AWS Solutions Architect—Associate Level

Lesson 2: Designing Highly Available, Cost-efficient, Fault-tolerant Scalable Systems

simplilearn

What You'll Learn

Cloud versus Traditional Infrastructure

AWS Well-Architected Framework

Planning and Designing Cloud Infrastructure
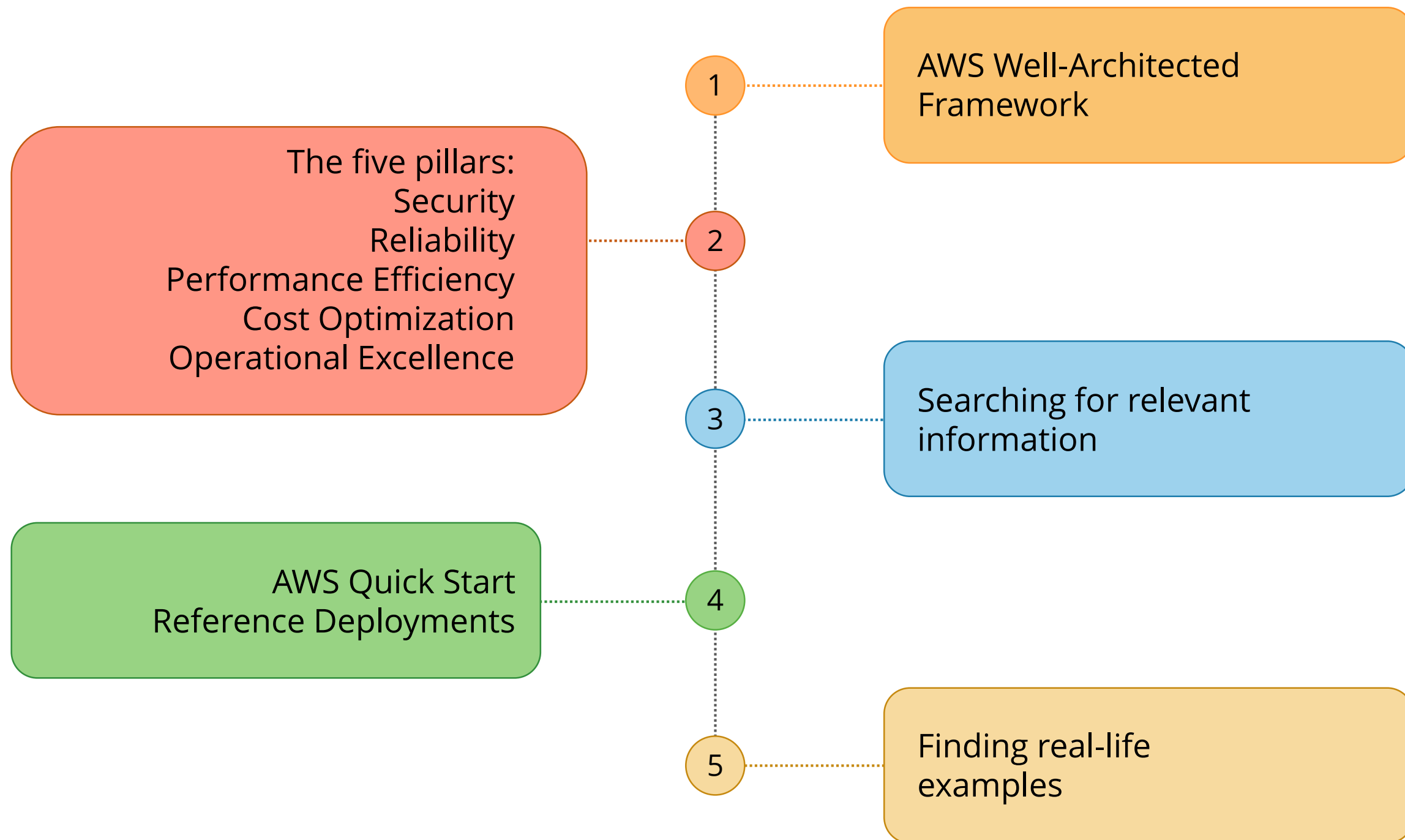
AWS Monitoring and Logging Tools

Hybrid Cloud Infrastructure

2

How to Design Cloud Services

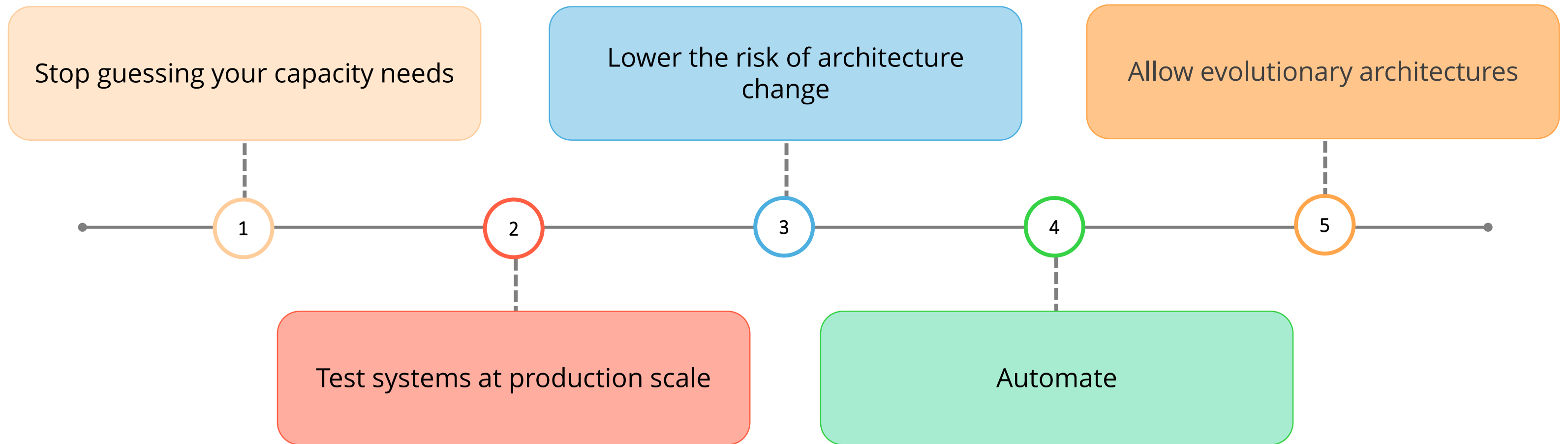Overview of the AWS cloud design principles

Designing Cloud Services

In this section you'll learn about

1 — AWS Well-Architected Framework

The five pillars:
Security
Reliability
Performance Efficiency
Cost Optimization
Operational Excellence — **2**

3 — Searching for relevant information

AWS Quick Start
Reference Deployments — **4**

5 — Finding real-life examples

4

simpli·learn

AWS Well-Architected Framework

Cloud infrastructure is different from traditional on-premise infrastructure.

The five principles of AWS Well-Architected Framework are the following:

Stop guessing your capacity needs

Lower the risk of architecture change

Allow evolutionary architectures

1

2

3

4

5

Test systems at production scale

Automate

The AWS framework helps you take an informed decision.

5

Stop Guessing Your Capacity Needs

AWS helps you eliminate the guesswork in your infrastructure capacity needs. You can use as much or as little capacity as you need and automatically scale up and down as required.

Database

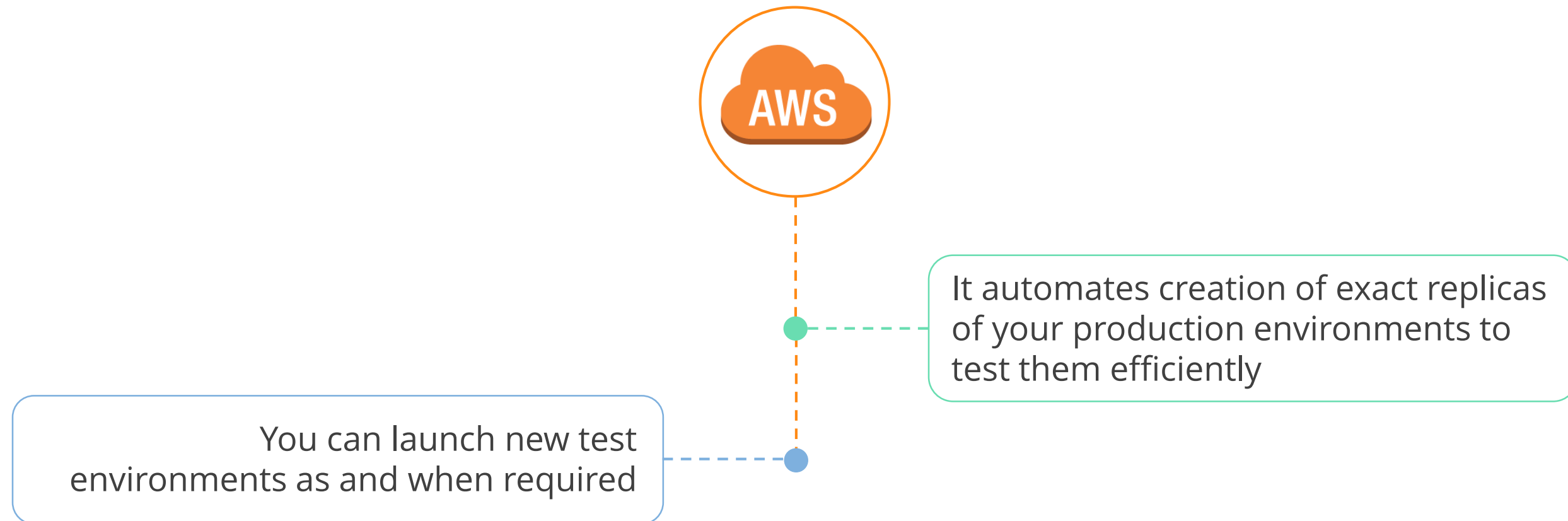Networking

Expenditure

Performance

6
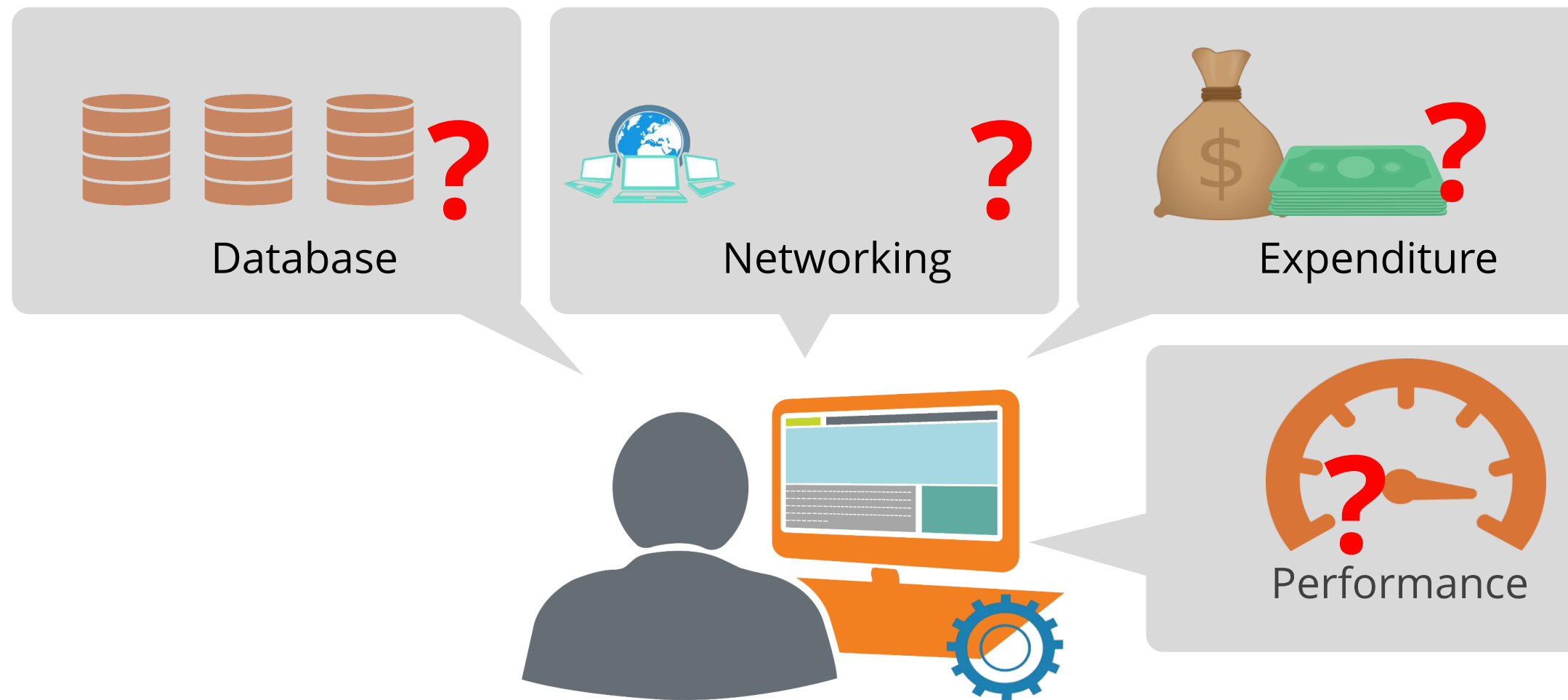
Test Systems at Production Scale

In traditional environments it is difficult to test new products due to high cost or unavailability of resources.

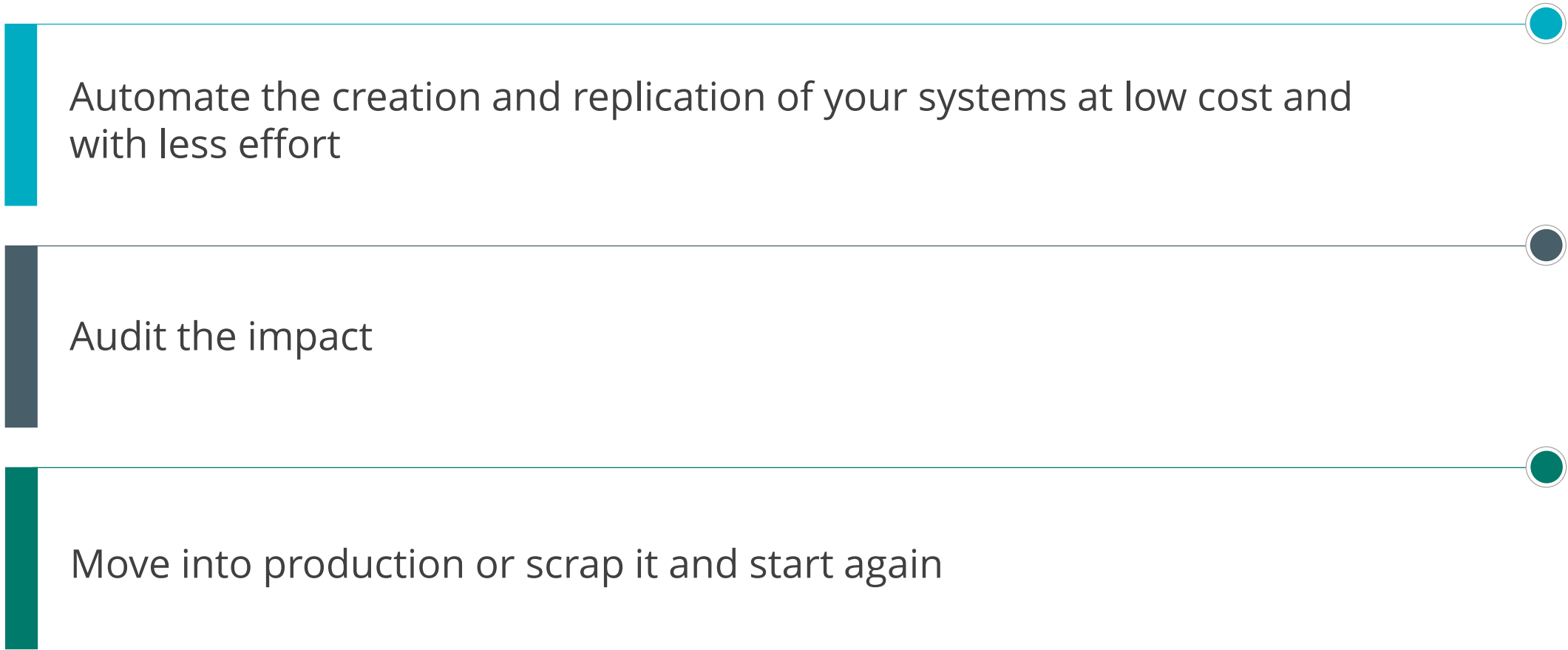AWS Cloud allows you to create duplicate environments when you require them.

Lower the Risk of Architecture Change

AWS lowers the risk of architecture change because

AWS

It automates creation of exact replicas of your production environments to test them efficiently

You can launch new test environments as and when required

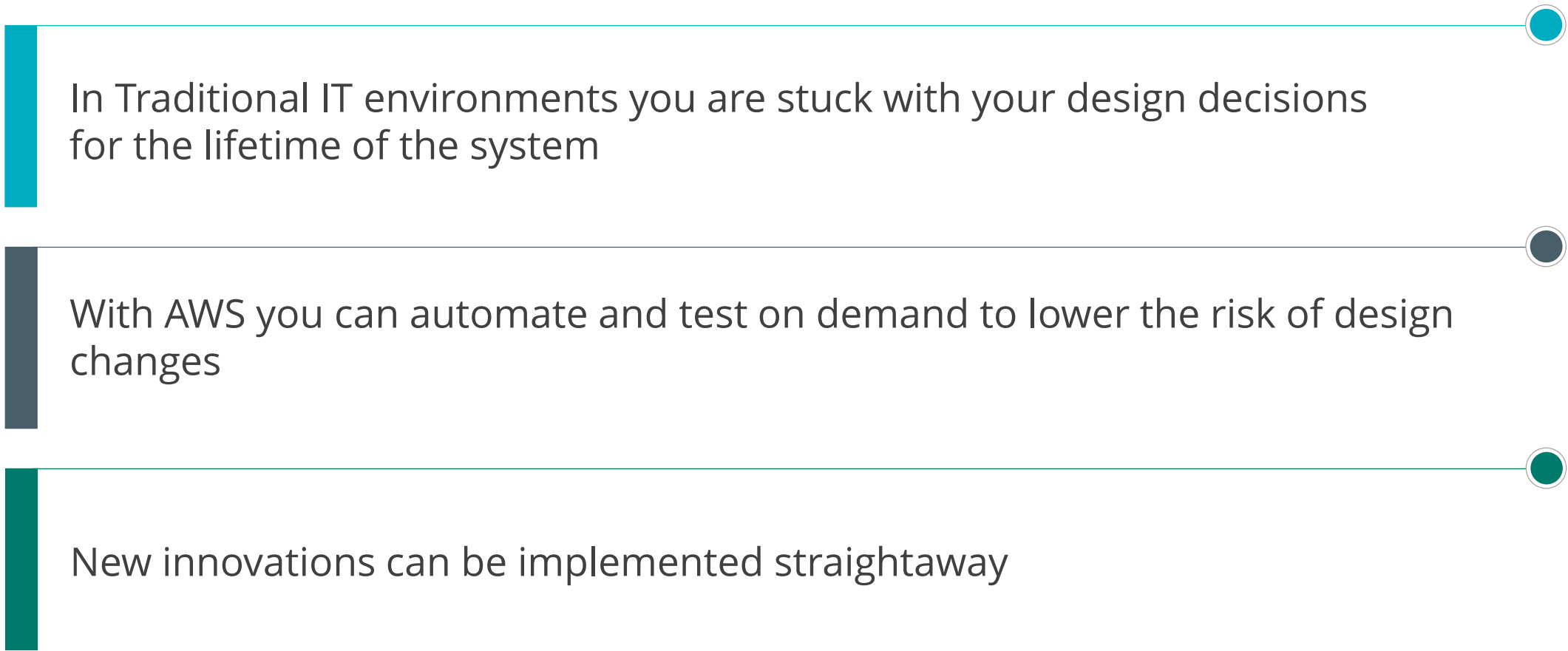simplilearn

Automation

AWS allows you to

Automate the creation and replication of your systems at low cost and with less effort

Audit the impact

Move into production or scrap it and start again

9

Evolutionary Architectures

In Traditional IT environments you are stuck with your design decisions for the lifetime of the system

With AWS you can automate and test on demand to lower the risk of design changes

New innovations can be implemented straightaway

Five Pillars

The AWS Well-Architected Framework is based on five pillars:

Security	Reliability	Performance Efficiency	Cost Optimization	Operational Excellence

Security

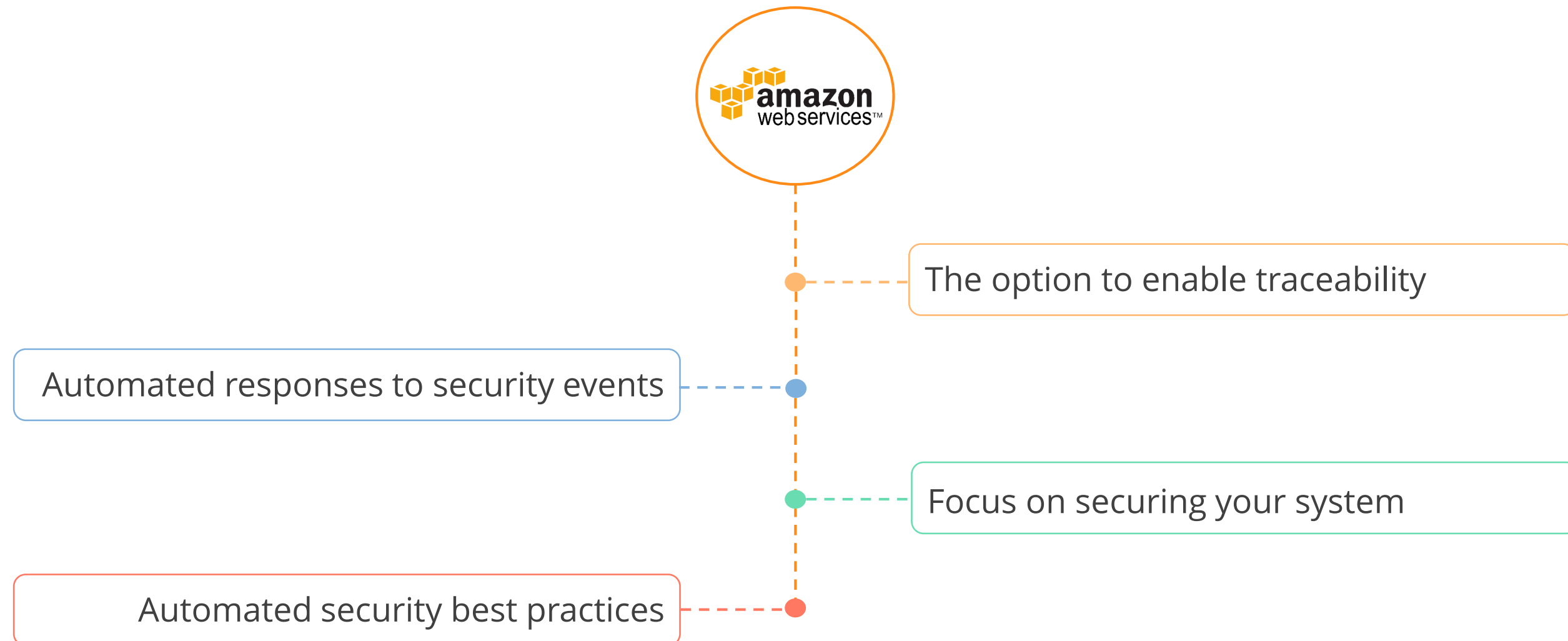| Security | Reliability | Performance Efficiency | Cost Optimization | Operational Excellence |

Amazon defines Security as, "The ability to protect information, systems, and assets while delivering business value through risk assessments and mitigation strategies."
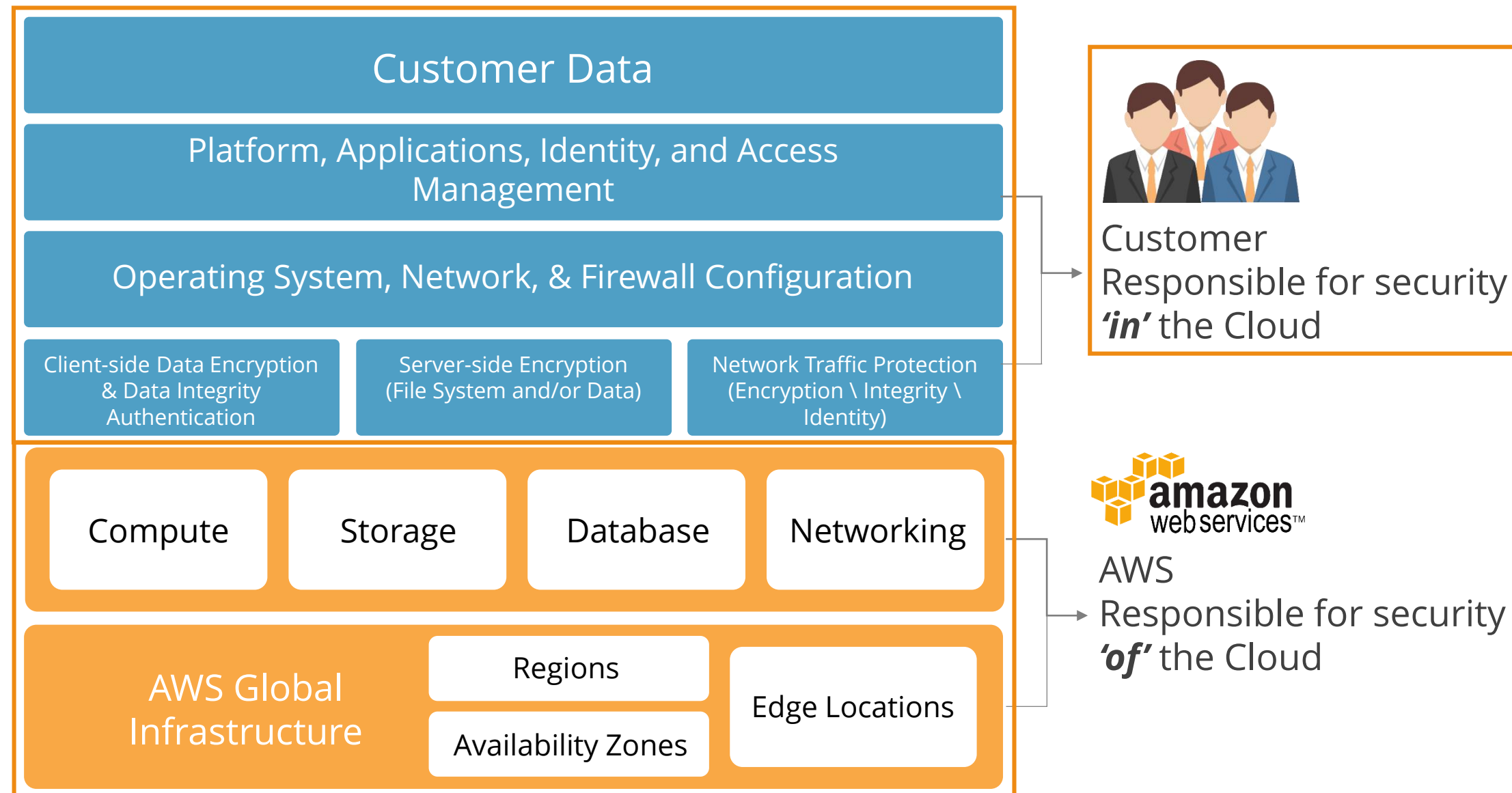
simpli·learn

