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

SUP308

Optimize your AWS workloads with best-practice guidance

Ram Cherukuri (he/him)
General Manager, Trusted Advisor
AWS

Kartik Cheekoty (he/him)
Principal Product Manager, Trusted Advisor
AWS

Alistair McLaurin (he/him)
AWS Cloud Platform Lead
HSBC

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

Session overview

- Overview of AWS Support
- Best-practice guidance with AWS Support
  - AWS Trusted Advisor
  - AWS Trusted Advisor Priority
  - AWS Well-Architected
  - AWS re:Post
  - AWS Support app in Slack
  - AWS Health
- Review how HSBC optimizes their workloads

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

AWS Premium Support plans

Developer	Business	Enterprise On-Ramp	Enterprise Support
			Training 500 online training credits per year per customer Additional credits at a 30% discount
		Account and Billing ✓ Concierge Support (reactive)	**Account and Billing** ✓ Full Concierge Support (proactive)
		Architecture Review (1 per year)	**Architecture Review** ✓
		Cost Optimization (1 per year)	**Cost Optimization** ✓
		Operational Review (1 per year)	**Operational Review** ✓
		Technical Account Manager (On-Demand TAM Pool)	**Technical Account Manager** (Designated TAM)
		Strategic Business Review (2 per year)	**Strategic Business Review** ✓
		TAM-assisted case escalation (2 per month)	**TAM-assisted case escalation** ✓
		TAM Office Hours (1 per month)	**TAM Office Hours** ✓
	Infrastructure Event Management (Available for additional fee)	**Infrastructure Event Management** (1 per year)	**Infrastructure Event Management** ✓
Case Severity/Response Times ✓ General guidance: < 24 hours System impaired: < 12 hours	**Case Severity/Response Times** ✓ General guidance: < 24 hours System impaired: < 12 hours Production system impaired: < 4 hours Production system down:< 1 hour	**Case Severity/Response Times** ✓ General guidance: < 24 hours System impaired: < 12 hours Production system impaired: < 4 hours Production system down: < 1 hour **Business-critical system down: < 30 mins**	**Case Severity/Response Times** ✓ General guidance: < 24 hours System impaired: < 12 hours Production system impaired: < 4 hours Production system down: < 1 hour **Business-critical system down: < 15 mins**
AWS Service Guidance ✓ General guidance	**AWS Service Guidance** ✓ Contextual guidance based on your use case	**AWS Service Guidance**✓ Consultative review and guidance based on your applications	**AWS Service Guidance** ✓ Consultative review and guidance based on your applications

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

What is optimization?

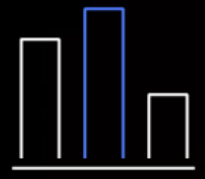
op·ti·mize /ˈäp-tə-ˌmīz/

Process or methodology of making something as effective or functional as possible

- Continuous, not a one-time activity
- Driven by business objectives
- Requires investment
- Areas – cost, performance, sustainability, and security

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

Common questions around optimization



Are we using resources **efficiently**?



How do we **gain** better **visibility**?



How do we build **accountability**?

Where do we **start**?

aws

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

AWS Trusted Advisor



Trusted Advisor

Home of AWS best-practice recommendations

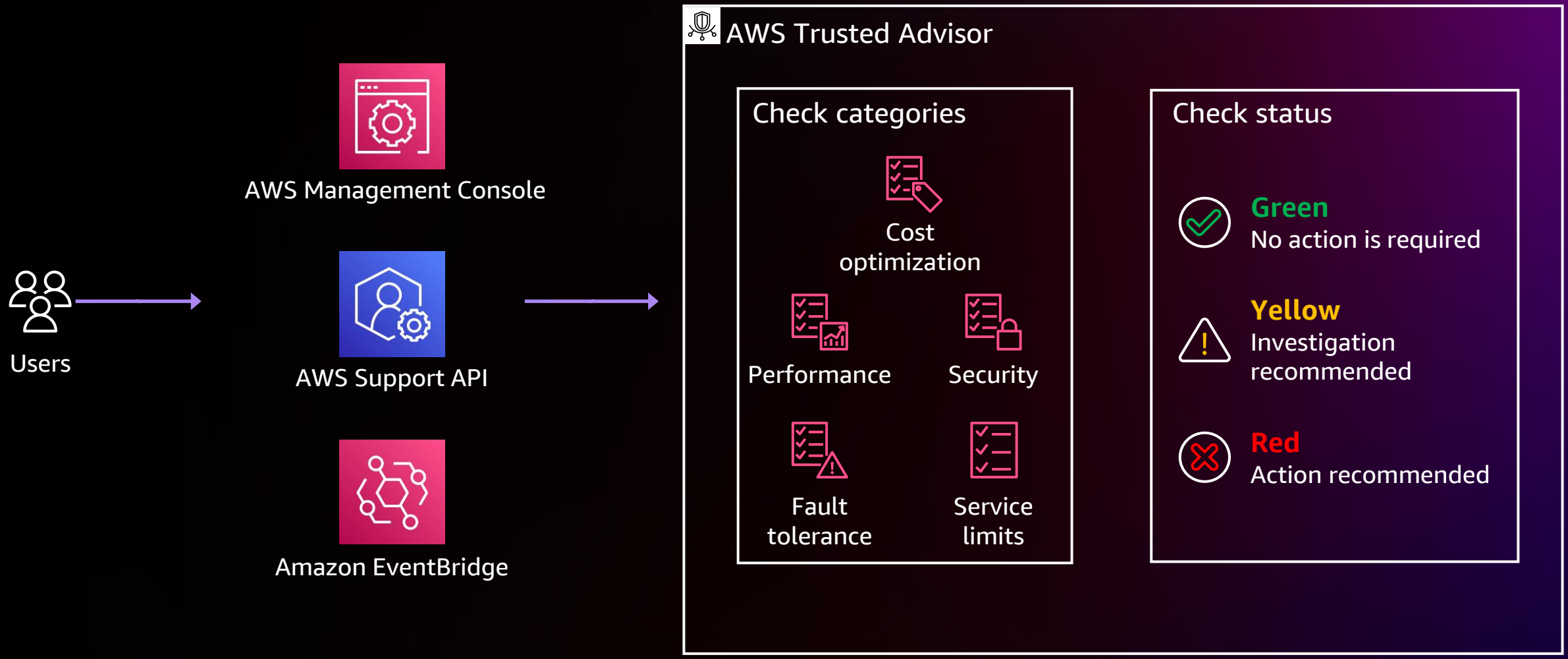
Recommendations to help you follow AWS best practices to optimize your services and resources

- Inspects your AWS environment and makes recommendations
- Saves money, improves resilience and performance, closes security gaps
- Trusted Advisor checks for all accounts in your AWS Organizations

aws

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

Codified best practices



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

Coverage of best practices

Services with AWS Trusted Advisor best practices


Amazon EC2


AWS Lambda

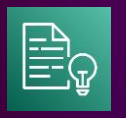

Amazon EC2 Auto Scaling


AWS Compute Optimizer


Amazon DynamoDB


Amazon ElastiCache


Amazon RDS


Amazon Comprehend


Amazon Kinesis


Amazon OpenSearch Service


Amazon VPC


Amazon Redshift


Amazon CloudFront



Amazon Route 53


AWS Direct Connect


Elastic Load Balancing


AWS Identity and Access Management (IAM)


Amazon SES


AWS CloudTrail


AWS WA Tool


AWS Security Hub


AWS Cost Explorer


Amazon S3


Amazon EBS

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

AWS Trusted Advisor Priority

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

What is Trusted Advisor Priority?

- Prioritized, context-aware recommendations and critical risks
- Curated by AWS account team members
- IT leaders and central governance teams
- Management account and delegated administrator account of your AWS Organizations
- Available for AWS customers on Enterprise Support

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

Where do recommendations come from?

Trusted Advisor Priority recommendations can come from one of two sources



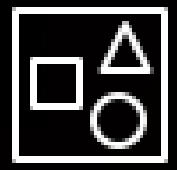
Trusted Advisor



Your AWS account team

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

What context is applied in Trusted Advisor Priority?



Knowledge of production accounts



Critical workloads



What's important to you



Outcomes you wish to achieve

aws

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

Recommendations shared on Trusted Advisor Priority

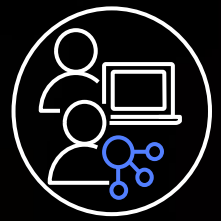
- Prioritized programmatic checks
- Account team-identified opportunities
  - Single points of failure such as single Availability Zone deployments
  - End-of-life notifications
- Support engagements like a Cost Optimization Workshop

Detect **Raise awareness** **Provide prescriptive guidance**

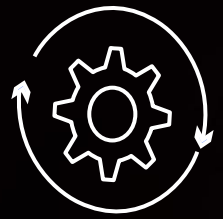
aws

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

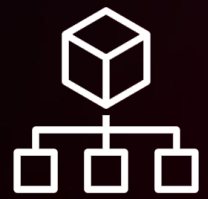
Benefits of Trusted Advisor Priority



Prioritized and custom guidance



Closed-loop feedback tracking



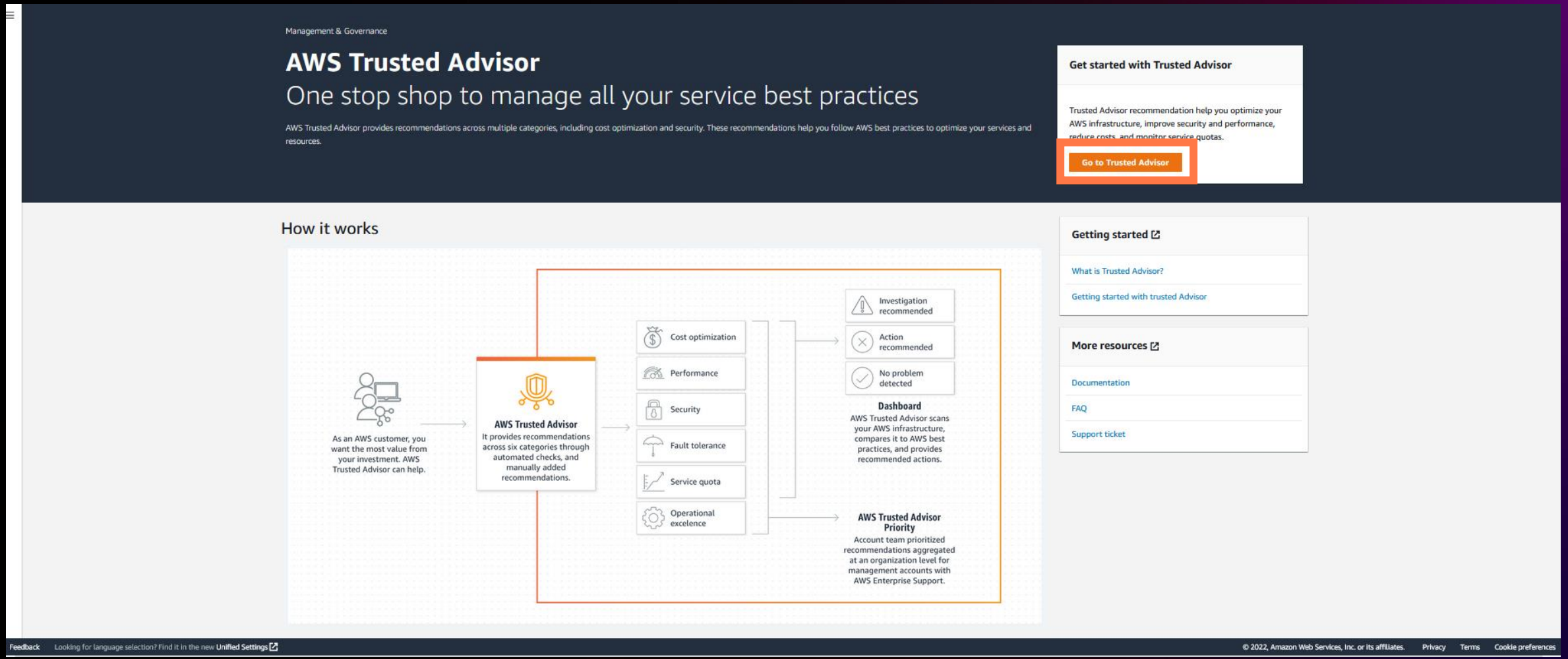
Multi-account/ AWS Organizations views



Historical view of actions taken

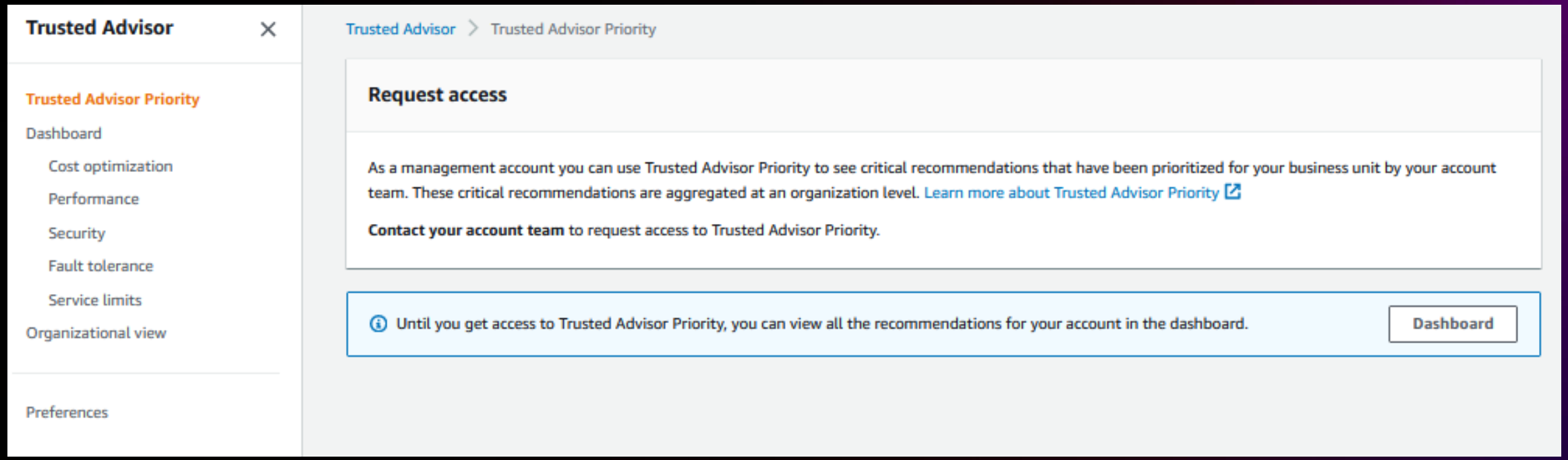
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How do I get started?



