INNOVATE

AI/ML EDITION

24 February 2022

Setting up secure & well-governed ML environments on AWS

Michael Stringer
Senior Domain Solutions Architect - Security, AWS

Agenda

Agenda

1. Six key security considerations when deploying Machine Learning workloads on AWS

Agenda

1. Security considerations when deploying Machine Learning workloads on AWS
2. Overview of fundamental services and configuration settings to use in your environment

Agenda

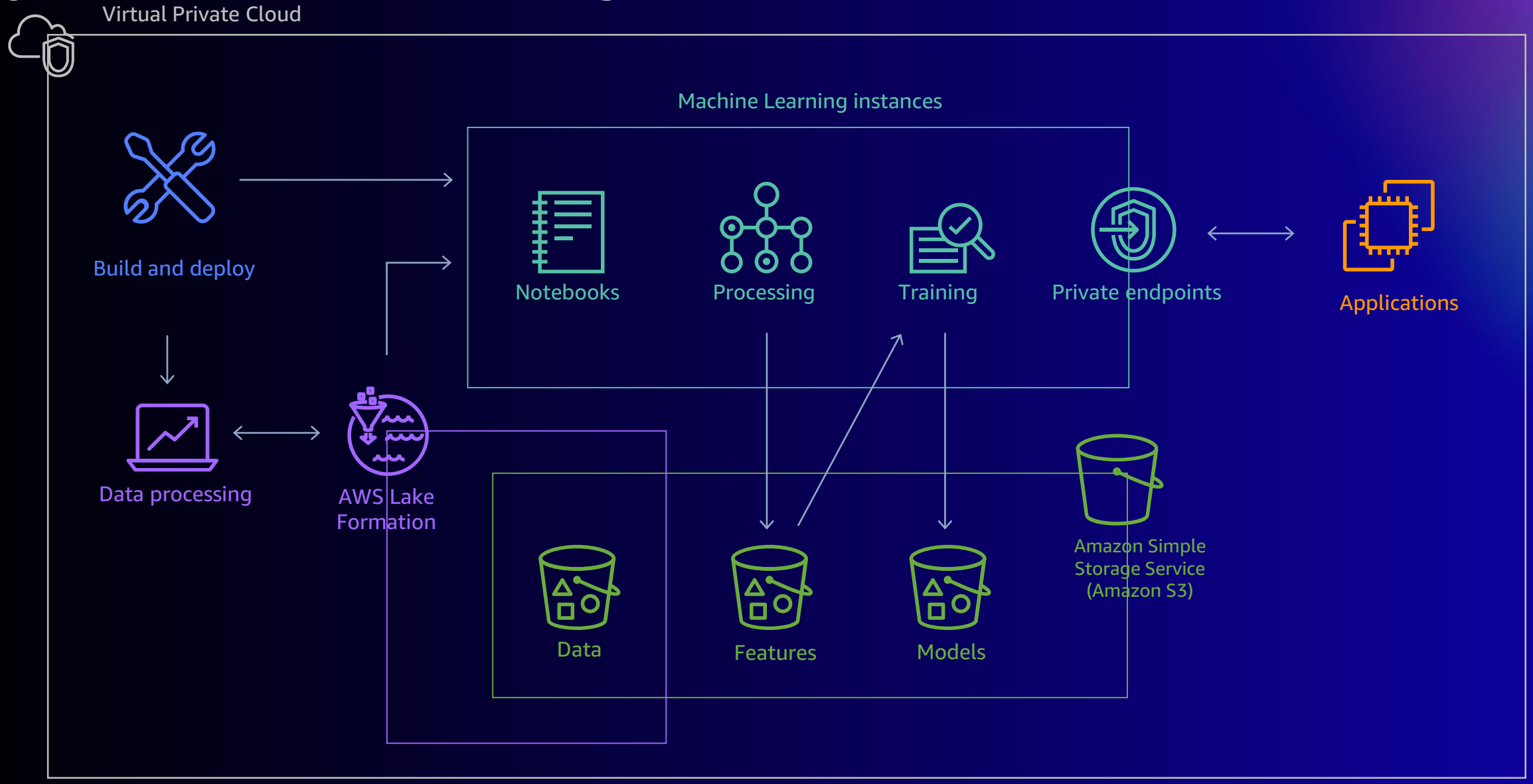
1. Security considerations when deploying Machine Learning workloads on AWS
2. Overview of fundamental services and configuration settings to use in your environment
3. Demonstration – where to find and configure security controls in the AWS Console

Agenda

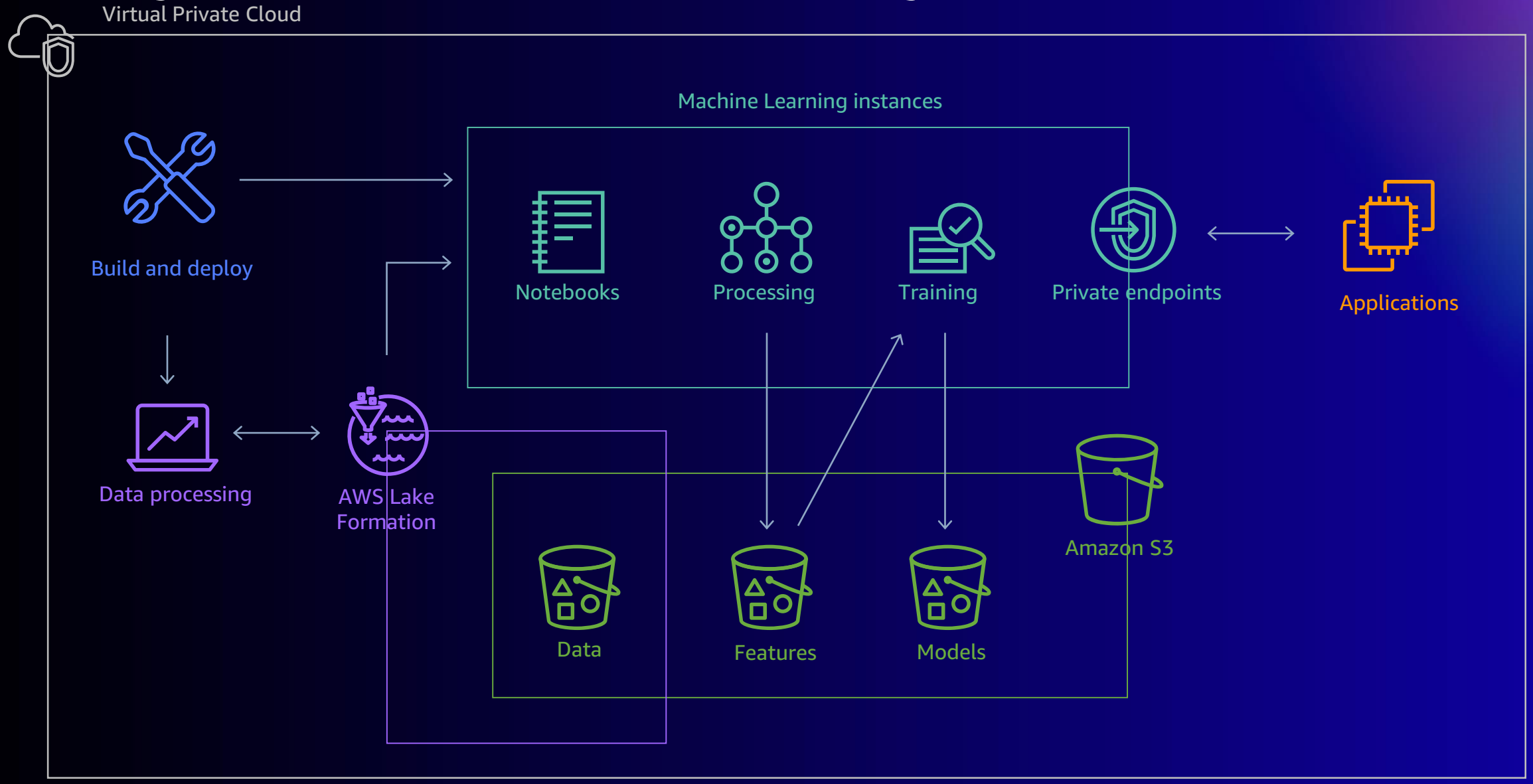
1. Security considerations when deploying Machine Learning workloads on AWS
2. Overview of fundamental services and configuration settings to use in your environment
3. Demonstration – where to find and configure security controls in the AWS Console
4. Wrap-up of the presentation / links to resources for further reading

