AWS re:Invent

STG401

Manage objects and optimize cost with Amazon S3 and Amazon S3 Glacier

Dave Rocamora
Solutions Architect
Amazon Web Services

Peter Schmiedeskamp
Analyst
Amazon Web Services

AWS re:Invent

aws

Agenda

Introduction

Amazon S3 operations and the S3 API

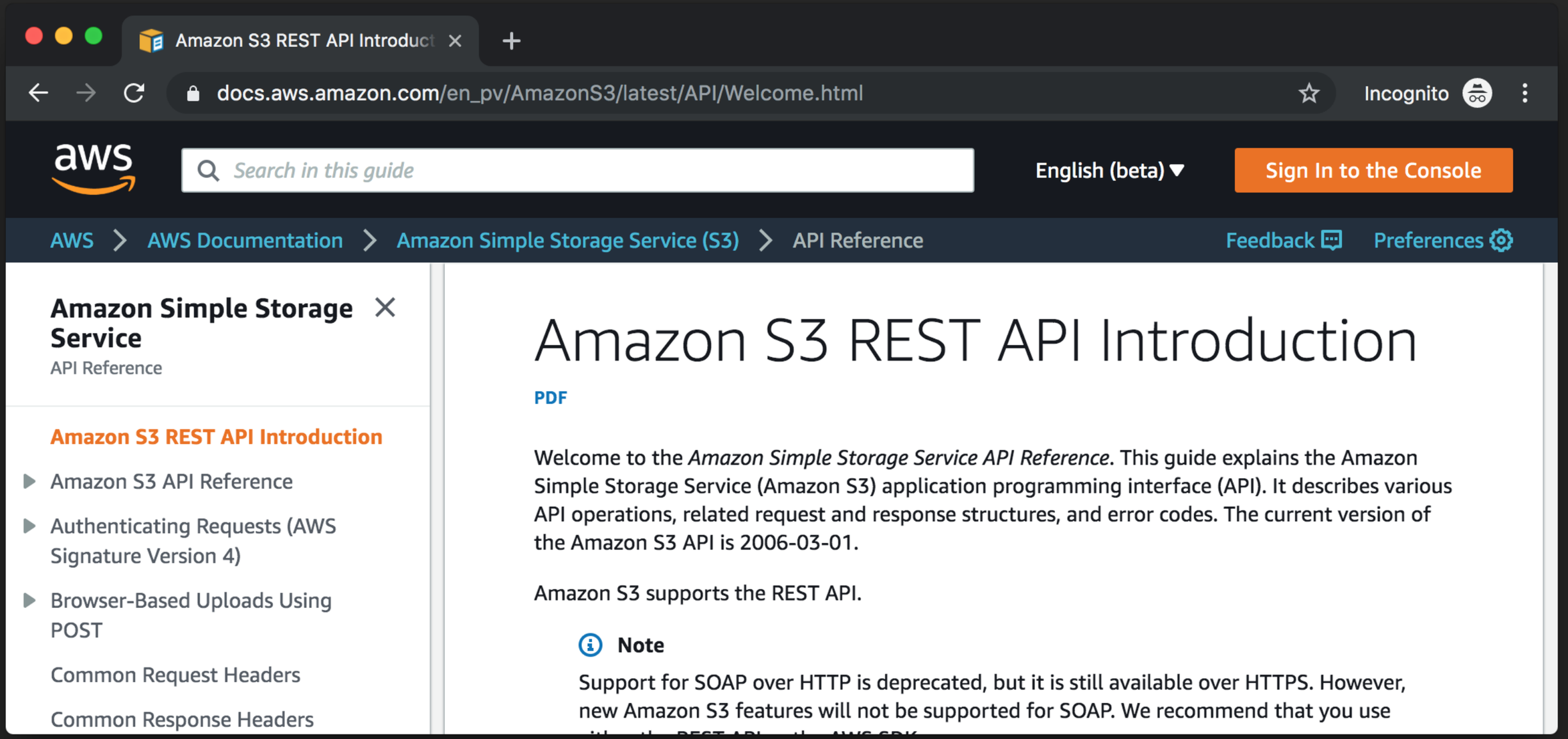
Right-sizing objects

Multi-account considerations

How this is going to work

- We are providing an AWS account and an Amazon EC2 instance to do this workshop.
- You will use Amazon EC2 Instance Connect or SSH to access the instance.
- This workshop is interactive. We present an exercise, you do it, we discuss.
- Support staff is here to help.