Security

AWS provides numerous security options, such as:

The option to enable traceability

Automated responses to security events

Focus on securing your system

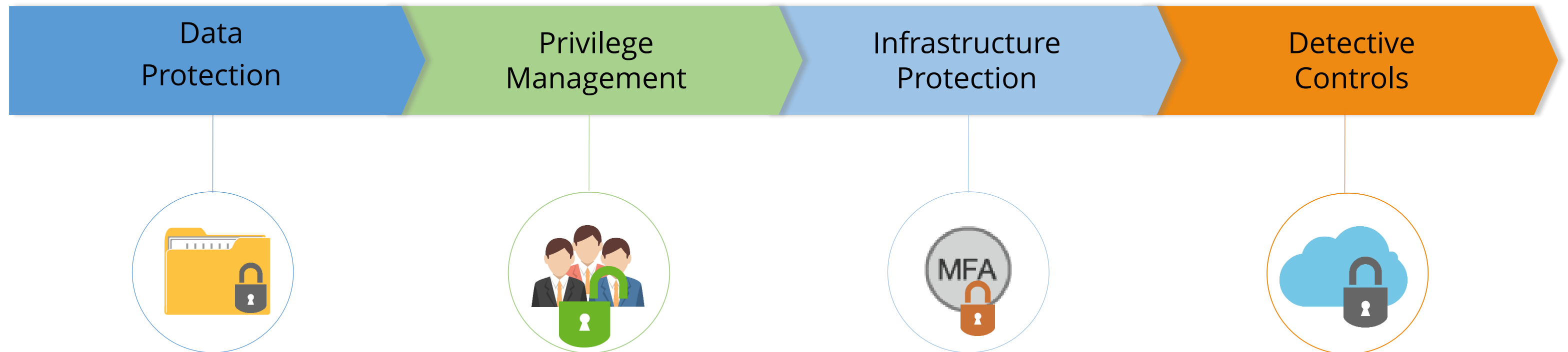Automated security best practices

AWS Shared Responsibility Model

The AWS shared responsibility model is divided into two sections—Security *'in'* the Cloud and Security *'of'* the cloud.
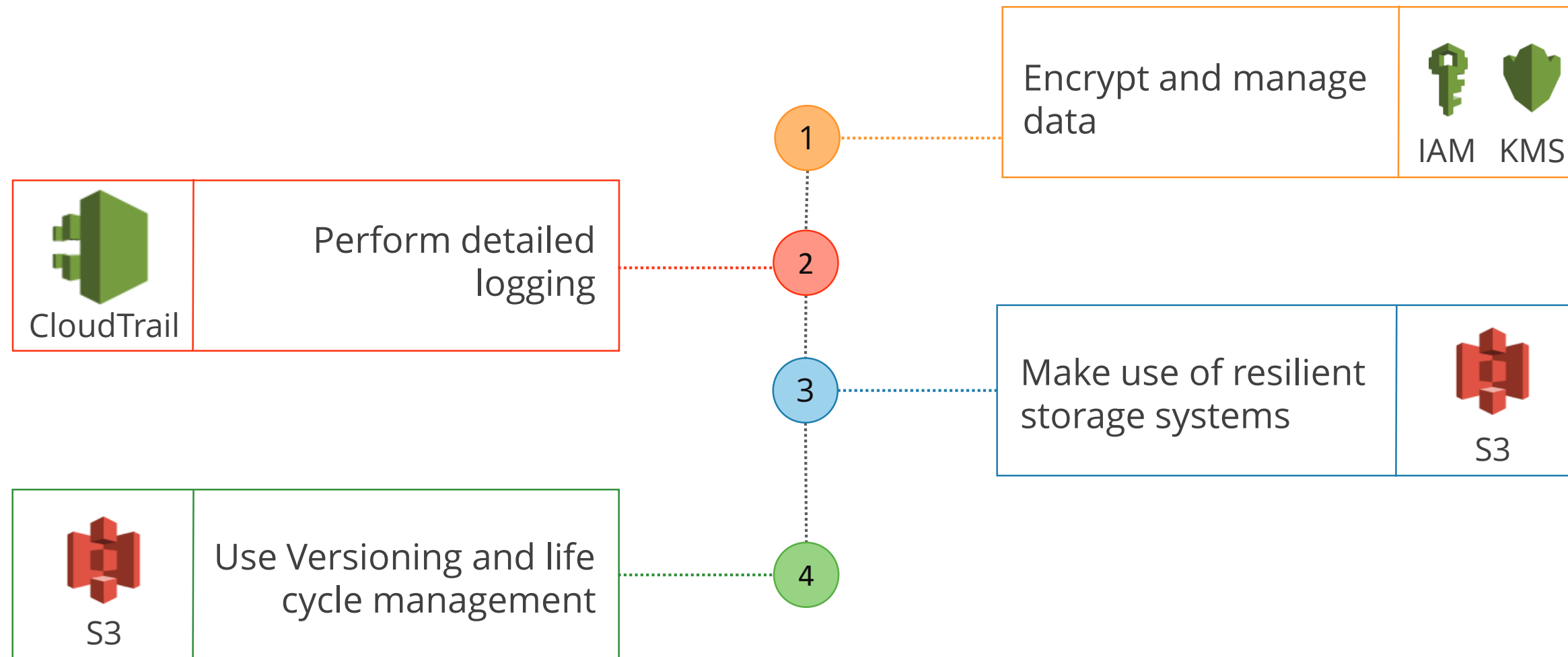
Customer Data

Platform, Applications, Identity, and Access Management

Operating System, Network, & Firewall Configuration

| Client-side Data Encryption & Data Integrity Authentication | Server-side Encryption (File System and/or Data) | Network Traffic Protection (Encryption \ Integrity \ Identity) |

Customer
Responsible for security *'in'* the Cloud

| Compute | Storage | Database | Networking |

AWS Global Infrastructure

| Regions |
| Availability Zones |

| Edge Locations |

amazon web services™

AWS
Responsible for security *'of'* the Cloud

14

simpli learn

Security in the Cloud

Security in the cloud is composed of four areas:

| Data Protection | Privilege Management | Infrastructure Protection | Detective Controls |

15

Data Protection

Before architecting any system use the following practices to ensure security of the AWS account:

1 — Encrypt and manage data
IAM   KMS

2 — Perform detailed logging
CloudTrail

3 — Make use of resilient storage systems
S3

4 — Use Versioning and life cycle management
S3

AWS S3 has 11 9s of durability, that is, if you store 10,000 objects with Amazon S3, you can on an average expect to incur a loss of a single object once every 10,000,000 years.
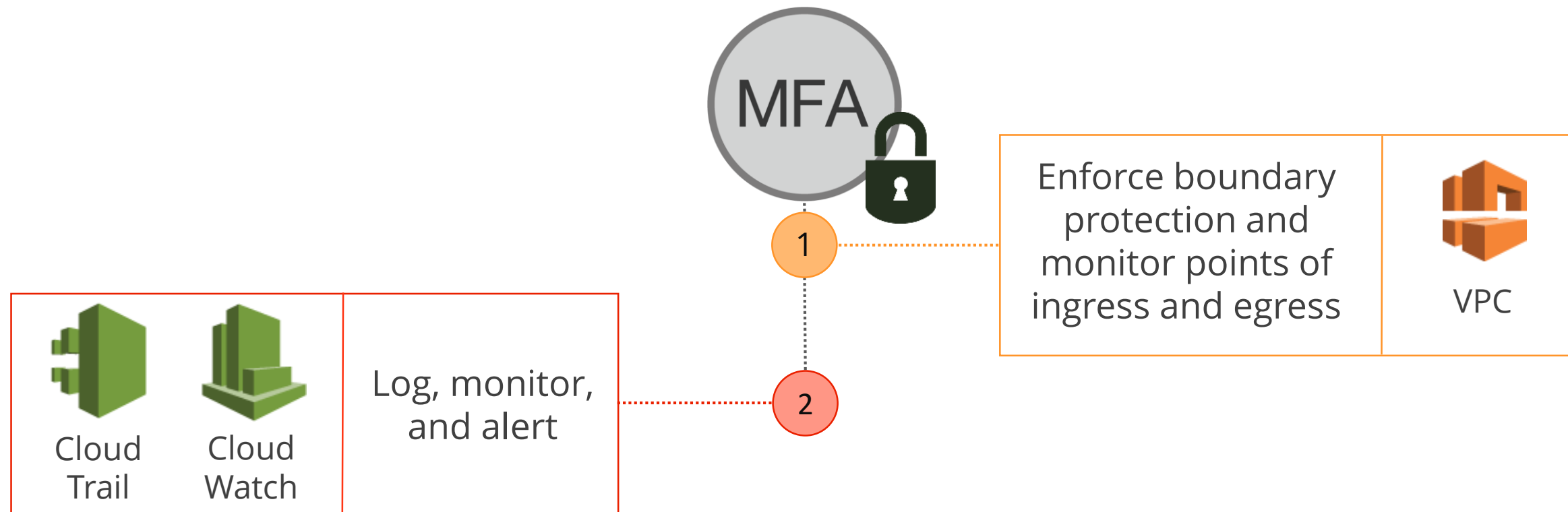
simpli·learn

Privilege Management

A central part of an information security program is to ensure only authorized and authenticated users access your resources in a way that is acceptable. To ensure compliance use
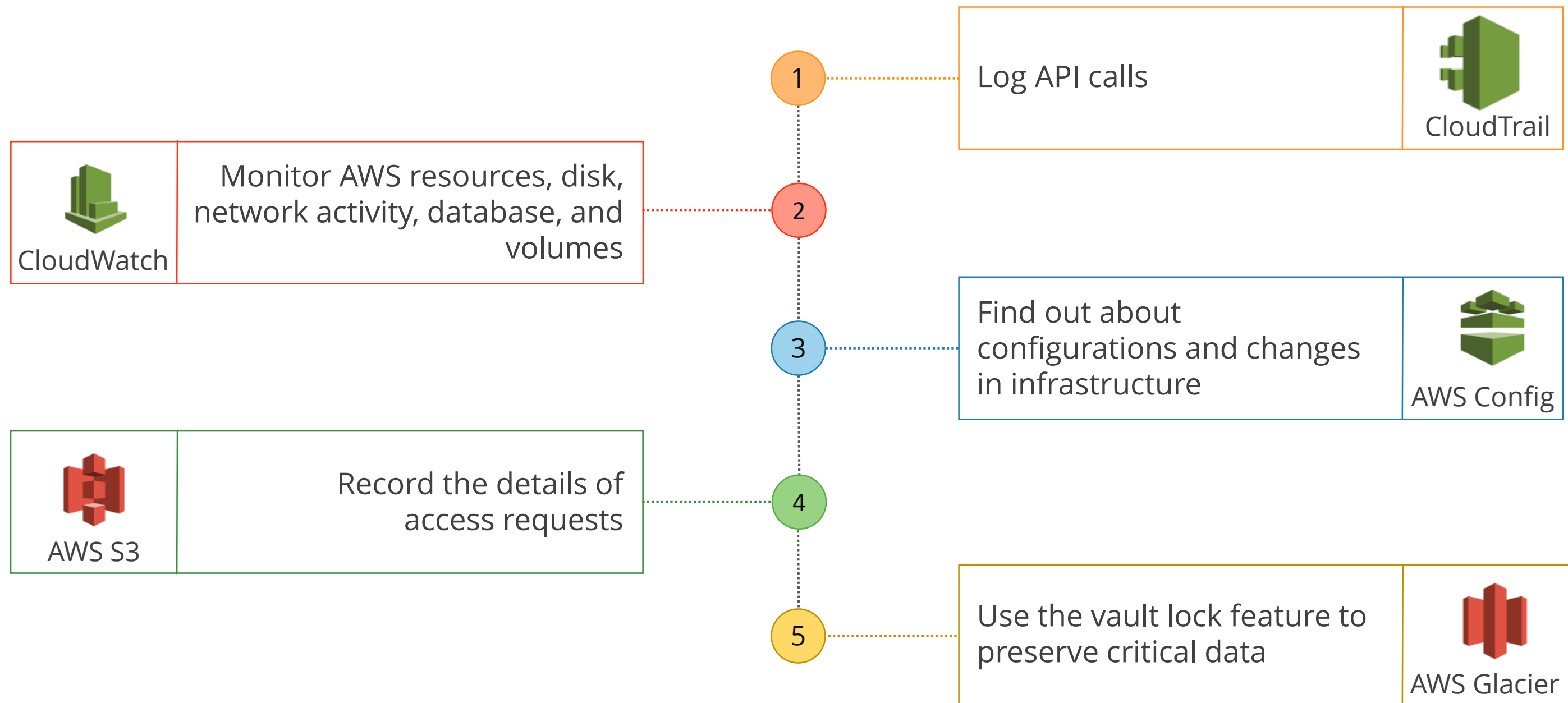
Access Control List (ACL)

Password Management

1

2

3

Role-Based Access Controls (RBAC)

17
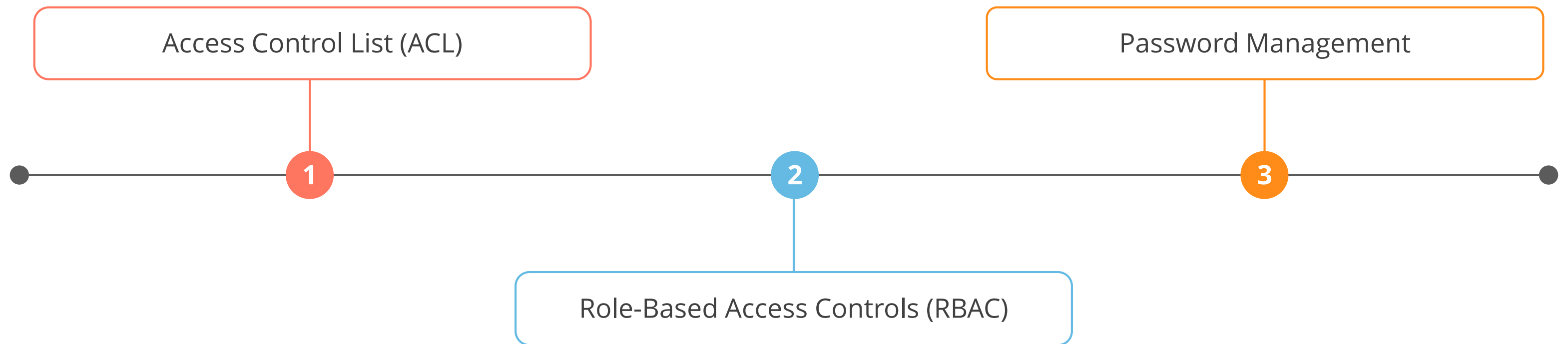
Infrastructure Protection

Use multiple layers of defence and multi-factor authentication in all types of environments. AWS implements "stateful" and "stateless" packet inspection by using AWS native technologies or partner products and services available through the AWS Marketplace.

MFA

1 Enforce boundary protection and monitor points of ingress and egress — VPC

2 Log, monitor, and alert — Cloud Trail, Cloud Watch

Detective Controls

To detect or identify a security breach for both a quality support process and a compliance obligation use "Detective Controls", such as

1 — Log API calls — CloudTrail

2 — Monitor AWS resources, disk, network activity, database, and volumes — CloudWatch

3 — Find out about configurations and changes in infrastructure — AWS Config

4 — Record the details of access requests — AWS S3

5 — Use the vault lock feature to preserve critical data — AWS Glacier

19

Reliability
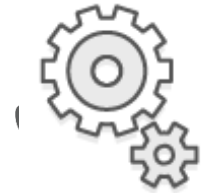
Security	Reliability	Performance Efficiency	Cost Optimization	Operational Excellence

Amazon defines Reliability as, "The ability of a system to recover from infrastructure or service failures, dynamically acquire computing resources to meet demand, and mitigate disruptions such as misconfigurations or transient network issues."

simpli·learn

Reliability in the Cloud

Reliability in the Cloud allows you to perform the following:

AWS

Test recovery procedures

Automatically recover from failure

Scale horizontally to increase aggregate system availability

Stop guessing capacity

simplilearn
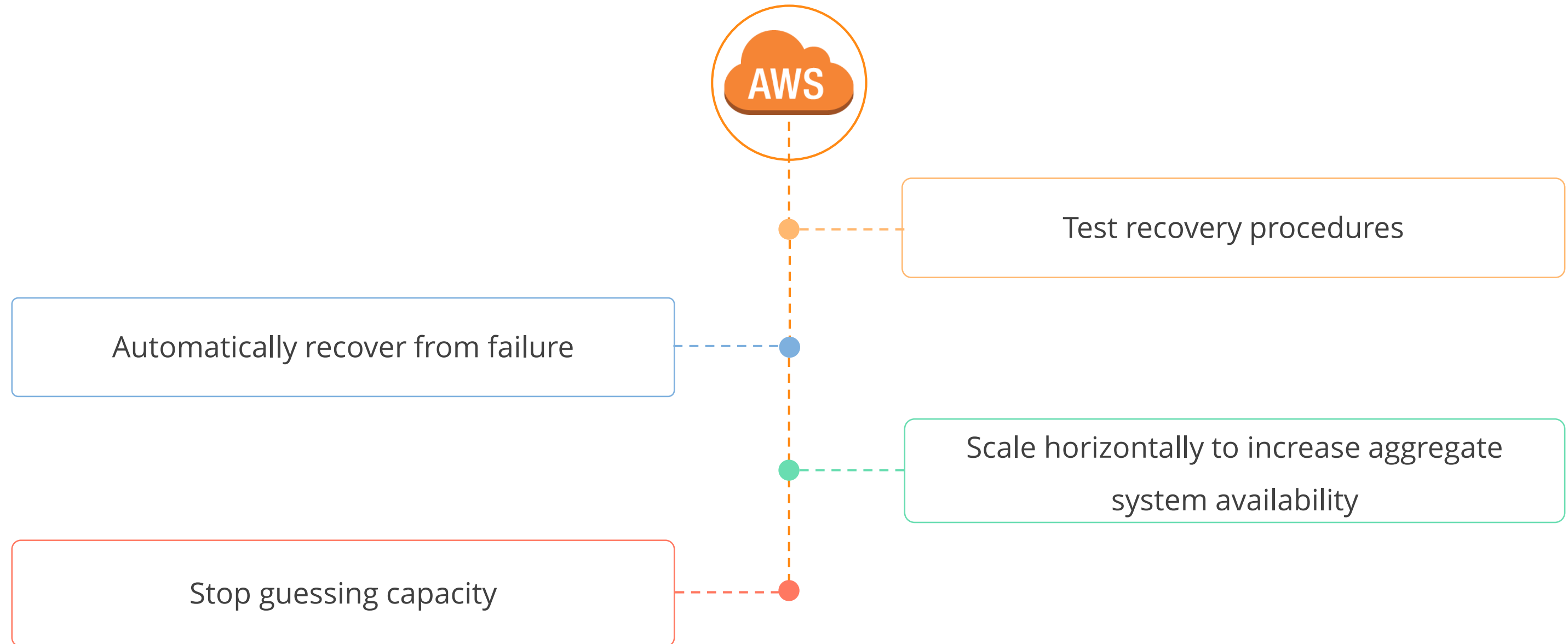
Reliability in the Cloud (contd.)

Reliability in the cloud is composed of three areas:

Foundations	Change Management	Failure Management

Foundations

Before architecting any system, foundations that impact reliability should be in place such as:

| Sufficient Bandwidth | Sufficient Compute Capacity | Trained Staff | Support Contracts |
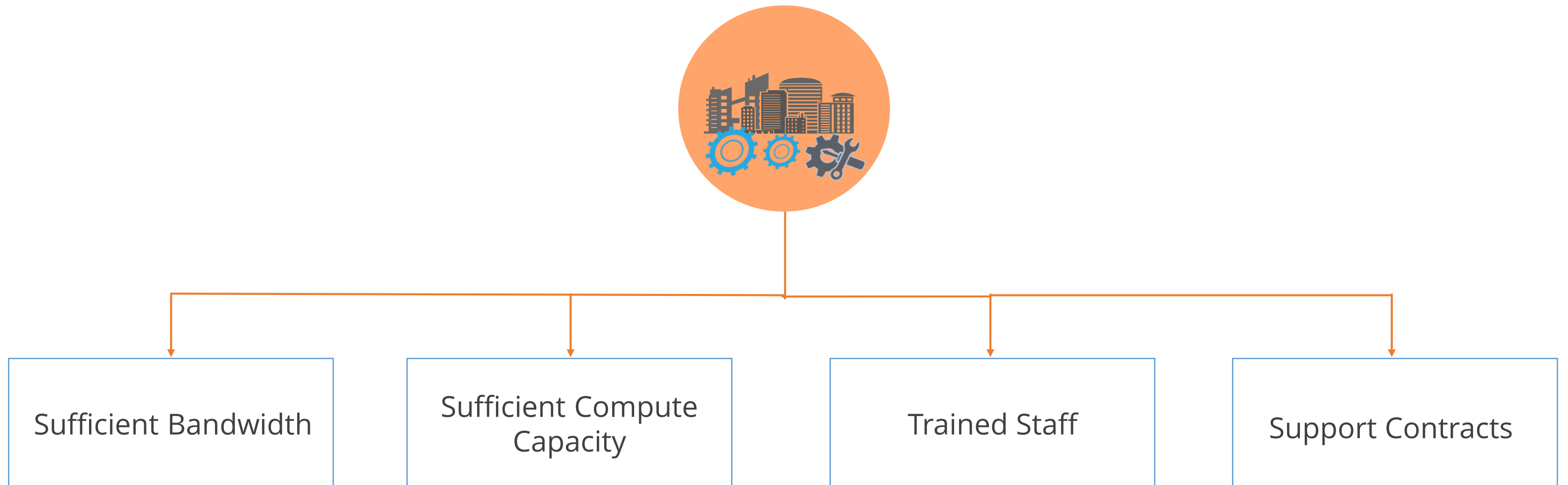
Change Management

Being aware of how change affects a system allows you to:

| Plan Proactively | Monitor | Automate responses to KPIs |

24
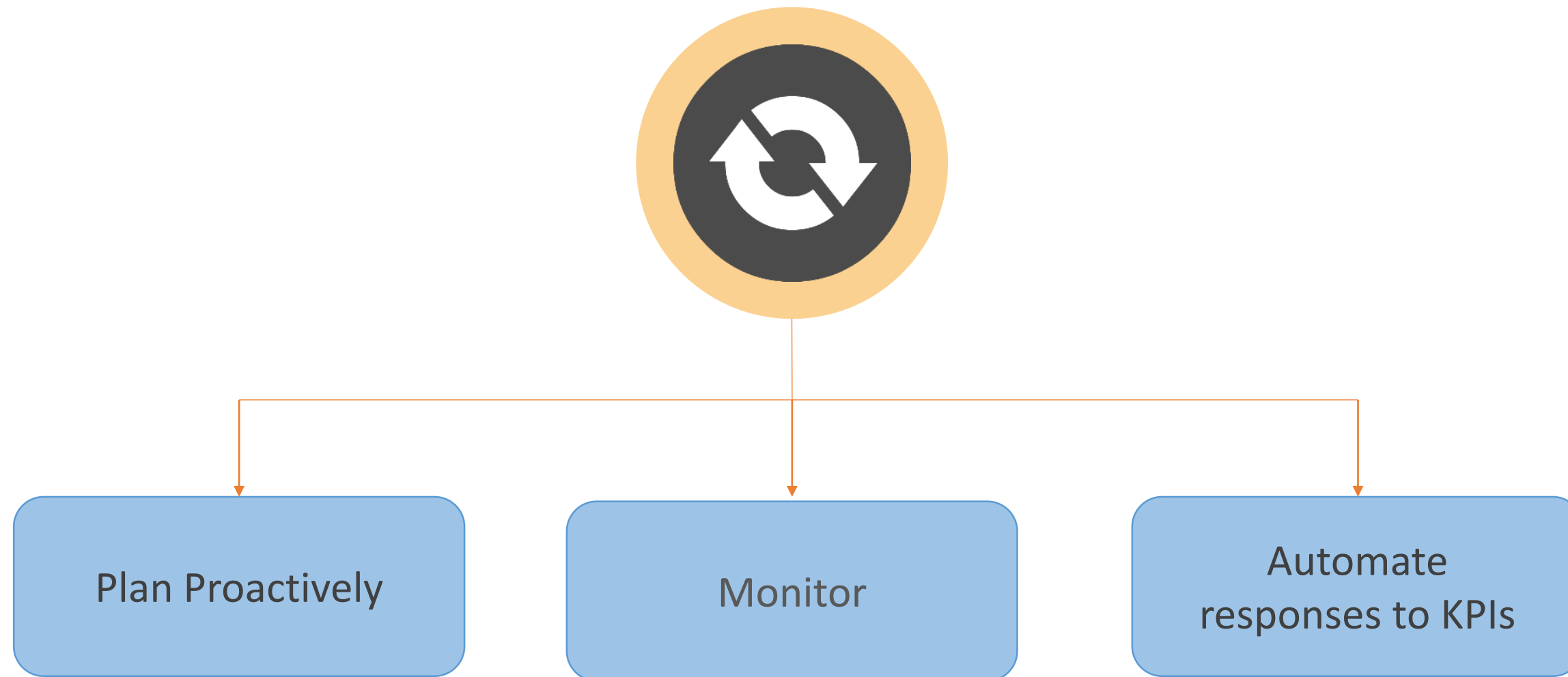
Failure Management

A key to managing failure is the frequent and automated testing of systems for failure and thorough recovery. The cloud enables you to:

CloudWatch

Set up automation to react to monitoring data in case of failure

Launch temporary copies of your system to verify recovery processes

CloudFormation

Store all your data for future use

AWS S3

25
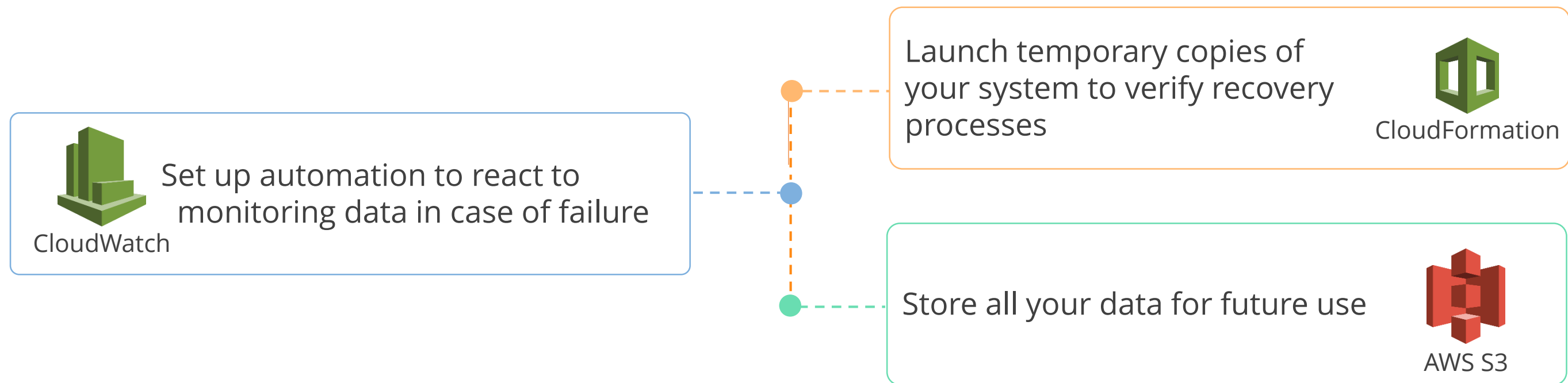
Performance Efficiency

Security	Reliability	Performance Efficiency	Cost Optimization	Operational Excellence

Amazon defines Performance Efficiency as, "The ability to use computing resources efficiently to meet system requirements, and to maintain that efficiency as demand changes and technologies evolve."