Security considerations

Typical Machine Learning architecture

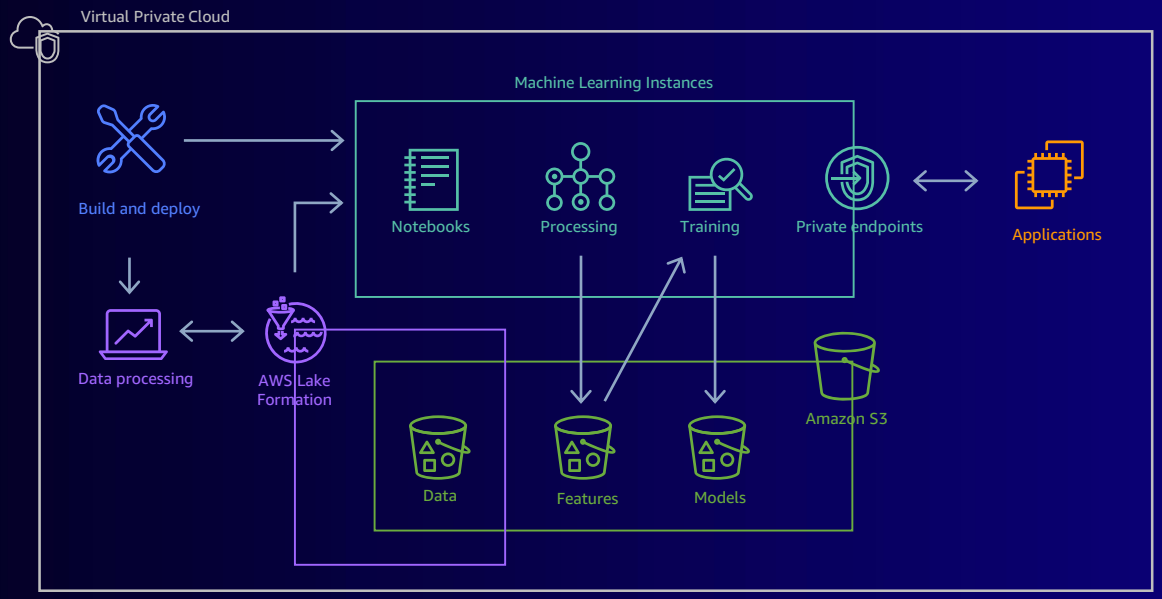


Security considerations for ML deployments



Security considerations for ML deployments

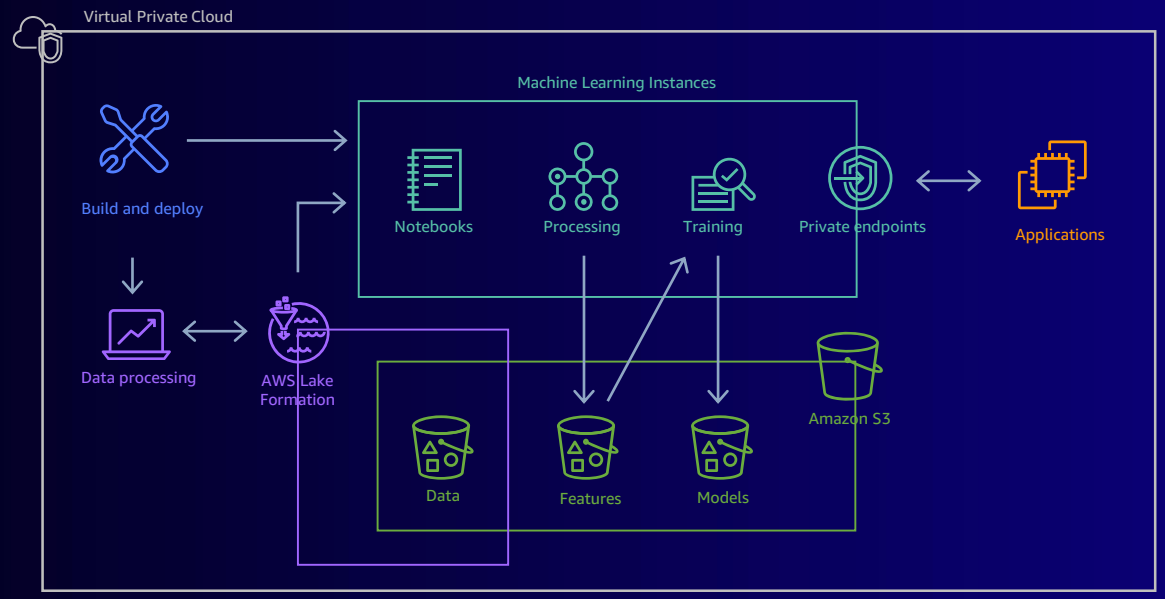
AWS account structure



Security considerations for ML Deployments

Logging and auditing

AWS account structure

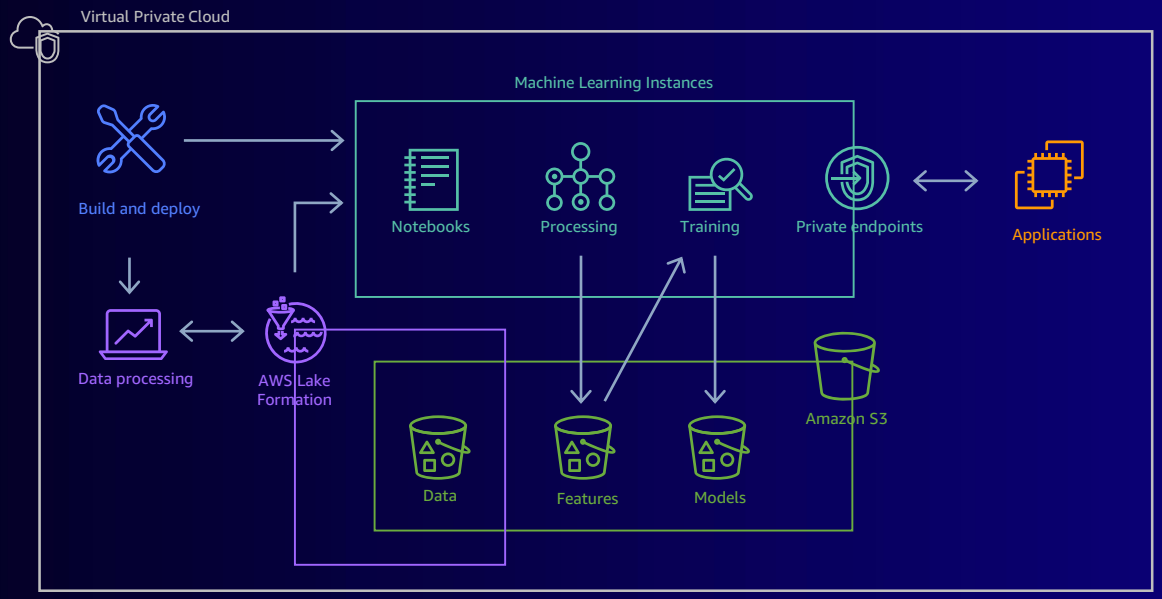


Security considerations for ML deployments

Logging and auditing

Identity and authorization

AWS account structure



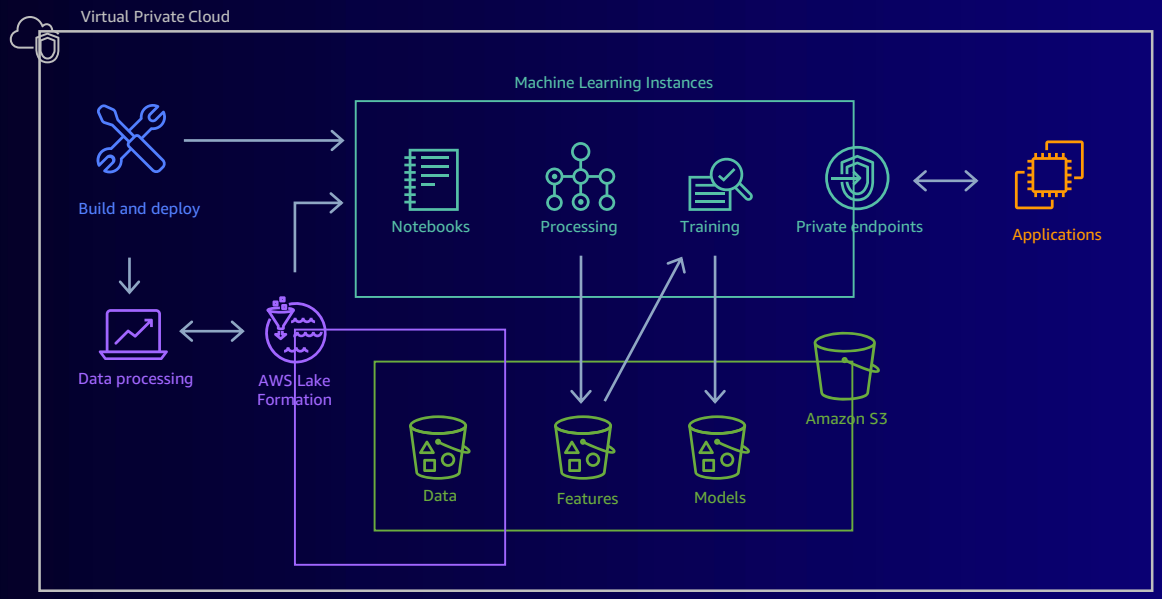
Security considerations for ML deployments

Logging and auditing

Identity and authorization

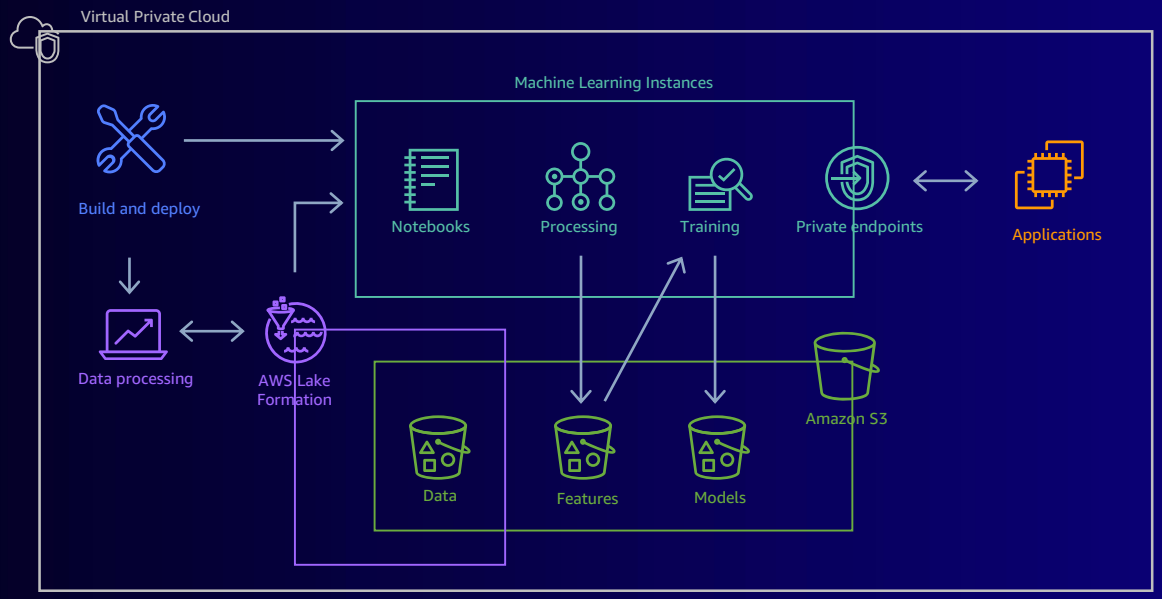
Network controls

AWS account structure



Security considerations for ML deployments

AWS account structure



Logging and auditing

Identity and authorization

Network controls

Encryption

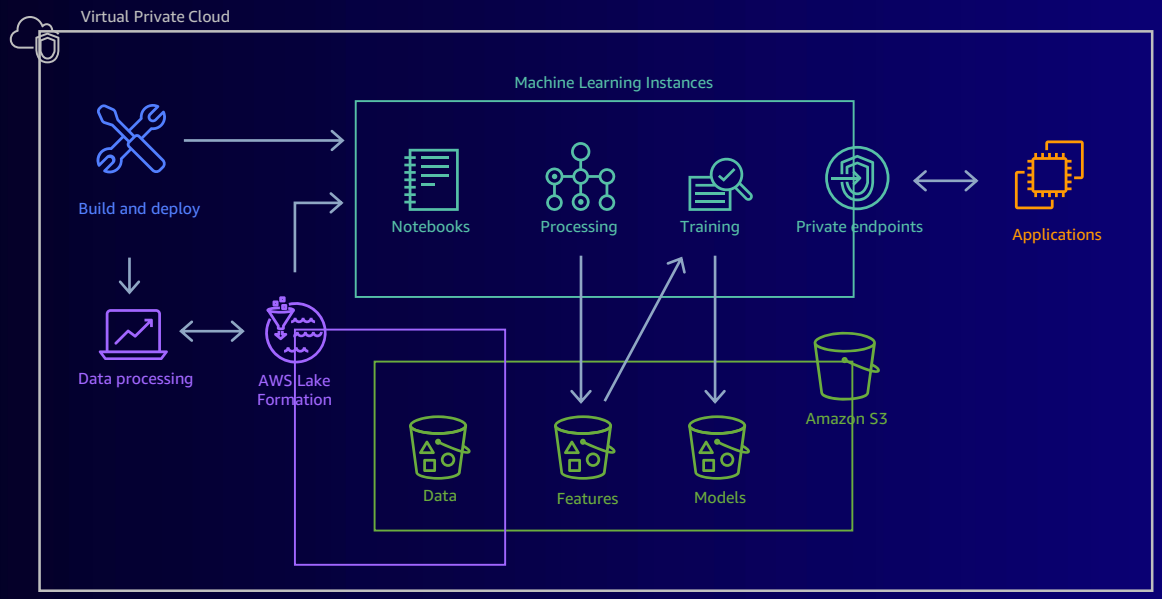
Security considerations for ML deployments

AWS account structure

Logging and auditing

Identity and authorization

Network controls



Encryption

User access

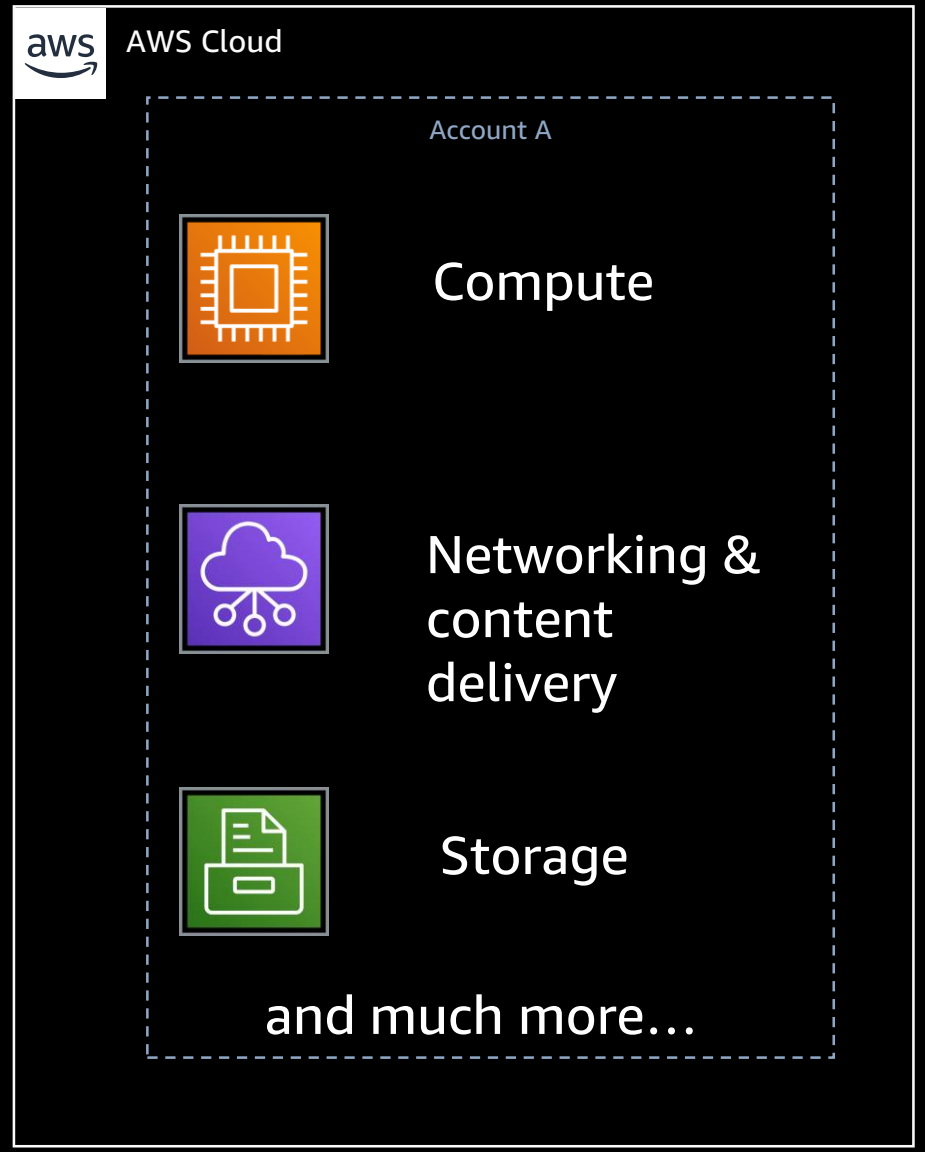
AWS account structure and guardrails

What is an AWS account?