Amazon S3 is an API



Costs to consider

Storage

Requests

Storage management

Data transfer

S3 transfer acceleration

Replication

Your time

Compute time

Costs to consider

Storage

Requests

Storage management

Data transfer

S3 transfer acceleration

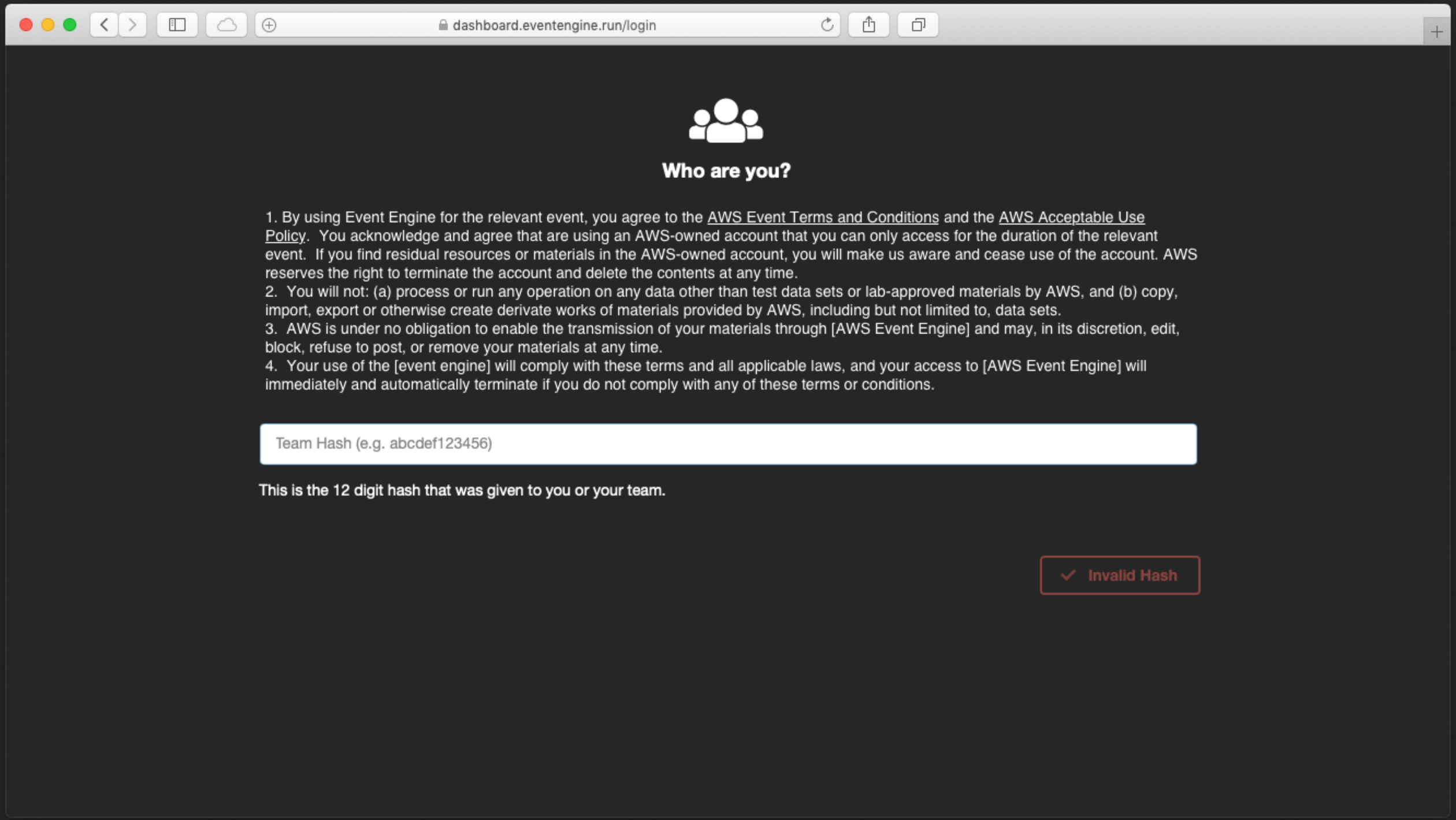
Replication

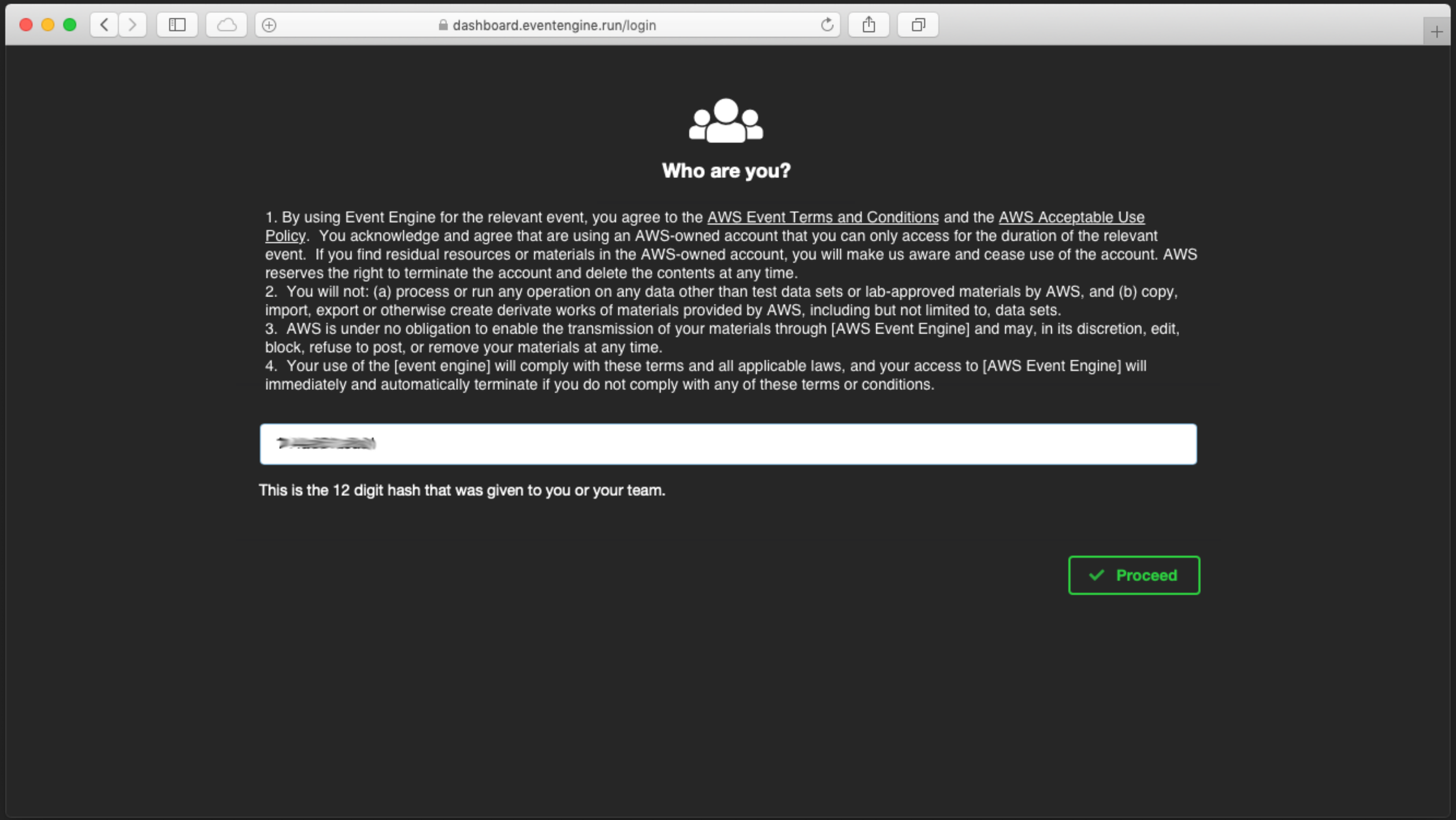
Your time

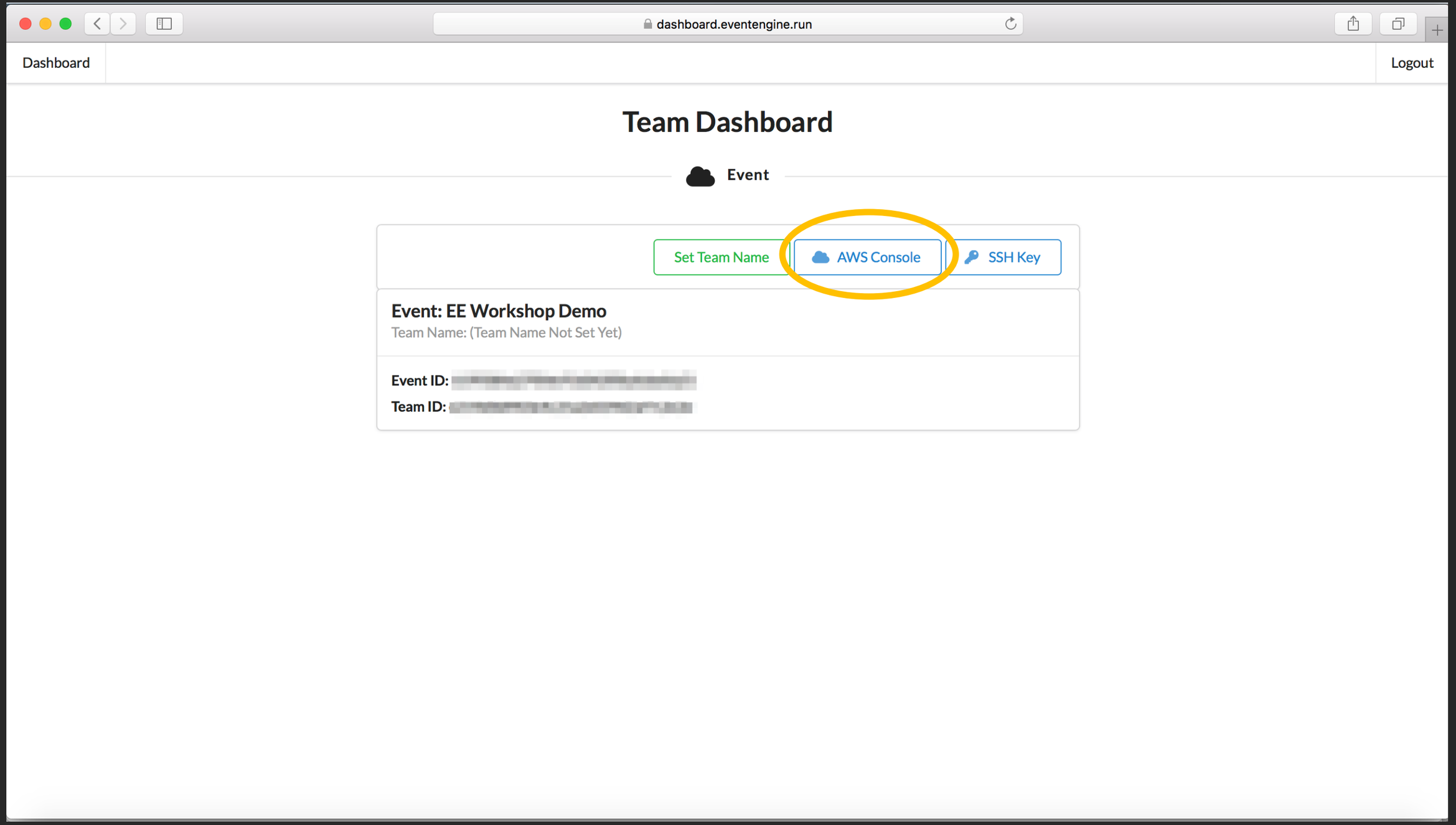
Compute time

Workshop AWS account

https://dashboard.eventengine.run


dashboard.eventengine.run/login
Who are you?
1. By using Event Engine for the relevant event, you agree to the AWS Event Terms and Conditions and the AWS Acceptable Use Policy. You acknowledge and agree that are using an AWS-owned account that you can only access for the duration of the relevant event. If you find residual resources or materials in the