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

PEX310

Optimizing your multi-tenant SaaS architecture

Tod Golding

Senior Principal Solutions Architect,
AWS SaaS Factory
AWS

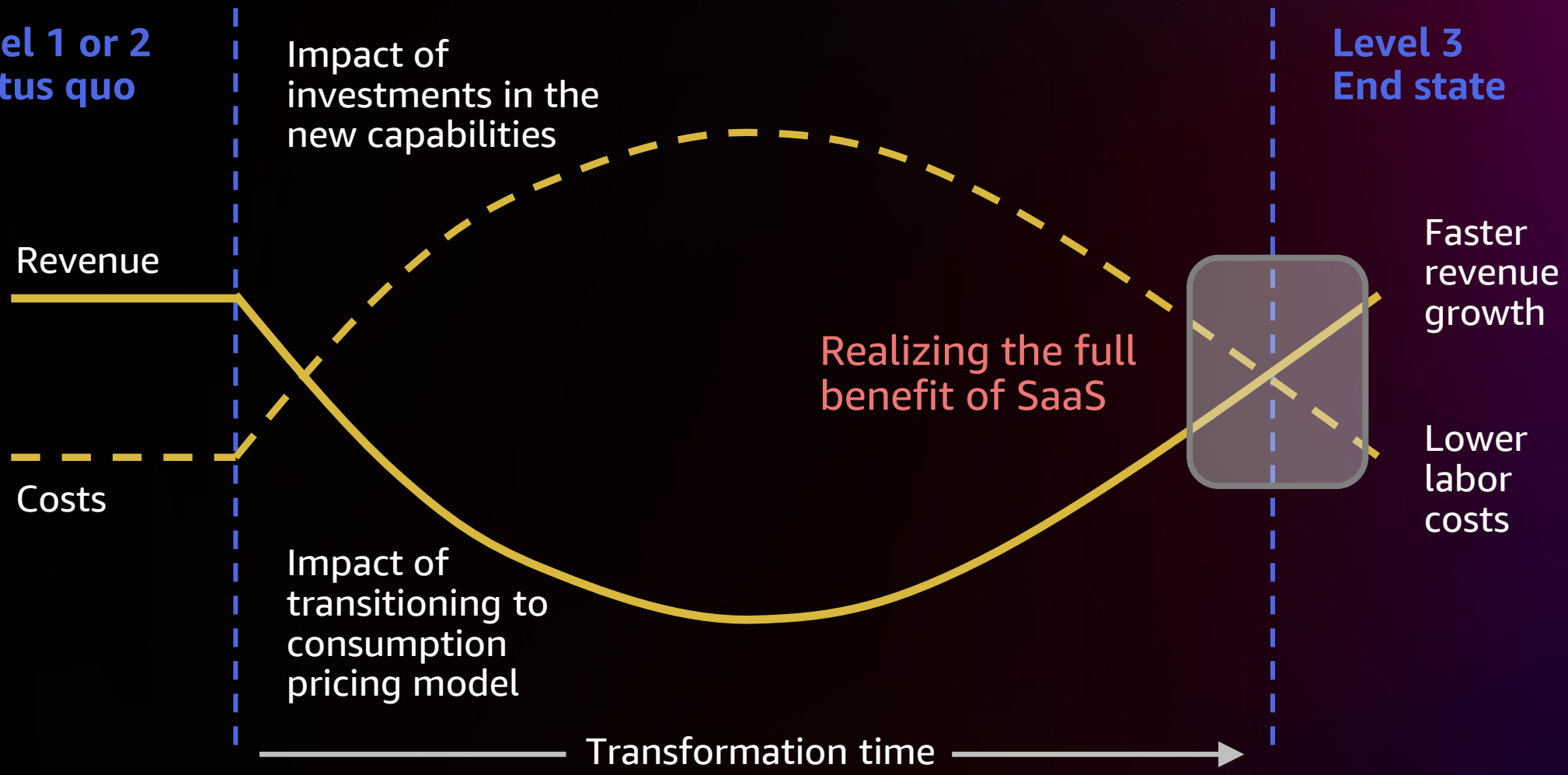
Peter Yang

Senior Partner Solutions Architect,
AWS SaaS Factory
AWS

aws

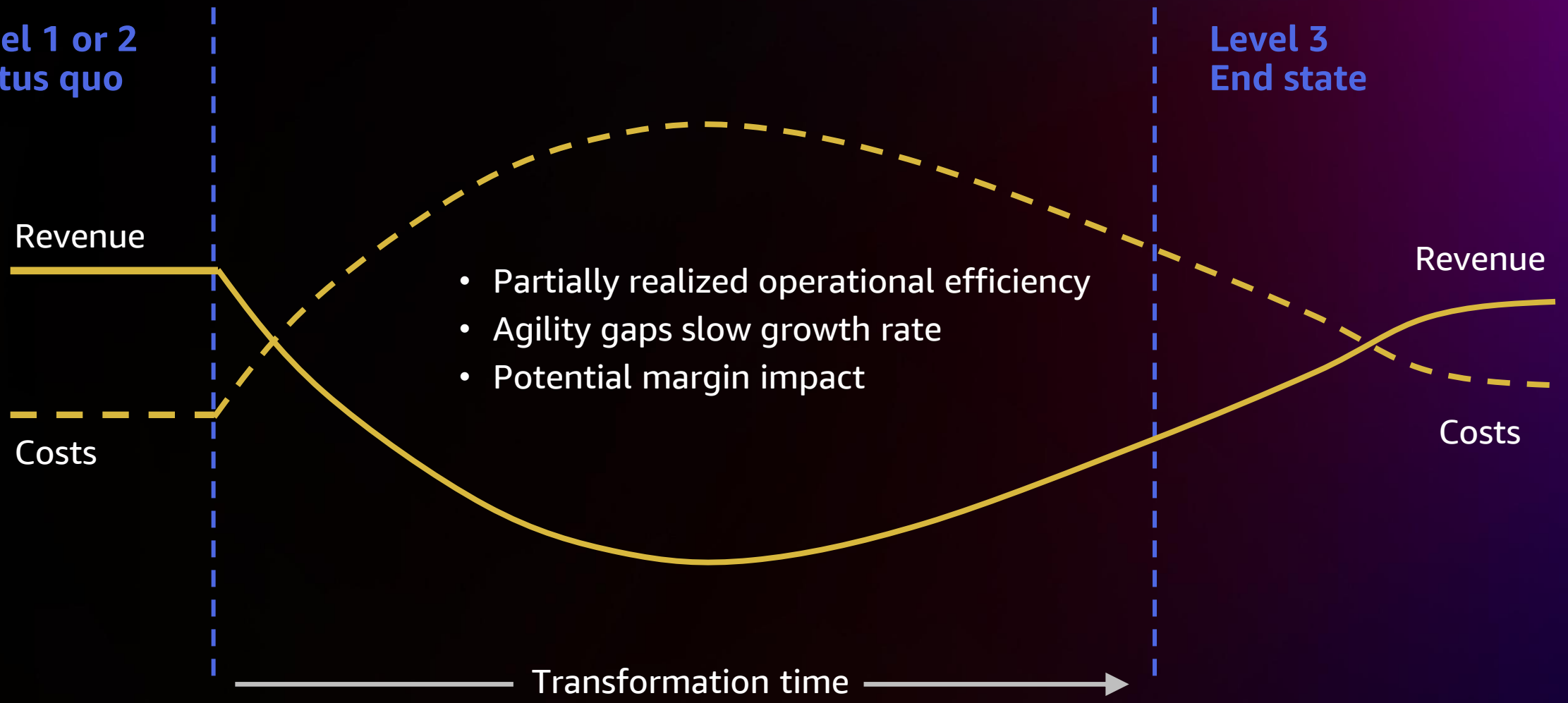
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The ideal SaaS journey



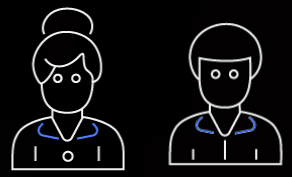
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The reality for some organizations