Each AWS account:

- Is a resource container for AWS cloud services
- Is an explicit security boundary
- Is a container for cost tracking and billing
- Is a mechanism to enforce limits and thresholds
    - e.g. Service quotas and API thresholds

- Over time, organizations add more accounts to support more applications and services



Scaling to a multi-account model

Many teams

Rapid innovation with resources provisioned quickly and exclusively for each team

Billing

Simplify billing where resources used within an AWS account can be allocated to the business unit that is responsible for that account

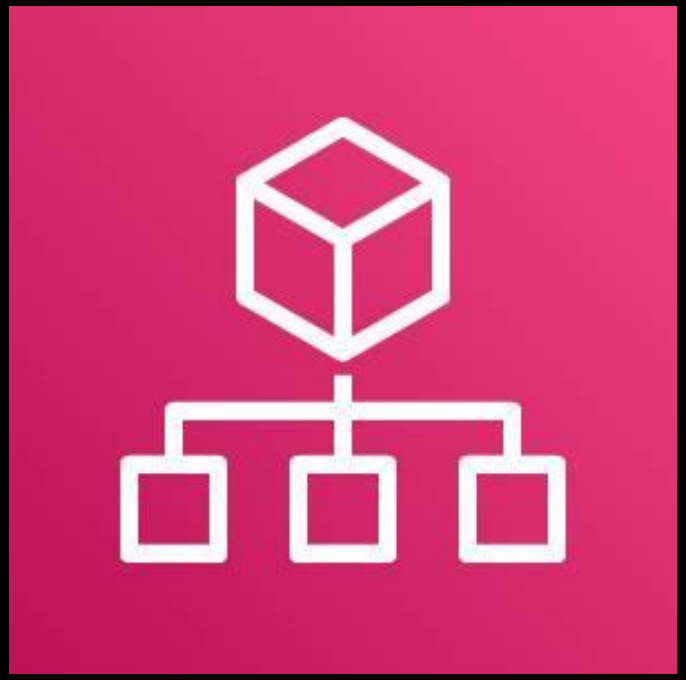
Business process

Organize AWS accounts to reflect business processes with different operational, regulatory, and budgetary requirements

Isolation & security

Tight security boundaries enforced by built-in isolation between accounts, and consolidation for workloads with similar risk profiles

AWS Organizations



Provides you tools to centrally govern and manage your cloud environment

- Quickly scale by creating accounts and allocate resources
- Customize your environment by applying governance policies
- Secure and audit your environment
- Manage costs and identify cost-saving measures

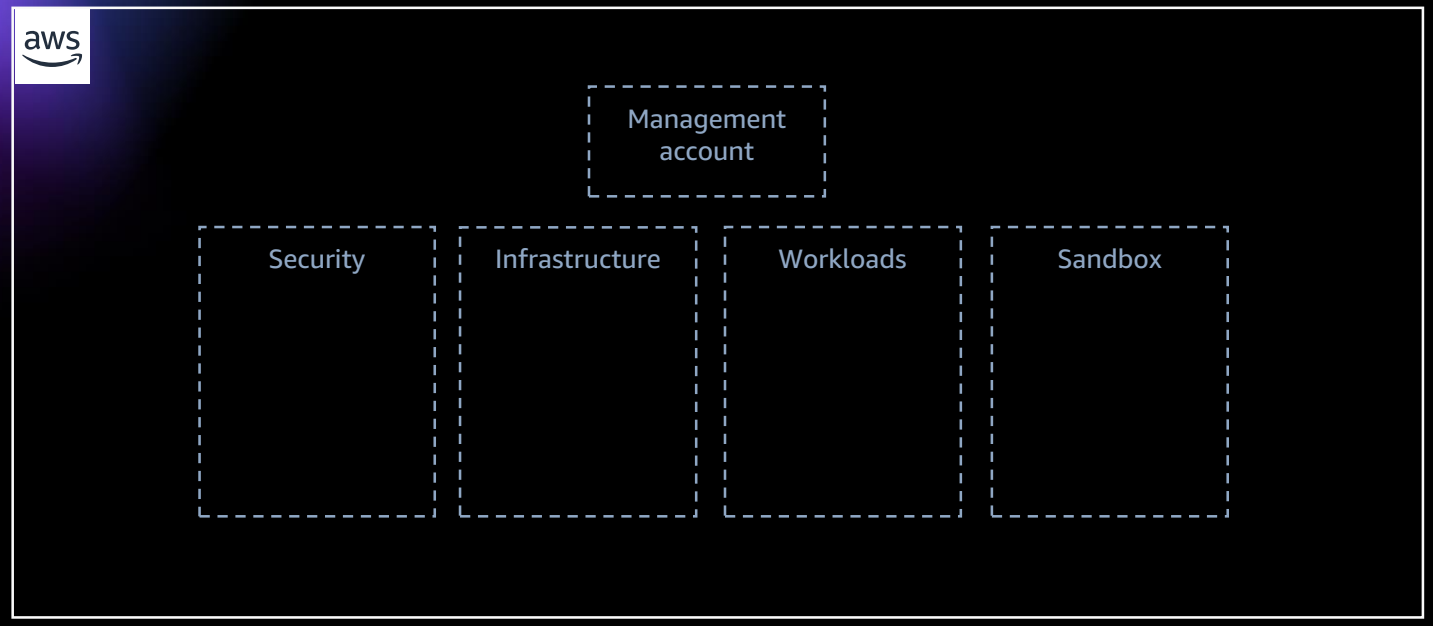
Service control policies (SCPs)

- Enable you to control which AWS service APIs are accessible:
  - ✓ Define the list of APIs that are allowed – allowlisting
  - ✓ Define the list of APIs that must be blocked – denylisting
- SCPs are:
  - ✓ Invisible to all users in the child account, including root
  - ✓ Applied to all users in the child account, including root
- Permissions:
  - ✓ Intersection between the SCP and IAM permissions
  - ✓ IAM Access Analyzer is SCP aware

SCP Example: Prevent AWS CloudTrail from being disabled

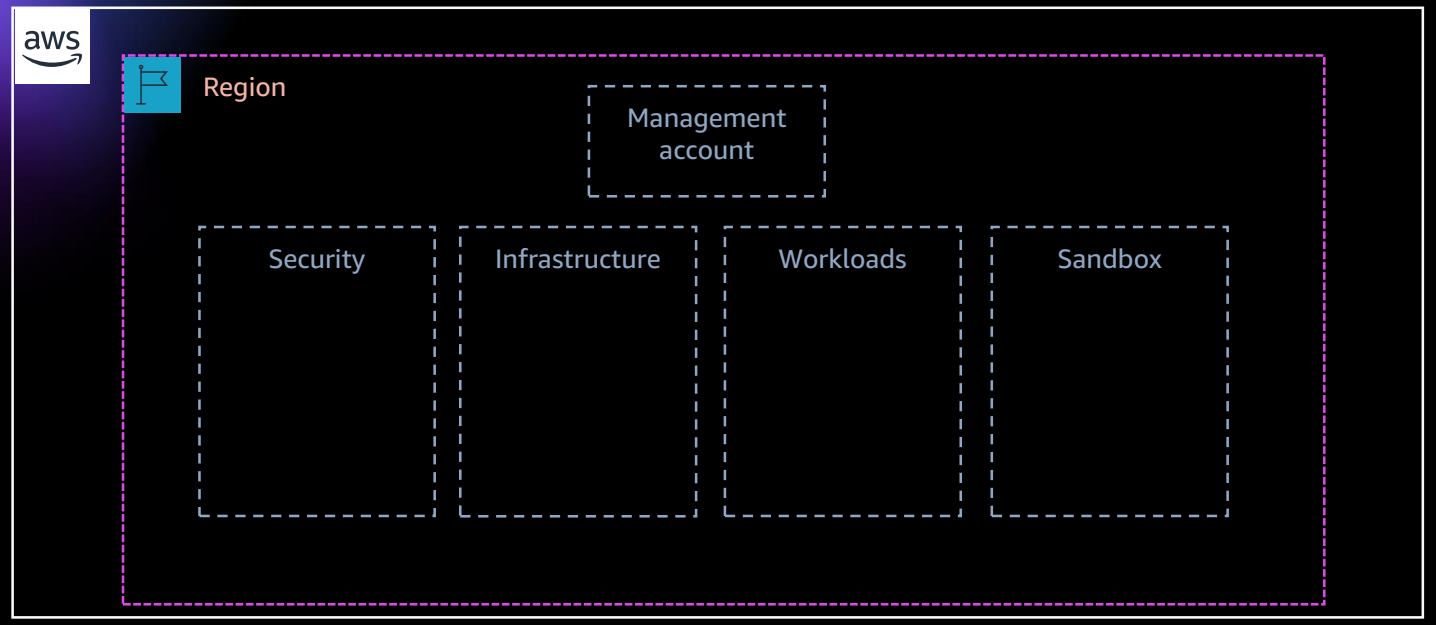
```
{
  "Version": "2012-10-17",
  "Statement": [
     {
      "Effect": "Deny",
      "Action": ”cloudtrail:StopLogging",
      "Resource": "*"
     }
  ]
}
```

Create a new organization



Create a new organization with four OUs

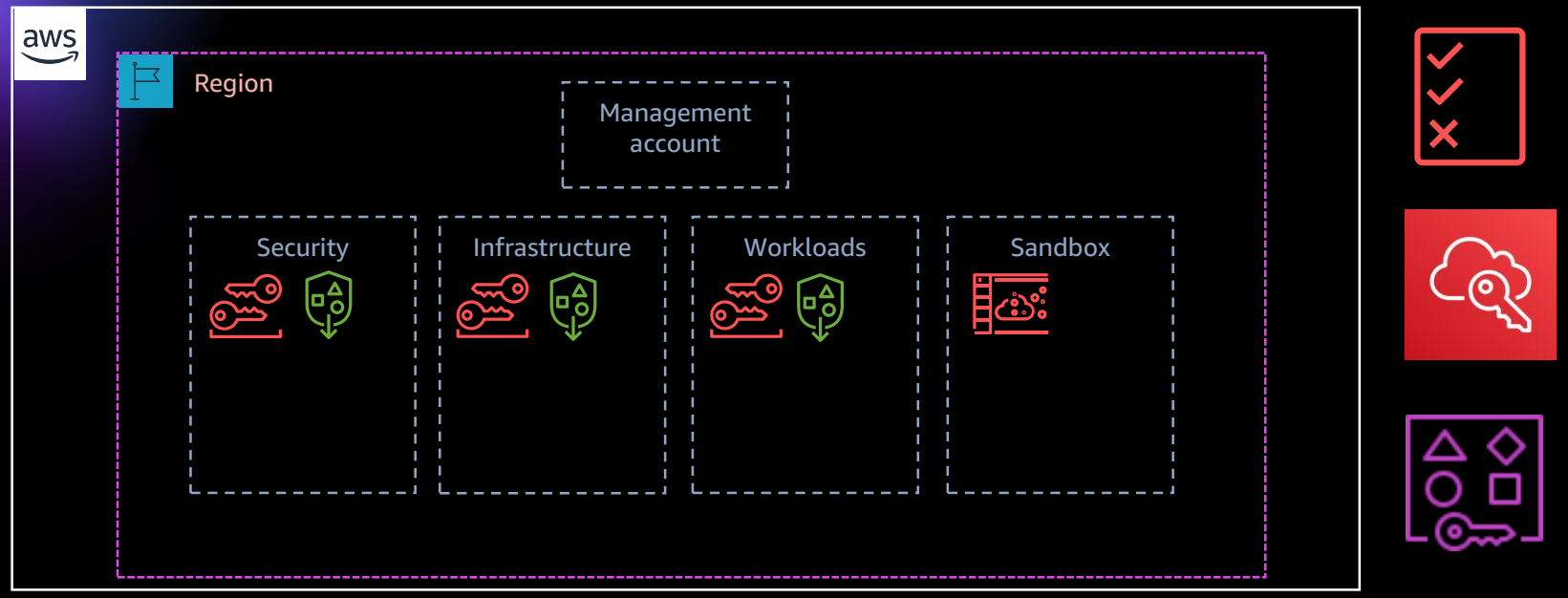
Operate workloads in specific regions



Apply a region-based SCP to the organization

Future instances/workloads can only be deployed in approved regions

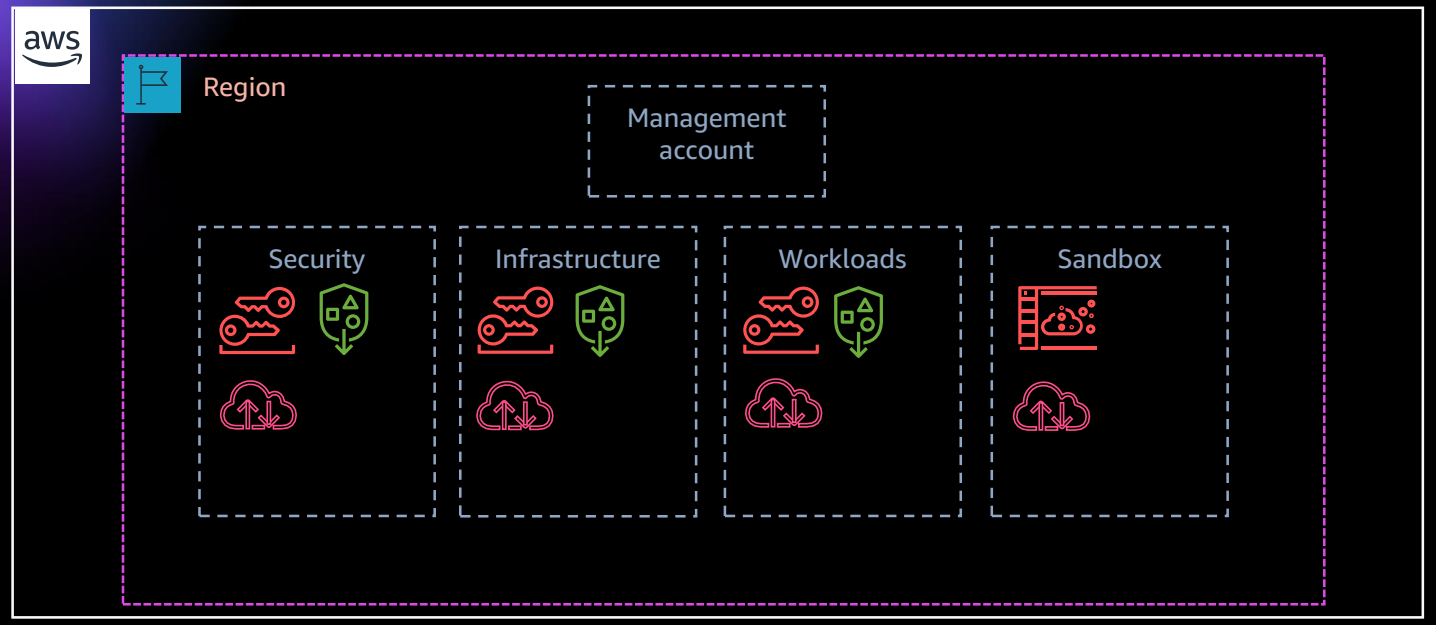
Provide access and resources for developers

