14-848: CLOUD INFRASTRUCTURE

AWS OVERVIEW * LECTURE 21 * FALL 2019 * KESDEN

PUTTING THE PIECES TOGETHER

- We've looked at building a Cloud from the bottom up
- Today we look at what a fully fleshed out cloud looks like
  - We use Amazon AWS as an example
- Beginning next class we look at some special topics
  - Things that are important to Cloud Infrastructure
  - But didn't fit neatly at one level or another

AWS SERVICES

- **Computing and Cluster Management**
- Storage
- Database Services
- Networking
- Developer Tools
- Identity, Directory, Etc
- Management
- Domain-Specific

EC2: ELASTIC COMPUTE CLOUD

- You're likely very familiar with this one by now ☺
- Provides a nice friendly way to organize, configure, deploy, and pull back virtual machine instances
- Provides the ability to choose from global locations to be near users as well as to manage risk and legal jurisdictions

ECS: EC2 CONTAINER SERVICE

- As you'd expect…
  - Lets you manage the deployment of containers across an EC2 cluster
- ECR: ECS Container Registry
  - Fully managed Docker container registry

USAGE MODELS: BATCH, LAMBDA, AND SPOT INSTANCES

- *AWS Batch*:
  - Makes it easy to allocate and manage instances to do batch (not real time) data processing.
  - Manages the cluster
  - Bids for *Spot Instances or allocates EC2 instances*
- *AWS Lambda*
  - *Code run in response to trigger*
  - *Pay-as-you-go, not per instance, etc.*
- *Spot Instances* provide "excess" EC2 capacity at a discount
  - Very cheap, but can get suspended at any time
  - More for computing than providing services

SIMPLIFYING COMMON DEPLOYMENT: LIGHTSAIL AND BEANSTALK

- *Lightsail*
  - Essentially offers a simplified interface for EC2 packaged with storage and optionally database services and a simplified pricing model
  - Designed to offer a simple interface again for those with straight-forward needs, after years of adding options gave EC2 a lot of knobs
- *Beanstalk*
  - Provides a deployment front-end for AWS, enabling one to integrate application deployment directly with AWS.

AWS AUTOSCALING

- Set up policies to automatically scale AWS instances in response to demand
  - Amazon EC2 instances
  - Spot instances,
  - Amazon ECS containers
  - Amazon DynamoDB tables and indexes (more soon)
  - Amazon Aurora Replicas (more soon)

AWS SERVICES

- Computing and Cluster Management
- **Storage**
- Database Services
- Networking
- Developer Tools
- Identity, Directory, Etc
- Management
- Domain-Specific

SIMPLE STORAGE SERVICE (S3)

- Not a file system, but simpler
  - Flat name space, no hierarchy
  - Just bucket-based key-value storage
- "Eleven 9s Availability (99.999999999%)
- Stand-alone service
  - Can be used within or outside of the AWS ecosystem.

ELASTIC BLOCK STORAGE (EBS)

- General-purpose file system
  - Everything you'd expect from a general-purpose file system
- Available from within EC2 (and derived services)
  - Basically provides an external file system to EC2 instances
- Can choose backing store
  - SSDs
  - Striped SSDs (better performance)
  - Disk (cheapest)

GLACIER

- Designed for archival storage
  - Very high, but configurable latency
  - A few minutes, a few hours, or half a day (or so)
- "Active archive"
  - Can query data in place, just not quickly.
- Stores *archives* in *vaults*
  - Archive: Intended to be a .zip or .tar file, but can be anything
  - *Vault*: Names container. Unit for organization, protections, policies, etc.
- *"Eleven 9s availability" (99.999999999)*
  - Automatically replicated to 3+ availability zones

AWS STORAGE GATEWAY

- Uses a dedicated machine or VM as a local gateway to AWS storage
  - Provides a standard interface, e.g. NFS
  - Acts as a cache
  - Manages bandwidth for transfer, e.g. compression at gateway
- Storage is modeled as
  - Files
  - Volumes, e.g. like NFS
  - Virtual tapes, e.g. for backups.

AWS SERVICES

- Computing and Cluster Management
- Storage
- **Database Services**
- Networking
- Developer Tools
- Identity, Directory, Etc
- Management
- Domain-Specific

AWS SERVICES

- Computing and Cluster Management
- Storage
- **Database Services**
- Networking
- Developer Tools
- Identity, Directory, Etc
- Management
- Domain-Specific

AURORA

- Amazon's Database Engine
- Think of it as Amazon's version of MySQL
  - Oracle and Amazon might each object to that comparison ☺
  - MySQL compatible, for easy migration of apps
- Essentially, it is Amazon's database engine
  - It produces some nice benchmarks and reportedly optimizes some things that have been sore spots for MySQL.