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

Use metrics to drive optimization



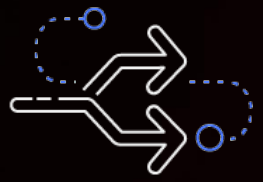
Profiling
tenant experience



Analyzing
tenant costs



Assessing
operational efficiency



Evaluating
organizational agility

- Use data to find refinement opportunities
- Measure the efficiency of your environment
- Examine how your architecture is supporting business goals
- You need to know what you don’t know

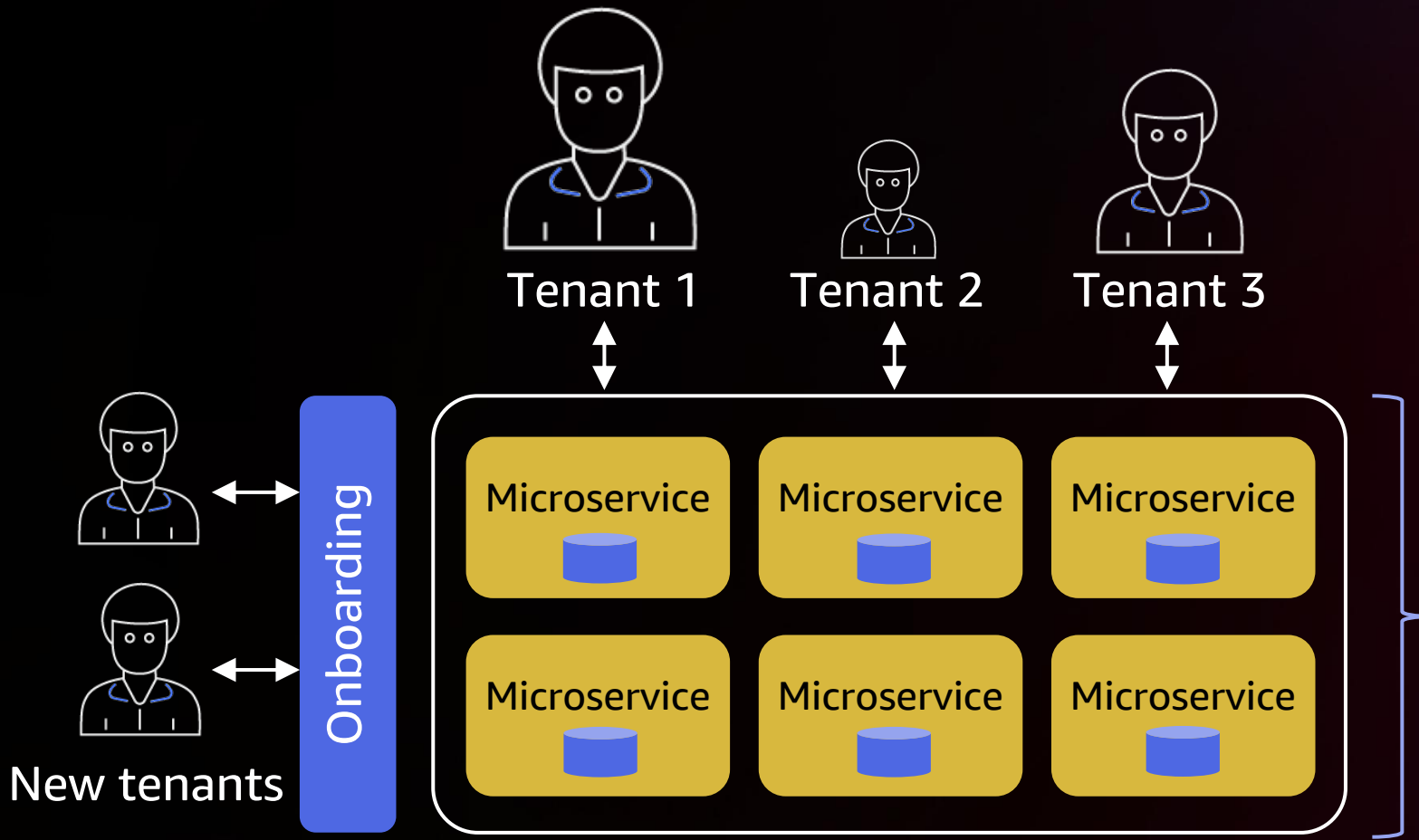
aws

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

Think small to think big

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

Supporting a diverse range of experiences

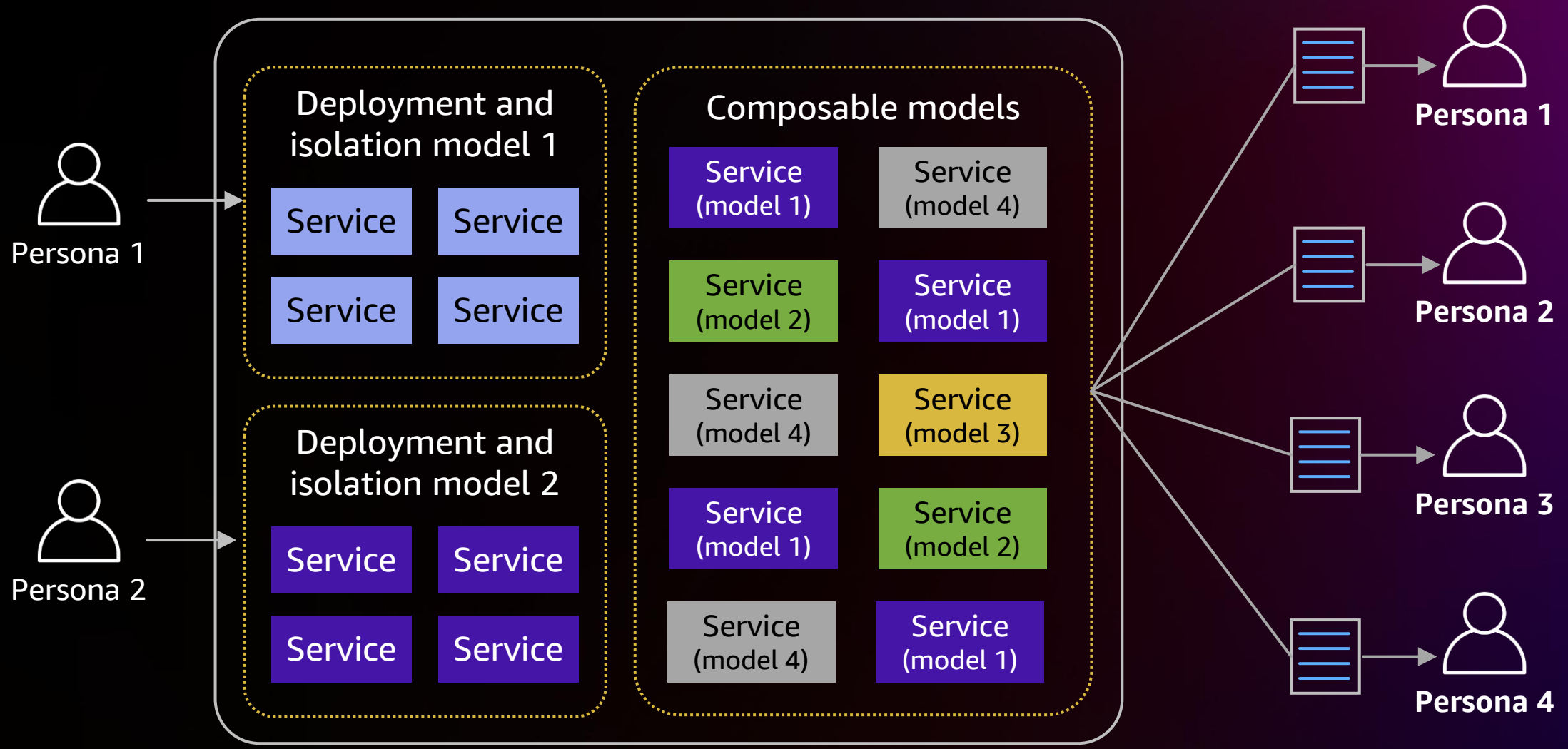


- Multiple tiers/personas
- Shared resources
- Noisy neighbor
- Unpredictable activity patterns
- Varying tenant profiles (isolation)
- Challenging scaling policies
- Greater sensitivity to the scope of impact

