

# **AWS Solution Architect Associate Certification Training – Module 5**

## **5. Introduction to Amazon Web Services**

### **History and Evolution of AWS**

Amazon has a long history of using a decentralized IT infrastructure. This arrangement enabled our development teams to access compute and storage resources on demand, and it has increased overall productivity and agility. By 2005, Amazon had spent over a decade and millions of dollars building and managing the large-scale, reliable, and efficient IT infrastructure that powered one of the world's largest online retail platforms. Amazon launched Amazon Web Services (AWS) so that other organizations could benefit from Amazon's experience and investment in running a large-scale distributed, transactional IT infrastructure. AWS has been operating since 2006, and today serves hundreds of thousands of customers worldwide. Today Amazon.com runs a global web platform serving millions of customers and managing billions of dollars' worth of commerce every year.

Using AWS, you can requisition compute power, storage, and other services in minutes and have the flexibility to choose the development platform or programming model that makes the most sense for the problems they're trying to solve. You pay only for what you use, with no up-front expenses or long-term commitments, making AWS a cost-effective way to deliver applications.

Here are some of examples of how organizations, from research firms to large enterprises, use AWS today:

- A large enterprise quickly and economically deploys new internal applications, such as HR solutions, payroll applications, inventory management solutions, and online training to its distributed workforce.
- An e-commerce website accommodates sudden demand for a "hot" product caused by viral buzz from Facebook and Twitter without having to upgrade its infrastructure.
- A pharmaceutical research firm executes large-scale simulations using computing power provided by AWS.
- Media companies serve unlimited video, music, and other media to their worldwide customer base.

### **Overview of AWS data centers across the globe**

AWS pioneered cloud computing in 2006, creating cloud infrastructure that allows you to securely build and innovate faster. AWS is continuously innovating the design and systems of their data centers to protect them from man-made and natural risks. Then AWS implement controls, build automated systems, and undergo third-party audits to confirm security and compliance. As a result, the most highly-regulated organizations in the world trust AWS every day. Take a virtual tour of one of our data centers to learn about our security approach to protect the data of millions of active monthly customers.

## PERIMETER LAYER

AWS data center physical security begins at the Perimeter Layer. This Layer includes a number of security features depending on the location, such as security guards, fencing, security feeds, intrusion detection technology, and other security measures. Scroll down to learn more about the types of security measures we have in place within the Perimeter Layer of the data centers we operate around the world.

**ACCESS IS SCRUTINIZED:** AWS restricts physical access to people who need to be at a location for a justified business reason. Employees and vendors who have a need to be present at a data center must first apply for access and provide a valid business justification. The request is reviewed by specially designated personnel, including an area access manager. If access is granted, it is revoked once necessary work is completed.

**ENTRY IS CONTROLLED AND MONITORED:** Entering the Perimeter Layer is a controlled process. We staff our entry gates with security officers and employ supervisors who monitor officers and visitors via security cameras. When approved individuals are on site, they are given a badge that requires multi-factor authentication and limits access to pre-approved areas.

**AWS DATA CENTER WORKERS ARE SCRUTINIZED, TOO:** AWS employees who routinely need access to a data center are given permissions to relevant areas of the facility based on job function. But their access is regularly scrutinized, too. Staff lists are routinely reviewed by an area access manager to ensure each employee's authorization is still necessary. If an employee doesn't have an ongoing business need to be at a data center, they have to go through the visitor process.

**MONITORING FOR UNAUTHORIZED ENTRY:** We are continuously watching for unauthorized entry on our property, using video surveillance, intrusion detection, and access log monitoring systems. Entrances are secured with devices that sound alarms if a door is forced or held open.

**AWS SECURITY OPERATIONS CENTERS MONITORS GLOBAL SECURITY:** AWS Security Operations Centers are located around the world and are responsible for monitoring, triaging, and executing security programs for our data centers. They oversee physical access management and intrusion detection response while also providing global, 24/7 support to the on-site data center security teams. In short, they support our security with continuous monitoring activities such as tracking access activities, revoking access permissions, and being available to respond to and analyze a potential security incident.