AWS-owned account, you will make us aware and cease use of the account. AWS reserves the right to terminate the account and delete the contents at any time.
2. You will not: (a) process or run any operation on any data other than test data sets or lab-approved materials by AWS, and (b) copy, import, export or otherwise create derivate works of materials provided by AWS, including but not limited to, data sets.
3. AWS is under no obligation to enable the transmission of your materials through [AWS Event Engine] and may, in its discretion, edit, block, refuse to post, or remove your materials at any time.
4. Your use of the [event engine] will comply with these terms and all applicable laws, and your access to [AWS Event Engine] will immediately and automatically terminate if you do not comply with any of these terms or conditions.
Team Hash (e.g. abcdef123456)
This is the 12 digit hash that was given to you or your team.
Invalid Hash


dashboard.eventengine.run/login
Who are you?
1. By using Event Engine for the relevant event, you agree to the AWS Event Terms and Conditions and the AWS Acceptable Use Policy. You acknowledge and agree that are using an AWS-owned account that you can only access for the duration of the relevant event. If you find residual resources or materials in the AWS-owned account, you will make us aware and cease use of the account. AWS reserves the right to terminate the account and delete the contents at any time.
2. You will not: (a) process or run any operation on any data other than test data sets or lab-approved materials by AWS, and (b) copy, import, export or otherwise create derivate works of materials provided by AWS, including but not limited to, data sets.
3. AWS is under no obligation to enable the transmission of your materials through [AWS Event Engine] and may, in its discretion, edit, block, refuse to post, or remove your materials at any time.
4. Your use of the [event engine] will comply with these terms and all applicable laws, and your access to [AWS Event Engine] will immediately and automatically terminate if you do not comply with any of these terms or conditions.
This is the 12 digit hash that was given to you or your team.
Proceed


dashboard.eventengine.run
Dashboard
Logout
Team Dashboard
Event
Set Team Name
AWS Console
SSH Key
Event: EE Workshop Demo
Team Name: (Team Name Not Set Yet)
Event ID:
Team ID:

AWS Console Login

Remember to only use "[redacted]" as your region, unless otherwise directed by the event operator.

