



Introduction to AWS

Εθνικό Μετσόβιο Πολυτεχνείο
28 Μαΐου 2025

Νικηφόρος Μποτής (nbotis@amazon.com)

Sr. Solutions Architect, Δημόσιος Τομέας Ελλάδος/Κύπρου/Μάλτας





amazon go

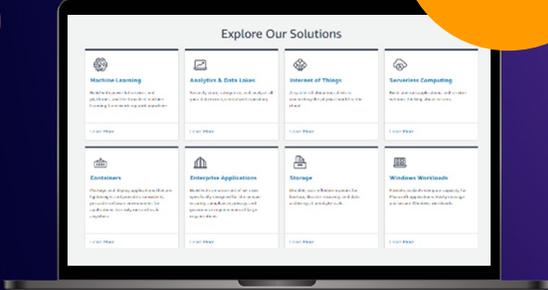
No Lines.
No Checkout.
(No, Seriously.)

- 1994 Founded
- 1995 Amazon.com
- 1998 Added CDs & DVDs
- 2006 Amazon Web Services
- 2007 Kindle/Prime
- 2011 Video
- 2012 Groceries
- 2014 Alexa/Echo
- 2015 Bookstores
- 2017 Amazon Go
- 2019 Prime Air – Drone Delivery
- 2023 Kuiper – LOE Satellite

Innovations from Amazon



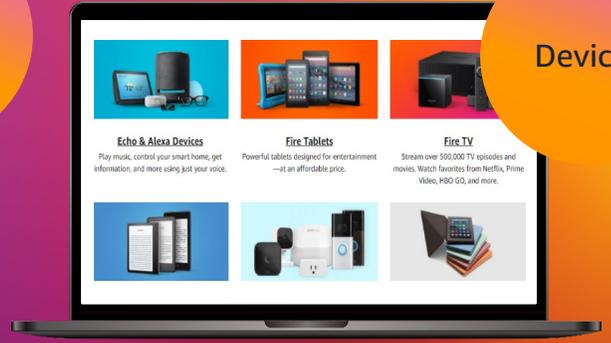
E-commerce



Robotics



Devices



Streaming content



Just Walk Out



Scout



Amazon One



Kuiper



Consumables



Amazon's Secret Recipe of Innovation

$$f(\text{innovation}) = (\text{org} * \text{arch}) \quad (\text{mechanisms} * \text{culture})$$

Amazon's Secret Recipe of Innovation



Mechanisms

- Press Release
- FAQ
- Visuals



Culture

- Leadership Principles
- Hire builders
- Accept failure

(mechanisms * culture)

$$f(\text{innovation}) = (\text{org} * \text{arch})$$



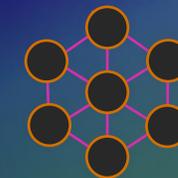
Innovation

- Customer obsession
- Invent on behalf of customers
- Long-term thinking



Organization

- Two pizzas
- Full ownership
- Full accountability



Architecture

- Microservices
- Self-service tools
- Automation

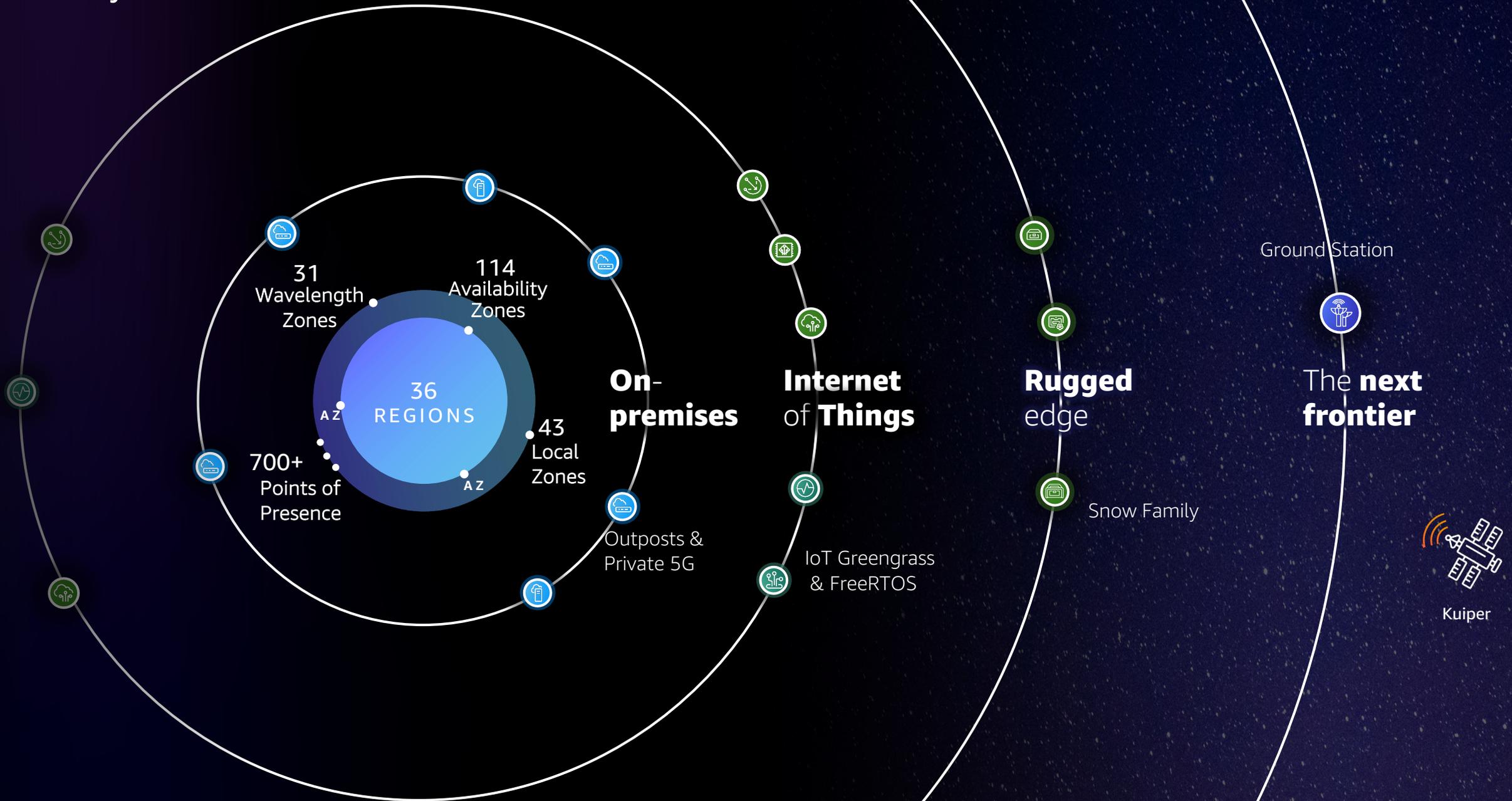
The most secure, extensive, and reliable Global Cloud Infrastructure