Performance Efficiency in the Cloud

AWS provides products such as NoSQL, Media Transcoding, Machine Learning as a service which increase performance efficiency and allows you to:

AWS

Democratize advanced technologies

Go global in minutes

Use server-less architectures

Experiment often

simpli·learn
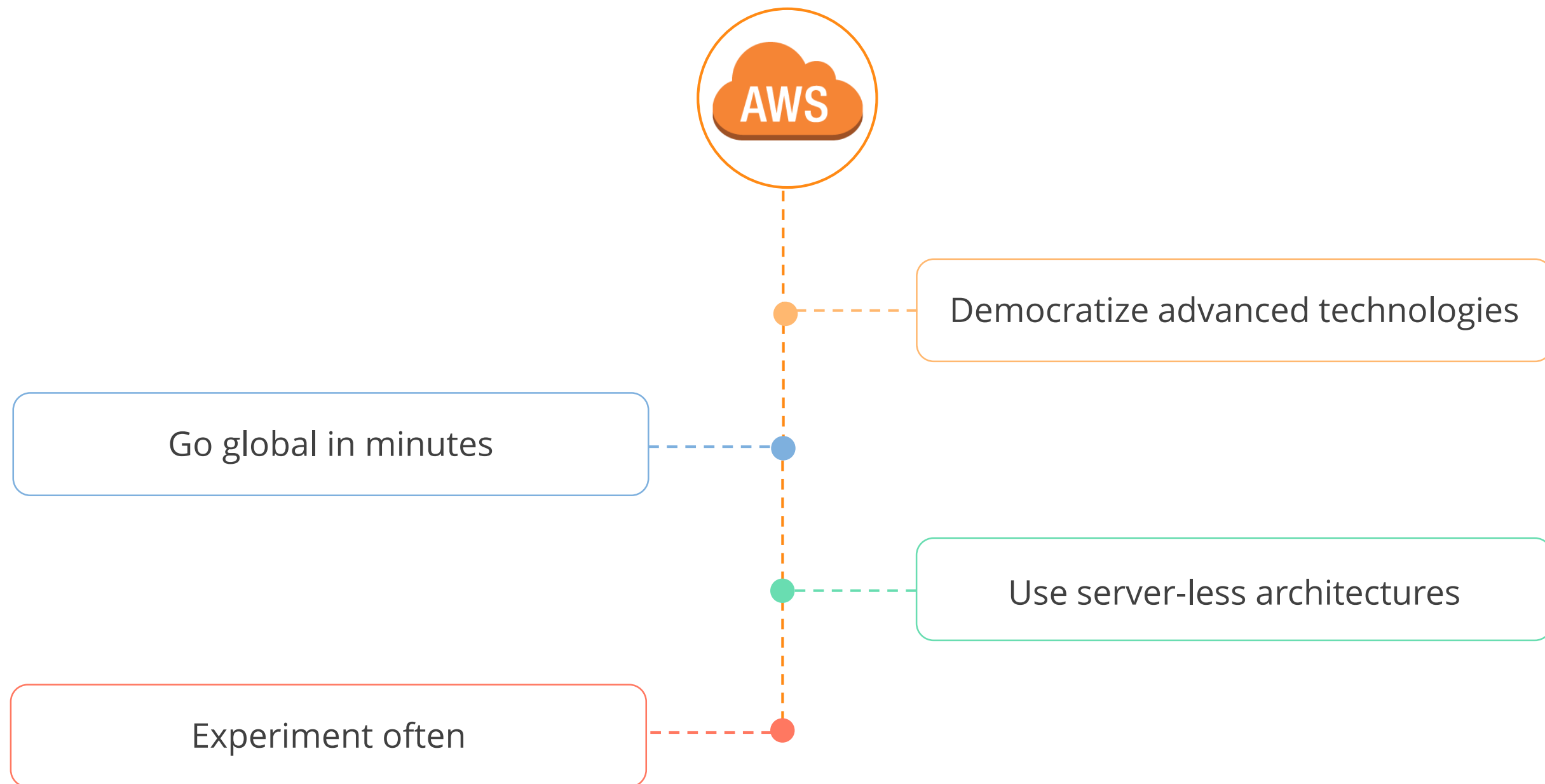
Performance Efficiency in the Cloud

Performance efficiency in the cloud is composed of four areas:

| Compute | Storage | Database | Space-time Trade-off |
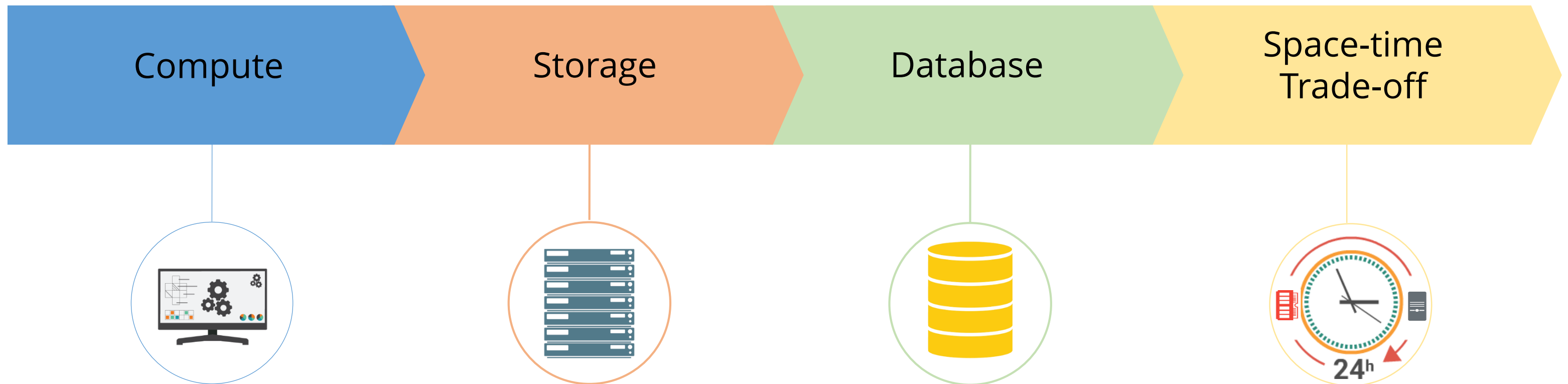
Compute

The Cloud helps compute optimal server configuration, which varies based on application design, usage patterns, and configuration settings.

Allows you to run a web application

AWS Lambda

Allows you to execute code without running an instance

AWS Lambda

Making estimates in advance can lead to incorrect server configurations and low performance efficiency.

Storage

The optimal storage solution for a particular system varies according to your need, whether you need:

Block

File

Object storage

Type of throughput required

Frequency of access availability

Availability and durability constraints

30
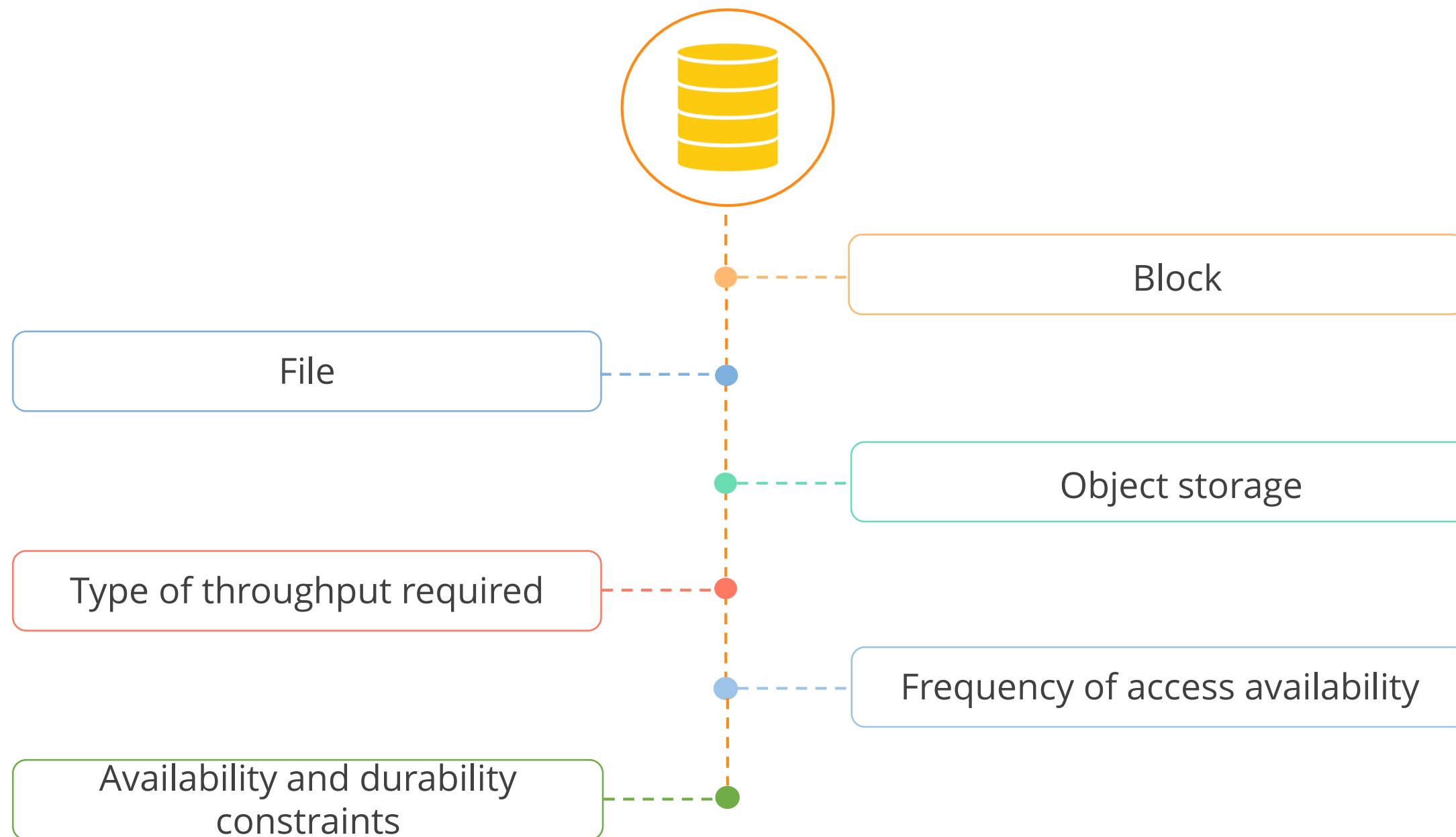
Database

The optimal database solution for a particular system can vary based on requirements for consistency, availability, partition tolerance, and latency. Amazon provides numerous options:

Provides single-digit millisecond latency at any scale

Dynamo DB

Install your own databases from templates

Aurora   AMI

Fully managed relational database

RDS

Change the number or type of nodes per your performance or capacity needs

Redshift
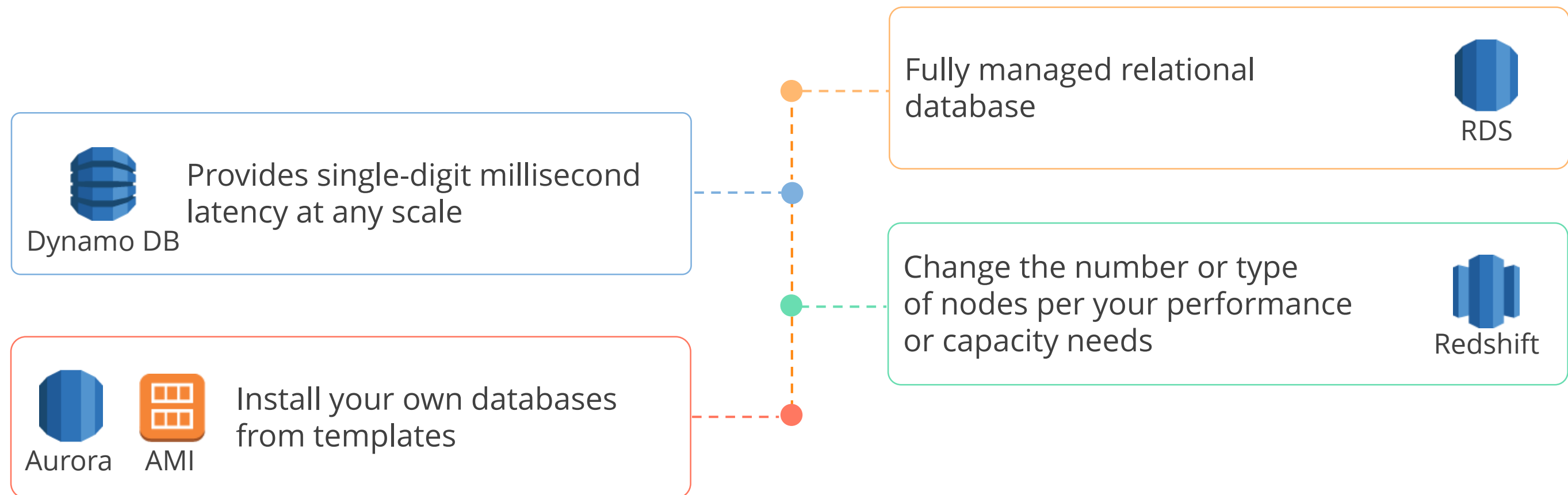
ℹ️ Selecting the wrong database solution can lead to low performance.

Space-time Trade-off

Space (memory or storage) is used to reduce processing time (compute) or time is increased to reduce space requirements. AWS provides the option to maximize one or the other.
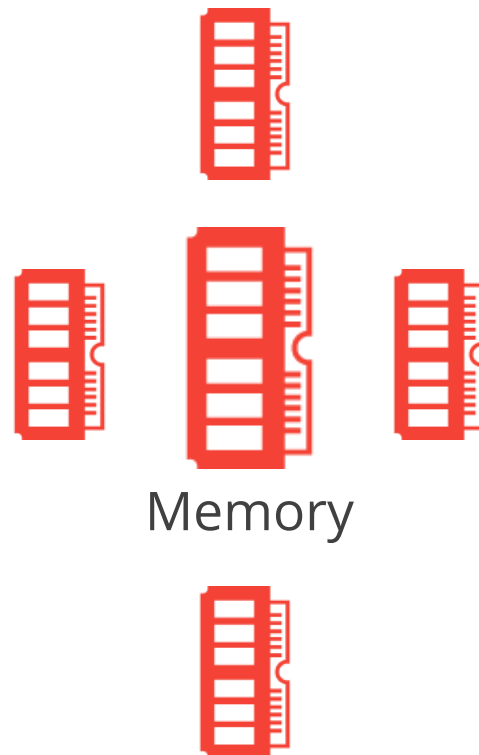
Memory

Time

Monitor your processes regularly to identify any degradation in performance.

Cost Optimization

Security	Reliability	Performance Efficiency	Cost Optimization	Operational Excellence

Amazon defines Cost Optimizations as, "The ability to avoid or eliminate unneeded cost or suboptimal resources."

simpl_ilearn

Cost Optimization in the Cloud

AWS Cloud has a number of ways in which you can provide cost optimization, such as:

Transparently attribute expenditure

Use managed services to reduce cost of ownership

Trade capital expense for operating expense

Benefit from economies of scale

Stop spending money on data center operations

simpli·learn
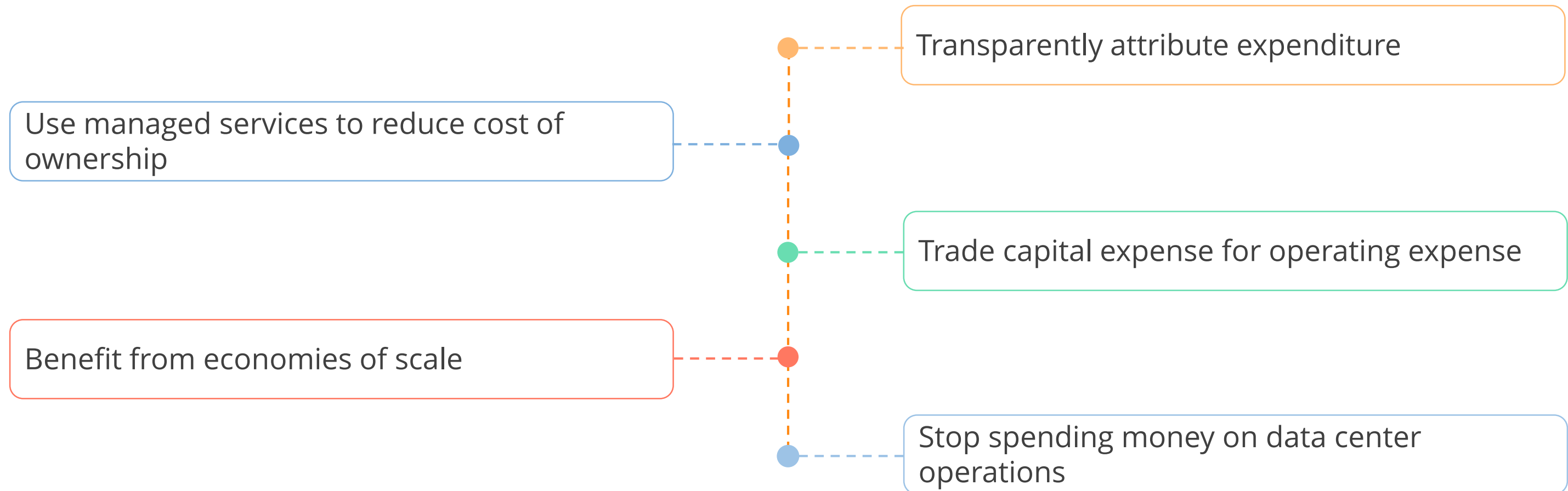
Cost Optimization in the Cloud

Cost Optimization in the cloud is composed of four areas:

Matched Supply and Demand	Cost-effective Resources	Expenditure Awareness	Optimizing over Time
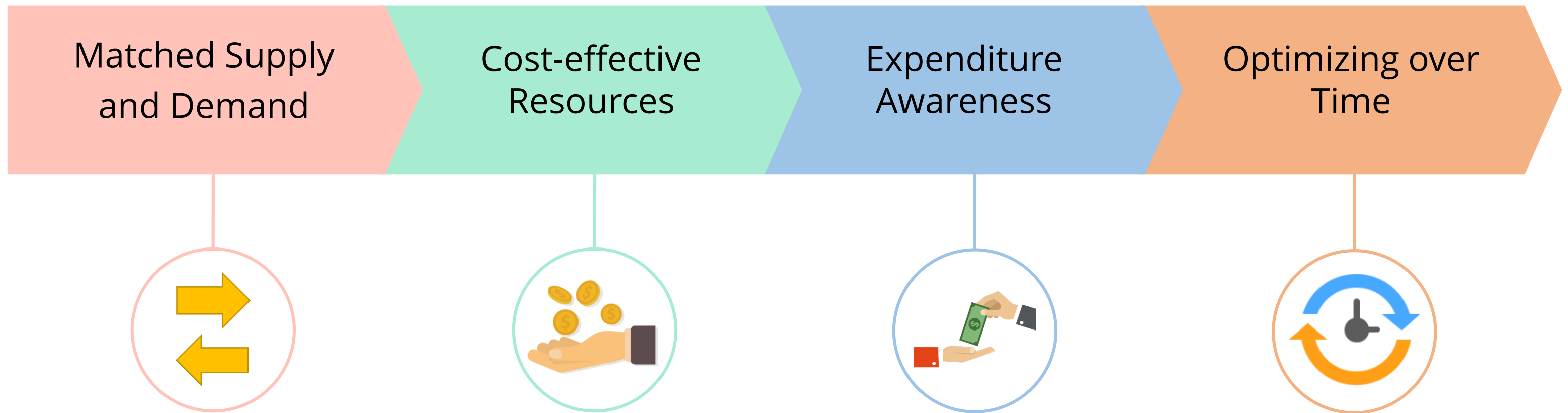
Matched Supply and Demand

Matching supply to demand delivers the lowest costs for a system, but sufficient capacity is needed to cope with demand and failures. AWS automatically provisions resources to match demand using:

Auto scaling approaches

Time-based approaches

Event-driven approaches

Queue-based approaches

simpli·learn
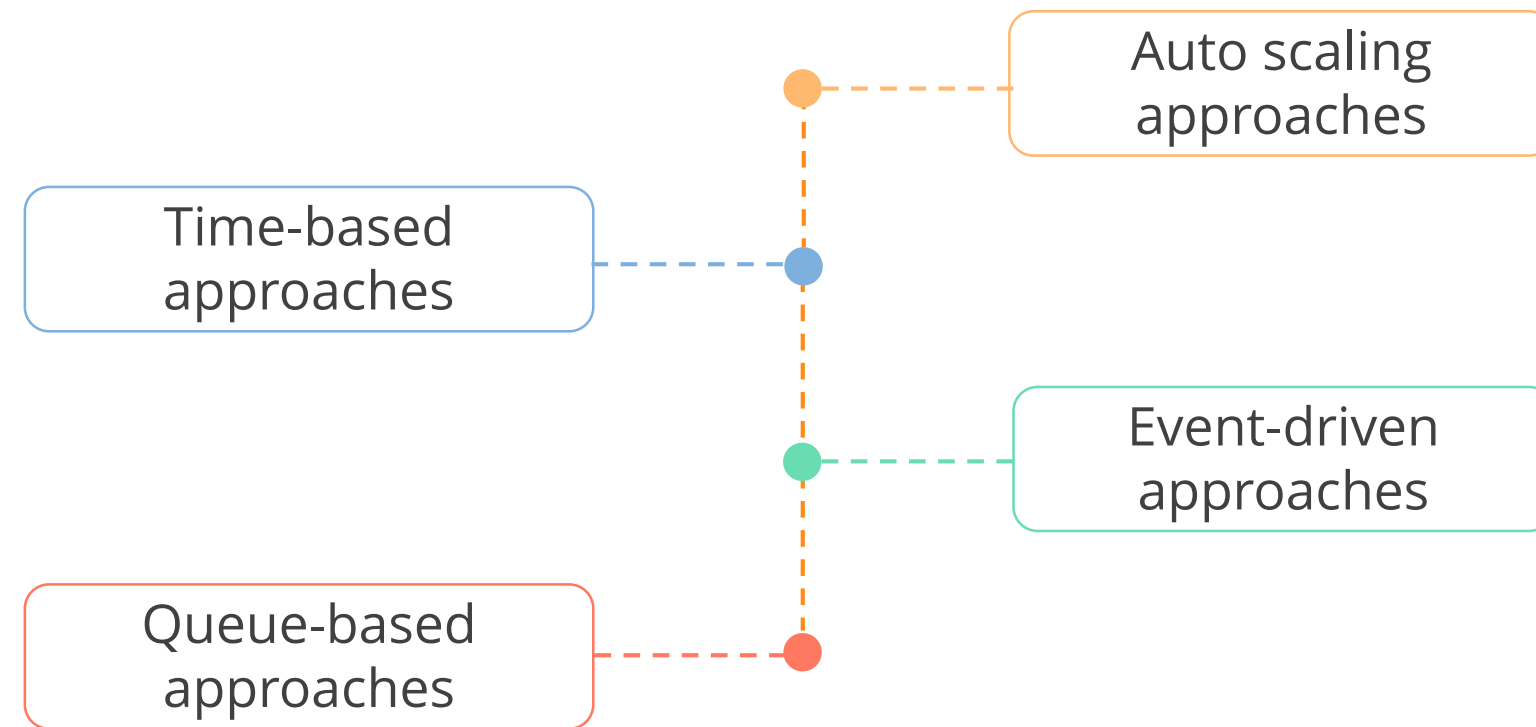
Cost-effective Resources

The key to cost saving is using appropriate instances and resources for your system. Some of the services provided by AWS to reduce cost are:

On-demand Instances

Reserved Instances

Spot Instances

Amazon RDS

DynamoDB

37
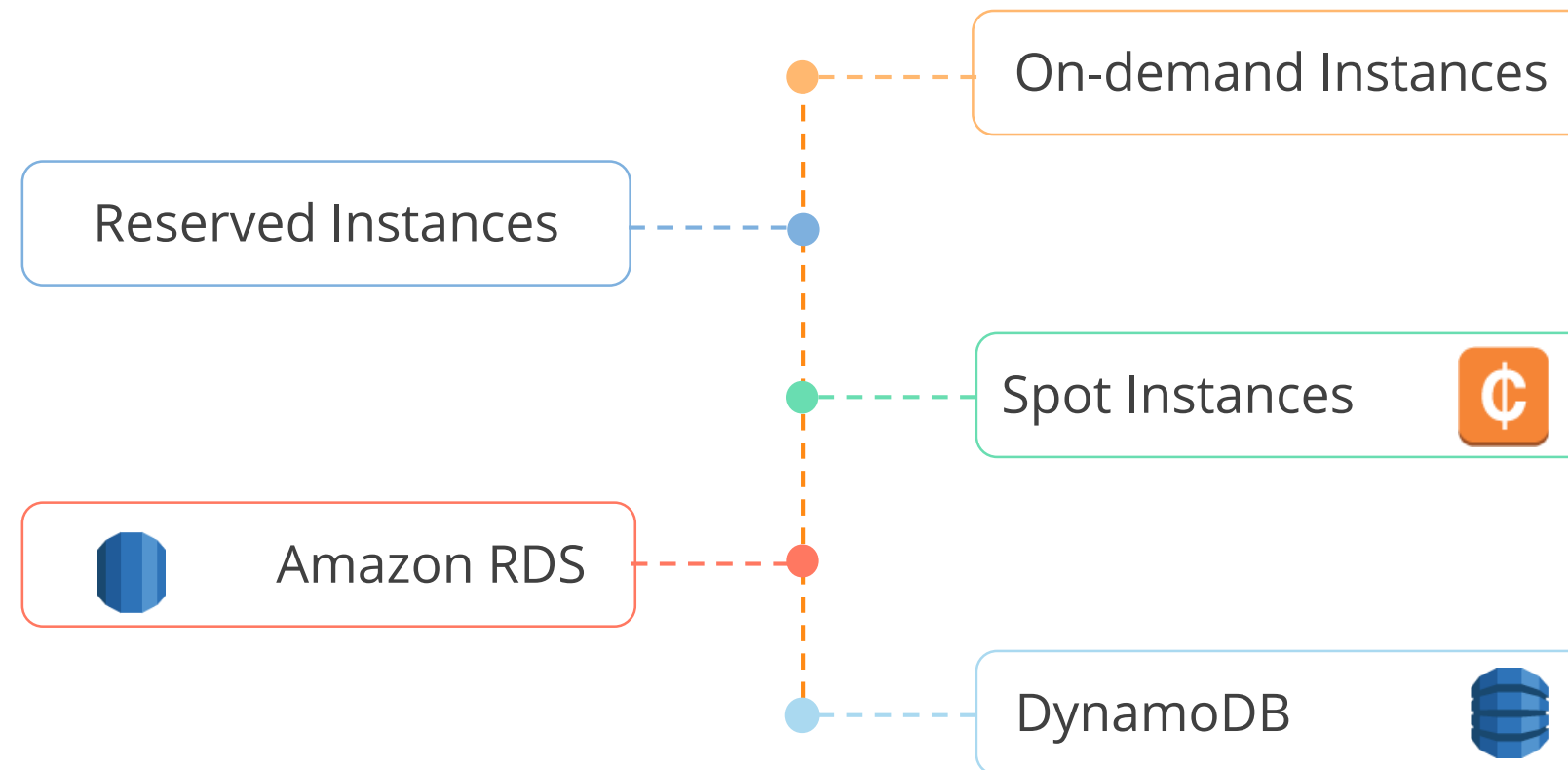
Expenditure Awareness

To categorize and track AWS costs for resources, you can use:

Cost Allocation Tags

The Cost Allocation Report

Billing Alerts

The AWS Simple Monthly Calculator
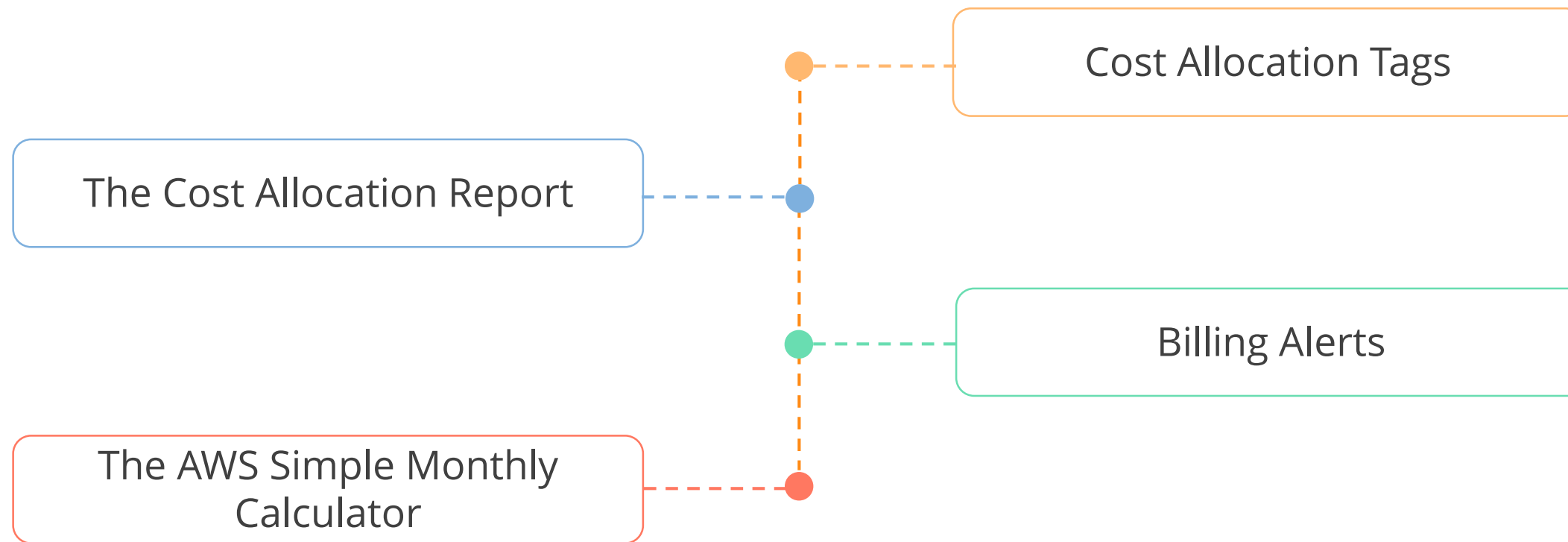
Optimizing over Time

AWS always releases new products and services; therefore, it is a good idea to reassess the existing setup to see if it is the most cost effective.

