AWS Solutions Architecture Notes

Your Name

Date: May 3, 2025

Contents

Chapter 1

Course Introduction

Chapter 2

Identity & Federation

2.1 IAM - What should you know by now

- Users: Long term credentials - Groups - Roles: short-term credentials, uses STS - EC2 Instance Roles: Uses the EC2 metadata service. One role of a time per instance - Service Roles: API Gateways, CodeDeploy etc - Cross Account roles - Policies - AWS Managed - Customer Managed - Inline Policies - Resource Based Policies (S3 Bucket, SQS queues, etc.....)

2.1.1 IAM Policies Deep Dive

- Anatomy of a policy: JSON doc with Effect, Action, Resource, Conditions, Policy Variables - Explicit DENY has precedence over ALLOW - Best practice: use least privilege for maximum security - Access Advisor: See permissions granted and when last accessed - Access Analyser: Analyse resources that are shared with external entity - Navigate Examples at:

2.1.2 IAM AWS Managed Policies - Administrator Access example

Listing 2.1: Sample JSON Data

```
{
"Version": "2012-10-17",
"Statement": [
{
"Effect": "Allow",
"Action": "*",
"Resource": "*"
}
]
}
```

2.1.3 IAM Policies Conditions

2.1.4 IAM Policies Variables and Tags

2.1.5 IAM Roles vs Resource Based Policies

- Attach a policy to a resource (example: S3 bucket policy) versus attaching of a using a role as a proxy - When you assume a role (user, application or service), you give up your original permissions and take the permissions assigned to the role. - When using a resource-based policy, the principal doesn't have to give up any permissions - Example: User in account A needs to scan a DynamoDB table in Account A and dump it in an S3 bucket in Account B

2.1.6 IAM Permissions Boundaries

- IAM permission boundaries are supported for users and roles (not groups) - Advanced feature to use a managed policy to set the maximum permissions an IAM entity can get - Can be used in combinations of AWS organisations SCP

2.1.7 Use cases

- Delegate responsibilities to non administrators within their permission boundaries, for example create create new IAM users
- Allow developers to self-assign policies and managed their own permissions while making sure they can't 'escalate' their privileges (i.e make themselves admin)
- Useful to restrict one specific user (instead of a whole account using Organisations and SCP)

2.1.8 IAM Access Analyser

- Find out which resources are shared externally - S3 Buckets - IAM Roles - KMS Keys - Lambda Functions and Layers - SQS queues - Secrets Manager Secrets - Define Zone ofTrust = AWS Account or AWS Organisation - Access outside of zone trusts =¿ findings

2.1.9 IAM Access Analyser

- IAM Access Analyser Policy Validation - Validates your policy against IAM policy grammar and best practices - General warnings, security warnings errors suggestions - Provides actionable recommendations - IAM Access Analyser Policy Generation - Generates IAM policy based on access activity - CloudTrail logs is reviewed to generate the policy with the fine-grained permissions and the appropriate Actions and Services - Reviews CloudTrail logs for up to 90 days

2.2 STS

2.2.1 Using STS to Assume a Role

- Define an IAM Role within your account or cross-account - Define which principals can access this IAM role - Using AWS STS (Security Token Service) to retrieve credentials and impersonate

the IAM role you have access to (AssumeRole API) - Temporary credentials can be valid between 15 to 12 hours.

2.2.2 Assuming a Role with STS

- Provide access for an IAM user in one AWS account that you own to access resources in another account that you own. - Provide access to IAM users in AWS accounts owned by third parties - Provide access for services offered by AWS to AWS resources - Provide access for externally authenticated users (identity federation) - Ability to revoke active sessions and credentials for a role (by adding a policy using a time statement - AWSRevokeOlderSessions)

When you assume a role (user, application or service), you give up your original permissions and take the permissions assigned to the role.

2.2.3 Providing Access to an IAM User in Yours or Another AWS Account That You Own

- You can grant your IAM users permissions to switch to roles within your AWS account or to roles defined in other AWS accounts that you own. - Benefits: - You must explicitly grant your users permission to assume the role - Your users must actively switch to the role using the AWS management console or assume the role using the AWS CLI or AWS API - You can add multi-factor authentication (MFA) protection to the role so that only users who sign in with an MFA device can assume the role. - Least Privilege + auditing using CloudTrail

2.2.4 Providing Access to AWS Accounts Owned by Third Parties

- Zone of trust = accounts, organisations that you own - Outside Zone ofTrust = 3rd parties - Use IAM Access Analyser to find out which resources are exposed - For granting access to a thrid party: - The third part AWS account ID - An External ID (secret between you and the third party) - To uniquely associate with the role between you and 3rd party - Must be provided when defining the trust and when assuming the role - Must be chosen by the third party - Define permissions in the IAM policy

Session Tags in STS

- Tags that you pass when you assume an IAM Role or federate user in STS - aws:PrincipalTag Condition - Compares the tags attached to the principal making the request with the tag you specified in the policy. - Example: Allow a principal to pass session tags only if the principal making the reqeust has the specified tags.