Infrastructure allows customers
to run workloads

Customers have the same access and
capabilities no matter where you are

240+ fully featured services from
data centers globally

The Everywhere Cloud



AWS Global Network

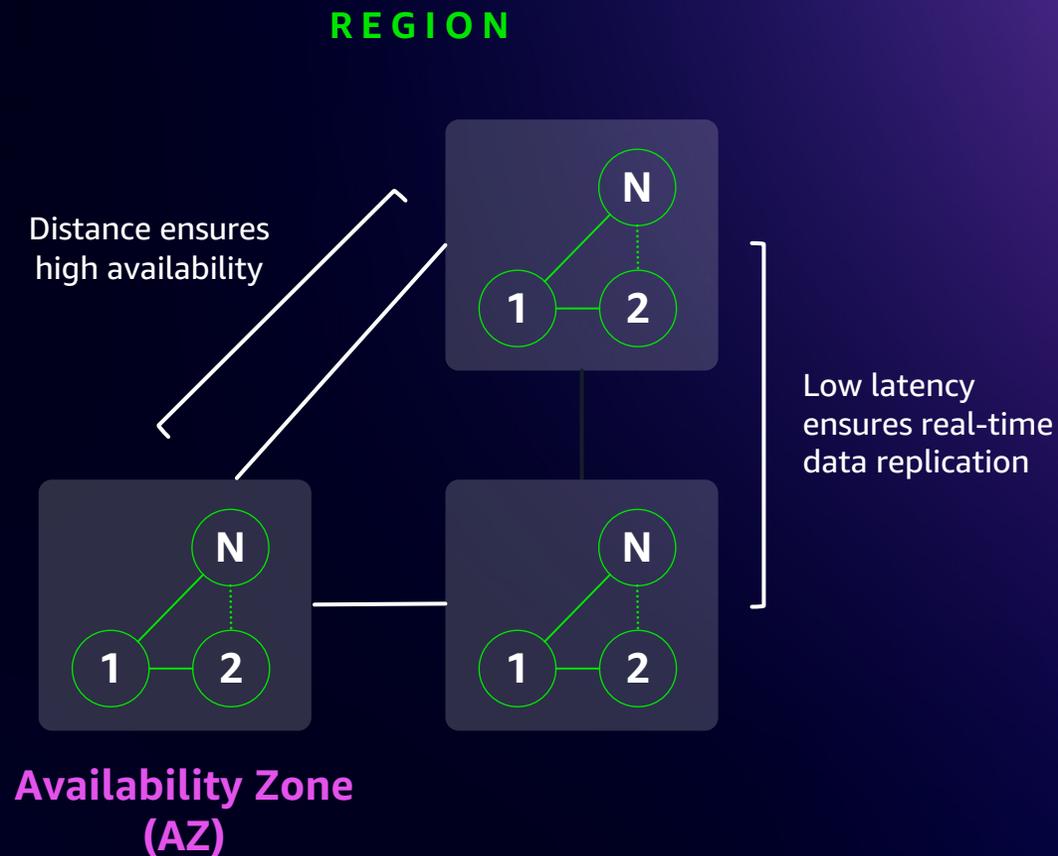
400 Gbps

Network backbone

5,000,000 km

Terrestrial fiber optic cable

Availability Zones design for resiliency



100K+ servers at scale

ISOLATED PARTITION

Fully isolated with one or more data centers

POWER

Highly available, fault tolerant, and scalable

DISTANCE

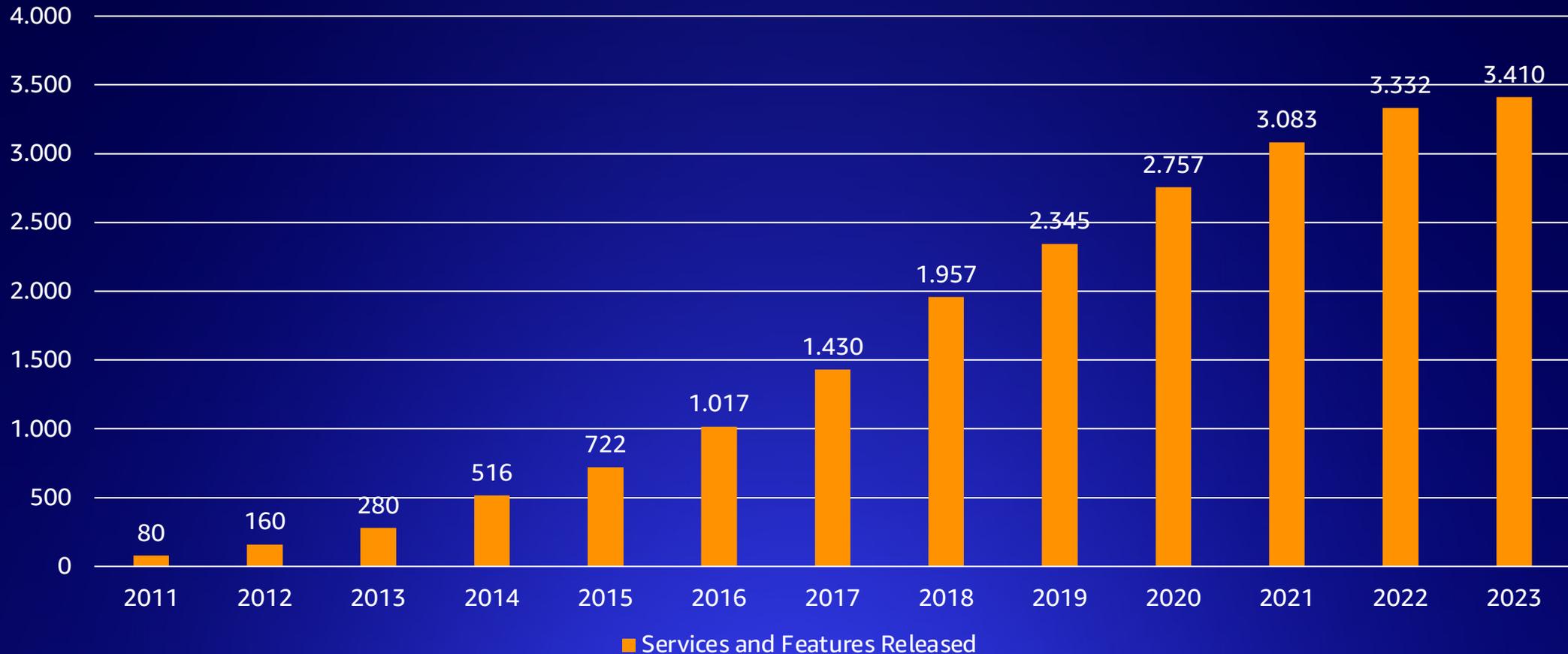
Physically separated by a meaningful distance – all within 60 miles (100km) of each other

INTERCONNECTION

Data centers connected via fully redundant and isolated metro fiber

AWS Innovation Paces - 90% of features are requested from customers

new services / features release



#3410 in 2023, 65.57 new features launched per week

240+

SERVICES

Broad and Deep Functionality



AWS Can Help – Breadth and Depth of Services

ANALYTICS

ANALYTICS	STREAMING
DATA EXCHANGE	ETL
DATA LAKE	HADOOP/SPARK
DATA PIPELINES	INTERACTIVE SQL QUERIES
DATA WAREHOUSE	VISUALIZATIONS
ELASTICSEARCH	

AR + VR

AR/VR EXPERIENCES

AWS COST MANAGEMENT

ANALYZE AWS COSTS
COST & USAGE BUDGETS
COST & USAGE REPORTS
RESERVED INSTANCES REPORTING

APPLICATION INTEGRATION

EMAIL	SEARCH
MESSAGE BROKER	TRANSCODING
QUEUEING & NOTIFICATIONS	WORKFLOW

BUSINESS APPLICATIONS

EMAIL & CALENDARING	UNIFIED COMMUNICATIONS
ONLINE MEETINGS	VOICE-ENABLED WORKPLACE
SHARING & COLLABORATION	

BLOCKCHAIN

BLOCKCHAIN TEMPLATES
LEDGER DATABASE
MANAGED BLOCKCHAIN

CUSTOMER ENGAGEMENT

CONTACT CENTER
TARGETING
USER ENGAGEMENT ACROSS CHANNELS

COMPUTE

COMPUTE	CONTAINERS
AUTO SCALING	CONTAINER SERVICE
BATCH JOBS	MANAGED KUBERNETES
EVENT-DRIVEN SERVERLESS COMPUTING	STORE & RETRIEVE DOCKER IMAGES
INSTANCE TYPES	
MANAGED VIRTUAL PRIVATE SERVERS	
MANAGED REPOSITORY FOR SERVERLESS APPS	
RUN & MANAGE WEB APPS	
SERVERLESS COMPUTE	
VIRTUAL SERVERS	

DATABASE

REALTIONAL DATABASES	PURPOSE-BUILT DATABASES
HIGH-PERFORMANCE RELATIONAL DATABASE BUILT FOR THE CLOUD	DOCUMENT DATABASE
MANAGED MARIADB	GRAPH DATABASE
MANAGED MYSQL	IN-MEMORY CACHING
MANAGED ORACLE	KEY-VALUE STORE DATABASE
MANAGED POSTGRESQL	LEDGER DATABASE
MANAGED SQL SERVER	TIME SERIES DATABASE

DEVELOPER TOOLS

ANALYZE & DEBUG
APPLICATION LIFECYCLE MANAGEMENT
AUTHORING
BUILD & TEST
CONTAINERS
DEVOPS RESOURCE MANAGEMENT
ONE-CLICK APP DEVELOPMENT
PATCHING
PIPELINE ORCHESTRATION
RESOURCE TEMPLATES
TRIGGERS

END USER COMPUTING

APP STREAMING	MOBILE ACCESS
DESKTOP COMPUTING	STORAGE & COLLABORATION

HYBRID ARCHITECTURE

AWS SERVICES ON PREMISES	INTEGRATED NETWORKING
DATA INTEGRATION	INTEGRATED RESOURCE & DEPLOYMENT MANAGEMENT
INTEGRATED DEVICES & EDGE SYSTEMS	VMWARE CLOUD ON AWS
INTEGRATED IDENTITY & ACCESS	

GAME TECH

CROSS-PLATFORM 3D GAME ENGINE
GAME SERVER HOSTING

INFRASTRUCTURE

AVAILABILITY ZONES
CUSTOM HARDWARE
DATA CENTER INFRASTRUCTURE
GLOBAL NETWORK BACKBONE
POINTS OF PRESENCE
POWER INFRASTRUCTURE
REGIONS

INTERNET OF THINGS (IOT)