Kinesis Firehose

ElasticSearch

Time Optimization

AWS

Kinesis

QuickSight

Operational Excellence

Security	Reliability	Performance Efficiency	Cost Optimization	Operational Excellence

Operational Excellence: Operational practices and procedures used to manage production workloads.

simpli learn

Cost Optimization in the Cloud

Perform operations with code

Align operations processes to business objectives

Make regular, small, incremental changes

Test for responses to unexpected events

Learn from operational events and failures

Keep operations procedures current

41

Operational Excellence in the Cloud

Operational Excellence in the cloud is composed of three areas:

Preparation	Operations	Responses
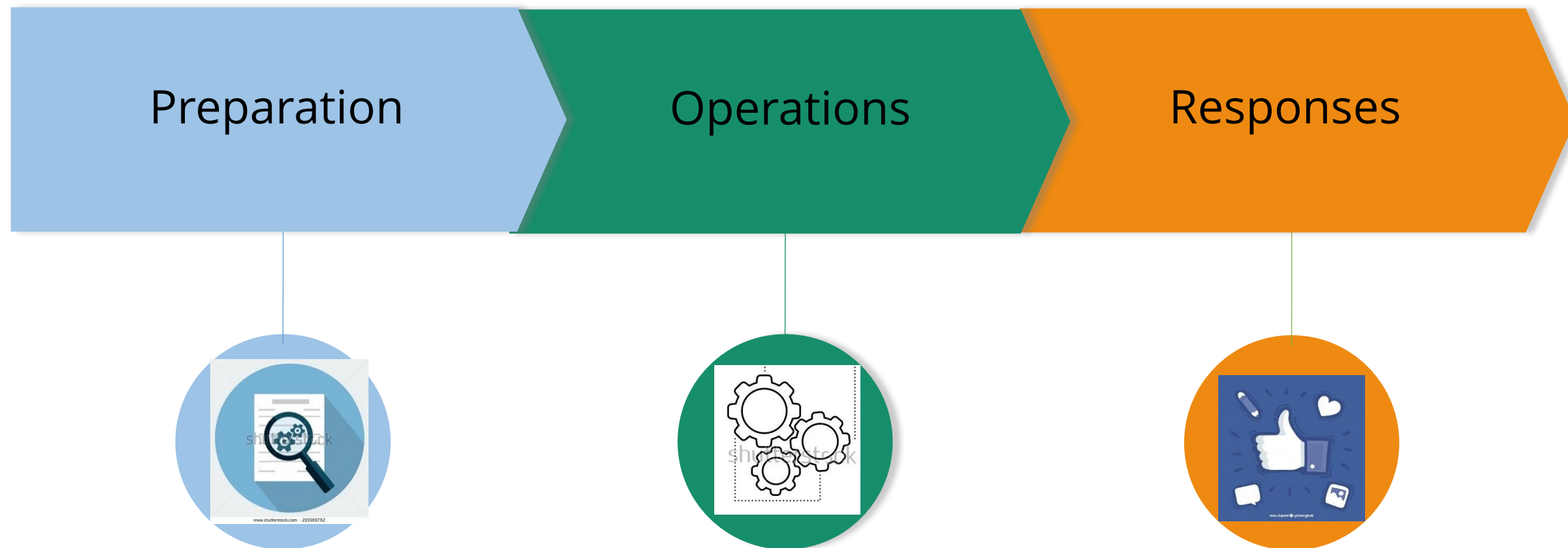
Preparation

Operations checklists will ensure that workloads are ready for production operation, and prevent unintentional production promotion without effective preparation.

Environments, architecture, and the configuration parameters for resources within them, should be documented in a way that allows components to be easily identified for tracking and troubleshooting. Changes to configuration should also be trackable and automated.

1

2

5

4

3

What needs to be in place?
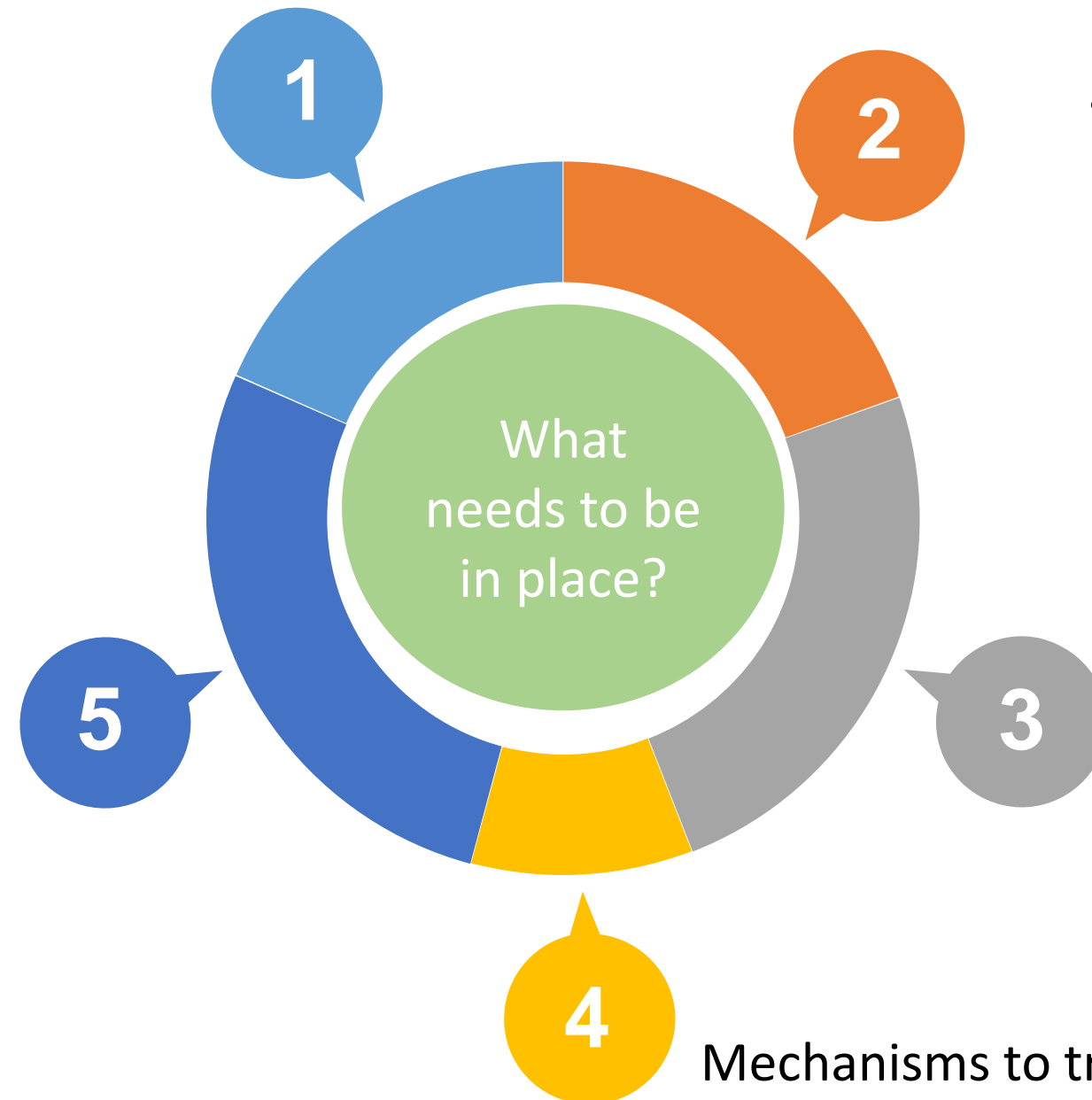
Runbooks & Playbooks: Operations teams should be able to perform normal daily tasks using runbooks, as well as guidance for responding to unexpected operational events (playbooks).

Routine reviews of business cycle events that can drive changes in operations should also be performed. All runbooks and playbooks should be fully tested so that gaps or challenges can be identified and any potential risk can be mitigated.

Mechanisms to track and learn from failures should be in place.

43

simpli·learn

Preparation (contd.)

- AWS services such as AWS CloudFormation can be used to ensure that environments contain all required resources when deployed in production, which reduces the opportunity for human error.
- Implementing Auto Scaling will allow workloads to automatically respond when business-related events affect operational needs.
- Services like AWS Config with the AWS Config rules feature create mechanisms to automatically track and respond to changes in your AWS workloads and environments.
- AWS Trusted advisor can highlight any areas for improvement.

Operations

Operations should be standardized and manageable on a routine basis.

Automation
Small frequent changes can be made easily
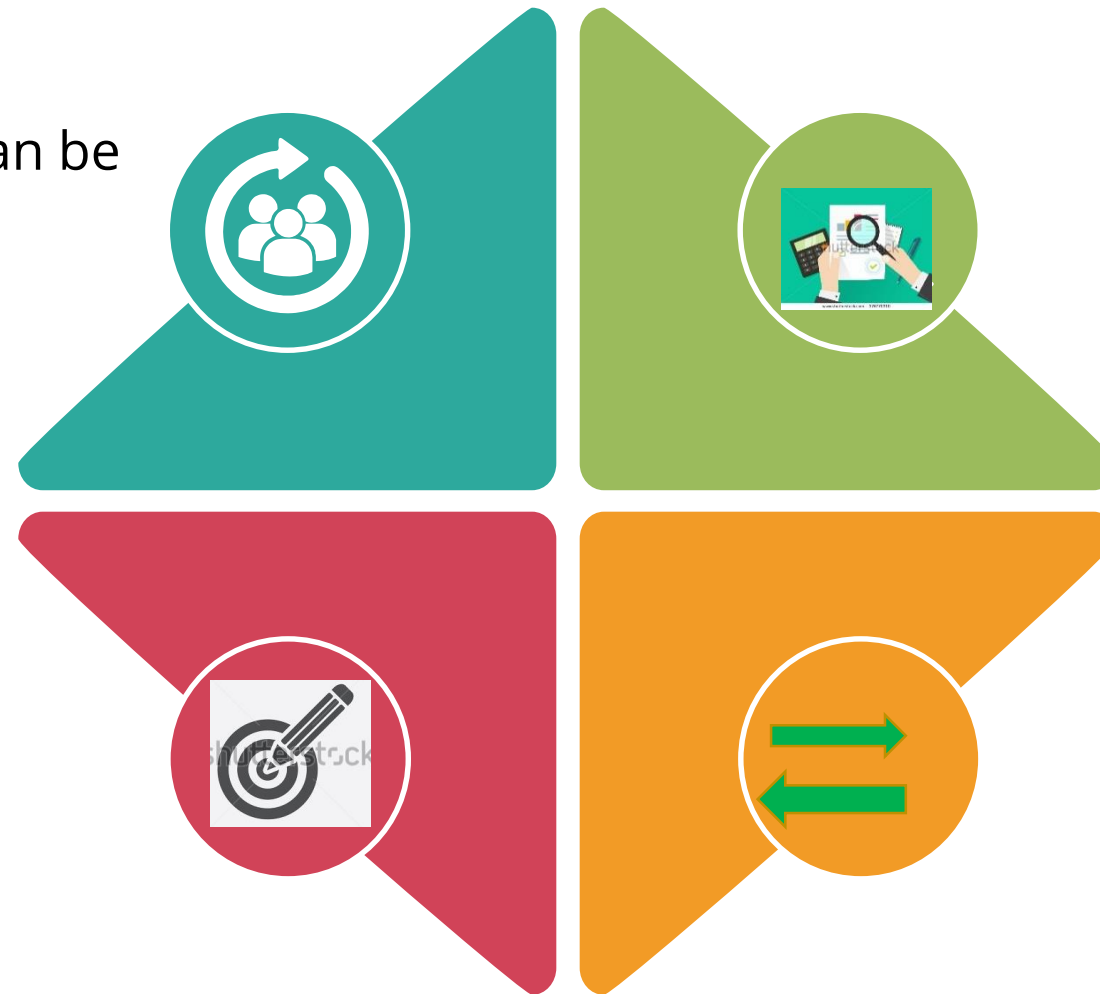
Regular QA

Defined Mechanisms
- Track
- Audit
- Roll Back
- Review

Changes are NOT
- Large and infrequent
- Require scheduled downtime
- Require manual execution

Operations (contd.)

- In AWS you can set up a continuous integration / continuous deployment (CI/CD) pipeline (e.g., source code repository, build systems, deployment and testing automation).
- Release management processes, whether manual or automated, should be tested and be based on small incremental changes, and tracked versions.
- Change quality assurance should include risk mitigation strategies such as Blue/Green, Canary, and A/B testing. Operations checklists should be used to evaluate a workload's readiness for production. Aggregate logs for centralized monitoring and alerts.
- Make sure alerts trigger automated responses, including notification and escalations. Also design monitors for anomalies, not just failures.

Responses

Responses to unexpected operational events should be automated.

Automation	Alerts	QA	Response	Tools

- Mitigation,
- Remediation
- Rollback
- Recovery

- Timely
- Invoke Escalations

- Rollback failed deployments

- CloudWatch
- CloudFormation

- Centrally monitor workloads
- Alerts and notifications

Accessing More Information

The AWS Architecture Center provides guidance and application architecture best practices to build highly scalable and reliable applications in the AWS cloud.

AWS Reference Architectures

The AWS Reference Architecture Datasheets provide architectural guidance to build an application on the AWS cloud infrastructure.

AWS Whitepapers

The technical AWS whitepapers cover all AWS related topics such as architecture, security, and economics.

Architecture Whitepapers from AWS

AWS Well-Architected Framework

This paper describes the AWS Well-Architected Framework, which enables customers to assess and improve their cloud-based architectures and better understand the business impact of their design decisions. We address general design principles as well as specific best practices and guidance in four conceptual areas that we define as the pillars of the Well-Architected Framework. Download Whitepaper (PDF).

AWS Cloud Architecture Best Practices Whitepaper

The cloud reinforces some old concepts of building highly scalable Internet architectures and introduces some new concepts that entirely change the way applications are built and deployed. To leverage the full benefit of the Cloud, including its elasticity and scalability, it is important to understand AWS services, features, and best practices. This whitepaper provides a technical overview of all AWS services and highlights various application architecture best practices to help you design efficient, scalable cloud architectures. Download Whitepaper (PDF).

Building Fault-Tolerant Applications on AWS Whitepaper

AWS provides you with the necessary tools, features and geographic regions that enable you to build reliable, affordable fault-tolerant systems that operate with a minimal amount of human interaction. This whitepaper discusses all the fault-tolerant features that you can use to build highly reliable and highly available applications in the AWS Cloud. Download Whitepaper (PDF).

Using AWS for Disaster Recovery Whitepaper

In the event of a disaster, you can quickly launch resources in Amazon Web Services (AWS) to ensure business continuity. The paper highlights relevant AWS features and services that you can leverage for your DR processes and shows example scenarios on how to recover from a disaster. It further provides recommendations on how you can improve your DR plan and leverage the full potential of AWS for your Disaster Recovery processes. Download Whitepaper (PDF).

AWS Quick Start Reference Deployments

You can rapidly deploy a fully functional environment for a number of enterprise software applications using the AWS CloudFormation templates.

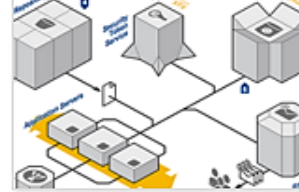## AWS Quick Start Reference Deployments

AWS Quick Start reference deployments help you rapidly deploy fully functional software on the AWS cloud, following AWS best practices for security and availability. An AWS CloudFormation template automates the deployment, and a deployment guide describes the architecture and implementation in detail. Quick Starts are modular and customizable; you can layer additional functionality on top or modify them for your own implementations. Use Quick Starts to deploy the following solutions on AWS:

Lync Server 2013

Build a small or medium-sized Microsoft Lync Server 2013 environment on AWS with high availability and disaster recovery. The guide also provides guidance for larger deployments.
View guide: HTML | PDF

Exchange Server 2013

Deploy Microsoft Exchange Server 2013 with Active Directory Domain Services in a highly available architecture on AWS, choosing a new or existing Amazon VPC.
View guide: HTML | PDF

Windows PowerShell DSC

Build a Microsoft Windows PowerShell DSC pull or push server environment on AWS, using Active Directory and Remote Desktop Gateway.
View guide: HTML | PDF

SharePoint Server 2013

Deploy Microsoft SharePoint Server 2013 on AWS, using SQL Server AlwaysOn Availability Groups with WSFC as the database tier.
View guide: HTML | PDF

Remote Desktop Gateway

Build a secure remote administration solution on AWS, using Remote Desktop Gateway and RDP to access Windows-based instances.
View guide: HTML | PDF

Active Directory Domain Services

Deploy Active Directory Domain Services (AD DS) and Domain Name Server (DNS) on AWS, and choose from three deployment scenarios.
View guide: HTML | PDF

SQL Server with WSFC

Implement a high availability solution with Windows Server Failover Clustering (WSFC) and SQL Server AlwaysOn Availability Groups on AWS.
View guide: HTML | PDF

Web Application Proxy and AD FS

Deploy Web Application Proxy and Active Directory Federation Services (AD FS) into a new or existing AWS infrastructure, following AWS best practices.
View guide: HTML | PDF

Case Studies

AWS maintains a large list of case studies and success stories from their clients. These case studies can be used to explain how, and why some of the largest and most successful companies use AWS for their business.

Case Studies & Customer Success Stories, Powered by the AWS Cloud

AWS case studies and success stories showcase why customers chose AWS, what they're running in the cloud, and what business benefits they have achieved after using AWS. Common topics include Analytics, Big Data, Enterprise, Government & Education, Startups, and Web Apps. You can find an alphabetical listing of all AWS customer case studies here.

Netflix	Airbnb	Nokia	Yelp
Watch the Video »	Learn More »	Watch the Video »	Watch the Video »
Expedia	Adobe	Pinterest	Zynga
Learn More »	Watch the Video »	Watch the Video »	Watch the Video »

simplilearn

Knowledge Check

| KNOWLEDGE CHECK | The AWS Well-Architected Framework is designed to help you _____. |

a. Stop guessing your capacity needs

b. Test systems at production scale

c. Lower the risk of architecture change

d. Increase the amount of administration required

| KNOWLEDGE CHECK | The AWS Well-Architected Framework is designed to help you _____. |

a. Stop guessing your capacity needs

b. Test systems at production scale

c. Lower the risk of architecture change

d. Increase the amount of administration required

The correct answer is   **a), b), and c)**

The AWS Well-Architected Framework is designed to help you understand the pros and cons of the decisions you make while building systems on AWS. So you can stop guessing your capacity needs, test systems at production scale, lower the risk of architecture change, automate to make architectural experimentation easier, and allow for evolutionary architectures.

KNOWLEDGE CHECK	The AWS Well-Architected Framework is based on which of the following five pillars:

a. Security, Reliability, Performance Efficiency, Cost Optimization, and Operational Excellence

b. Security, Redundancy, Performance Efficiency, Cost Optimization, and Operational Excellence

c. Security, Reliability, Environmental Efficiency, Cost Optimization, and Operational Excellence

d. Security, Redundancy, Performance Efficiency, Resource Optimization, and Operational Excellence

The AWS Well-Architected Framework is based on which of the following five pillars:

a. Security, Reliability, Performance Efficiency, Cost Optimization, and Operational Excellence

b. Security, Redundancy, Performance Efficiency, Cost Optimization, and Operational Excellence

c. Security, Reliability, Environmental Efficiency, Cost Optimization, and Operational Excellence

d. Security, Redundancy, Performance Efficiency, Resource Optimization, and Operational Excellence

The correct answer is **a)**

The AWS Well-Architected Framework is based on Security, Reliability, Performance Efficiency, Cost Optimization, and Operational Excellence.

Planning and Designing

The principles involved in planning and designing cloud infrastructure

Planning and Designing

In this section you'll learn about:

Scaling

Loose
Coupling

Redundancy

Cost
Optimization

Automation

Security
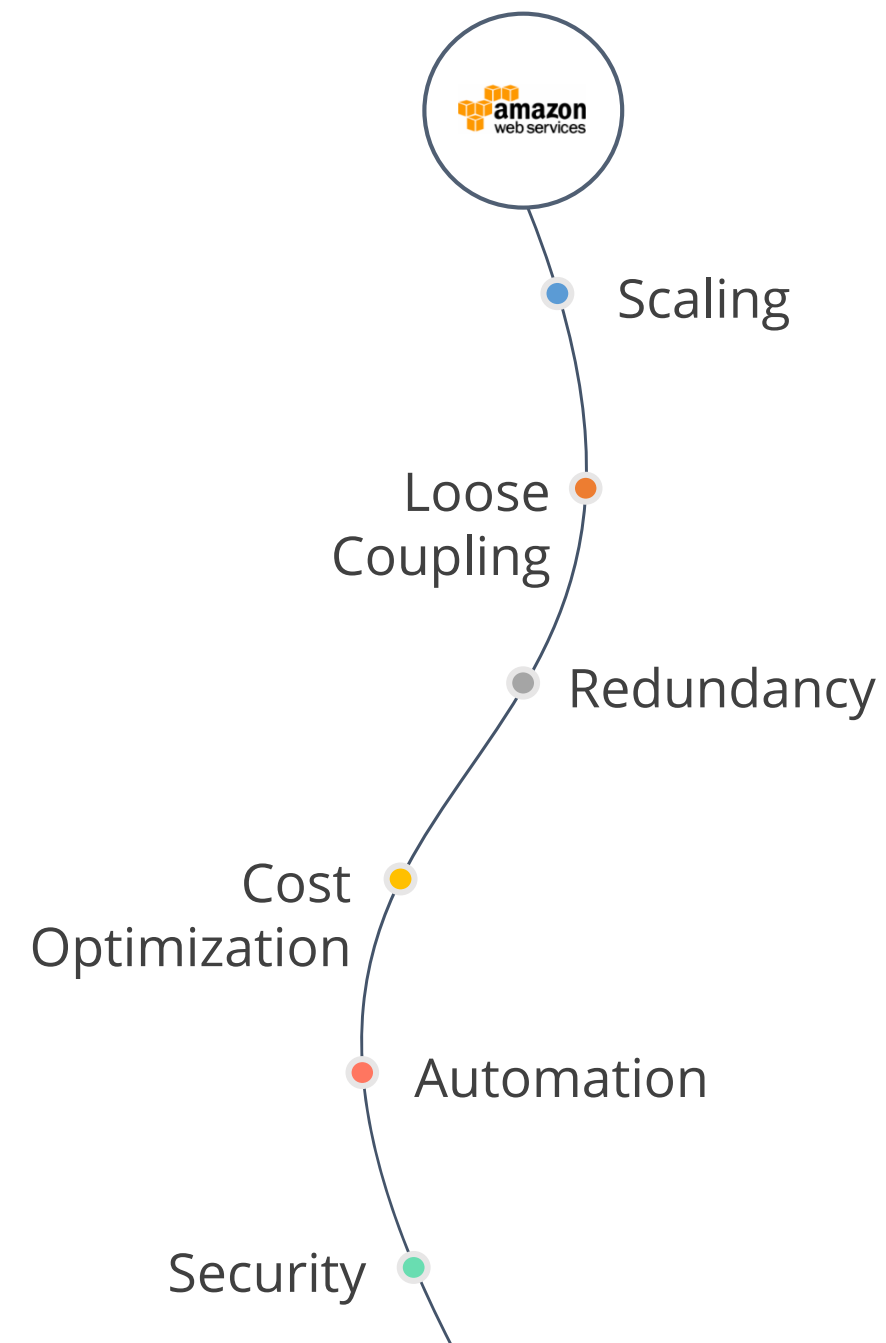
Scalability

Cloud computing provides virtually unlimited on-demand capacity so you can scale whenever you need to. There are two ways to scale—vertically and horizontally.

Vertical Scaling

Vertical scaling means increasing the specifications of an individual resource. For example, increasing the memory and CPU on a server.

Memory   Memory

CPU   CPU

> ℹ️ Vertical scaling can eventually hit a limit and sometimes prove expensive.

simpli learn

Horizontal Scaling

Horizontal scaling means increasing the number of resources rather than the specifications of a resource. For example, adding additional web servers to help spread the load of traffic hitting your application.

