Overview of Amazon Web Services

AWS Whitepaper

Overview of Amazon Web Services: AWS Whitepaper

Table of Contents

Overview of Amazon Web Services

Publication date: **April 2017** (*Document Details* (p. 42))

Abstract

The AWS Cloud provides a broad set of infrastructure services, such as computing power, storage options, networking and databases that are delivered as a utility: on-demand, available in seconds, with pay-as-you-go pricing. From data warehousing to deployment tools, directories to content delivery, over 90 AWS services are available. New services can be provisioned quickly, without upfront capital expense. This allows enterprises, start-ups, small and medium-sized businesses, and customers in the public sector to access the building blocks they need to respond quickly to changing business requirements. This whitepaper provides you with an overview of the benefits of the AWS Cloud and introduces you to the services that make up the platform.

Introduction

In 2006, Amazon Web Services (AWS) began offering IT infrastructure services to businesses in the form of web services—now commonly known as cloud computing. One of the key benefits of cloud computing is the opportunity to replace up-front capital infrastructure expenses with low variable costs that scale with your business. With the cloud, businesses no longer need to plan for and procure servers and other IT infrastructure weeks or months in advance. Instead, they can instantly spin up hundreds or thousands of servers in minutes and deliver results faster.

Today, AWS provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around the world.

What Is Cloud Computing?

Cloud computing is the on-demand delivery of compute power, database storage, applications, and other IT resources through a cloud services platform via the Internet with pay-as-you-go pricing. Whether you are running applications that share photos to millions of mobile users or you're supporting the critical operations of your business, a cloud services platform provides rapid access to flexible and low-cost IT resources. With cloud computing, you don't need to make large upfront investments in hardware and spend a lot of time on the heavy lifting of managing that hardware. Instead, you can provision exactly the right type and size of computing resources you need to power your newest bright idea or operate your IT department. You can access as many resources as you need, almost instantly, and only pay for what you use.

Cloud computing provides a simple way to access servers, storage, databases and a broad set of application services over the Internet. A cloud services platform such as Amazon Web Services owns and maintains the network-connected hardware required for these application services, while you provision and use what you need via a web application.

Six Advantages of Cloud Computing

- **Trade capital expense for variable expense** – Instead of having to invest heavily in data centers and servers before you know how you're going to use them, you can pay only when you consume computing resources, and pay only for how much you consume.
- **Benefit from massive economies of scale** – By using cloud computing, you can achieve a lower variable cost than you can get on your own. Because usage from hundreds of thousands of customers is aggregated in the cloud, providers such as AWS can achieve higher economies of scale, which translates into lower pay as-you-go prices.
- **Stop guessing about capacity** – Eliminate guessing on your infrastructure capacity needs. When you make a capacity decision prior to deploying an application, you often end up either sitting on expensive idle resources or dealing with limited capacity. With cloud computing, these problems go away. You can access as much or as little capacity as you need, and scale up and down as required with only a few minutes' notice.
- **Increase speed and agility** – In a cloud computing environment, new IT resources are only a click away, which means that you reduce the time to make those resources available to your developers from weeks to just minutes. This results in a dramatic increase in agility for the organization, since the cost and time it takes to experiment and develop is significantly lower.
- **Stop spending money running and maintaining data centers** – Focus on projects that differentiate your business, not the infrastructure. Cloud computing lets you focus on your own customers, rather than on the heavy lifting of racking, stacking, and powering servers.
- **Go global in minutes** – Easily deploy your application in multiple regions around the world with just a few clicks. This means you can provide lower latency and a better experience for your customers at minimal cost.

Types of Cloud Computing

Cloud computing provides developers and IT departments with the ability to focus on what matters most and avoid undifferentiated work such as procurement, maintenance, and capacity planning. As cloud computing has grown in popularity, several different models and deployment strategies have emerged to help meet specific needs of different users. Each type of cloud service and deployment method provides you with different levels of control, flexibility, and management. Understanding the differences between Infrastructure as a Service, Platform as a Service, and Software as a Service, as well as what deployment strategies you can use, can help you decide what set of services is right for your needs.

Cloud Computing Models

Infrastructure as a Service (IaaS)

Infrastructure as a Service (IaaS) contains the basic building blocks for cloud IT and typically provide access to networking features, computers (virtual or on dedicated hardware), and data storage space. IaaS provides you with the highest level of flexibility and management control over your IT resources and is most similar to existing IT resources that many IT departments and developers are familiar with today.

Platform as a Service (PaaS)

Platform as a Service (PaaS) removes the need for your organization to manage the underlying infrastructure (usually hardware and operating systems) and allows you to focus on the deployment and management of your applications. This helps you be more efficient as you don't need to worry about resource procurement, capacity planning, software maintenance, patching, or any of the other undifferentiated heavy lifting involved in running your application.

Software as a Service (SaaS)

Software as a Service (SaaS) provides you with a completed product that is run and managed by the service provider. In most cases, people referring to Software as a Service are referring to end-user applications. With a SaaS offering you do not have to think about how the service is maintained or how the underlying infrastructure is managed; you only need to think about how you will use that particular piece of software. A common example of a SaaS application is web-based email which you can use to send and receive email without having to manage feature additions to the email product or maintain the servers and operating systems that the email program is running on.

Cloud Computing Deployment Models

Cloud

A cloud-based application is fully deployed in the cloud and all parts of the application run in the cloud. Applications in the cloud have either been created in the cloud or have been migrated from an existing infrastructure to take advantage of the benefits of cloud computing. Cloud-based applications can be built on low-level infrastructure pieces or can use higher level services that provide abstraction from the management, architecting, and scaling requirements of core infrastructure.

Hybrid

A hybrid deployment is a way to connect infrastructure and applications between cloud-based resources and existing resources that are not located in the cloud. The most common method of hybrid deployment is between the cloud and existing on-premises infrastructure to extend, and grow, an organization's infrastructure into the cloud while connecting cloud resources to the internal system. For more information on how AWS can help you with your hybrid deployment, please visit our hybrid page.

On-premises

The deployment of resources on-premises, using virtualization and resource management tools, is sometimes called the "private cloud." On-premises deployment doesn't provide many of the benefits of cloud computing but is sometimes sought for its ability to provide dedicated resources. In most cases this deployment model is the same as legacy IT infrastructure while using application management and virtualization technologies to try and increase resource utilization.

Global Infrastructure

AWS serves over a million active customers in more than 190 countries. We are steadily expanding global infrastructure to help our customers achieve lower latency and higher throughput, and to ensure that their data resides only in the AWS Region they specify. As our customers grow their businesses, AWS will continue to provide infrastructure that meets their global requirements.

The AWS Cloud infrastructure is built around AWS Regions and Availability Zones. An AWS 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. These Availability Zones offer you the ability to operate production applications and databases that are more highly available, fault tolerant, and scalable than would be possible from a single data center. The AWS Cloud operates 42 Availability Zones within 16 geographic Regions around the world, with five more Availability Zones and two more Regions coming online in 2017.

Each Amazon Region is designed to be completely isolated from the other Amazon Regions. This achieves the greatest possible fault tolerance and stability. Each Availability Zone is isolated, but the Availability Zones in a Region are connected through low-latency links. AWS provides you with the flexibility to place instances and store data within multiple geographic regions as well as across multiple Availability Zones within each AWS Region. Each Availability Zone is designed as an independent failure zone. This means that Availability Zones are physically separated within a typical metropolitan region and are located in lower risk flood plains (specific flood zone categorization varies by AWS Region). In addition to discrete uninterruptable power supply (UPS) and onsite backup generation facilities, they are each fed via different grids from independent utilities to further reduce single points of failure. Availability Zones are all redundantly connected to multiple tier-1 transit providers.

Security and Compliance

Security

Cloud security at AWS is the highest priority. As an AWS customer, you will benefit from a data center and network architecture built to meet the requirements of the most security-sensitive organizations. Security in the cloud is much like security in your on-premises data centers—only without the costs of maintaining facilities and hardware. In the cloud, you don't have to manage physical servers or storage devices. Instead, you use software-based security tools to monitor and protect the flow of information into and of out of your cloud resources.

An advantage of the AWS Cloud is that it allows you to scale and innovate, while maintaining a secure environment and paying only for the services you use. This means that you can have the security you need at a lower cost than in an on-premises environment.

As an AWS customer you inherit all the best practices of AWS policies, architecture, and operational processes built to satisfy the requirements of our most security-sensitive customers. Get the flexibility and agility you need in security controls.

The AWS Cloud enables a shared responsibility model. While AWS manages security **of** the cloud, you are responsible for security **in** the cloud. This means that you retain control of the security you choose to implement to protect your own content, platform, applications, systems, and networks no differently than you would in an on-site data center.

AWS provides you with guidance and expertise through online resources, personnel, and partners. AWS provides you with advisories for current issues, plus you have the opportunity to work with AWS when you encounter security issues.

You get access to hundreds of tools and features to help you to meet your security objectives. AWS provides security-specific tools and features across network security, configuration management, access control, and data encryption.

Finally, AWS environments are continuously audited, with certifications from accreditation bodies across geographies and verticals. In the AWS environment, you can take advantage of automated tools for asset inventory and privileged access reporting.

Benefits of AWS Security

- **Keep Your Data Safe**: The AWS infrastructure puts strong safeguards in place to help protect your privacy. All data is stored in highly secure AWS data centers.
- **Meet Compliance Requirements**: AWS manages dozens of compliance programs in its infrastructure. This means that segments of your compliance have already been completed.
- **Save Money**: Cut costs by using AWS data centers. Maintain the highest standard of security without having to manage your own facility
- **Scale Quickly**: Security scales with your AWS Cloud usage. No matter the size of your business, the AWS infrastructure is designed to keep your data safe.

Compliance

AWS Cloud Compliance enables you to understand the robust controls in place at AWS to maintain security and data protection in the cloud. As systems are built on top of AWS Cloud infrastructure,

compliance responsibilities will be shared. By tying together governance-focused, audit-friendly service features with applicable compliance or audit standards, AWS Compliance enablers build on traditional programs. This helps customers to establish and operate in an AWS security control environment.

The IT infrastructure that AWS provides to its customers is designed and managed in alignment with best security practices and a variety of IT security standards. The following is a partial list of assurance programs with which AWS complies:

- SOC 1/ISAE 3402, SOC 2, SOC 3
- FISMA, DIACAP, and FedRAMP
- PCI DSS Level 1
- ISO 9001, ISO 27001, ISO 27018

AWS provides customers a wide range of information on its IT control environment in whitepapers, reports, certifications, accreditations, and other third-party attestations. More information is available in the Risk and Compliance whitepaper and the AWS Security Center.

