AWS BASICS

Marko Baric
@markobaric

WHY CLOUD?

#1
AGILE INVESTING

No upfront investment
No capacity planning
Spend as you grow
Reduced running costs

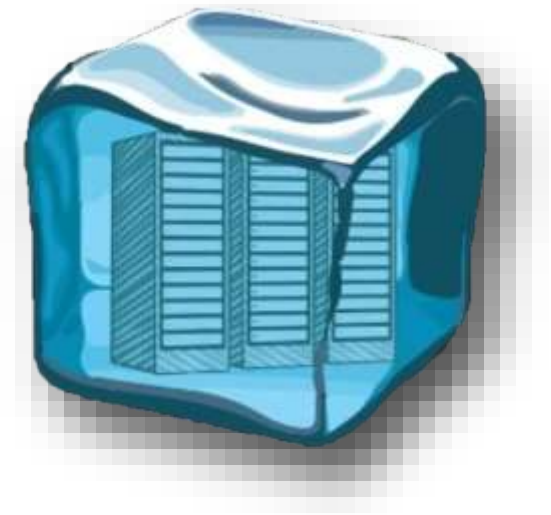
#2
INNOVATION

Experiment
Innovate and refactor rapidly
Fail often without risk

#3
FOCUS ON BUSINESS PROBLEM

No need to build fundamental services (computing, storage, database, messaging, analysis, etc.)
Instead build your app

RUNNING IN TRADITIONAL DATACENTERS



- Need to predict capacity
- Procurement of servers takes months to complete
- Large investment upfront - might not pay off
- On-premises datacenters provide basic compute and storage only
- Once you build it, you're frozen in time

