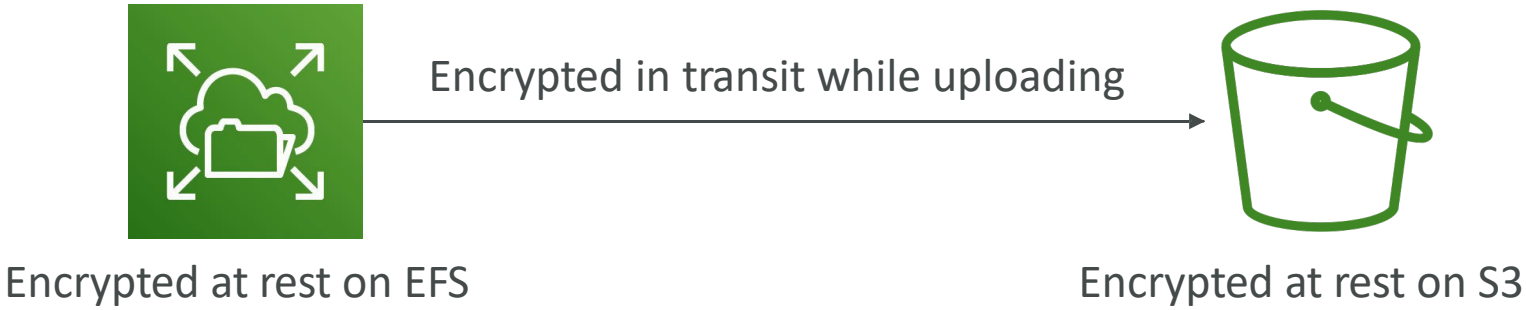
KMS

Data at rest vs. Data in transit



- At rest: data stored or archived on a device
  - On a hard disk, on a RDS instance, in S3 Glacier Deep Archive, etc.
- In transit (in motion): data being moved from one location to another
  - Transfer from on-premises to AWS, EC2 to DynamoDB, etc.
  - Means data transferred on the network
- We want to encrypt data in both states to protect it!
- For this we leverage encryption keys

AWS KMS (Key Management Service)

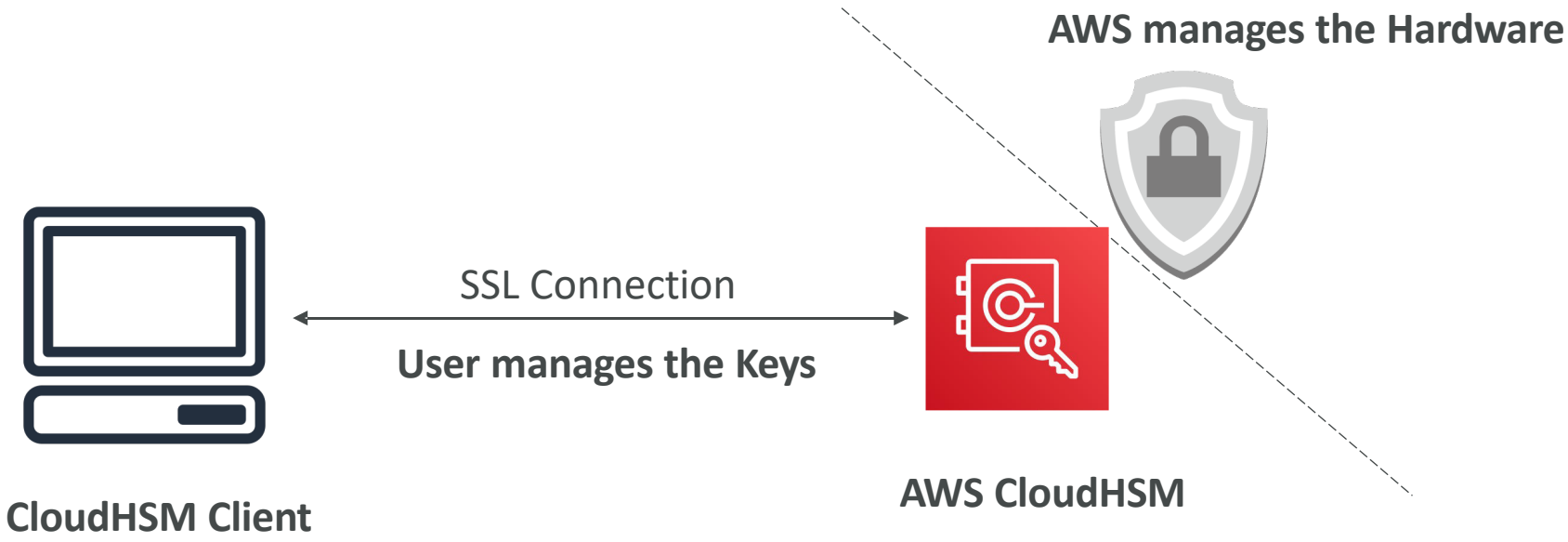
- Anytime you hear “encryption” for an AWS service, it’s most likely KMS
- KMS = AWS manages the encryption keys for us
- Encryption Opt-in:
  - EBS volumes: encrypt volumes
  - S3 buckets: Server-side encryption of objects
  - Redshift database: encryption of data
  - RDS database: encryption of data
  - EFS drives: encryption of data
- Encryption Automatically enabled:
  - CloudTrail Logs
  - S3 Glacier
  - Storage Gateway

CloudHSM

- KMS => AWS manages the software for encryption
- CloudHSM => AWS provisions encryption hardware
- Dedicated Hardware (HSM = Hardware Security Module)
- You manage your own encryption keys entirely (not AWS)
- HSM device is tamper resistant, FIPS 140-2 Level 3 compliance

Sample HSM device

CloudHSM Diagram



Types of Customer Master Keys: CMK

- Customer Managed CMK:
  - Create, manage and used by the customer, can enable or disable
  - Possibility of rotation policy (new key generated every year, old key preserved)
  - Possibility to bring-your-own-key
- AWS managed CMK:
  - Created, managed and used on the customer's behalf by AWS
  - Used by AWS services (aws/s3, aws/ebs, aws/redshift)
- AWS owned CMK:
  - Collection of CMKs that an AWS service owns and manages to use in multiple accounts
  - AWS can use those to protect resources in your account (but you can't view the keys)
- CloudHSM Keys (custom keystore):
  - Keys generated from your own CloudHSM hardware device
  - Cryptographic operations are performed within the CloudHSM cluster