Amazon Web Services Cloud Platform

AWS consists of many cloud services that you can use in combinations tailored to your business or organizational needs. This section introduces the major AWS services by category. To access the services, you can use the AWS Management Console, the Command Line Interface, or Software Development Kits (SDKs).

Topics

AWS Management Console

Access and manage Amazon Web Services through the AWS Management Console, a simple and intuitive user interface. You can also use the AWS Console Mobile App to quickly view resources on the go.

AWS Command Line Interface

The AWS Command Line Interface (CLI) is a unified tool to manage your AWS services. With just one tool to download and configure, you can control multiple AWS services from the command line and automate them through scripts.

Software Development Kits

Our Software Development Kits (SDKs) simplify using AWS services in your applications with an Application Program Interface (API) tailored to your programming language or platform.

Services

Compute Services

Topics

Amazon EC2

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

The Amazon EC2 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 (called Amazon EC2 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 and system administrators the tools to build failure resilient applications and isolate themselves from common failure scenarios.

Benefits

Elastic Web-Scale Computing

Amazon EC2 enables you to increase or decrease capacity within minutes, not hours or days. You can commission one, hundreds, or even thousands of server instances simultaneously. Because this is all controlled with web service APIs, your application can automatically scale itself up and down depending on its needs.

Completely Controlled

You have complete control of your Amazon EC2 instances. You have root access to each one, and you can interact with them as you would any machine. You can stop your Amazon EC2 instance while retaining the data on your boot partition, and then subsequently restart the same instance using web service APIs. Instances can be rebooted remotely using web service APIs.

Flexible Cloud Hosting Services

You can choose among multiple instance types, operating systems, and software packages. Amazon EC2 allows you to select the memory configuration, CPU, instance storage, and boot partition size that are optimal for your choice of operating system and application. For example, your choice of operating systems includes numerous Linux distributions and Microsoft Windows Server.

Integrated

Amazon EC2 is integrated with most AWS services, such as Amazon Simple Storage Service (Amazon S3), Amazon Relational Database Service (Amazon RDS), and Amazon Virtual Private Cloud (Amazon VPC) to provide a complete, secure solution for computing, query processing, and cloud storage across a wide range of applications.

Reliable

Amazon EC2 offers a highly reliable environment where replacement instances can be rapidly and predictably commissioned. The service runs within Amazon's proven network infrastructure and data centers. The Amazon EC2 Service Level Agreement (SLA) commitment is 99.95% availability for each Region.

Secure

Amazon EC2 works in conjunction with Amazon VPC to provide security and robust networking functionality for your compute resources.

- Your compute instances are located in a VPC with an IP address range that you specify. You decide which instances are exposed to the Internet and which remain private.
- Security groups and network access control lists (ACLs) allow you to control inbound and outbound network access to and from your instances.
- You can connect your existing IT infrastructure to resources in your VPC using industry-standard encrypted IPsec virtual private network (VPN) connections.
- You can provision your Amazon EC2 resources as Dedicated Instances. Dedicated Instances are Amazon EC2 instances that run on hardware dedicated to a single customer for additional isolation.
- You can provision your Amazon EC2 resources on Dedicated Hosts, which are physical servers with EC2 instance capacity fully dedicated to your use. Dedicated Hosts can help you address compliance requirements and reduce costs by allowing you to use your existing server-bound software licenses.

Inexpensive

Amazon EC2 passes on to you the financial benefits of Amazon's scale. You pay a very low rate for the compute capacity you actually consume. See Amazon EC2 Instance Purchasing Options for a more detailed description.

- **On-Demand Instances**—With On-Demand instances, you pay for compute capacity by the hour with no long-term commitments. You can increase or decrease your compute capacity depending on the demands of your application and only pay the specified hourly rate for the instances you use. The use of On-Demand instances frees you from the costs and complexities of planning, purchasing, and maintaining hardware and transforms what are commonly large fixed costs into much smaller variable costs. On-Demand instances also remove the need to buy "safety net" capacity to handle periodic traffic spikes.
- **Reserved Instances**—Reserved Instances provide you with a significant discount (up to 75%) compared to On-Demand instance pricing. You have the flexibility to change families, operating system types, and tenancies while benefitting from Reserved Instance pricing when you use Convertible Reserved Instances.
- **Spot Instances**—Spot Instances allow you to bid on spare Amazon EC2 computing capacity. Since Spot instances are often available at a discount compared to On-Demand pricing, you can significantly reduce the cost of running your applications, grow your application's compute capacity and throughput for the same budget, and enable new types of cloud computing applications.

Amazon EC2 Container Service

Amazon EC2 Container Service (ECS) is a highly scalable, high-performance container management service that supports Docker containers. It allows you to easily run applications on a managed cluster of Amazon EC2 instances. Amazon ECS eliminates the need for you to install, operate, and scale your own cluster management infrastructure. With simple API calls, you can launch and stop Docker-enabled applications, query the complete state of your cluster, and access many familiar features like security groups, Elastic Load Balancing (p. 22), Amazon Elastic Block Store (Amazon EBS) (p. 15) volumes, and AWS Identity and Access Management (IAM) (p. 26) roles. You can use Amazon ECS to schedule the placement of containers across your cluster based on your resource needs and availability requirements. You can also integrate your own scheduler or third-party schedulers to meet business- or application-specific requirements.

Amazon EC2 Container Registry

Amazon EC2 Container Registry (ECR) is a fully-managed Docker container registry that makes it easy for developers to store, manage, and deploy Docker container images. Amazon ECR is integrated with Amazon EC2 Container Service (ECS) (p. 12), simplifying your development to production workflow. Amazon ECR eliminates the need to operate your own container repositories or worry about scaling the underlying infrastructure. Amazon ECR hosts your images in a highly available and scalable architecture, allowing you to reliably deploy containers for your applications. Integration with AWS Identity and Access Management (IAM) (p. 26) provides resource-level control of each repository. With Amazon ECR, there are no upfront fees or commitments. You pay only for the amount of data you store in your repositories and data transferred to the Internet.

Amazon Lightsail

Amazon Lightsail is designed to be the easiest way to launch and manage a virtual private server with AWS. Lightsail plans include everything you need to jumpstart your project – a virtual machine, SSD-based storage, data transfer, DNS management, and a static IP address – for a low, predictable price.

AWS Batch

AWS Batch enables developers, scientists, and engineers to easily and efficiently run hundreds of thousands of batch computing jobs on AWS. AWS Batch dynamically provisions the optimal quantity and type of compute resources (e.g., CPU or memory-optimized instances) based on the volume and specific resource requirements of the batch jobs submitted. With AWS Batch, there is no need to install and manage batch computing software or server clusters that you use to run your jobs, allowing you to focus on analyzing results and solving problems. AWS Batch plans, schedules, and executes your batch computing workloads across the full range of AWS compute services and features, such as Amazon EC2 and Spot Instances.

AWS Elastic Beanstalk

AWS Elastic Beanstalk is an easy-to-use service for deploying and scaling web applications and services developed with Java, .NET, PHP, Node.js, Python, Ruby, Go, and Docker on familiar servers such as Apache, Nginx, Passenger, and Internet Information Services (IIS).

You can simply upload your code, and AWS Elastic Beanstalk automatically handles the deployment, from capacity provisioning, load balancing, and 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.

AWS Lambda

AWS Lambda lets you run code without provisioning or managing servers. You pay only for the compute time you consume—there is no charge when your code is not running. With Lambda, you can run code for virtually any type of application or backend service—all with zero administration. Just upload your code, and Lambda takes care of everything required to run and scale your code with high availability. You can set up your code to automatically trigger from other AWS services, or you can call it directly from any web or mobile app.

Auto Scaling

Auto Scaling helps you maintain application availability and allows you to scale your Amazon EC2 capacity up or down automatically according to conditions that you define. You can use Auto Scaling to help ensure that you are running your desired number of Amazon EC2 instances. Auto Scaling can also automatically increase the number of Amazon EC2 instances during demand spikes to maintain performance and decrease capacity during lulls to reduce costs. Auto Scaling is well suited both to applications that have stable demand patterns and applications that experience hourly, daily, or weekly variability in usage.

Storage

Topics

Amazon S3

Amazon Simple Storage Service (Amazon S3) is object storage with a simple web service interface to store and retrieve any amount of data from anywhere on the web. It is designed to deliver 99.999999999% durability, and scales past trillions of objects worldwide.

You can use Amazon S3 as primary storage for cloud-native applications; as a bulk repository, or "data lake," for analytics; as a target for backup and recovery and disaster recovery; and with serverless computing.

It's simple to move large volumes of data into or out of Amazon S3 with Amazon's cloud data migration options. Once data is stored in Amazon S3, it can be automatically tiered into lower cost, longer-term cloud storage classes like Amazon S3 Standard - Infrequent Access and Amazon Glacier for archiving.

Amazon S3 Features

Amazon S3 provides the most feature-rich object storage platform available in the cloud today.

- **Simple**: Amazon S3 is simple to use with a web-based management console and mobile app. Amazon S3 also provides full REST APIs and SDKs for easy integration with third-party technologies.

- **Durable**: Amazon S3 provides durable infrastructure to store important data and is designed for durability of 99.999999999% of objects. Your data is redundantly stored across multiple facilities and multiple devices in each facility.

- **Scalable**: With Amazon S3, you can store as much data as you want and access it when needed. You can stop guessing your future storage needs and scale up and down as required, dramatically increasing business agility.

- **Secure**: Amazon S3 supports data transfer over SSL and automatic encryption of your data once it is uploaded. You can also configure bucket policies to manage object permissions and control access to your data using IAM (p. 26).

- **Available**: Amazon S3 Standard is designed for up to 99.99% availability of objects over a given year and is backed by the Amazon S3 Service Level Agreement, ensuring that you can rely on it when needed. You can also choose an AWS Region to optimize for latency, minimize costs, or address regulatory requirements.

- **Low Cost**: Amazon S3 allows you to store large amounts of data at a very low cost. Using lifecycle policies, you can set policies to automatically migrate your data to Standard - Infrequent Access and Amazon Glacier as it ages to further reduce costs.

- **Simple Data Transfer**: Amazon provides multiple options for cloud data migration, and makes it simple and cost-effective for you to move large volumes of data into or out of Amazon S3. You can choose from network-optimized, physical disk-based, or third-party connector methods for import to or export from Amazon S3.