Web Server
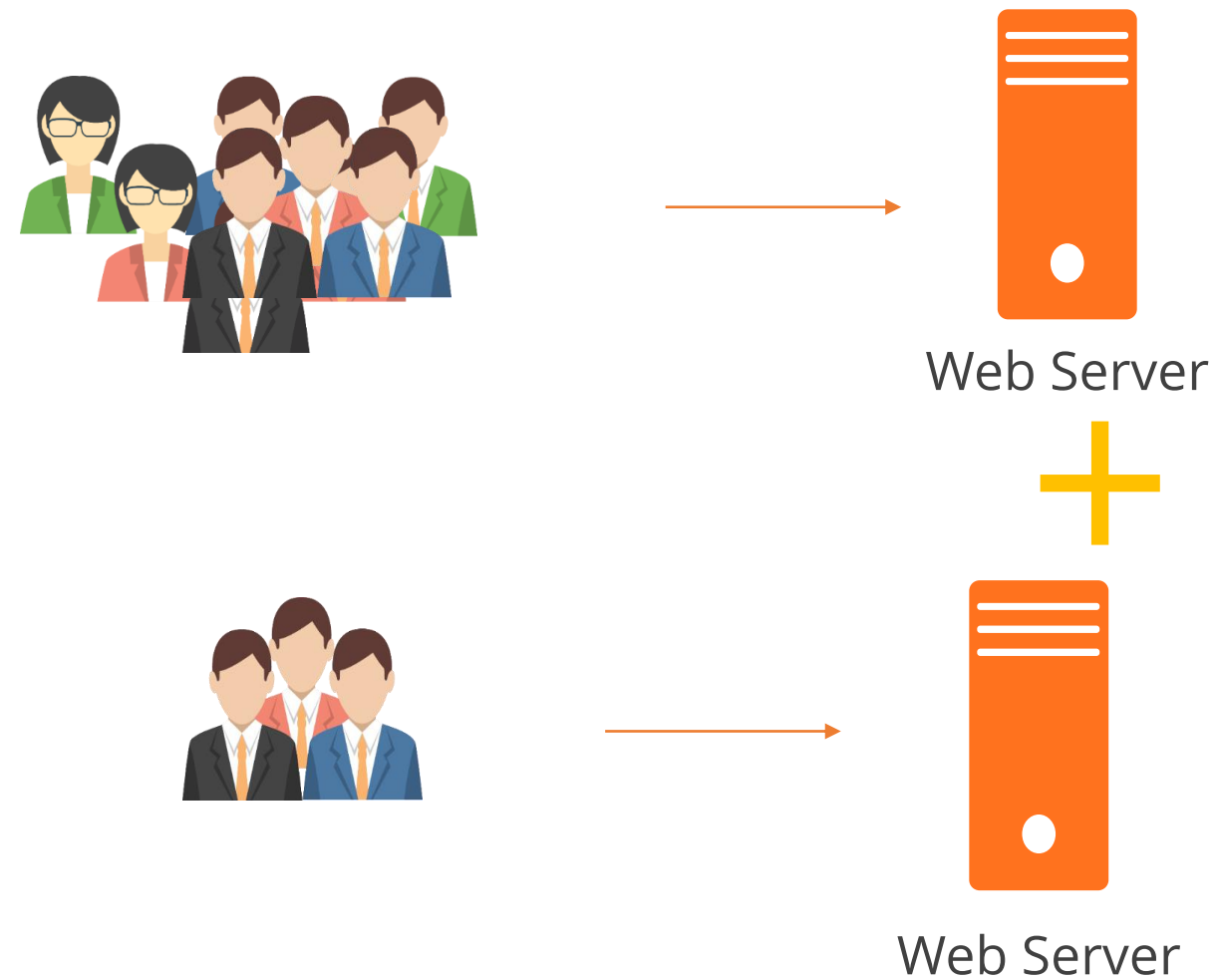
Web Server

Not every architecture can distribute their workload to multiple resources.

Stateless Applications

A stateless application is one that needs no knowledge of previous interactions and stores no session information. Example, a webserver that provides the same web page to any end user.

Web Server
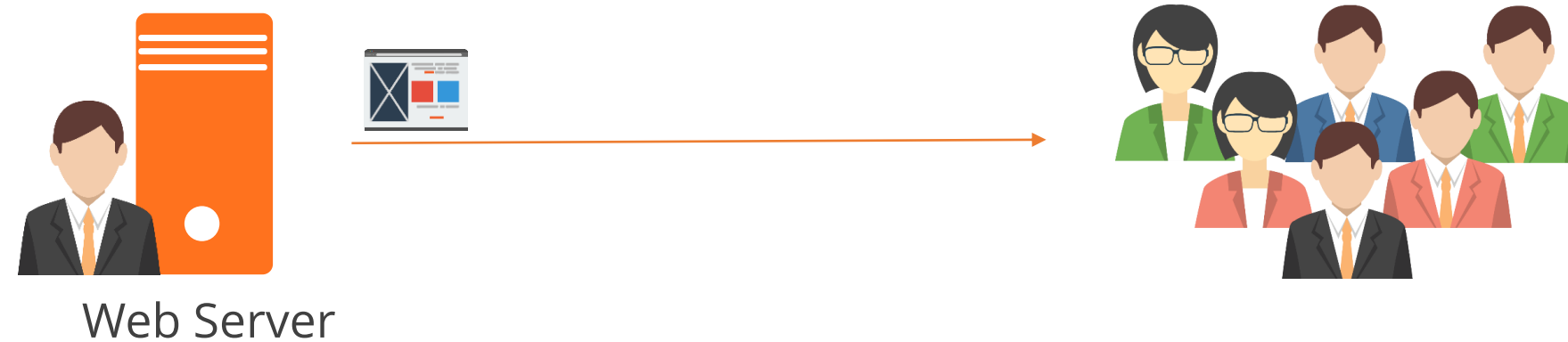
Push Model

A popular way to distribute workload across multiple resources is by using a load balancer, such as the AWS Elastic Load Balancer.

Elastic Load Balancer
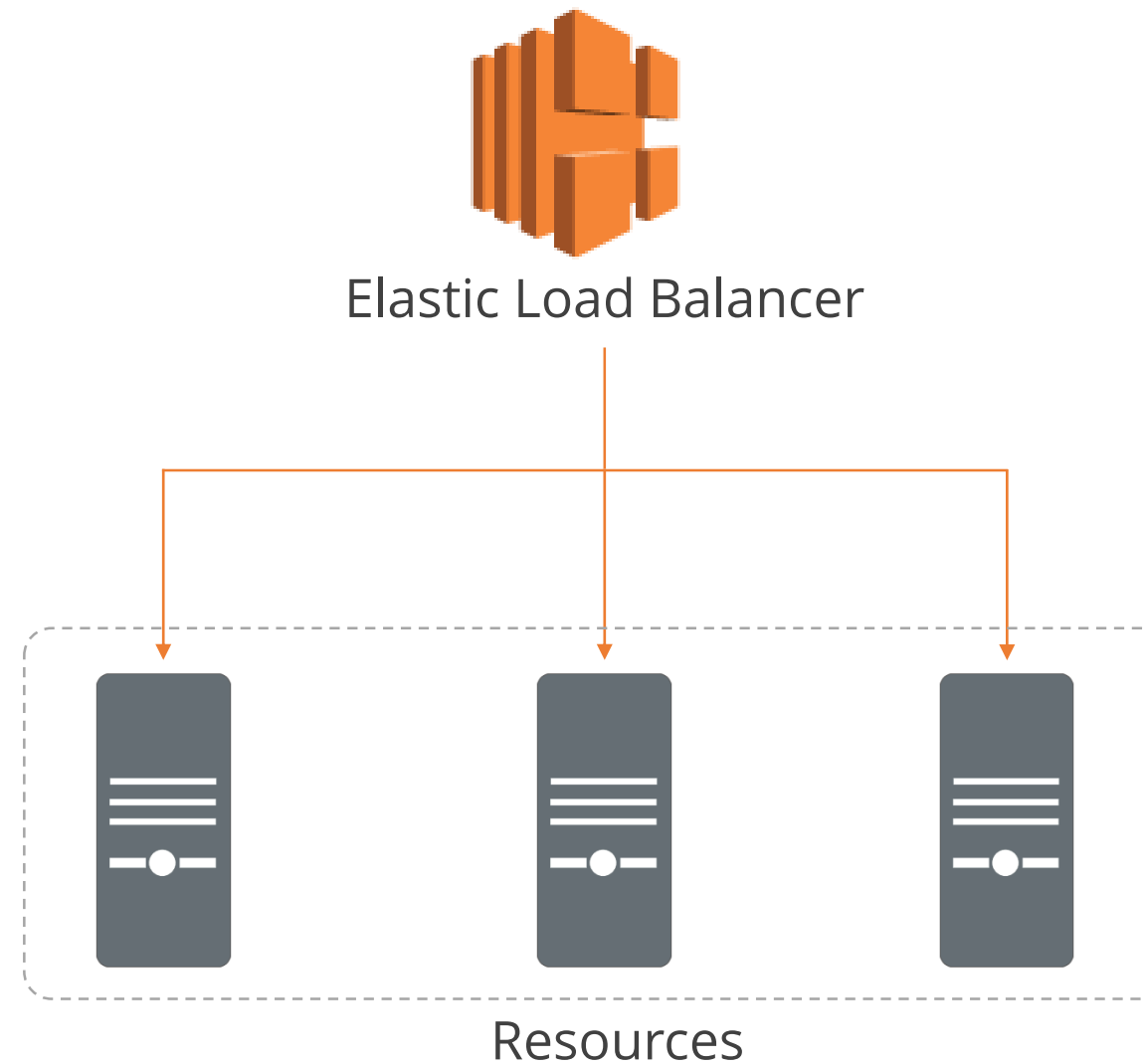
Resources

Pull Model

In the pull model, tasks that need to be performed can be stored as messages in a queue and multiple compute resources can pull and process the messages in a distributed fashion.

Message Queue

Task 1

Task 2

Task 3

Resources

Resources

Resources

Results
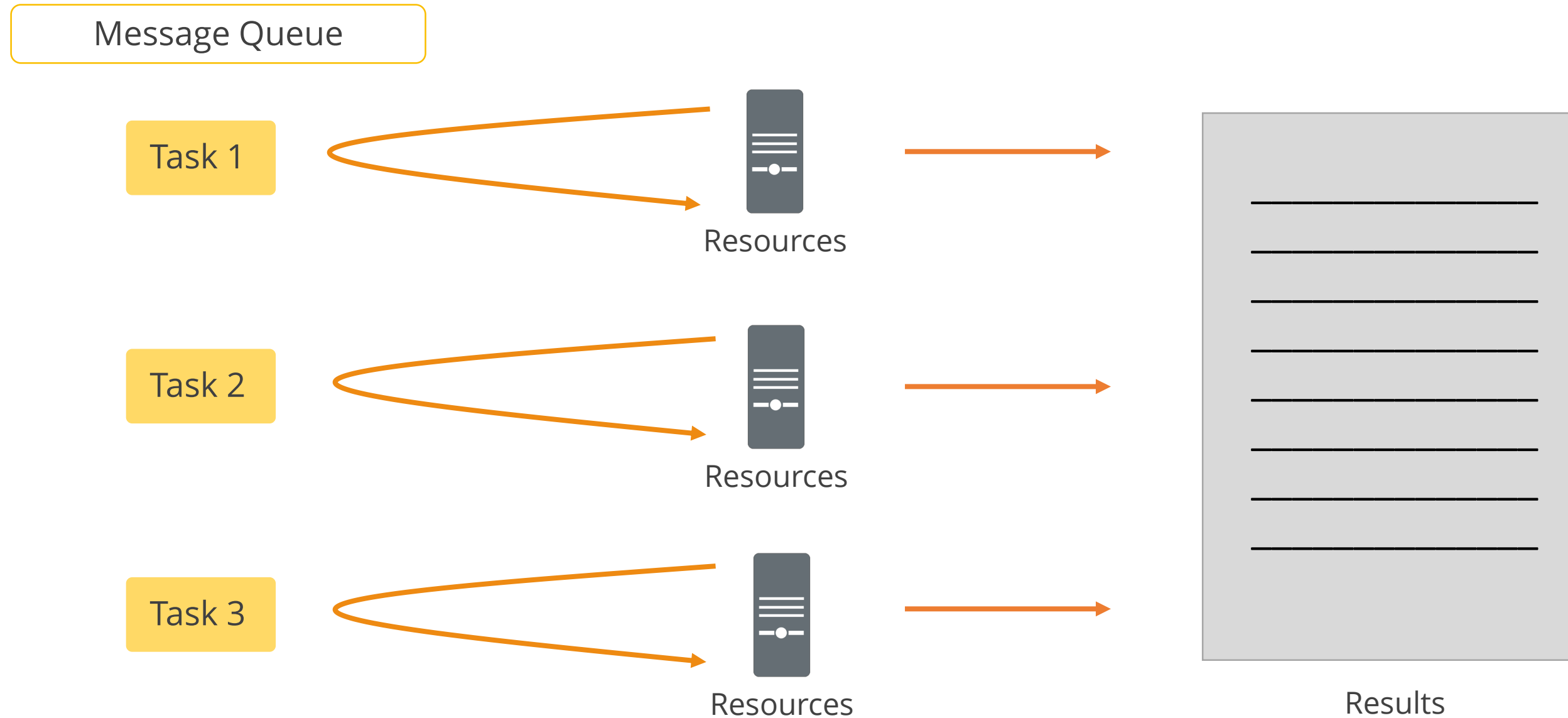
Stateless Components

Components in the architecture can be made stateless by not storing anything on the local file system and instead storing user or session based information in a database (DynamoDB or MySQL), or on a shared storage layer (Amazon S3 or EFS).

DynamoDB   MYSQL   Amazon S3   EFS

simpli·learn
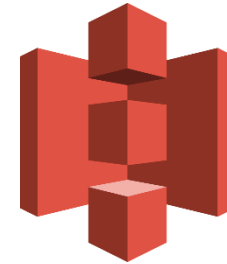
Stateful Applications

Some layers of the architecture cannot be turned into stateless components. For example, databases, which are stateful by definition or applications designed to run on a single server.

Amazon RDS

DynamoDB

67

Distributed Processing

In situations which require large amounts of data to be processed, a distributed processing approach should be used. If a task was to run on a single compute resource, it would max out the resources and take a long time to complete. But if the task is divided into smaller fragments of work, then each of the tasks can be executed across a larger set of compute resources.

Task   CPU

simplilearn
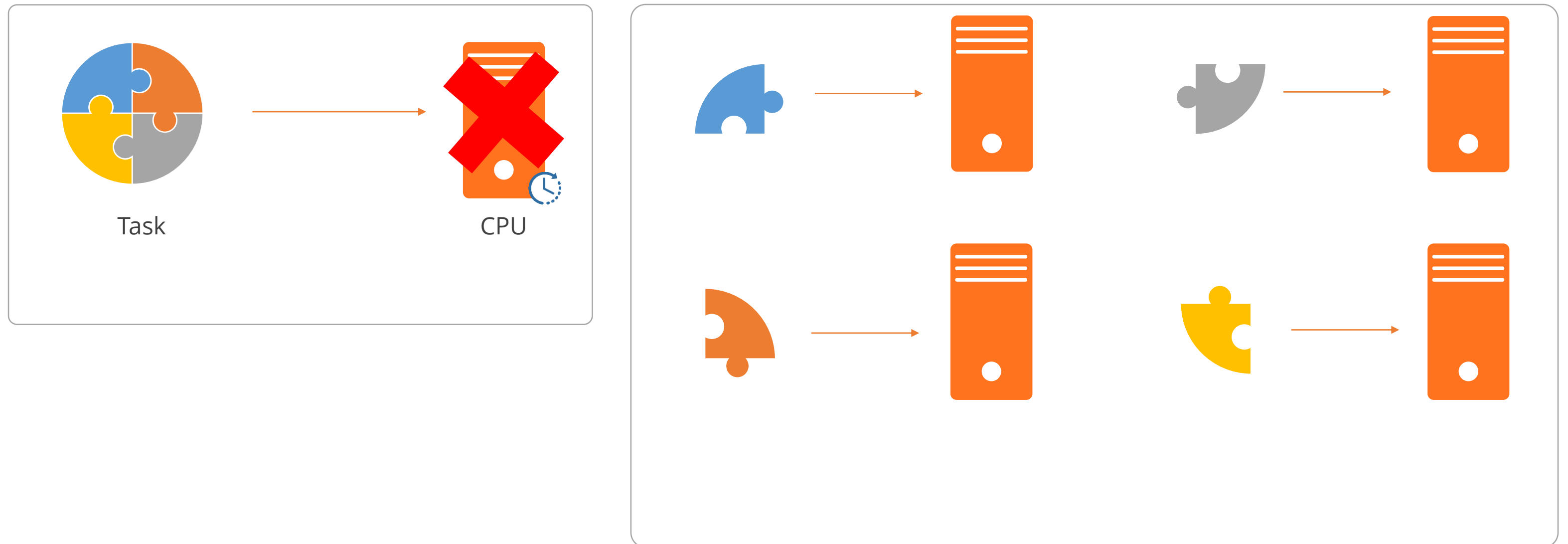
Disposable Resources

Cloud computing completely changes the mindset of an IT infrastructure environment. With cloud computing all infrastructure is temporary or disposable. New instances can be launched when required and can be disposed when the requirement ends.

Automate Compute Resource Initiation

The AWS features that make new environment creation an automated and repeatable process are:

Bootstrapping

Golden Images

Hybrid Approach

70
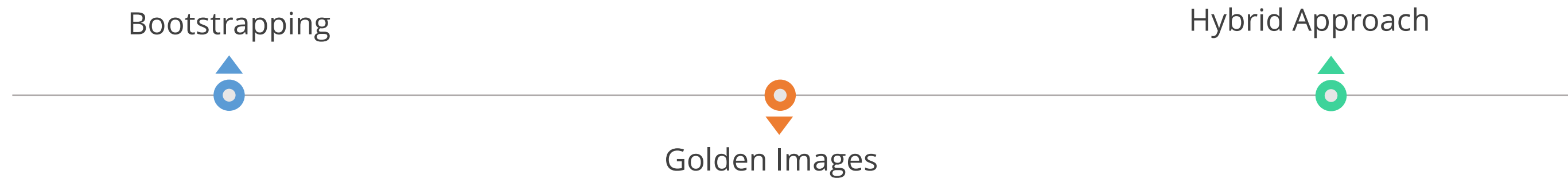
Bootstrapping

You can launch AWS resources with a default configuration and execute automated scripts to install software or copy data to bring those resources to the required state.

Data   0 1   EC2   Web Server
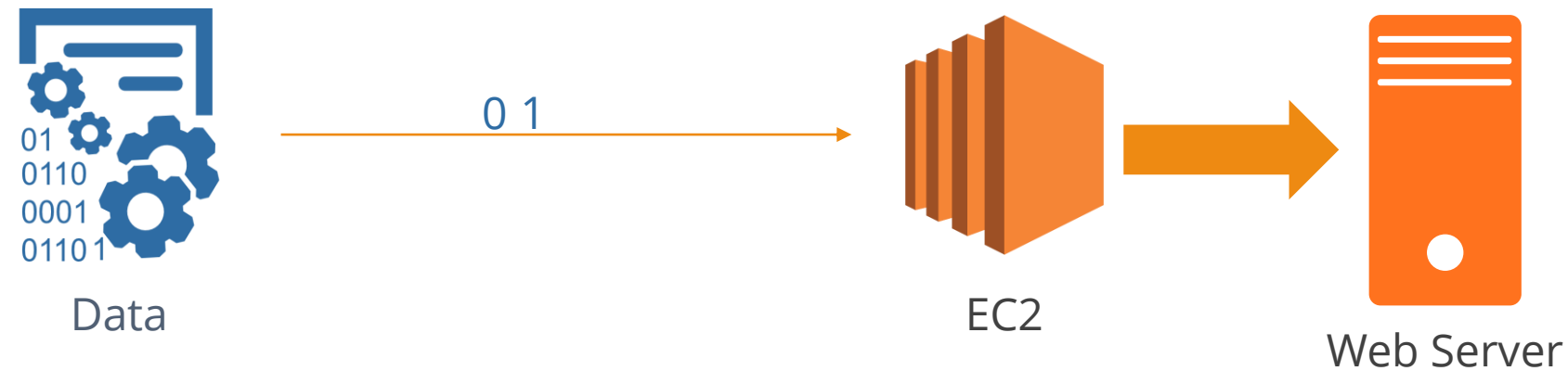
With AWS you can achieve Bootstrapping with your own scripts, Chef/Puppet, OpsWork lifecycle events, or CloudFormation.

Golden Images

Golden Images mean:

- You take snapshots of EC2 instances, RDS instances, or EBS volumes which can launch new instances.
- EC2 instances can be customized and saved as Amazon machine images (AMIs) and then you can launch as many instances as you want.

EC2 → Snapshot → New EC2 Instance

RDS → Snapshot → New RDS Instance

EBS → Snapshot → New EBS Instance

EC2 → AMI → Multiple EC2 Instances

simpli learn
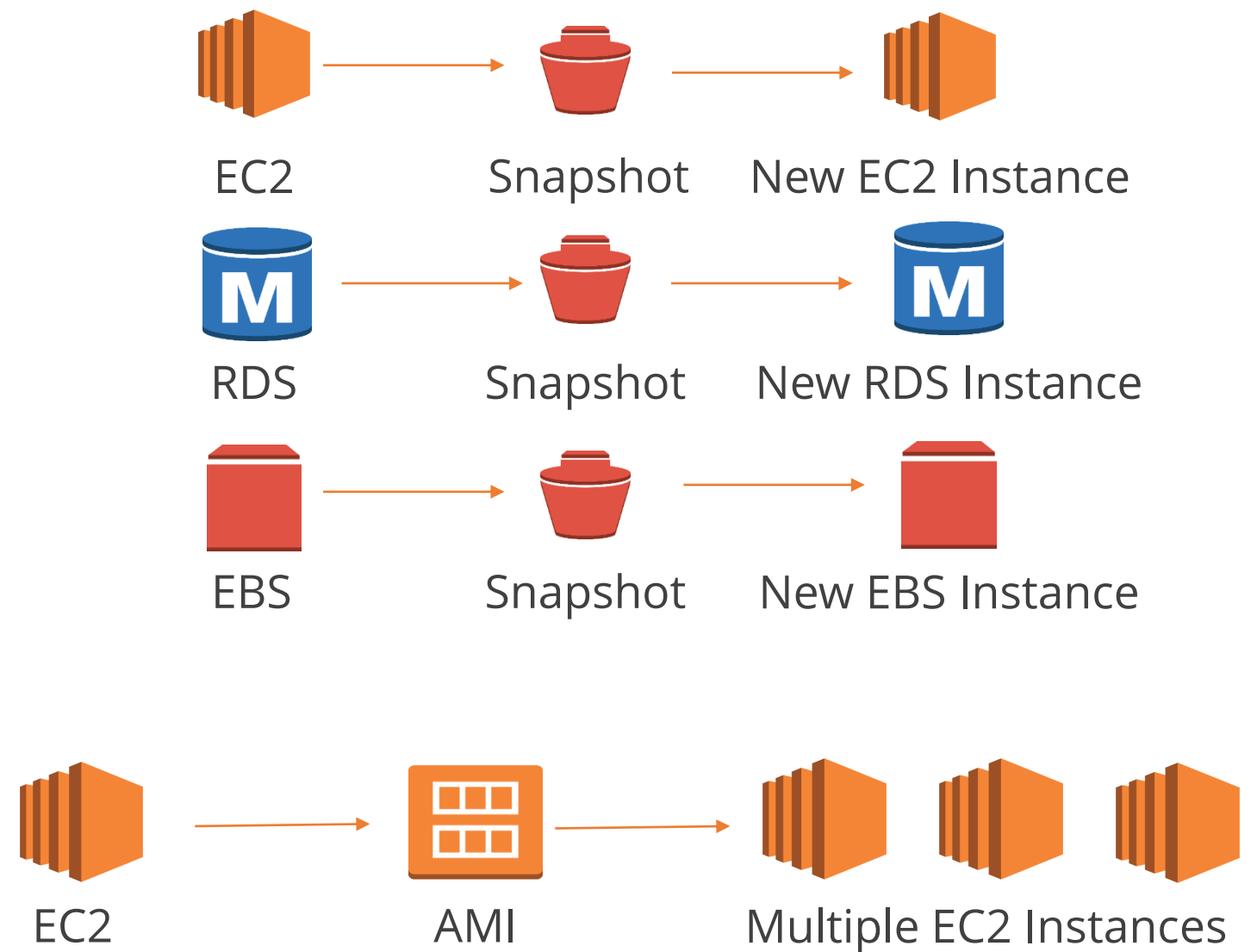
Hybrid Approach

Utilize both bootstrapping and golden images to automate your compute launch processes.