STS Important APIs

- AssumeRole: access a role within your account or cross-account - AssumeRoleWithSAML: return credentials for users logged with SAML - AssumeRoleWithWebIdentity: return creds for users logged with an IdP - Example providers include Amazon Cognito, Login with Amazon, Facebook, Google or any OpenID Connect-compatible identity provider. - AWS Recommends using Cognito - GetSessionToken: for MFA from a user or AWS account root user - GetFederationToken: obtain temporary creds for a federated user, usually a proxy app that will give the creds for a federated user, usually a proxy app that will give the creds to a distributed app inside a corporate network.

2.2.5 Identity Federation in AWS

- Give users outside of AWS permissions to access AWS resources in your account - You don't need to create IAM Users (user management is outside AWS) - Use cases: - A corporate has its own identity system (e.g Active Directory) - Web / Mobile application that needs access to AWS resources - Identity Federation can have many flavours: - SAML 2.0 - Custom Identity Broker - Web Identity Federation With(out) Amazon Cognito - IAM Identity Centre

SAML 2.0 Federation

- Security Assertion Markup Language 2.0 (SAML 2.0) - Open standard used by many identity providers (e.g ADFS) - Supports integration with Microsoft Active Directory Federations Services (ADFS) - Or any SAML 2.0 - compatible IdPs with AWS - Access the AWS Console, AWS CLI or AWS API using temporary credentials - No need to create IAM Users for each of your employees - Need to setup a trust between AWS IAM and SAML 2.0 Identity Provider (both ways) - Under-the-hood: Uses the STS API AssumeRoleWithSAML - SAML 2.0 Federation is the "old way", IAM Identity Center Federation is the new managed and simpler way

SAML 2.0 Federation - AWS API Access

SAML 2.0 Federation - AWS Console Access

SAML 2.0 Federation - Active Directory DS (ADFS)

Custom Identity Broker Application

- Use only if identity provider is NOT compatible with SAML 2.0 - The identity broker is NOT compatible with SAML 2.0 - The identity broker must determine the appropriate IAM Role - Uses the STS API AssumeRole or GetFederationToken

Web Identity Federation - Without Cognito

- Not recommended by AWS - use cognito instead

Web Identity Federation - With Cognito

- Preferred over for Web Identity Federation - Create IAM Roles using Cognito with the least privilege needed - Build trust between the OIDC IdP and AWS - Cognito benefits: - Supports anonymous users - Supports MFA - Data Synchronisation - Cognito replaces a Token Vending Machine (TVM)

Web Identity Federation - IAM Policy

- After being authenticated with Web Identity Federation, you can identify the user with an IAM policy variable - Example:

2.2.6 AWS Directory Services (AD)?

- Found on any Windows Server with AD Domain Services - Database of objects: User Accounts, Computers, Printers, File Shares, Security Groups - Centralised security management, create account, assign permissions - Objects are organised in trees - A group of trees

What is ADFS (AS Federation Services)?

- ADFS provides Single Sign-On across applications - SAML across 3rd party: AWS Console, Dropbox, Officer365, etc.....

AWS Directory Services

- AWS Managed Microsoft AD - Create your own AD in AWS, manage users, locally supports MFA - Establish ”trust” connection with your own on premises - AD Connector - Directory Gateway (proxy) to redirect to on-premises AD, supports MFA - Users are managed on the on-premise AD - Simple AD - AD-compatible managed directory on AWS - Cannot be joined with on-premise AD

AWS Directory Services and AWS Managed Microsoft AD

- Managed Service: Microsoft AD in your AWS VPC - EC2 Windows Instances: - EC2 Windows instances can join the domain and run traditional AD applications (sharepoint, etc) - Seamlessly Domain Join Amazon EC2 Instances from Multiple Accounts and VPCs - Integrations - RDS for SQL Server, AWS Workspaces, Quicksight..... - AWS SSO to provide access to third party applications - Standalone repository in AWS or jointed to on-premises AD - Multi AZ deployment of AD in 2 AZ, of DC (Domain Controllers) can be increased for scaling - Automated backups - Automated Multi-Region replication of your directory