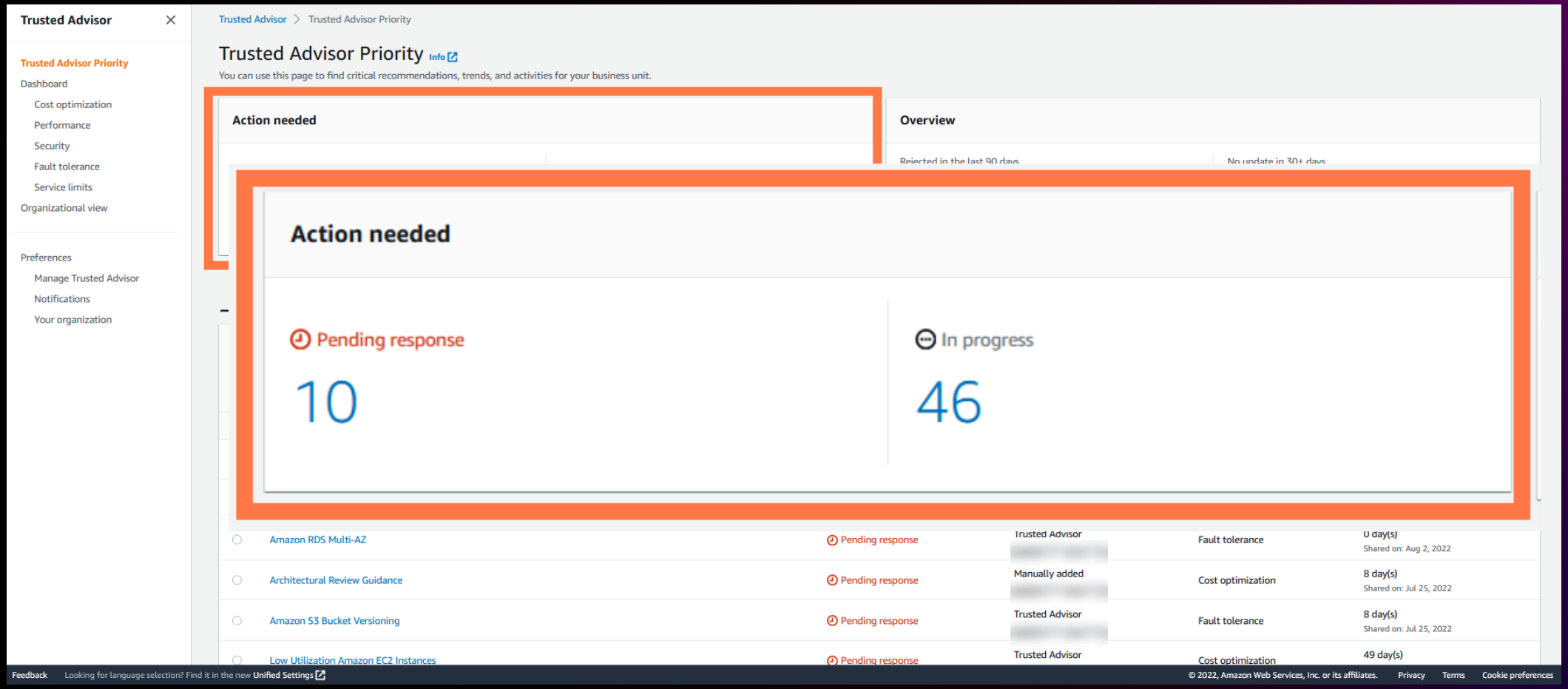
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How do I get started?



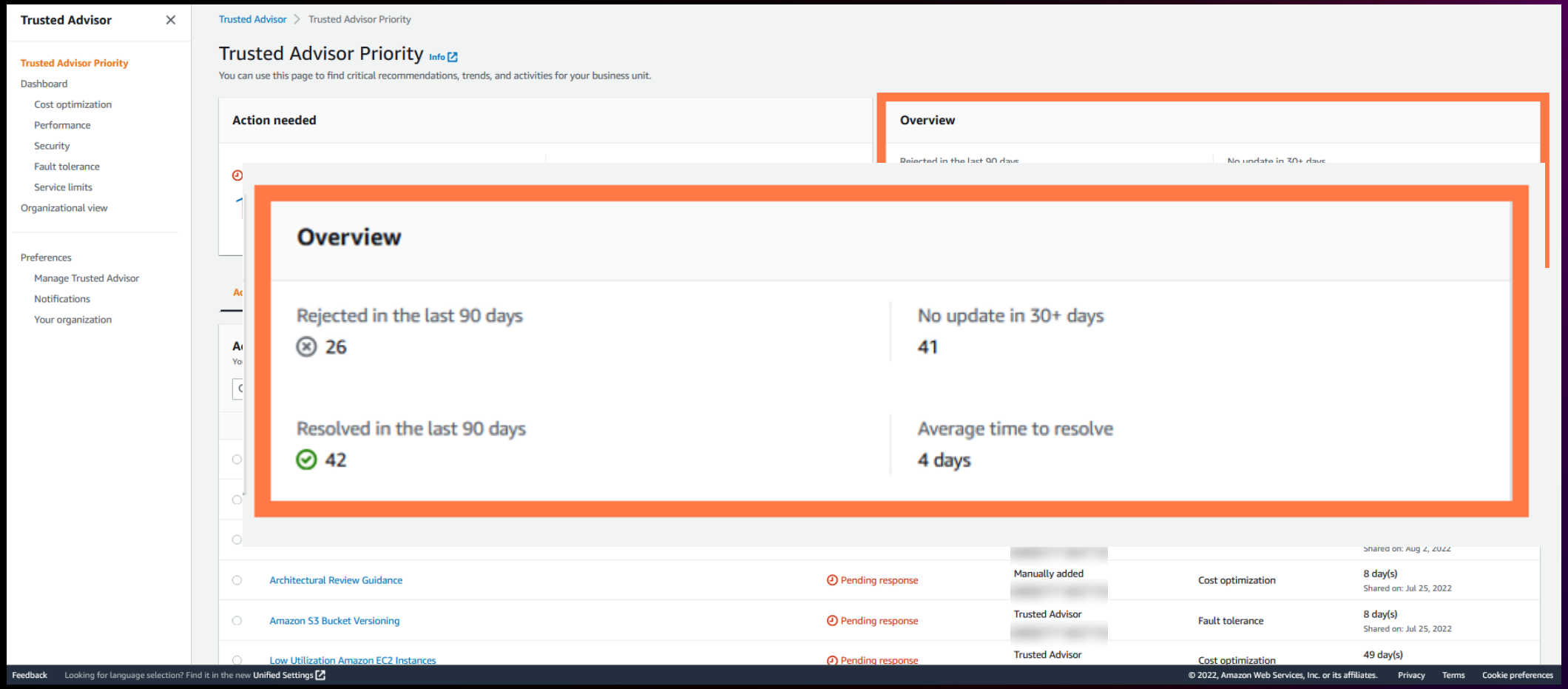
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Call to action on Trusted Advisor Priority



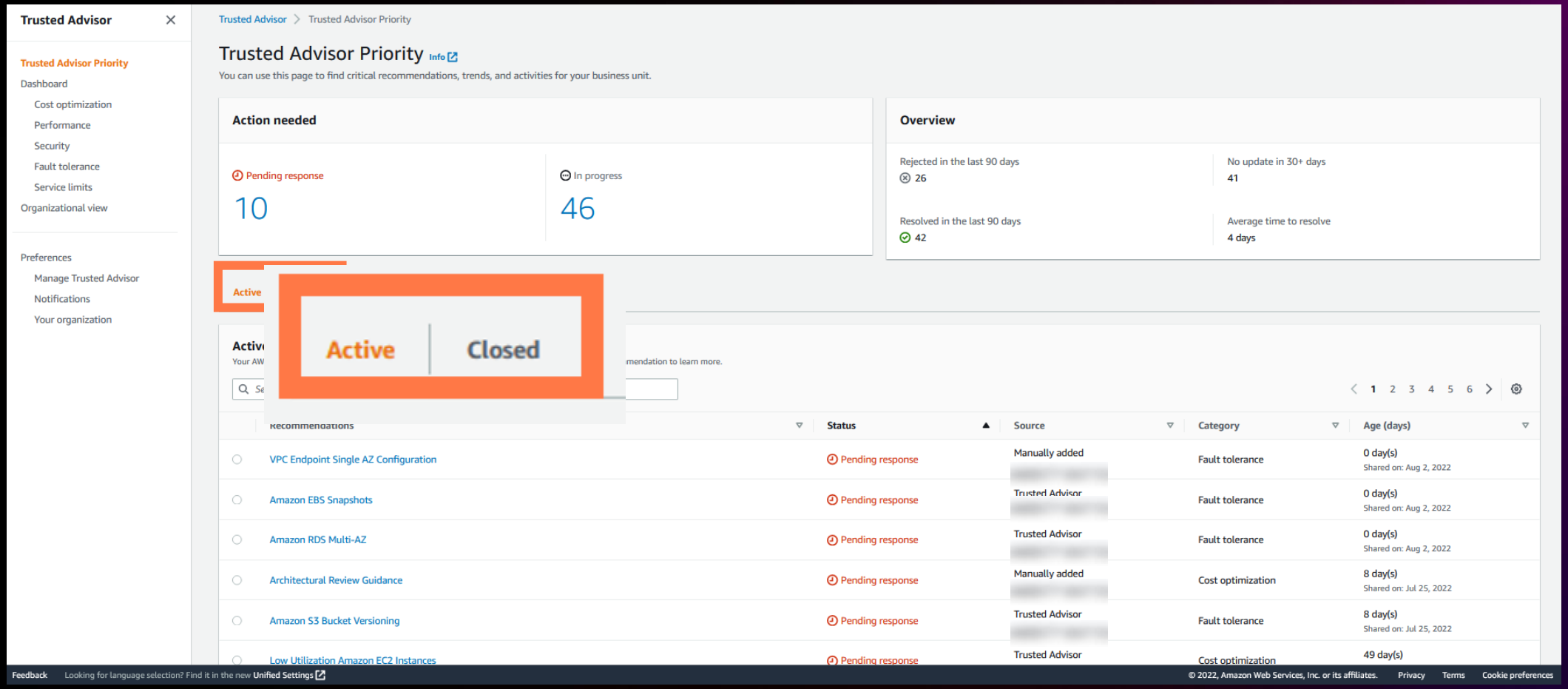
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Trusted Advisor Priority overview



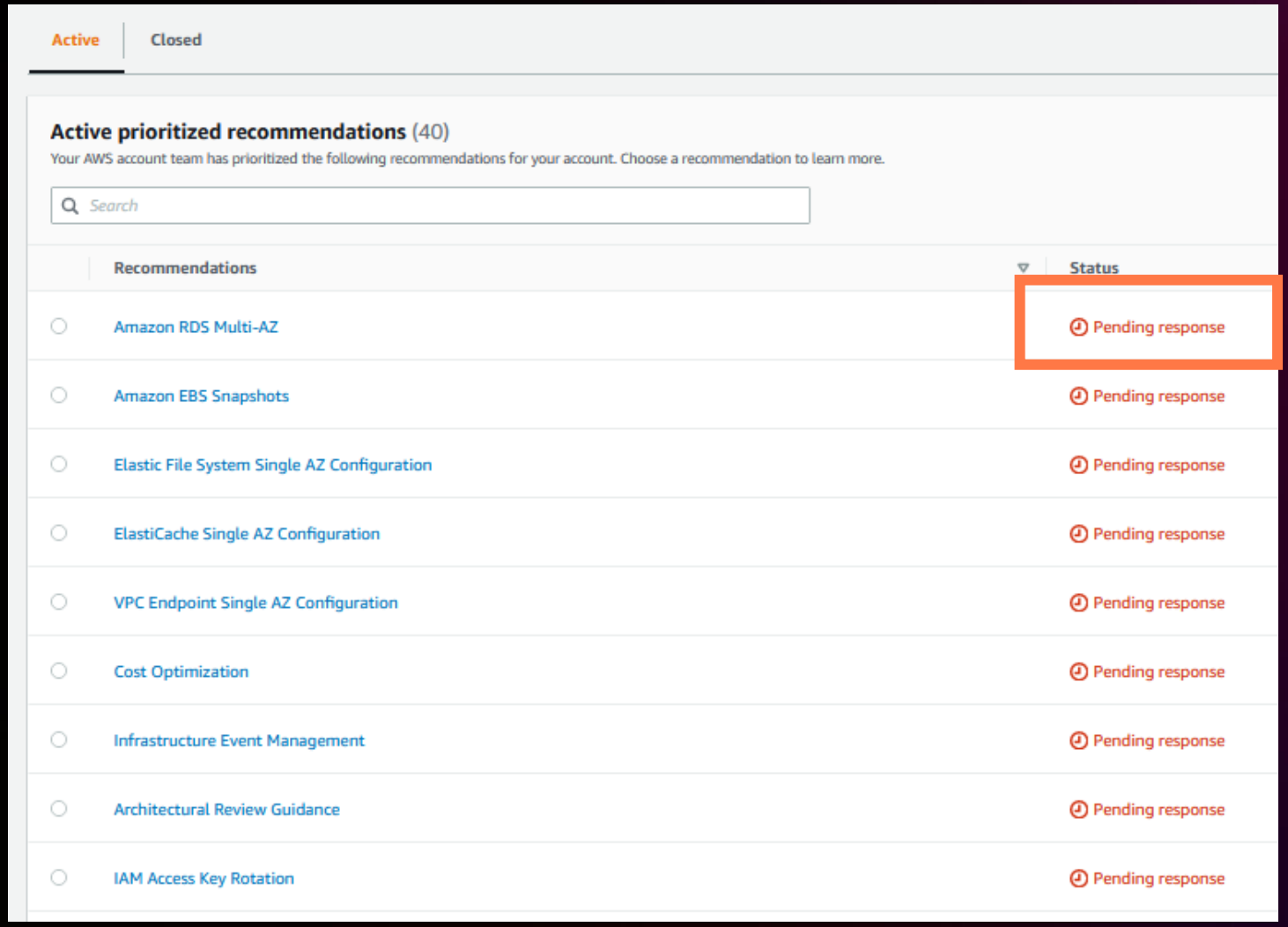
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Active and closed recommendations