EC2 → Snapshot → New EC2 Instance ← Data

Web Server
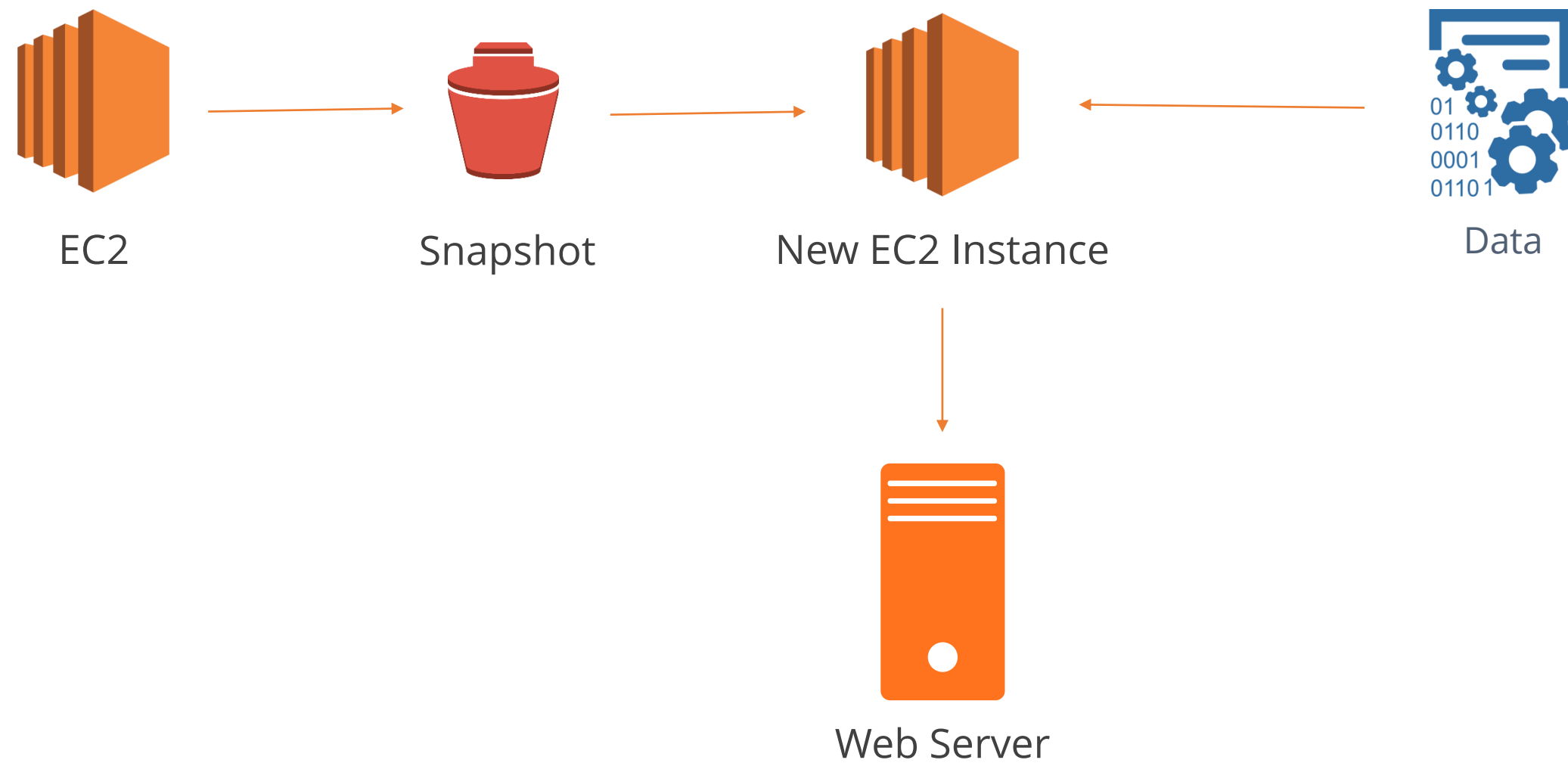
Infrastructure as Code

AWS assets are programmable—you can apply all the techniques discussed earlier to entire environments and not just individual resources.
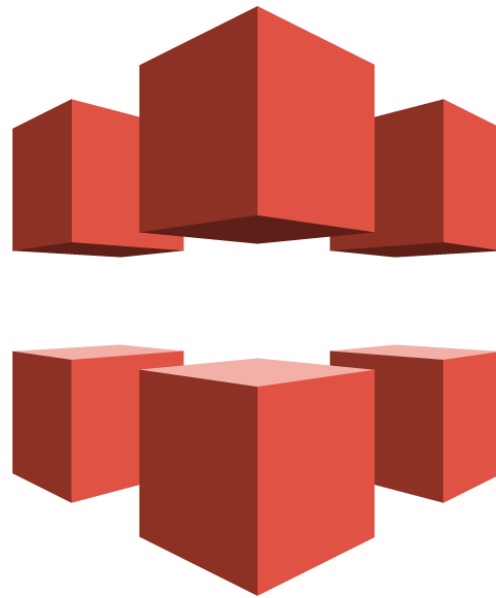
Amazon CloudFront

AWS Automation Services

AWS allows you to reduce the level of manual interaction in your environment. You can react to a variety of events without any manual effort by using AWS automation services, such as:

EC2 Auto Recovery

CloudWatch Alarms and Events

AWS Elastic Beanstalk

Auto Scaling

OpsWorks LifeCycle Events

simpli learn
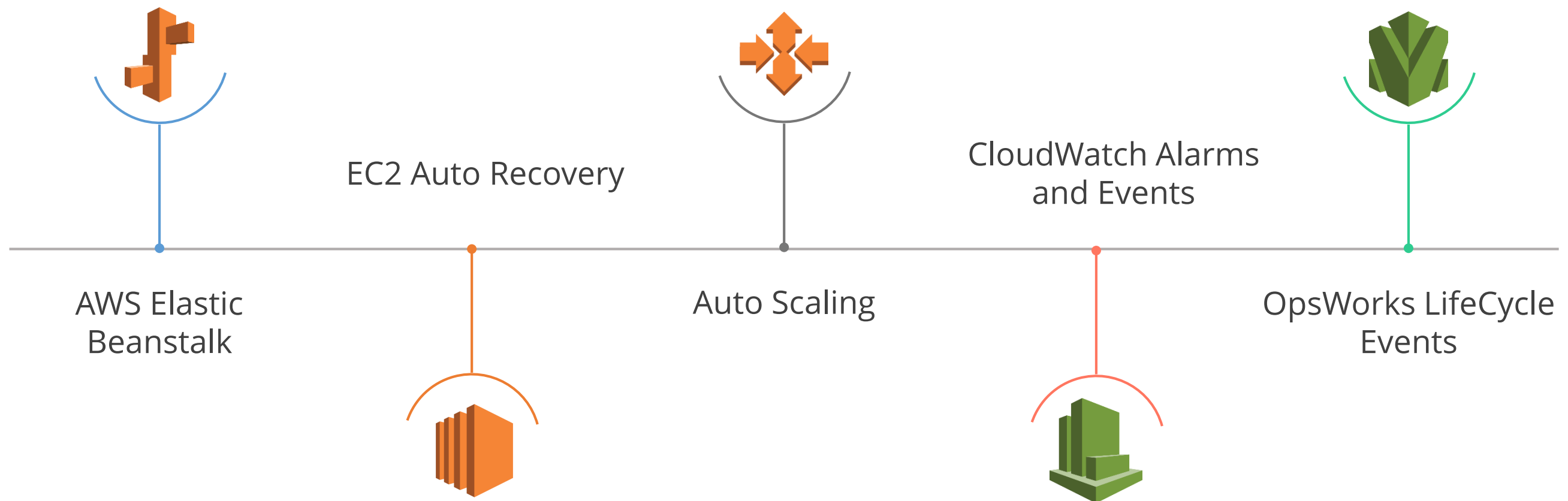
Loose Coupling

Design applications that comprise smaller, loosely coupled components so that there is no single point of failure.

Storage

Web Server

Web Server
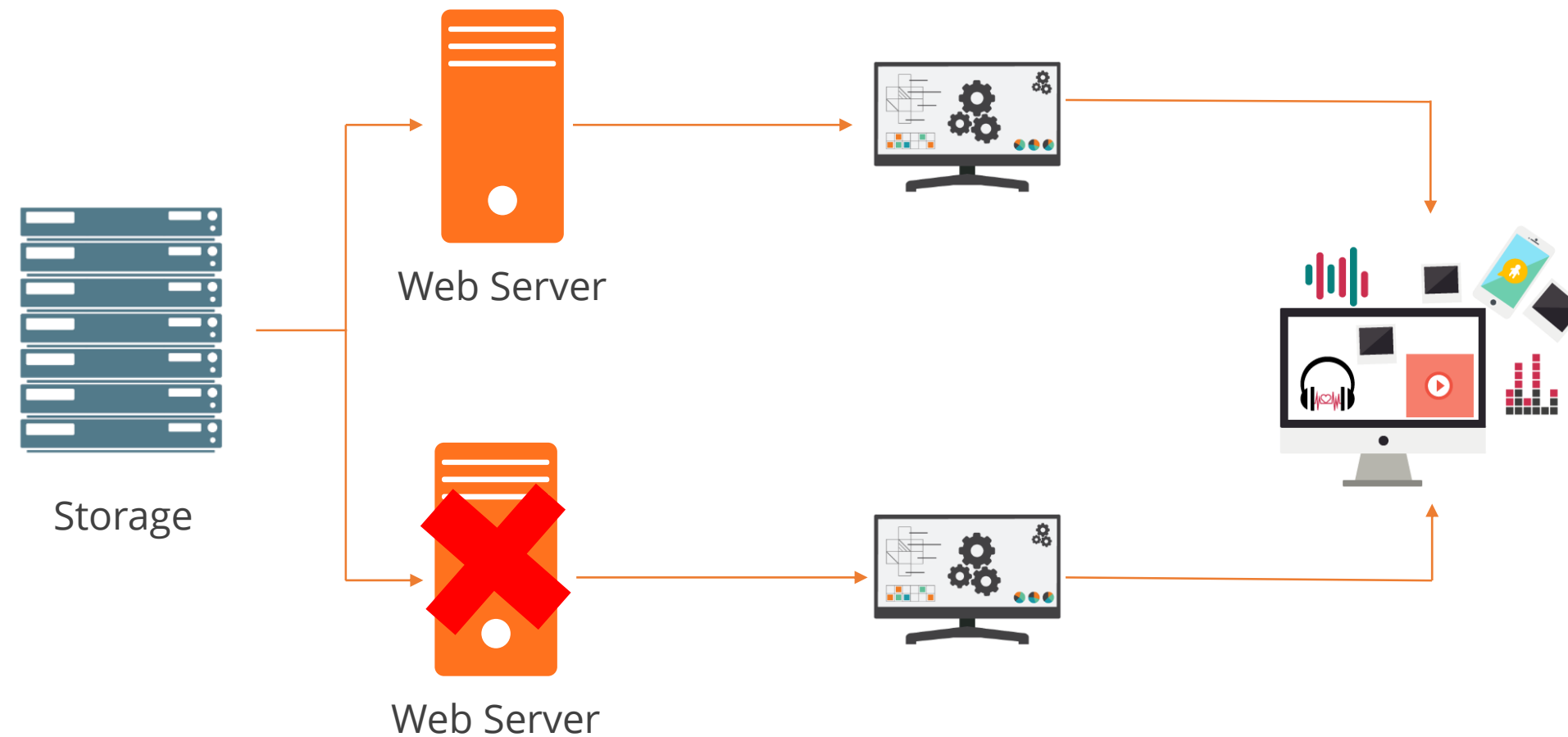
Well-Defined Interfaces

Ensure that components interact with each other through RESTful APIs.

Resources   API Gateway   Resources   API Gateway   Resources

Service Discovery

Loose coupling ensures services interact with each other without any prior knowledge of their existence.

Amazon Elastic Load Balancing
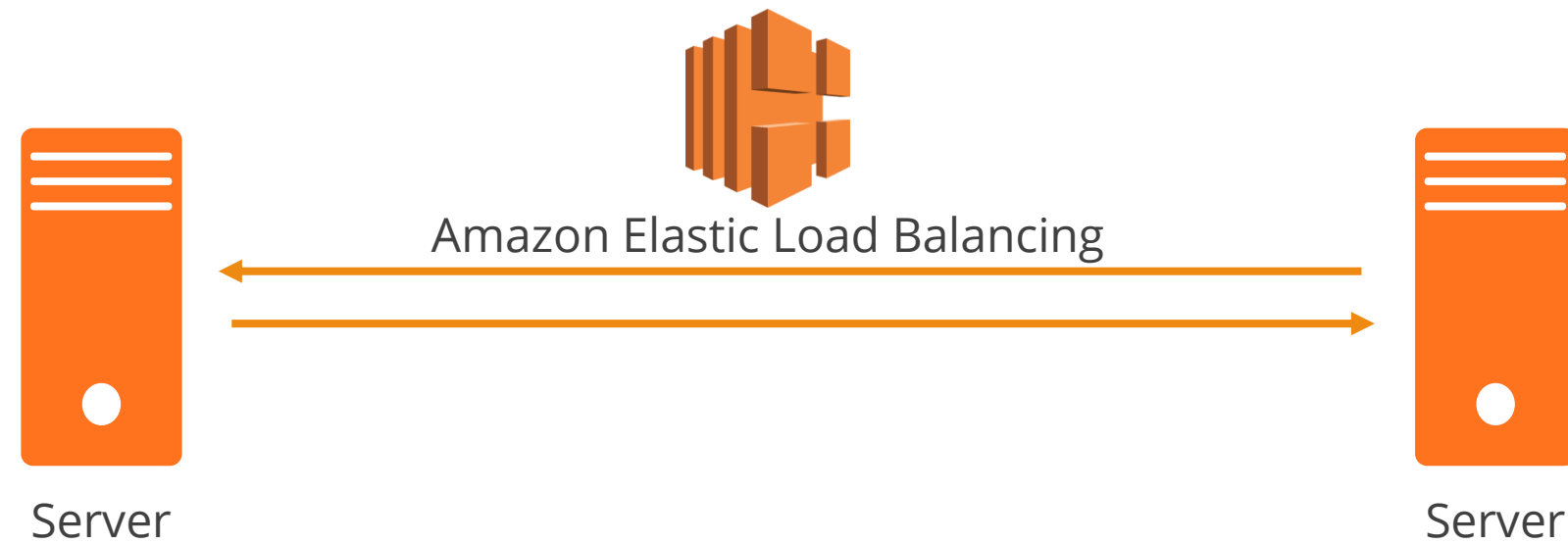
Server

Server

Asynchronous Integration

Asynchronous integration involves the use of an intermediate storage layer, such as an SQS queue. This approach means that when Server A completes its action, it sends a notification to SQS. This way the compute resources are decoupled and not directly linked to each other.

Resource A   Resource B   Resource C

Amazon SQS   Amazon SQS
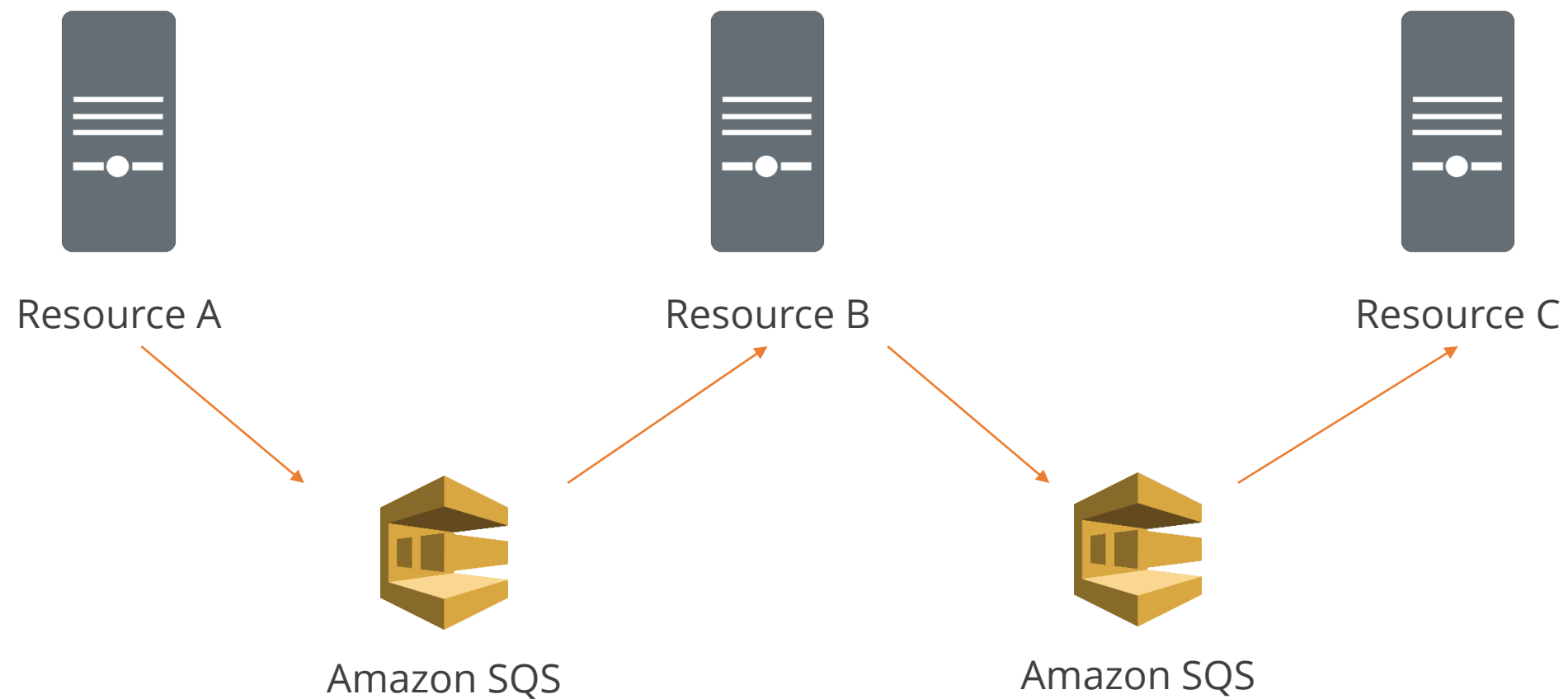
Graceful Failure

Applications are designed to fail gracefully, as this allows other resources to continue service without causing a complete outage.

Route 53

Primary Website

Backup Website

80
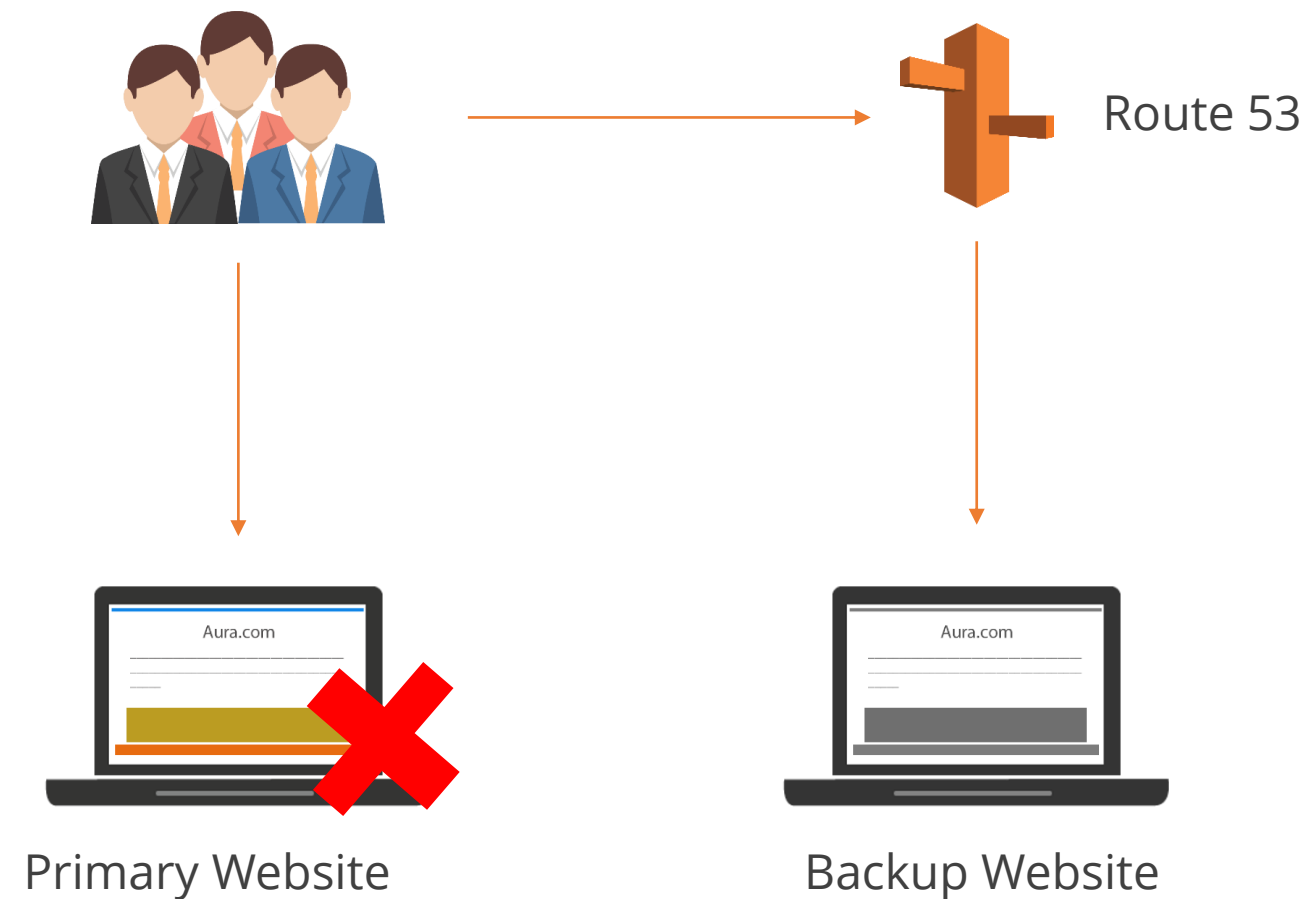
Services, Not Servers

To help your organization grow, AWS provides a suite of services to lower IT costs, such as:
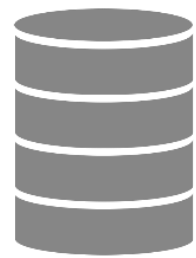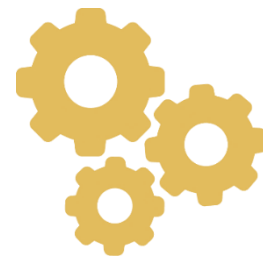
Database   Machine Learning   Analytics   Search   Email
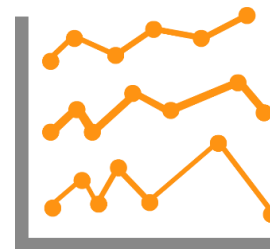
Databases

With AWS database, usage is not restricted by constraints on licensing, support capabilities, and hardware availabilities. AWS offers fully managed, easily scalable services such as:

DynamoDB   ElasticCache   RDS   Redshift   MySQL DB   OracleDB

82

Redundancy

Redundancy removes single points of failure from systems. This is achieved using multiple resources for the same tasks in standby or active mode.

Resource   Resource

Task
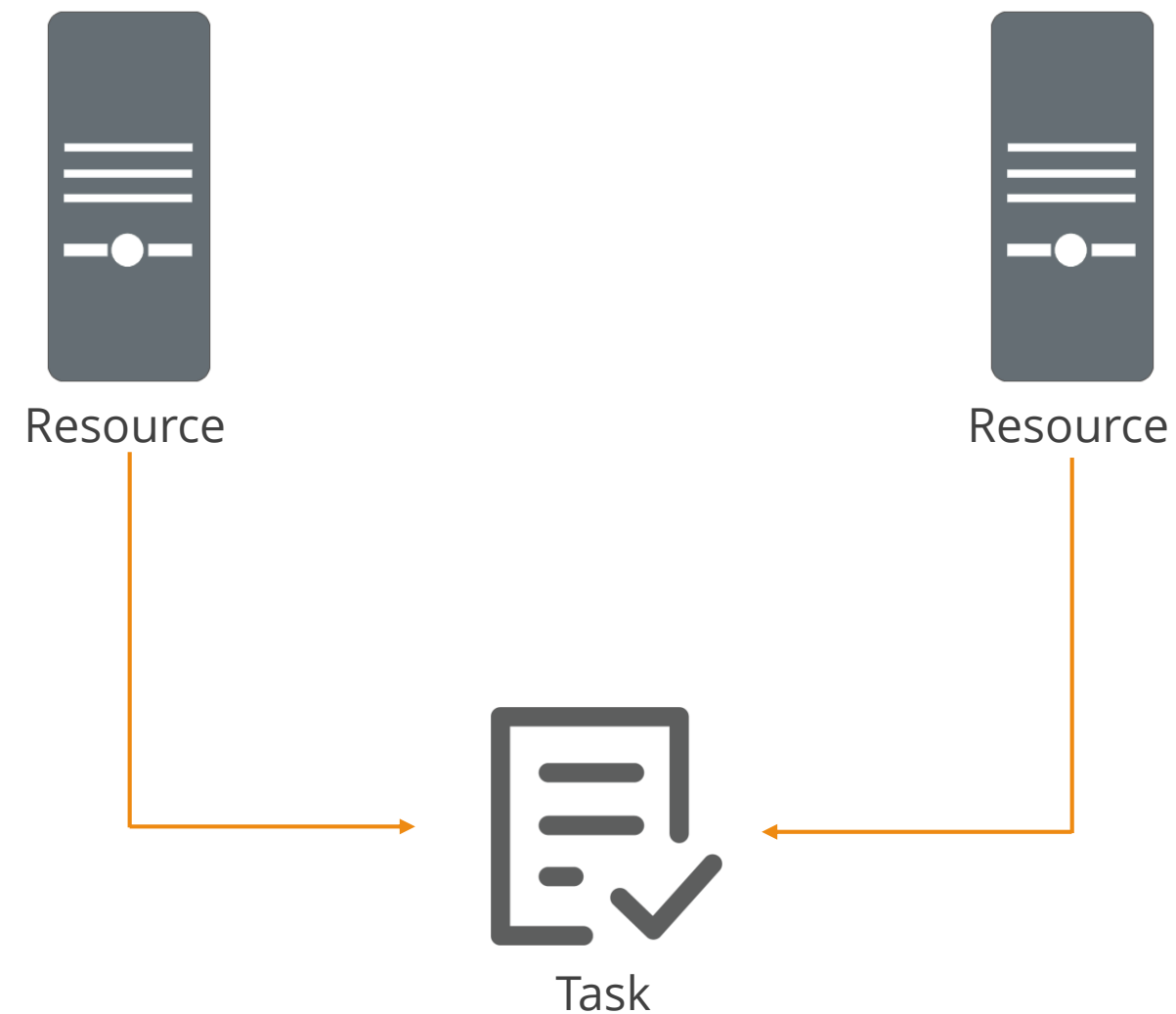
Standby Redundancy

Standby redundancy is used for stateful components like databases. The standby resource becomes the primary resource by "failing over" to the primary database.

Database   Standby Database

Active Redundancy

Active redundancy is where multiple redundant compute resources share requests and absorb the loss of one or more failing instances. For example, multiple web servers sitting behind a load balancer.

Elastic Load Balancer

Web Server   Web Server   Web Server

Task

85
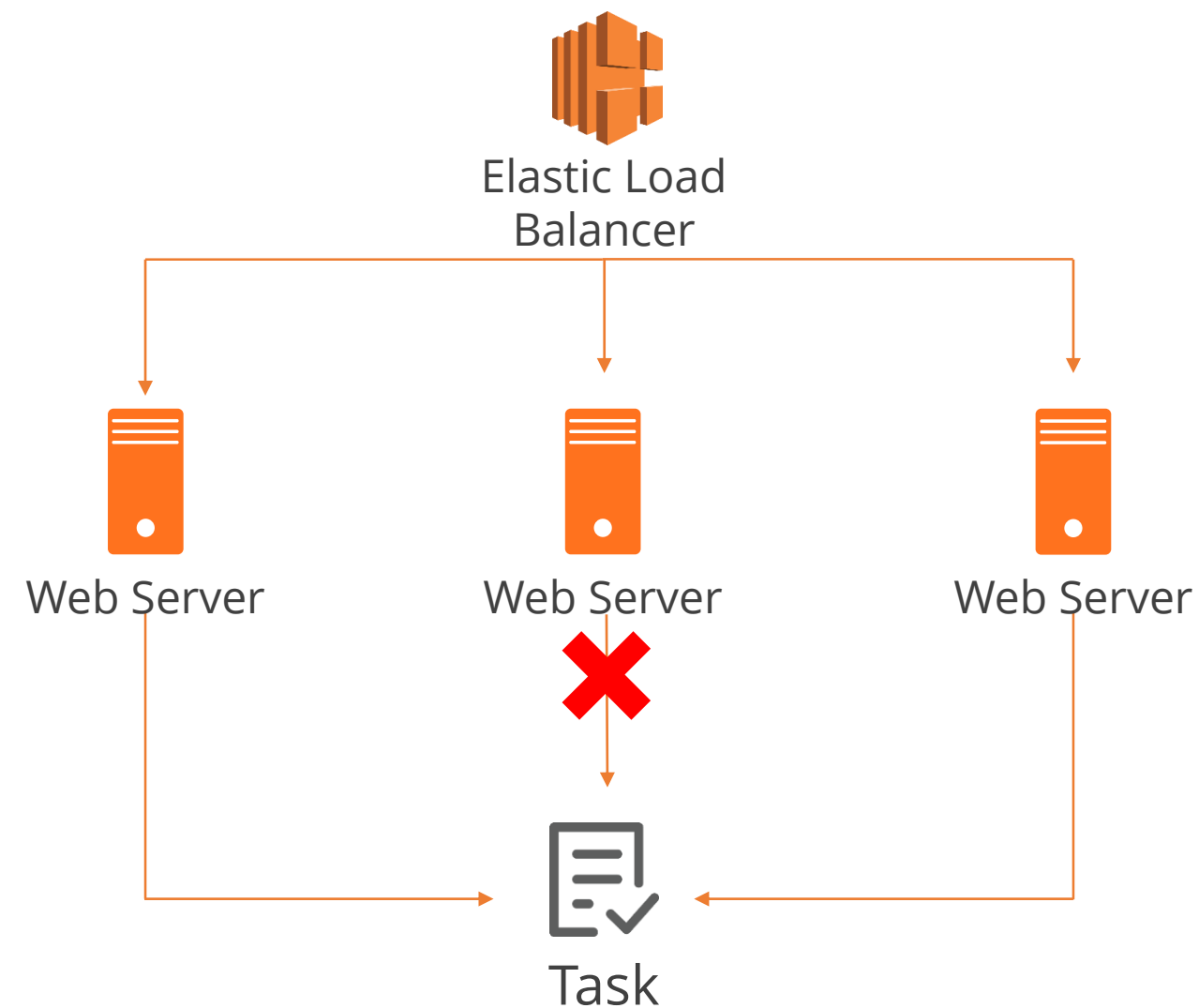
Failure Detection

Automatic failure detection allows you to react to an outage automatically without the need for manual intervention. Services like ELB and Route53 let you configure health checks to automatically route traffic to healthy resources.