RULES ENGINE
DEVICE ANALYTICS
DEVICE GATEWAY
DEVICE SDK
DEVICE SHADOWS
EVENT DETECTION & RESPONSE
LOCAL COMPUTE
LOCAL DATA COLLECTION
MANAGEMENT & SECURITY
MICROCONTROLLER OPERATING SYSTEM
REGISTRY
VISUAL APPLICATIONS DEVELOPMENT

MACHINE LEARNING

ML FRAMEWORKS	SAGEMAKER
DEEP LEARNING AMIS & CONTAINERS	AUTOMATIC MODEL TUNING
HARDWARE ACCELERATION	DATA LABELING
ML AT THE EDGE	HOSTED NOTEBOOKS
TENSORFLOW, PYTORCH, MXNET	ML MARKETPLACE
	MODEL HOSTING
	MODEL OPTIMIZATION
	MODEL TRAINING
AI SERVICES	PRE-BUILT ALGORITHMS
CHATBOTS	TOPIC MODELING
ENTITY EXTRACTION	DEEP LEARNING MODELS
FACE ANALYTICS	REINFORCEMENT LEARNING
FACE SEARCH	SPOT INSTANCES
FORECASTING	BATCH PREDICTIONS
IMAGE LABELING	REAL-TIME PREDICTIONS
NATURAL LANGUAGE PROCESSING	
PERSONALIZATION & RECOMMENDATION	
SENTIMENT ANALYSIS	
SPEECH TRANSCRIPTION	
TEXT & DATA EXTRACTION	
TEXT TO SPEECH	
TRANSLATION	
VIDEO & IMAGE ANALYSIS	
CONTENT MODERATION	

MANAGEMENT & GOVERNANCE

ACTIVITY & API USAGE TRACKING	MONITORING
CHATBOT	PROVISIONING
CONFIGURATION TRACKING	RESOURCE TEMPLATES
GOVERNANCE	SECURITY RECOMMENDATIONS
INVENTORY TRACKING	SERVER MANAGEMENT
LICENSE MANAGER	SERVICE CATALOG
MANAGE POLICIES	SYSTEMS MANAGER
MANAGE RESOURCES	

MARKETPLACE

ANALYTICS	MACHINE LEARNING
DATA PRODUCTS	NETWORKING
DATABASES	OPERATING SYSTEMS
DEVOPS	SECURITY
IOT	STORAGE

MEDIA SERVICES

LIVE VIDEO TRANSPORT	VIDEO PERSONALIZATION & MONETIZATION
MEDIA STORAGE	VIDEO PROCESSING & DELIVERY
TRANSCODING	VIDEO STREAMING ANALYSIS
VIDEO ORIENTATION & PACKAGING	

MIGRATION & TRANSFER

APPLICATION MIGRATION
DATABASE MIGRATION
EXABYTE-SCALE MIGRATION
ONLINE DATA TRANSFER
SCHEMA CONVERSION
SERVER MIGRATION
TRANSFER FOR SFTP

MOBILE

API GATEWAY	MOBILE APP TESTING
DEVELOPMENT FRAMEWORK	SINGLE INTEGRATED CONSOLE
IDENTITY	SYNC
MOBILE ANALYTICS	TARGETED PUSH NOTIFICATIONS

NETWORKING & CONTENT DELIVERY

APPLICATION DELIVERY
DEDICATED NETWORK CONNECTION
DOMAIN NAME SYSTEM
LOAD BALANCING
MONITOR APIS
MONITOR MICROSERVICES
NETWORK TOPOLOGY
NETWORKING HUB
PRIVATE CONNECTION TO APPS
SCALE VPC & ACCOUNT CONNECTIONS
SERVICE DISCOVERY
VIRTUAL PRIVATE CLOUD

ROBOTICS

SATELLITE

SATELLITE OPERATIONS

SECURITY, IDENTITY, & COMPLIANCE

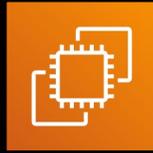
ACCESS CONTROL
ASSESSMENT & REPORTING
CONFIGURATION COMPLIANCE
DATA PROTECTION
DDOS PROTECTION
IDENTITY MANAGEMENT
KEY MANAGEMENT & STORAGE
MONITORING & LOGGING
RESOURCE MANAGEMENT
THREAT DETECTION
WEB APPLICATION FIREWALL

STORAGE

ARCHIVE STORAGE
BACKUP & RESTORE
BLOCK STORAGE
DATA TRANSFER
EDGE PROCESSING & COMPUTING
FILE STORAGE
HIGH-PERFORMANCE FILE SYSTEM
HYBRID CLOUD STORAGE
OBJECT STORAGE
WINDOWS FILE SYSTEM

CUSTOMER ENABLEMENT

ACCOUNT MANAGEMENT
DASHBOARD PERSONALIZATION
ENTERPRISE SUPPORT
EXPERTS MARKETPLACE
OPTIMIZATION GUIDANCE
PARTNER ECOSYSTEMS
PROFESSIONAL SERVICES
SECURITY & BILLING REPORTS
SOLUTIONS MANAGEMENT
TRAINING & CERTIFICATION