A one-size-fits-all footprint limits your options

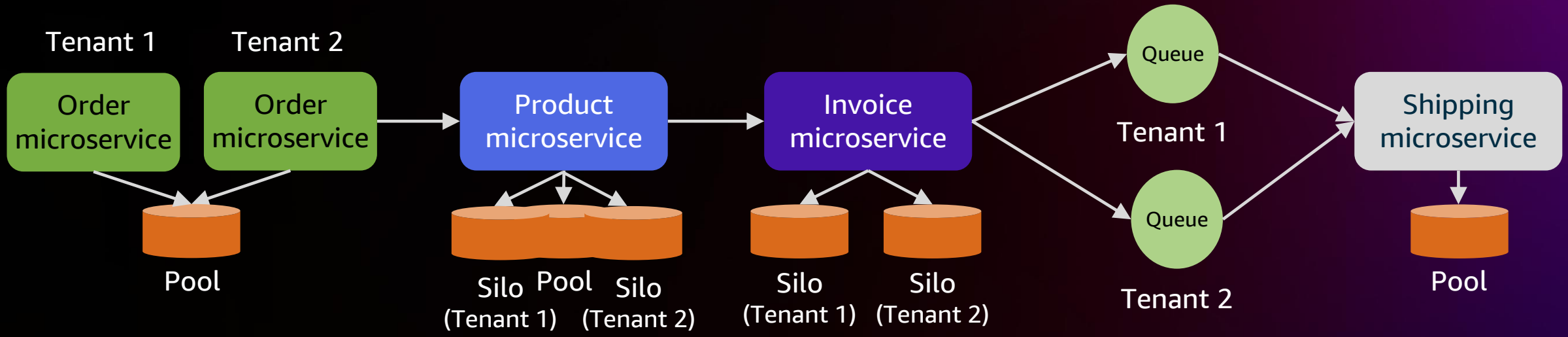
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Focus on flexibility and composability