## INFRASTRUCTURE LAYER

The Infrastructure Layer is the data center building and the equipment and systems that keep it running. Components like back-up power equipment, the HVAC system, and fire suppression equipment are all part of the Infrastructure Layer. These devices and systems help protect servers and ultimately your data. Scroll down to learn more about the types of security measures we deploy in the Infrastructure Layer of our data centers.

**LAYER-BY-LAYER ACCESS REVIEW:** Like other layers, access to the Infrastructure Layer is restricted based on business need. By implementing a layer-by-layer access review, the right to enter every layer is not granted by default. Access to any particular layer is only granted if there is a specific need to access that specific layer.

**MAINTAINING EQUIPMENT IS A PART OF REGULAR OPERATIONS:** AWS teams run diagnostics on machines, networks, and backup equipment to ensure they're in working order now and in an emergency. Routine maintenance checks on data center equipment and utilities are part of our regular operations.

**EMERGENCY-READY BACKUP EQUIPMENT:** Water, power, telecommunications, and internet connectivity are designed with redundancy, so we can maintain continuous operations in an emergency. Electrical power systems are designed to be fully redundant so that in the event of a disruption, uninterruptible power supply units can be engaged for certain functions, while generators can provide backup power for the entire facility. People and systems monitor and control the temperature and humidity to prevent overheating, further reducing possible service outages.

## **DATA LAYER**

The Data Layer is the most critical point of protection because it is the only area that holds customer data. Protection begins by restricting access and maintaining a separation of privilege for each layer. In addition, we deploy threat detection devices, video surveillance and system protocols, further safeguarding this layer. Scroll down to learn more about some of the security measures we have in place within the Data Layer.

**TECHNOLOGY AND PEOPLE WORK TOGETHER FOR ADDED SECURITY:** There are mandatory procedures to obtain authorization to enter the Data Layer. This includes review and approval of a person's access application by authorized individuals. Meanwhile, threat and electronic intrusion detection systems monitor and automatically trigger alerts of identified threats or suspicious activity. For example, if a door is held or forced open an alarm is triggered. We deploy security cameras and retain footage in alignment with legal and compliance requirements.

**PREVENTING PHYSICAL AND TECHNOLOGICAL INTRUSION:** Access points to server rooms are fortified with electronic control devices that require multi-factor authorization. We're also prepared to prevent technological intrusion. AWS servers can warn employees of any attempts to remove data. In the unlikely event of a breach, the server is automatically disabled.

**SERVERS AND MEDIA RECEIVE EXACTING ATTENTION:** Media storage devices used to store customer data are classified by AWS as Critical and treated accordingly, as high impact, throughout their life-cycle. We have exacting standards on how to install, service, and eventually destroy the devices when they are no longer useful. When a storage device has reached the end of its useful life, AWS decommissions media using techniques detailed in NIST 800-88. Media that stored customer data is not removed from AWS control until it has been securely decommissioned.

**THIRD-PARTY AUDITORS VERIFY OUR PROCEDURES AND SYSTEMS:** AWS is audited by external auditors on more than 2,600 requirements throughout the year. When third-party auditors inspect our data centers they do a deep dive to confirm we're following established rules needed to obtain our security certifications. Depending on the compliance program and its requirements, external auditors may interview AWS employees about how they handle and dispose of media. Auditors may also watch security camera feeds and observe entrances and hallways throughout a data center. And they often examine equipment such as our electronic access control devices and security cameras.

## ENVIRONMENTAL LAYER

The Environmental Layer is dedicated to environmental considerations from site selection and construction to operations and sustainability. AWS carefully chooses our data center locations to mitigate environmental risk, such as flooding, extreme weather, and seismic activity. Scroll down to learn more about the types of security measures we have in place within the Environmental Layer.

**PREPARED FOR THE UNEXPECTED:** AWS proactively prepares for potential environmental threats, like natural disasters and fire. Installing automatic sensors and responsive equipment are two ways we safeguard our data centers. Water-detecting devices can alert employees to problems as automatic pumps work to remove liquid and prevent damage. Similarly, automatic fire detection and suppression equipment reduces risk and can notify AWS employees and firefighters of a problem.