- **Integrated**: Amazon S3 is deeply integrated with other AWS services to make it easier to build solutions that use a range of AWS services. Integrations include Amazon CloudFront (p. 21), Amazon CloudWatch (p. 23), Amazon Kinesis (p. 29), Amazon RDS (p. 17), Amazon Glacier (p. 15), Amazon EBS (p. 15), Amazon DynamoDB (p. 17), Amazon Redshift (p. 30), Amazon Route 53 (p. 21), Amazon EMR (p. 29), Amazon VPC (p. 20), Amazon Key Management Service (KMS) (p. 28), and AWS Lambda (p. 13).

- **Easy to Manage**: Amazon S3 Storage Management features allow you to take a data-driven approach to storage optimization, data security, and management efficiency. These enterprise-class capabilities give you data about your data, so you can manage your storage based on that personalized metadata.

Amazon Elastic Block Store

Amazon Elastic Block Store (Amazon EBS) provides persistent block storage volumes for use with Amazon EC2 instances in the AWS Cloud. Each Amazon EBS volume is automatically replicated within its Availability Zone to protect you from component failure, offering high availability and durability. Amazon EBS volumes offer the consistent and low-latency performance needed to run your workloads. With Amazon EBS, you can scale your usage up or down within minutes—all while paying a low price for only what you provision.

Amazon EBS Features

- **High Performance Volumes**: Choose between solid-state disk (SSD)-backed or hard disk drive (HDD)-backed volumes that can deliver the performance you need for your most demanding applications.
- **Availability**: Each Amazon EBS volume is designed for 99.999% availability and automatically replicates within its Availability Zone to protect your applications from component failure.
- **Encryption**: Amazon EBS encryption provides seamless support for data-at-rest and data-in-transit between EC2 instances and EBS volumes.
- **Access Management**: Amazon's flexible access control policies allow you to specify who can access which EBS volumes ensuring secure access to your data.
- **Snapshots**: Protect your data by creating point-in-time snapshots of EBS volumes, which are backed up to Amazon S3 for long-term durability.

Amazon Elastic File System

Amazon Elastic File System (Amazon EFS) provides simple, scalable file storage for use with Amazon EC2 instances in the AWS Cloud. Amazon EFS is easy to use and offers a simple interface that allows you to create and configure file systems quickly and easily. With Amazon EFS, storage capacity is elastic, growing and shrinking automatically as you add and remove files, so your applications have the storage they need, when they need it.

When mounted on Amazon EC2 instances, an Amazon EFS file system provides a standard file system interface and file system access semantics, allowing you to seamlessly integrate Amazon EFS with your existing applications and tools. Multiple EC2 instances can access an Amazon EFS file system at the same time, allowing Amazon EFS to provide a common data source for workloads and applications running on more than one EC2 instance.

You can mount your Amazon EFS file systems on your on-premises data center servers when connected to your VPC with AWS Direct Connect. You can mount your Amazon EFS file systems on on-premises servers to migrate data sets to EFS, enable cloud bursting scenarios, or backup your on-premises data to EFS.

Amazon EFS is designed for high availability and durability, and provides performance for a broad spectrum of workloads and applications, including big data and analytics, media processing workflows, content management, web serving, and home directories.

Amazon Glacier

Amazon Glacier is a secure, durable, and extremely low-cost storage service for data archiving and long-term backup. You can reliably store large or small amounts of data for as little as $0.004 per gigabyte per month, a significant savings compared to on-premises solutions. To keep costs low yet suitable for varying retrieval needs, Amazon Glacier provides three options for access to archives, from a few minutes to several hours.

AWS Storage Gateway

The AWS Storage Gateway service seamlessly enables hybrid storage between on-premises storage environments and the AWS Cloud. It combines a multi-protocol storage appliance with highly efficient network connectivity to Amazon cloud storage services, delivering local performance with virtually unlimited scale. You can use it in remote offices and data centers for hybrid cloud workloads involving migration, bursting, and storage tiering.

Database

Topics

Amazon Aurora

Amazon Aurora is a MySQL and PostgreSQL compatible relational database engine that combines the speed and availability of high-end commercial databases with the simplicity and cost-effectiveness of open source databases. Amazon Aurora provides up to five times better performance than MySQL with the security, availability, and reliability of a commercial database at one tenth the cost.

Benefits

- **High Performance**: Amazon Aurora provides 5 times the throughput of standard MySQL or twice the throughput of standard PostgreSQL running on the same hardware. This consistent performance is on par with commercial databases, at one-tenth of the cost. On the largest Amazon Aurora instance, you can achieve up to 500,000 reads and 100,000 writes per second. You can further scale read operations using read replicas that have very low 10 ms latency.
- **Highly Secure**: Amazon Aurora provides multiple levels of security for your database. These include network isolation using Amazon VPC, encryption at rest using keys you create and control through AWS Key Management Service (KMS), and encryption of data in transit using SSL. On an encrypted Amazon Aurora instance, data in the underlying storage is encrypted, as are the automated backups, snapshots, and replicas in the same cluster.
- **MySQL and PostgreSQL Compatible**: The Amazon Aurora database engine is fully compatible with MySQL 5.6 using the InnoDB storage engine. This means the code, applications, drivers, and tools you already use with your MySQL databases can be used with Amazon Aurora with little or no change. This also allows for easy migration of existing MySQL databases using standard MySQL import and export tools or using MySQL binlog replication. We're now previewing PostgreSQL compatible Amazon Aurora database instances, supporting the SQL dialect and functionality of PostgreSQL 9.6.
- **Highly Scalable**: You can scale your Amazon Aurora database from an instance with 2 vCPUs and 4 GiB of memory up to an instance with 32 vCPUs and 244 GiB of memory. You can also add up to 15 low latency read replicas across three Availability Zones to further scale read capacity. Amazon Aurora automatically grows storage as needed, from 10 GB up to 64 TB.
- **High Availability and Durability**: Amazon Aurora is designed to offer greater than 99.99% availability. Recovery from physical storage failures is transparent, and instance failover typically requires less than 30 seconds. Amazon Aurora's storage is fault-tolerant and self-healing. Six copies of your data are replicated across three Availability Zones and continuously backed up to Amazon S3.
- **Fully Managed**: Amazon Aurora is a fully managed database service. You no longer need to worry about database management tasks such as hardware provisioning, software patching, setup,

configuration, monitoring, or backups. Amazon Aurora automatically and continuously monitors and backs up your database to S3, enabling granular point-in-time recovery.

Amazon RDS

Amazon Relational Database Service (Amazon 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 up to focus on your applications and business. Amazon RDS provides you six familiar database engines to choose from, including Amazon Aurora (p. 16), PostgreSQL, MySQL, MariaDB, Oracle, and Microsoft SQL Server.

Benefits

- **Fast and Easy to Administer**: Amazon RDS makes it easy to go from project conception to deployment. Use the AWS Management Console, the AWS RDS Command Line Interface, or simple API calls to access the capabilities of a production-ready relational database in minutes. No need for infrastructure provisioning, and no need for installing and maintaining database software.
- **Highly Scalable**: You can scale your database's compute and storage resources with only a few mouse clicks or an API call, often with no downtime. Many Amazon RDS engine types allow you to launch one or more Read Replicas to offload read traffic from your primary database instance.
- **Available and Durable**: Amazon RDS runs on the same highly reliable infrastructure used by other Amazon Web Services. When you provision a Multi-AZ DB instance, Amazon RDS synchronously replicates the data to a standby instance in a different Availability Zone (AZ). Amazon RDS has many other features that enhance reliability for critical production databases, including automated backups, database snapshots, and automatic host replacement.
- **Secure**: Amazon RDS makes it easy to control network access to your database. Amazon RDS also lets you run your database instances in Amazon VPC, which enables you to isolate your database instances and to connect to your existing IT infrastructure through an industry-standard encrypted IPsec VPN. Many Amazon RDS engine types offer encryption at rest and encryption in transit.
- **Inexpensive**: You pay very low rates and only for the resources you actually consume. In addition, you benefit from the option of On-Demand pricing with no up-front or long-term commitments, or even lower hourly rates using our Reserved Instance pricing.

Amazon DynamoDB

Amazon DynamoDB is a fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale. It is a fully managed database and supports both document and key-value data models. Its flexible data model and reliable performance make it a great fit for mobile, web, gaming, ad-tech, Internet of Things (IoT), and many other applications.

Benefits

- **Fast, Consistent Performance**: Amazon DynamoDB is designed to deliver consistent, fast performance at any scale for all applications. Average service-side latencies are typically single-digit milliseconds. As your data volumes grow and application performance demands increase, Amazon DynamoDB uses automatic partitioning and SSD technologies to meet your throughput requirements and deliver low latencies at any scale.
- **Highly Scalable**: When you create a table, simply specify how much request capacity you require. If your throughput requirements change, simply update your table's request capacity using the AWS Management Console or the Amazon DynamoDB APIs. Amazon DynamoDB manages all the scaling behind the scenes, and you are still able to achieve your prior throughput levels while scaling is underway.

- **Fully Managed**: Amazon DynamoDB is a fully managed cloud NoSQL database service. You simply create a database table, set your throughput, and let the service handle the rest. You no longer need to worry about database management tasks such as hardware or software provisioning, setup and configuration, software patching, operating a reliable, distributed database cluster, or partitioning data over multiple instances as you scale.
- **Event-Driven Programming**: Amazon DynamoDB integrates with AWS Lambda (p. 13) to provide Triggers that enable you to architect applications that automatically react to data changes.
- **Fine-grained Access Control**: Amazon DynamoDB integrates with AWS IAM (p. 26) for fine-grained access control for users within your organization. You can assign unique security credentials to each user and control each user's access to services and resources.
- **Flexible:** Amazon DynamoDB supports both document and key-value data structures, giving you the flexibility to design the best architecture that is optimal for your application.

Amazon ElastiCache

Amazon ElastiCache is a web service that makes it easy to deploy, operate, and scale an in-memory cache in the cloud. The service improves the performance of web applications by allowing you to retrieve information from fast, managed, in-memory caches, instead of relying entirely on slower disk-based databases.

Amazon ElastiCache supports two open-source in-memory caching engines:

- Redis - a fast, open source, in-memory data store and cache. Amazon ElastiCache for Redis is a Redis-compatible in-memory service that delivers the ease-of-use and power of Redis along with the availability, reliability, and performance suitable for the most demanding applications. Both single-node and up to 15-shard clusters are available, enabling scalability to up to 3.55 TiB of in-memory data. ElastiCache for Redis is fully managed, scalable, and secure. This makes it an ideal candidate to power high-performance use cases such as web, mobile apps, gaming, ad-tech, and IoT.
- Memcached - a widely adopted memory object caching system. ElastiCache is protocol compliant with Memcached, so popular tools that you use today with existing Memcached environments will work seamlessly with the service.