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

Find the right mix of scale, cost, and experience



Siloed compute and pooled storage

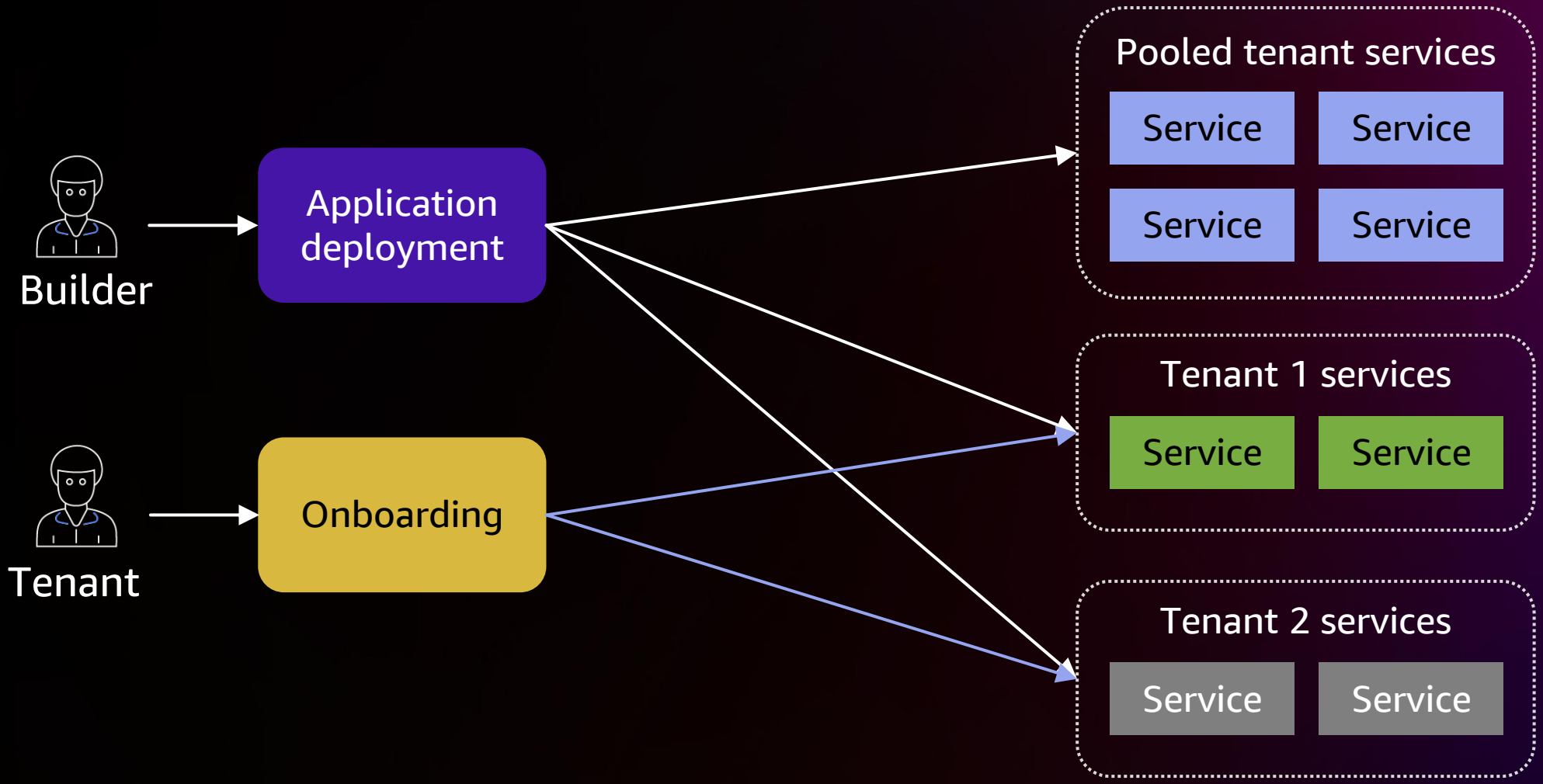
Pooled compute and siloed storage

Pooled compute and siloed storage

Pooled compute and storage consuming siloed queues

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

Trading deployment complexity for flexibility

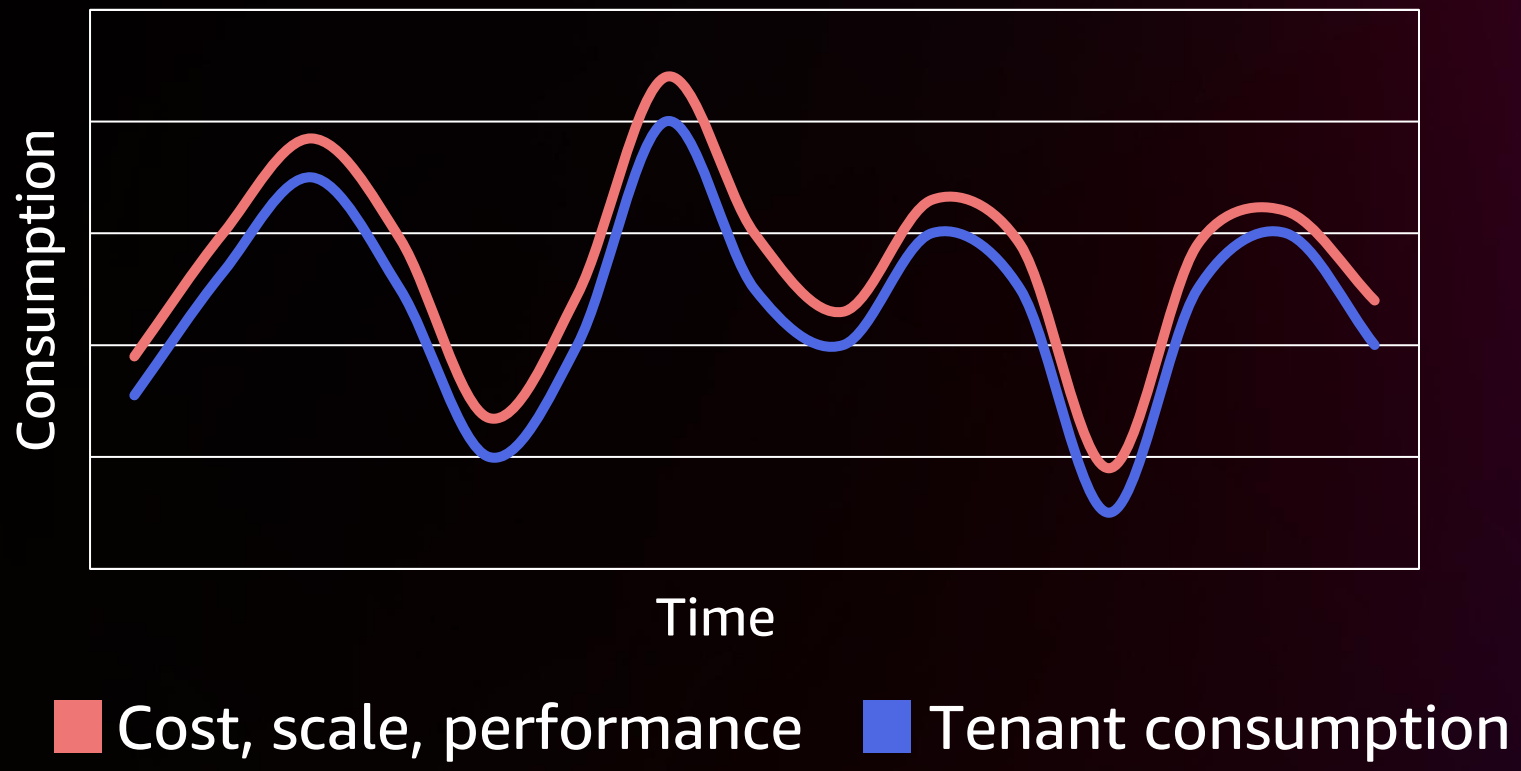


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Aligning consumption with activity

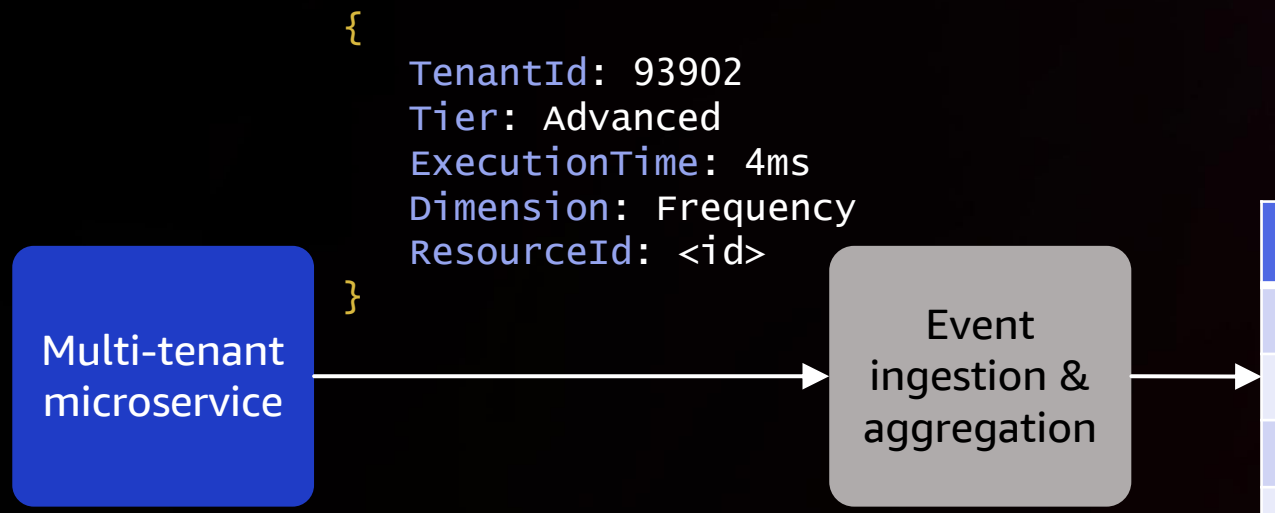
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The SaaS optimization dream



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

The foundation: Consumption metrics



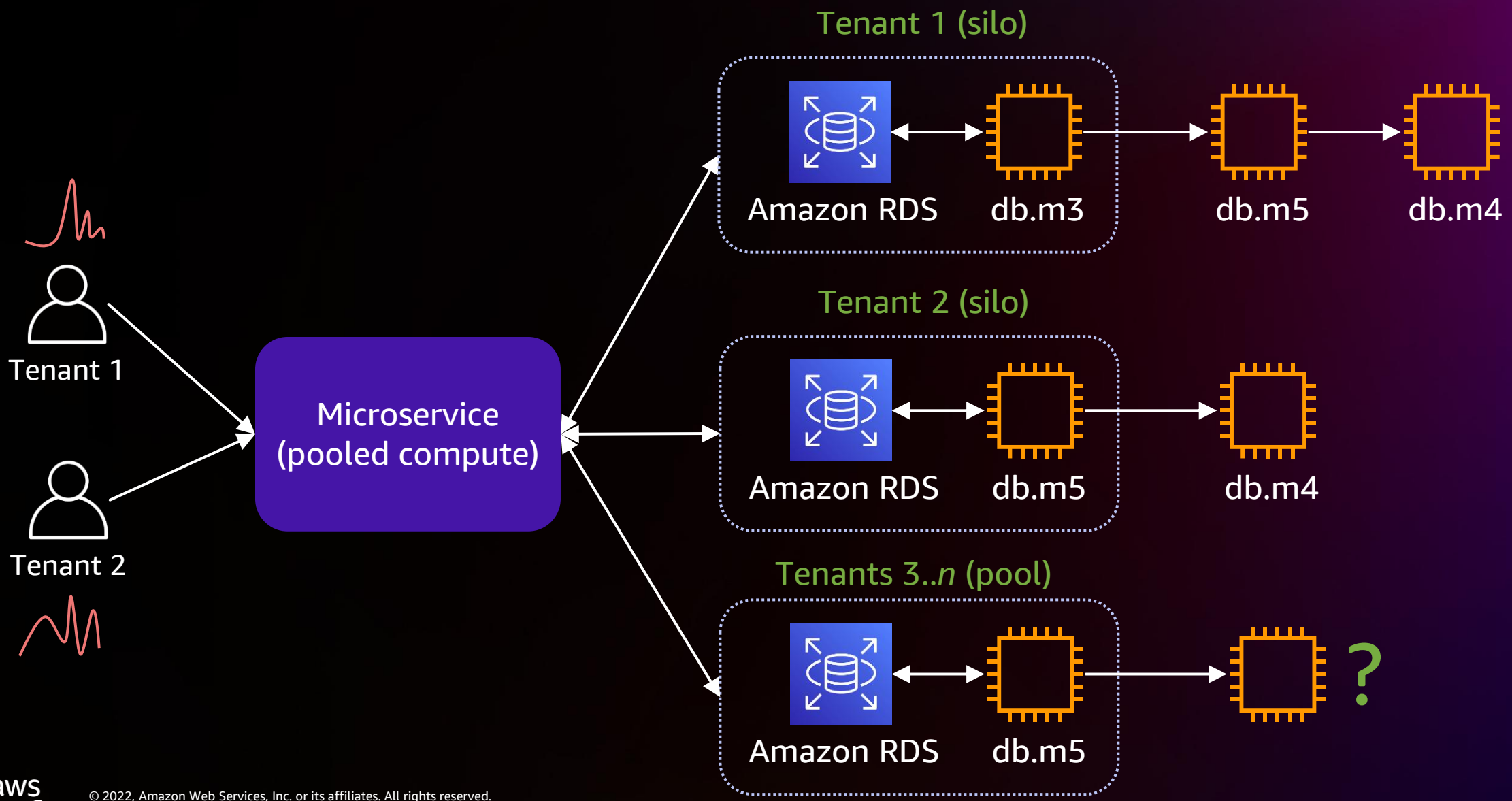
Resource-level consumption

Tenant	ResourceId	Allocation
19393	arn:aws:s3:::my_bucket/tenant19393	11%
19393	arn:aws:rds:us-west-1:111:db:mydb	4%
84719	arn:aws:s3:::my_bucket/tenant84719	8%
84719	arn:aws:rds:us-west-1:111:db:mydb	12%