**HIGH AVAILABILITY THROUGH MULTIPLE AVAILABILITY ZONES:** Unlike virtually every other technology infrastructure provider, each AWS Region has multiple Availability Zones. Each Availability Zone consists of one or more data centers, are physically separated from one another, and have redundant power and networking. Customers who care about high availability and performance of their applications can deploy applications across multiple Availability Zones in the same region for fault tolerance and low latency. Availability Zones are connected to each other with fast, private fiber-optic networking, enabling you to easily architect applications that automatically fail-over between Availability Zones without interruption.

**SIMULATING DISRUPTIONS & MEASURING OUR RESPONSE:** The AWS Business Continuity Plan is an operations process guide outlining how to avoid and lessen disruptions due to natural disasters with detailed steps to take before, during, and after an event. To mitigate and prepare for the unexpected, AWS tests the Business Continuity Plan regularly with drills that simulate different scenarios. We document how our people and processes perform, then debrief on lessons learned and any corrective actions that may be needed to improve our response rate. We are trained and ready to rebound from disruptions quickly, which includes a methodical recovery process to minimize further downtime due to errors.

**GREENER IN THE AWS CLOUD:** In addition to addressing environmental risks, we also incorporate sustainability considerations into our data center design. AWS has a long-term commitment to use 100% renewable energy. When companies move to the AWS Cloud from on-premises infrastructure, they typically reduce carbon emissions by 88% because our data centers can offer environmental economies of scale. Organizations generally use 77% fewer servers, 84% less power, and tap into a 28% cleaner mix of solar and wind power in the AWS Cloud versus their own data centers.

## Regions and Availability Zones

The AWS Cloud spans 60 Availability Zones within 20 geographic regions around the world, with announced plans for 12 more Availability Zones and four more AWS Regions in Bahrain, Cape Town, Hong Kong SAR, and Milan.

The AWS Global infrastructure is built around Regions and Availability Zones (AZs). AWS Regions provide multiple, physically separated and isolated Availability Zones which are connected with low latency, high throughput, and highly redundant networking. These Availability Zones offer AWS customers an easier and more effective way to design and operate applications and databases, making them more highly available, fault tolerant, and scalable than traditional single datacenter infrastructures or multi-datacenter infrastructures. For customers who specifically need to replicate their data or applications over greater geographic distances, there are AWS Local Regions. An AWS Local Region is a single datacenter designed to complement an existing AWS Region. Like all AWS Regions, AWS Local Regions are completely isolated from other AWS Regions.

### Global Infrastructure



**High Availability through Multiple Availability Zones:** Unlike virtually every other technology infrastructure provider, each AWS Region has multiple Availability Zones and data centers. As we've learned from running the leading cloud infrastructure technology platform since 2006, customers who care about the availability and performance of their applications want to deploy these applications across multiple Availability Zones in the same region for fault tolerance and low latency. Availability Zones are connected to each other with fast, private fiber-optic networking, enabling you to easily architect applications that automatically fail-over between Availability Zones without interruption.

**Improving Continuity With Replication Between Regions:** In addition to replicating applications and data across multiple data centers in the same Region using Availability Zones, you can also choose to increase redundancy and fault tolerance further by replicating data between geographic Regions. You can do so using both private, high speed networking and public internet connections to provide an additional layer of business continuity, or to provide low latency access across the globe.

**Meeting Compliance and Data Residency Requirements:** You retain complete control and ownership over the region in which your data is physically located, making it easy to meet regional compliance and data residency requirements.

**Geographic Expansion:** AWS has announced plans to expand with 12 new Availability Zones in four new geographic Regions: Bahrain, Cape Town, Hong Kong SAR, and Milan.

### **Overview of AWS products and Services**

Amazon Web Services offers a broad set of global cloud-based products including compute, storage, databases, analytics, networking, mobile, developer tools, management tools, IoT, security and enterprise applications. These services help organizations move faster, lower IT costs, and scale. AWS is trusted by the largest enterprises and hottest start-ups to power a wide variety of workloads including: web and mobile applications, game development, data processing and warehousing, storage, archive and many others.