Migration

Topics

AWS Application Discovery Service

AWS Application Discovery Service helps systems integrators quickly and reliably plan application migration projects by automatically identifying applications running in on-premises data centers, their associated dependencies, and their performance profiles.

Planning data center migrations can involve thousands of workloads that are often deeply interdependent. Application discovery and dependency mapping are important early first steps in the migration process, but these tasks are difficult to perform at scale due to the lack of automated tools.

AWS Application Discovery Service automatically collects configuration and usage data from servers, storage, and networking equipment to develop a list of applications, how they perform, and how they are interdependent. This information is retained in encrypted format in an AWS Application Discovery Service database, which you can export as a CSV or XML file into your preferred visualization tool or cloud migration solution to help reduce the complexity and time in planning your cloud migration.

AWS Database Migration Service

AWS Database Migration Service helps you migrate databases to AWS easily and securely. The source database remains fully operational during the migration, minimizing downtime to applications that rely on the database. The AWS Database Migration Service can migrate your data to and from most widely used commercial and open-source databases. The service supports homogenous migrations such as Oracle to Oracle, as well as heterogeneous migrations between different database platforms, such as Oracle to Amazon Aurora or Microsoft SQL Server to MySQL. It also allows you to stream data to Amazon Redshift from any of the supported sources including Amazon Aurora, PostgreSQL, MySQL, MariaDB, Oracle, SAP ASE, and SQL Server, enabling consolidation and easy analysis of data in the petabyte-scale data warehouse. AWS Database Migration Service can also be used for continuous data replication with high availability.

AWS Server Migration Service

AWS Server Migration Service (SMS) is an agentless service which makes it easier and faster for you to migrate thousands of on-premises workloads to AWS. AWS SMS allows you to automate, schedule, and track incremental replications of live server volumes, making it easier for you to coordinate large-scale server migrations.

See also AWS Database Migration Service (p. 19).

AWS Snowball

AWS Snowball is a petabyte-scale data transport solution that uses secure appliances to transfer large amounts of data into and out of AWS. The use of Snowball addresses common challenges with large-scale data transfers including high network costs, long transfer times, and security concerns. Transferring data with Snowball is simple, fast, secure, and can be as little as one-fifth the cost of high-speed Internet.

With Snowball, you don't need to write any code or purchase any hardware to transfer your data. Simply create a job in the AWS Management Console and a Snowball appliance will be automatically shipped to you. Once it arrives, attach the appliance to your local network, download and run the Snowball client to establish a connection, and then use the client to select the file directories that you want to transfer to the appliance. The client will then encrypt and transfer the files to the appliance at high speed. Once the transfer is complete and the appliance is ready to be returned, the E Ink shipping label will automatically update and you can track the job status using the Amazon Simple Notification Service (SNS) (p. 36), text messages, or directly in the console.

Snowball uses multiple layers of security designed to protect your data including tamper-resistant enclosures, 256-bit encryption, and an industry-standard Trusted Platform Module (TPM) designed to ensure both security and full chain of custody of your data. Once the data transfer job has been processed and verified, AWS performs a software erasure of the Snowball appliance.

AWS Snowball Edge

AWS Snowball Edge is a 100 TB data transfer device with on-board storage and compute capabilities. You can use Snowball Edge to move large amounts of data into and out of AWS, as a temporary storage tier for large local datasets, or to support local workloads in remote or offline locations.

Snowball Edge connects to your existing applications and infrastructure using standard storage interfaces, streamlining the data transfer process and minimizing setup and integration. Multiple Snowball Edge devices can be clustered together to form a local storage tier and process your data on-premises, helping ensure your applications continue to run even when they are not able to access the cloud.

AWS Snowmobile

AWS Snowmobile is an exabyte-scale data transfer service used to move extremely large amounts of data to AWS. You can transfer up to 100 PB per Snowmobile, a 45-foot long ruggedized shipping container, pulled by a semi-trailer truck. Snowmobile makes it easy to move massive volumes of data to the cloud, including video libraries, image repositories, or even a complete data center migration. Transferring data with Snowmobile is secure, fast, and cost effective.

After an initial assessment, a Snowmobile will be transported to your data center, and AWS personnel will configure it for you so it can be accessed as a network storage target. When your Snowmobile is on site, AWS personnel will work with your team to connect a removable, high-speed network switch from the Snowmobile to your local network. Then you can begin your high-speed data transfer from any number of sources within your data center to the Snowmobile. After your data is loaded, the Snowmobile is driven back to AWS where your data is imported into Amazon S3 or Amazon Glacier.

AWS Snowmobile uses multiple layers of security designed to protect your data including dedicated security personnel, GPS tracking, alarm monitoring, 24/7 video surveillance, and an optional escort security vehicle while in transit. All data is encrypted with 256-bit encryption keys managed through AWS KMS (p. 28) and designed to ensure both security and full chain of custody of your data.

Networking and Content Delivery

Topics

Amazon VPC

Amazon Virtual Private Cloud (Amazon VPC) lets you provision a logically isolated section of the 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 use both IPv4 and IPv6 in your VPC for secure and easy access to resources and applications.

You can easily customize the network configuration for your 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 EC2 instances in each subnet.

Additionally, you can create a hardware virtual private network (VPN) connection between your corporate data center and your VPC and leverage the AWS Cloud as an extension of your corporate data center.

Amazon CloudFront

Amazon CloudFront is a global content delivery network (CDN) service that accelerates delivery of your websites, APIs, video content, or other web assets. It integrates with other AWS products to give developers and businesses an easy way to accelerate content to end users with no minimum usage commitments.

Amazon CloudFront can be used to deliver your entire website, including dynamic, static, streaming, and interactive content using a global network of edge locations. Requests for your content are automatically routed to the nearest edge location, so content is delivered with the best possible performance. Amazon CloudFront is optimized to work with other services in AWS, such as Amazon S3 (p. 14), Amazon EC2 (p. 10), Elastic Load Balancing (p. 22), and Amazon Route 53 (p. 21). Amazon CloudFront also works seamlessly with any non-AWS origin server that stores the original, definitive versions of your files. Like other AWS products, there are no long-term contracts or minimum monthly usage commitments for using Amazon CloudFront—you pay only for as much or as little content as you actually deliver through the content delivery service.

Amazon Route 53

Amazon Route 53 is a highly available and scalable cloud Domain Name System (DNS) web service. It is designed to give developers and businesses an extremely reliable and cost-effective way to route end users to Internet applications by translating human readable names, such as www.example.com, into the numeric IP addresses, such as 192.0.2.1, that computers use to connect to each other. Amazon Route 53 is fully compliant with IPv6 as well.

Amazon Route 53 effectively connects user requests to infrastructure running in AWS—such as EC2 instances, Elastic Load Balancing load balancers, or Amazon S3 buckets—and can also be used to route users to infrastructure outside of AWS. You can use Amazon Route 53 to configure DNS health checks to route traffic to healthy endpoints or to independently monitor the health of your application and its endpoints. Amazon Route 53 traffic flow makes it easy for you to manage traffic globally through a variety of routing types, including latency-based routing, Geo DNS, and weighted round robin—all of which can be combined with DNS Failover in order to enable a variety of low-latency, fault-tolerant architectures. Using Amazon Route 53 traffic flow's simple visual editor, you can easily manage how your end users are routed to your application's endpoints—whether in a single AWS Region or distributed around the globe. Amazon Route 53 also offers Domain Name Registration—you can purchase and manage domain names such as example.com and Amazon Route 53 will automatically configure DNS settings for your domains.

AWS Direct Connect

AWS Direct Connect makes it easy to establish a dedicated network connection from your premises to AWS. Using AWS Direct Connect, you can establish private connectivity between AWS and your data center, office, or co-location environment, which in many cases can reduce your network costs, increase bandwidth throughput, and provide a more consistent network experience than Internet-based connections.

AWS Direct Connect lets you establish a dedicated network connection between your network and one of the AWS Direct Connect locations. Using industry standard 802.1Q virtual LANS (VLANs), this dedicated connection can be partitioned into multiple virtual interfaces. This allows you to use the same connection to access public resources, such as objects stored in Amazon S3 using public IP address space, and private resources such as EC2 instances running within a VPC using private IP address space, while maintaining network separation between the public and private environments. Virtual interfaces can be reconfigured at any time to meet your changing needs.

Elastic Load Balancing

Elastic Load Balancing (ELB) automatically distributes incoming application traffic across multiple EC2 instances. It enables you to achieve greater levels of fault tolerance in your applications, seamlessly providing the required amount of load balancing capacity needed to distribute application traffic.

Elastic Load Balancing offers two types of load balancers that both feature high availability, automatic scaling, and robust security. These include the Classic Load Balancer that routes traffic based on either application or network level information, and the Application Load Balancer that routes traffic based on advanced application-level information that includes the content of the request.The Classic Load Balancer is ideal for simple load balancing of traffic across multiple EC2 instances, while the Application Load Balancer is ideal for applications needing advanced routing capabilities, microservices, and container-based architectures. Application Load Balancer offers the ability to route traffic to multiple services or load balance across multiple ports on the same EC2 instance.

Developer Tools

Topics

AWS CodeCommit

AWS CodeCommit is a fully managed source control service that makes it easy for companies to host secure and highly scalable private Git repositories. AWS CodeCommit eliminates the need to operate your own source control system or worry about scaling its infrastructure. You can use AWS CodeCommit to securely store anything from source code to binaries, and it works seamlessly with your existing Git tools.

AWS CodeBuild

AWS CodeBuild is a fully managed build service that compiles source code, runs tests, and produces software packages that are ready to deploy. With CodeBuild, you don't need to provision, manage, and scale your own build servers. CodeBuild scales continuously and processes multiple builds concurrently, so your builds are not left waiting in a queue. You can get started quickly by using prepackaged build environments, or you can create custom build environments that use your own build tools.

AWS CodeDeploy

AWS CodeDeploy is a service that automates code deployments to any instance, including EC2 instances and instances running on premises. AWS CodeDeploy makes it easier for you to rapidly release new features, helps you avoid downtime during application deployment, and handles the complexity of updating your applications. You can use AWS CodeDeploy to automate software deployments, eliminating the need for error-prone manual operations. The service scales with your infrastructure so you can easily deploy to one instance or thousands.

AWS CodePipeline

AWS CodePipeline is a continuous integration and continuous delivery service for fast and reliable application and infrastructure updates. CodePipeline builds, tests, and deploys your code every time there is a code change, based on the release process models you define. This enables you to rapidly and reliably deliver features and updates. You can easily build out an end-to-end solution by using our pre-built plugins for popular third-party services like GitHub or by integrating your own custom plugins into any stage of your release process.