AWS Microsoft Managed AD - Integrations

Connect to on-premise AD

- Ability to connect your on-premise Active Directory to AWS Managed Microsoft AD - Must establish a Direct Connection (DX) or VPN connection - Can setup three kinds of forest trust - One-way trust: AWS -¿ On-Premise - One-way trust: On-Premise -¿ AWS - Two-way forest: trust - AWS -¿ On-Premise - Forest trust is different than synchronisation (replication is not supported)

Solution Achitecture: Active Directory Replication

- You may want to create a replica of your AD on EC2 in the cloud to minimise latency of in case DX or VPN goes down - Establish trust between the AWS Manged Microsoft AD and EC2

AWS Directory Services AD Connector

- AD Connector is a directory gateway to redirect director requests to your on premises Microsoft Active Directory - No caching capability - Managed users solely on-premise, no possibility of setting up a trust - VPN or Direct Connect - Doesn’t work with SQL Server, doesn’t do seamless joining, can’t share director.

AWS Directory Services Simple AD

- Simple AD is an inexpensive Active Directory-compatible service with the common directory features - Supports joining EC2 instances, manage users and groups - Does not support MFA, RDS SQL server, AWS SSO - Small: 500 users, large: 5000 users - Powered by Samba 4, compatible with Microsoft AD - lower cost, low scale, basic AD compatible or LDAP compatibility - No trust relationship

2.2.7 AWS Orgnisations

AWS Organisations - OrgnizationAccountAccessRole

- IAM role which grants full administrator permissions in the Member account to the Management account - Used to perform admin tasks in the Member accounts (e.g - creating IAM users) - Could be assumed by IAM users in the Management account - Automatically added to all new Member account created with AWS organisations - Must be created manually if you invite an existing Member account

Multi Account Strategies

- Create accounts per department, per cost centre, per dev / test / prod, based on regulatory restrictions (using SCP), for better resource isolation (ex:VPC), to have separate per-account service limits isolated account for logging. - Multi Account vs. On Account MultiVPC - Use tagging standards for billing purposes - Enable CloudTrail on all accounts, send logs to central S3 account - Send CloudWatch logs to central logging account - Strategy to create an account for security

Orgnisational Units (OU) - Examples

AWS Organisation - Feature Modes

- Consolidated billing features: - Consolidated Billing across all accounts - single payment method - Pricing benefits from aggregated usage (volume discount for EC2, S3.....)

All Features (Default)

- Includes consolidated billing features, SCP - Invited accounts must approve enabling all features - Ability to apply an SCP to prevent member accounts from leaving the org - Can't switch back to Consolidated Billing Features only

AWS Organisations - Reserved Instances

- For billing purposes, the consolidated billing features of AWS organisations treats all the accounts in the organisation as one account. - The means that all accounts in the organisation can receive the hourly cost benefit of Reserved Instances that are purchased by any other account. - The payer account (Management account) of an organisation can turn off Reserved Instance (RI) discount and Saving Plans discount sharing for any accounts in that organisation, including the payer account. - This means that RIs and Saving Plans discounts aren't shared between any accounts that have sharing turned off. - To share an RI or Savings Plans discount with an account, both accounts must have sharing turned on.

AWS Organisation - Moving Accounts

- Remove the member account from the AWS organisation - Send an invite to the member account from the AWS organisation - Accept the invite to the new Organisation from the member account.

Service Control Policies (SCP)

- Define allowlist or blocklist IAM actions - Applied at the OU or Account level - Does not apply to the Management Account - SCP is applied to all the Users and Roles in the account, including Root user - The SCP does not affect Service-linked roles - Service-linked roles enable other AWS services to integrate with AWS organisations and can't be restricted by SCP's - SCP must have an explicit Allow from the root of each OU in the direct path to the target account (does not allow anything by default) - Use cases: - Restrict access to certain services (for example: can't use EMR) - Enforce PCI compliance by explicitly disabling services.