**Amazon EC2:** Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers.

Amazon EC2's simple web service interface allows you to obtain and configure capacity with minimal friction. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment. Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change. Amazon EC2 changes the economics of computing by allowing you to pay only for capacity that you actually use. Amazon EC2 provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios.

**Amazon Virtual Private Cloud (VPC):** Amazon Virtual Private Cloud (Amazon VPC) lets you provision a logically isolated section of the Amazon Web Services (AWS) Cloud where you can launch AWS resources in a virtual network that you define. You have complete control over your virtual networking environment, including selection of your own IP address range, creation of subnets, and configuration of route tables and network gateways.

You can easily customize the network configuration for your Amazon VPC. For example, you can create a public-facing subnet for your web servers that has access to the Internet, and place your backend systems such as databases or application servers in a private-facing subnet with no Internet access. You can leverage multiple layers of security, including security groups and network access control lists, to help control access to Amazon EC2 instances in each subnet.

**AWS Elastic Beanstalk:** AWS Elastic Beanstalk is an easy-to-use service for deploying and scaling web application and services developed with Java, .Net, PHP, Node.js, Python, Ruby, GO and Docker on familiar servers such as Apache, Nginx, Passenger, and IIS.

You can simply upload your code and Elastic Beanstalk automatically handles the deployment, from capacity provisioning, load balancing, auto-scaling to application health monitoring. At the same time, you retain full control over the AWS resources powering your application and can access the underlying resources at any time.

**Auto Scaling:** Auto Scaling allows you to scale your Amazon EC2 capacity up or down automatically according to conditions that you define. With Auto Scaling, you can ensure that the number of Amazon EC2 instances you're using increases seamlessly during demand spikes to maintain performance, and decreases automatically during demand lulls to minimize costs. Auto Scaling is particularly well suited for applications that experience hourly, daily, or weekly variability in usage. Auto Scaling is enabled by Amazon CloudWatch and available at no additional charge beyond Amazon CloudWatch fees.

**Elastic Load Balancing:** Elastic Load Balancing automatically distributes incoming application traffic across multiple Amazon EC2 instances. It enables you to achieve even greater fault tolerance in your applications, seamlessly providing the amount of load balancing capacity needed in response to incoming application traffic. Elastic Load Balancing detects unhealthy instances within a pool and automatically reroutes traffic to healthy instances until the unhealthy instances have been restored. Customers can enable Elastic Load Balancing within a single Availability Zone or across multiple zones for even more consistent application performance.

**Amazon S3:** Amazon S3 is storage for internet. It is designed to make web scale computing easier for developers.

Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the web. It gives any developer access to the same highly scalable, reliable, secure, fast, inexpensive infrastructure that Amazon uses to run its own global network of web sites. The service aims to maximize benefits of scale and to pass those benefits on to developers.

**Amazon Elastic Block Store (EBS):** Amazon Elastic Block Store (EBS) provides block level storage volumes for use with Amazon EC2 instances. Amazon EBS volumes are network-attached, and persist independently from the life of an instance. Amazon EBS provides highly available, highly reliable, predictable storage volumes that can be attached to a running Amazon EC2 instance and exposed as a device within the instance. Amazon EBS is particularly suited for applications that require a database, file system, or access to raw block level storage.

**Amazon Glacier:** Amazon Glacier is an extremely low-cost storage service that provides secure and durable storage for data archiving and backup. In order to keep costs low, Amazon Glacier is optimized for data that is infrequently accessed and for which retrieval times of several hours are suitable. Amazon Glacier changes the game for data archiving and backup as you pay nothing upfront, pay a very low price for storage, and can scale your usage up or down as needed.

**Amazon Storage Gateway:** The Amazon Storage Gateway is a service connecting an on-premises software appliance with cloud-based storage to provide seamless and secure integration between an



organization's on-premises IT environment and AWS's storage infrastructure. The service allows you to securely store data in the AWS cloud for scalable and cost-effective storage. The AWS Storage Gateway supports industry-standard storage protocols that work with your existing applications. It provides low-latency performance by maintaining frequently accessed data on-premises while storing all of your data in Amazon Simple Storage Service (Amazon S3) or Amazon Glacier.