Login Link

Open AWS Console | Copy Login Link

Credentials / CLI Snippets

Mac / Linux | Windows

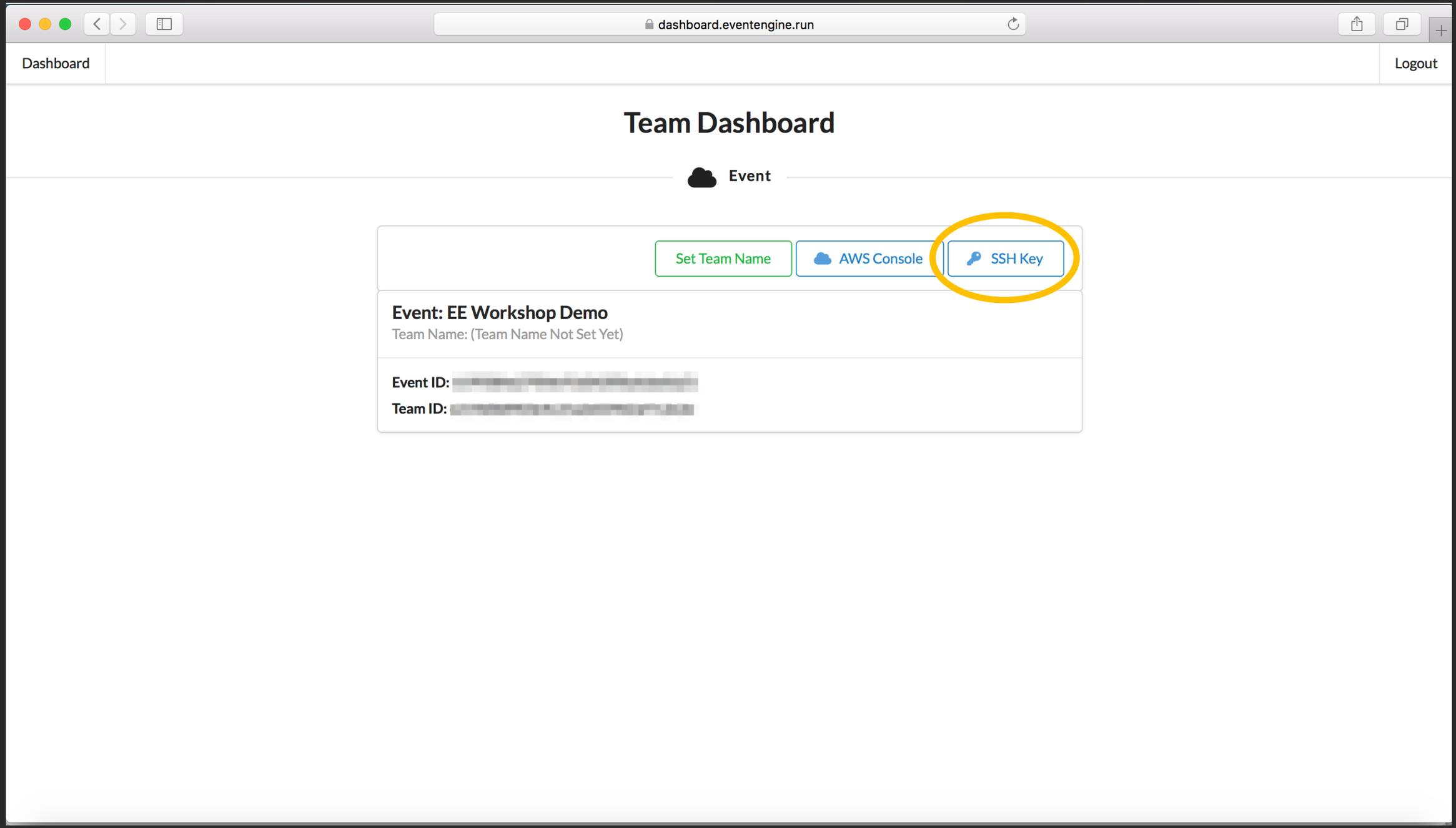
Mac or Linux

```
export AWS_DEFAULT_REGION=[redacted]
export AWS_ACCESS_KEY_ID=[redacted]
export AWS_SECRET_ACCESS_KEY=[redacted]
export AWS_SESSION_TOKEN=[redacted]
```

How do I use the AWS CLI?