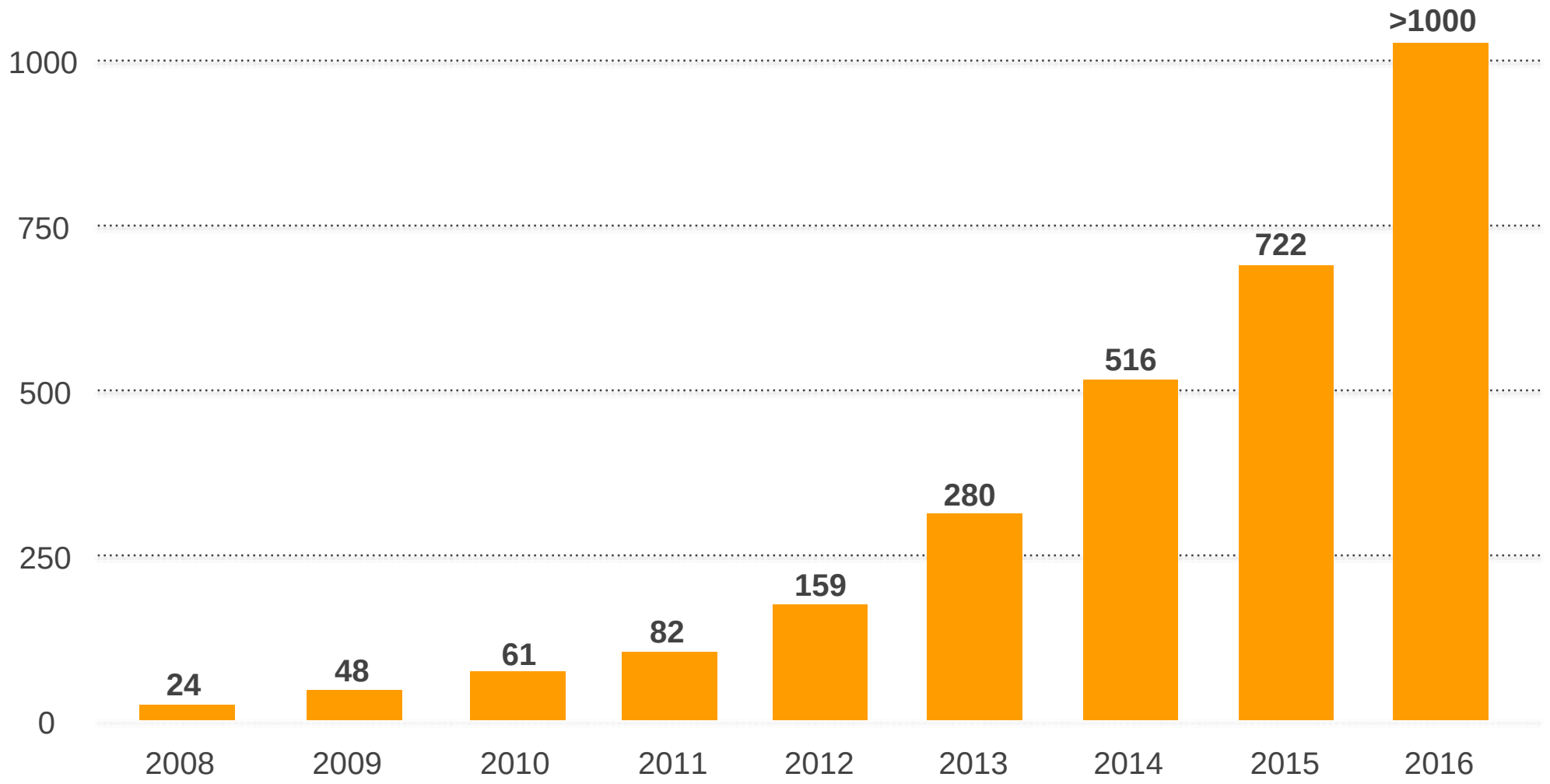
RUNNING IN CLOUD

- No upfront costs
- Pay as you grow
- Vast ecosystem of platform services
- Ready to be used immediately - no delays for procurement
- More secure
- More robust
- Built with all best practices in mind
- Constantly improved and extended with new features

You can't be competitive today
if you're not running in cloud !

INCREDIBLE
PACE OF INNOVATION

NUMBER OF NEW SERVICES AND FEATURES RELEASED



>2,400

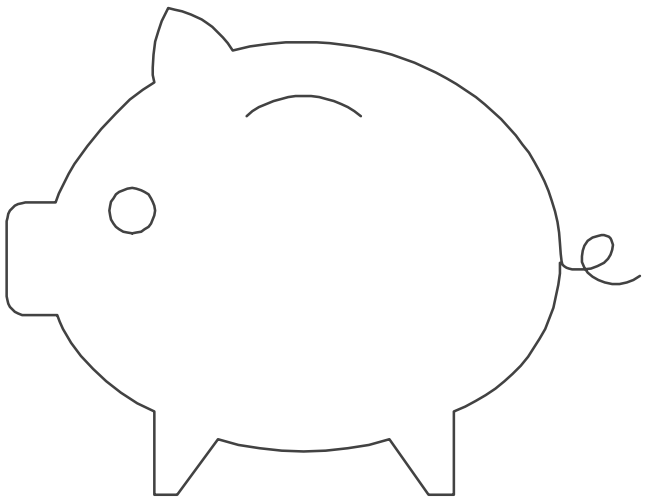
Services and Features

55% growth

year-over-year

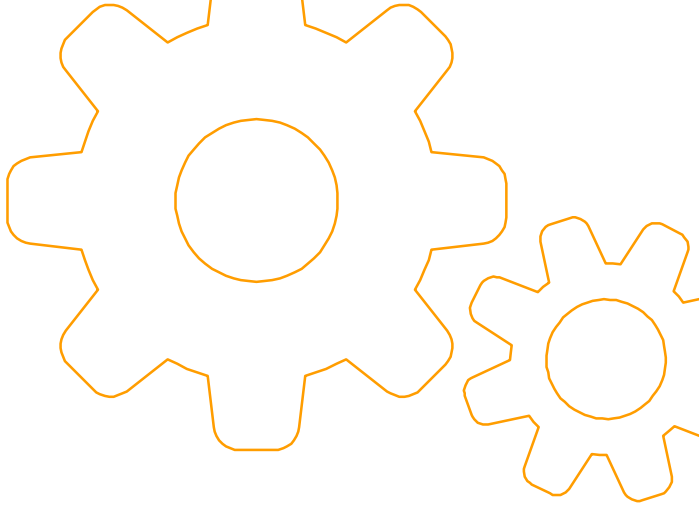
$11.1 billion

revenue in 2016



GARTNER MAGIC QUADRANT

- "furthest completeness of vision"
- "highest ability to execute"
- "broadest range of customers and use cases"