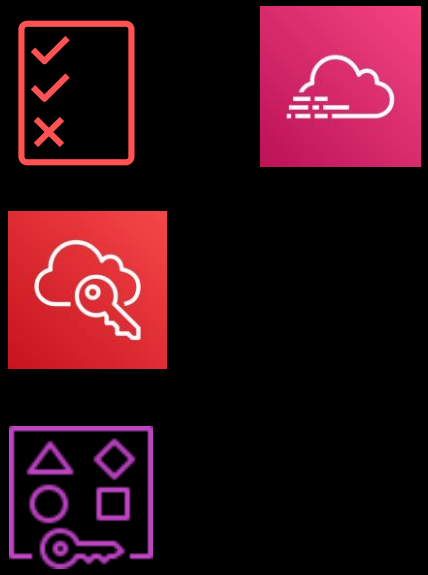


Enable AWS Single Sign On (SSO) for access
Create a Sandbox OU for test accounts
Use Resource Access Management to share subnets across accounts

Developers have access to resources and a space to build

Ensure all actions are logged for auditing

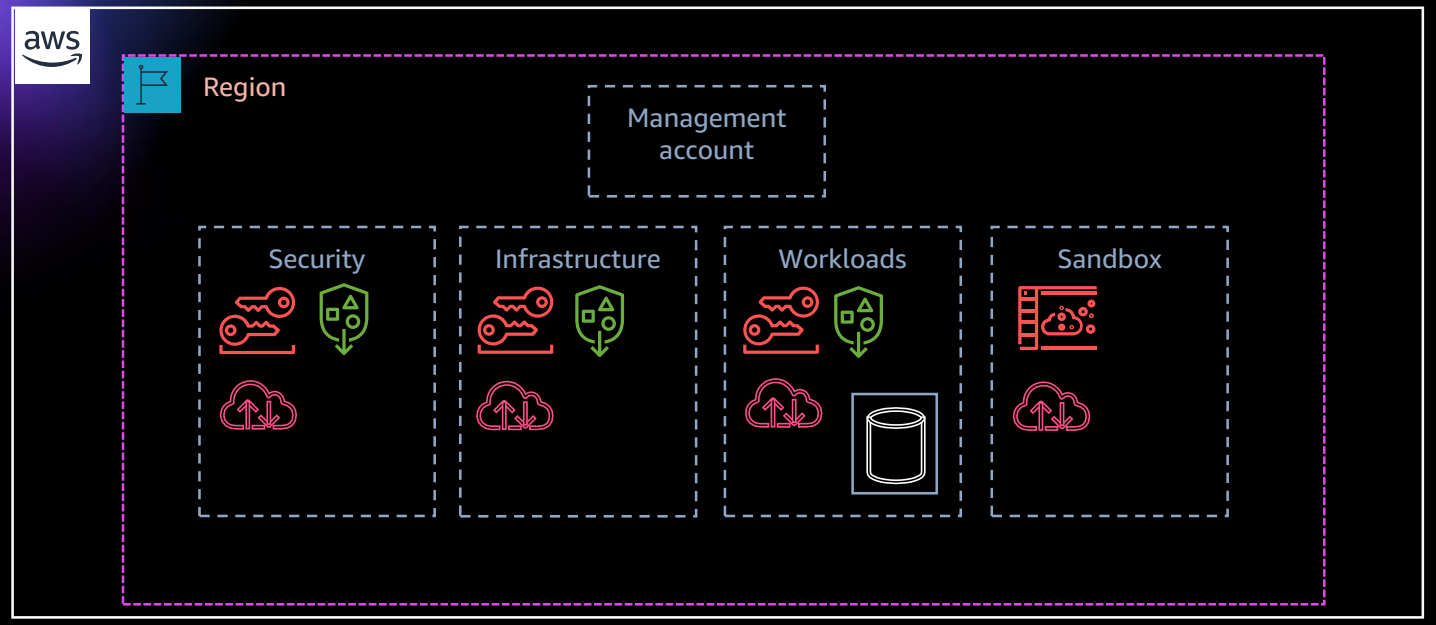


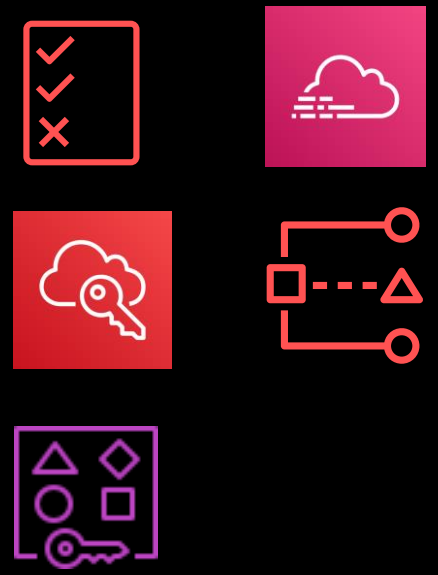


Enable AWS CloudTrail to create a searchable log of all cloud activity from the organization

Logging (and log activity) cannot be turned off or modified by users

Secure customer data





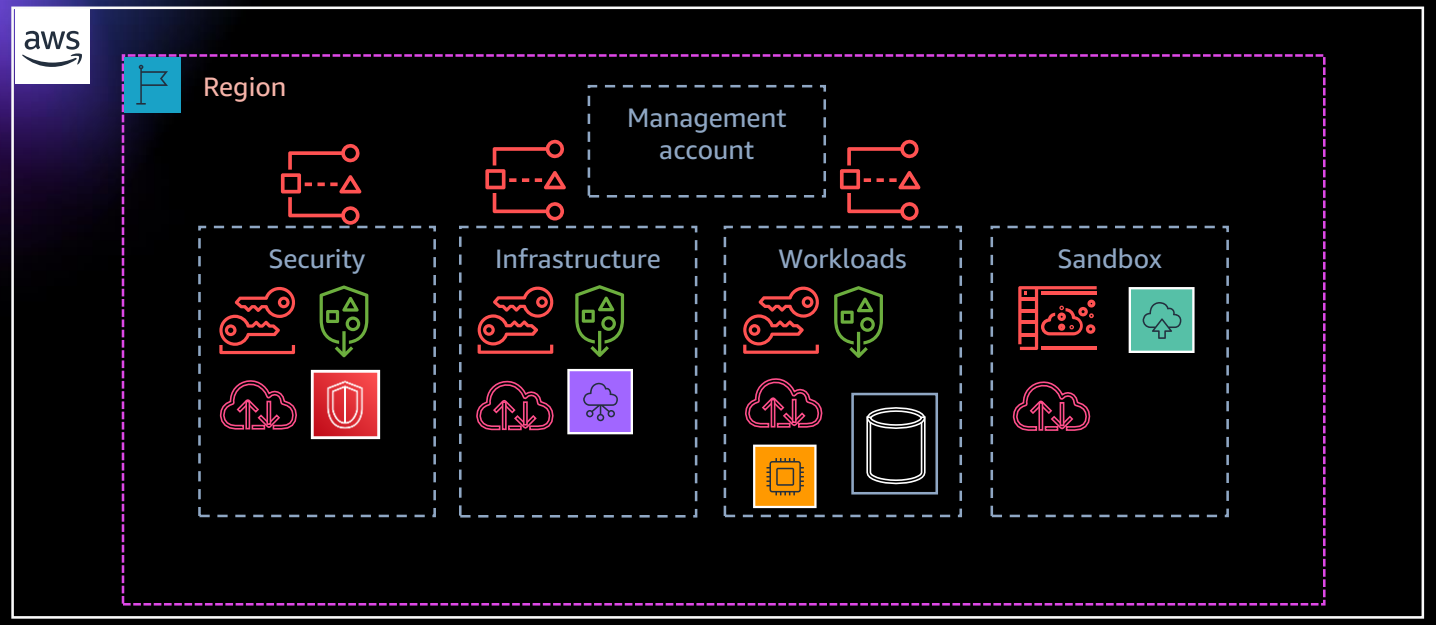
Isolate customer data in an Amazon S3 Bucket to an account with limited access

Apply an SCP, preventing changes to the Amazon S3 bucket visibility

Customer data is isolated and secure

aws

Test AWS services in a controlled environment

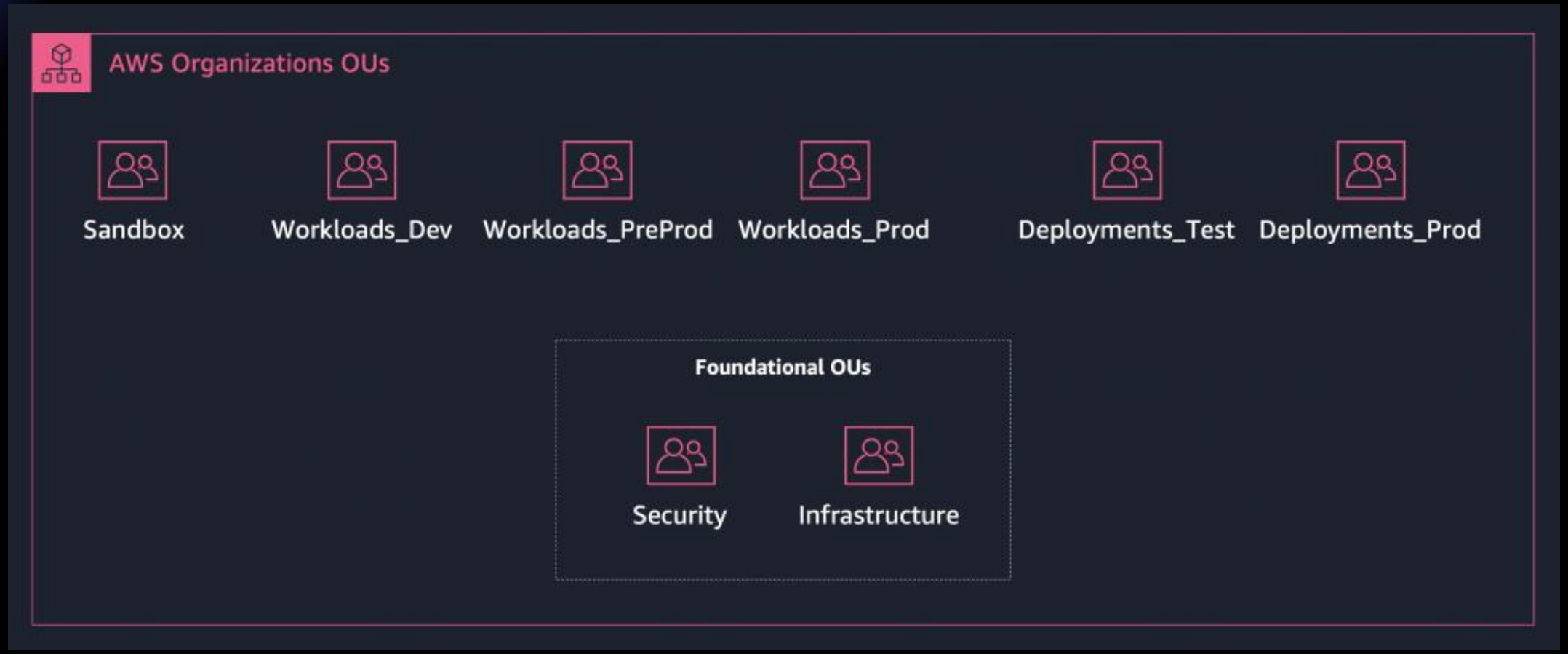


Create SCPs for each OU, defining services used in each environment

Allow use of new services once they are fully vetted

Safely deploy new services for teams

Suggested OU structure for ML workloads



“Seems complicated – is there a simpler way?”

