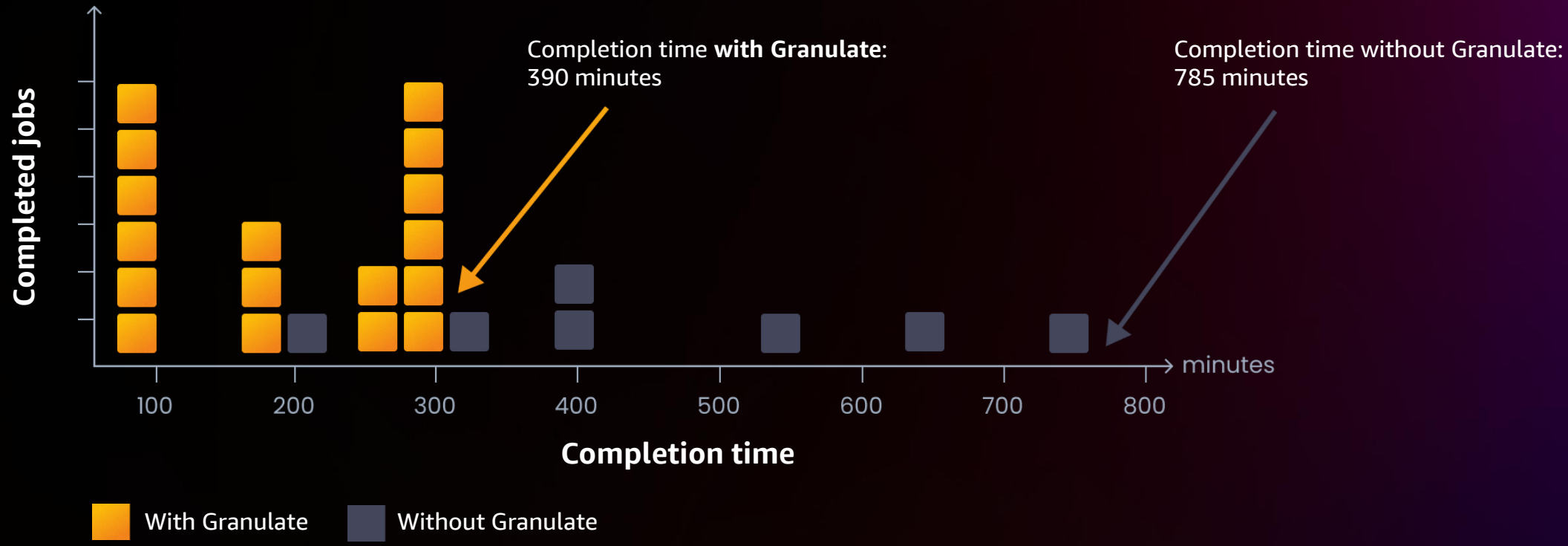
45% Job completion time reduction

aws

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

Mobility and Autonomous Driving

Massive Reduction In Average Job Completion Time and Costs In Less Than 21 Days



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

Online Gaming

About

Industry: **Gaming**
Employees: **800+**
Cloud: **AWS**
Use Case: **Java, Node.js**

Customer Leverages Granulate To Improve Performance and Reduce AWS Costs

Customer halts migration, leveraging IWO to improve cost-efficiency

40% Cost reduction

50% Job completion time reduction

40% Throughput Increase

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

Online Gaming

About

Industry: **Gaming**
Employees: **800+**
Cloud: **AWS**
Use Case: **Java, Node.js**

Kafka Consumers Reduced CPU Utilization and Increased Throughput

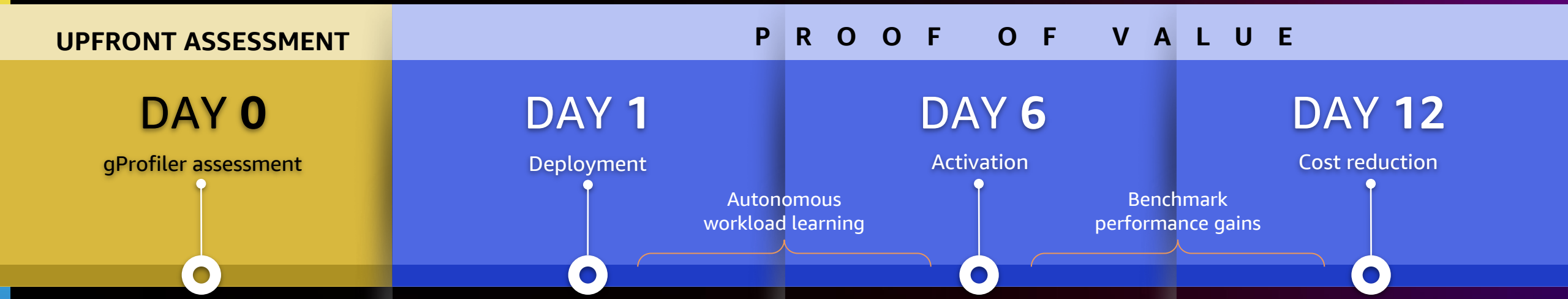


50% Reduction In CPU Utilization

40% Job completion time reduction

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

Customer Journey



30-minute DevOps

Run Granulate open-source gProfiler on several VMs

Get upfront analysis of expected performance improvement and cost reduction

1-hour DevOps

Granulate deployment on selected workloads

No code changes

1-hour meeting to review

Immediate performance improvements upon activation

Capacity management

Use performance improvements to downsize clusters count and reduce cost

NO ENGINEERING RESOURCES REQUIRED

Complimentary implementation and benchmark support included

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

Key Takeaways

- Understand your pyramid position
- Evaluate your current sweet spot
- Strive for integrate automated optimization to your process
- Save costs and improve your application performance
- Improve your sweet spot

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

Thank you!



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

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