SECRETS MANAGER

AWS Secrets Manager

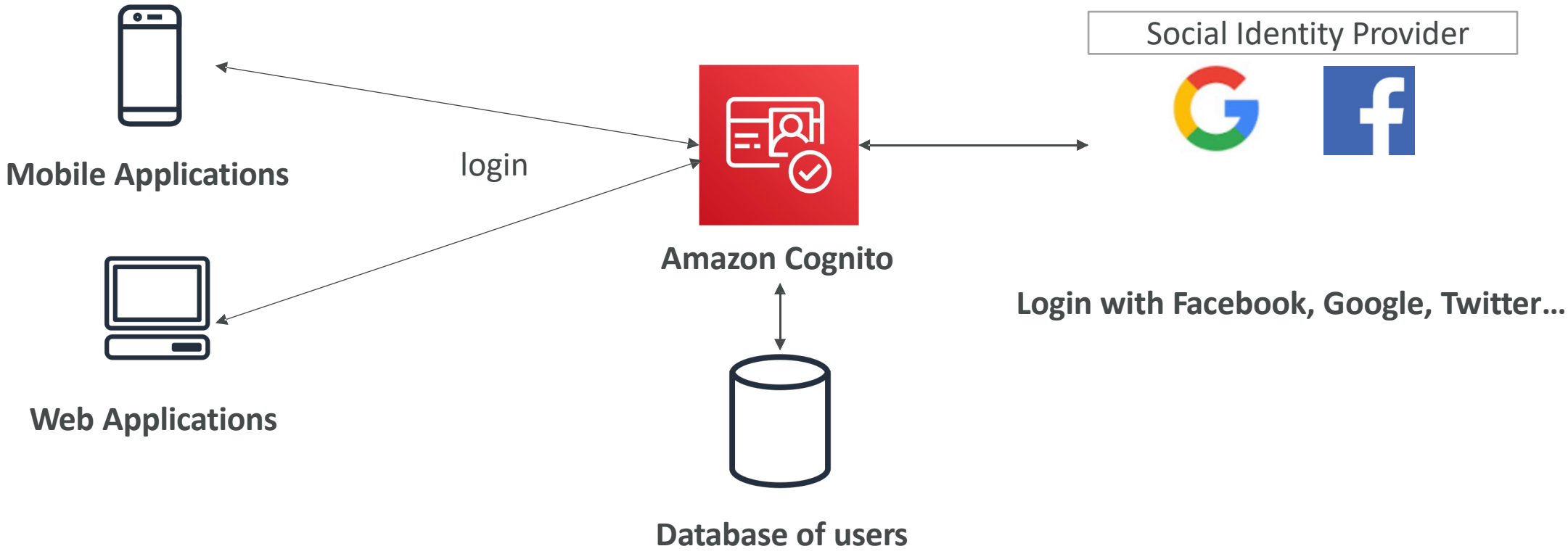
- Newer service, meant for storing secrets
- Capability to force rotation of secrets every X days
- Automate generation of secrets on rotation (uses Lambda)
- Integration with Amazon RDS (MySQL, PostgreSQL, Aurora)
- Secrets are encrypted using KMS

- Mostly meant for RDS integration

COGNITO

Amazon Cognito (simplified)

- Identity for your Web and Mobile applications users (potentially millions)
- Instead of creating them an IAM user, you create a user in Cognito



CLOUD FORMATION

What is CloudFormation

- CloudFormation is a declarative way of outlining your AWS Infrastructure, for any resources (most of them are supported).
- For example, within a CloudFormation template, you say:
  - I want a security group
  - I want two EC2 instances using this security group
  - I want an S3 bucket
  - I want a load balancer (ELB) in front of these machines

- Then CloudFormation creates those for you, in the right order, with the exact configuration that you specify

Benefits of AWS CloudFormation (1/2)

- Infrastructure as code
  - No resources are manually created, which is excellent for control
  - Changes to the infrastructure are reviewed through code

- Cost
  - Each resources within the stack is tagged with an identifier so you can easily see how much a stack costs you
  - You can estimate the costs of your resources using the CloudFormation template
  - Savings strategy: In Dev, you could automation deletion of templates at 5 PM and recreated at 8 AM, safely

Benefits of AWS CloudFormation (2/2)

- Productivity
  - Ability to destroy and re-create an infrastructure on the cloud on the fly
  - Automated generation of Diagram for your templates!
  - Declarative programming (no need to figure out ordering and orchestration)

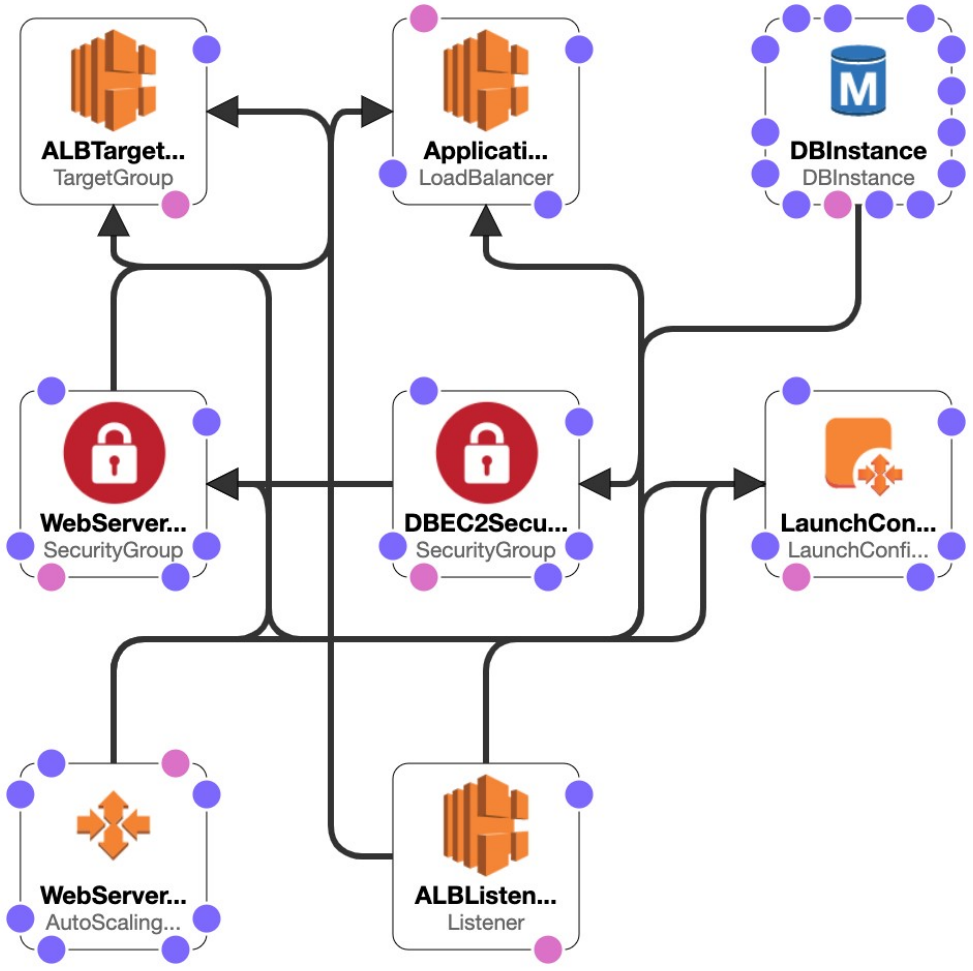
- Don’t re-invent the wheel
  - Leverage existing templates on the web!
  - Leverage the documentation

- Supports (almost) all AWS resources:
  - Everything we’ll see in this course is supported