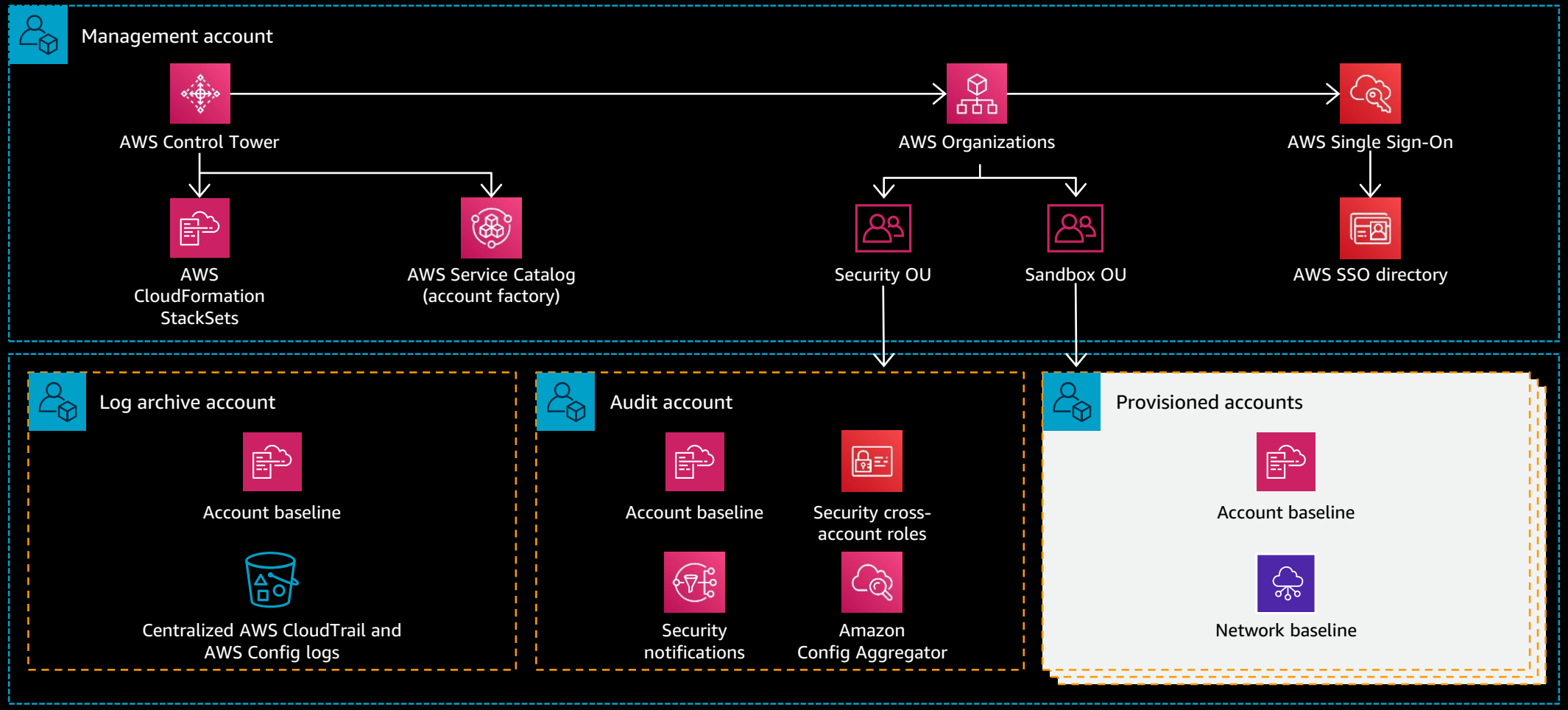
Managing your multi-account environment

AWS Organizations give you native tools to build your environment

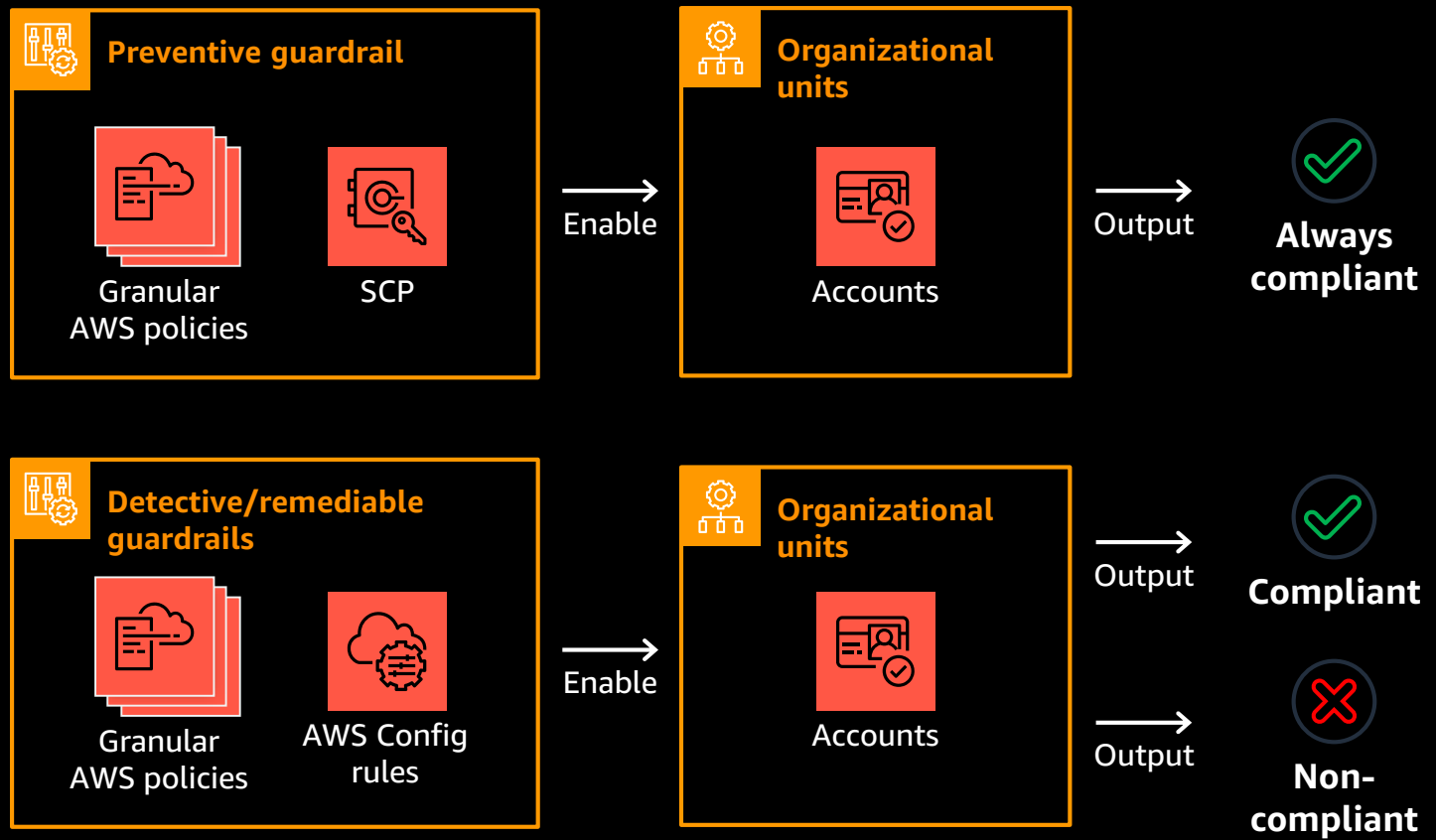
If you'd like to jump-start your AWS environment using a simple UI and built-in best practices, we recommend AWS Control Tower

AWS Control Tower

Use AWS Control Tower for accounts



Establish guardrails

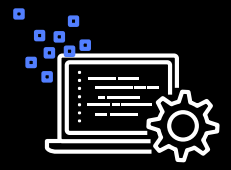


Suggested guardrails for ML environments

Guardrail	Type	Requirement
Enable MFA for the root user	Detective	Strongly recommended
Disallow public read access to Amazon S3	Detective	Strongly recommended
Enable AWS Config in all available regions	Preventive	Mandatory
Disallow deletion of log archive	Preventive	Mandatory
Enable AWS CloudTrail in all available regions	Preventive	Mandatory
Disallow Amazon S3 buckets that are not versioning enabled	Detective	Elective
Disallow changes to bucket policy for Amazon S3 buckets	Detective	Elective

Logging and auditing

AWS CloudTrail



Capture

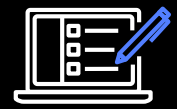
Record activity as CloudTrail events

Store

Retain events logs in secure Amazon S3 bucket

Act

Trigger actions when important events are detected



Review

Analyze recent events and logs with Amazon Athena or Amazon CloudWatch Logs Insights