- Capture consumption for resources (compute, storage, etc.)
- Capture data about features and usage patterns
- Connect consumption to specific microservices
- View consumption in multiple contexts (tenant, tier, service)

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

The right-sizing challenge



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

Making right-sizing someone else's problem

Storage


Amazon Aurora Serverless


Amazon DynamoDB


Amazon S3

Compute


Amazon EKS (AWS Fargate)


Amazon ECS (AWS Fargate)


AWS Lambda

Analytics


Amazon Athena


Amazon QuickSight


Amazon EMR (EMR Serverless)

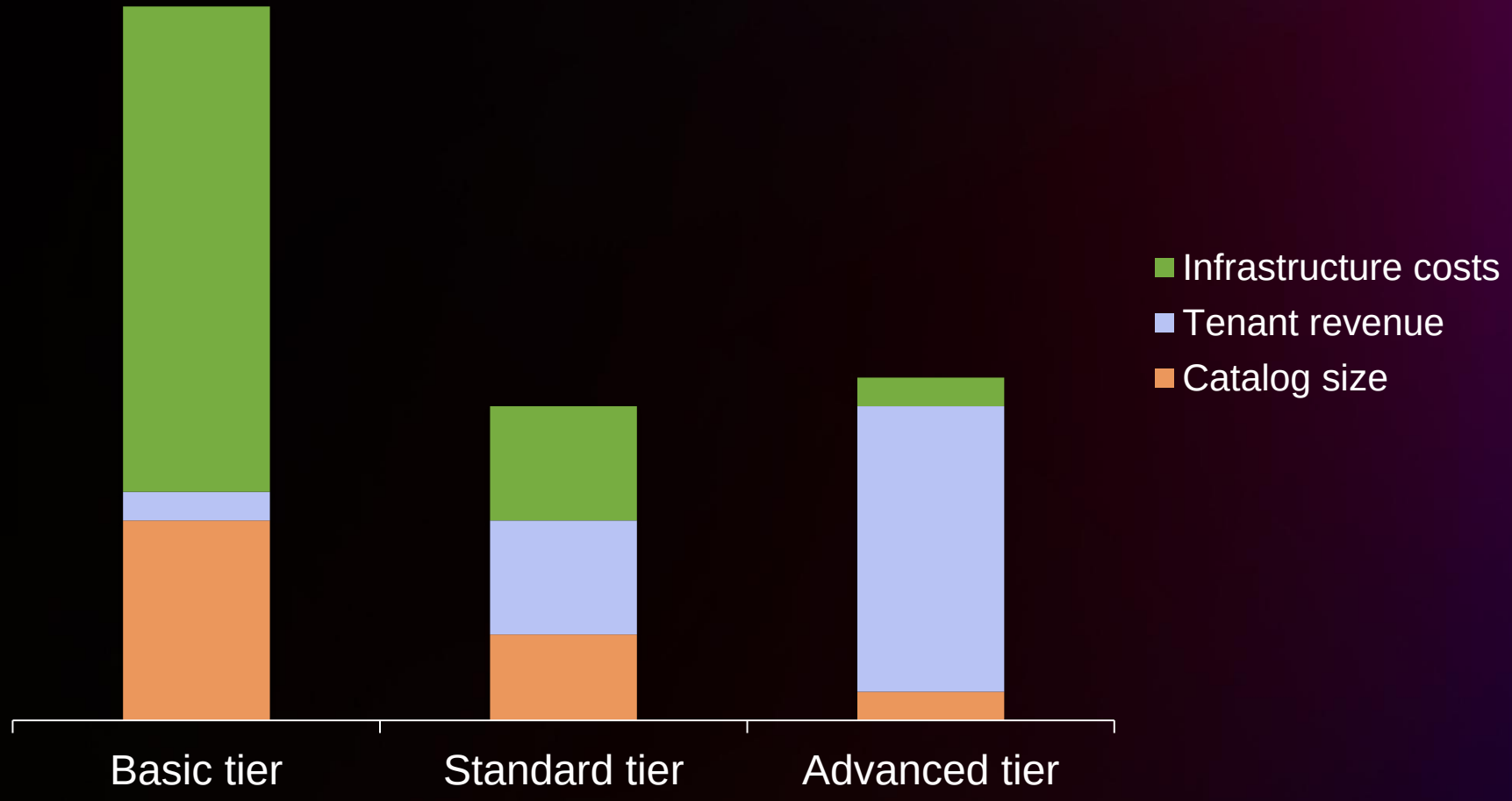
aws

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

Correlating cost with consumption

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

The consumption reality



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

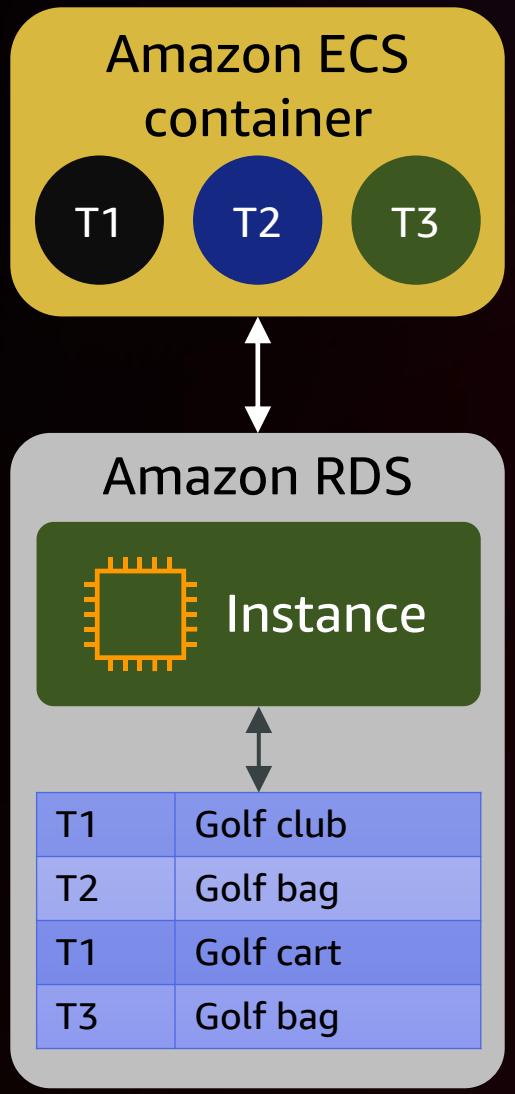
Cost: Thinking beyond margins

Cost units

Compute

Storage (compute)

Storage (disk)



Consumption units

Tenant 1 = 10%
Tenant 2 = 72%
Tenant 3 = 18%

Tenant 1 = 38%
Tenant 2 = 35%
Tenant 3 = 27%

Tenant 1 = 58%
Tenant 2 = 17%
Tenant 3 = 25%

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

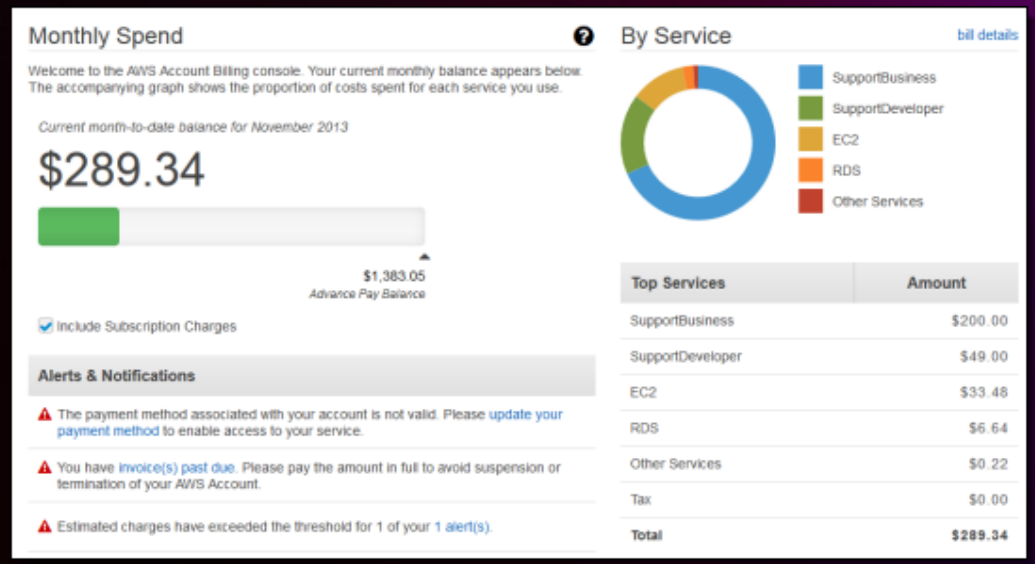
Correlating consumption and cost

Captured tenant allocations

- Amazon EC2
  - Tenant 1 = 80%
  - Tenant 2 = 20%
- Amazon DynamoDB
  - Tenant 1 = 54%
  - Tenant 2 = 46%
- Amazon SQS
  - Tenant 1 = 73%
  - Tenant 2 = 27%
- Amazon RDS
  - Tenant 1 = 61%
  - Tenant 2 = 39%