CloudFormation Stack Designer

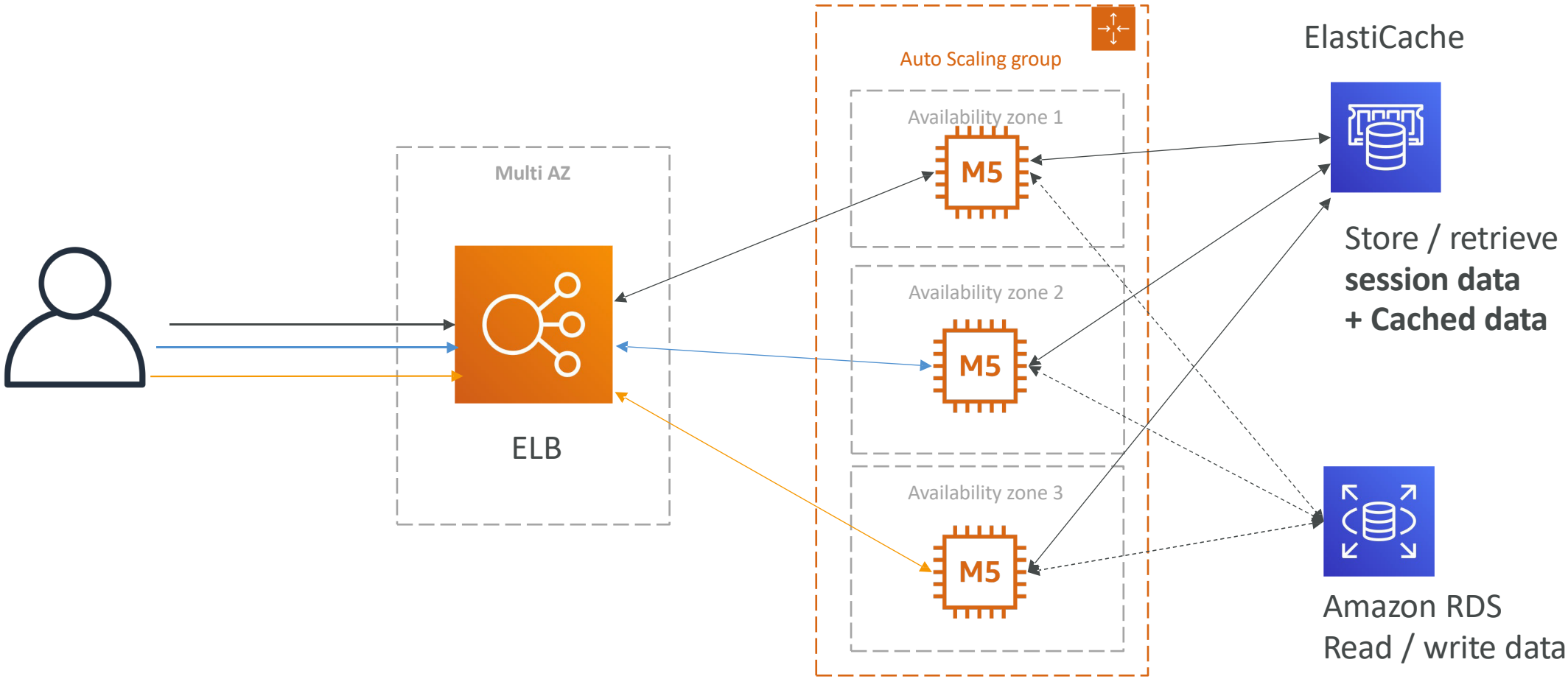
- Example: WordPress CloudFormation Stack

- We can see all the resources
- We can see the relations between the components



ELASTIC BEANSTALK

Typical architecture: Web App 3-tier



Developer problems on AWS

- Managing infrastructure
- Deploying Code
- Configuring all the databases, load balancers, etc
- Scaling concerns

- Most web apps have the same architecture (ALB + ASG)
- All the developers want is for their code to run!
- Possibly, consistently across different applications and environments

AWS Elastic Beanstalk Overview

- Elastic Beanstalk is a developer centric view of deploying an application on AWS
- It uses all the component's we've seen before: EC2, ASG, ELB, RDS, etc...
- But it's all in one view that's easy to make sense of!
- We still have full control over the configuration

- Beanstalk = Platform as a Service (PaaS)
- Beanstalk is free but you pay for the underlying instances

Elastic Beanstalk

- Managed service
  - Instance configuration / OS is handled by Beanstalk
  - Deployment strategy is configurable but performed by Elastic Beanstalk
  - Capacity provisioning
  - Load balancing & auto-scaling
  - Application health-monitoring & responsiveness

- Just the application code is the responsibility of the developer

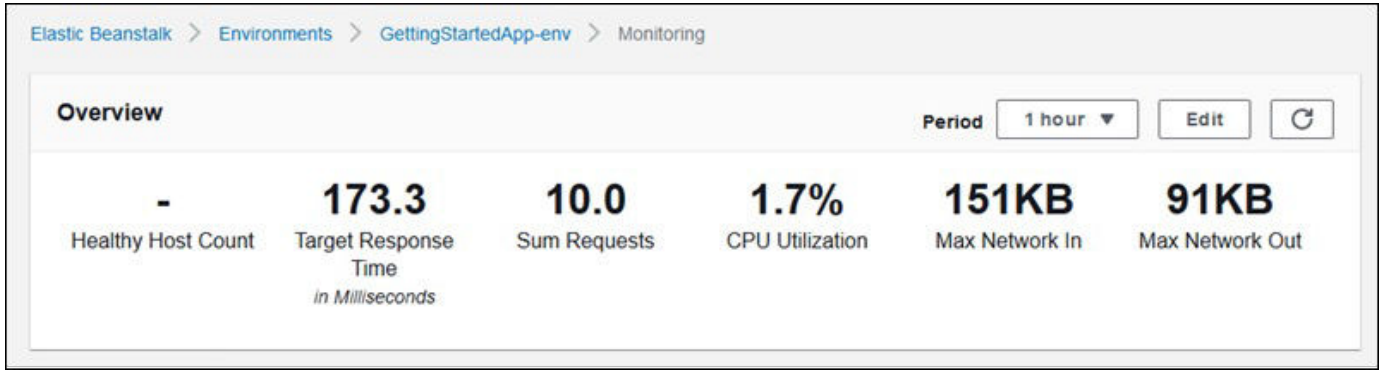
- Three architecture models:
  - Single Instance deployment: good for dev
  - LB + ASG: great for production or pre-production web applications
  - ASG only: great for non-web apps in production (workers, etc..)