Amazon EC2



Amazon RDS



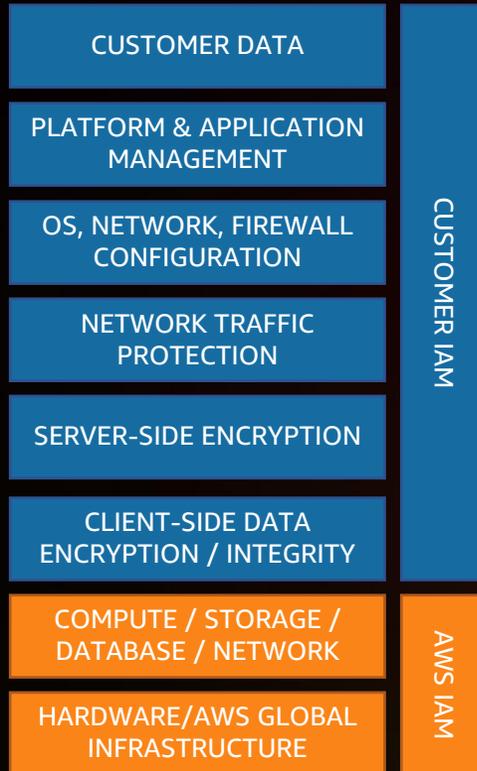
Amazon S3



AWS KMS



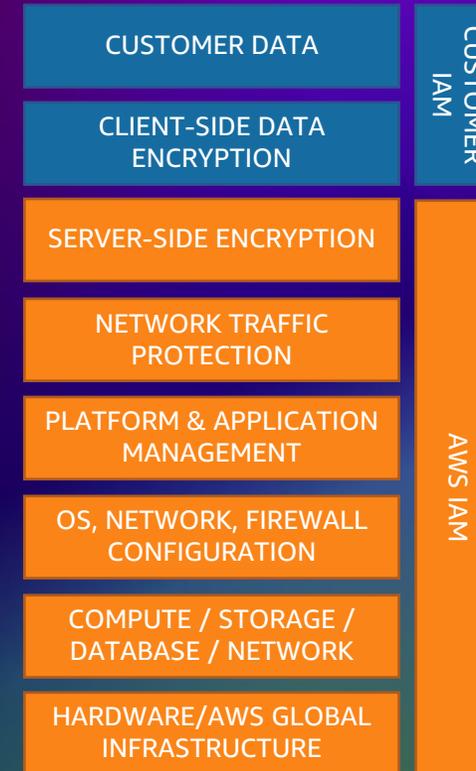
Amazon DynamoDB



INFRASTRUCTURE SERVICES



CONTAINER SERVICES



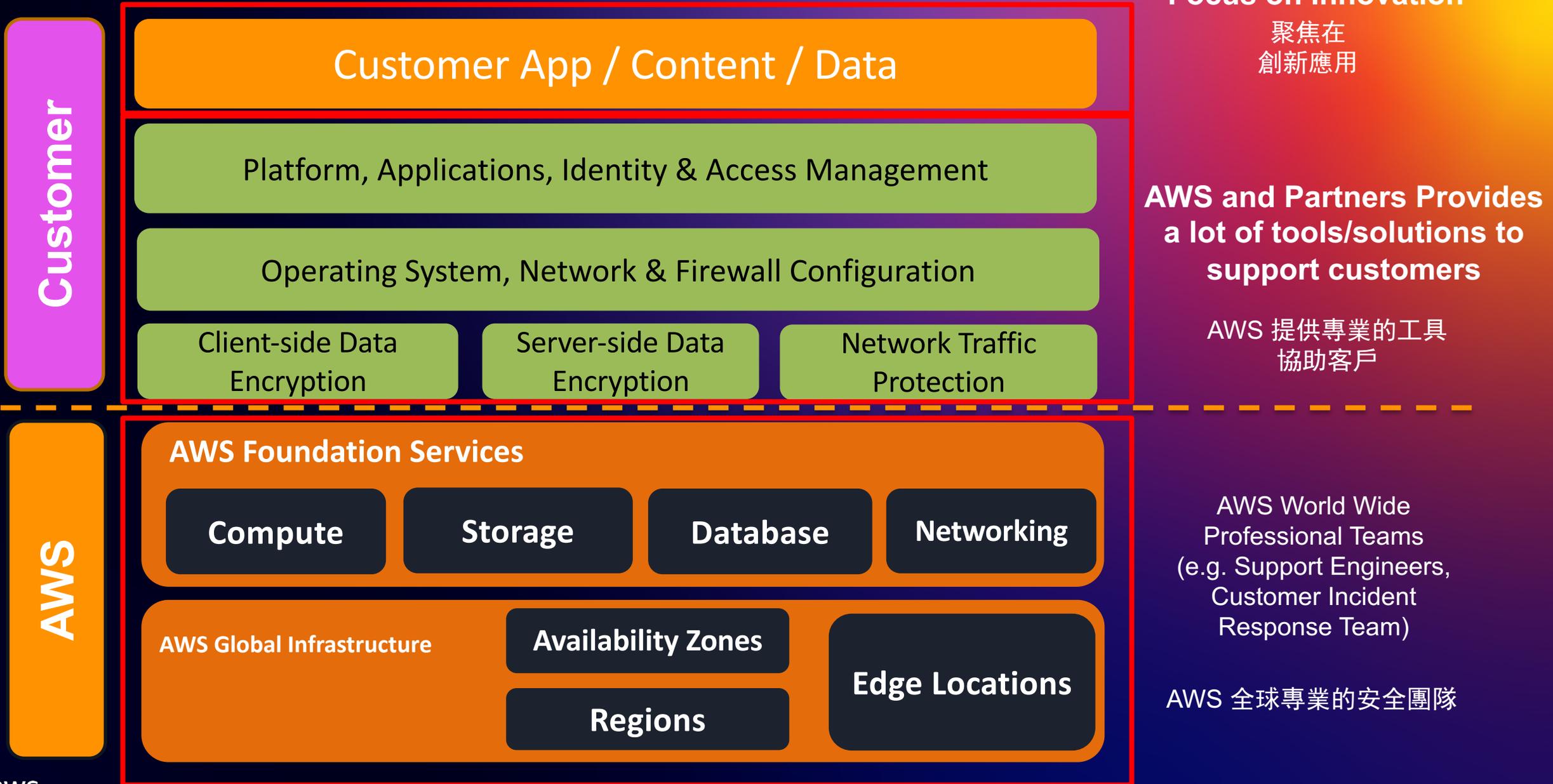
ABSTRACTED SERVICES

Less Customizable
+
Less Customer Responsibility
+
More Best Practices built-in

More Customizable
+
More Customer Responsibility



Shared Responsibility Model



Focus on Innovation

聚焦在
創新應用

**AWS and Partners Provides
a lot of tools/solutions to
support customers**

AWS 提供專業的工具
協助客戶

AWS World Wide
Professional Teams
(e.g. Support Engineers,
Customer Incident
Response Team)

AWS 全球專業的安全團隊

Shared Responsibility Model

AWS Well-Architected

Learn, measure, and build using architectural best practices

AWS Architecture Center This is My Architecture AWS Answers AWS Solutions Case Studies Cloud Security

Security **IN** the
Cloud

Managed by
customers

AWS Cloud Compliance

Assurance programs for finance, healthcare, government and more.

I'd like information about Compliance in the Cloud »

Compliance Cloud Security Assurance Programs Resources Latest News Testimonials

AWS Artifact

No cost, self-service portal for on-demand access to AWS' compliance reports.