SCP Hierarchy

- Management Account - Can do anything (no SCP apply) - Account A - Can do anything - EXCEPT S3 (explicit Deny from Sandbox OU) - EXCEPT EC2 (explicit deny) - Account B and C - Can do anything - EXCEPT S3 (explicit Deny from Sandbox OU) - Account D - Can access EC2 - Prod OU and Account E and F

SCP Examples - Blocklist and Allowlist strategies

IAM Policy Evaluation Logic

Restricting Tags with IAM policies

- You can restrict specific tags on AWS resources - Using the aws:TagKeys Condition Key - Validate the Tag Keys attached to a resource against the Tag Keys in the IAM Policy - Example: Allow IAM users to create EBS Volumes only if it has the "Env" and "CostCenter" Tags - Use either ForAllValues (must have all keys) or ForAnyValue (must have any of these keys at a minimum)

Using SCP to restrict creating resources without appropriate tags

- Prevent IAM Users/Roles in the affected member accounts from creating resources if they don't have a specific Tag

AWS Organisations - Tag Policies

- Helps you standardise tags across resources in an AWS organisation - Ensure consistent tags, audit tagged resources, maintain proper resources categorisation - You define Tag keys and their allowed values - Helps with AWS Cost Allocation Tags and Attribute-based Access Control - Prevent any non-compliant tagging operations on specified services and resources - Generate a report that lists all tagged/ non compliant resources - Use Amazon EventBridge to monitor non-compliant tags

AWS Organisation - AI Service Opt-out Policies

- Certain AWS AI services may use your content for continuous improvement of Amazon AI/ML services - Example: Amazon Lex, Amazon Comprehend, Amazon Polly..... - You can opt-out of having your content stored or used by AWS AI services - Create an Opt-out Policy that enforces this setting across all Member accounts and AWS Regions - You can opt-out all AI services or selected services - Can be attached to Organisation Root, specific OU or individual Member account

AWS Organisations - Backup policies

- AWS Backup enables you to create Backup Plans that define how to backup your AWS resources - JSON Documents that define backup plans across an AWS Organisation - Gives you granular control over backup up your resources (e.g backup frequency, time window, backup region,.....) - Can be attached to Organisation Root, specific OU or individual Member account - Immutable backup plans appear in Member accounts (view only)

Using SCP to Deny a Region aws:ReqeustRegion

2.2.8 AWS IAM Identity Center

AWS IAM Identity Center -successor to AWS Single Sign-On-

- One login (single sign-on) for all your - AWS accounts in AWS organisations - Business cloud applications (e.g, Salesforce, Box, Microsoft 365) - SAML2.0 enabled applications - EC2 windows instances

- Identity Providers - Built-in identity store in IAM identity center - 3rd party Active Directory (AD), OneLogin, Okta

- AWS IAM Identity Center - Login Flow

AWS IAM Identity Center

AWS IAM Identity Center - Fine-grained Permissions and Assignments

- Multi-Account Permissions - Manage access across AWS accounts in your AWS Organisation - Permission Sets - a collection of one or more IAM policies assigned to users and groups to define AWS access - Application Assignments - SSO access to many SAML 2.0 business applications (Salesforce, Box, Microsoft 365) - Provide required URLs, certificates and metadata - Attribute-Based Access Control (ABAC) - Fine-grained permissions based on users' attributes stored in IAM identity center identity store - Example: Cost center, title, locale - Use case: Define permission once, then modify AWS access by changing the attributes

AWS Control Tower

- Easy way to setup and govern a secure and compliant multi-account AWS environment based on best practices - Benefits: - Automate the set up of your environment in a few clicks - Automate ongoing policy management using guardrails - Detect policy violations and remediate them

- Monitor compliance through an interactive dashboard - AWS ControlTower runs on top AWS Organisations: - It automatically sets up AWS Organisations to organise accounts and implement SCPs (Service Control Policies)

AWS Controller Tower - Account Factory

- Automates account provisioning and deployments - Enables you to create pre-approved baselines and configuration options for AWS accounts in your organisation (e.g VPC default configuration, subnets, region, ...) - Uses AWS service catalog to provision new AWS accounts

AWS Control Tower - Detect and Remediate Policy Violations

- Guardrail - Provide ongoing governance for your Control Tower environment (AWS Accounts) - Preventive - using SCPs (e.g Disallow Creation of Access Keys for the Root User) - Detective - users AWS Config (e.g Detect Whether MFA for the Root User is Enabled) - Example: identify non-compliant resources (e.g, untagged resources)

AWS Control Tower - Guardrails Levels

- Mandatory - Automatically enabled and enforced by AWS control tower - Example: Disallow public Read access to the Log Archive account - Strongly Recommended - Based on AWS best practices (optional) - Example: Enable encryption for EBS volumes attached to EC2 instances - Elective - Commonly used by enterprises (optional) - Examples: Disallow delete actions without MFA in S3 buckets