Elastic Load Balancer

Health check
HTTP 200 - Yes

Resources

Health check
HTTP 200 - No

Resources

Launch new healthy resource

Resources

Health check
HTTP 200 - Yes

Resources

simpli·learn
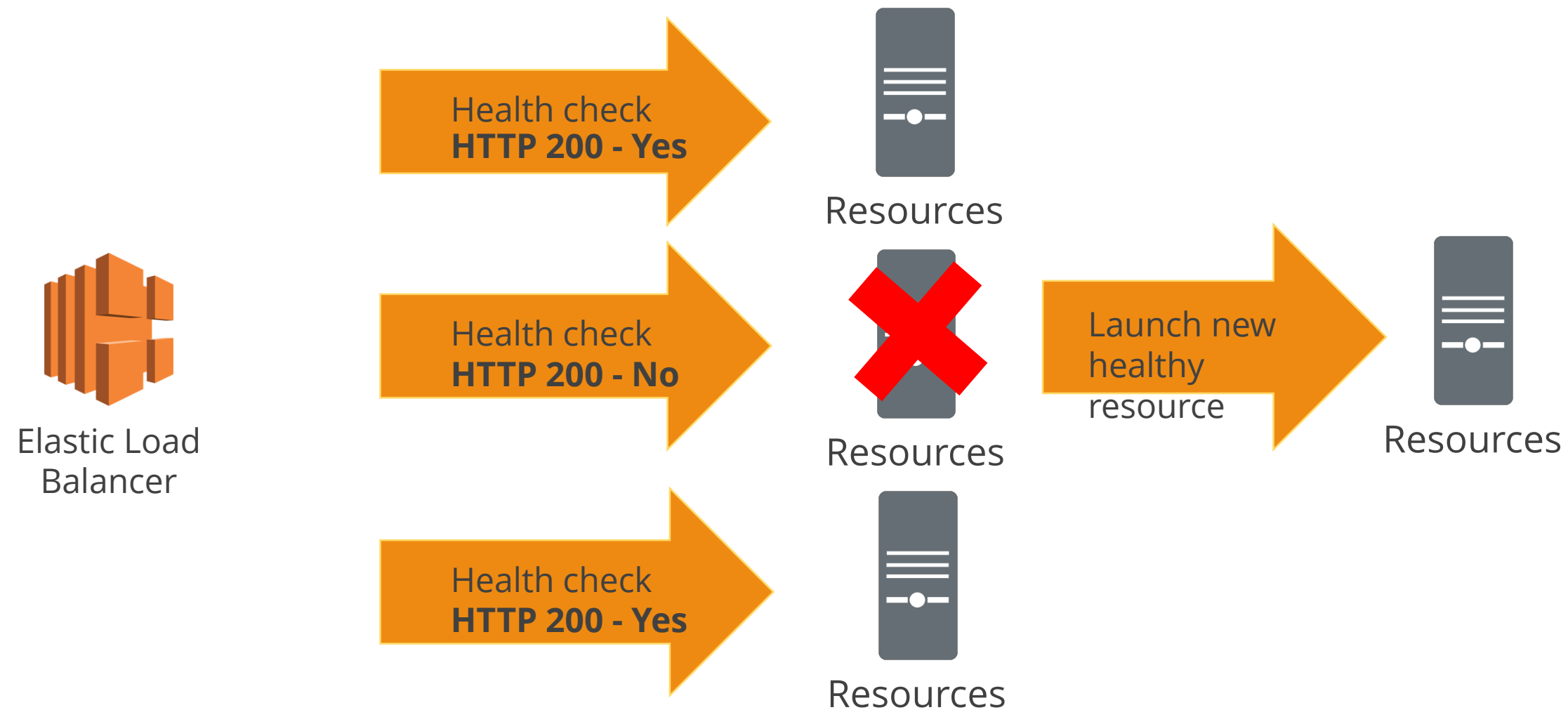
Durable Data Storage

Replicating your data to other sites or resources protects its availability and integrity. The three ways to replicate data are:

Synchronous

Asynchronous

Quorum Based
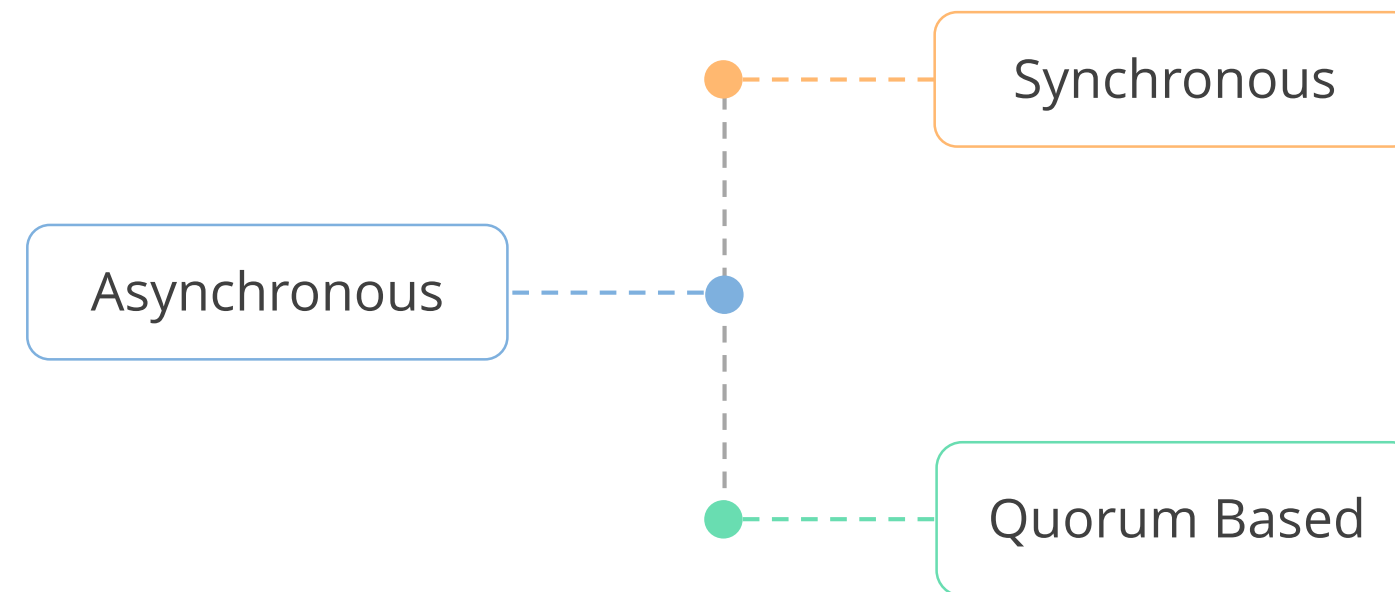
Synchronous Replication

Synchronous replication ensures that data has been durably stored on both the primary and replication locations. Any write operation will be acknowledged as complete when this has taken place.

Write operation →

Acknowledgement sent ←

Write operation acknowledged ←

User

Primary Resource

Replication Resource

Asynchronous Replication

Asynchronous replication decouples the primary node from the replications so a write operation doesn't wait for any acknowledgement.

Write operation

Acknowledgement sent

User

Primary Resource

Replication Resource

simpli·learn

Quorum Replication

Quorum-based replication is a mix of both synchronous and asynchronous replication.

Write operation

Acknowledgement sent

Write operation acknowledged

User

Primary Resource

Replication Resources
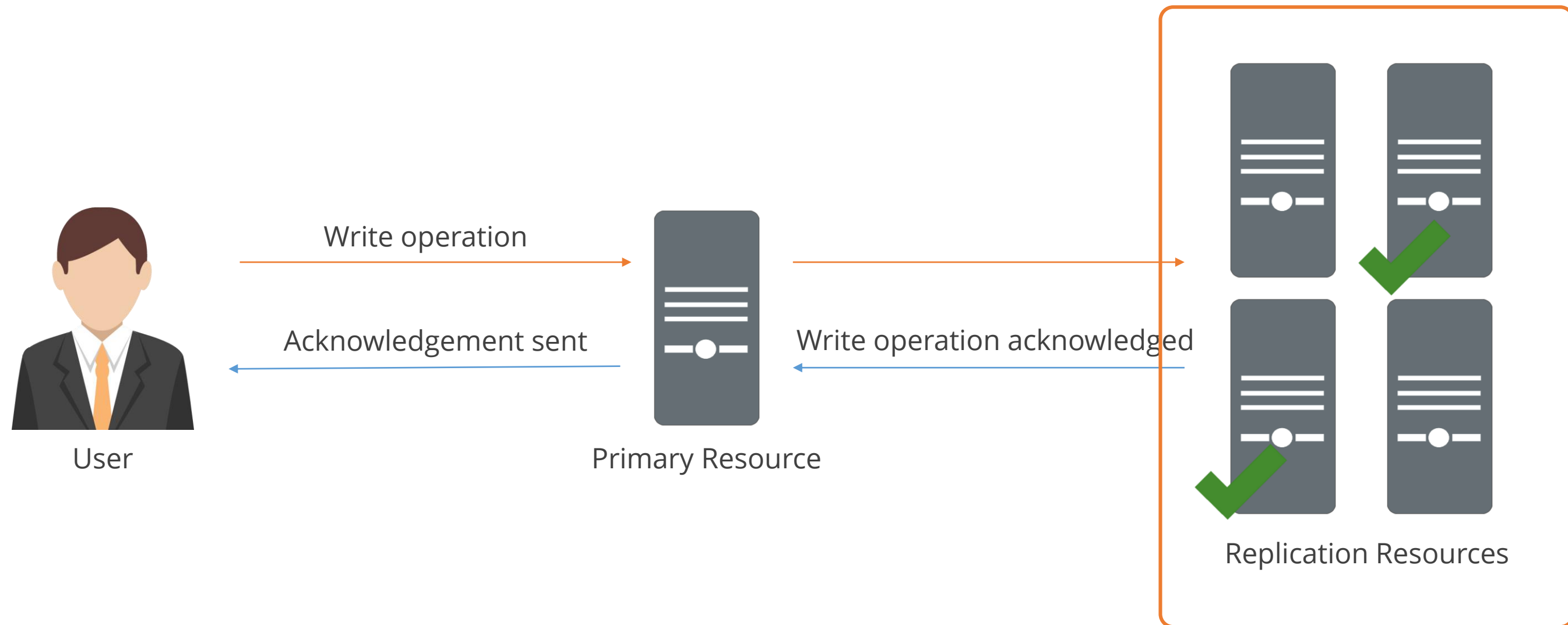
Data Center Resilience

A traditional data center failover involves failing over all your resources to a secondary distant data center. Due to the distance between the two data centers, synchronous replication is often impractical, slow, and involves data loss and as such is not tested very often.

Primary Datacenter   Secondary Datacenter

91

Data Center Resilience

AWS data centers are configured to provide multiple Availability Zones in each region with low latency network connectivity. This means replicating your data across data centers in a synchronous manner and as a result the failover becomes simpler.

Availability Zone

Availability Zone

Availability Zone

Region

Availability Zone A

Availability Zone B

Availability Zone C

92

simplilearn
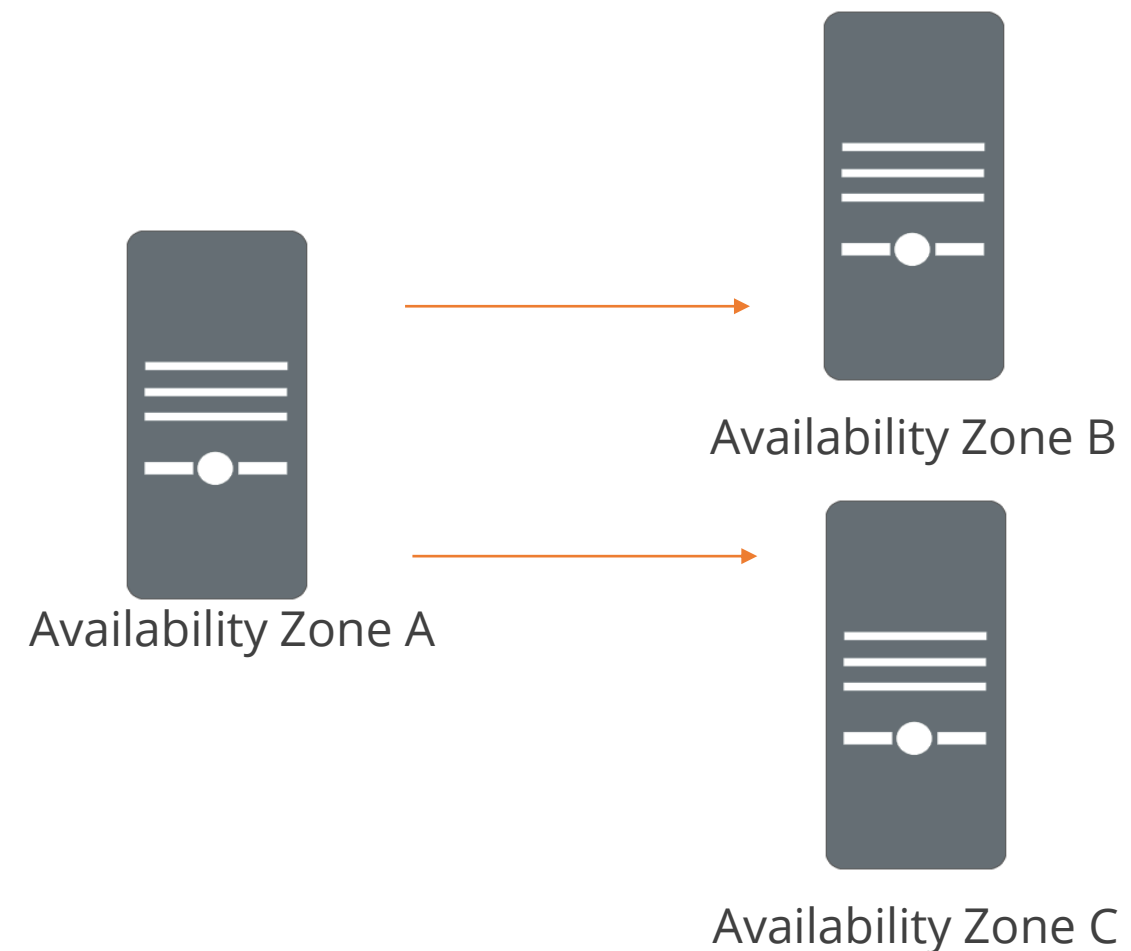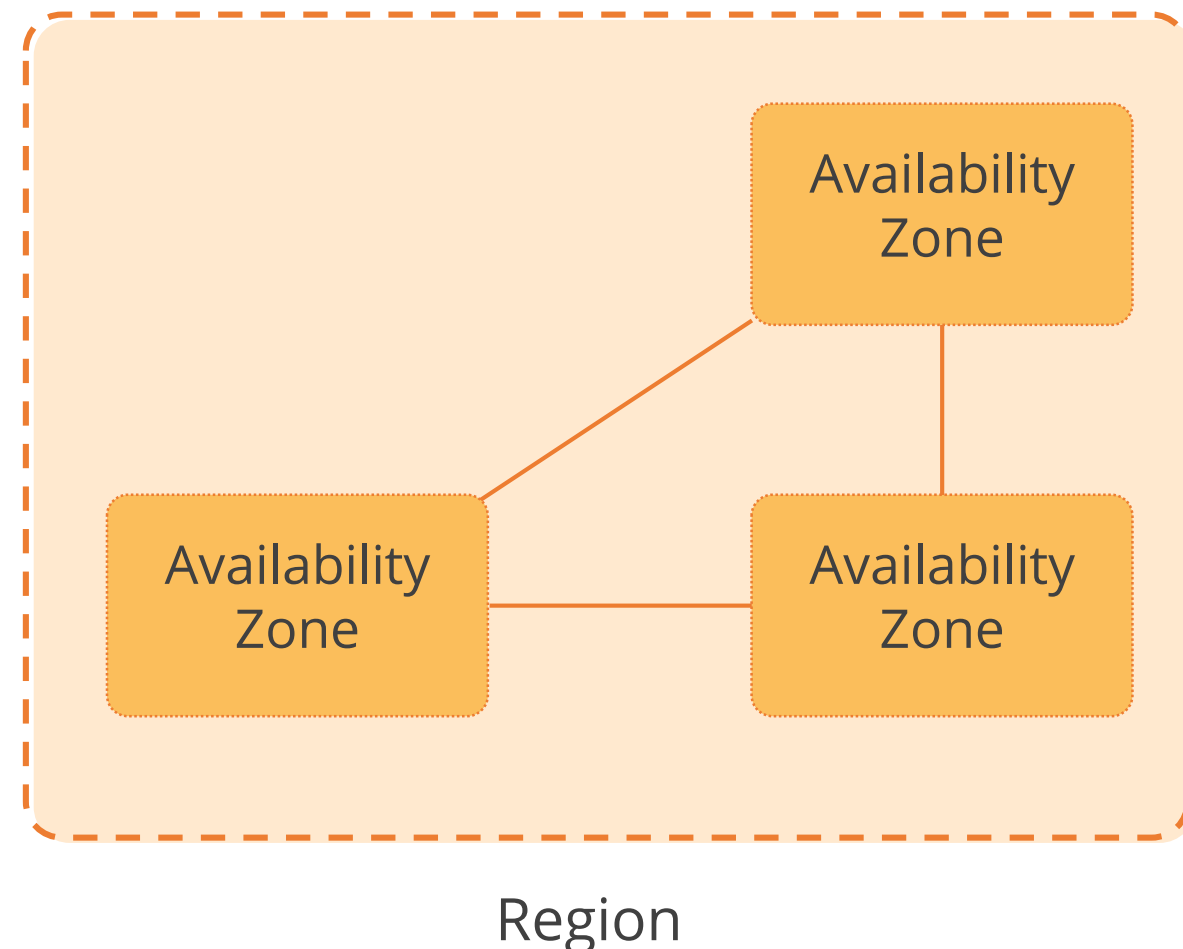
Shuffle Sharding/Fault Isolation

Shuffle Sharding is a practice that sends a few of the requests to some of the resources. This ensures if one shard of resources is infected or down, the other shard of resources will be up and running.
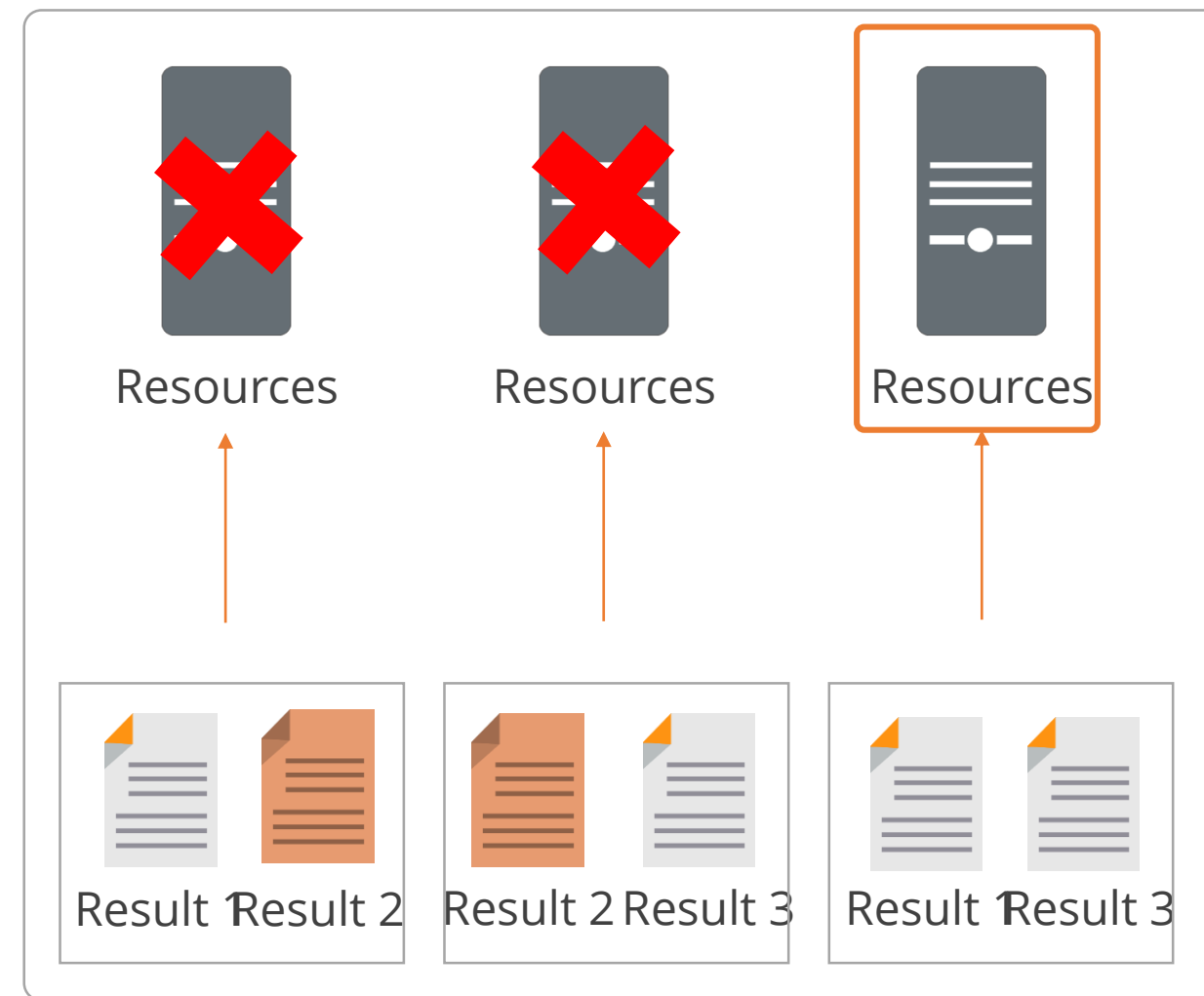
Optimize for Cost

AWS economies of scale offer organizations huge opportunities to make cost savings.

94

Right Sizing

AWS allows you to select the most cost effective resource and configuration to fit your requirements. A wide variety of instance types can be chosen from:

Amazon EC2

RDS

Redshift

Elastic Search

95

Elasticity

AWS offers many elasticity options to help you save money.
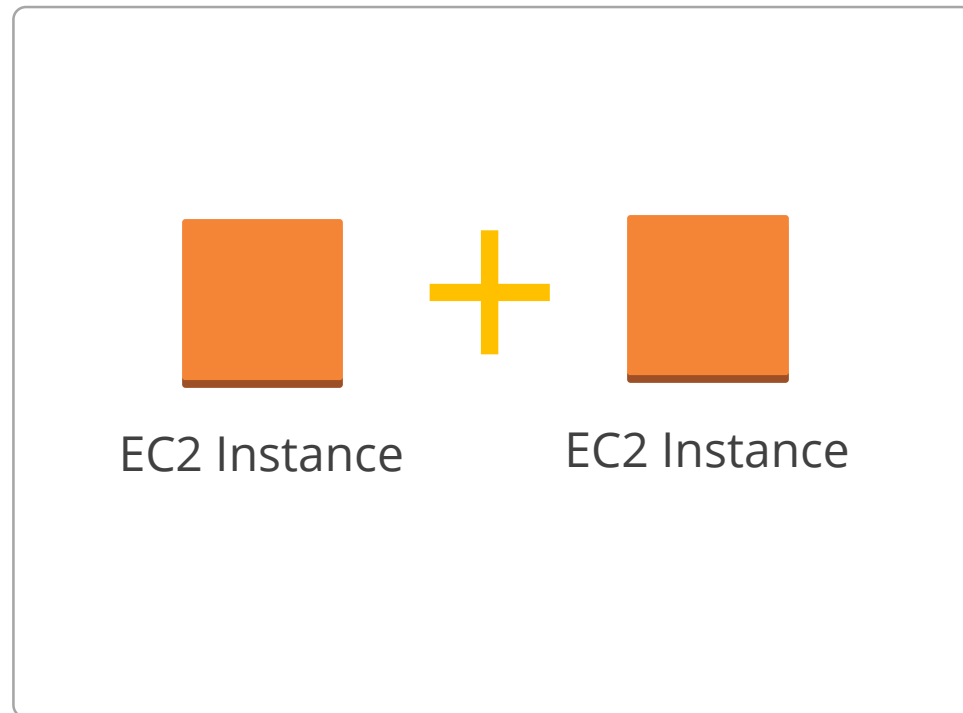
Horizontal Scaling
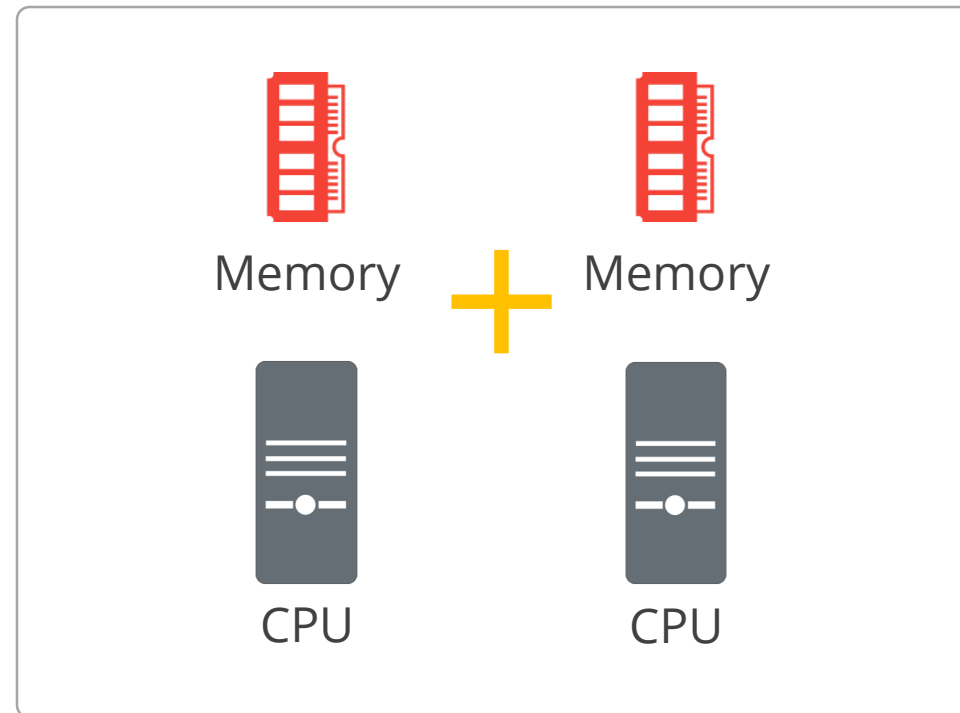
Vertical Scaling
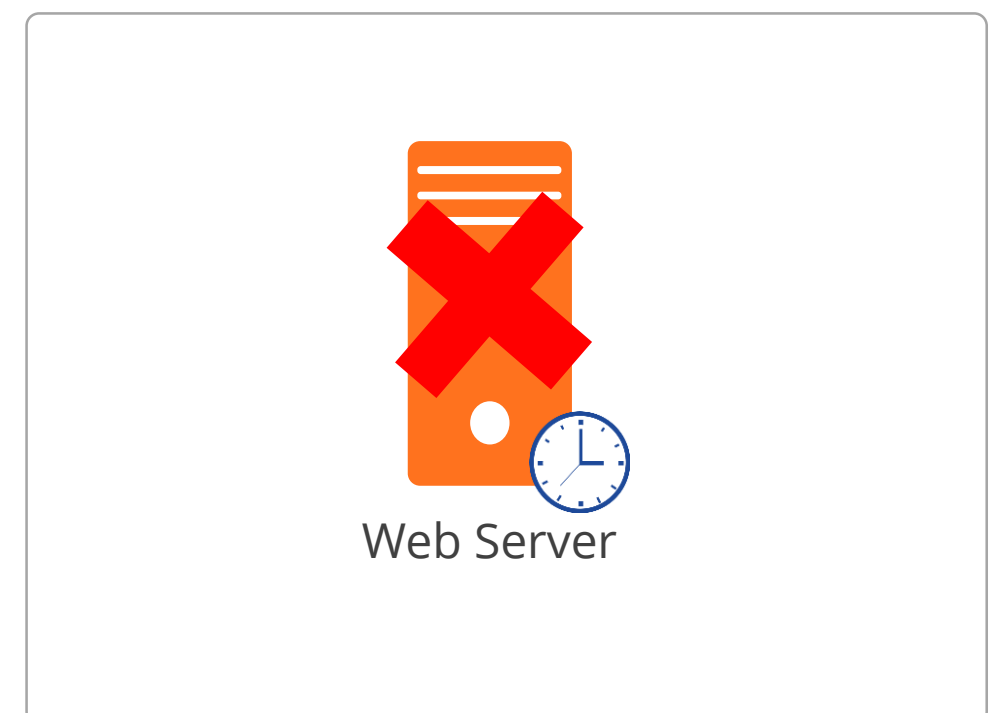
Shutting down non-production servers

96

Purchasing Options—On-demand

EC2 on-demand instance pricing means you only pay for what you use with no long term commitments.