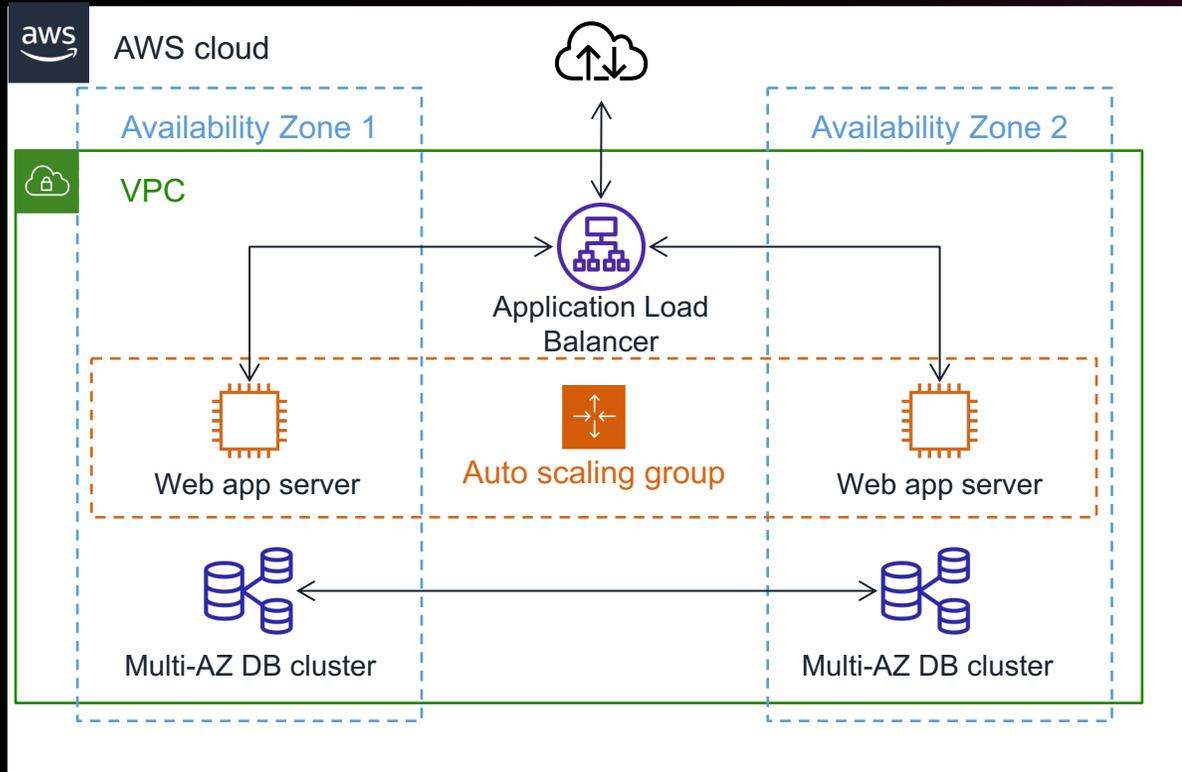
Start for Free with AWS Artifact

Artifact Getting Started FAQ Documentation Compliance Security

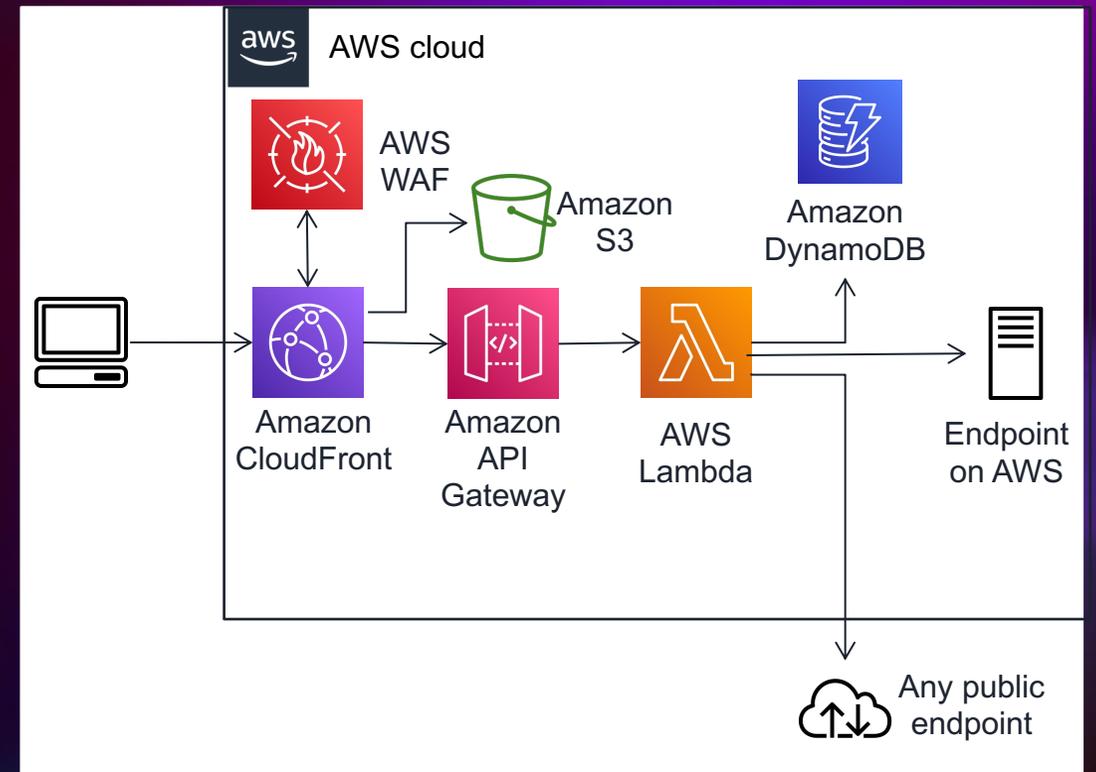
Security **OF** the
Cloud

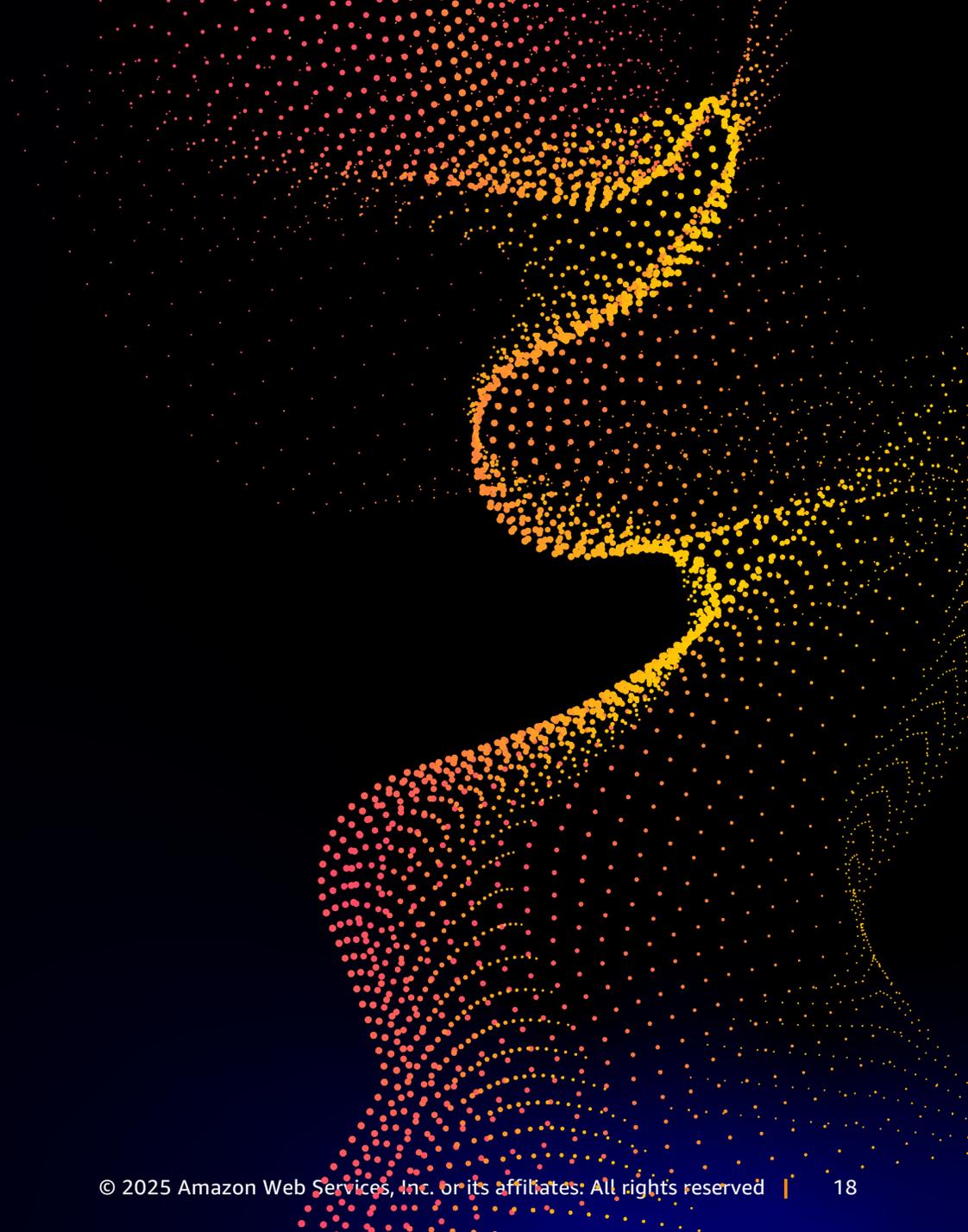
Managed by
AWS

Three tier web application



Serverless web application

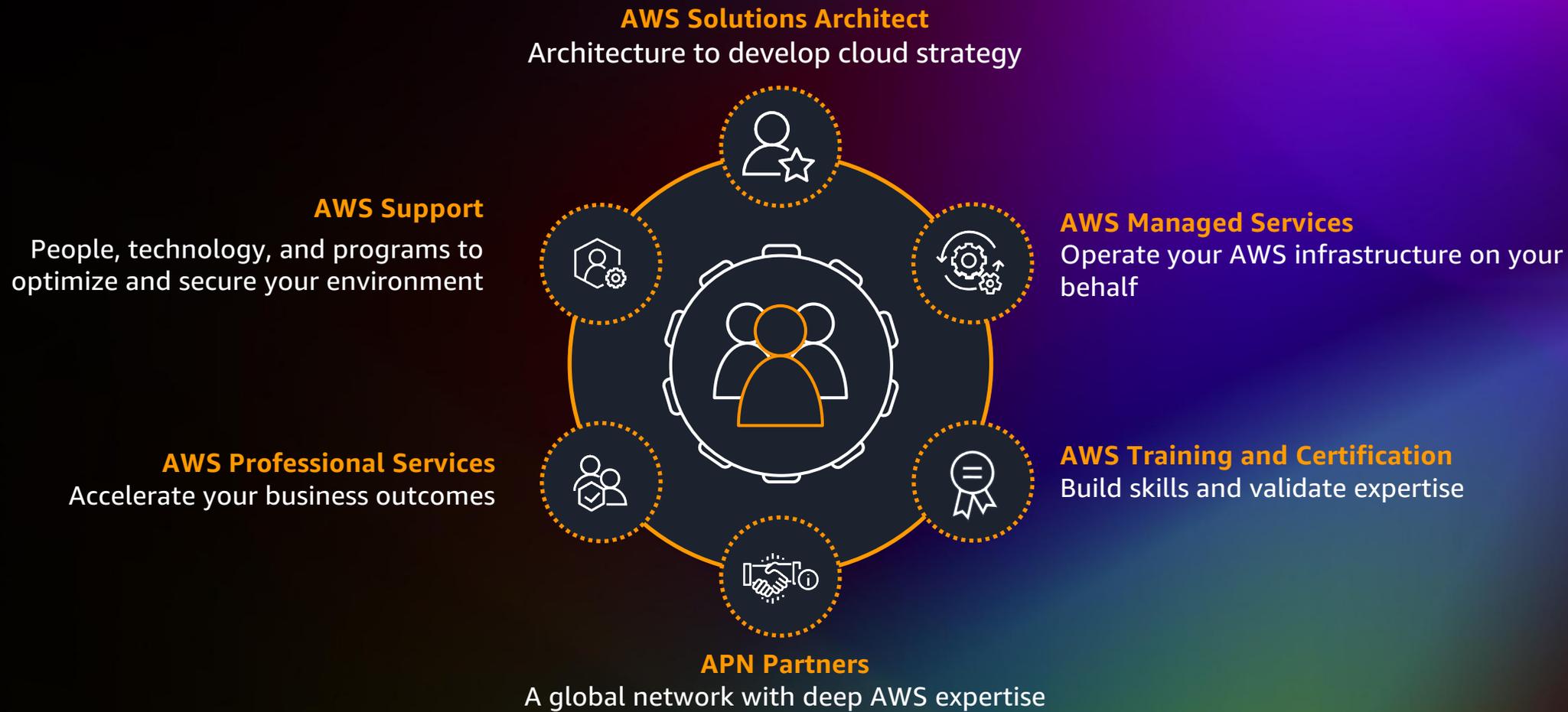




Top 1 Question from Customers:
How Can We Start?

Support from Local Teams

Migrate and build faster in the cloud



Learn by Role

Build cloud skills to advance your career



Solutions Architect



Cloud Practitioner



Developer



DevOps Engineer



Machine Learning



Operations

Learn by Solution

Build cloud skills to solve real-world problems



Advanced Networking



Data Analytics



Databases



AWS for Games



Machine Learning



Media Services



Security



Serverless



Storage

Ευχαριστώ!