AWS X-Ray

AWS X-Ray helps developers analyze and debug distributed applications in production or under development, such as those built using a microservices architecture. With X-Ray, you can understand how your application and its underlying services are performing so you can identify and troubleshoot the root cause of performance issues and errors. X-Ray provides an end-to-end view of requests as they travel through your application, and shows a map of your application's underlying components. You can use X-Ray to analyze both applications in development and in production, from simple three-tier applications to complex microservices applications consisting of thousands of services.

Management Tools

Topics

Amazon CloudWatch

Amazon CloudWatch is a monitoring service for AWS Cloud resources and the applications you run on AWS. You can use Amazon CloudWatch to collect and track metrics, collect and monitor log files, set alarms, and automatically react to changes in your AWS resources. Amazon CloudWatch can monitor AWS resources such as Amazon EC2 instances, Amazon DynamoDB tables, and Amazon RDS DB instances, as well as custom metrics generated by your applications and services, and any log files your applications generate. You can use Amazon CloudWatch to gain system-wide visibility into resource utilization, application performance, and operational health. You can use these insights to react and keep your application running smoothly.

Amazon EC2 Systems Manager

Amazon EC2 Systems Manager is a management service that helps you automatically collect software inventory, apply operating system (OS) patches, create system images, and configure Windows and Linux

operating systems. These capabilities help you define and track system configurations, prevent drift, and maintain software compliance of your EC2 and on-premises configurations. By providing a management approach that is designed for the scale and agility of the cloud but extends into your on-premises data center, EC2 Systems Manager makes it easier for you to seamlessly bridge your existing infrastructure with AWS.

EC2 Systems Manager is easy to use. Simply access EC2 Systems Manager from the EC2 Management Console, select the instances you want to manage, and define the management tasks you want to perform. EC2 Systems Manager is available now at no cost to manage both your EC2 and on-premises resources.

EC2 Systems Manager contains the following tools:

- **Run Command:** Provides a simple way of automating common administrative tasks like remotely executing shell scripts or PowerShell commands, installing software updates, or making changes to the configuration of OS, software, EC2 and instances and servers in your on-premises data center.
- **State Manager:** Helps you define and maintain consistent OS configurations such as firewall settings and anti-malware definitions to comply with your policies. You can monitor the configuration of a large set of instances, specify a configuration policy for the instances, and automatically apply updates or configuration changes.
- **Inventory:** Helps you collect and query configuration and inventory information about your instances and the software installed on them. You can gather details about your instances such as installed applications, DHCP settings, agent detail, and custom items. You can run queries to track and audit your system configurations.
- **Maintenance Window:** Lets you define a recurring window of time to run administrative and maintenance tasks across your instances. This ensures that installing patches and updates, or making other configuration changes does not disrupt business-critical operations. This helps improve your application availability.
- **Patch Manager:** Helps you select and deploy operating system and software patches automatically across large groups of instances. You can define a maintenance window so that patches are applied only during set times that fit your needs. These capabilities help ensure that your software is always up to date and meets your compliance policies.
- **Automation:** Simplifies common maintenance and deployment tasks, such as updating Amazon Machine Images (AMIs). Use the Automation feature to apply patches, update drivers and agents, or bake applications into your AMI using a streamlined, repeatable, and auditable process.
- **Parameter Store:** Provides an encrypted location to store important administrative information such as passwords and database strings. The Parameter Store integrates with AWS KMS to make it easy to encrypt the information you keep in the Parameter Store.

AWS CloudFormation

AWS CloudFormation gives developers and systems administrators an easy way to create and manage a collection of related AWS resources, provisioning and updating them in an orderly and predictable fashion.

You can use the AWS CloudFormation sample templates or create your own templates to describe your AWS resources, and any associated dependencies or runtime parameters, required to run your application. You don't need to figure out the order for provisioning AWS services or the subtleties of making those dependencies work. CloudFormation takes care of this for you. After the AWS resources are deployed, you can modify and update them in a controlled and predictable way, in effect applying version control to your AWS infrastructure the same way you do with your software. You can also visualize your templates as diagrams and edit them using a drag-and-drop interface with the AWS CloudFormation Designer.

AWS CloudTrail

AWS CloudTrail is a web service that records AWS API calls for your account and delivers log files to you. The recorded information includes the identity of the API caller, the time of the API call, the source IP address of the API caller, the request parameters, and the response elements returned by the AWS service.

With CloudTrail, you can get a history of AWS API calls for your account, including API calls made using the AWS Management Console, AWS SDKs, command line tools, and higher-level AWS services (such as AWS CloudFormation (p. 24)). The AWS API call history produced by CloudTrail enables security analysis, resource change tracking, and compliance auditing.

AWS Config

AWS Config is a fully managed service that provides you with an AWS resource inventory, configuration history, and configuration change notifications to enable security and governance. The Config Rules feature enables you to create rules that automatically check the configuration of AWS resources recorded by AWS Config.

With AWS Config, you can discover existing and deleted AWS resources, determine your overall compliance against rules, and dive into configuration details of a resource at any point in time. These capabilities enable compliance auditing, security analysis, resource change tracking, and troubleshooting.

AWS OpsWorks

AWS OpsWorks is a configuration management service that uses Chef, an automation platform that treats server configurations as code. OpsWorks uses Chef to automate how servers are configured, deployed, and managed across your EC2 instances or on-premises compute environments. OpsWorks has two offerings, AWS OpsWorks for Chef Automate and AWS OpsWorks Stacks.

AWS Service Catalog

AWS Service Catalog allows organizations to create and manage catalogs of IT services that are approved for use on AWS. These IT services can include everything from virtual machine images, servers, software, and databases to complete multi-tier application architectures. AWS Service Catalog allows you to centrally manage commonly deployed IT services and helps you achieve consistent governance and meet your compliance requirements, while enabling users to quickly deploy only the approved IT services they need.

AWS Trusted Advisor

AWS Trusted Advisor is an online resource to help you reduce cost, increase performance, and improve security by optimizing your AWS environment. Trusted Advisor provides real-time guidance to help you provision your resources following AWS best practices.

AWS Personal Health Dashboard

AWS Personal Health Dashboard provides alerts and remediation guidance when AWS is experiencing events that might affect you. While the Service Health Dashboard displays the general status of AWS services, Personal Health Dashboard gives you a personalized view into the performance and availability of the AWS services underlying your AWS resources. The dashboard displays relevant and timely information to help you manage events in progress, and provides proactive notification to help you plan for scheduled activities. With Personal Health Dashboard, alerts are automatically triggered by changes

in the health of AWS resources, giving you event visibility and guidance to help quickly diagnose and resolve issues.

AWS Managed Services

AWS Managed Services provides ongoing management of your AWS infrastructure so you can focus on your applications. By implementing best practices to maintain your infrastructure, AWS Managed Services helps to reduce your operational overhead and risk. AWS Managed Services automates common activities such as change requests, monitoring, patch management, security, and backup services, and provides full-lifecycle services to provision, run, and support your infrastructure. Our rigor and controls help to enforce your corporate and security infrastructure policies, and enables you to develop solutions and applications using your preferred development approach. AWS Managed Services improves agility, reduces cost, and unburdens you from infrastructure operations so you can direct resources toward differentiating your business.

Security, Identity, and Compliance

Topics
- Amazon Cloud Directory (p. 26)
- AWS Identity and Access Management (p. 26)
- Amazon Inspector (p. 27)
- AWS Certificate Manager (p. 27)
- AWS CloudHSM (p. 27)
- AWS Directory Service (p. 27)
- AWS Key Management Service (p. 28)
- AWS Organizations (p. 28)
- AWS Shield (p. 28)
- AWS WAF (p. 28)

Amazon Cloud Directory

Amazon Cloud Directory enables you to build flexible, cloud-native directories for organizing hierarchies of data along multiple dimensions. With Cloud Directory, you can create directories for a variety of use cases, such as organizational charts, course catalogs, and device registries. While traditional directory solutions, such as Active Directory Lightweight Directory Services (AD LDS) and other LDAP-based directories, limit you to a single hierarchy, Cloud Directory offers you the flexibility to create directories with hierarchies that span multiple dimensions. For example, you can create an organizational chart that can be navigated through separate hierarchies for reporting structure, location, and cost center.

Amazon Cloud Directory automatically scales to hundreds of millions of objects and provides an extensible schema that can be shared with multiple applications. As a fully-managed service, Cloud Directory eliminates time-consuming and expensive administrative tasks, such as scaling infrastructure and managing servers. You simply define the schema, create a directory, and then populate your directory by making calls to the Cloud Directory API.

AWS Identity and Access Management

AWS Identity and Access Management (IAM) enables you to securely control access to AWS services and resources for your users. Using IAM, you can create and manage AWS users and groups, and use permissions to allow and deny their access to AWS resources. IAM allows you to do the following:

- Manage IAM users and their access: You can create users in IAM, assign them individual security credentials (access keys, passwords, and multi-factor authentication devices), or request temporary security credentials to provide users access to AWS services and resources. You can manage permissions in order to control which operations a user can perform.
- Manage IAM roles and their permissions: You can create roles in IAM and manage permissions to control which operations can be performed by the entity, or AWS service, that assumes the role. You can also define which entity is allowed to assume the role.
- Manage federated users and their permissions: You can enable identity federation to allow existing identities (users, groups, and roles) in your enterprise to access the AWS Management Console, call AWS APIs, and access resources, without the need to create an IAM user for each identity.

Amazon Inspector

Amazon Inspector is an automated security assessment service that helps improve the security and compliance of applications deployed on AWS. Amazon Inspector automatically assesses applications for vulnerabilities or deviations from best practices. After performing an assessment, Amazon Inspector produces a detailed list of security findings prioritized by level of severity.

To help you get started quickly, Amazon Inspector includes a knowledge base of hundreds of rules mapped to common security best practices and vulnerability definitions. Examples of built-in rules include checking for remote root login being enabled, or vulnerable software versions installed. These rules are regularly updated by AWS security researchers.

AWS Certificate Manager

AWS Certificate Manager is a service that lets you easily provision, manage, and deploy Secure Sockets Layer/Transport Layer Security (SSL/TLS) certificates for use with AWS services. SSL/TLS certificates are used to secure network communications and establish the identity of websites over the Internet. AWS Certificate Manager removes the time-consuming manual process of purchasing, uploading, and renewing SSL/TLS certificates. With AWS Certificate Manager, you can quickly request a certificate, deploy it on AWS resources such as Elastic Load Balancing load balancers or Amazon CloudFront distributions, and let AWS Certificate Manager handle certificate renewals. SSL/TLS certificates provisioned through AWS Certificate Manager are free. You pay only for the AWS resources you create to run your application.