aws

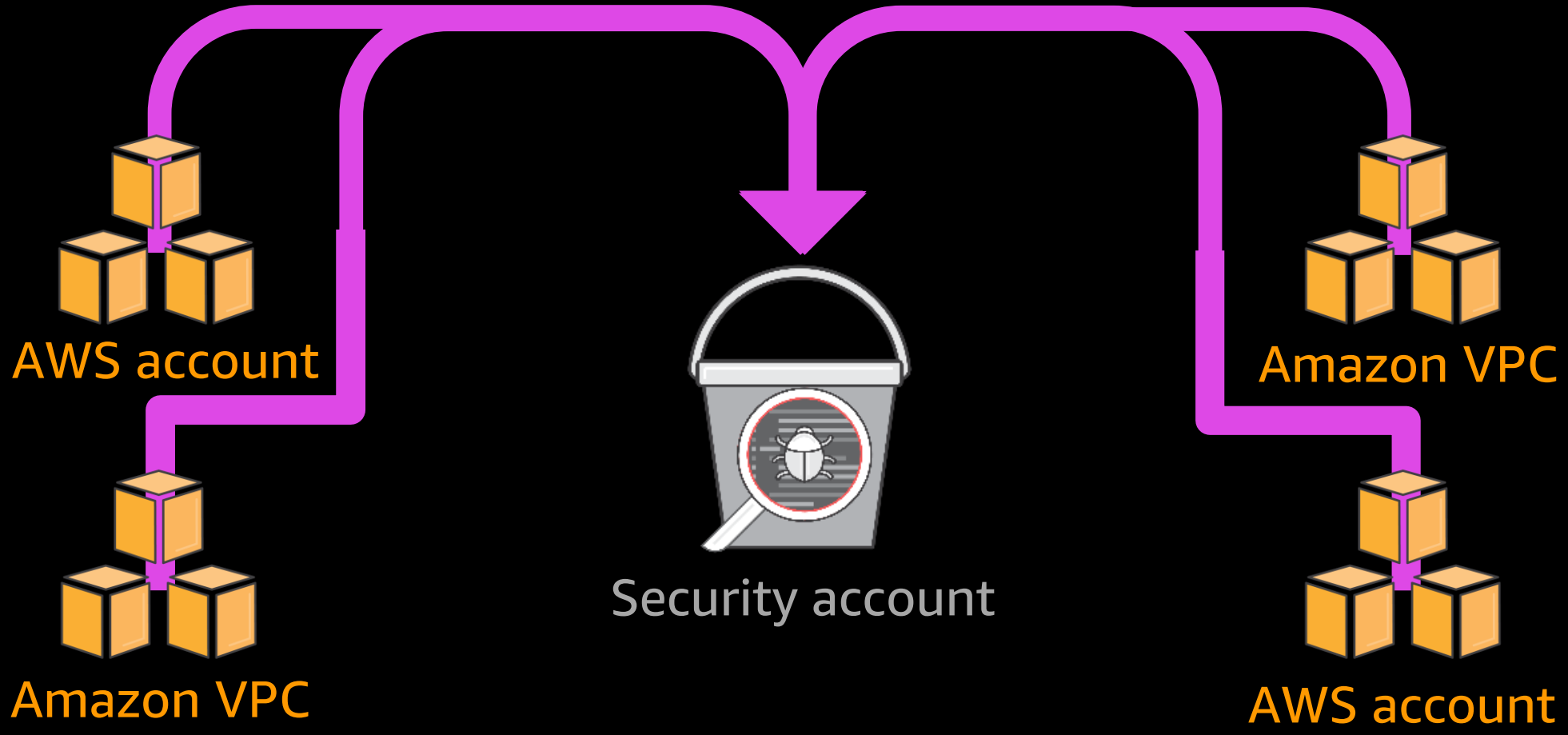
AWS CloudTrail

WEB SERVICE THAT RECORDS AWS API CALLS FOR YOUR ACCOUNT AND DELIVERS LOGS

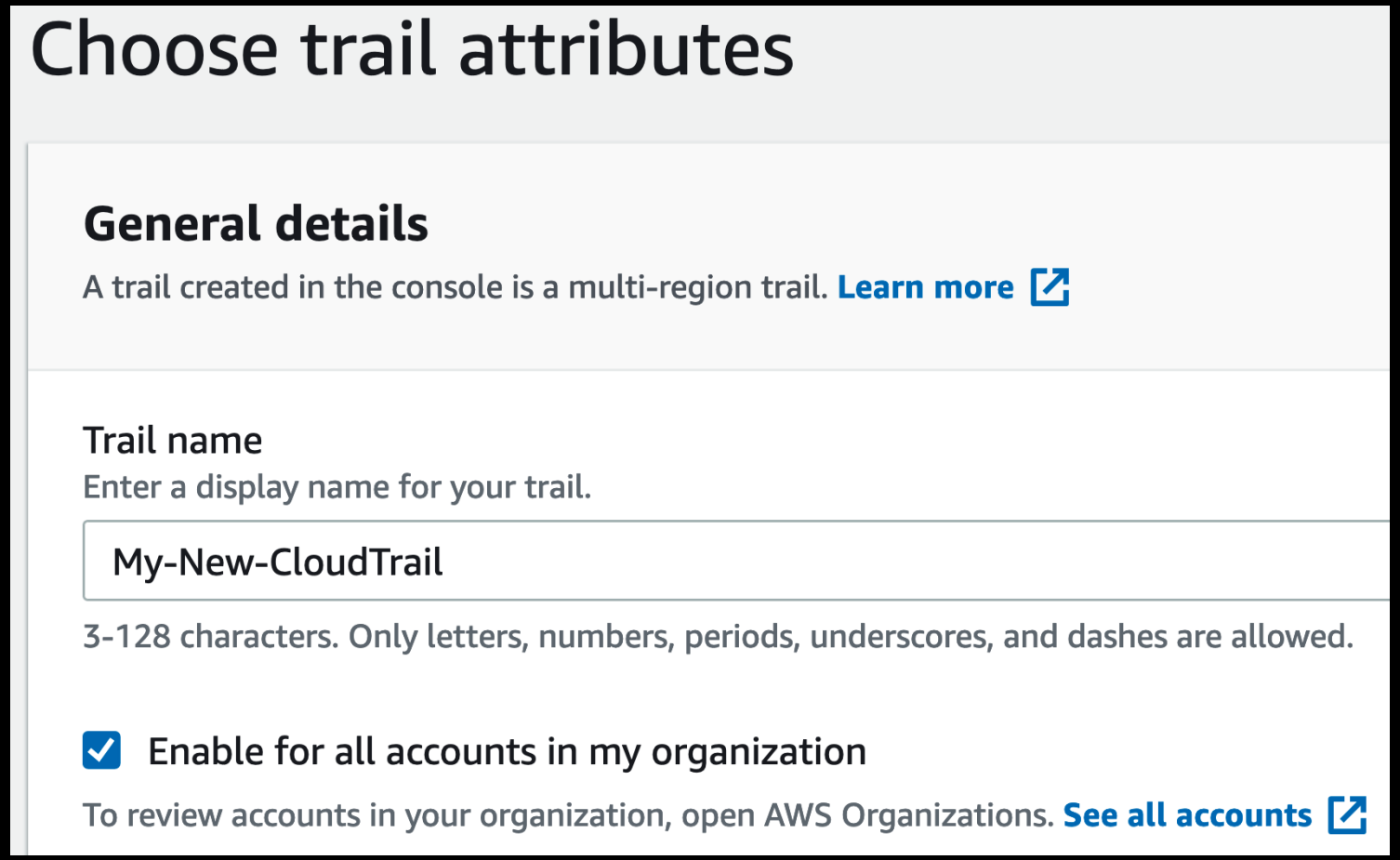
Who?	When?	What?	Where to?	Where from?
Bill	3:27pm	Launch instance	us-west-2	72.21.198.64
Alice	8:19am	Added Bob to admin group	us-east-1	127.0.0.1
Steve	2:22pm	Deleted Amazon DynamoDB table	eu-west-1	205.251.233.176


```
{
"Records": [
        {
        "eventVersion": "1.0",
        "userIdentity": {
            "type": "IAMUser",
            "principalId": "EX_PRINCIPAL_ID",
            "arn": "arn:aws:iam::123456789012:user/Alice",
            "accountId": "123456789012",
            "accessKeyId": "EXAMPLE_KEY_ID",
            "userName": "Alice",
            "sessionContext": {
                "attributes": {
                    "mfaAuthenticated": "false",
                    "creationDate": "2014-03-25T18:45:11Z"
                }
            }
        },
        "eventTime": "2014-03-25T21:08:14Z",
        "eventSource": "iam.amazonaws.com",
        "eventName": "AddUserToGroup",
        "awsRegion": "us-east-1",
        "sourceIPAddress": "127.0.0.1",
        "userAgent": "AWSConsole",
        "requestParameters": {
            "userName": "Bob",
            "groupName": "admin"
        },
        "responseElements": null
    },
    ...additional entries
  ]
}
```

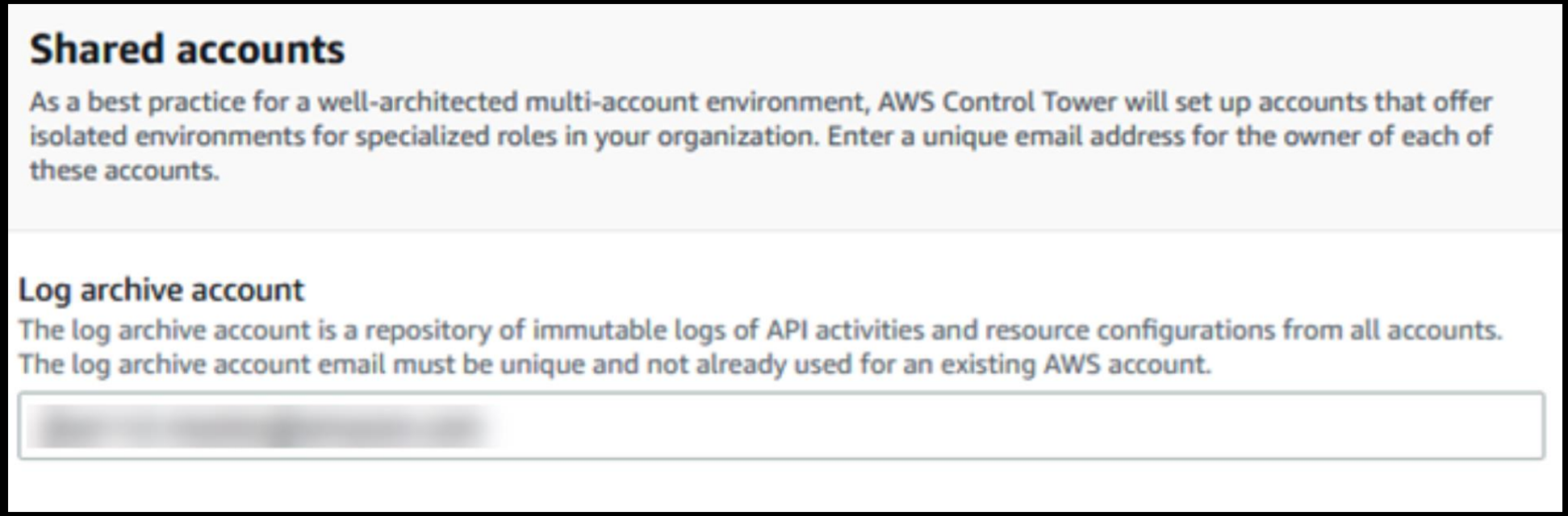
Centralize AWS CloudTrail logs



CloudTrail integration with AWS Organizations



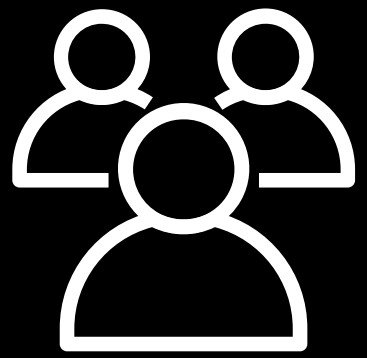
CloudTrail integration with AWS Control Tower



Identity and authorization

Identity, access, and resource management

Who



Identity management

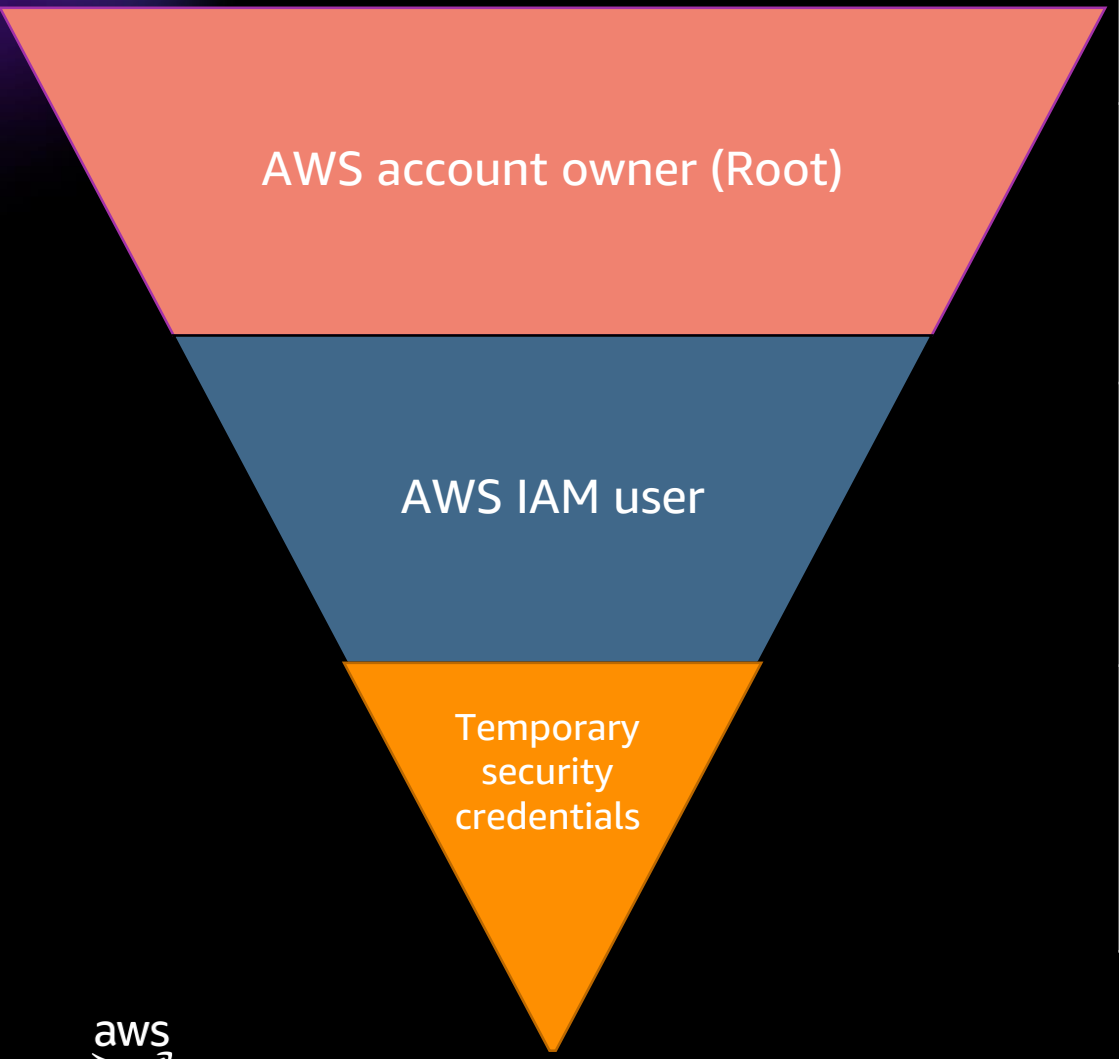
Can access

Access management

What

Resource management

AWS IAM hierarchy of privileges



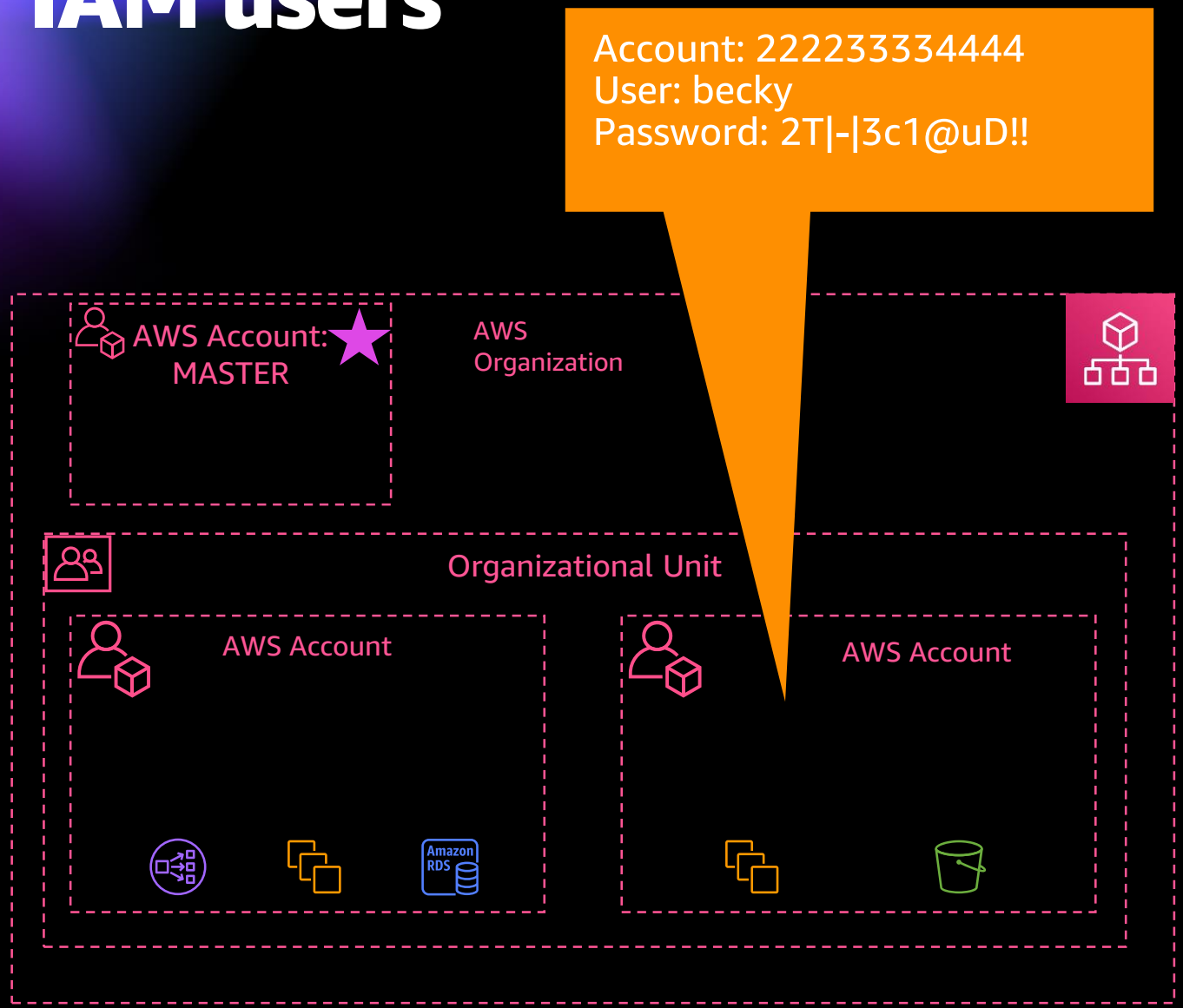
Permissions	Example

Unrestricted access to all enabled services and resources.	`Action: *`
	`Effect: Allow`
	`Resource: *`
	`(implicit)`