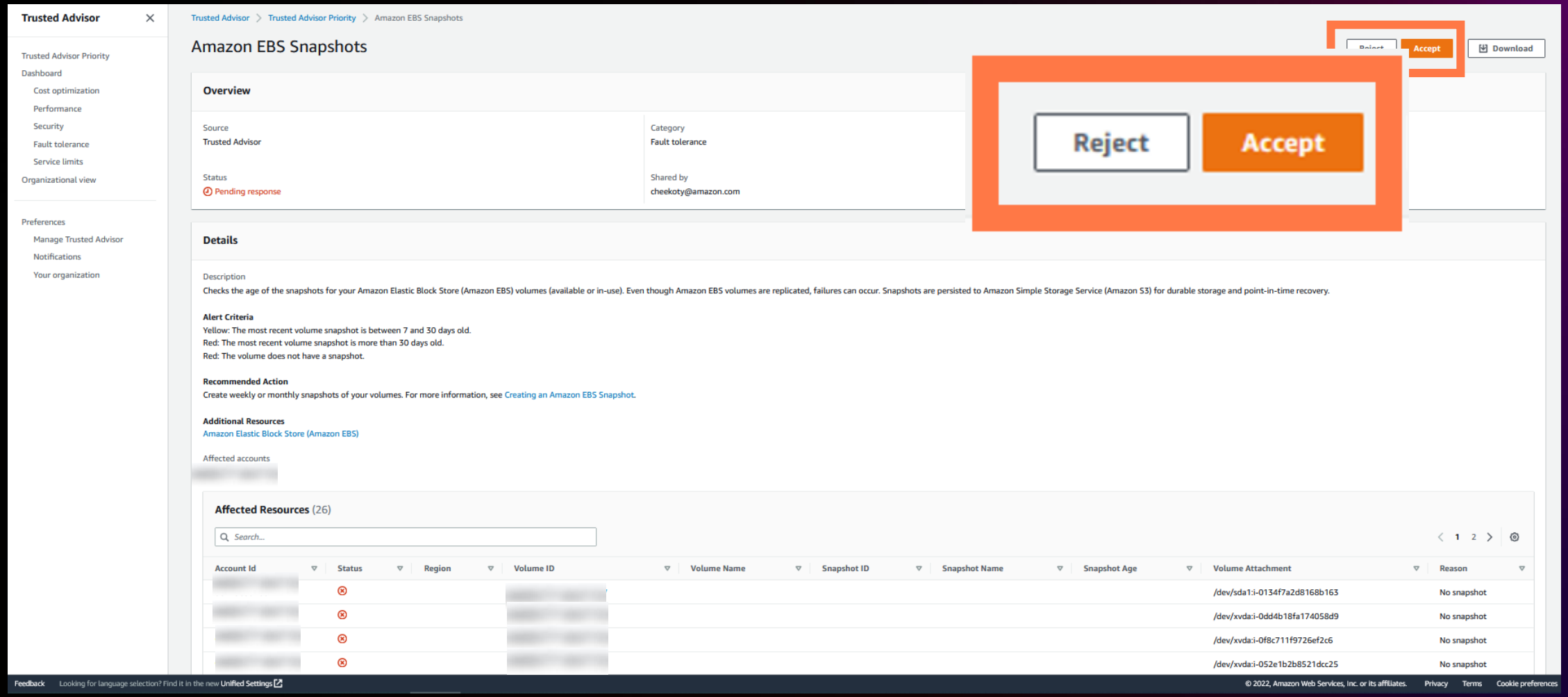
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Recommendation status



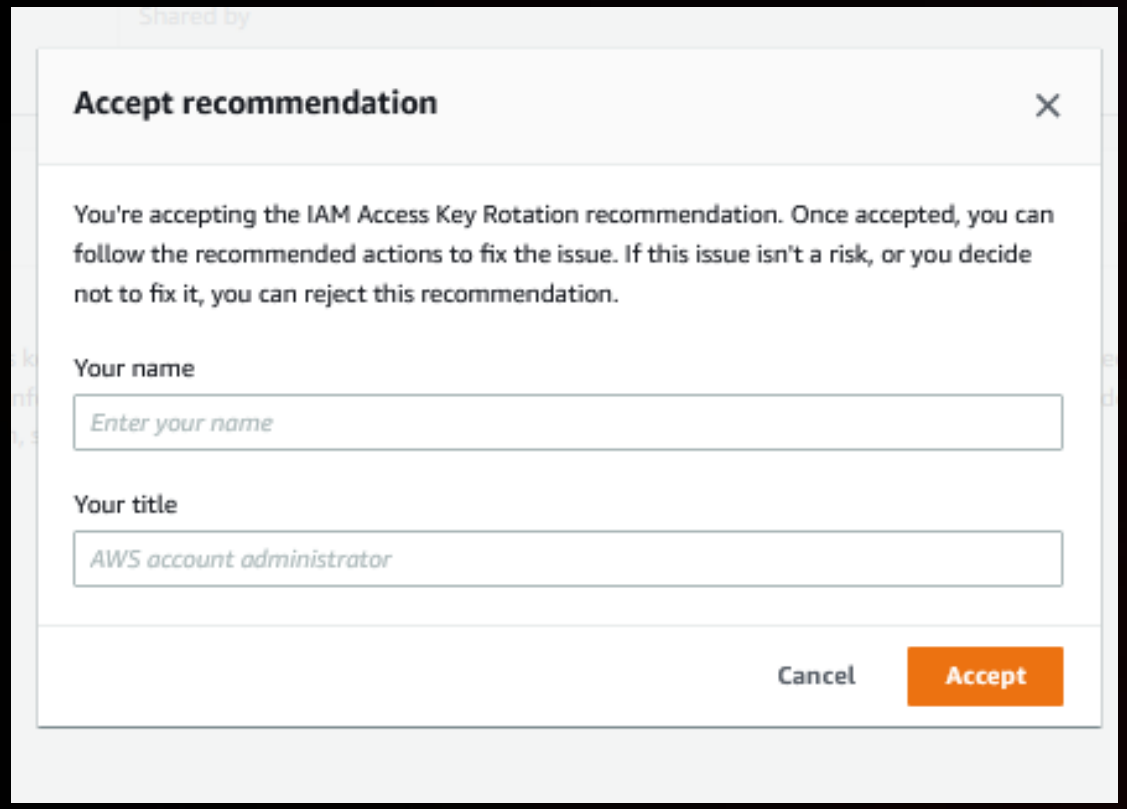
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

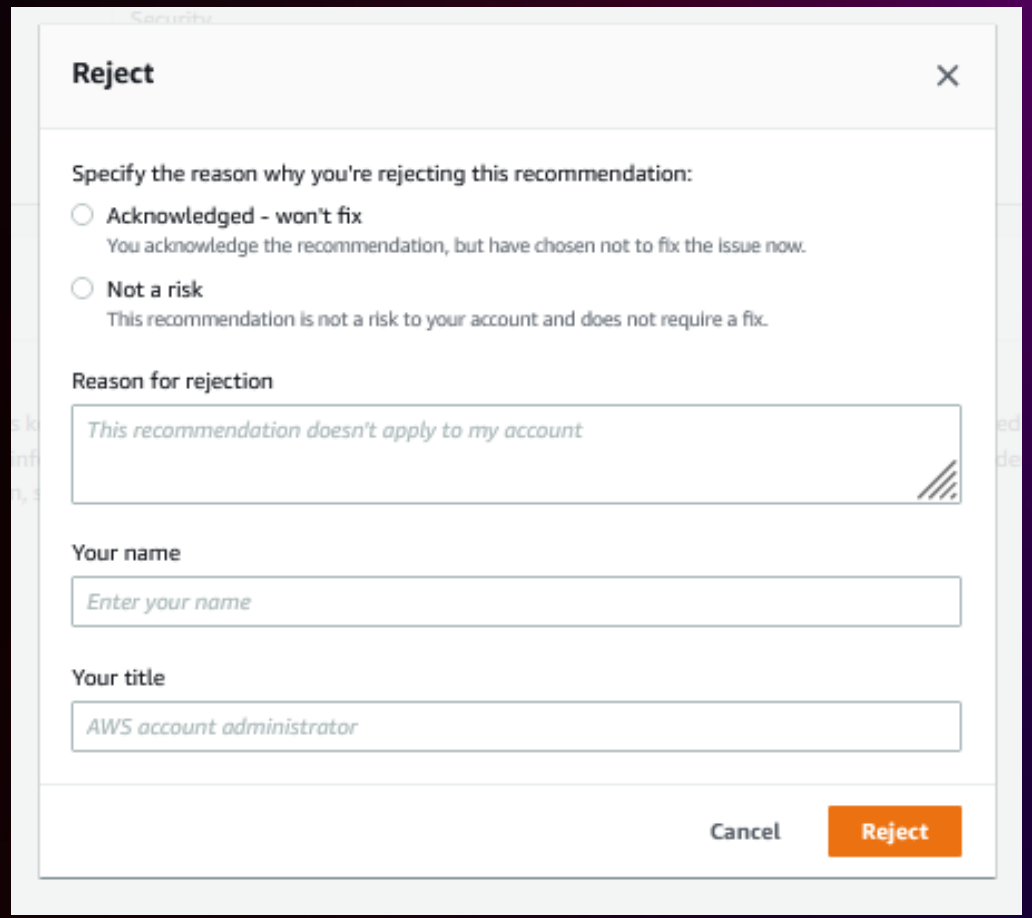
Accept/reject a prioritized recommendation



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

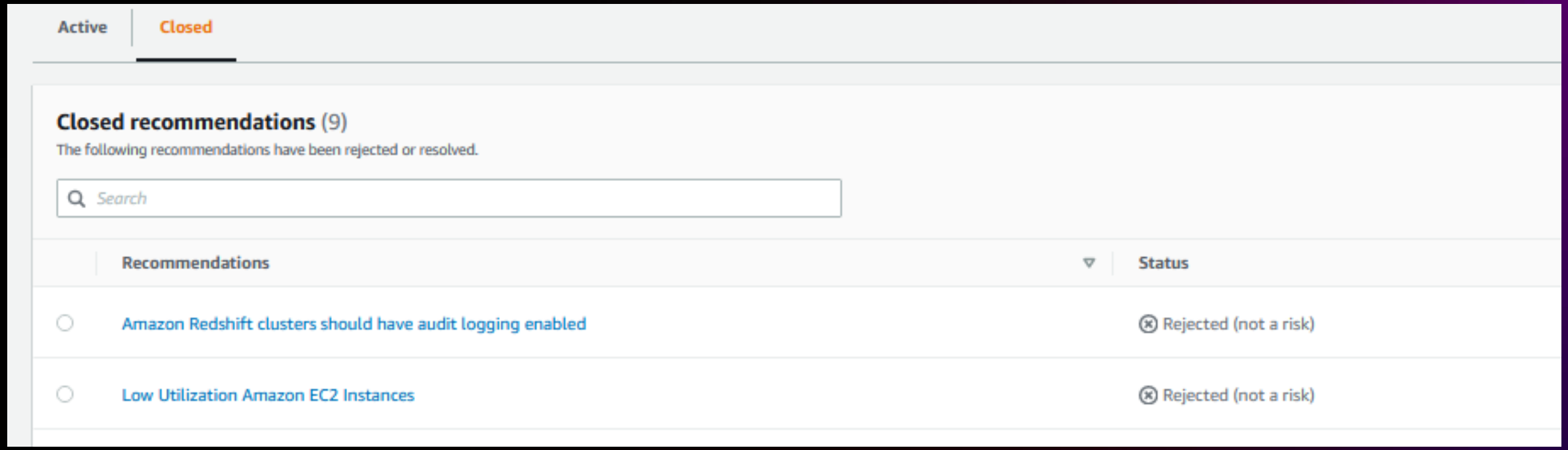
Accept/reject a prioritized recommendation





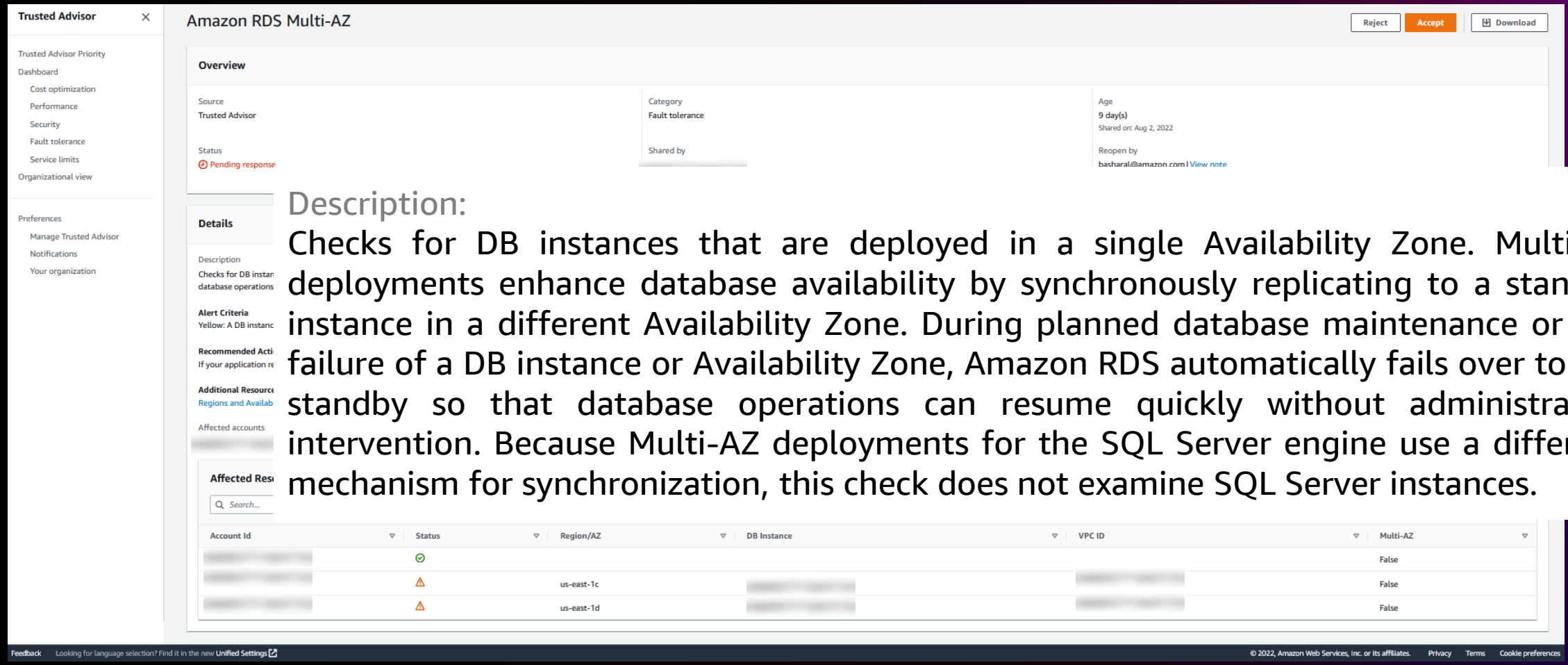
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Closed recommendations