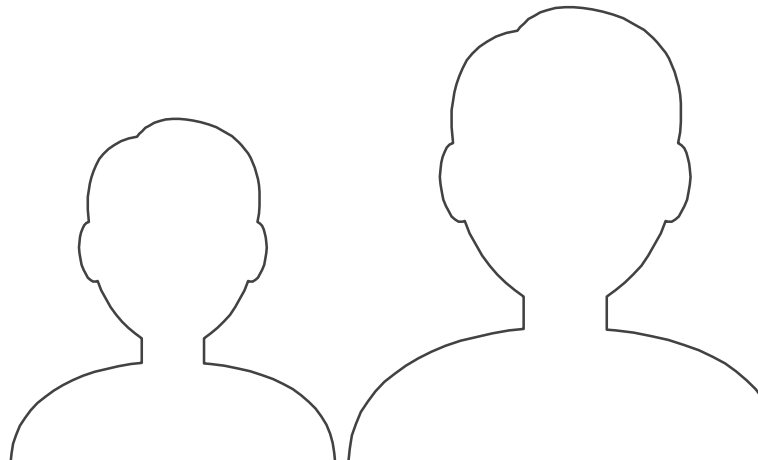
AWS runs

10x more compute power

than all other cloud providers combined!

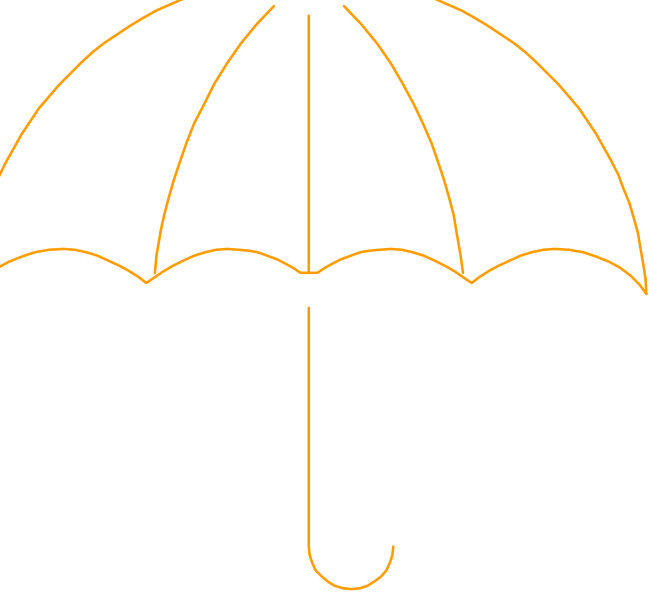
>1 million

active customers (companies)



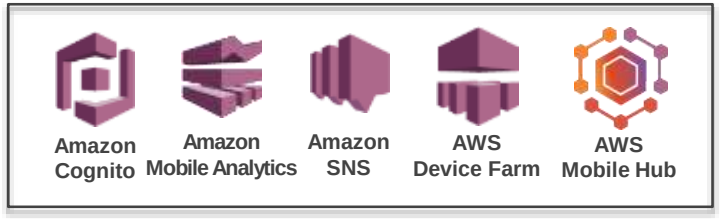
STARTUP CUSTOMERS

ENTERPRISE CUSTOMERS



How broad is AWS platform?