1

Infrastructure costs



2

Tenant 1 cost = $156.12

Tenant 2 cost = $133.22

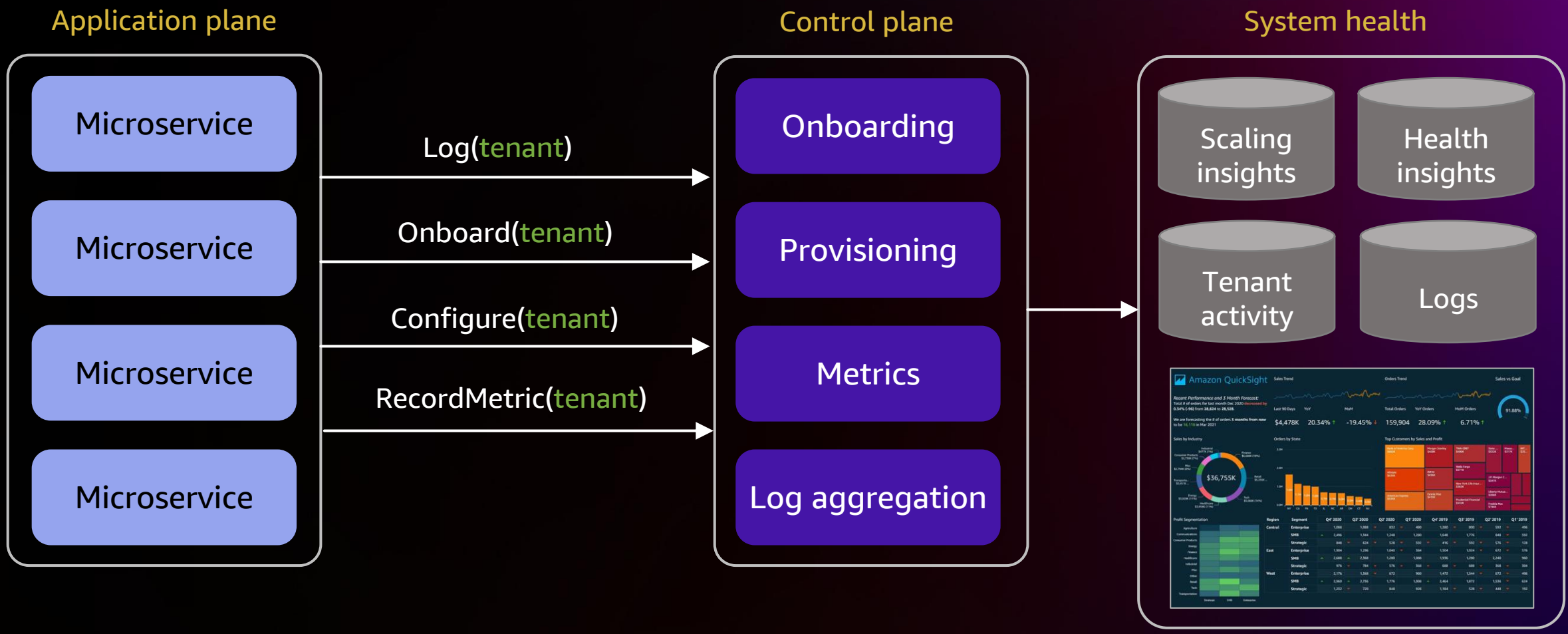
Cost per tenant

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

Extending your operational scope

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

It starts with tenant-aware system health



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

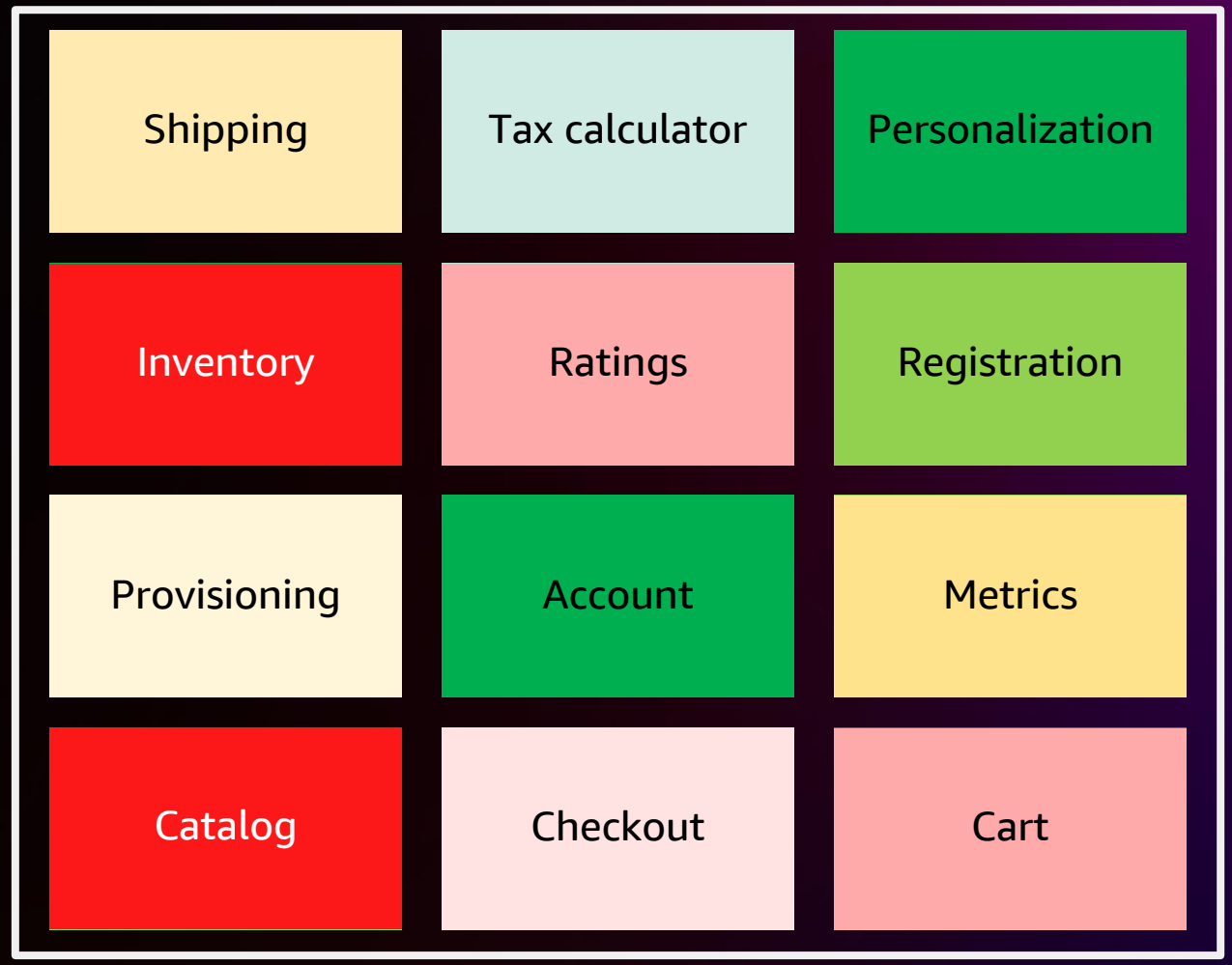
A tenant-aware view of system health

Activity type

- ○ Number of requests
- ○ Latency of calls
- ○ Storage consumption
- ○ SLA threshold violations
- ◉ Most active services
- ○ Least active services