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

Recommendation #1: Improve resiliency

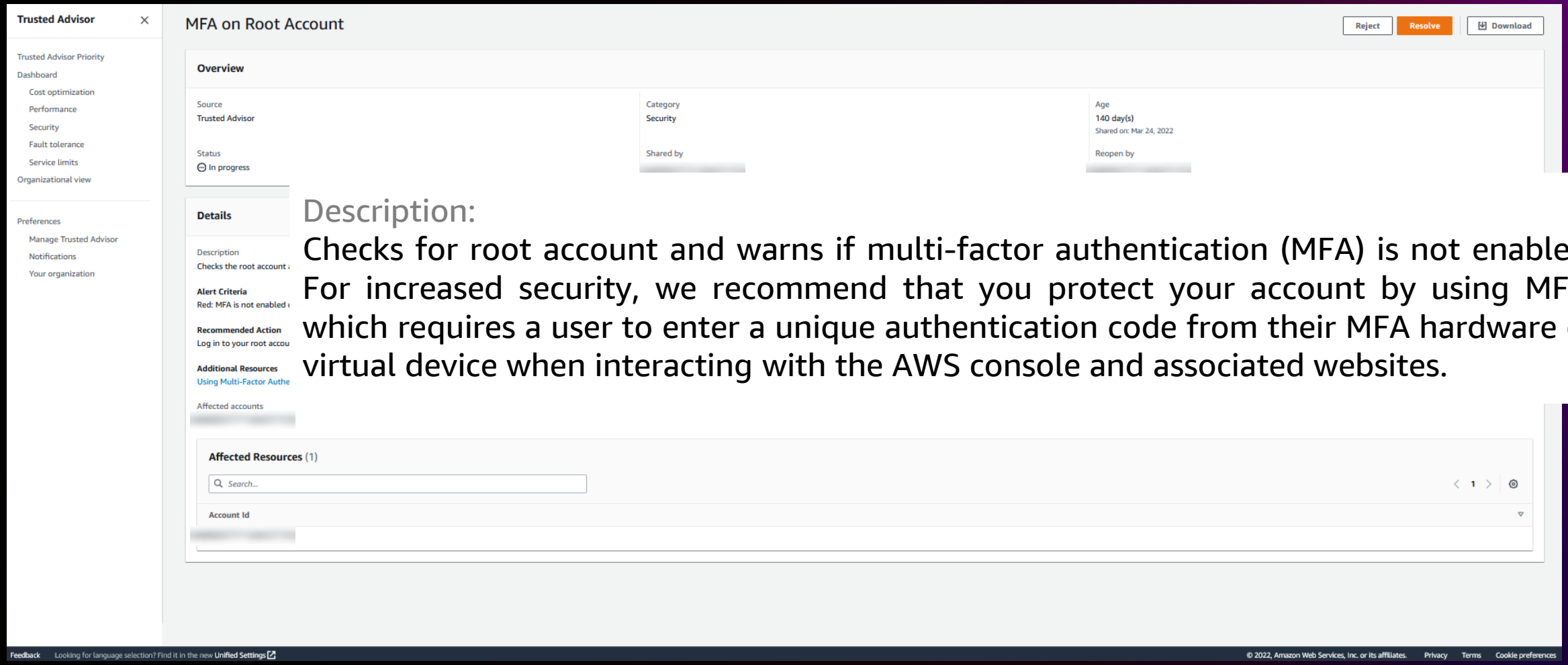


Description:

Checks for DB instances that are deployed in a single Availability Zone. Multi-AZ deployments enhance database availability by synchronously replicating to a standby instance in a different Availability Zone. During planned database maintenance or the failure of a DB instance or Availability Zone, Amazon RDS automatically fails over to the standby so that database operations can resume quickly without administrative intervention. Because Multi-AZ deployments for the SQL Server engine use a different mechanism for synchronization, this check does not examine SQL Server instances.

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

Recommendation #2: Improve security

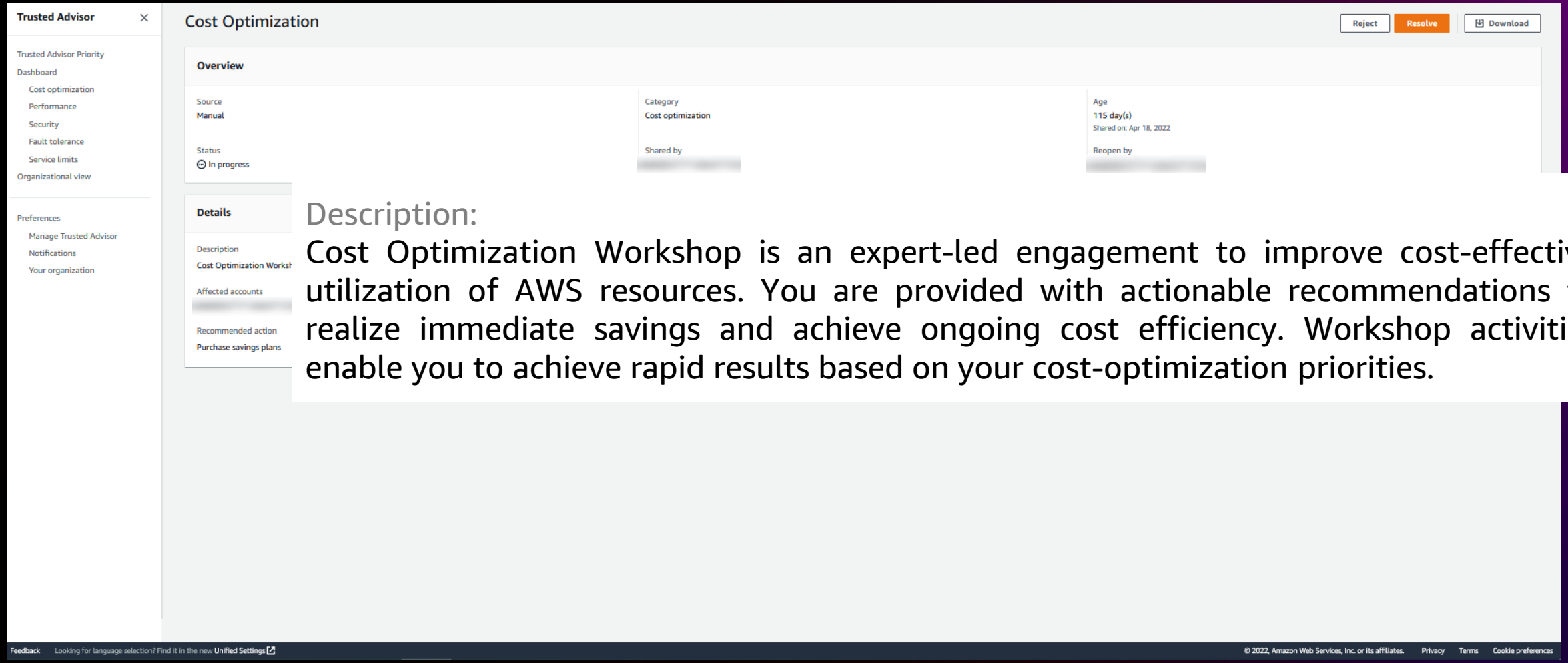


Description:

Checks for root account and warns if multi-factor authentication (MFA) is not enabled. For increased security, we recommend that you protect your account by using MFA, which requires a user to enter a unique authentication code from their MFA hardware or virtual device when interacting with the AWS console and associated websites.

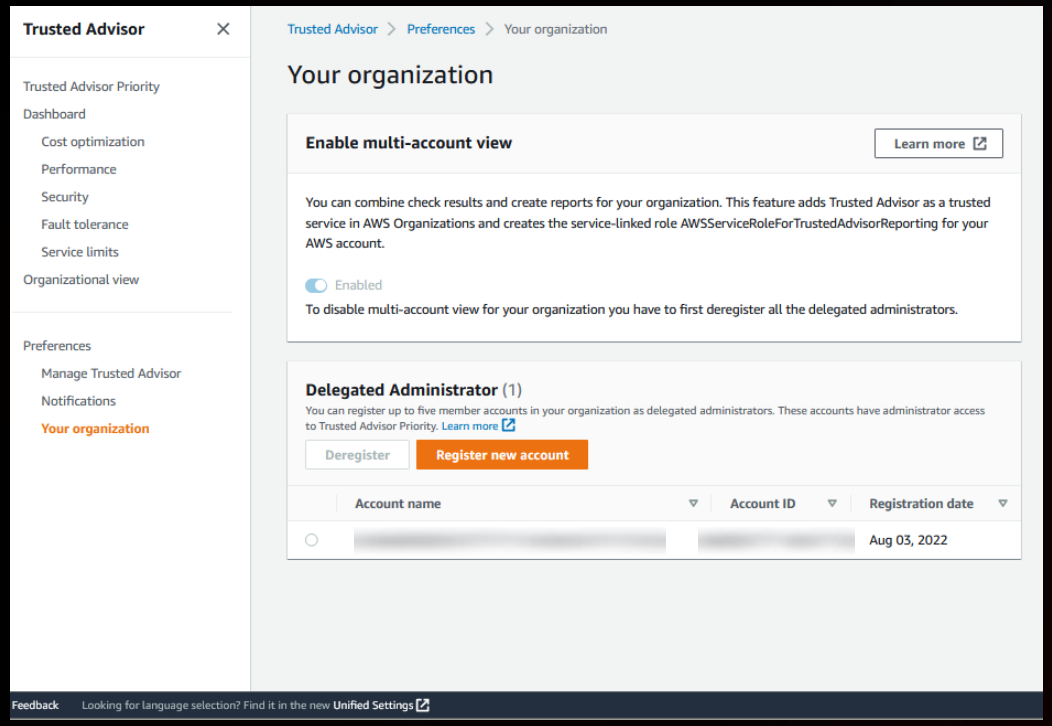
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

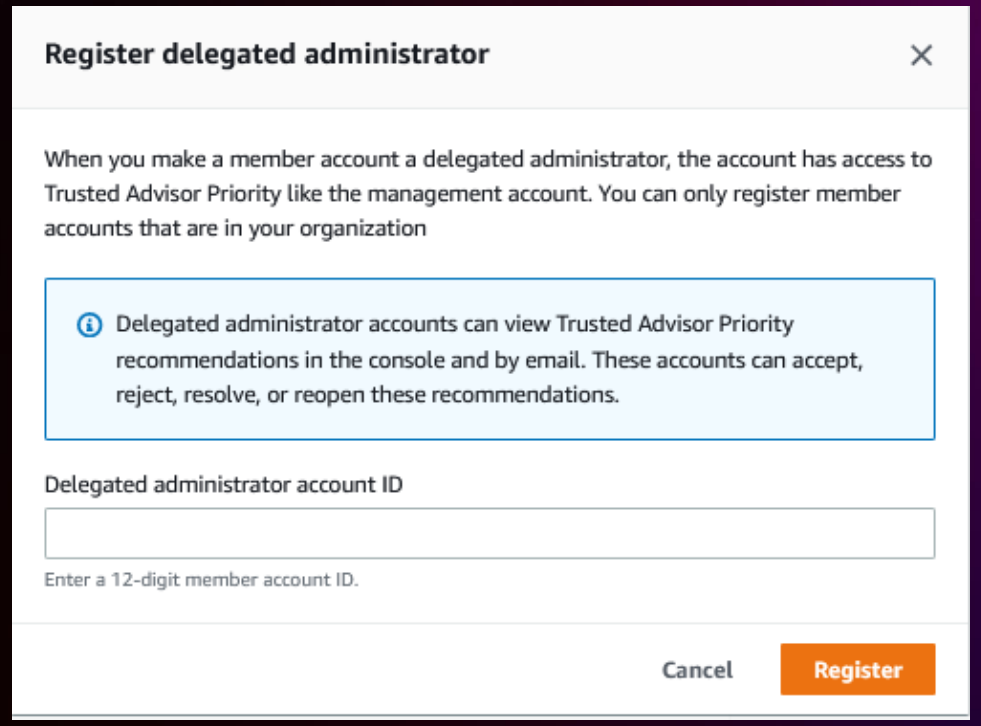
Recommendation #3: Improve cost efficiency