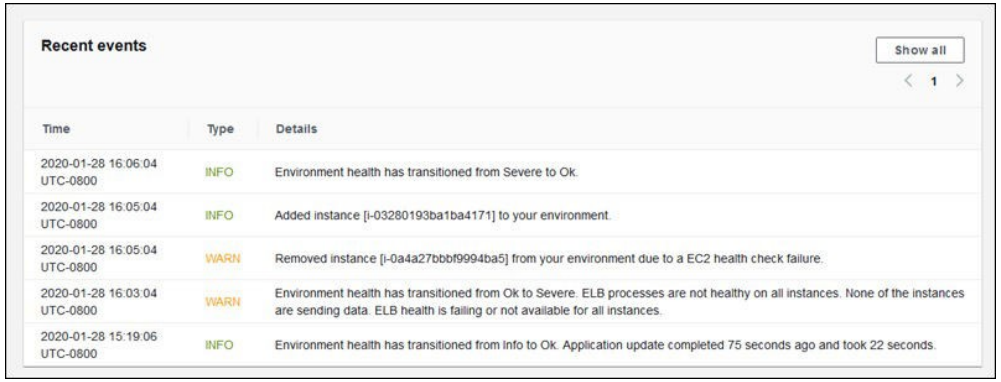
Elastic Beanstalk

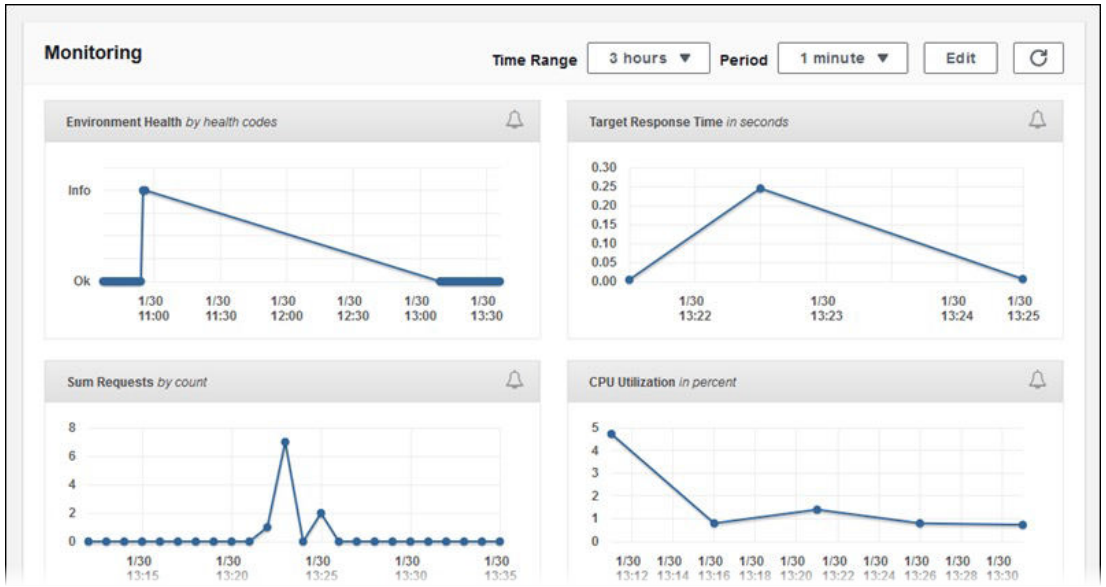
- Support for many platforms:
  - Go
  - Java SE
  - Java with Tomcat
  - .NET on Windows Server with IIS
  - Node.js
  - PHP
  - Python
  - Ruby
  - Packer Builder
  - Single Container Docker
  - Multi-Container Docker
  - Preconfigured Docker
- If not supported, you can write your custom platform (advanced)

Elastic Beanstalk – Health Monitoring

- Health agent pushes metrics to CloudWatch
- Checks for app health, publishes health events







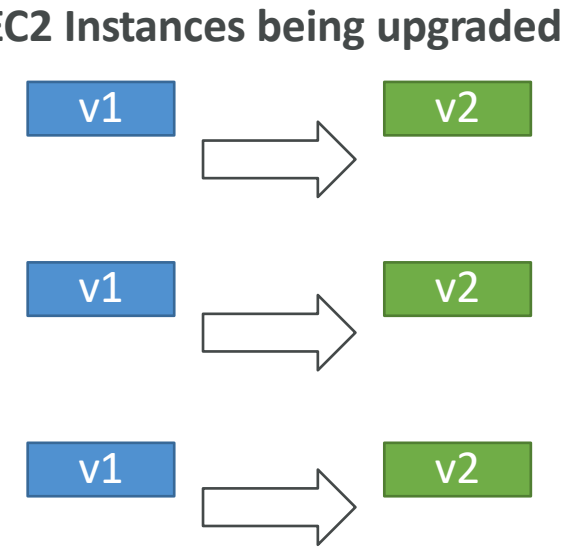
CODE - COMMIT, BUILD, PIPELINE, ARTIFACT

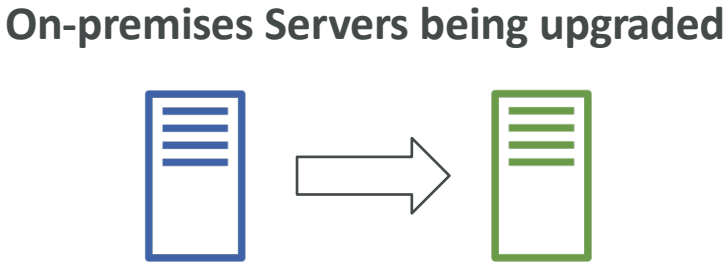
AWS CodeDeploy

- We want to deploy our application automatically

- Works with EC2 Instances
- Works with On-Premises Servers
- Hybrid service

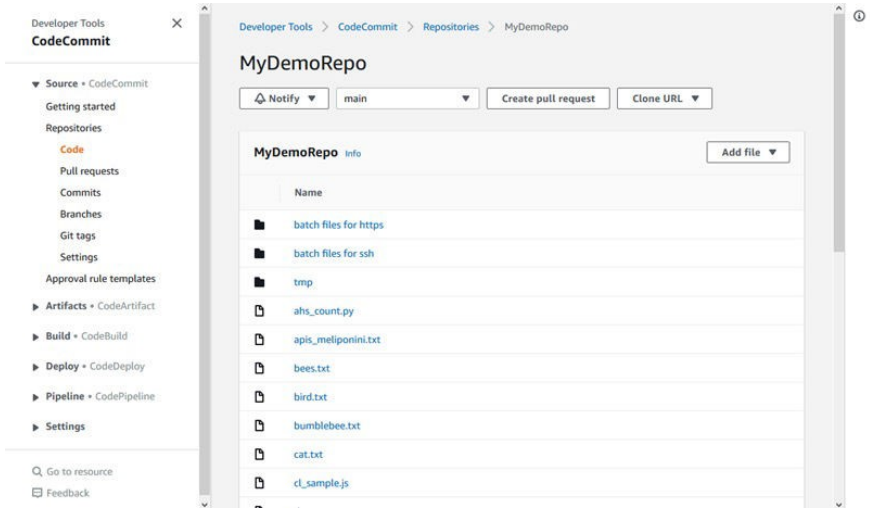
- Servers / Instances must be provisioned and configured ahead of time with the CodeDeploy Agent





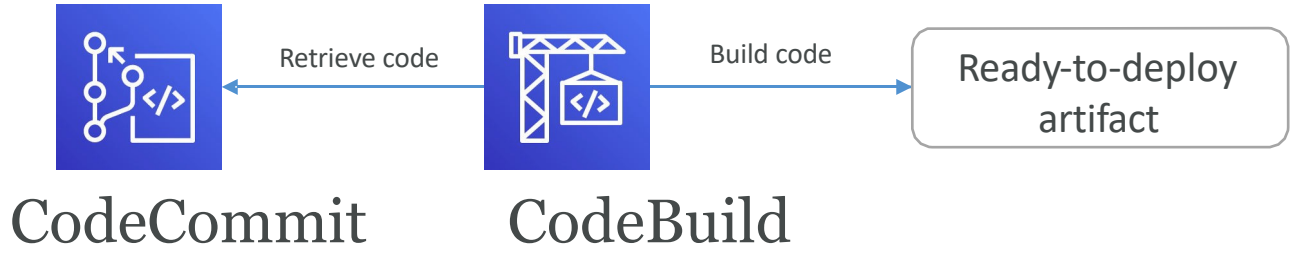
AWS CodeCommit

- Before pushing the application code to servers, it needs to be stored somewhere
- Developers usually store code in a repository, using the Git technology
- A famous public offering is GitHub, AWS' competing product is CodeCommit
- CodeCommit:
  - Source-control service that hosts Git-based repositories
  - Makes it easy to collaborate with others on code
  - The code changes are automatically versioned
- Benefits:
  - Fully managed
  - Scalable & highly available
  - Private, Secured, Integrated with AWS