2.3 AWS Resource Access Manager (RAM)

- Share AWS resources that you own with other AWS accounts - Share with any account or within your Organisation - Avoid resource duplication! - VPC Subnets - Allow to have all the resources launched in the same subnets - Must be from the same AWS organisations - Cannot share security groups and defaultVPC - Participants can manage their own resources in there - Participants can't view, modify, delete resources that belong to other participants or the owner - AWS Transit Gateway - Route 53 (Resolver Rules, DNS Firewall Rule Groups) - License Manager Configurations

AWS Resource Access Manager (RAM)

- Aurora DB Clusters - ACM Private Cerficiate Authority - CodeBuild Project - EC2 (Dedicated Hosts, Capacity Reservation) - AWS Glue (Catalog, Database, Table) - AWS Network Firewall Policies - AWS Resources groups - Systems Manager Incident Manager (Contacts, Response Plans) - AWS Outposts (Outpost, Site)

Resource Access Manager - VPC example

- Each account..... - Is responsible for it own resources - Cannot view modify or delete other resources in other accounts - Network is shared so..... - Anything deployed in the VPC can talk to other resources in the VPC - Applications are accessed easily across accounts, using a private

IP - Security groups from other accounts can be referenced for maximum security - Use cases - Applications within the same trust boundaries - Applications with a high degree of interconnectivity

Resource Access Manager Managed Prefix List

- A set of one or more CIDR blocks - Makes it easier to configure and maintain Security Groups and Route Tables - Customer-Managed Prefix List - Set of CIDRs that you define and manage by you - Can be shared with other AWS accounts or AWS Organisation - Modify to update many security groups at once - AWS-Managed Prefix List - Set of CIDRs for AWS services - You can't create, modify, share or delete them.

Resource Access Manger Route 53 Outbound Resolver

- Helps you scale forwarding rules to your DNS in case you have multiple accounts and VPC

2.4 Summary of Identity and Federation

- Users and Accounts all in AWS
- AWS Organisations
- AWS Control Tower to setup secure and compliant multi-account AWS environment (best practices)
- Federation with SAML
- Federation without SAML with a custom IdP (GetFederationToken)
- IAM Identity Center to connect to multiple AWS Accounts (Organisation) and SAML apps
- Web Identity Federation (not recommended)
- Cognito for most web and mobile applications (has anonymous mode, MFA)
- AWS Directory Service:
  - Managed Microsoft AD - standalone or setup trust AD with on-premises, has MFA, seamless joins, RDS integration
  - AD Connector - proxy requests to on-premises
  - Simple AD - standalone and cheap AD-compatible with no MFA, no advanced capabilities
- AWS RAM to share resource (example VPC subnets)

Chapter 3

Security

Chapter 4

Compute & Load Balancing

Chapter 5

Storage

Chapter 6

Caching

Chapter 7

Databases

Chapter 8

Service Communication

Chapter 9

Data Engineering

Chapter 10

Monitoring

Chapter 11

Deployment and Instance Management

Chapter 12

Cost Control

12.1 Cost Allocation Tags

- With Tags we can track resources that relate to each other
- With Cost Allocation Tags we can enable detailed costing reports
- Just like Tags, but they show up as columns in Reports
- AWS Generated Cost Allocation Tags
  - Automatically applied to the resource you create
  - Start with Prefix aws: (e.g. aws: createdBy)
  - They're not applied to resources created before the activation
- User tags
  - Defined by the user
  - Start with Prefix user:
- Cost Allocation Tags just appear in the Billing Console
- Takes up to 24 hours for the tags to show up in the report

12.2 AWS Tag Editor

- Allows you to managed tags of multiple resources at once - You can add/update/delete tags - Search tagged/untagged resources in all AWS Regions

12.3 Trusted Advisor

- No need to install anything - high level AWS account assessment - Analyse your AWS accounts and provides recommendation - Cost Optimisation - Performance - Security - Fault Tolerance - Service Limits - Operational Excellence - Core Checks and recommendations - all customers - Can enable weekly email notification from the console - Full Trusted Advisor - Available for Business and Enterprise support plans - Ability to set CloudWatch alarms when reaching limits - Programmatic Access using AWS support API

Column1	Basic Support
AWS Trusted Advisor Best Practice Checks	7 Core Checks
Enhanced Technical Support	24x7 customer service, documentation, whitepapers and support
Case Severity / Response Times	Data12
Data16	Data17