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

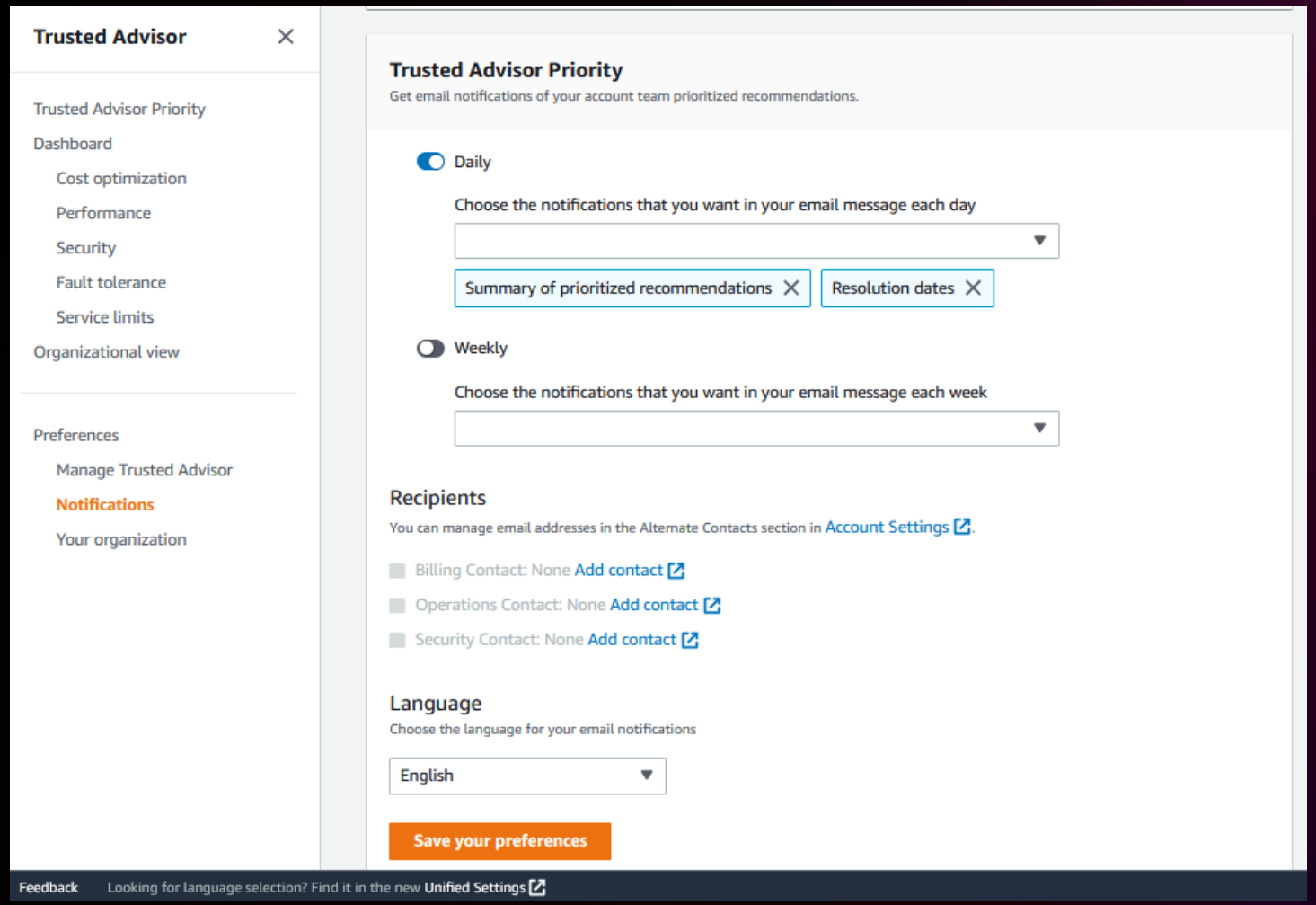
Delegated administrator





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

Email notifications



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

Key takeaways

- Trusted Advisor dashboard – available in all accounts
- Trusted Advisor Priority – multi-account view; prioritized recommendations for the IT leader persona
- Form a leadership team for Trusted Advisor
- Reach out to your AWS account team to get started on Trusted Advisor Priority!

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

Customer feedback on TA Priority



“We liked the whole framework as we could now get resource level details from the identified risks and also track all the identified risks within our team from a single dashboard. It is a high value add to us as this helps us to easily collaborate with our internal stakeholders and AWS Account team.”

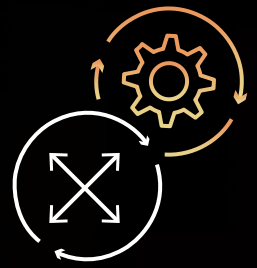
Head of CloudOps
A5 Labs

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

AWS Well-Architected Framework

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

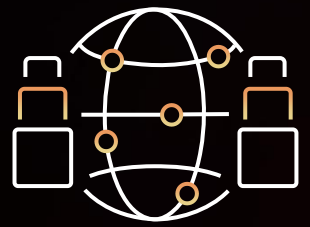
Six pillars of AWS Well-Architected Framework



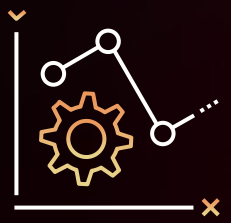
Operational excellence



Security



Reliability



Performance efficiency



Cost optimization



Sustainability

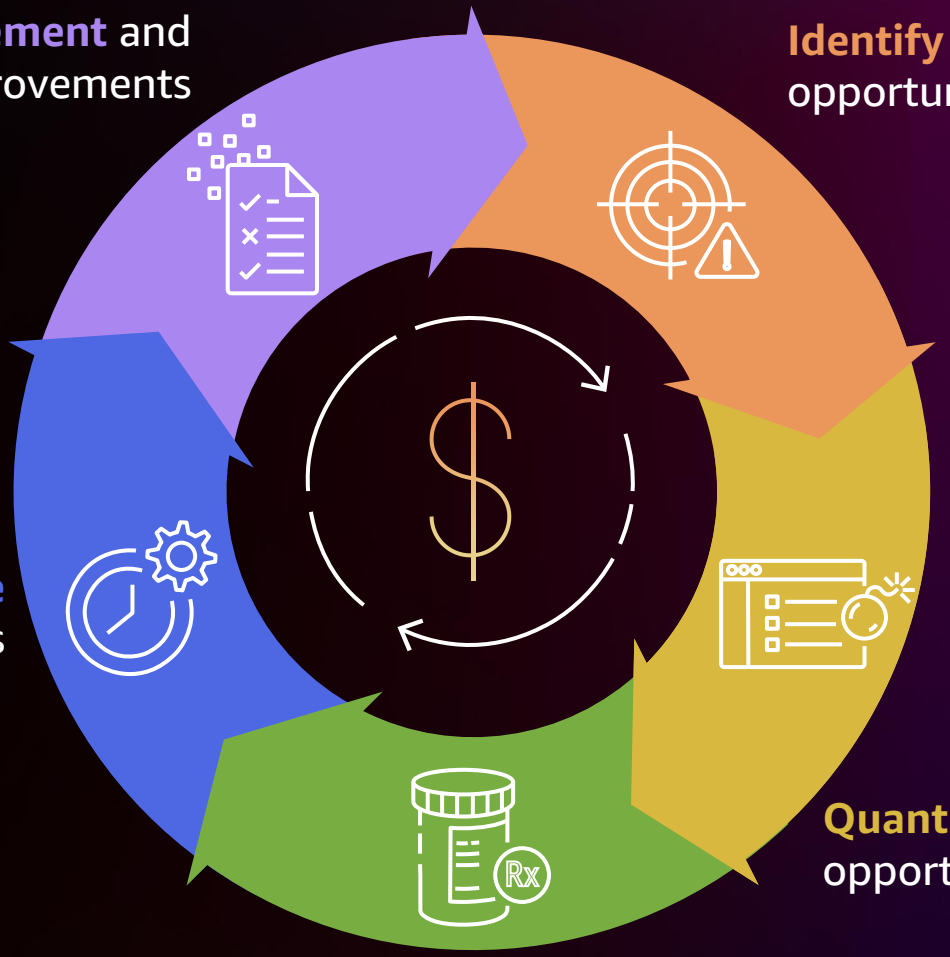
aws

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

Optimization review



AWS Well-Architected review

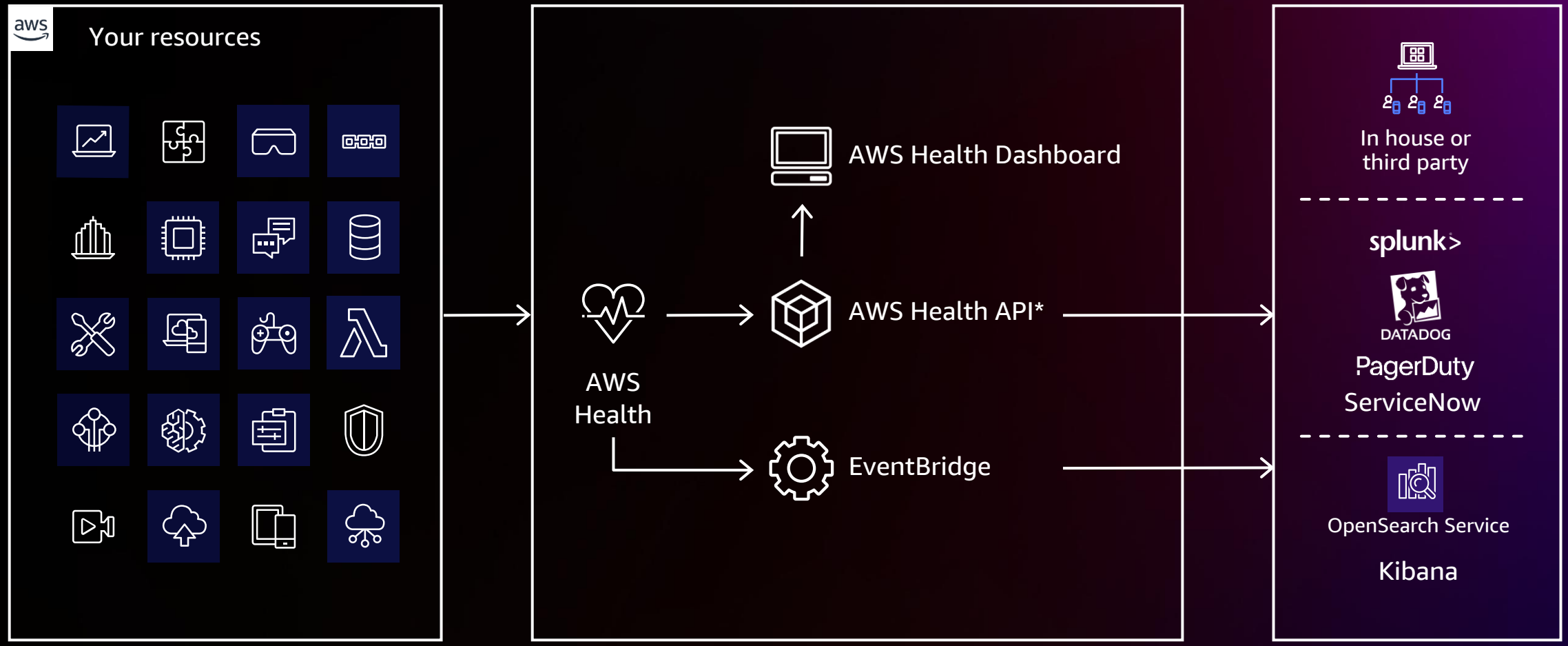


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS Health

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

Transparent visibility into your resource performance and availability



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

* AWS Health API is only available for premium support customers

AWS re:Post

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

re:Post is a community-driven, knowledge-sharing site



A single source of truth

AWS continuously publishes verified Q&A pairs to address common questions



Expert-reviewed answers

Experts review community-provided answers for accuracy



Better engagement

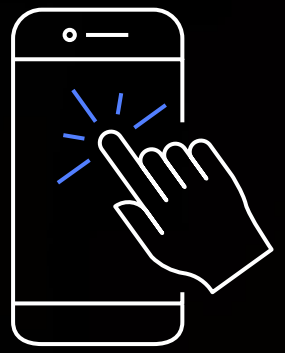
Built-in mechanisms to increase the likelihood of responses

re:Post launched at re:Invent 2021 and replaced AWS Forums

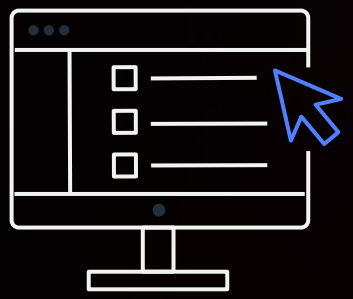
aws

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

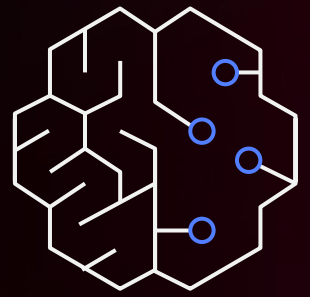
Easily find the answers you're looking for



Login is not required to browse content



Content is organized by topics and tags



Machine learning drives smart content recommendations

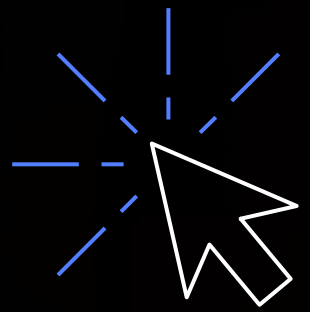


Follow specific questions, topics, and community experts

aws

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

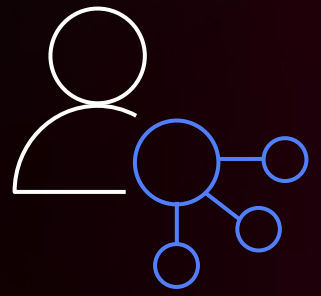
Get guidance you can trust



Log in to re:Post with your AWS credentials



Ask and answer questions; upvote and downvote responses



Questions are automatically shared with domain experts

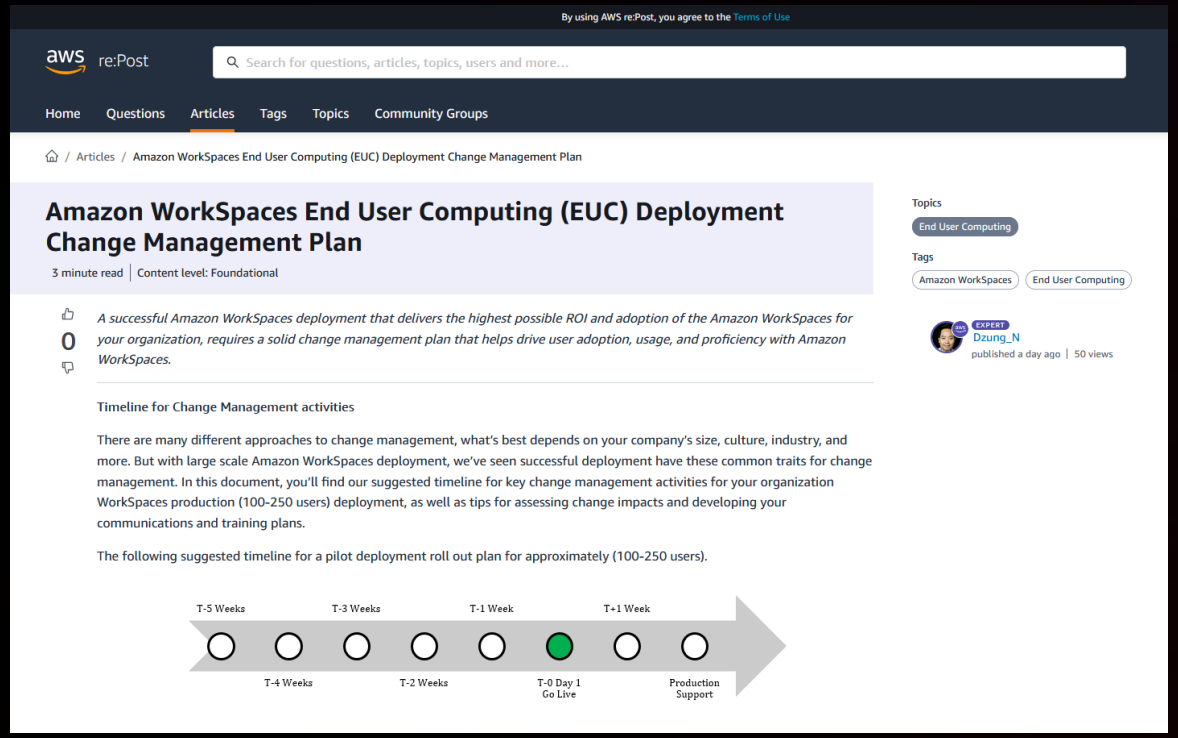


Answers are marked as reviewed by experts and AWS employees

aws

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

New in AWS re:Post



- Community-generated articles
- New way for experts to share knowledge
- Diagrams, how-tos, best practices
- To design, develop, and optimize

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

AWS Support app in Slack

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

AWS Support app in Slack



- Your team members can now work alongside AWS Support engineers
- Identify root causes and resolve issues
- No need to move between your Slack workspace and the AWS Support Center console

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

AWS Support app in Slack



Create, update, search for, and resolve support cases



Start a live chat session with support agents and members from your team



Request service quota increases

aws

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



How does HSBC optimize its workloads?