AWS CodeBuild

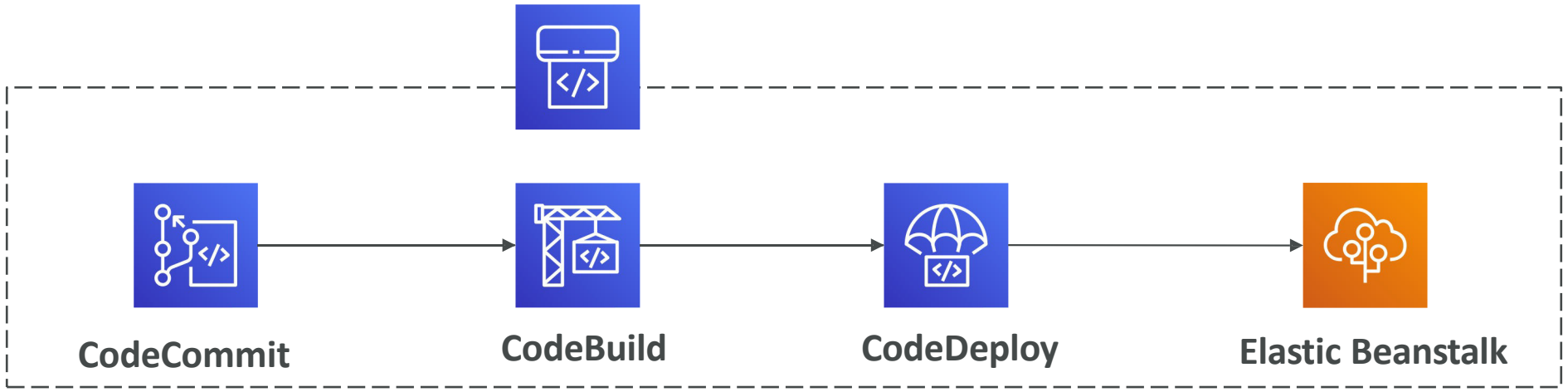
- Code building service in the cloud (name is obvious)
- Compiles source code, run tests, and produces packages that are ready to be deployed (by CodeDeploy for example)



- Benefits:
  - Fully managed, serverless
  - Continuously scalable & highly available
  - Secure
  - Pay-as-you-go pricing – only pay for the build time

AWS CodePipeline

- Orchestrate the different steps to have the code automatically pushed to production
  - Code => Build => Test => Provision => Deploy
  - Basis for CICD (Continuous Integration & Continuous Delivery)
- Benefits:
  - Fully managed, compatible with CodeCommit, CodeBuild, CodeDeploy, Elastic Beanstalk, CloudFormation, GitHub, 3rd-party services (GitHub…) & custom plugins…
  - Fast delivery & rapid updates



AWS CodeArtifact

- Software packages depend on each other to be built (also called code dependencies), and new ones are created
- Storing and retrieving these dependencies is called artifact management
- Traditionally you need to setup your own artifact management system
- CodeArtifact is a secure, scalable, and cost-effective artifact management for software development
- Works with common dependency management tools such as Maven, Gradle, npm, yarn, twine, pip, and NuGet
- Developers and CodeBuild can then retrieve dependencies straight from CodeArtifact