Νικηφόρος Μποτής (nbotis@amazon.com)

Sr. Solutions Architect, Δημόσιος Τομέας Ελλάδος/Κύπρου/Μάλτας

AI chat applications at work



Amazon Q

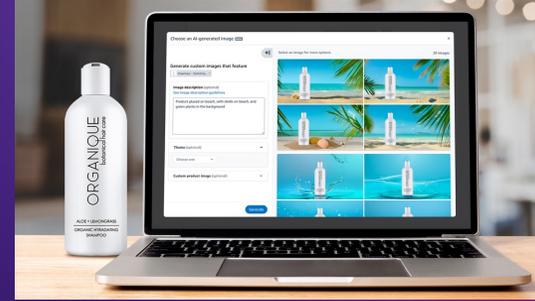
8 ways Amazon is using generative AI to make life easier



A more conversational Alexa with generative AI



Generative AI improves product listings



Generative AI creates more engaging advertisements



Customers can pay with their palm



Enjoy review highlights with generative AI



Generative AI helps predict the blitz on 'Thursday Night Football'



Amazon Go - eliminate checkout lines



Generative AI makes prescriptions easier to read

Source: [8 ways Amazon is using generative AI to make life easier, from a more conversational Alexa to a better reviews experience](#)



Swami Sivasubramanian  · Following

VP, AI and Data

1mo · Edited · 



With the general availability of Amazon Q, I wanted to give you the backstory on how we brought Q Apps to customers (hint: it started with our employees) and highlight some of the new capabilities that our customers are excited about.

Internally, we have something called LLM playground, where employees can experiment with different models. The playground has been wildly popular among Amazon users both in technical and non-technical roles. One of the things we noticed was non-developers started 1/building apps for functionality, 2/using those apps, and 3/sharing them with each other. Employees took this even further and maintained internal “how-to” guides for these apps on our internal Amazon portals/wikis. For instance, data center engineering and operations teams built an app—using natural language—to read equipment manuals and help debug when they are repairing equipment. What they wanted from us was to elevate apps from being a first class concept to having the ability to build on top of their business data (with built-in security and access controls) so that everyone could build an app right from Q Business and then publish their apps to broader App catalog. Interestingly, this was a pattern we started seeing from many of our external customers as well who wanted to use GenAI to push beyond getting insights. This led to Q Apps (available in preview).

Amazon Q Apps to customers (hint: it started with our employees)

Internally, we have something called LLM playground, where employees can experiment with different models.

The playground has been wildly popular among Amazon users both in technical and non-technical roles. One of the things we noticed was non-developers started 1/building apps for functionality, 2/using those apps, and 3/sharing them with each other.

For instance, data center engineering and operations teams built an app—using natural language—to read equipment manuals and help debug when they are repairing equipment.



Amazon Q Apps to customers (hint: **it started with our employees**)

Internally, we have something called **LLM playground**, where employees can experiment with different models.

The playground has been wildly popular among Amazon users both in **technical and non-technical roles**. One of the things we noticed was non-developers started **1/building apps for functionality, 2/using those apps, and 3/sharing them with each other**.

For instance, **data center engineering and operations** teams built an app—using natural language—to read equipment manuals and help debug when they are repairing equipment.



Bottom Up



Playground & Experiment



Break silos between tech and non-tech.
Building, using, and sharing



Create values & new skills

Top 5 Question from Customers:

What's the benefits of Cloud?

Benefits of Cloud

**Operational
Resilience**

**Data Residency
and Compliance**

**Breadth and Depth
of Services**

**Low Network
Latency**

Samsung migrates **1.1 billion users** across three continents from Oracle to Amazon Aurora with AWS Database Migration Service

SAMSUNG

60 ms latency

or less for 90% of the time

44%

reduced monthly database costs

“The scalability of Amazon Aurora is the best benefit – especially if we focus on the cost.”

Salva Jung

Principal Architect and Engineering Manager

Customers realized an **88% ROI**

79% Improved speed-to-market

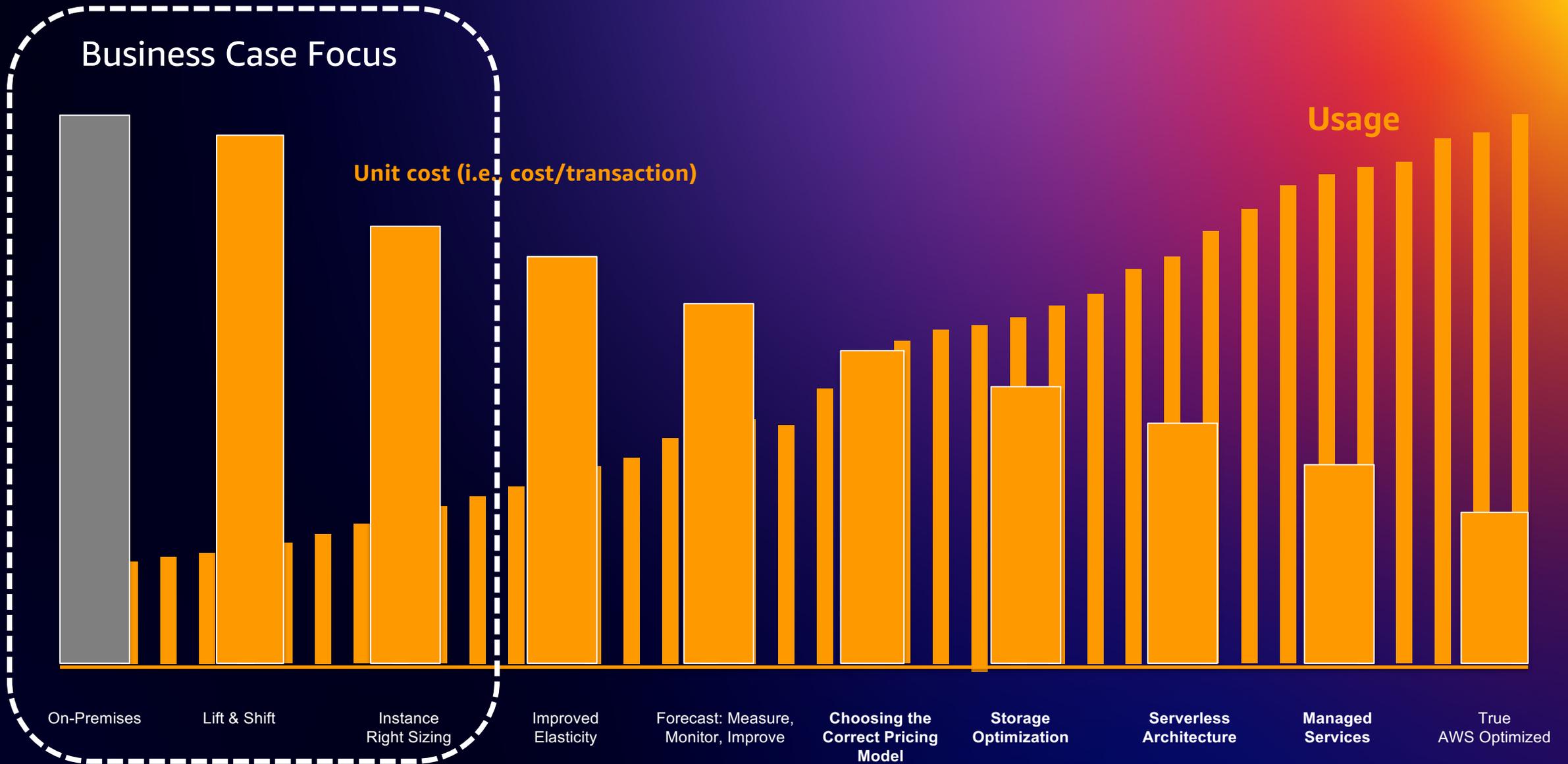
10% Reduced cost of legacy application maintenance

6% Reduced cost of new application maintenance

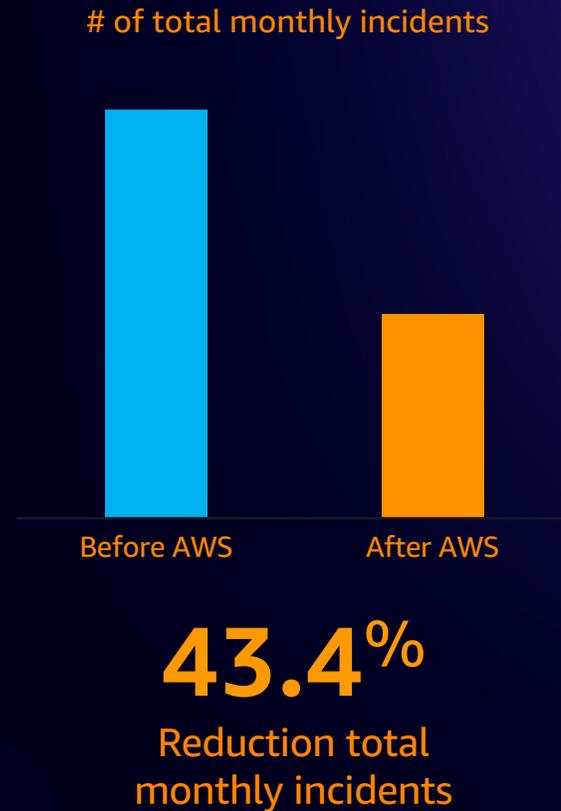
5% Accelerated customer onboarding

FORRESTER®

Cloud Economics



AWS customers reduce monthly incidents



- High infrastructure availability and durability; multi-Availability Zone fail over
- Automatic configuration and resource management; centrally manage users and credentials
- 36.1% drop in security incidents.