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

HSBC introduction



- HSBC serves more than 40 million customers and operates in 64 countries and territories
- We support around 1.3 million businesses worldwide
- There were **1.28 billion logins** to our personal banking mobile apps in 2020

aws

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

HSBC AWS case study

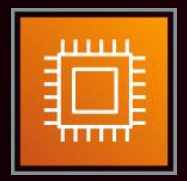
The HSBC AWS environment today consists of:



> 1,000 accounts



> 400 production apps



> 80 approved services



7 AWS Regions

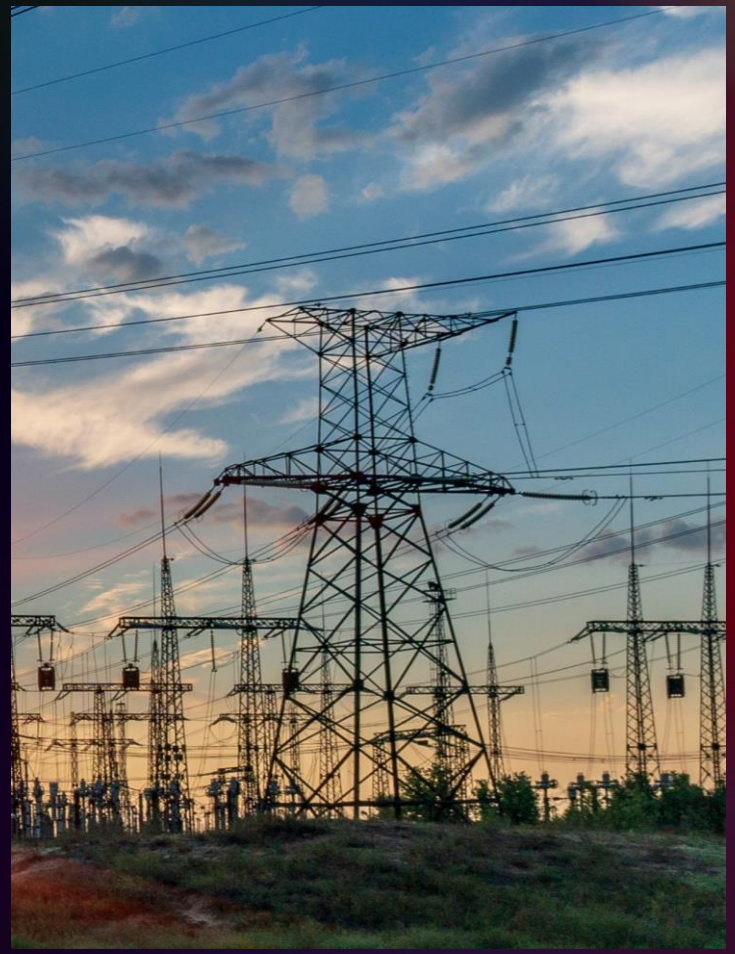
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Managing risk

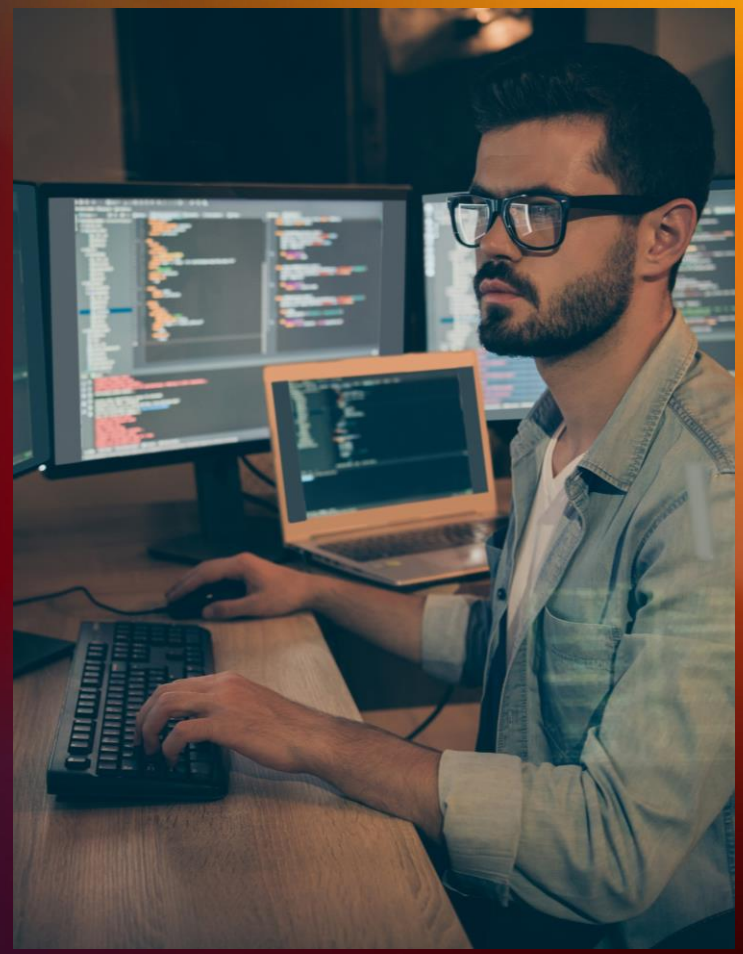
Natural disaster



Technical failure



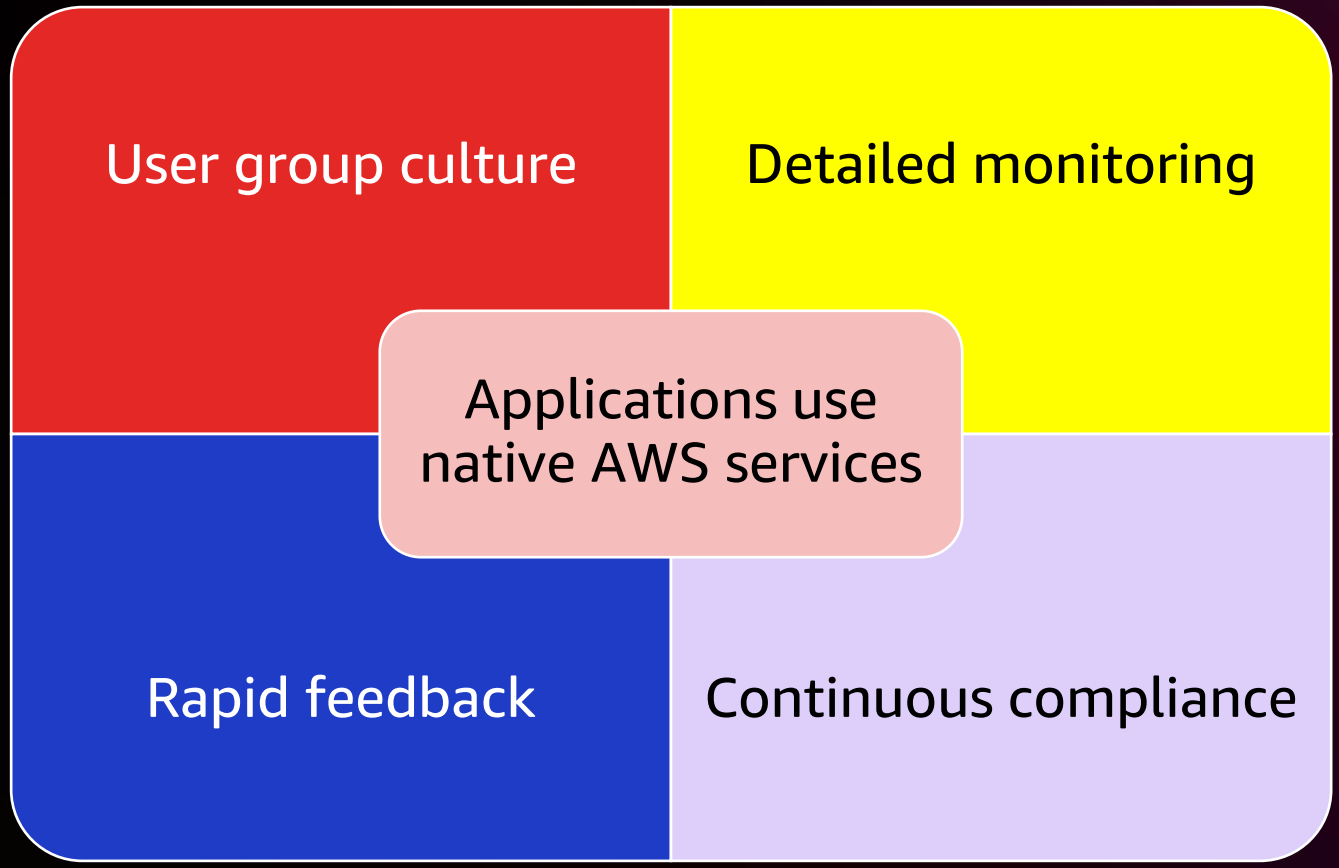
Human actions



Whether it’s natural, technical, or human, we’re managing risk

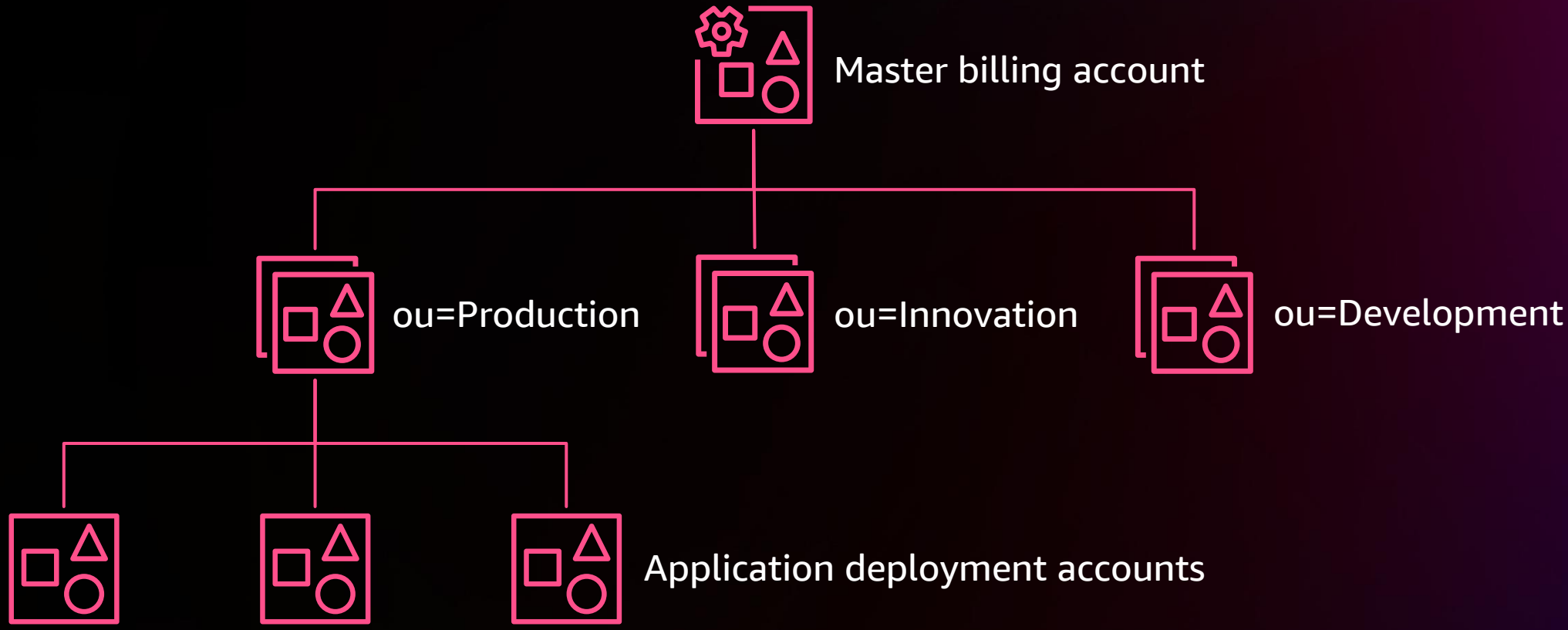
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The HSBC AWS team operating model