Table 12.1: Example Table with 4 Rows and 5 Columns

Trusted Advisor - Good to Know

- Can check if an S3 bucket is made public - But cannot check for S3 objects that are public inside of your bucket - Use Amazon EventBridge / S3 Events instead / AWS Config Rules

- Service Limits - Limits can only be monitored in Trusted Advisor (cannot be changed) - Cases must be created manually in AWS Support Centre to increase limits - OR use the AWS Service quotas service

12.4 AWS Service Quotas

- Notify you when you're close to a service quota value threshold - Create CloudWatch Alarms on the Service Quotas console - Example: Lambda concurrent executions - Helps you know if you need to request a quota increase or shutdown resources before limit is reached

12.5 EC2 Launch Types and Savings Plans

- On Demand Instances - short workload, predictable pricing, reliable. - Spot Instances - short workloads for check, can lose instances (not reliable) - Reserved: (Minimum 1 year) - Reserved Instances - long workloads - Convertible Reserved Instances - long workloads with flexible instances - Dedicated Instances: no other customers will share you hardware - Dedicated Hosts: book an entire physical server, control instance placement - Great for software licenses that operate at the core, or socket level - Can define host affinity so that instance reboots are kept on the same host

12.5.1 AWS Savings Plan

- New pricing model to get a discount based on long-term usage - Commit to a certain type of usage: ex $10 per hour for 1 to 3 years - Any usage beyond the savings plan is billed at the on-demand price

- EC2 Instance Savings plan (72% - same discount as Standard RIs) - Select instance family and locked to a specific region - Flexible across size, OS (Windows to Linux) tenancy. (dedicated or default) - Compute Savings Plan - Ability to move between instance family, region, compute type and OS and tenancy - SageMaker Savings plan (up to 64% off)

12.6 S3 Cost Savings

12.6.1 S3 Storage Classes

- Amazon S3 Standard - General Purpose - Amazon S3 Standard-Infrequent Access (IA) - Amazon S3 One Zone-Infrequent Access - Amazon S3 Glacier Instant Retrieval - Amazon S3 Glacier Flexible Retrieval - Amazon S3 Glacier Deep Archive - Amazon S3 Intelligent Tiering

Can move between classes manually or using S3 lifecycle configurations

12.6.2 S3 - Other Cost Savings

- S3 Lifecycle Rules: transition objects between tiers - Compress Objects - to save space - S3 Requester Pays: - In general, bucket owners pay for all Amazon S3 storage and data transfer costs associated with their bucket - With Requester Pays buckets, the requester instead of the bucket owner pays the cost of the request and the data downloaded from the bucket - The bucket owner always pays the cost of storing data - Helpful when you want to share large datasets with other accounts - If an IAM role is assumed the owner account of that role pays for the request

12.7 S3 Storage Classes - Reminder

12.7.1 S3 Storage Classes

- Amazon S3 Standard - General Purpose - Amazon S3 Standard-Infrequent Access (IA) - Amazon S3 One Zone-Infrequent Access - Amazon S3 Glacier Instant Retrieval - Amazon S3 Glacier Flexible Retrieval - Amazon S3 Glacier Deep Archive - Amazon S3 Intelligent Tiering

Can move between classes manually or using S3 lifecycle configurations

12.7.2 S3 Durability and Availalbility

- Durability: - High Durability (99.9999999999, 11 nines) of objects across multiple AZ - If you store 10,000,000 object with Amazon S3, you can on average expect to incur a loss of a single object once every 10,000 years (nice....but basic math) - Same for all storage classes

- Availability - Measures how readily available a service is - Varies depending on storage class - Example: S3 standard has 99.99% availability = not available 53 minutes a year

S3 Standard - General Purpose - 99.99% Availability - Used for frequently accessed data - Low latency and high throughput - Sustain 2 concurrent facility faiures - Use Cases: Big Data analytics, mobile and gaming applications, content and distribution

S3 Storage Classes - Infrequent Access - For data that is less frequently accessed, but requires rapid access when needed - Lower cost than S3 standard

Amazon S3 Standard-Infrequent Access (S3 Standard-IA) - 99.9% Availability - Use cases: Disaster Recovery, backups

Amazon S3 One Zone-Infrequent Access (S3 One Zone-IA) - High durability (99.999999999) in a single AZ; data lost when AZ is destroyed - 99.5% Availability - Use Cases: Storing secondary backup copies of on-premise data or data you can recreate