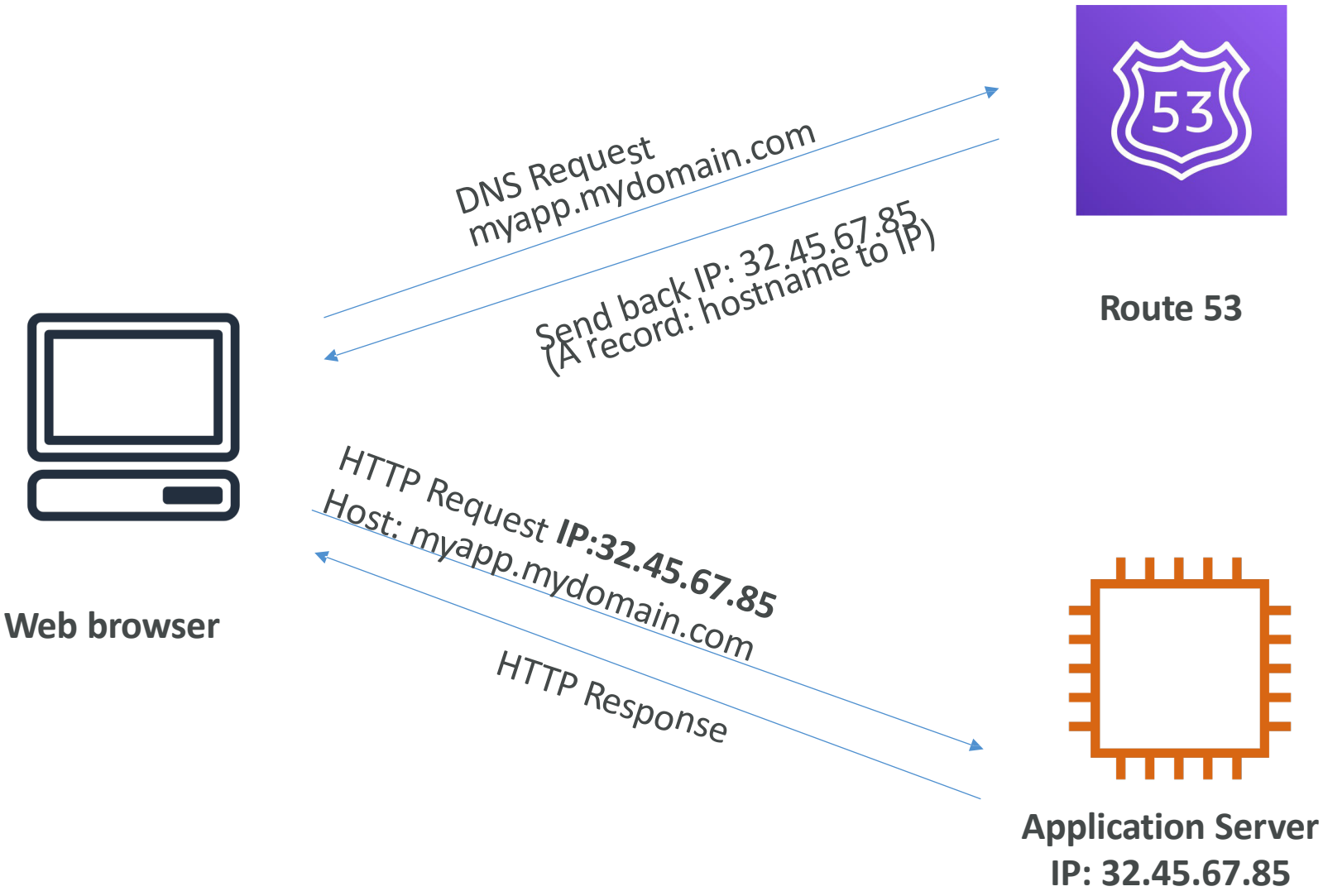
ROUTE 53

Amazon Route 53 Overview

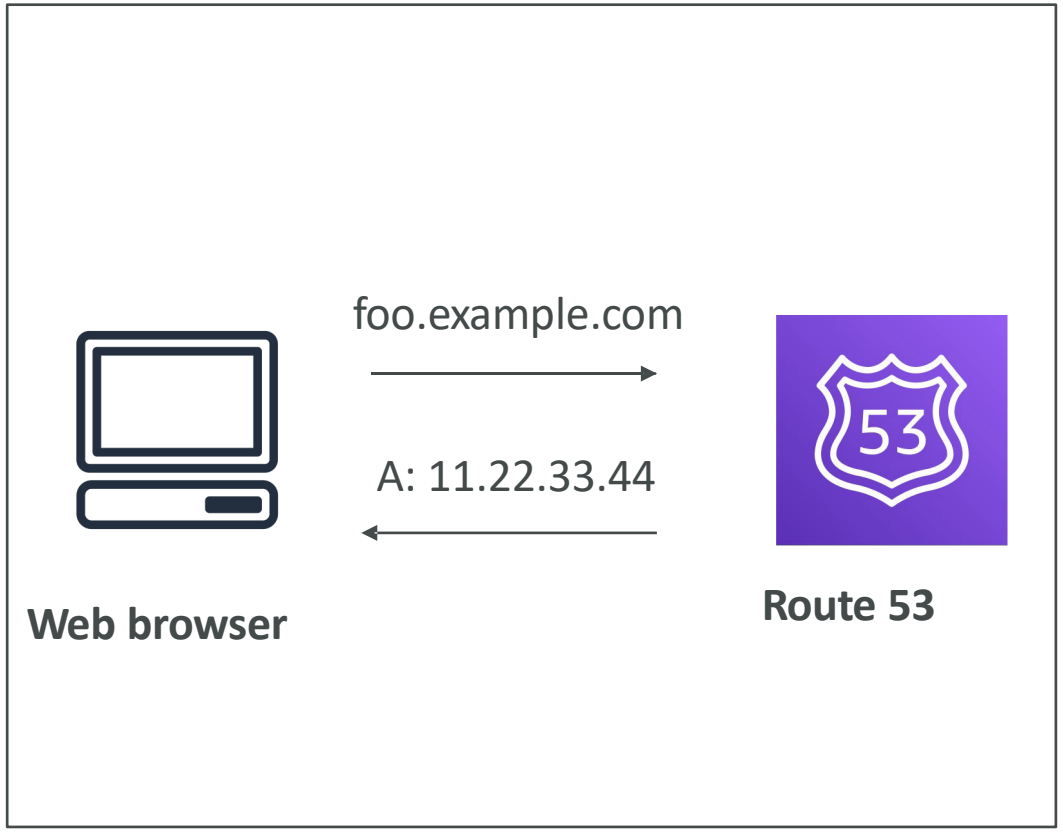
- Route53 is a Managed DNS (Domain Name System)
- DNS is a collection of rules and records which helps clients understand how to reach a server through URLs.

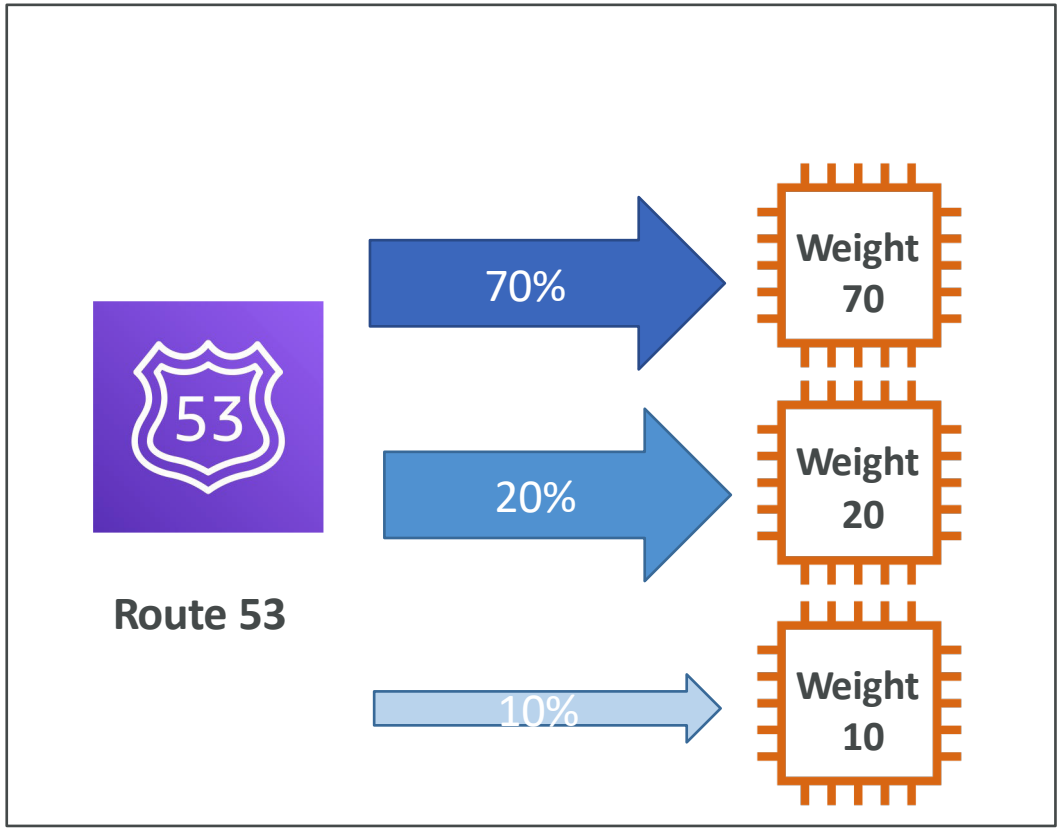
- In AWS, the most common records are:
  - www.google.com => 12.34.56.78 == A record (IPv4)
  - www.google.com => 2001:0db8:85a3:0000:0000:8a2e:0370:7334 == AAAA IPv6
  - search.google.com => www.google.com == CNAME: hostname to hostname
  - example.com => AWS resource == Alias (ex: ELB, CloudFront, S3, RDS, etc...)