AWS CloudHSM

The AWS CloudHSM service helps you meet corporate, contractual, and regulatory compliance requirements for data security by using dedicated Hardware Security Module (HSM) appliances within the AWS Cloud. The AWS CloudHSM service allows you to protect your encryption keys within HSMs designed and validated to government standards for secure key management. You can securely generate, store, and manage the cryptographic keys used for data encryption such that they are accessible only by you. AWS CloudHSM helps you comply with strict key management requirements without sacrificing application performance. CloudHSM instances are provisioned inside your VPC with an IP address that you specify, providing simple and private network connectivity to your EC2 instances. AWS provides dedicated and exclusive (single tenant) access to CloudHSM instances, isolated from other AWS customers.

AWS Directory Service

AWS Directory Service for Microsoft Active Directory (Enterprise Edition), also known as AWS Microsoft AD, enables your directory-aware workloads and AWS resources to use managed Active Directory in the AWS Cloud. The AWS Microsoft AD service is built on actual Microsoft Active Directory and does not require you to synchronize or replicate data from your existing Active Directory to the cloud. You

can use standard Active Directory administration tools and take advantage of built-in Active Directory features such as Group Policy, trusts, and single sign-on. With Microsoft AD, you can easily join Amazon EC2 and Amazon RDS for SQL Server instances to a domain, and use AWS Enterprise IT applications such as Amazon WorkSpaces with Active Directory users and groups.

AWS Key Management Service

AWS Key Management Service (KMS) is a managed service that makes it easy for you to create and control the encryption keys used to encrypt your data. This service uses HSMs to protect the security of your keys. AWS Key Management Service is integrated with several other AWS services to help you protect the data you store with these services. AWS Key Management Service is also integrated with AWS CloudTrail to provide you with logs of all key usage to help meet your regulatory and compliance needs.

AWS Organizations

AWS Organizations allows you to create groups of AWS accounts that you can use to more easily manage security and automation settings. With Organizations, you can centrally manage multiple accounts to help you scale. You can control which AWS services are available to individual accounts, automate new account creation, and simplify billing.

AWS Shield

AWS Shield is a managed Distributed Denial of Service (DDoS) protection service that safeguards web applications running on AWS. AWS Shield provides always-on detection and automatic inline mitigations that minimize application downtime and latency, so there is no need to engage AWS Support to benefit from DDoS protection. There are two tiers of AWS Shield: Standard and Advanced.

All AWS customers benefit from the automatic protections of AWS Shield Standard, at no additional charge. AWS Shield Standard defends against most common, frequently occurring network and transport layer DDoS attacks that target your website or applications.

For higher levels of protection against attacks targeting your web applications running on ELB, Amazon CloudFront, and Amazon Route 53 resources, you can subscribe to AWS Shield Advanced. In addition to the common network and transport layer protections that come with Standard, AWS Shield Advanced provides additional detection and mitigation against large and sophisticated DDoS attacks, near real-time visibility into attacks, and integration with AWS WAF, a web application firewall. AWS Shield Advanced also gives you access to the AWS DDoS Response Team (DRT) and protection against DDoS related spikes in your ELB, CloudFront, or Route 53 charges.

AWS WAF

AWS WAF is a web application firewall that helps protect your web applications from common web exploits that could affect application availability, compromise security, or consume excessive resources. AWS WAF gives you control over which traffic to allow or block to your web application by defining customizable web security rules. You can use AWS WAF to create custom rules that block common attack patterns, such as SQL injection or cross-site scripting, and rules that are designed for your specific application. New rules can be deployed within minutes, letting you respond quickly to changing traffic patterns. Also, AWS WAF includes a full-featured API that you can use to automate the creation, deployment, and maintenance of web security rules.

Analytics

Topics
- Amazon Athena (p. 29)

Amazon Athena

Amazon Athena is an interactive query service that makes it easy to analyze data in Amazon S3 using standard SQL. Athena is serverless, so there is no infrastructure to manage, and you pay only for the queries that you run.

Athena is easy to use. Simply point to your data in Amazon S3, define the schema, and start querying using standard SQL. Most results are delivered within seconds. With Athena, there's no need for complex extract, transform, and load (ETL) jobs to prepare your data for analysis. This makes it easy for anyone with SQL skills to quickly analyze large-scale datasets.

Amazon EMR

Amazon EMR provides a managed Hadoop framework that makes it easy, fast, and cost-effective to process vast amounts of data across dynamically scalable EC2 instances. You can also run other popular distributed frameworks such as Apache Spark, HBase, Presto, and Flink in Amazon EMR, and interact with data in other AWS data stores such as Amazon S3 and Amazon DynamoDB.

Amazon EMR securely and reliably handles a broad set of big data use cases, including log analysis, web indexing, data transformations (ETL), machine learning, financial analysis, scientific simulation, and bioinformatics.

Amazon CloudSearch

Amazon CloudSearch is a managed service in the AWS Cloud that makes it simple and cost-effective to set up, manage, and scale a search solution for your website or application. Amazon CloudSearch supports 34 languages and popular search features such as highlighting, autocomplete, and geospatial search.

Amazon Elasticsearch Service

Amazon Elasticsearch Service makes it easy to deploy, operate, and scale Elasticsearch for log analytics, full text search, application monitoring, and more. Amazon Elasticsearch Service is a fully managed service that delivers Elasticsearch's easy-to-use APIs and real-time capabilities along with the availability, scalability, and security required by production workloads. The service offers built-in integrations with Kibana, Logstash, and AWS services including Amazon Kinesis Firehose (p. 30), AWS Lambda (p. 13), and Amazon CloudWatch (p. 23) so that you can go from raw data to actionable insights quickly.

Amazon Kinesis

Amazon Kinesis is a platform for streaming data on AWS, offering powerful services to make it easy to load and analyze streaming data, and also providing the ability for you to build custom streaming

data applications for specialized needs. Web applications, mobile devices, wearables, industrial sensors, and many software applications and services can generate staggering amounts of streaming data—sometimes terabytes per hour—that need to be collected, stored, and processed continuously. Amazon Kinesis services enable you to do that simply and at a low cost.

Amazon Kinesis currently offers three services: Amazon Kinesis Firehose, Amazon Kinesis Analytics, and Amazon Kinesis Streams.

Amazon Kinesis Firehose

Amazon Kinesis Firehose is the easiest way to load streaming data into AWS. It can capture, transform, and load streaming data into Amazon Kinesis Analytics, Amazon S3, Amazon Redshift, and Amazon Elasticsearch Service, enabling near real-time analytics with existing business intelligence tools and dashboards you're already using today. It is a fully managed service that automatically scales to match the throughput of your data and requires no ongoing administration. It can also batch, compress, and encrypt the data before loading it, minimizing the amount of storage used at the destination and increasing security.

You can easily create a Firehose delivery stream from the AWS Management Console, configure it with a few clicks, and start sending data to the stream from hundreds of thousands of data sources to be loaded continuously to AWS—all in just a few minutes.

Amazon Kinesis Analytics

Amazon Kinesis Analytics is the easiest way to process streaming data in real time with standard SQL without having to learn new programming languages or processing frameworks. Amazon Kinesis Analytics enables you to create and run SQL queries on streaming data so that you can gain actionable insights and respond to your business and customer needs promptly.

Amazon Kinesis Analytics takes care of everything required to run your queries continuously and scales automatically to match the volume and throughput rate of your incoming data.

Amazon Kinesis Streams

Amazon Kinesis Streams enables you to build custom applications that process or analyze streaming data for specialized needs. Amazon Kinesis Streams can continuously capture and store terabytes of data per hour from hundreds of thousands of sources such as website clickstreams, financial transactions, social media feeds, IT logs, and location-tracking events. With Amazon Kinesis Client Library (KCL), you can build Amazon Kinesis Applications and use streaming data to power real-time dashboards, generate alerts, implement dynamic pricing and advertising, and more. You can also emit data from Amazon Kinesis Streams to other AWS services such as Amazon S3 (p. 14), Amazon Redshift (p. 30), Amazon EMR (p. 29), and AWS Lambda (p. 13).

Amazon Redshift

Amazon Redshift is a fast, fully managed, petabyte-scale data warehouse that makes it simple and cost-effective to analyze all your data using your existing business intelligence tools. Start small for $0.25 per hour with no commitments and scale to petabytes for $1,000 per terabyte per year, less than a tenth of the cost of traditional solutions. Customers typically see 3x compression, reducing their costs to $333 per uncompressed terabyte per year.

Amazon Redshift uses a variety of innovations to obtain very high query performance on datasets ranging in size from a hundred gigabytes to a petabyte or more. It uses columnar storage, data compression, and zone maps to reduce the amount of I/O needed to perform queries. Amazon Redshift has a massively parallel processing (MPP) data warehouse architecture, parallelizing and distributing SQL operations to take advantage of all available resources. The underlying hardware is designed for high

performance data processing, using local attached storage to maximize throughput between the CPUs and drives, and a 10GigE mesh network to maximize throughput between nodes.

With a few clicks in the console or a simple API call, you can easily change the number or type of nodes in your data warehouse and scale up all the way to a petabyte or more of compressed user data. Dense Storage (DS) nodes allow you to create very large data warehouses using hard disk drives (HDDs) for a very low price point. Dense Compute (DC) nodes allow you to create very high performance data warehouses using fast CPUs, large amounts of RAM and SSDs. While resizing, Amazon Redshift allows you to continue to query your data warehouse in read-only mode until the new cluster is fully provisioned and ready for use.

Amazon QuickSight

Amazon QuickSight is a fast, cloud-powered business analytics service that makes it easy to build visualizations, perform ad-hoc analysis, and quickly get business insights from your data. Using our cloud-based service you can easily connect to your data, perform advanced analysis, and create stunning visualizations and rich dashboards that can be accessed from any browser or mobile device.

AWS Data Pipeline

AWS Data Pipeline is a web service that helps you reliably process and move data between different AWS compute and storage services, as well as on-premises data sources, at specified intervals. With AWS Data Pipeline, you can regularly access your data where it's stored, transform and process it at scale, and efficiently transfer the results to AWS services such as Amazon S3 (p. 14), Amazon RDS (p. 17), Amazon DynamoDB (p. 17), and Amazon EMR (p. 29).

AWS Data Pipeline helps you easily create complex data processing workloads that are fault tolerant, repeatable, and highly available. You don't have to worry about ensuring resource availability, managing inter-task dependencies, retrying transient failures or timeouts in individual tasks, or creating a failure notification system. AWS Data Pipeline also allows you to move and process data that was previously locked up in on-premises data silos.