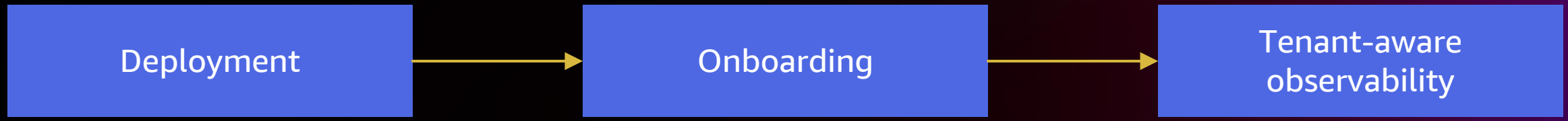
Tenant filters

- ◉ All tenants
- ◉ Tenant tier (advanced)
- ○ Individual tenant



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

A broader view of operational health



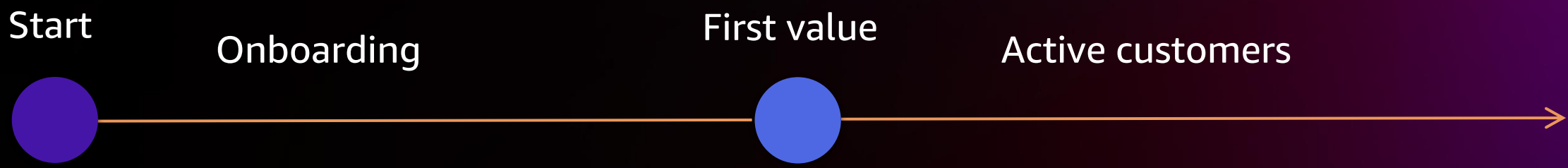
- Cycle time
- Release frequency
- Mean time to detection (MTTD)
- Mean time to recovery (MTTR)
- Defect escape rate
- Failed deployments

- Friction
- Time to value
- Scale
- Fragility
- Failure rate
- Testing/validation

- Tier/tenant metrics
- Proactive alarms
- Tenant activity
- Tenant health

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

Measuring tenant value and engagement



Heavy activity
Brand new customer

Minimal activity
Still onboarding at 30 days

No activity
Still onboarding at 60 days

Heavy activity
Utilizing all features of a license

Declining activity
Partial feature utilization

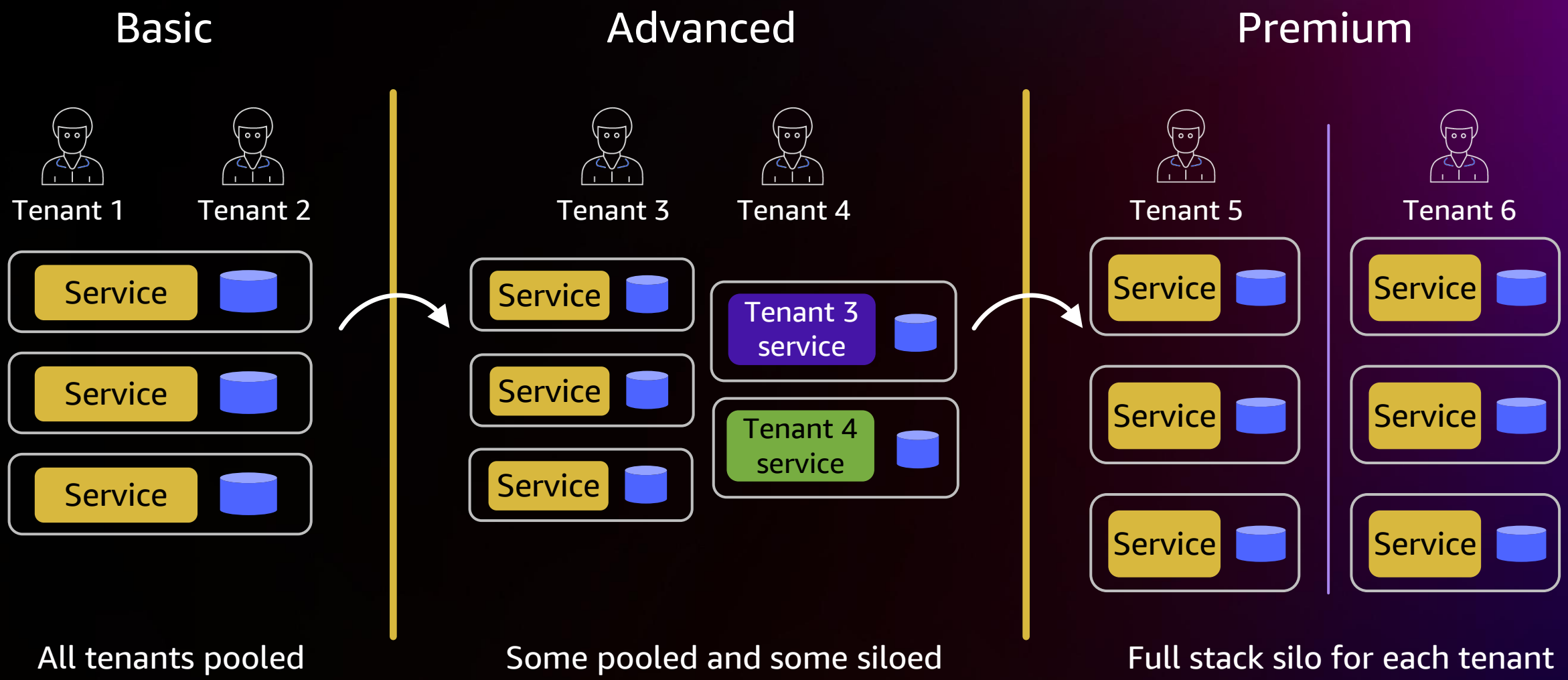
Significant decline in activity
Very limited feature utilization

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

Tiers: Connecting value with cost

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

Tiering: Establishing value boundaries



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

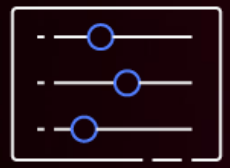
Finding your domain's best tiering boundary