Route 53 – Diagram for A Record

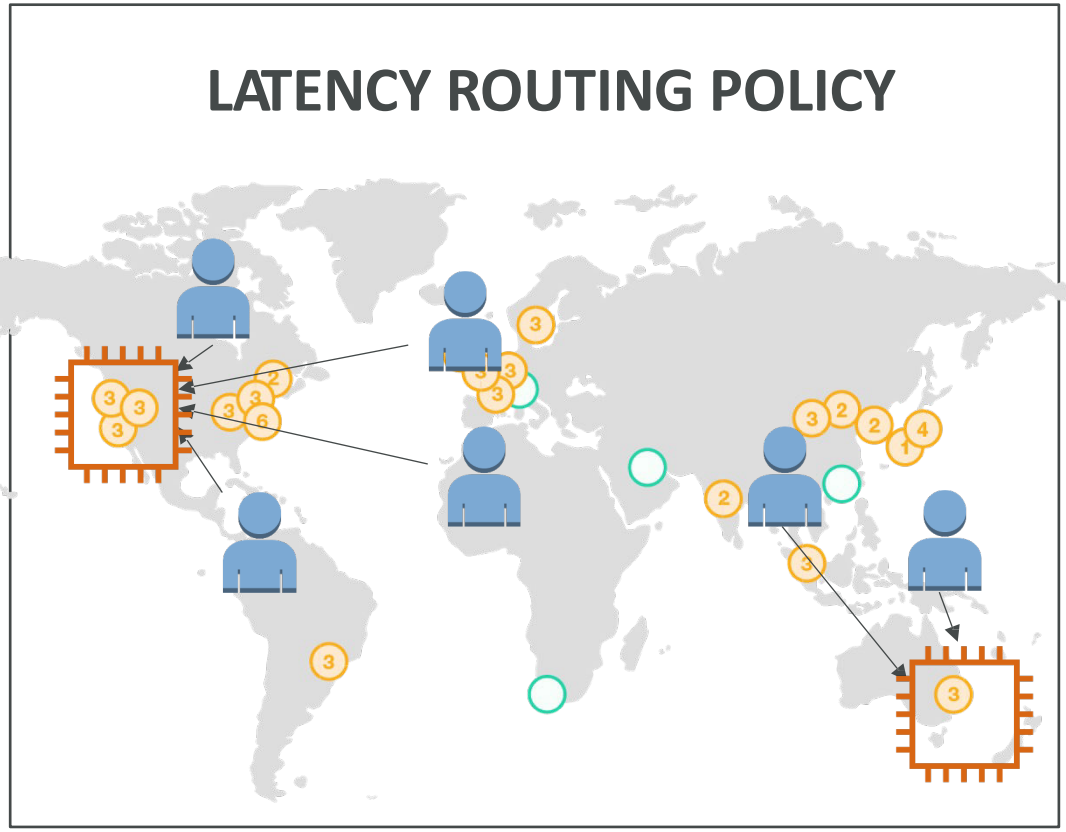


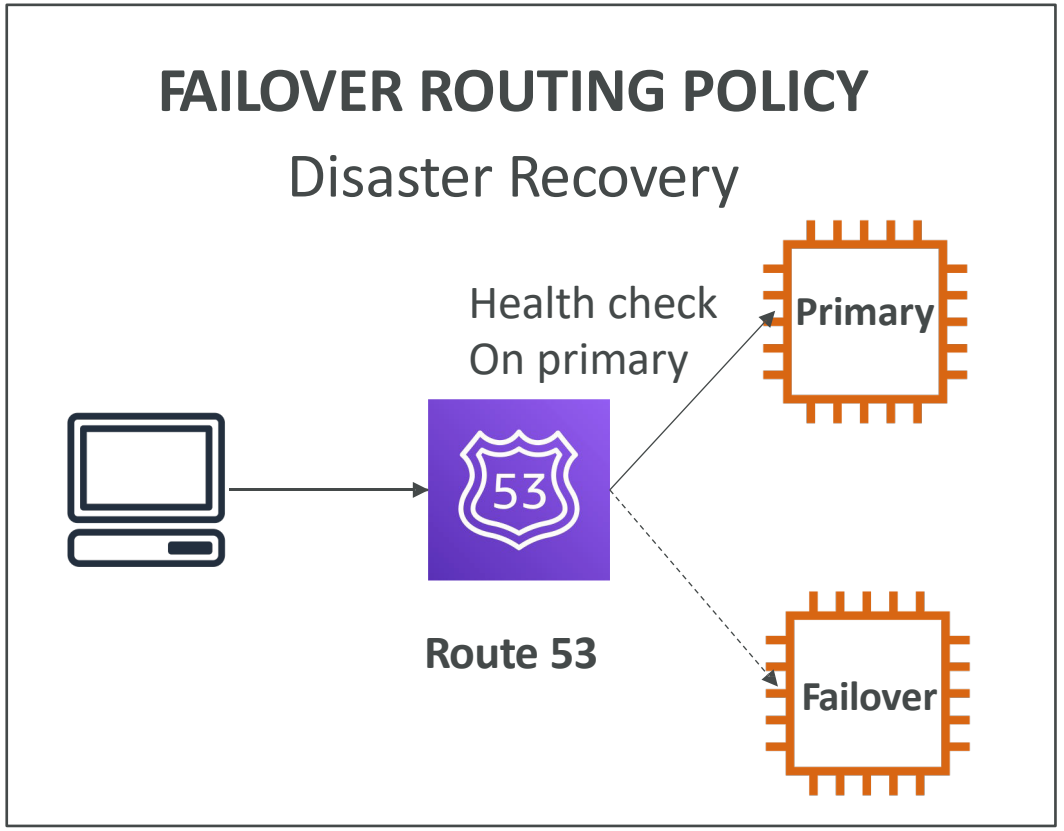
Route 53 Routing Policies





Route 53 Routing Policies

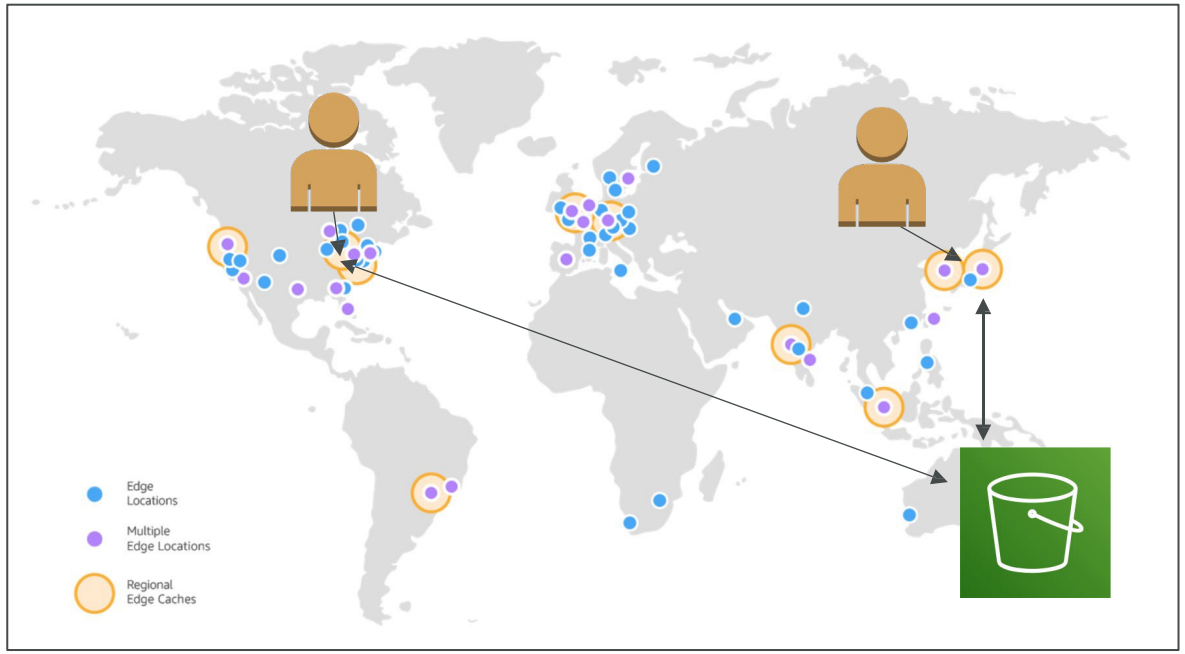




CLOUDFRONT

AWS CloudFront

- Content Delivery Network (CDN)
- Improves read performance, content is cached at the edge
- Improves users experience
- 216 Point of Presence globally (edge locations)
- DDoS protection (because worldwide), integration with Shield, AWS Web Application Firewall