Mobile Services

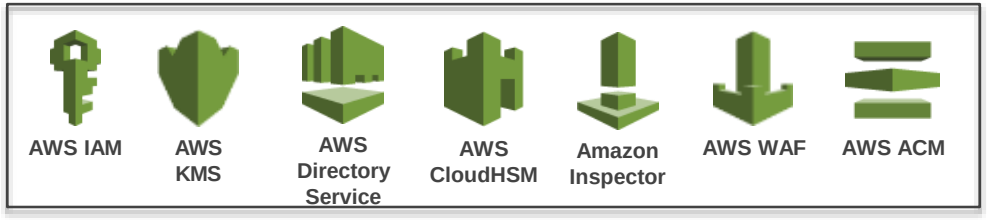


Application Services

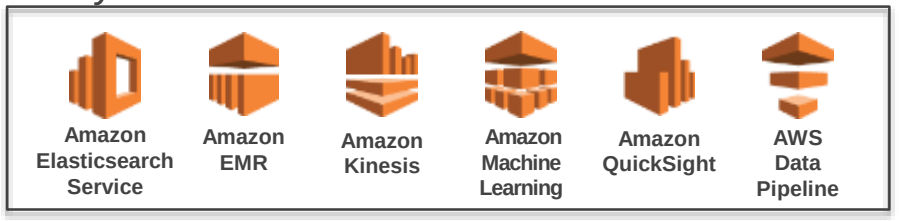
Enterprise Applications

IoT

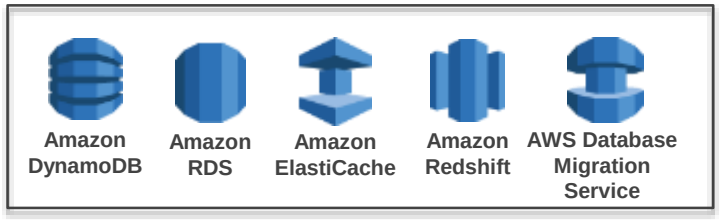
Security and Identity



Analytics

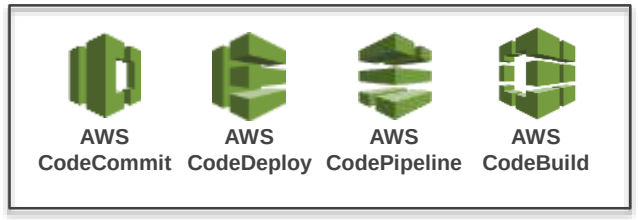


Database

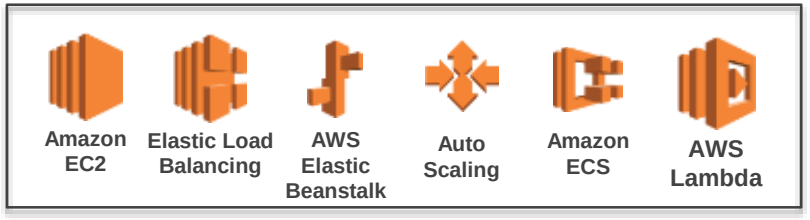


Management tools

Developer tools



Compute

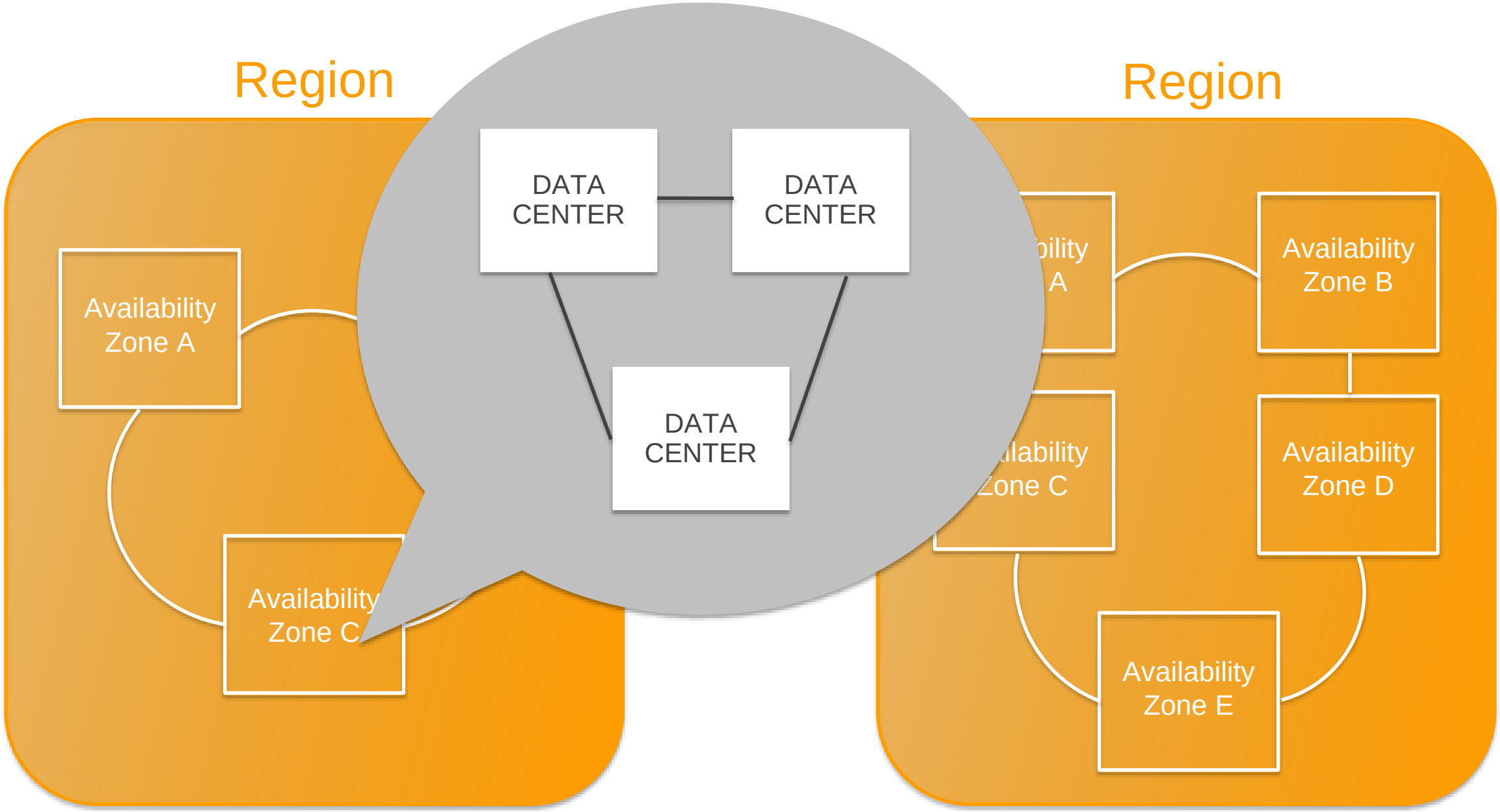


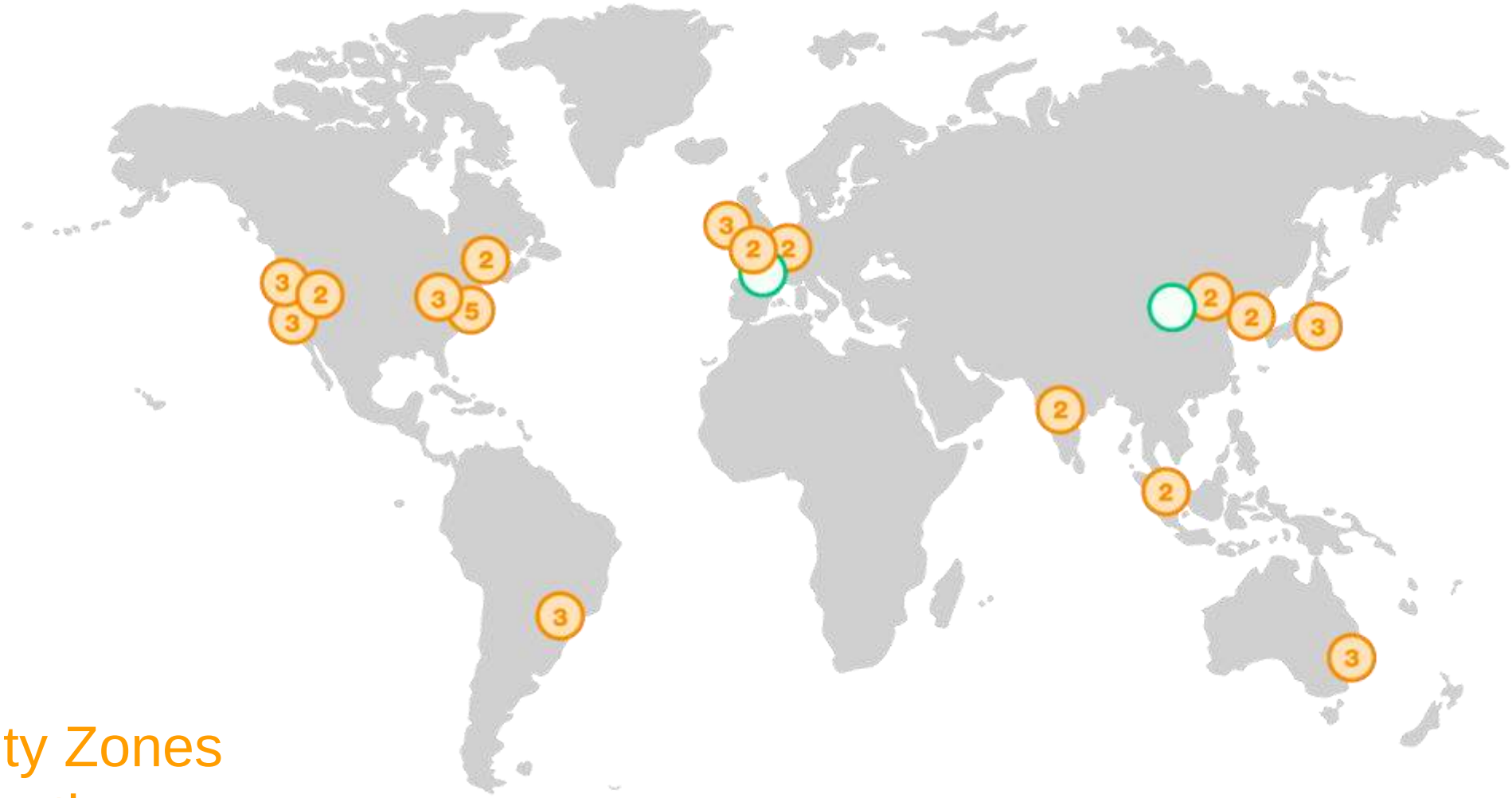
Networking

Storage and Content Delivery

Now let's go technical 

AWS GLOBAL INFRASTRUCTURE


Region
Region
DATA CENTER
DATA CENTER
DATA CENTER
Availability Zone A
Availability Zone C
Availability Zone B
Availability Zone D
Availability Zone E



16 Regions
42 Availability Zones
69 Edge Locations

OVERVIEW OF SERVICES

...get ready for abbreviation hell

COMPUTE SERVICES

Amazon EC2

Amazon EC2

- Virtual servers in the cloud