Access restricted by Group and User policies	`Action:`
	`['s3:*','sts:Get*']`
	`Effect: Allow`
	`Resource: *`

Access restricted by generating identity and further by policies used to generate token	`Action: [ 's3:Get*' ]`
	`Effect: Allow`
	`Resource:`
	`'arn:aws:s3:::mybucket/*'`

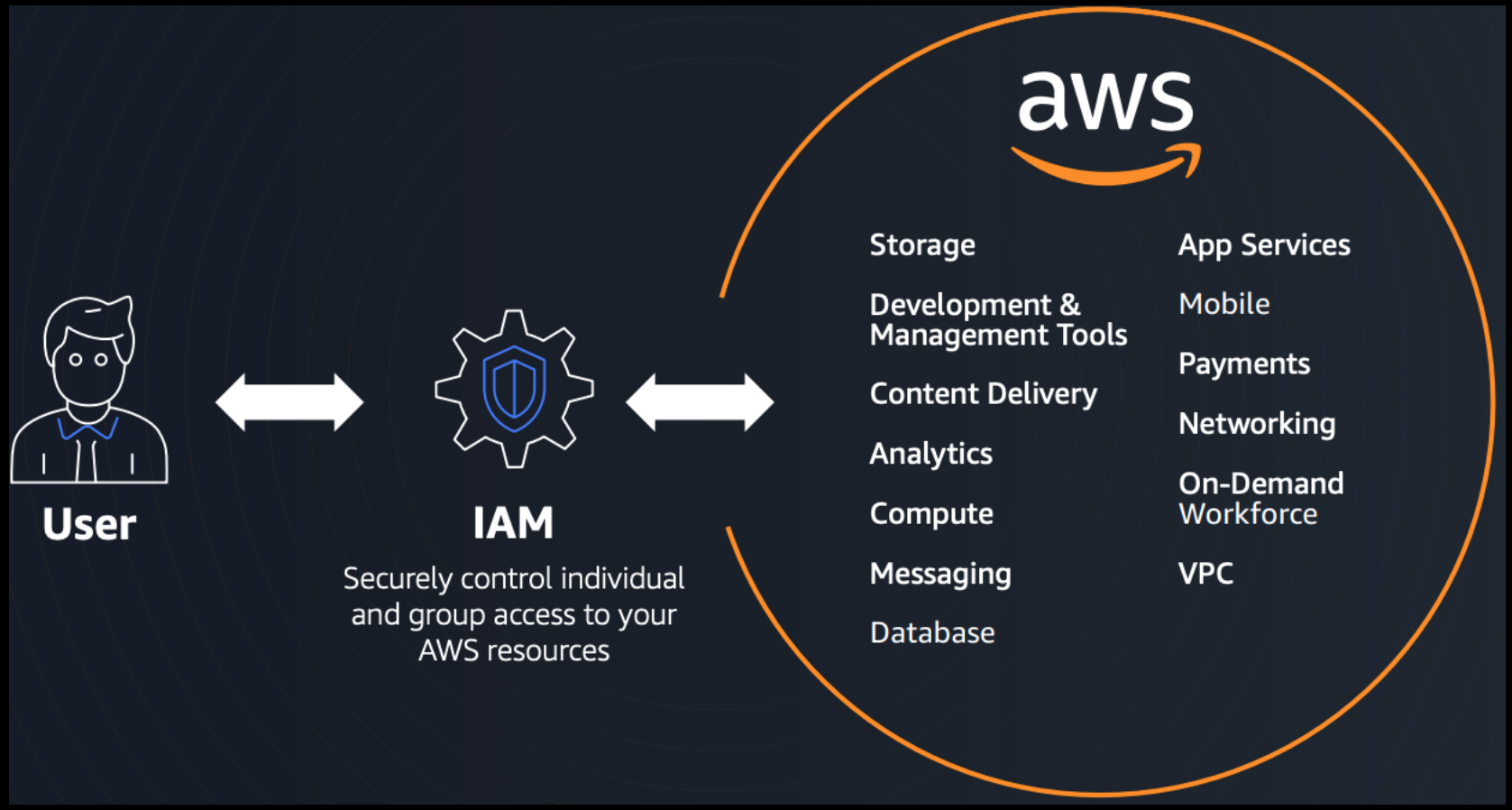
IAM users



Works best when you have:

- A relatively small number of users (limit is 5,000)
- One AWS account, or a relatively small number of them
- A need for long-term credentials
- No user directory, or no ability to connect your directory to AWS
- Your very first AWS account

Utilize least privilege IAM roles and policies



Common IAM permissions for ML environments

Analysts / data engineers

Development

- Launch apps / notebooks
- Data access
- Processing jobs
- Code repo access
- Package repo access

Data scientists / ML engineers

Training

- Training job
- Hyperparameter tuning job
- Transform job
- Auto ML job
- Experiments / trials / components
- Amazon Elastic Container Registry (Amazon ECR) access

DevOps engineers

Deployment

- Create endpoint
- Transform job
- Invoke endpoint
- Monitoring job

Permissions to encrypt/decrypt data, training artifacts and models

Permissions to create elastic network interfaces (ENIs) during training and hosting

Permissions to pass role to a service

Permissions to create storage volume and manage users

Validate IAM roles

Look for overly permissive roles

- Use IAM Access Analyzer
- Third-party/open source tools

Detect and remove unused roles

- Implement continuous monitoring of role activity using AWS Config

https://aws.amazon.com/iam/features/analyze-access/

Federate access

Easy Management

Easily manage AWS account and role access at scale

CLI

SSO from Command Line Interface (CLI)

One logon

One Sign-In experience for cloud business applications

Existing Identity Provider

Bring your own identities, or create them natively

First Party Application Integration

One Sign-In for integrated applications. AWS IoT SiteWise Monitor, Amazon SageMaker Notebooks

AWS Control Tower orchestrates AWS Single Sign-On to centralize identity and access

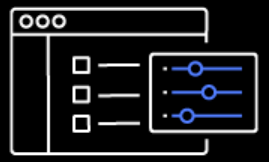
- AWS SSO provides default directory for identity
- AWS SSO also allows federated access management across all accounts in your organization
- Preconfigured groups (such as AWS Control Tower administrators, auditors and AWS Service Catalog end users)
- Preconfigured permission sets (e.g., admin, read-only, write)
- AWS SSO integrates with third-party IDP (Microsoft Azure AD, Ping, Okta)

Network controls

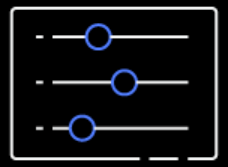
Amazon VPC - Virtual Private Cloud

- Provision a logically isolated section of the AWS Cloud where you can launch AWS resources in a virtual network that you define