**Amazon Relational Database Service (RDS):** Amazon Relational Database Service (RDS) makes it easy to set up, operate, and scale a relational database in the cloud. It provides cost-efficient and resizable capacity while managing time-consuming database administration tasks, freeing you to focus on your applications and business.

Amazon RDS gives you access to several familiar database engines, including Amazon Aurora, MySQL, PostgreSQL, MariaDB, Oracle, and SQL Server. This means that the code, applications, and tools you already use with your existing databases can be used with Amazon RDS.

**Amazon DynamoDB:** DynamoDB is a fast, fully managed NoSQL database service that makes it simple and cost-effective to store and retrieve any amount of data, and serve any level of request traffic. All data items are stored on Solid State Drives (SSDs) for high availability and durability.

With DynamoDB, you can offload the administrative burden of operating and scaling a highly available distributed database cluster, while paying a low price for only what you use.

**AWS Management Console:** The Management Console provides a simple web interface for Amazon Web Services. You can log in using your AWS account name and password. If you've enabled AWS Multi-Factor Authentication, you will be prompted for your device's authentication code.

### **AWS account plans and split-ups**

All customers receive Basic support included with your AWS account. All plans, including Basic Support, provide access to customer service, AWS documentation, whitepapers, and support forums.

For access to technical support and additional Support resources, AWS offer plans to fit your unique needs.

	Basic	Developer	Business	Enterprise
<b>Customer Service and Communities</b>	Business hours* access to customer service, documentation, whitepapers, and support forums	Business hours* access to customer service, documentation, whitepapers, and support forums	Business hours* access to customer service, documentation, whitepapers, and support forums	Business hours* access to customer service, documentation, whitepapers, and support forums
<b>Best Practices</b>	Access to 2 core Trusted Advisor checks	Access to 2 core Trusted Advisor checks	Access to full set of Trusted Advisor checks	Access to full set of Trusted Advisor checks
<b>Technical Support</b>		Business hours* access to Cloud Support Associates via email	24x7 access to Cloud Support Engineers via email, chat & phone	24x7 access to Sr. Cloud Support Engineers via email, chat & phone
<b>Launch Support</b>			Infrastructure Event Management (Available for additional fee)	Infrastructure Event Management (Included)
<b>Programmatic Case Management</b>			AWS Support API	AWS Support API
<b>Third-Party Software Support</b>			Interoperability & configuration guidance and troubleshooting	Interoperability & configuration guidance and troubleshooting
<b>Operations Support</b>				Operational reviews, recommendations, and reporting
<b>Account Assistance</b>				Assigned Support Concierge

Proactive Guidance			Designated Technical Account Manager	
			(¥ 599 per month as minimum fee)	(¥ 100,000 per month as minimum fee)
Pricing**	Included	¥ 299 per month	<ul style="list-style-type: none"> <li>• 10% of monthly AWS usage for the first ¥ 0–¥ 60,000</li> <li>• 7% of monthly AWS usage from ¥ 60,000–¥ 500,000</li> <li>• 5% of monthly AWS usage from ¥ 500,000–¥ 1,500,000</li> <li>• 3% of monthly AWS usage over ¥ 1,500,000</li> </ul>	<ul style="list-style-type: none"> <li>• 10% of monthly AWS usage for the first ¥ 0–¥ 1,000,000</li> <li>• 7% of monthly AWS usage from ¥ 1,000,000–¥ 3,000,000</li> <li>• 5% of monthly AWS usage from ¥ 3,000,000–¥ 6,000,000</li> <li>• 3% of monthly AWS usage over ¥ 6,000,000</li> </ul>

### Multiple AWS accounts for billing purpose:

Amazon Web Services (AWS) enables customers to achieve significant gains in productivity, innovation, and cost reduction when they move to the AWS cloud. AWS offers a variety of services and features that allow for flexible control of cloud computing resources and also of the AWS account(s) managing those resources. These options help to ensure proper cost allocation, agility, and security, however customers are sometimes unsure of how to best implement an account strategy—especially when working with multiple AWS accounts. This webpage provides customers with account-level considerations, best practices, and high-level strategic guidance to help customers use AWS Organizations to structure and manage multiple AWS accounts for billing purposes.