AWS Glue

AWS Glue is a fully managed ETL service that makes it easy to move data between your data stores. AWS Glue simplifies and automates the difficult and time-consuming tasks of data discovery, conversion, mapping, and job scheduling. AWS Glue guides you through the process of moving your data with an easy-to-use console that helps you understand your data sources, prepare the data for analytics, and load it reliably from data sources to destinations.

AWS Glue is integrated with Amazon S3 (p. 14), Amazon RDS (p. 17), and Amazon Redshift (p. 30), and can connect to any Java Database Connectivity (JDBC)-compliant data store. AWS Glue automatically crawls your data sources, identifies data formats, and then suggests schemas and transformations, so you don't have to spend time hand-coding data flows. You can then edit these transformations, if necessary, using the tools and technologies you already know, such as Python, Spark, Git, and your favorite integrated developer environment (IDE), and share them with other AWS Glue users. AWS Glue schedules your ETL jobs and provisions and scales all the infrastructure required so your ETL jobs run quickly and efficiently at any scale. There are no servers to manage, and you pay only for resources consumed by your ETL jobs.

Artificial Intelligence

Topics
- Amazon Lex (p. 32)

Amazon Lex

Amazon Lex is a service for building conversational interfaces into any application using voice and text. Lex provides the advanced deep learning functionalities of automatic speech recognition (ASR) for converting speech to text, and natural language understanding (NLU) to recognize the intent of the text, to enable you to build applications with highly engaging user experiences and lifelike conversational interactions. With Amazon Lex, the same deep learning technologies that power Amazon Alexa are now available to any developer, enabling you to quickly and easily build sophisticated, natural language, conversational bots ("chatbots").

Speech recognition and natural language understanding are some of the most challenging problems to solve in computer science, requiring sophisticated deep learning algorithms to be trained on massive amounts of data and infrastructure. Amazon Lex democratizes these deep learning technologies by putting the power of Alexa within reach of all developers. Harnessing these technologies, Amazon Lex enables you to define entirely new categories of products made possible through conversational interfaces.

Amazon Polly

Amazon Polly is a service that turns text into lifelike speech. Polly lets you create applications that talk, enabling you to build entirely new categories of speech-enabled products. Polly is an Amazon artificial intelligence (AI) service that uses advanced deep learning technologies to synthesize speech that sounds like a human voice. Polly includes 47 lifelike voices spread across 24 languages, so you can select the ideal voice and build speech-enabled applications that work in many different countries.

Amazon Polly delivers the consistently fast response times required to support real-time, interactive dialog. You can cache and save Polly's speech audio to replay offline or redistribute. And Polly is easy to use. You simply send the text you want converted into speech to the Polly API, and Polly immediately returns the audio stream to your application so your application can play it directly or store it in a standard audio file format, such as MP3.

With Polly, you only pay for the number of characters you convert to speech, and you can save and replay Polly's generated speech. Polly's low cost per character converted, and lack of restrictions on storage and reuse of voice output, make it a cost-effective way to enable Text-to-Speech everywhere.

Amazon Rekognition

Amazon Rekognition is a service that makes it easy to add image analysis to your applications. With Rekognition, you can detect objects, scenes, and faces in images. You can also search and compare faces. The Amazon Rekognition API enables you to quickly add sophisticated deep-learning-based visual search and image classification to your applications.

Amazon Rekognition is based on the same proven, highly scalable, deep learning technology developed by Amazon's computer vision scientists to analyze billions of images daily for Prime Photos. Amazon Rekognition uses deep neural network models to detect and label thousands of objects and scenes in your images, and we are continually adding new labels and facial recognition features to the service.

The Amazon Rekognition API lets you easily build powerful visual search and discovery into your applications. With Amazon Rekognition, you only pay for the images you analyze and the face metadata you store. There are no minimum fees, and there are no upfront commitments.

Amazon Machine Learning

Amazon Machine Learning (Amazon ML) is a service that makes it easy for developers of all skill levels to use machine learning technology. Amazon Machine Learning provides visualization tools and wizards that guide you through the process of creating machine learning models without having to learn complex ML algorithms and technology. Once your models are ready, Amazon Machine Learning makes it easy to obtain predictions for your application using simple APIs, without having to implement custom prediction generation code or manage any infrastructure.

Amazon Machine Learning is based on the same proven, highly scalable, ML technology used for years by Amazon's internal data scientist community. The service uses powerful algorithms to create ML models by finding patterns in your existing data. Then, Amazon Machine Learning uses these models to process new data and generate predictions for your application.

Amazon Machine Learning is highly scalable and can generate billions of predictions daily, and serve those predictions in real time and at high throughput.

Mobile Services

Topics
- AWS Mobile Hub (p. 33)
- Amazon Cognito (p. 34)
- Amazon Pinpoint (p. 34)
- AWS Device Farm (p. 34)
- AWS Mobile SDK (p. 34)
- Amazon Mobile Analytics (p. 35)

AWS Mobile Hub

AWS Mobile Hub provides an integrated console experience that you can use to quickly create and configure powerful mobile app backend features and integrate them into your mobile app. You create a project by selecting features to add to your app.

The features and AWS services that are supported by Mobile Hub are constantly evolving. Currently they include:

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

When you build your project for iOS Objective-C, iOS Swift, or Android, Mobile Hub automatically provisions and configures all of the AWS service resources that your app's features require. Mobile Hub

then guides you through integrating the features into your app code and downloading a fully working quickstart app project that demonstrates those features.

After your mobile app is built, you can use Mobile Hub to test your app, then monitor and visualize how it is being used.

Amazon Cognito

Amazon Cognito lets you easily add user sign-up and sign-in to your mobile and web apps. With Amazon Cognito, you also have the option to authenticate users through social identity providers such as Facebook, Twitter, or Amazon, with SAML identity solutions, or by using your own identity system. In addition, Amazon Cognito enables you to save data locally on users' devices, allowing your applications to work even when the devices are offline. You can then synchronize data across users' devices so that their app experience remains consistent regardless of the device they use.

With Amazon Cognito, you can focus on creating great app experiences instead of worrying about building, securing, and scaling a solution to handle user management, authentication, and sync across devices.

Amazon Pinpoint

Amazon Pinpoint makes it easy to run targeted campaigns to drive user engagement in mobile apps. Amazon Pinpoint helps you understand user behavior, define which users to target, determine which messages to send, schedule the best time to deliver the messages, and then track the results of your campaign.

Targeted push notifications based on app usage trends and user behavior have become a popular approach for mobile app user engagement because response rates are often several times higher than traditional email marketing campaigns. By using targeted push notifications, you can increase message relevance and effectiveness, measure engagement, and continually improve your campaigns.

Getting started with Amazon Pinpoint is easy. First, AWS Mobile Hub guides you through the process to integrate the AWS Mobile SDK with your app. Next, you define your target segments, campaign message, and specify the delivery schedule. Once your campaign is running, Pinpoint provides metrics so you can run analytics and track the impact of your campaign.

With Amazon Pinpoint, there are no upfront setup costs, and no fixed monthly cost. You only pay for the number of users your campaign targets, the messages you send, and the events you collect, so you can start small and scale as your application grows.

AWS Device Farm

AWS Device Farm is an app testing service that lets you test and interact with your Android, iOS, and web apps on many devices at once, or reproduce issues on a device in real time. View video, screenshots, logs, and performance data to pinpoint and fix issues before shipping your app.

AWS Mobile SDK

The AWS Mobile SDK helps you build high quality mobile apps quickly and easily. It provides easy access to a range of AWS services, including AWS Lambda (p. 13), Amazon S3 (p. 14), Amazon DynamoDB (p. 17), Amazon Mobile Analytics (p. 35), Amazon Machine Learning (p. 33), Elastic Load Balancing (p. 22), Auto Scaling (p. 13), and more.

The AWS Mobile SDK includes libraries, code samples, and documentation for iOS, Android, Fire OS, and Unity so you can build apps that deliver great experiences across devices and platforms.

Amazon Mobile Analytics

With Amazon Mobile Analytics, you can measure app usage and app revenue. By tracking key trends such as new versus returning users, app revenue, user retention, and custom in-app behavior events, you can make data-driven decisions to increase engagement and monetization for your app. You can view key charts in the Mobile Analytics console and automatically export your app event data to Amazon S3 and Amazon Redshift to run custom analysis.

Application Services

Topics
- AWS Step Functions (p. 35)
- Amazon API Gateway (p. 35)
- Amazon Elastic Transcoder (p. 35)
- Amazon SWF (p. 35)

AWS Step Functions

AWS Step Functions makes it easy to coordinate the components of distributed applications and microservices using visual workflows. Building applications from individual components that each perform a discrete function lets you scale and change applications quickly. Step Functions is a reliable way to coordinate components and step through the functions of your application. Step Functions provides a graphical console to arrange and visualize the components of your application as a series of steps. This makes it simple to build and run multi-step applications. Step Functions automatically triggers and tracks each step, and retries when there are errors, so your application executes in order and as expected. Step Functions logs the state of each step, so when things do go wrong, you can diagnose and debug problems quickly. You can change and add steps without even writing code, so you can easily evolve your application and innovate faster. AWS Step Functions manages the operations and underlying infrastructure for you to help ensure your application is available at any scale.

Amazon API Gateway

Amazon API Gateway is a fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale. With a few clicks in the AWS Management Console, you can create an API that acts as a "front door" for applications to access data, business logic, or functionality from your back-end services, such as workloads running on Amazon EC2, code running on AWS Lambda, or any web application. Amazon API Gateway handles all the tasks involved in accepting and processing up to hundreds of thousands of concurrent API calls, including traffic management, authorization and access control, monitoring, and API version management.

Amazon Elastic Transcoder

Amazon Elastic Transcoder is media transcoding in the cloud. It is designed to be a highly scalable, easy-to-use, and cost-effective way for developers and businesses to convert (or transcode) media files from their source format into versions that will play back on devices like smartphones, tablets, and PCs.

Amazon SWF

Amazon Simple Workflow (Amazon SWF) helps developers build, run, and scale background jobs that have parallel or sequential steps. You can think of Amazon SWF as a fully-managed state tracker and

task coordinator in the cloud. If your application's steps take more than 500 milliseconds to complete, you need to track the state of processing. If you need to recover or retry if a task fails, Amazon SWF can help you.