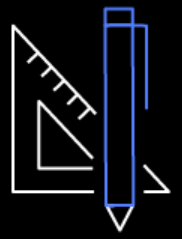
Bring your own network



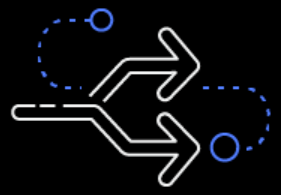
IP addresses



Subnets



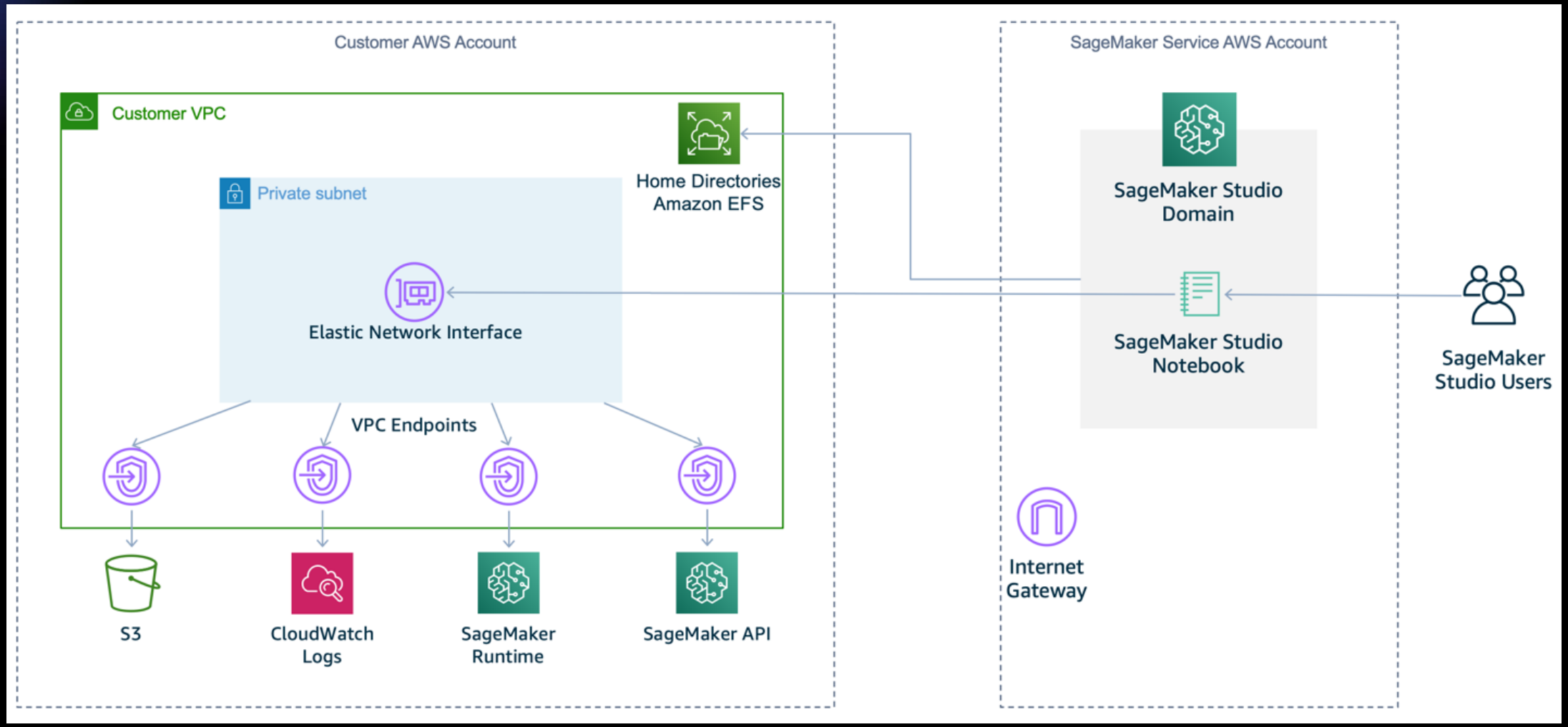
Network topology



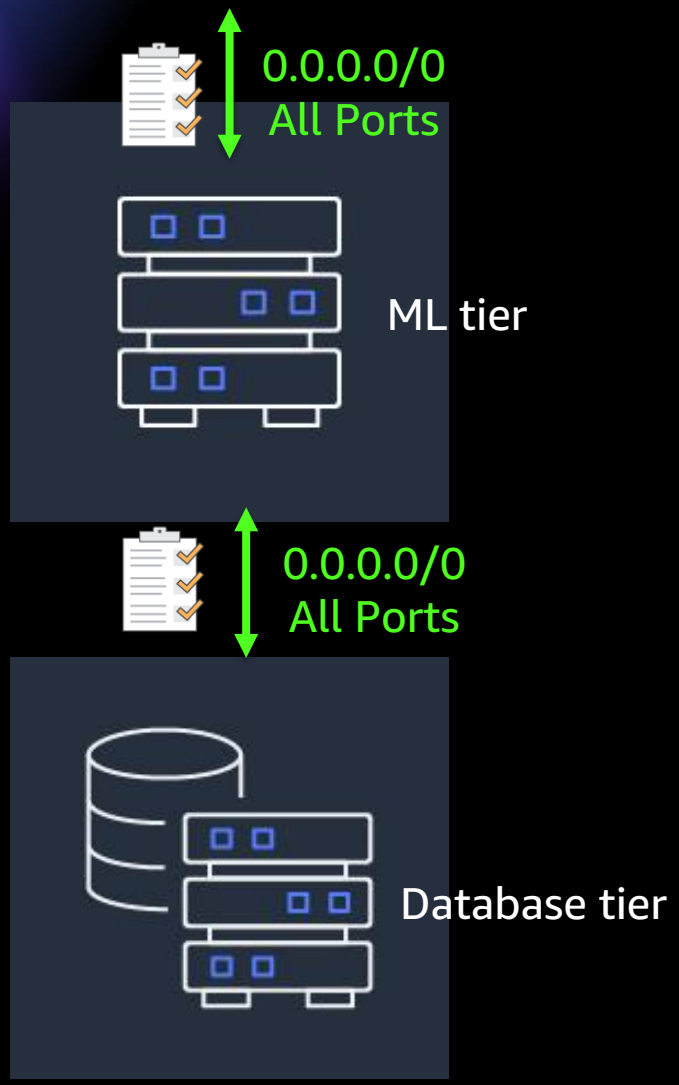
Routing rules

Security rules

Deploy ML workloads in a VPC

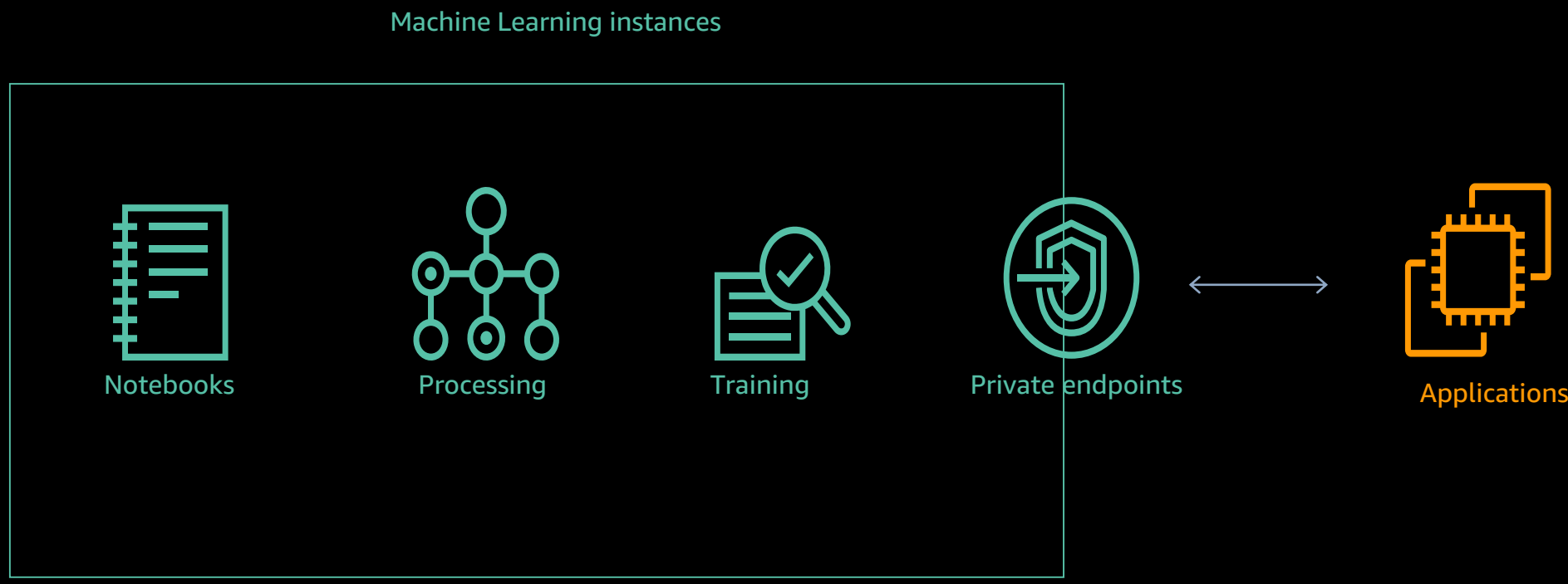


Limit security groups



Ensuring that only the required ports are open and the connection is enabled from known network ranges is a foundational approach to security

Use AWS PrivateLink to connect applications



Encryption

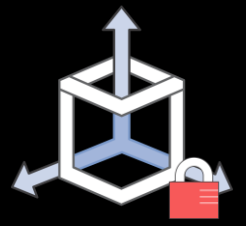
AWS Key Management Service (AWS KMS)

- **Easily create and control the keys used to encrypt or digitally sign your data**
- **Fully managed** - You control access to your encrypted data by defining permissions to use keys while AWS KMS enforces your permissions and handles the durability and physical security of your keys.
- **Centralized key management** – A single control point to manage keys and define policies consistently across integrated AWS services and your own applications.
- **Manage encryption for AWS services** – Integrated with AWS services to simplify using your keys to encrypt data across your AWS workloads.
- **Encrypt data in your applications** – Integrated with the AWS Encryption SDK to enable you to used KMS-protected data encryption keys to encrypt locally within your applications.
- **Built-in auditing** – Integrated with AWS CloudTrail to record all API requests.
- **Compliance** - The security and quality controls in AWS KMS have been certified under multiple compliance schemes to simplify your own compliance obligations.

ML data protection – encryption

Volumes encryption

Bucket encryption and deny policies

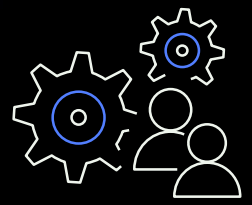


Inter-container traffic encryption

- Many customers also require the use of customer managed KMS keys (CMKs) for at-rest encryption versus AWS managed keys
- In addition to at-rest encryption within ML components, enable Amazon S3 default encryption and use deny policies to prevent unencrypted uploads
- ML services such as Amazon SageMaker support use of CMKs for uploading outputs back to Amazon S3
- Disable root access on any notebook instances

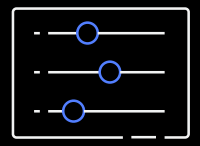
User access

Amazon SageMaker Studio



Collaboration at scale

Share notebooks without tracking code dependencies

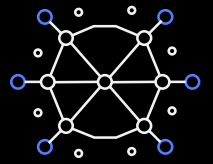


Easy experiment management

Organize, track, and compare thousands of experiments

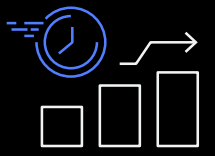
Automatic model generation

Get accurate models with full visibility and control without writing code



Higher quality ML models

Automatically debug errors, monitor models, and maintain high quality



Increased productivity

Code, build, train, deploy, and monitor in a unified visual interface

Demo

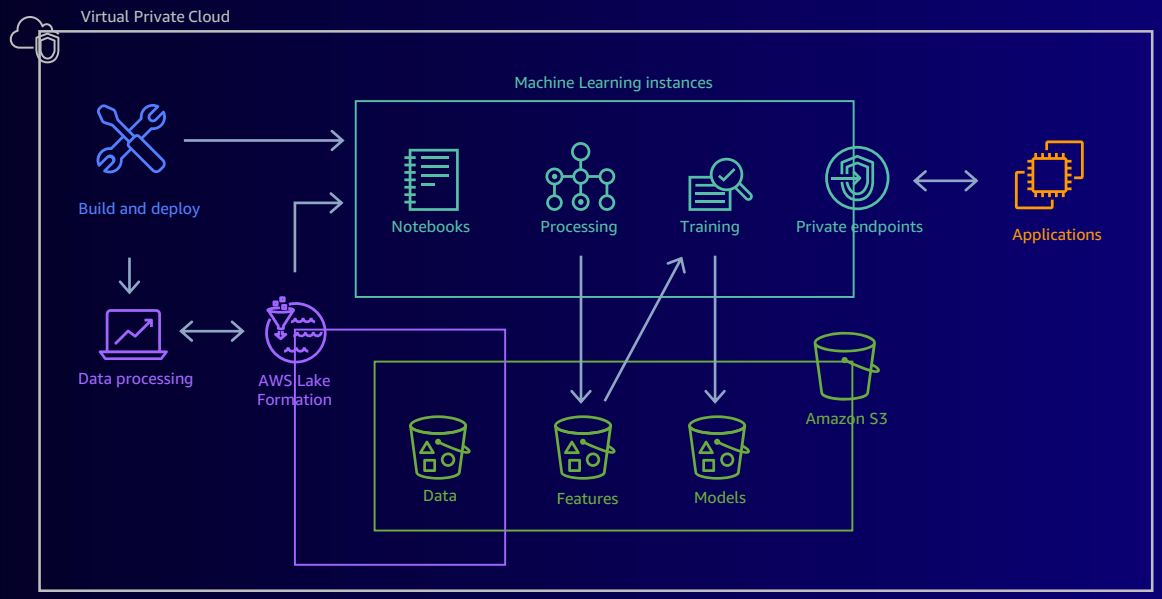
Recap

AWS account structure

Logging and auditing

Identity and authorization

Network controls

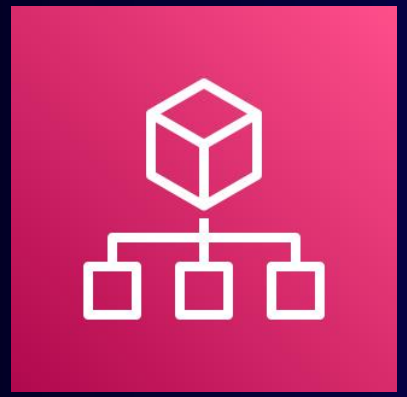


Encryption

User access

Recap

AWS account structure



AWS Organizations

AWS Control Tower

Recap

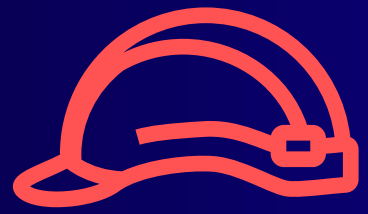
Logging and auditing

AWS CloudTrail

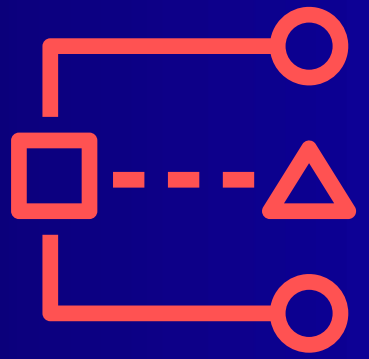
Recap

Identity and authorization

AWS Identity and Access Management (IAM)



Role



AWS IAM Access Analyzer

AWS Single Sign-On

Recap

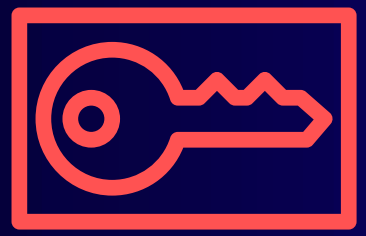
Network controls

Amazon Virtual Private Cloud (Amazon VPC)

PrivateLink

Recap

AWS Key Management Service (AWS KMS)



AWS Security Token Service (AWS STS)

Encrypted data

Encryption

Recap

Amazon
SageMaker Studio

User access

Further reading

Setting up secure, well-governed machine learning environments on AWS:
https://aws.amazon.com/blogs/mt/setting-up-machine-learning-environments-aws/

7 ways to improve security of your machine learning workflows:
https://aws.amazon.com/blogs/security/7-ways-to-improve-security-of-your-machine-learning-workflows/

Machine Learning best practices in financial services:
https://aws.amazon.com/blogs/machine-learning/machine-learning-best-practices-in-financial-services/

Visit the AI & Machine Learning resource hub for more resources

Dive deeper into these resources, get inspired and learn how you can use AI and machine learning to accelerate your business outcomes.

- The machine learning journey e-book
- 7 leading machine learning use cases e-book
- A strategic playbook for data, analytics, and machine learning e-book
  Accelerate machine learning innovation with the right cloud services & infrastructure e-book
- Choosing the right compute infrastructure for machine learning e-book
- Improving service and reducing costs in contact centers e-book
- Why ML is essential in your fight against online fraud e-book
- … and more!

https://bit.ly/3mwi59V

Visit resource hub

AWS Machine Learning (ML) Training and Certification

AWS is how you build machine learning skills

Courses built on the curriculum leveraged by Amazon's own teams. Learn from the experts at AWS.

aws.training/machinelearning

Flexibility to learn your way

Learn online with on-demand digital courses or live with virtual instructor-led training, plus hands-on labs and opportunities for practical application.

explore.skillbuilder.aws/learn

Validate your expertise

Demonstrate expertise in building, training, tuning, and deploying machine learning models with an industry-recognized credential.

aws.amazon.com/certification

aws

Thank you for attending AWS Innovate – AI/ML Edition

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.


aws-apj-marketing@amazon.com


twitter.com/AWSCloud


facebook.com/AmazonWebServices

youtube.com/user/AmazonWebServices


slideshare.net/AmazonWebServices


twitch.tv/aws

Thank you!

Michael Stringer