Checkout the AWS CLI documentation here: https://docs.aws.amazon.com/cli/latest/userguide/cli-chap-welcome.html

OK


dashboard.eventengine.run
Dashboard
Logout
Team Dashboard
Event
Set Team Name
AWS Console
SSH Key
Event: EE Workshop Demo
Team Name: (Team Name Not Set Yet)
Event ID:
Team ID:

Amazon S3 operations and the S3 API

AWS re:Invent

aws

Exercise 1: Getting and putting objects

- Get this object: `$SRC/readme.md`
- Take a look at it (`more, less, cat`)
- Put it here: `$DEST/readme.md`
- Copy `$SRC/bigobject` to `$DEST/bigobject`

Exercise 2: Copying a few objects

- There are 1,500 objects at `$SRC/challenge_1/`
- Copy them to `$DEST/challenge_1/`

Exercise 3: Exploring ListObjectsV2

- Use the `list-bench` program to explore listing objects in the $SRC and $DEST bucket
- Look at some suggestions in the README

Exercise 4: Copying a LOT of objects

- There are 1,000,000 objects at $SRC/challenge_2/
- Copy them to $DEST
- There is a manifest of the source objects at $SRC/manifests/manifest-us-east-1.csv

Exercise 5: Using other storage classes

- Put readme.md at `$DEST/readme.md.s3ia` and make sure it is stored in Infrequent Access
- Put readme.md at `$DEST/readme.md.glacier` and make sure it is stored in Amazon S3 Glacier
- Restore `$DEST/readme.md.glacier`

Right-sizing objects

What is the right object size?

	Small Objects	Big Objects
Organization	Built-in	Depends on the format
Cost	More requests, per object fees, minimum object size fees	More data transfer (maybe)
Performance	More small requests	Larger requests, wasted data transfer (maybe)

Exercise 6: Combine 3 objects

- `$SRC/challenge_3` contains 3 objects
- Concatenate them into one object at `$DEST/combined`

Exercise 7: Download part of the combined object

- Download the 2nd part of the `$DEST/combined` object
- Download a range of bytes from `$DEST/combined`
- The range is in the README