Consumption



Isolation



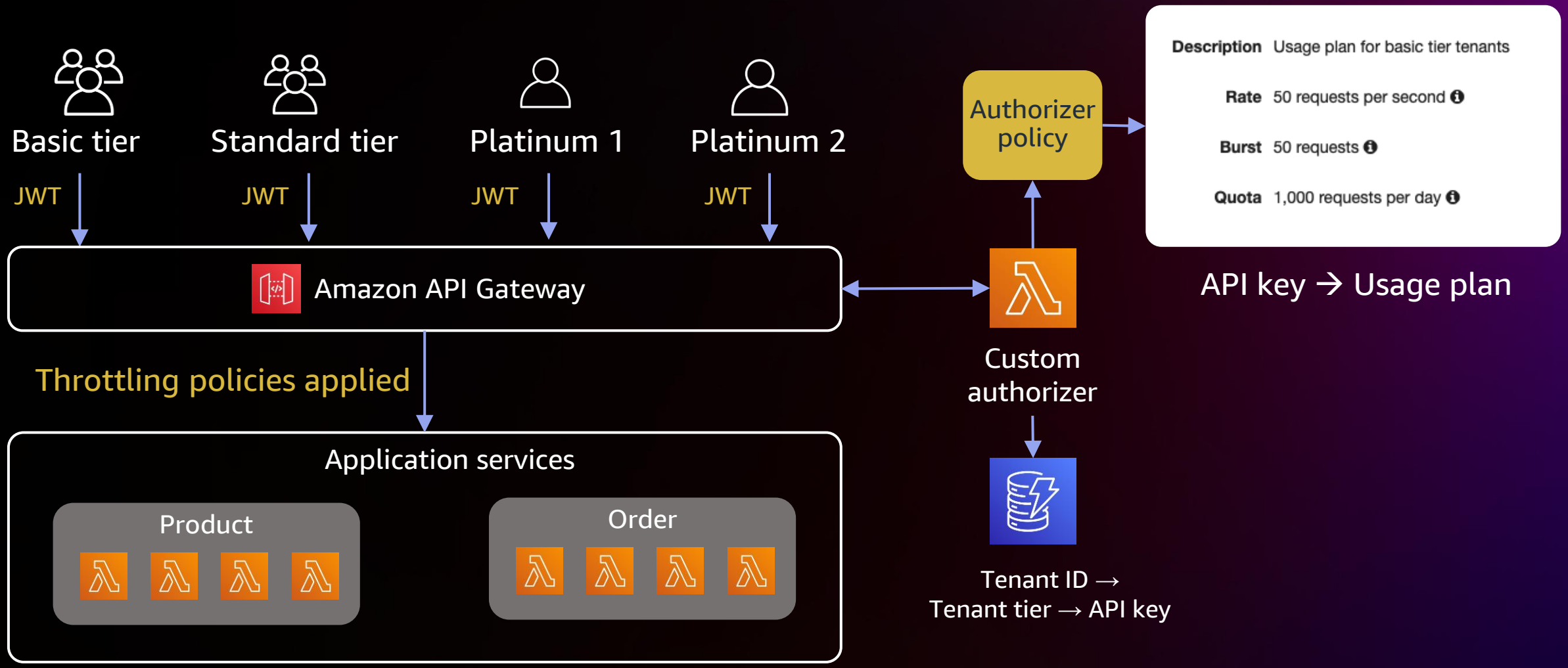
Features



SLAs

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

Applying throttling to tiering



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

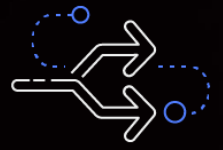
Validating your multi-tenant design

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

Continually validate your multi-tenant design



Noisy neighbor design



Onboarding automation



Tiering strategies



Isolation policies



Operational policies and insights



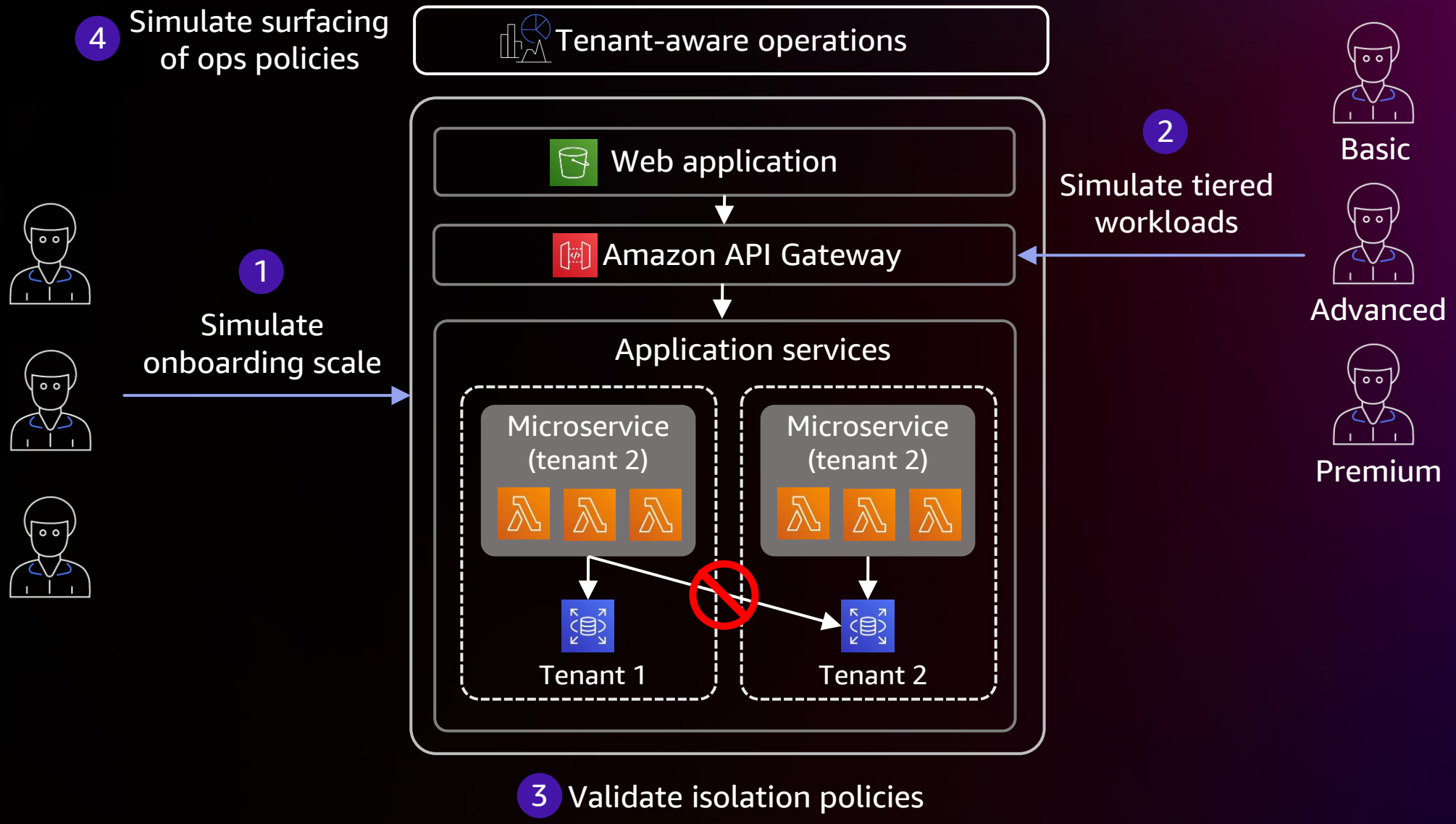
Tenant lifecycle management

How do you test and validate these elements of your SaaS architecture?

aws

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

Simulating multi-tenant environments



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

Takeaways

- You must be data driven to optimize your SaaS environment
- Optimization is a holistic strategy that touches the entire business
- Take a fine-grained approach to decomposing your system
- More knobs and dials give you and the business more options
- Tiering can play an important role in optimization
- Operations isn’t just about health – it’s often at the center of your optimization strategy
- Cost data will play a significant role in shaping your optimization model
- Understand what your tenants are doing and why

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

More SaaS sessions

Breakout sessions

- SAS305 – SaaS architecture patterns: From concept to implementation
- SAS405 – SaaS microservices deep dive: Simplifying multi-tenant development
- SAS306 – SaaS migration: Inside a real-world multi-tenant transformation
- SAS302 – Supporting extensibility in SaaS environments
- PEX310 – Optimizing your multi-tenant SaaS architecture

Workshops

- SAS403 – SaaS microservices deep dive: Multi-tenancy meets microservices
- SAS402 – Serverless meets SaaS: Inside a real-world serverless SaaS solution
- SAS401 – Amazon EKS SaaS: Building a working multi-tenant environment

Business session

- PEX209 – Building your SaaS journey on AWS

aws

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

More SaaS sessions

Chalk talks

- SAS307 – DevOps and SaaS: Applying automation in multi-tenant environments
- SAS303 – SaaS anywhere: Building SaaS solutions that run in hybrid models
- SAS301 – Multi-tenant meets ML: Building ML-based SaaS environments
- SAS304 – Solving the SaaS compliance puzzle
- PEX313 – The SaaS control plane: The heart of SaaS growth
- ARC403 – Amazon EKS SaaS deep dive: Inside a multi-tenant EKS solution
- ARC323 – Designing a multi-tenant SaaS tiering and throttling strategy
- SVS315 – Building multi-tenant applications with AWS Lambda and AWS Fargate

Builders' session

- ARC327 – How to optimize cost in your multi-tenant architecture

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

Additional resources

1

Subscribe to AWS SaaS Insights

Get monthly emails with bite-size advice and the latest updates

2

Explore the SaaS on AWS hub

Check out the SaaS on AWS page for more resources and insights

3

Discover resources for builders

Access our curated list of SaaS reference solutions, demos, tech events, and more

aws

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

Thank you!

Tod Golding
todg@amazon.com

Peter Yang
pcyang@amazon.com



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

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