RELATIONAL DATABASE SERVICE (RDS)

- Provides for scalable deployment of databases in the cloud
  - Not a database engine
  - Works with popular database engines, e.g. MySQL, Aurora, Oracle, PostgresSQL
- Manages deployment and scaling
  - E.g. deployment of read-only replicas
  - Easy to control access, directly and with Amazon's VPN service

DYNAMODB

- We talked about this briefly earlier
- Amazon's NoSQL database
- Automatic partitioning and scaling
- Fine-grained access control
- Event driven programming with *Lambda*

AMAZON ELASTICACHE

- We talked about this extensively earlier
- In-Memory cache service
- Uses Redis or Memcached

AWS SERVICES

- Computing and Cluster Management
- Storage
- Database Services
- **Networking**
- Developer Tools
- Identity, Directory, Etc
- Management
- Domain-Specific

VIRTUAL PRIVATE CLOUD (VPC)

- Exactly what one would expect
- Configure networks
  - Public, internet facing
  - Private, access controlled
  - VPN connections, e.g. to company site or data center

CLOUDFRONT

- Amazon's Content Delivery Network (CDN) service
- Deploy content across 144 points of presence (POPs) across globe
- Route requests to nearest deployment
- Works with AWS and on-AWS servers
- Works with many types of content
- Lambdas (event-triggered code) can execute in response to queries

ROUTE 53 DNS

- Amazon's DNS Service
- Does Everything DNS should
- And more
  - Geo-routing
  - Load balancing routing
  - Fail-over routing
  - Etc.

ELASTIC LOAD BALANCING (EBS)

- Distribute traffic across EC2 instances
- Classic load balancer
  - Bases decisions mostly upon TCP header
  - Can also look at very few parts of HTTP(S) header, e.g. X-Forwarded
  - Can handle HTTP-based sticky sessions
- Network load balancer
  - What the classic load balancer can do
  - Plus look at the IP header
- Application load balancer
  - What the network layer can do
  - Plus look much more deeply into HTTP(S) headers, work with TLS, etc.

WEB APPLICATION FIREWALL (WAF)

- Rules-based protection from common attacks
  - E.g. SQL injection
  - X-site scripting

AWS SERVICES

- Computing and Cluster Management
- Storage
- Database Services
- Networking
- **Developer Tools**
- Identity, Directory, Etc
- Management
- Domain-Specific

AWS DEVELOPER TOOLS

- AWS CodeCommit
  - Hosts private GIT repositories
- AWS CodeBuild
  - Hosts build environment
  - Compiles, builds, runs tests, etc.
- AWS CodeDeploy
  - Manages deployment, production vs test, dependencies, etc
- AWS CodePipeline
  - Continuous integration
  - Build-test-deploy according to defined pipeline
- AWS X-Ray
  - Trace requests through AWS services for debugging and optimization

AWS SERVICES

- Computing and Cluster Management
- Storage
- Database Services
- Networking
- Developer Tools
- **Management**
- Domain-Specific

MANAGEMENT

- Amazon CloudWatch
  - Monitor AWS instances, e.g. provide metrics based upon activity and logs
- Amazon EC2 Systems Manager
  - Manage instances, apply patches, update software, inventory software, etc.
- AWS CloudFormation
  - Given a template, sorts out dependencies, and allocates necessary resources.
  - Basically creates instances of packages
- AWS CloudTrail
  - Logs API calls as they travel through system and provides logs for analysis

AWS SERVICES

- Computing and Cluster Management
- Storage
- Database Services
- Networking
- Developer Tools
- **Identity, Directory, Etc.**
- Management
- Domain-Specific

IDENTITY, DIRECTORY, ETC.

- Cloud Directory
  - Like a multi-hierarchy LDAP.
  - Basically a way to organize bits of information into trees
- Active Directory
  - Microsoft AD within AWS
- Identity and Access Management (IAM) + Organizations
  - Manage users, groups, roles, and permissions
- Certificate Manager
  - Create and destroy SSL certs
- Hardware Security Module (HSM)
  - Single tenant appliances within VPN to manage keys

MANAGEMENT, *CONT*

- Config
  - Provides an audit tool for collecting information about resources and configurations
- OpsWorks
  - Chef based configuration, deployment, and management of servers
- ServiceCatalog
  - Organizes and provides a set of services for use by an organization.
  - Basically, lets an organization provide access-controlled use of its services to its constituents
- AWS Service Heath Dashboard and Personal Health Dashboard
  - Information about AWS outages and problems
  - Information about impact upon organization's deployments, in particular
- Managed Services
  - Amazon and 3rd party "Partners" provides the people that move your deployment to AWS and manage it

AWS SERVICES

- Computing and Cluster Management
- Storage
- Database Services
- Networking
- Developer Tools
- Identity, Directory, Etc.
- Management
- **Domain-Specific**

GAMING

- Gaming
  - GameLift:
    - Manages game servers.
    - Matching players to instances
    - Scales game servers for low latency
    - Protection from DDOS
  - Lumberyard
    - Amazon's Game Engine

MOBILE

- App Analytics
- App Content Delivery
- Cloud Logic
- NoSQL Database
- Push Notifications
- User Data Storage
- User Sign-in
- Connectors
- Conversational Bots
- User Engagement

https://docs.aws.amazon.com/aws-technical-content/latest/aws-overview/aws-overview.pdf

DOMAIN SPECIFIC

- Analytics and AI
  - Toolkits to access, search, stream, and analyze data
- Integrating Applications
  - Many features
  - Notable ones might include notification, messaging, and queueing services
- IoT
  - Framework to more readily use Amazon services as back-end for IOT devices
  - Integrate local processing with cloud interaction, e.g. via local and cloud lambdas (event-triggered actions)

WRAPPING IT UP

- Major cloud providers provide a phenomenal number of services
- They are based upon the things we've talked about in class
- But, they extend higher into the application space
  - And may be rich enough to be domain specific
- They also involve a lot of management tools, visualization tools, and coordination tools that we haven't covered
  - We'll touch on some of these in future lectures