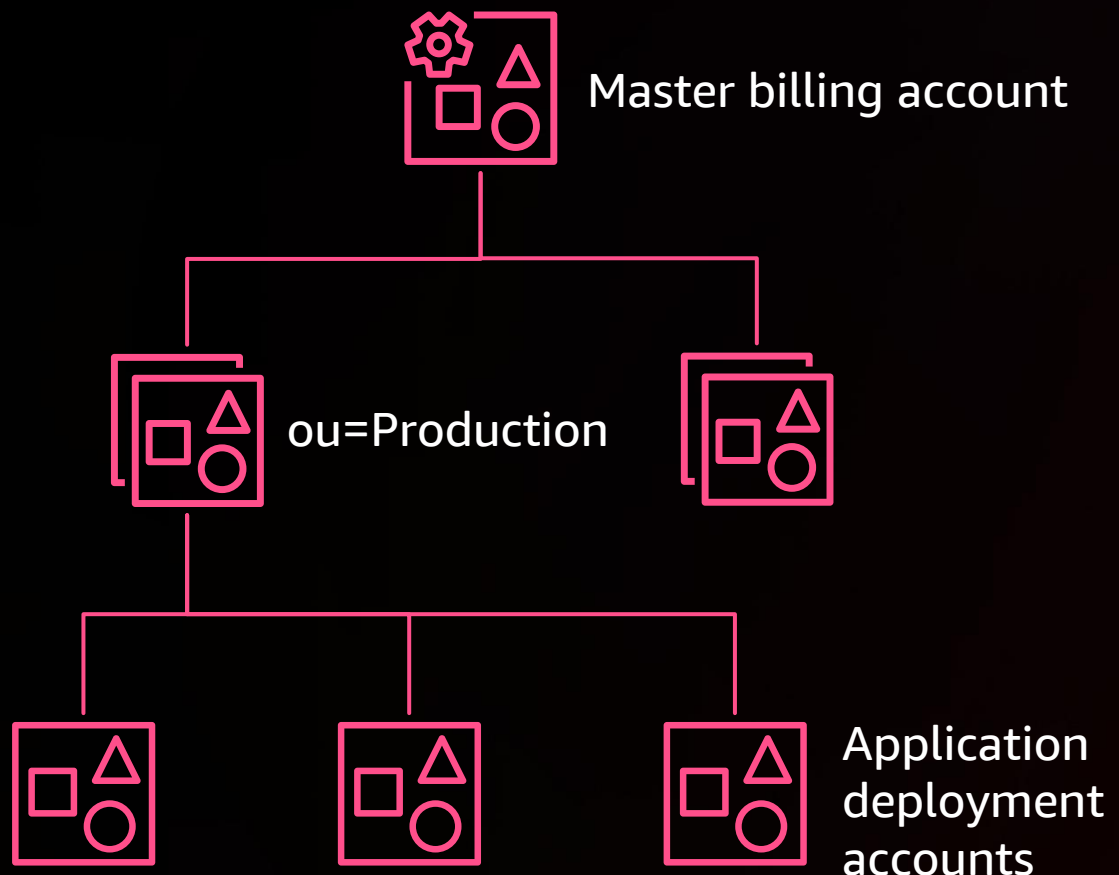
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Cost optimization: An organization view



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

Cost optimization: Organizational responsibilities



Central team

- Enterprise Discount Program
- Reserved Instances
- Savings plans
- Training and newsletters

Platform team

- Instance Scheduler
- Architecture review
- Landing zone and tooling optimization

Application team

- Optimization through service, OS, Region, and processor choices
- Continual refactoring with a cost lens

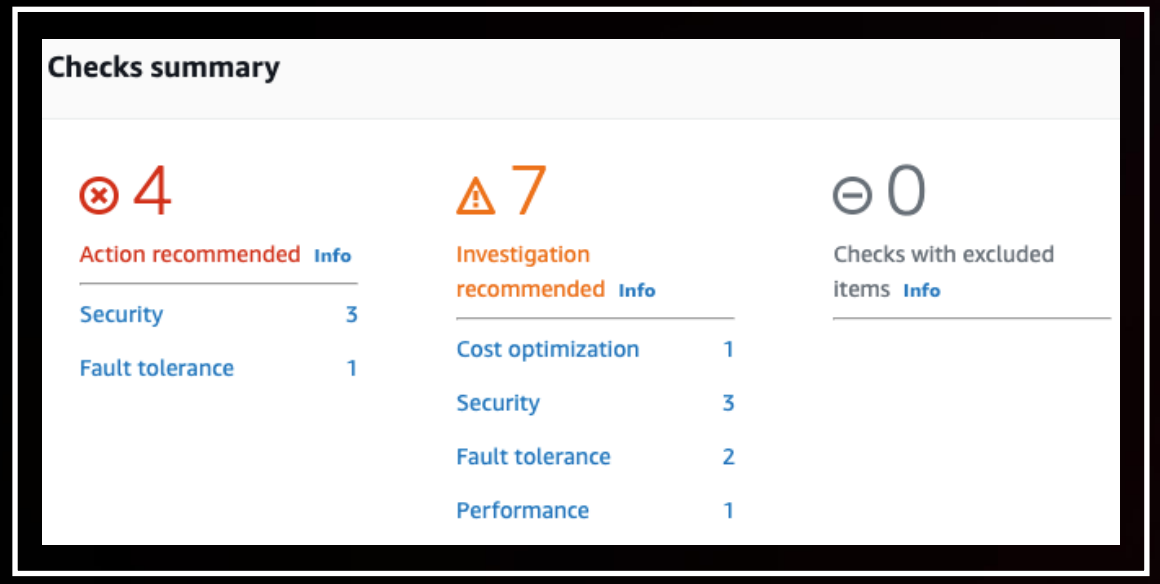
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

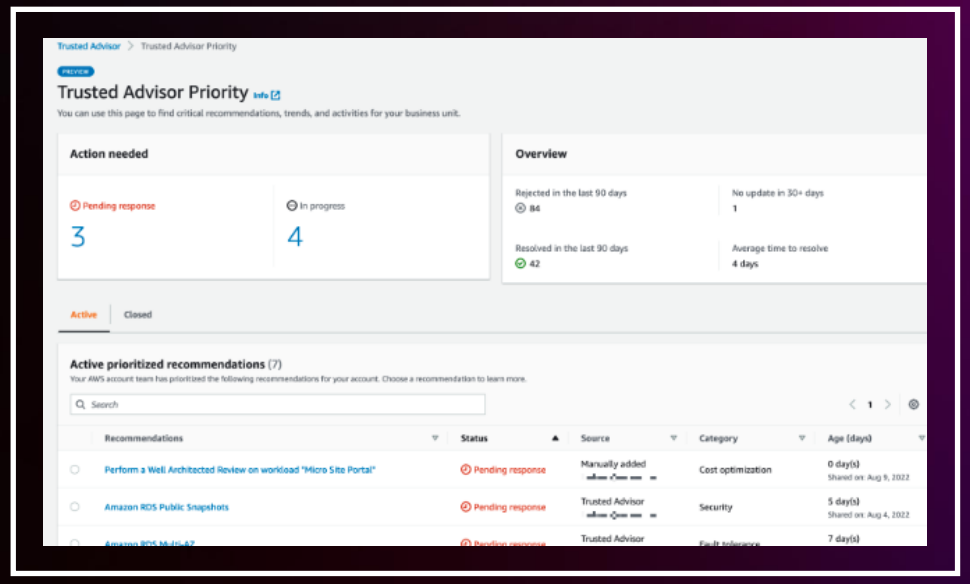
Cost optimization: Sources of data

	Sources of data
Central team Enterprise Discount Program Reserved Instances Savings plans Training and newsletters	AWS proactive services AWS Trusted Advisor Cloudability Billing console AWS account team
Platform team Instance Scheduler Architecture review Landing zone and tooling optimization	AWS blogs and service announcements Testing and development AWS TAM and SA partnership
Application team Optimization through service, OS, Region and processor choices Continual refactoring with a cost lens	Cost Optimization Workshops Near real-time cost analytics Blogs and monthly forum

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

Tools: Trusted Advisor and TA Priority





We encourage all our app teams to monitor Trusted Advisor

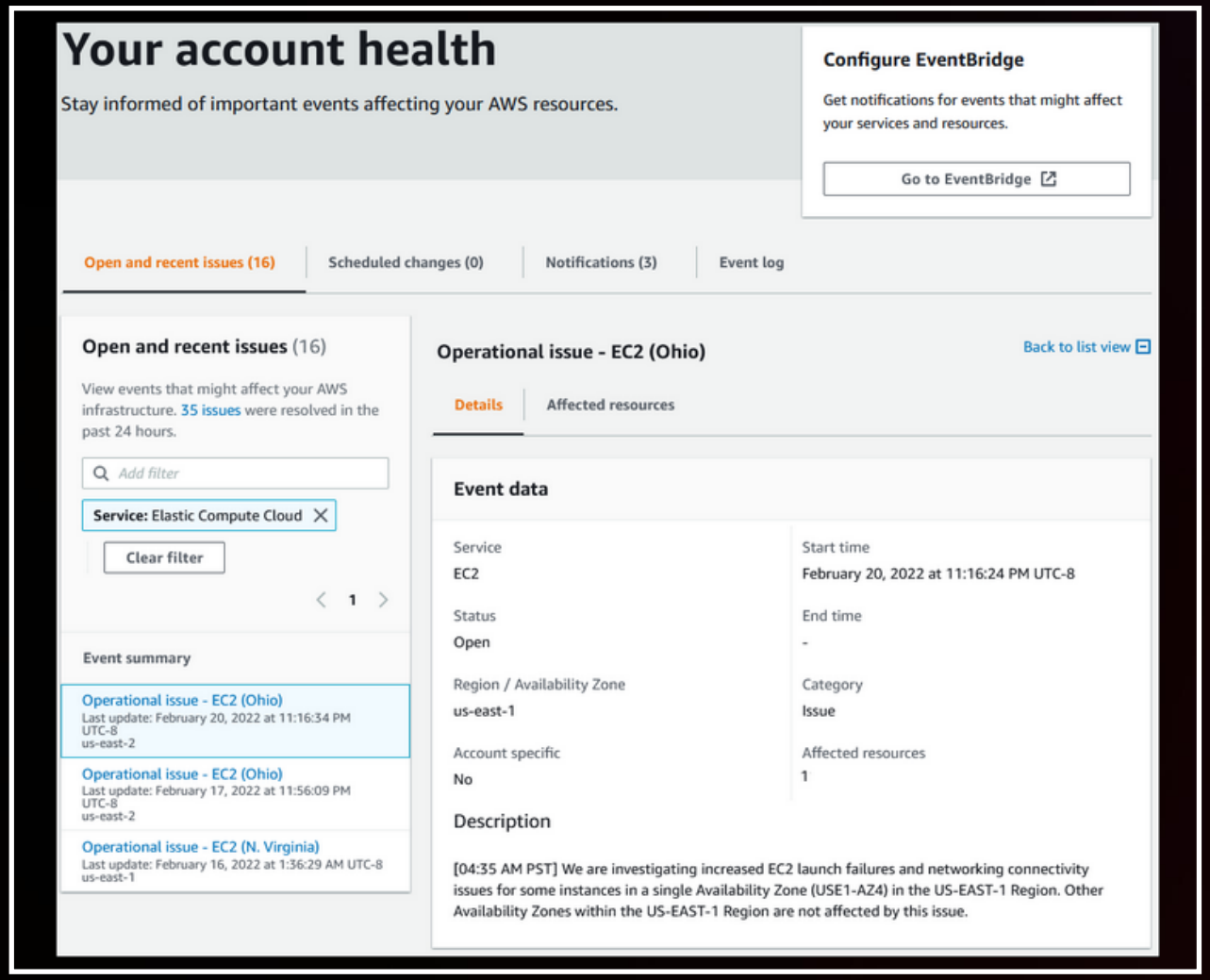
The platform team also monitor security and fault-tolerance checks centrally

We work with our AWS TAMs to share risk recommendations

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

Note: These screenshots were not taken from HSBC and are public images due to confidentiality

Tools: Personal Health Dashboard



Personal Health Dashboard communicates not only AWS service disruptions but also risks associated with technical debt

At HSBC, we link Personal Health Dashboard directly to incident management to rapidly communicate impact to our applications

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

Note: This screenshot was not taken from HSBC and is a public image due to confidentiality

Process: Well-Architected reviews



HSBC

As part of the go-live framework for every HSBC application, we have taken the AWS Well-Architected pillars and adapted the weighting of them to our environment and the business criticality of the application in question

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

Availability as a core architecture review pillar

Critical – Tier 0

Example: Customer internet banking SLA 99.9999%

Architect for high availability first; this may include multi-AZs and multi-Regions with sophisticated load balancing and failover

Continually cost-optimize but always with availability objective in mind



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

Availability as a core architecture review pillar



Medium – Tier 2

Example: Staff holiday booking system SLA 99.9%

Architect with a greater weight on the cost pillar; this may include making availability trade-offs to reduce costs

Failover models may be pilot light- or cold start-based, still exceeding an availability SLA but balancing this with cost targets

aws

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

Identifying cost-optimization opportunities

HSBC HAS INVESTED IN AWS ENTERPRISE SUPPORT

We have worked with our AWS technical account manager and proactive services

Cost-optimization reviews identified 20% cost-savings opportunities:

- Savings plans & RIs
- AWS Graviton for managed services
- CPU type

HSBC identified that dev/UAT workloads were being left switched on permanently



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

https://aws.amazon.com/premiumsupport/technology-and-programs/proactive-services/