Exercise 8: Using S3 Select

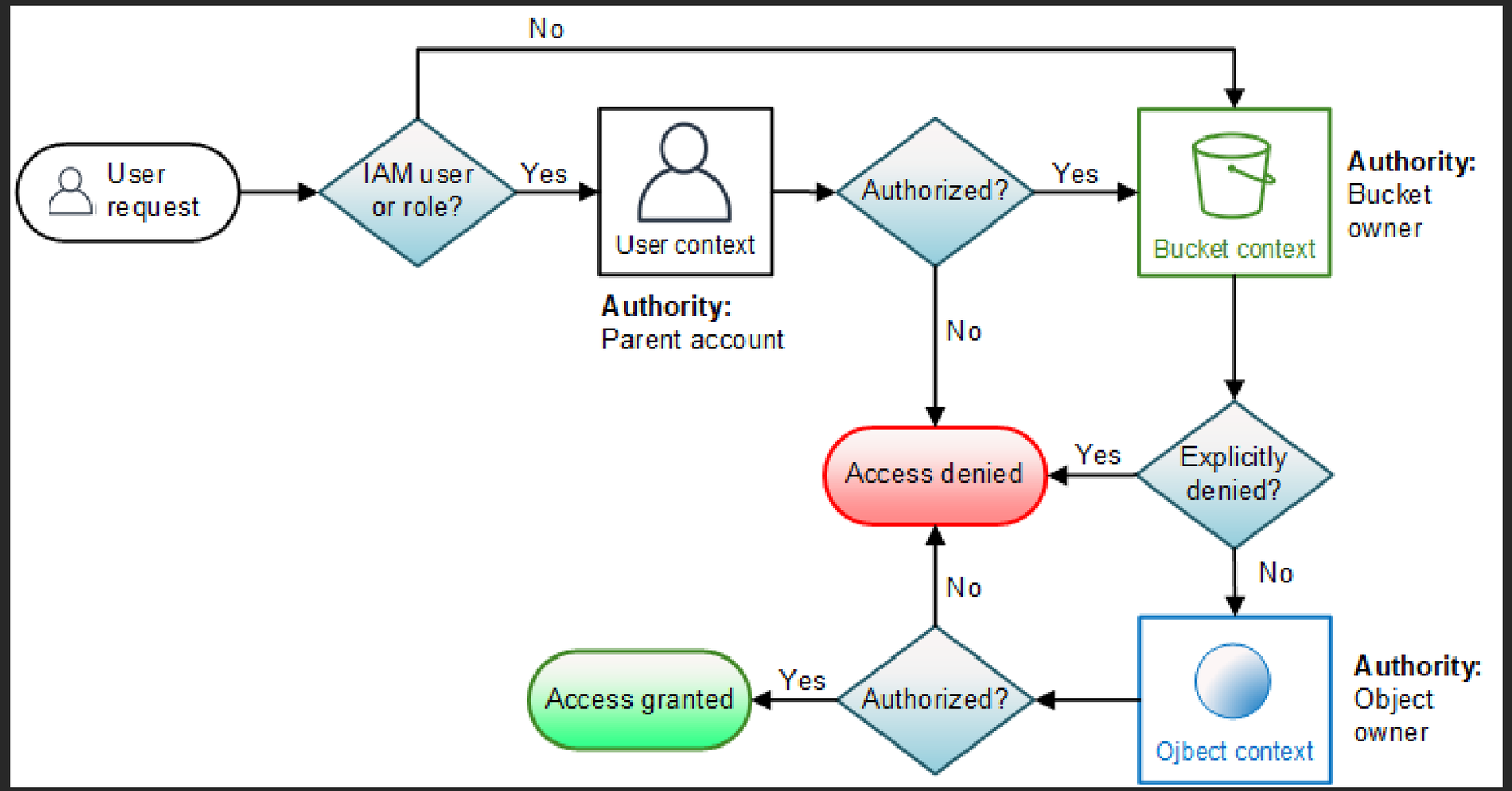
- Get the results for sensors in Nevada from the combined object
- Get the average FunFactor for all sensors in Washington state from the combined object
- Use the s3-select script to make this easier

Multi-account considerations

Exercise 9: PutObject from another account

- There are 2 AWS accounts (A and B) that want to work with you.
- Update the bucket policy on your S3 bucket to allow account A to put objects into your bucket and allow account B to read objects from your bucket.
- There is a policy template in the README. You will have to update it with account IDs for A and B.

How Amazon S3 authorizes a request



Exercise 10: Lock down your bucket

- Make sure there are no public objects in your bucket

Related breakouts

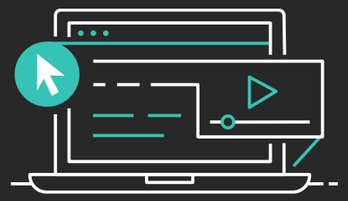
STG315 – Amazon S3 security settings & controls

WPS319 – Best practices for working with large-scale geospatial data

STG329 – Maximizing storage throughput and performance

Learn storage with AWS Training and Certification

Resources created by the experts at AWS to help you build cloud storage skills



45+ free digital courses cover topics related to cloud storage, including:

- Amazon S3
- AWS Storage Gateway
- Amazon S3 Glacier
- Amazon Elastic File Storage (Amazon EFS)
- Amazon Elastic Block Storage (Amazon EBS)

Classroom offerings, like Architecting on AWS, feature AWS expert instructors and hands-on activities

Visit aws.amazon.com/training/path-storage/

aws training and certification

Thank you!

Dave Rocamora
rocamora@amazon.com

Peter Schmiedeskamp
pws@amazon.com

AWS re:Invent

aws

Please complete the session survey in the mobile app.