Value of the Cloud



Cost savings (TCO)

What is it?

Infrastructure cost savings/ avoidance from moving to the cloud

20% reduction in IT infrastructure costs as a percentage of revenue



Staff productivity

What is it?

Efficiency improvement by function on a task-by-task basis

66% increase in VMs managed per server admin



Operational resilience

What is it?

Benefit of improving SLAs and reducing unplanned outage

69% reduction in unplanned downtime



Business agility

What is it?

Deploying new features/ applications faster and reducing errors

43% reduction in time-to-market for new application features



Sustainability

What is it?

Minimizing environmental impact of operations

88% reduction in workload carbon footprint

Tactical impact

Strategic impact

Safeguarding Ukraine's data to preserve its present and build its future

Challenge

In February 2022, the Russian government invaded Ukrainian territory and the Ukrainian government had an urgent need to safeguard critical data essential for government operation and its citizens' wellbeing by migrating data to cloud

Solution

The AWS Disaster Response and AWS Public Sector teams collaborated with the Ukrainian government to create Project Sunflower, which used AWS Snowball Edge devices to migrate critical government and commercial data to the cloud; the Snow team collaborated with the Project Sunflower team to make Snowball devices readily available for use by the WWPS SA team for the data migration

Result

- Snowball Edge successfully migrated PBs of critical government data into the AWS Cloud from Ukraine locations
- Enabled Ukrainian government and public to have access to critical data despite being in a war

"We used to assume that this is just how it is in war – everything gets destroyed and you have to rebuild from nothing. But by migrating to the safety and security of the cloud, the government and its citizen services prevail."

Liam Maxwell, Director of Government Transformation, WWPS



AWS Snowball Edge Compute

Availability Zones design for resiliency



Running in two zones you need to reserve enough spare capacity to rapidly switch the load for resilience

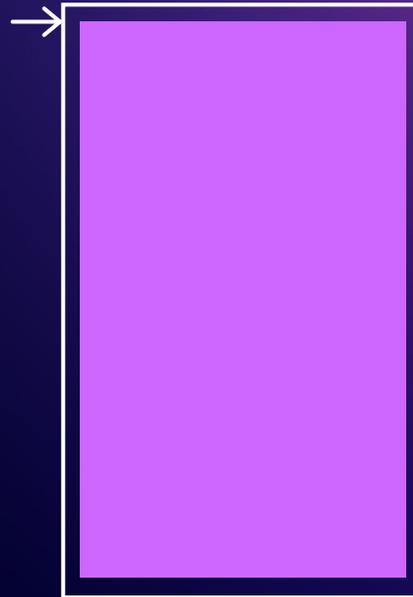
Availability Zones design for resiliency

AVAILABILITY ZONE
A

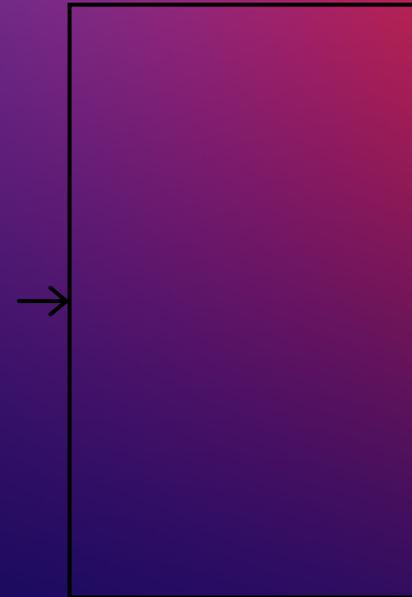


Reserved capacity is
sufficient to run
workload in one zone

AVAILABILITY ZONE
B



AVAILABILITY ZONE
C



Workload
requirement:
Immediate
failover

Running in two zones you need to reserve enough spare capacity
to rapidly switch the load for resilience

Availability Zones design for resiliency



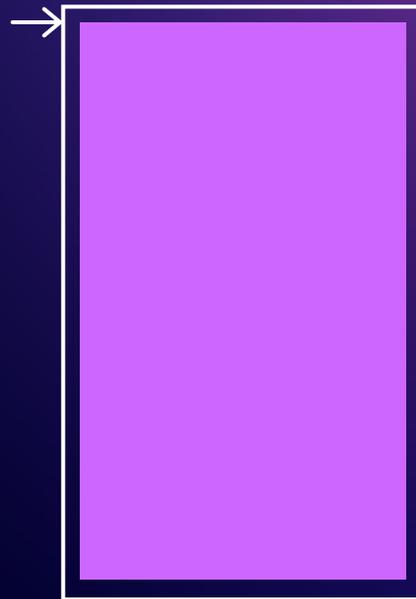
Running in three zones you need less reserved capacity to switch the load

Availability Zones design for resiliency

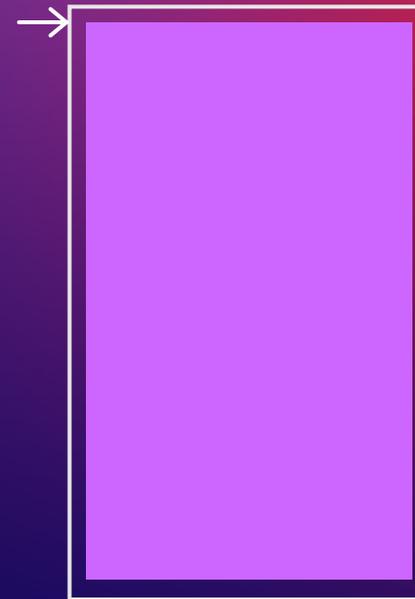
AVAILABILITY ZONE
A



AVAILABILITY ZONE
B



AVAILABILITY ZONE
C



Reserved capacity is
sufficient to run
workload in two zones

Workload
requirement:
Immediate
failover

Split the extra load over two zones;
higher average utilization, lower cost and lower impact

There are many different business drivers



Moderna - Hybrid Reference Cases



"AWS Storage Gateway lets us easily integrate data files from analytical instruments . . . and the transparent S3 storage lets us easily connect our cloud-based applications and leverage the powerful storage capabilities of S3. **With the AWS File Gateway, we can now unleash the full power of AWS on our instrument data.**"

Dave Johnson, PhD
Director of Informatics

Problem

Made a strategic decision to leverage AWS for processing and analyzing scientific data