Amazon S3 Glacier Storage Classes - Low-cost object storage meant for archiving / backup - Pricing: price for storage + object retrieval cost

- Amazon S3 Glacier Instant Retrieval - Millisecond retrieval, great for data accessed once a quarter - Minimum storage duration of 90 days - Amazon S3 Glacier Flexible Retrieval (Formerly

Amazon S3 Glacier) - Expedited (1 to 5 minutes), Standard (3 to 5 hours), Bulk (5 to 12 hours) - free - Minimum storage duration of 90 days - Amazon S3 Glacier Deep Archive - for long term storage - Standard (12 hours), Bulk (48 Hours) - Minimum Storage duration of 180 days

12.7.3 S3 Intelligent Tiering

- Small monthly monitoring and auto-tiering fee - Moves objects automatically between Access Tiers based on usage - There are no retrieval charges in S3 Intelligent-Tiering

- Frequent Access tier (automatic): default tier - Infrequent Access Tier (automatic): objects not accessed for 30 days - Archive Instant Access tier (automatic): objects not accessed for 90 days - Archive Access tier (optional): configurable from 90 days to 700+ days - Deep Archive Access tier (optional): config from 180 days to 700+ days

S3 Storage Classes Comparision

S3 Storage Classes - Price Comparison

12.8 AWS Budgets and Cost Explorer

12.8.1 AWS Budgets

- Create budget and send alarms when costs exceeds the budget - 4 Types of budgets: Usage, Cost, Reservation, Savings Plans - For Reserved Instances (RI) - Track utilisation - Supports EC2, ElastiCache, RDS, Redshift - Up to 5 SNS notifications per budget - Can filter by: Service, Linkedin Account, Tag, Purchase Option, Instance Type, Region, Availability Zone, API Operations, etc..... - Same options as AWS Cost Explorer - 2 budgets are free than $0.002 / day / budget

12.8.2 Budget Actions

- Run actions on your behalf when a budget exceeds a certain cost or usage threshold - Supports 3 actions types - Applying an IAM Policy to a user, group or IAM role - Applying Service Control Policy (SCP) to an OU - Stop EC2 or RDS Instances - Actions can be executed automatically or require a workflow approval process - Reduced unintentional overspending in your account.

12.8.3 Centralised Budget Management

12.8.4 DeCentralised Budget Management

12.8.5 Cost Explorer

- Visualise, understand and manage your AWS costs and usage over time - Create custom reports that analyse cost and usage data - Analyse you data at a high level: total costs and usage across all accounts - Or Monthly, hourly, resource level granularity - Choose an optimal Savings PLan (to lower prices on your bill) - Forcast usage up to 12 months based on previous usage

12.9 AWS Compute Optimiser

- Reduce costs and improve performance by recommending optimal AWS resources for you workloads - Helps you choose optimal configurations and right-size your workloads (over/under provisioned) - Uses Machine Learning to analyse your resources configurations and their utilisation CloudWatch metrics - Supported resources - EC2 Instances - EC2 Auto Scaling Groups - EBS volumes - Lambda functions - Lower your costs by up to 25% - Recommendations can be exported to S3

12.9.1 Computer Optimiser - CloudWatch Agent

- Needed to analyse Memory Utilisation - Not needed for CPU, NetworkIn/Out, DiskReadOps, DiskWriteOps

12.10 EC2 Reserved Instance

- Reserved Instances in an AWS Organisation - All accounts share the Reserved Instances and Savings Plan - The payer account (Management account) of an organisation can turn off Reserved Instance (RI) discount and Savings Plans discount sharing for any accounts in that organisation, including the payer account. - Renewal of Reserved Instances - You can queue (schedule or reserve ahead of time) you reserved instances - To renew a RI, just queue an RI purchase whenever the previous one expires

Chapter 13

Migration

Chapter 14

VPC

Chapter 15

Machine Learning

15.1 Rekognition Overview

- Find objects, people, text, scenes in images and videos using ML - Facial analysis and facial search to do user verification, people counting. - Create a database of ”familiar faces” or compare against celebrities - Use cases: - Labeling - Content Moderation - Text Detection - Face Detection and Analysis (gender, age range, emotions....) - Face Search and Verification - Celebrity Recognition - Pathing (ex: for sports game analysis)

15.1.1 Amazon Rekognition - Content Moderation

- Detect content that is inappropriate, unwanted or offensive (image and videos) - Used in social media, broadcast media, advertising and e-commerce situation to create a safer user experience - Set a Minimum Confidence Threshold for items that will be flagged - Flag sensitive content for manual review in Amazon Augmented AI (A2I) - Help comply with regulations