- Multiple operating systems supported:
  - Amazon Linux, Red Hat Enterprise, SUSE, Ubuntu, Windows Server
- Range of different machine flavours:
  - General purpose, memory optimised, I/O optimised, GPU optimised, FPGA
- Range of different instance types:
  - T2, M3, M4, C3, C4, R3, R4, X1, P2, G2, F1, I3, D2
- Range of different compute power:
  - nano, micro, small, medium, large, xlarge, 2xlarge, 3xlarge, etc....
  - From 0.5 GB RAM / 1 vCPU
  - To 1952 GB RAM / 128 vCPU
- Range of pricing models:
  - On-demand instances, reserved instances, spot instances, dedicated hosts, placement groups, etc.

Auto Scaling

- Scale in and out based on your own scaling policies
- Manual, dynamic scaling or scheduled scaling
- Instance health checks
- Monitoring
- Logging of scaling events
- Cooldown periods
- Lifecycle hooks

Elastic Load Balancing

Elastic
Load Balancing

- Distributes traffic across fleet of EC2 instances
- Highly scalable - can handle any load
- Instance health checks
- SSL offloading
- Supports IPv6, HTTP/2 and WebSockets

- Two flavours:
  - Classic balancing (on network level - layer 3)
  - Application balancing (on application level - layer 7)

AWS Lambda

AWS
Lambda

- *Function-as-a-service*
- Event-driven serverless compute service
- Pay only for execution time, never pay for idle time
- Automatically scales to handle ANY load
- Supports functions in JavaScript, C#, Java, Python
- Execution monitoring

NETWORKING SERVICES

Amazon VPC

Amazon
VPC

- Logically isolated section of AWS cloud
- Think of it as your own "private virtual data centre in the cloud"
- Define your own virtual networking, subnets, route tables, internet gateways, NATs, access lists, etc.
- Manage group of your resources as one unit