Centralized tools: Reserved Instances

REAL SAVINGS FIGURES FROM HSBC 2022 YTD

Amazon EC2	Amazon ElastiCache	Amazon Redshift	Amazon RDS
			
35% savings	**30% savings**	**27% savings**	**28% savings**

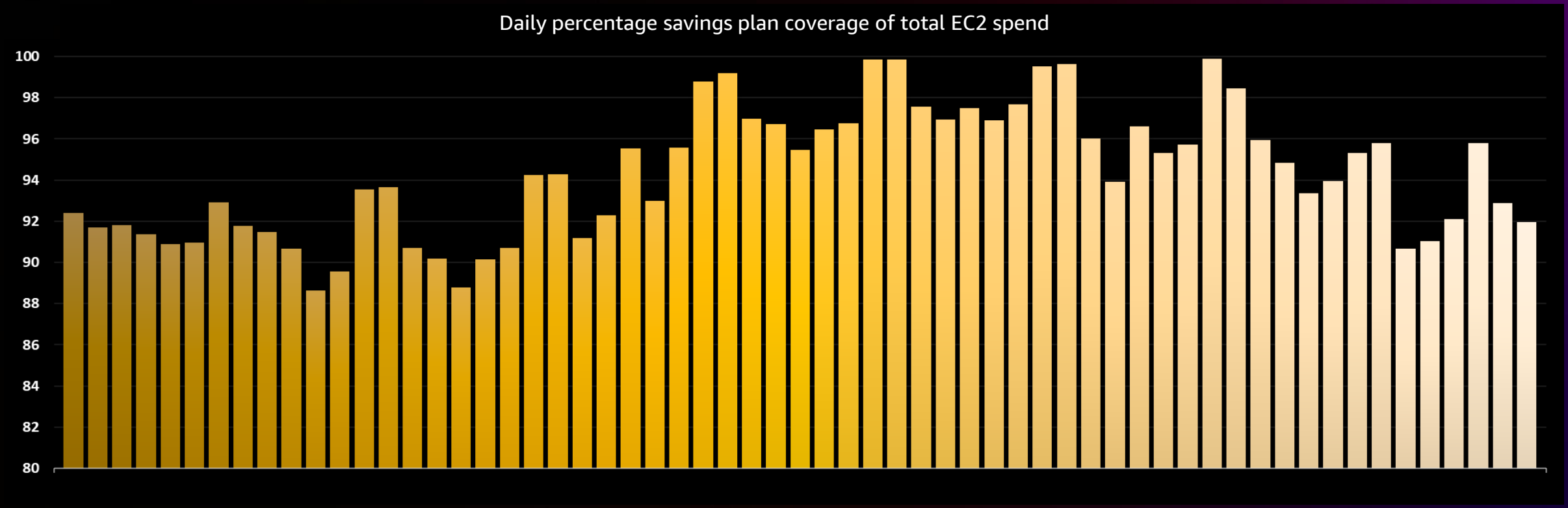
Mix commitment terms based on workload

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

Source: HSBC data, Year to Date 2022

Centralized tools: Savings plans

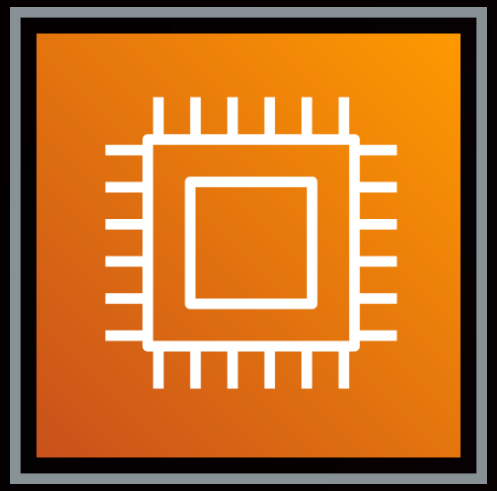
- Over several months we increased savings plans coverage to an average > 90%
- Data is reviewed monthly to purchase capacity in line with growth



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

Source: HSBC data, 61-day period in 2022, peak equals 100% coverage

Platform tools: Platform optimization



AWS Lambda



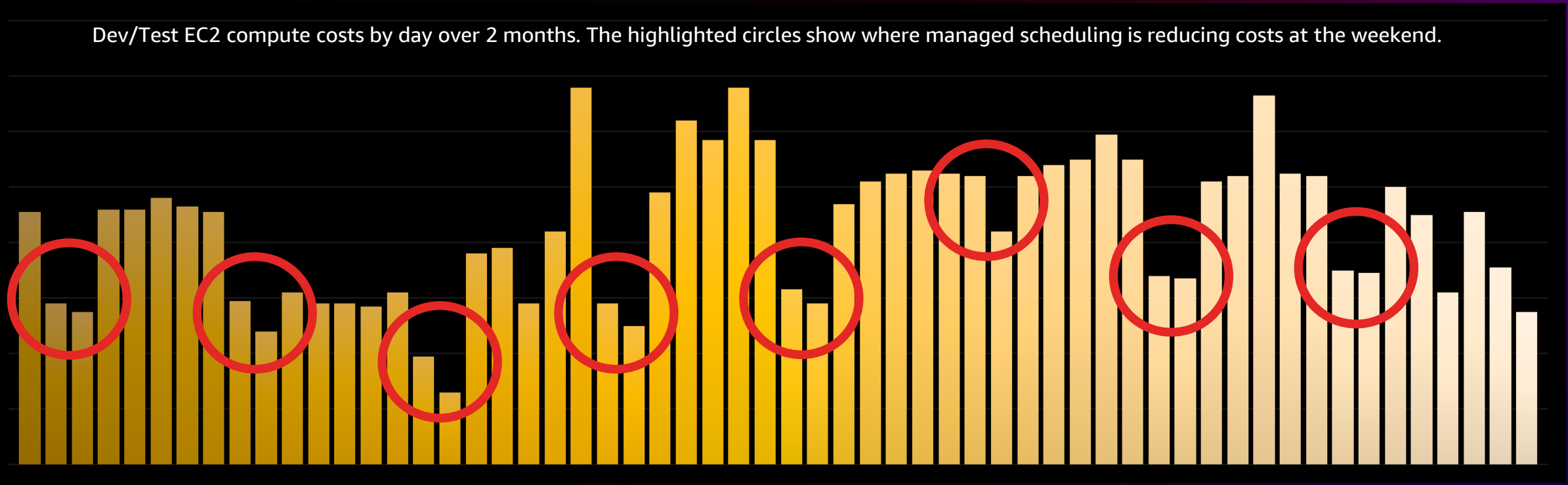
AWS Graviton

- HSBC runs over 100,000 serverless functions within the core platform
  By moving functions to Graviton we achieved a 20% to 30% savings
- Analyzing top 10 most expensive functions using AWS Lambda Power Tuning and then rewriting them for cost efficiency allowed us to achieve a further 20% savings

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

Platform tools: Instance Scheduler

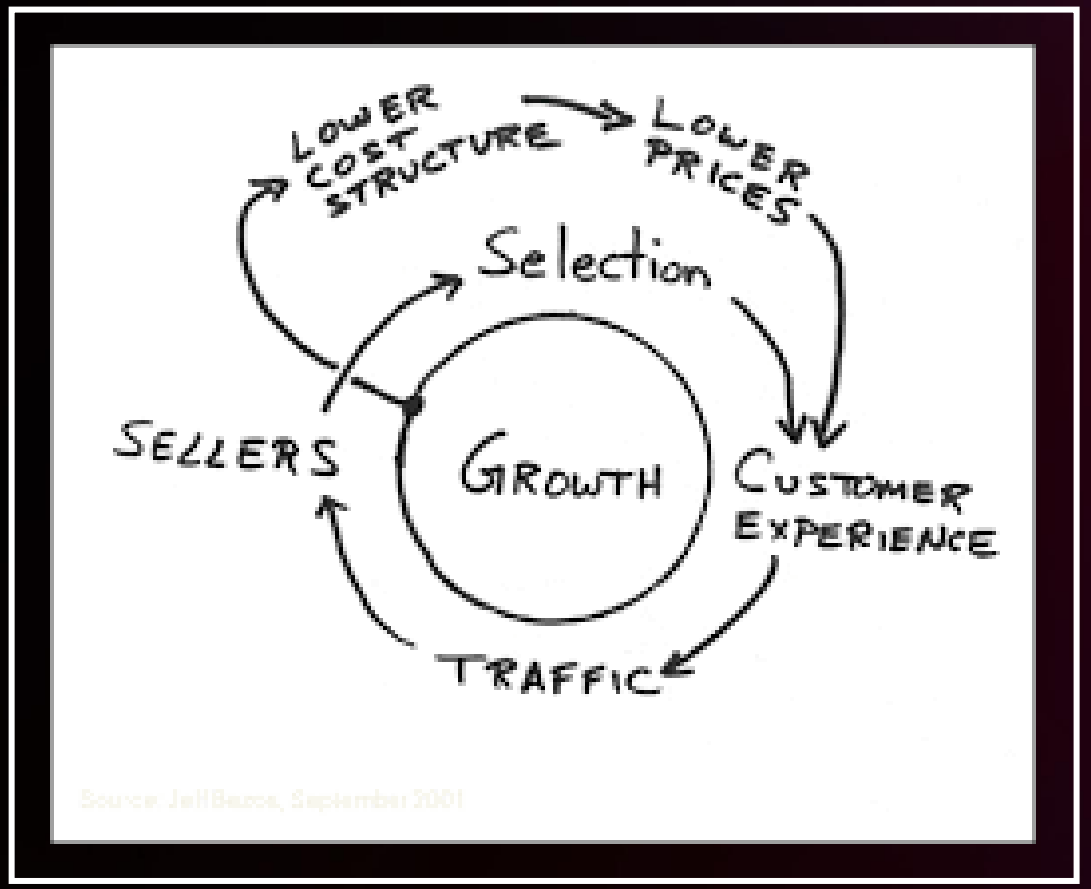
- The platform team runs AWS Instance Scheduler and Cloud Custodian to shut down EC2 compute and other resources at the weekend, unless customers opt out



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

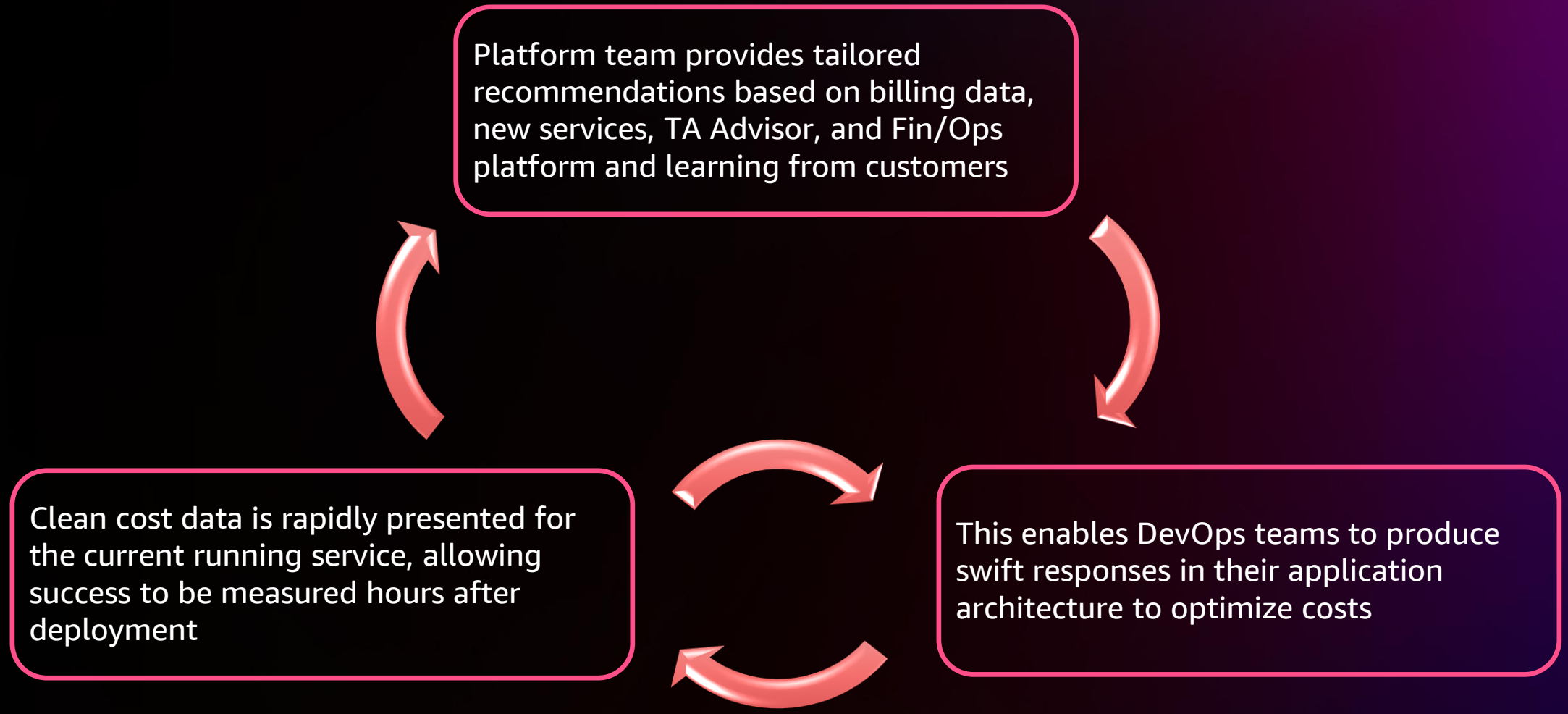
Source: HSBC data, 58-day period in 2022

Decentralization: Learning from Amazon.com



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

Decentralization: Fast feedback loops



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

Example: AWS Pricing Calculator

Processor choices make a difference for compute. AMD backed instances can offer an 8%–12% savings for the same compute power over Intel.

Prices vary by Region too, sometimes dramatically. AMD backed instances in the AWS India Region are currently c. 40% cheaper than US-East-1; pricing for a m6a.4xl instance below was taken from the AWS Pricing Calculator.

Monthly cost	Description	Region
217.67 USD	-	Asia Pacific (Mumbai)
336.78 USD	-	US East (N. Virginia)

Even if your production application is running in a specific geography, it may make commercial sense to run development and build tools in a Region with the lower costs.

aws

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

Decentralization: Creating an optimization culture

1. Educate teams on managing costs within the cloud. Enable anomaly detection and alarms as mistakes will happen.
2. Create communication channels. Allow app teams to communicate their enhancements.
3. Constantly refactor applications with cost and availability in mind.
4. Focus on the 0.01% improvements – scale makes small changes matter.
5. Focus on the Well-Architected pillars, but change emphasis by type of application.

aws

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

Call to action

- Trusted Advisor Priority – review and take actions on recommendations
- AWS Well-Architected Framework – review and evaluate workloads
- AWS re:Post – engage in community conversations
- AWS Support app in Slack – request assistance
- Integrate with AWS Health for alerts on operational events

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

Thank you!

Ram Cherukuri
cherukur@amazon.com

Kartik Cheekoty
cheekoty@amazon.com

Alistair McLaurin
alistair.james.mclaurin@hsbc.com



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

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