Accelerate moving scientific data that is generated on premises to the cloud

Solution

Deploy multiple Storage Gateway appliances to copy scientific data to Amazon S3

Process and analyze 100s of TBs of data in Amazon S3

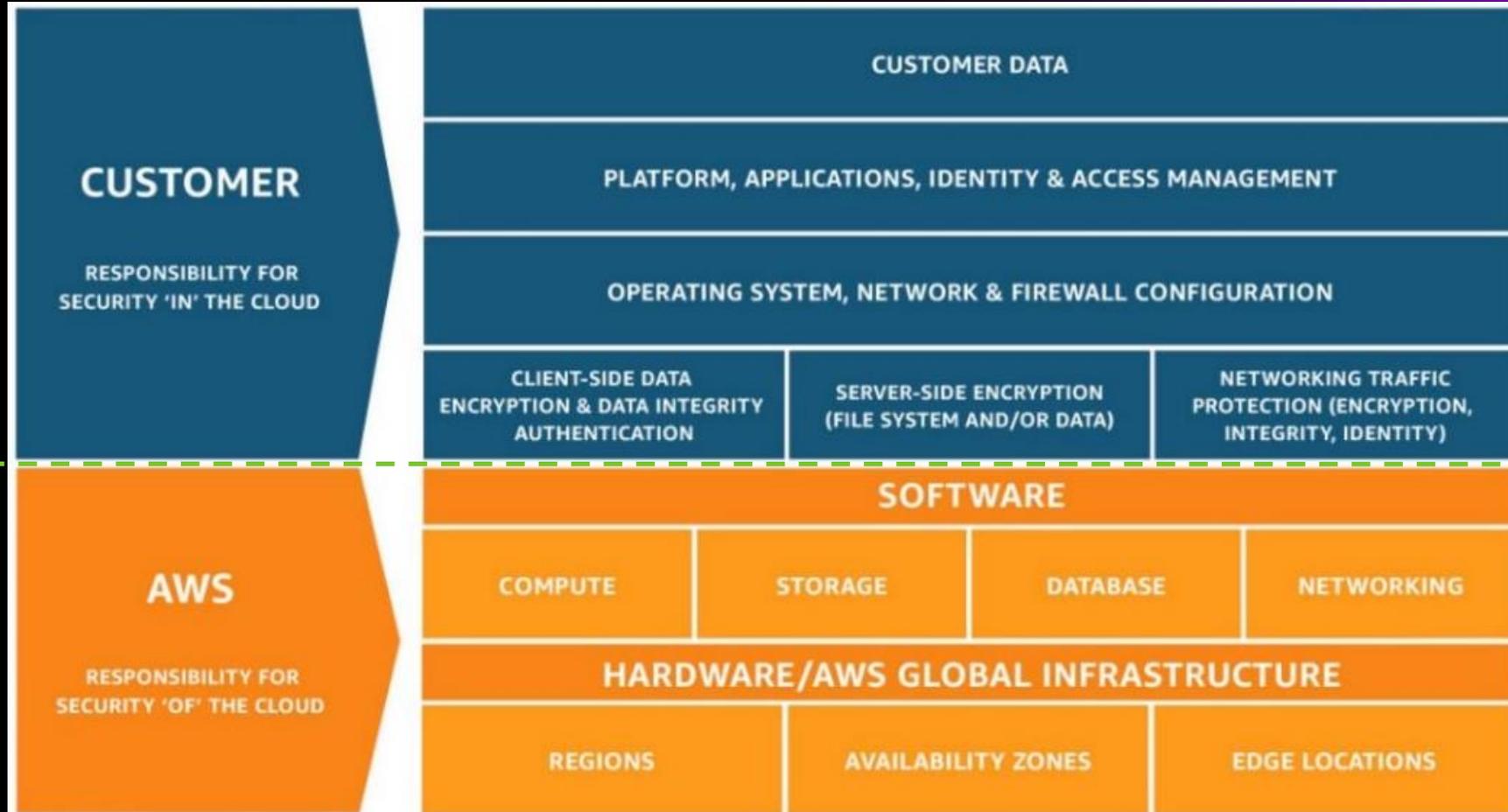
Outcome

On-premises scientific instruments get low-latency access to cloud storage

Lower cost and simplicity

Automation of data management

AWS Shared Responsibility Model



Security **IN** the Cloud
Managed by **customers**

Security **OF** the Cloud
Managed by **AWS**

Supporting resources

AWS Well-Architected

Learn, measure, and build using architectural best practices

AWS Architecture Center This is My Architecture AWS Answers AWS Solutions Case Studies Cloud Security

Security **IN** the
Cloud

Managed by
customers

AWS Cloud Compliance

Assurance programs for finance, healthcare, government and more.

I'd like information about Compliance in the Cloud »

Compliance Cloud Security Assurance Programs Resources Latest News Testimonials

AWS Artifact

No cost, self-service portal for on-demand access to AWS' compliance reports.

Start for Free with AWS Artifact

Artifact Getting Started FAQ Documentation Compliance Security

Security **OF** the
Cloud

Managed by
AWS

PROBLEM

Sought to revolutionize the drug discovery process and bring drugs to market faster to help patients as soon as possible

SOLUTION

Using digital applications and machine learning built on AWS, Moderna developed a new mRNA COVID-19 vaccine with unprecedented speed. AI algorithms enable drug development from rapid experimentation cycles to help design mRNA and DNA sequences to the automation of quality control analyses, saving countless hours of manual review to improvements across production processes and logistics.

IMPACT

Designed and shipped the first clinical batch for its coronavirus vaccine in just 42 days.

Allowed company to rapidly expand into other infectious disease targets and modalities like oncology and autoimmune disease – using the flexible, programmable nature of mRNA, powered by machine learning,

The Moderna logo is displayed in white lowercase letters on a dark blue background. The word "moderna" is written in a clean, sans-serif font. Below the text is a horizontal dashed line consisting of several short segments.

“ We can be far more secure in the cloud and achieve a **higher level of assurance at a much lower cost**, in terms of effort and dollars invested. We determined that **security in AWS is superior to our on-premises data center across several dimensions**, including patching, encryption, auditing and logging, entitlements, and compliance.”



—John Brady
CISO, FINRA

Nasdaq wanted to migrate their mission-critical equities and options markets to the cloud to prepare for the future. It requires the underlying infrastructure to be hyper-scalable, ultra-resilient, and highly performant

Deploying North American markets on a co-designed edge compute solution using AWS Outposts

End-to-end order-to-trade process occurs in low double-digit microseconds with a 10% performance improvement in network latency, creating a cloud migration blueprint for 130+ markets infrastructure customers

Moderna - Hybrid Reference Cases



"AWS Storage Gateway lets us easily integrate data files from analytical instruments . . . and the transparent S3 storage lets us easily connect our cloud-based applications and leverage the powerful storage capabilities of S3. **With the AWS File Gateway, we can now unleash the full power of AWS on our instrument data.**"

Dave Johnson, PhD
Director of Informatics

Problem

Made a strategic decision to leverage AWS for processing and analyzing scientific data

Accelerate moving scientific data that is generated on premises to the cloud

Solution

Deploy multiple Storage Gateway appliances to copy scientific data to Amazon S3

Process and analyze 100s of TBs of data in Amazon S3

Outcome

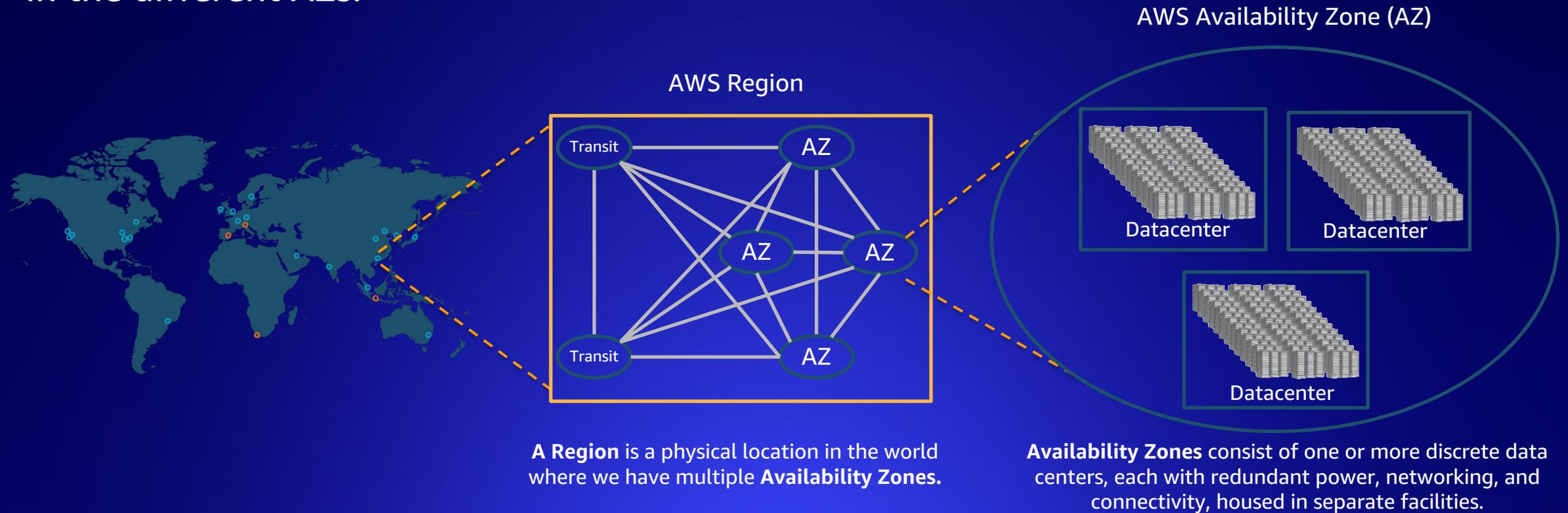
On-premises scientific instruments get low-latency access to cloud storage

Lower cost and simplicity

Automation of data management

AWS region design

AWS Regions are comprised of multiple AZs for **high availability**, **high scalability**, and **high fault tolerance**. Applications and data are replicated in real time and consistent in the different AZs.



A Region is a physical location in the world where we have multiple **Availability Zones**.

Availability Zones consist of one or more discrete data centers, each with redundant power, networking, and connectivity, housed in separate facilities.