Amazon Route 53

Amazon
Route 53

- Managed DNS service
- Public and private hosted zones
- Endpoint health checks
- Routing options:
  - latency based
  - geolocation based
  - weighted routing

STORAGE SERVICES

Amazon S3

Amazon S3

- Low latency, high throughput "object" storage
- Unlimited size
- Max object size for a singe object is 5TB
- 99.999999999% durability
- Cross-AZ / Cross-region replication
- Ideal for backup storage, bulk file storage, data lakes, analytics, etc.
- Encryption in-transit and at-rest
- Object versioning
- Object lifecycle management
- 3 storage classes: *standard*, *infrequent access* and *reduced redundancy*

Amazon Glacier

Amazon Glacier

- Petabyte-scale storage
- Ideal for infrequently accessed files (e.g. archives)
- Extremely low-cost
  - $0.004 per gigabyte per month

- Two price classes:
  - *Standard* - data retrieval time 3-5 hours
  - *Expedited* - data retrieval time 1-5 minutes

CONTENT DELIVERY SERVICES

Amazon CloudFront

Amazon
CloudFront

- Global content delivery network (CDN)
- Advanced security features (access policies, request signing, cookie signing)
- Supports IPv6, HTTP/2 and WebSockets
- Edge location available in any relevant place in the world
- You can even run application logic on the edge location

AWS WAF

AWS WAF

- Blocks traffic at the edge
- Protects from common web exploits
- Customisable filters based on:
  - request headers
  - request content
  - source IP
  - request size
- Automatic SQL injection protection, automatic XSS protection
- Can be set up on ELB or CloudFront
- AWS Shield - advanced DDoS detection and protection

DATABASE SERVICES

Amazon DynamoDB

Amazon
DynamoDB

- Fully managed NoSql data store
- Seamless scaling to support any load
- Supports any number of indexes with projected columns
- Automatic data partitioning
- Can stream data changes to other sources

Amazon RDS

Amazon
RDS

- Fully managed relational database service
- Supports MySQL, MariaDB, SQL Server, PostgreSQL, Oracle and Amazon Aurora
- Automatic backups with "point in time" restore
- Multi-AZ deployments with synchronous read replicas
- Automatic failover
- Advanced DB instance monitoring
- Amazon Aurora
  - five times the throughput of standard MySQL
  - data replicated to 6 replicas by default (up to 15)
  - supports databases up to 64 GB in size
  - continuous backup to S3

MANAGEMENT SERVICES

Amazon CloudWatch

Amazon
CloudWatch

- Central point of monitoring of logs, events and metrics
- Integrated with all other services
- Set alarms to automatically react to changes

Amazon CoudTrail

Amazon
CloudTrail

- Records every API call
- Essential for compliance auditing
- Streams logs to CloudWatch

Amazon CloudFormation

AWS
CloudFormation

- *Infrastructure-as-a-code* service
- Spin up entire stacks od resources with a single click
- Plays key role in deployment automation
- Stacks can be defined in JSON or YAML syntax

APPLICATION SERVICES

Amazon SQS

Amazon
SQS

- Fast, scalable message queuing
- Supports any volume of messages - scales automatically
- Cost-effective way to decouple cloud applications
- Two types of queues:
  - Standard
  - FIFO

Amazon SNS

Amazon
SNS

- Fast reliable push notification service
- Simple pub-sub model
- Can deliver to following subscribers:
  - Mobile notifications
  - SMS messages
  - SQS queues
  - HTTP/HTTPS endpoints
  - Lambda functions

ANALYTICS SERVICES

Amazon Kinesis

Amazon
Kinesis

- Reliable data streaming
- Supports TBs of data per hour from thousands of sources
- Emit stream data between AWS services
- Apply SQL queries directly on the stream data for analytics
- Deliver stream data directly to S3 or Redshift (warehouse)

DEVELOPER SERVICES

Developer Services


AWS CodeCommit


AWS CodeBuild


AWS CodeDeploy


AWS CodePipeline

- Services for continuous integration and deployment
- Integrated with the AWS ecosystem (e.g. OpsWorks)
- Works with external services (GitHub, Jenkins, Docker, etc.)

And about...

50

more services😮

No one will be running their datacenter
in few years, it's your last chance
to adapt.

THANK YOU!

Q & A