Messaging

Topics
- Amazon SQS (p. 36)
- Amazon SNS (p. 36)
- Amazon SES (p. 36)

Amazon SQS

Amazon Simple Queue Service (Amazon SQS) is a fast, reliable, scalable, fully managed message queuing service. Amazon SQS makes it simple and cost-effective to decouple the components of a cloud application. You can use Amazon SQS to transmit any volume of data, without losing messages or requiring other services to be always available. Amazon SQS includes standard queues with high throughput and at-least-once processing, and FIFO queues that provide FIFO (first-in, first-out) delivery and exactly-once processing.

Amazon SNS

Amazon Simple Notification Service (Amazon SNS) is a fast, flexible, fully managed push notification service that lets you send individual messages or to fan-out messages to large numbers of recipients. Amazon SNS makes it simple and cost effective to send push notifications to mobile device users, email recipients or even send messages to other distributed services.

With Amazon SNS, you can send notifications to Apple, Google, Fire OS, and Windows devices, as well as to Android devices in China with Baidu Cloud Push. You can use Amazon SNS to send SMS messages to mobile device users worldwide.

Beyond these endpoints, Amazon SNS can also deliver messages to Amazon Simple Queue Service (SQS), AWS Lambda functions, or to any HTTP endpoint.

Amazon SES

Amazon Simple Email Service (Amazon SES) is a cost-effective email service built on the reliable and scalable infrastructure that Amazon.com developed to serve its own customer base. With Amazon SES, you can send transactional email, marketing messages, or any other type of high-quality content to your customers. You can also use Amazon SES to receive messages and deliver them to an Amazon S3 bucket, call your custom code via an AWS Lambda function, or publish notifications to Amazon SNS. With Amazon SES, you have no required minimum commitments—you pay as you go, and you only pay for what you use.

See also Amazon Pinpoint (p. 34).

Business Productivity

Topics
- Amazon WorkDocs (p. 37)

- Amazon WorkMail (p. 37)
- Amazon Chime (p. 37)

Amazon WorkDocs

Amazon WorkDocs is a fully managed, secure enterprise storage and sharing service with strong administrative controls and feedback capabilities that improve user productivity.

Users can comment on files, send them to others for feedback, and upload new versions without having to resort to emailing multiple versions of their files as attachments. Users can take advantage of these capabilities wherever they are, using the device of their choice, including PCs, Macs, tablets, and phones. Amazon WorkDocs offers IT administrators the option of integrating with existing corporate directories, flexible sharing policies and control of the location where data is stored. You can get started using Amazon WorkDocs with a 30-day free trial providing 1 TB of storage per user for up to 50 users.

Amazon WorkMail

Amazon WorkMail is a secure, managed business email and calendar service with support for existing desktop and mobile email client applications. Amazon WorkMail gives users the ability to seamlessly access their email, contacts, and calendars using the client application of their choice, including Microsoft Outlook, native iOS and Android email applications, any client application supporting the IMAP protocol, or directly through a web browser. You can integrate Amazon WorkMail with your existing corporate directory, use email journaling to meet compliance requirements, and control both the keys that encrypt your data and the location in which your data is stored. You can also set up interoperability with Microsoft Exchange Server, making it easy to start using Amazon WorkMail.

Amazon Chime

Amazon Chime is a communications service that transforms online meetings with a secure, easy-to-use application that you can trust. Amazon Chime works seamlessly across your devices so that you can stay connected. You can use Amazon Chime for online meetings, video conferencing, calls, chat, and to share content, both inside and outside your organization. Amazon Chime frees you to work productively from anywhere.

Desktop & App Streaming

Topics
- Amazon WorkSpaces (p. 37)
- Amazon AppStream 2.0 (p. 38)

Amazon WorkSpaces

Amazon WorkSpaces is a fully managed, secure desktop computing service that runs on the AWS Cloud. Amazon WorkSpaces allows you to easily provision cloud-based virtual desktops and provide your users access to the documents, applications, and resources they need from any supported device, including Windows and Mac computers, Chromebooks, iPads, Fire tablets, Android tablets, and Chrome and Firefox web browsers. With just a few clicks in the AWS Management Console, you can deploy high-quality cloud desktops for any number of users. With Amazon WorkSpaces, you pay either monthly or hourly just for the Amazon WorkSpaces you launch, which helps you save money when compared to traditional desktops and on-premises virtual desktop infrastructure (VDI) solutions.

Amazon AppStream 2.0

Amazon AppStream 2.0 is a fully managed, secure application streaming service that allows you to stream desktop applications from AWS to any device running a web browser, without rewriting them. Amazon AppStream 2.0 provides you instant-on access to the applications you need, and a responsive, fluid user experience on the device of your choice.

Users today want access to their applications from anywhere, using their preferred device. To support this, organizations need to maintain multiple versions of their desktop applications, and take additional measures to secure applications and data. While browser-based applications solve device compatibility and security problems for IT, the reality is that organizations still have many traditional desktop applications that need to be supported. Organizations face a choice: either rewrite their applications to run natively in browsers, which is time consuming and expensive, or continue to maintain and support a complex catalog of desktop applications.

Amazon AppStream 2.0 gives you the benefits of native browser applications without the need to rewrite them. With Amazon AppStream 2.0, you can easily import your existing desktop applications to AWS and instantly start streaming them to an HTML5-compatible browser. You can maintain a single version of each of your apps, which makes application management easier. Your users always access the latest versions of their applications. Your applications run on AWS compute resources, and data is never stored on users' devices, which means they always get a high-performance, secure experience.

Unlike traditional on-premises solutions for desktop application streaming, Amazon AppStream 2.0 offers pay-as-you-go pricing, with no upfront investment and no infrastructure to maintain. You can scale instantly and globally, ensuring that your users always have the best possible experience.

Internet of Things (IoT)

Topics

AWS IoT Platform

AWS IoT is a managed cloud platform that lets connected devices easily and securely interact with cloud applications and other devices. AWS IoT can support billions of devices and trillions of messages, and can process and route those messages to AWS endpoints and to other devices reliably and securely. With AWS IoT, your applications can keep track of and communicate with all your devices, all the time, even when they aren't connected.

AWS IoT makes it easy to use AWS services like AWS Lambda (p. 13), Amazon Kinesis (p. 29), Amazon S3 (p. 14), Amazon Machine Learning (p. 33), and Amazon DynamoDB (p. 17) to build Internet of Things (IoT) applications that gather, process, analyze and act on data generated by connected devices, without having to manage any infrastructure.

AWS Greengrass

AWS Greengrass is software that lets you run local compute, messaging, and data caching for connected devices in a secure way. With AWS Greengrass, connected devices can run AWS Lambda functions, keep device data in sync, and communicate with other devices securely – even when not connected to the Internet. Using AWS Lambda, Greengrass ensures your IoT devices can respond quickly to local events, operate with intermittent connections, and minimize the cost of transmitting IoT data to the cloud.

AWS Greengrass seamlessly extends AWS to devices so they can act locally on the data they generate, while still using the cloud for management, analytics, and durable storage. With AWS Greengrass, you can use familiar languages and programming models to create and test your device software in the cloud, and then deploy it to your devices. AWS Greengrass can be programmed to filter device data and only transmit necessary information back to the cloud. AWS Greengrass authenticates and encrypts device data at all points of connection using the AWS IoT security and access management capabilities. This way data is never exchanged between devices when they communicate with each other and the cloud without proven identity.

AWS IoT Button

The AWS IoT Button is a programmable button based on the Amazon Dash Button hardware. This simple Wi-Fi device is easy to configure, and it's designed for developers to get started with AWS IoT (p. 38), AWS Lambda (p. 13), Amazon DynamoDB (p. 17), Amazon SNS (p. 36), and many other Amazon Web Services without writing device-specific code.

You can code the button's logic in the cloud to configure button clicks to count or track items, call or alert someone, start or stop something, order services, or even provide feedback. For example, you can click the button to unlock or start a car, open your garage door, call a cab, call your spouse or a customer service representative, track the use of common household chores, medications or products, or remotely control your home appliances.

The button can be used as a remote control for Netflix, a switch for your Philips Hue light bulb, a check-in/check-out device for Airbnb guests, or a way to order your favorite pizza for delivery. You can integrate it with third-party APIs like Twitter, Facebook, Twilio, Slack or even your own company's applications. Connect it to things we haven't even thought of yet.

Game Development

Topics
- Amazon GameLift (p. 39)
- Amazon Lumberyard (p. 39)

Amazon GameLift

Amazon GameLift is a managed service for deploying, operating, and scaling dedicated game servers for session-based multiplayer games. Amazon GameLift makes it easy to manage server infrastructure, scale capacity to lower latency and cost, match players into available game sessions, and defend from distributed denial-of-service (DDoS) attacks. You pay for the compute resources and bandwidth your games actually use, without monthly or annual contracts.

Amazon Lumberyard

Amazon Lumberyard is a free, cross-platform, 3D game engine for you to create the highest-quality games, connect your games to the vast compute and storage of the AWS Cloud, and engage fans on Twitch. By starting game projects with Lumberyard, you can spend more of your time creating great gameplay and building communities of fans, and less time on the undifferentiated heavy lifting of building a game engine and managing server infrastructure.

Next Steps

Reinvent how you work with IT by signing up for the AWS Free Tier, which enables you to gain hands-on experience with a broad selection of AWS products and services. Within the AWS Free Tier, you can test workloads and run applications to learn more and build the right solution for your organization. You can also contact AWS Sales and Business Development.

By signing up for AWS, you have access to Amazon's cloud computing services. Note: The sign-up process requires a credit card, which will not be charged until you start using services. There are no long-term commitments and you can stop using AWS at any time.

To help familiarize you with AWS, view these short videos that cover topics like creating an account, launching a virtual server, storing media and more. Learn about the breadth and depth of AWS on our general AWS Channel and Webinar Channel. Get hands on experience from our self-paced labs.

Conclusion

AWS provides building blocks that you can assemble quickly to support virtually any workload. With AWS, you'll find a complete set of highly available services that are designed to work together to build sophisticated scalable applications.

You have access to highly durable storage, low-cost compute, high-performance databases, management tools, and more. All this is available without up-front cost, and you pay for only what you use. These services help organizations move faster, lower IT costs, and scale. AWS is trusted by the largest enterprises and the 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.

Resources

- AWS Architecture Center
- AWS Whitepapers
- AWS Architecture Monthly
- AWS Architecture Blog
- This Is My Architecture videos
- AWS Answers
- AWS Documentation

Document Details

Contributors

The following individuals and organizations contributed to this document:

- Sajee Mathew, AWS Principal Solutions Architect

Document History

Date	Description
April 2017	Added new services and updated information throughout.
January 2014	First publication

AWS Glossary

For the latest AWS terminology, see the AWS Glossary in the *AWS General Reference*.