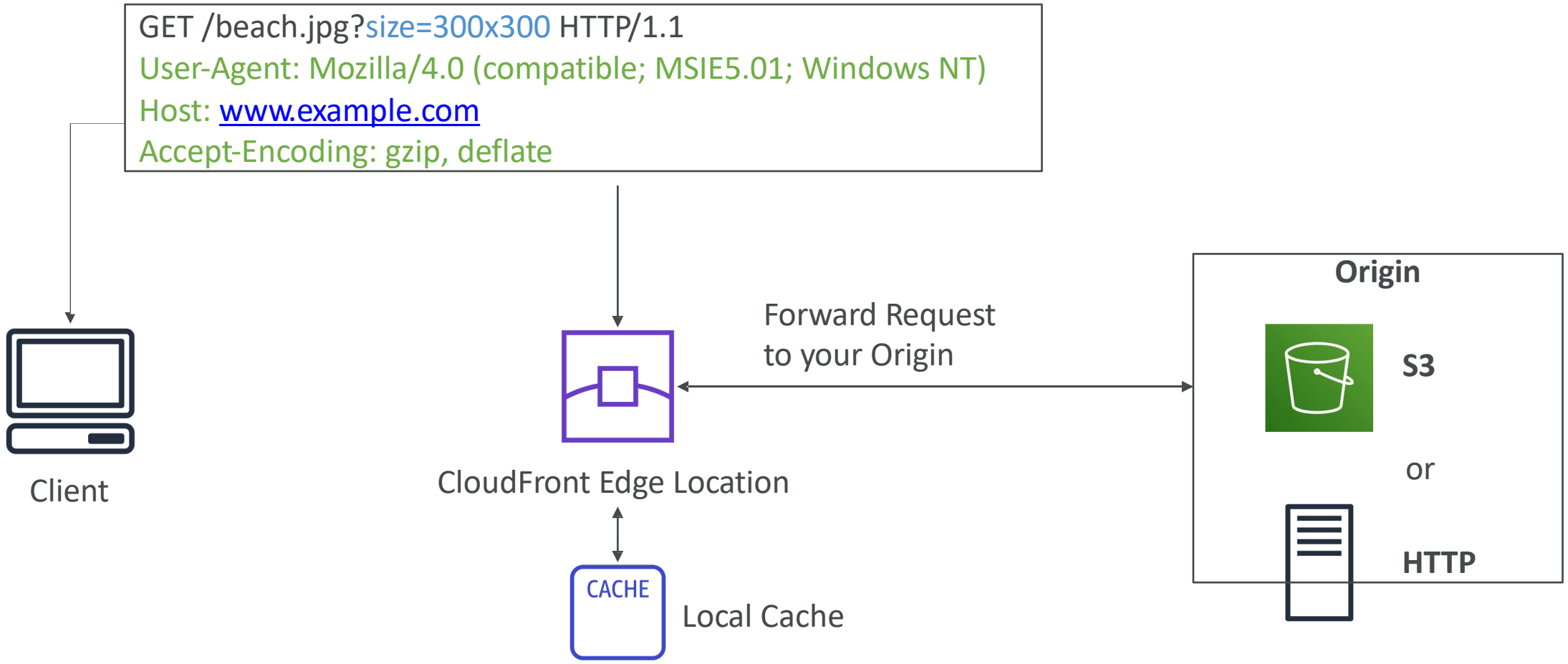
Source: https://aws.amazon.com/cloudfront/features/?nc=sn&loc=2

CloudFront – Origins

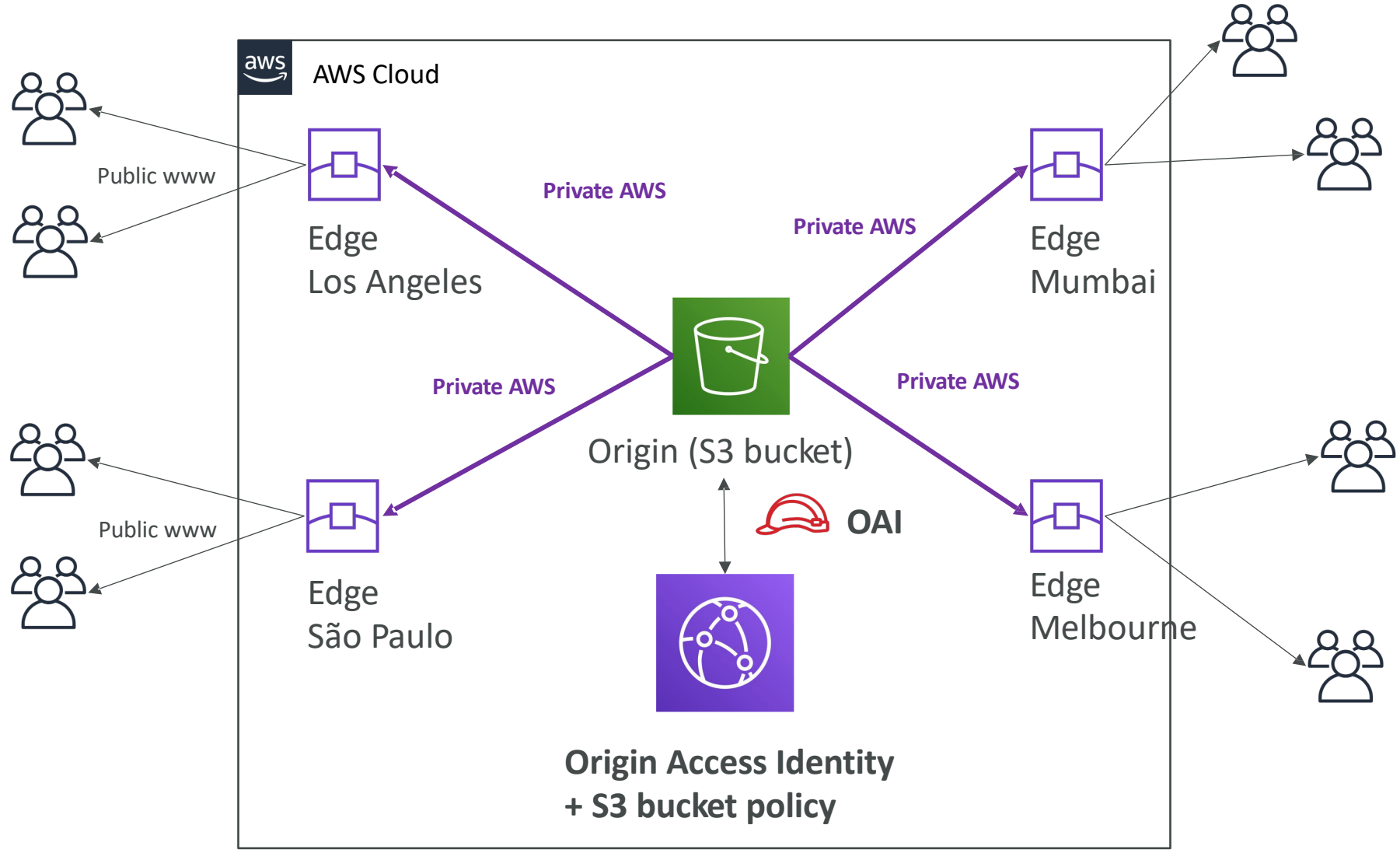
- S3 bucket
  - For distributing files and caching them at the edge
  - Enhanced security with CloudFront Origin Access Identity (OAI)
  - CloudFront can be used as an ingress (to upload files to S3)

- Custom Origin (HTTP)
  - Application Load Balancer
  - EC2 instance
  - S3 website (must first enable the bucket as a static S3 website)
  - Any HTTP backend you want

CloudFront at a high level



CloudFront – S3 as an Origin



CloudFront vs S3 Cross Region Replication

- CloudFront:
  - Global Edge network
  - Files are cached for a TTL (maybe a day)
  - Great for static content that must be available everywhere

- S3 Cross Region Replication:
  - Must be setup for each region you want replication to happen
  - Files are updated in near real-time
  - Read only
  - Great for dynamic content that needs to be available at low-latency in few regions