15.2 Transcribe Overview

- Automatically convert speech to text. - Uses a deep learning process called automatic speech recognition (ASR) to convert speech to text quickly and accurately. - Automatically remove Personally Identifiable Information (PII) using Redaction. - Supports Automatic Language Identification for multilingual audio - Use cases: - transcribe customer service calls - automate closed captioning and subtitling - generate metadata for media assets to create a fully searchable archive

15.3 Polly Overview

- Turn text into lifelike speech using deep learning - Allowing you to create applications that talk

15.3.1 Amazon Polly - Lexicon & SSML

- Customise the pronunciation of words with Pronunciation lexicons - Stylized words: St3ph4ne =¿ ”Stephane” - Acronyms: AWS =¿ ”Amazon Web Services” - Upload the lexicons and use them in the SynthesizeSpeech operation. - Generate speech from plain text or from documents marked up with Speech Synthesis Markup Language (SSML) - enables more customisation -

Emphasising specific words or phrases. - Using phonetic pronunciation. - Including breathing sounds, whispering. - Using the Newscaster speaking style

15.4 Translate Overview

- Natural and accurate language translation - Allows you to localise content - such as websites and applications - for international users and to easily translate large volumes of text efficiently.

15.5 Lex + Connect Overview

- Amazon Lex: (same technology that powers Alexa) - Automatic Speech Recognition (ASR) to convert speech to text - Natural Language Understanding to recognise the intent of text, callers - Helps build chatbots and call centre bots - Amazon Connect: - Receive calls, create contact flows, cloud-based virtual contact centre - Can integrate with other CRM systems of AWS - No upfront payments, 80% cheaper than traditional contact center solutions *Hmmmm.....*

15.6 Comprehend Overview

- For Natural Language Processing - NLP - Fully managed and serverless service - Uses machine learning to find insights and relationships in text - Language of the text - Extracts key phrases, places, people, brands or events - Understands how positive or negative the text is - Analyses text using tokenization and parts of speech - Automatically organises a collection of text files by topic - Sample use cases: - Analyse customer interactions (emails) to find what leads to a positive or negative experience - Create and groups articles by topics that Comprehend will uncover

15.7 Comprehend Medical Overview

- Amazon Comprehend Medical detects and returns useful information in unstructured clinical text: - Physician's notes - Discharge summaries - Test results - Case notes - Uses NLP to detect Protected Health Information (PHI) - DetectPHI API - Store your documents in Amazon S3, analyse real-time data with Kinesis Data Firehose or use Amazon Transcribe to transcribe patient narratives into text that can be analysed by Amazon Comprehend Medical.

15.8 SageMaker Overview

- Fully managed service for developers / data scientists to build ML models - Typically difficult to do all processes in one place + provision servers

15.9 Kendra Overview

- Fully managed document search service powered by Machine Learning - Extract answers from within a document (text, pdf, HTML, PowerPoint, MS Word, FAQs) - Natural language search

capabilities - Learn from user interactions / feedback to promote preferred results (Incremental Learning) - Ability to manually fine-tune search results (importance of data, freshness, customer)

15.10 Personalise Overview

- Fully managed ML-service to build apps with real-time personalised recommendations - Example: personalised product recommendations/re-ranking, customised direct marketing Example: User bought gardening tools, provide recommendations on the next one to buy. - Same technology used by Amazon.com - Integrates into existing websites, applications, SMS, email marketing systems..... - Implement in days, not months (don't need to build, train and deploy ML solutions) - Use cases; retail stores, media and entertainment

15.11 Textract Overview

- Automatically extracts text, handwriting and data from any scanned documents using AI and ML - Extract data from forms and tables - Read and process any type of document (PDFs, images) - Use cases: - Financial Services (e.g, invoices, financial reports) - Healthcare (e.g medical records, insurance claims) - Public Sector (e.g tax forms, ID documents, passports)

15.12 Machine Learning Summary

- Rekognition - face detection, labeling, celebrity recognition
- Transcribe - audio to text (ex: subtitles)
- Polly - text to audio
- Translate - translations
- Lex - build conversational bots - chatbots
- Comprehend - natural language processing
- SageMaker - Machine learning for every developer and data scientist
- Kendra - ML-powered search engine
- Personalise - real-time personalised recommendations
- Textract - detect text and data in documents

Chapter 16

Other Services

Chapter 17

Example Preparation