97
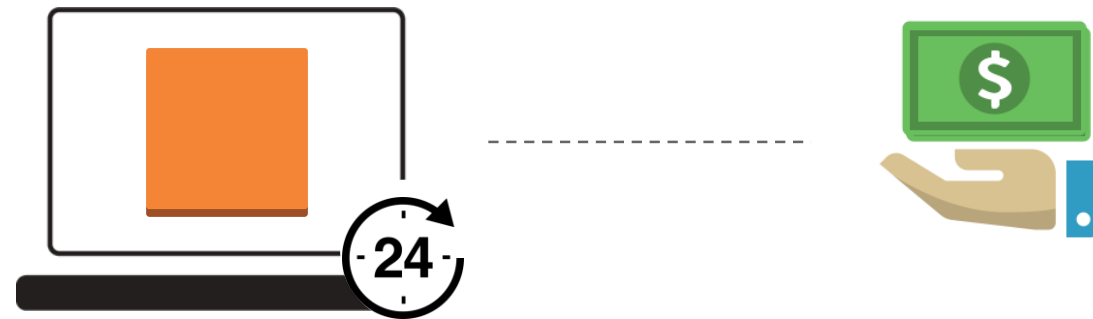
Purchasing Options—Reserved

AWS Trusted Advisor or AWS EC2 usage reports identify the resources that benefit from reserved capacity. Technically there is no difference between On-Demand EC2 instances and reserved instances. The only difference is the way you pay for it.

Purchasing Options—Spot

EC2 Spot Instances are ideal for workloads that have flexible start and end times as it allows bidding on spare EC2 computing capacity. Spot Instances are often available at significant discounts compared to on-demand pricing.

Spot Instance

Workloads
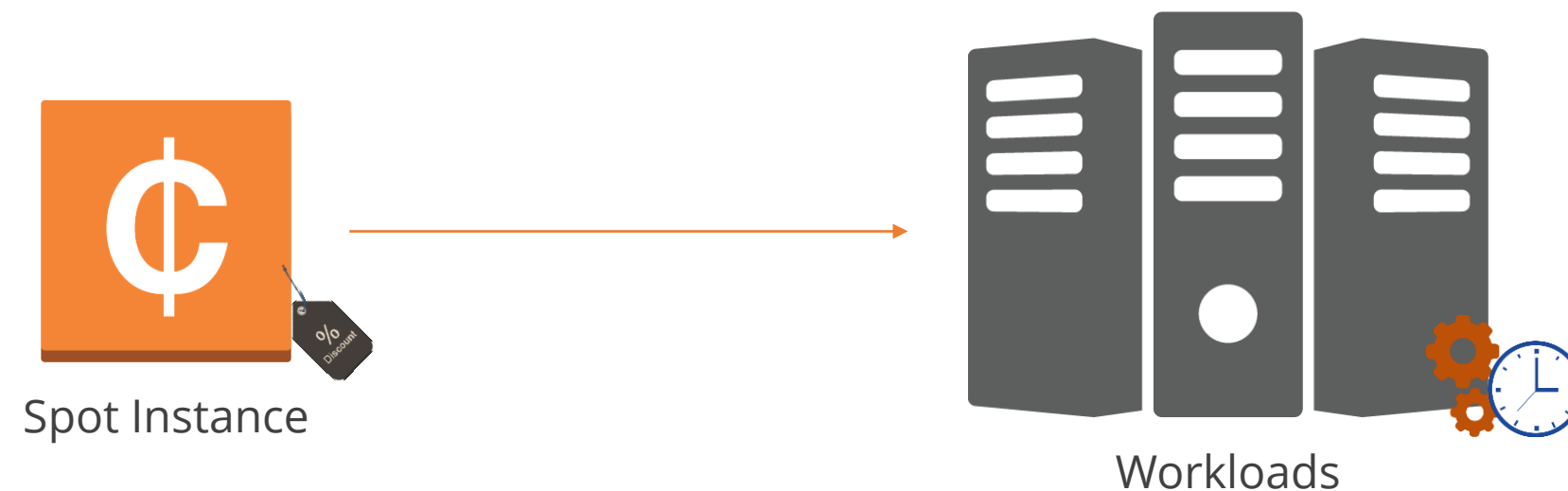
Purchasing Options—Spot Strategies

There are three strategies for Spot instances:

1 **Bidding**
Bid higher than spot market price to get cheaper overall price

2 **Mix Strategy**
Design applications that use a mixture of reserved, on-demand, and spot instances

3 **Spot Blocks**
Bid for fixed duration spot instances

100

Caching

Caching data means storing previously calculated data for future use so you don't have to recalculate it. There are two approaches:

1 **Application Caching**
- Applications store and retrieve information from fast, managed, in-memory caches. This way an application can look for results in the cache first, and if the data isn't there it can then calculate or retrieve the data and store it in the cache for subsequent requests.
- Amazon ElastiCache is a service that provides an in-memory cache in the cloud.

2 **Edge Caching**
- Static content such as images, videos, and dynamic content such as live video can be cached around the world using Edge Locations. This way users are served the content that is closest to them and it results in low latency response times.
- The principle applies to both downloading and uploading data.
- An example of Edge caching is Amazon CloudFront (CDN).

simpli·learn

Security

AWS offers you a range of products and services to ensure the security of your resources, such as:

Defense in Depth — Add multiple layers of protection to your resources.

Reduce Privileged Access — Give the users only that access which they require.

Security as Code — Capture all your security requirements in one script that you can deploy in new environments.

Real Time Auditing — Test and audit your environment in real time.

Knowledge Check

simplilearn

A Stateless application is one that _____.

a. requires knowledge of previous interactions but stores no session information

b. needs no knowledge of previous interactions and stores no session information

c. requires knowledge of previous interactions and stores session information

d. needs no knowledge of previous interactions but stores session information

A Stateless application is one that _____.

a. requires knowledge of previous interactions but stores no session information

b. needs no knowledge of previous interactions and stores no session information

c. requires knowledge of previous interactions and stores session information

d. needs no knowledge of previous interactions but stores session information

The correct answer is **b)**

A stateless application is one that needs no knowledge of previous interactions and stores no session information. For example, a webserver that provides the same web page to any end user.

Loose Coupling is desirable because _____.

a. it reduces the cost of your AWS resources

b. it stores previously calculated data for future use

c. it means the failure of one or more resources does not result in a service outage

d. it assists you to select resource and configuration to fit your requirements

KNOWLEDGE CHECK	Loose Coupling is desirable because _____.

a. it reduces the cost of your AWS resources

b. it stores previously calculated data for future use

c. it means the failure of one or more resources does not result in a service outage

d. it assists you to select resource and configuration to fit your requirements

The correct answer is **c)**

Applications should be designed so that they can be broken into smaller, loosely coupled components. The desired outcome is that a failure in one component should not cause other components to fail.

simpli learn

KNOWLEDGE CHECK

The three EC2 purchasing options that make cloud computing unique are _____.

a. On-Request, Auction, and Reserved pricing

b. On-Demand, Spot, and Permanent pricing

c. On-Request, Local, and Permanent pricing

d. On-Demand, Spot, and Reserved pricing

The three EC2 purchasing options that make cloud computing unique are _____.

a. On-Request, Auction, and Reserved pricing

b. On-Demand, Spot, and Permanent pricing

c. On-Request, Local, and Permanent pricing

d. On-Demand, Spot, and Reserved pricing

The correct answer is **d)**

EC2 on-demand instance pricing means you only pay for what you use with no long term commitments. Reserved enables you to commit to a defined period of 12-36 months to receive significantly discounted hourly rates compared to on-demand pricing. EC2 Spot Instances are ideal for workloads that have flexible start and end time as you are allowed to bid on spare EC2 computing capacity.

Monitoring and Logging

Overview of the tools available to enable AWS monitoring and logging

simplilearn

Monitoring and Logging

In this section you'll learn about:

1 — Amazon CloudWatch

Amazon CloudTrail — **2**

3 — Amazon Config
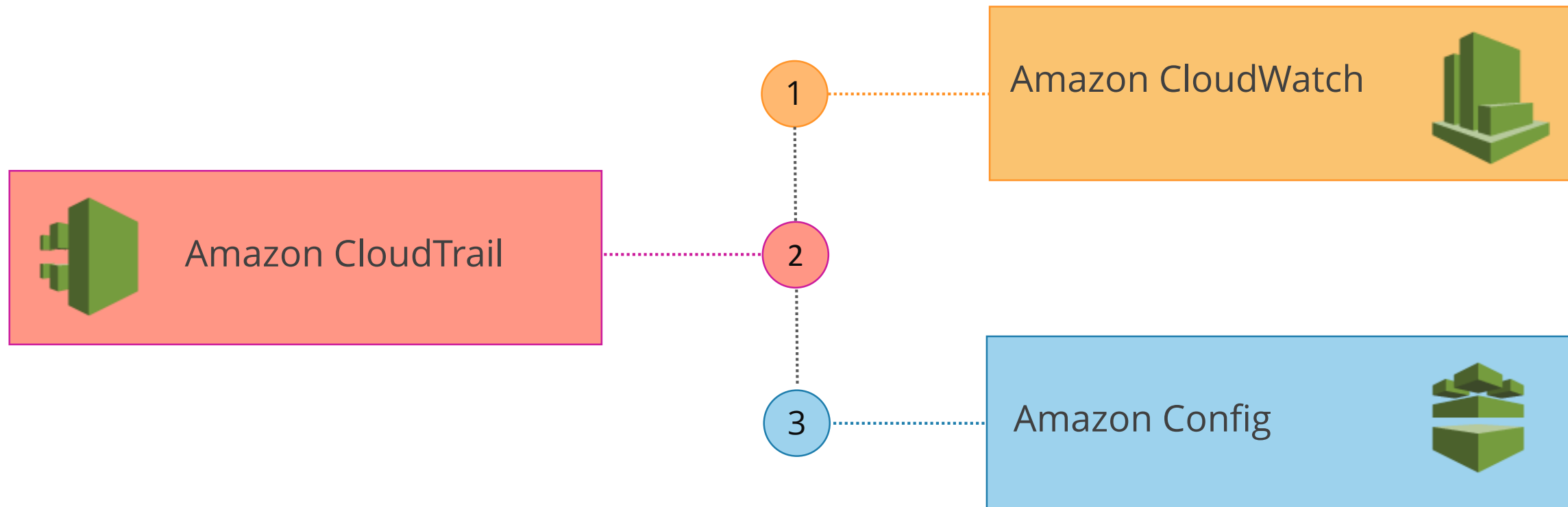
Amazon CloudWatch

With Amazon CloudWatch you can:

Monitor Amazon Web Services (AWS) resources and applications

1

Collect and track metrics

2

Use Alarms to send notifications

3

Automatically make changes to monitored resources based on defined rules

4

112

Amazon CloudWatch Events

Amazon CloudWatch Events deliver a stream of system events which alert about changes to AWS resources.

1 Alerts are sent to services such as AWS Lambda, Amazon SNS, Amazon SQS, and Amazon Kinesis Streams.

2 CloudWatch Events can be used to schedule events such as snapshot creation or instance reboot.

3 In addition to monitoring the built-in metrics, you can monitor your own custom metrics.

simpli·learn

Amazon CloudWatch Logs

Amazon CloudWatch Logs are used to monitor, store, and access application or system log files from Amazon EC2 instances, AWS CloudTrail, or other sources.

Store

Monitor

Access

CloudWatch Logs

114
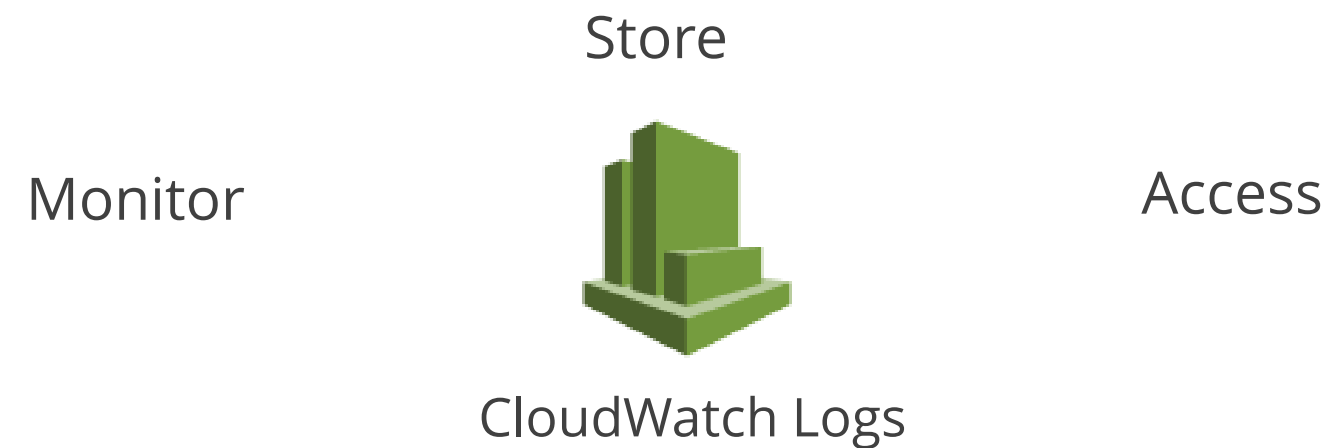
Amazon CloudWatch Limits

CloudWatch has limits for metrics, events, and logs. Some of the key limits are:

10 CloudWatch metrics per customer per month for free

10 alarms per customer per month for free

1,000,000 API requests per customer per month for free

1,000 Amazon SNS email notifications per customer per month for free

Unlimited number of custom metrics

Up to 5000 alarms per AWS account

Metric data is stored for two weeks

115

simplilearn
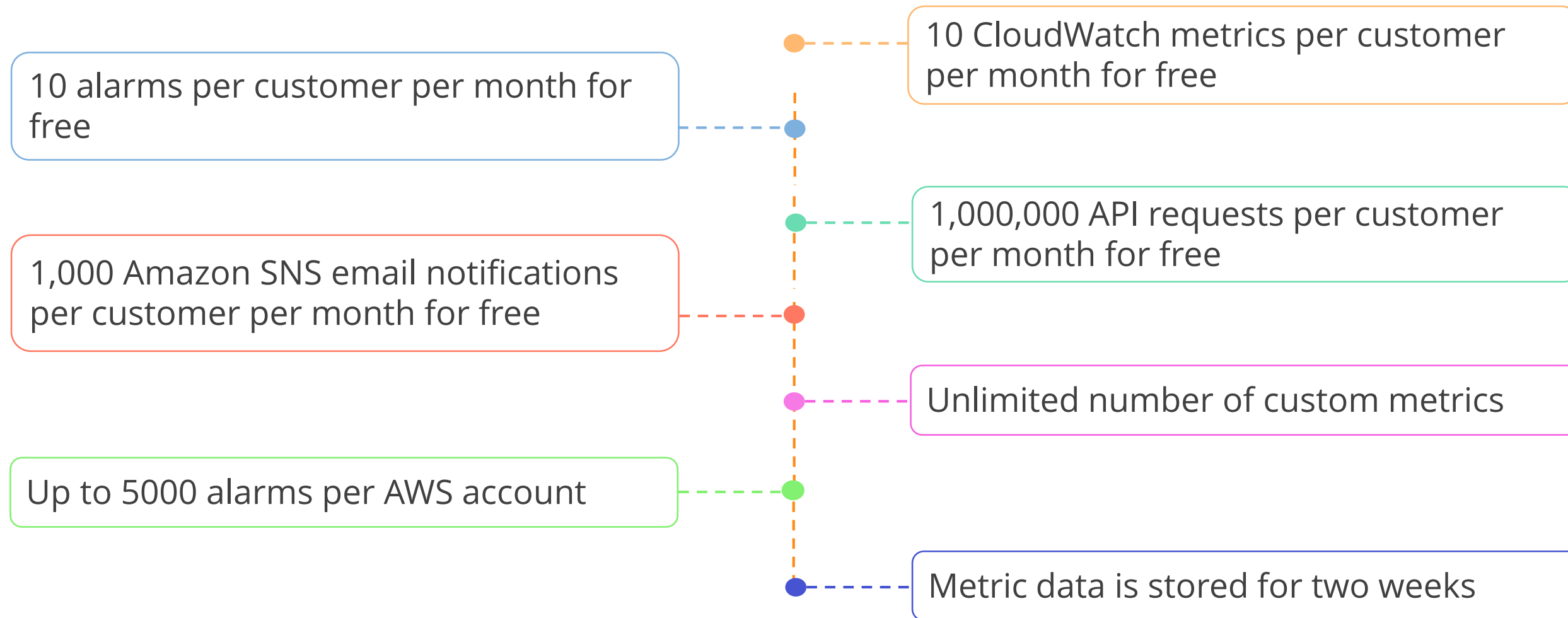
Amazon CloudTrail

Amazon CloudTrail is a web service that

CloudTrail

Records AWS API calls for your account

Stores and delivers log files for you

Provides a history of AWS API calls for your account

116

simplilearn
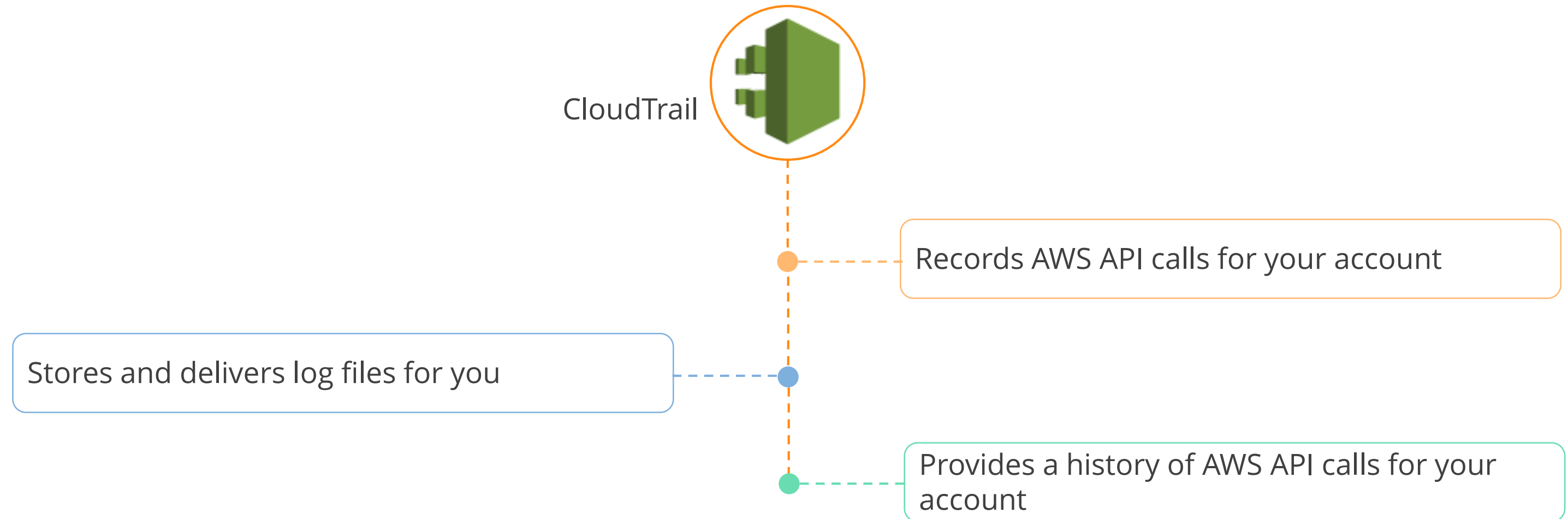
Increased Visibility

Amazon CloudTrail provides increased visibility. It helps you answer questions such as:

?
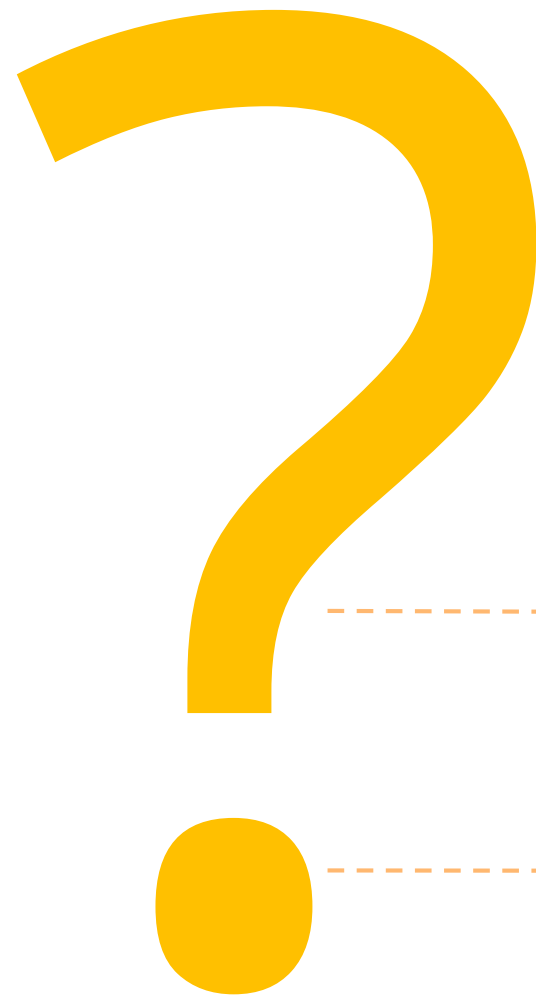
What actions did a given user take over a given time period?

Which user has taken actions on a given resource over a given time period?

What is the source IP address of a given activity?

Which activities failed due to inadequate permissions?

simpli learn

Durable and Inexpensive Log File Storage

CloudTrail uses Amazon S3 for log file storage and delivery, so log files are stored durably and inexpensively. You can use Amazon S3 lifecycle configuration rules to further reduce storage costs.

CloudTrail Log Files   Amazon S3   Log Files Amazon Glacier

Easy Administration

CloudTrail is a fully managed service. No installation is required; simply turn on CloudTrail for your account and start receiving CloudTrail log files in the Amazon S3 bucket that you specify.

CloudTrail

Log Files   S3 bucket

119
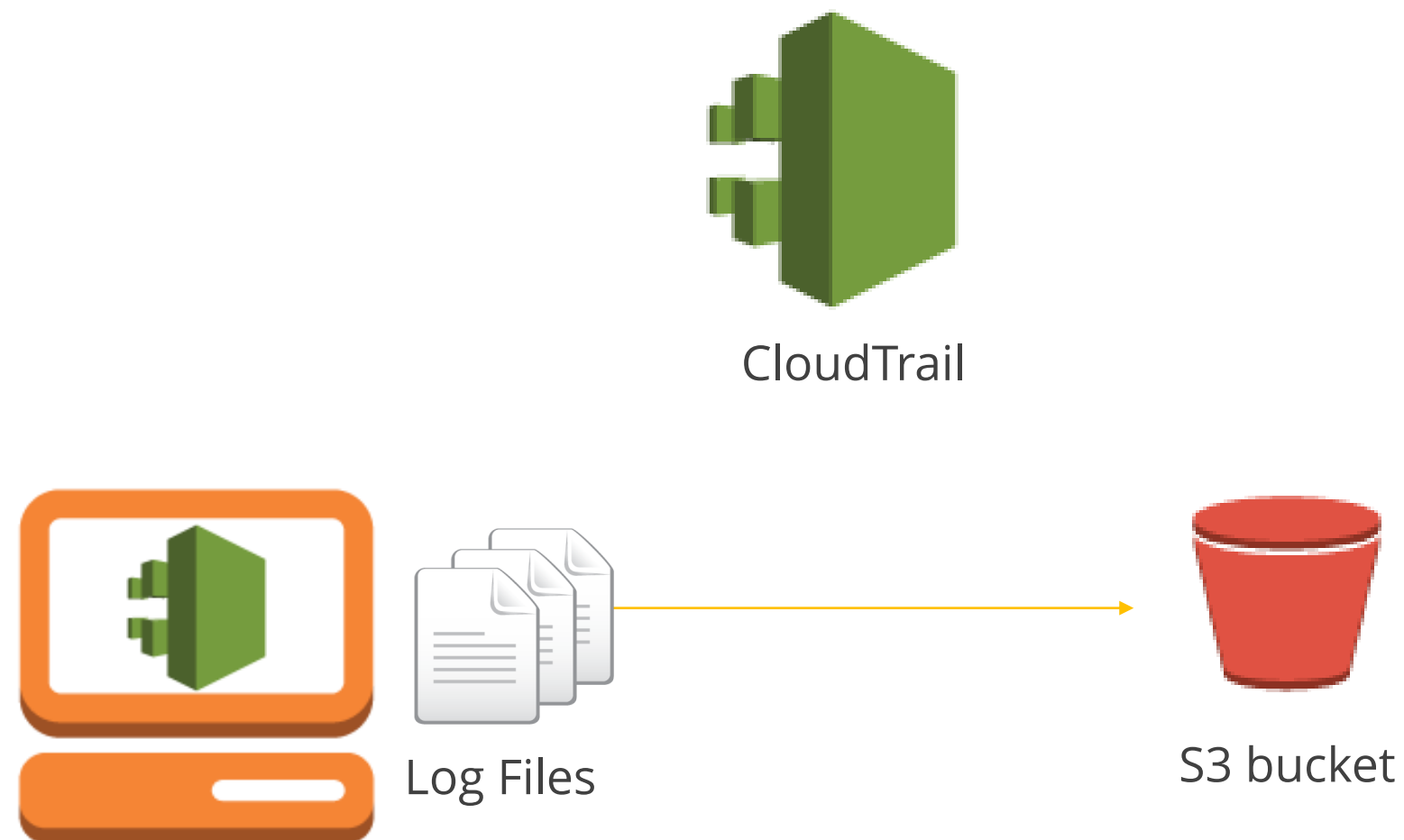
Notifications for Log File Delivery

CloudTrail uses the Amazon Simple Notification Service (SNS) to notify you when a new log file is delivered or a specific event has occurred.

CloudTrail   →   Amazon SNS   →   Security Audit Team
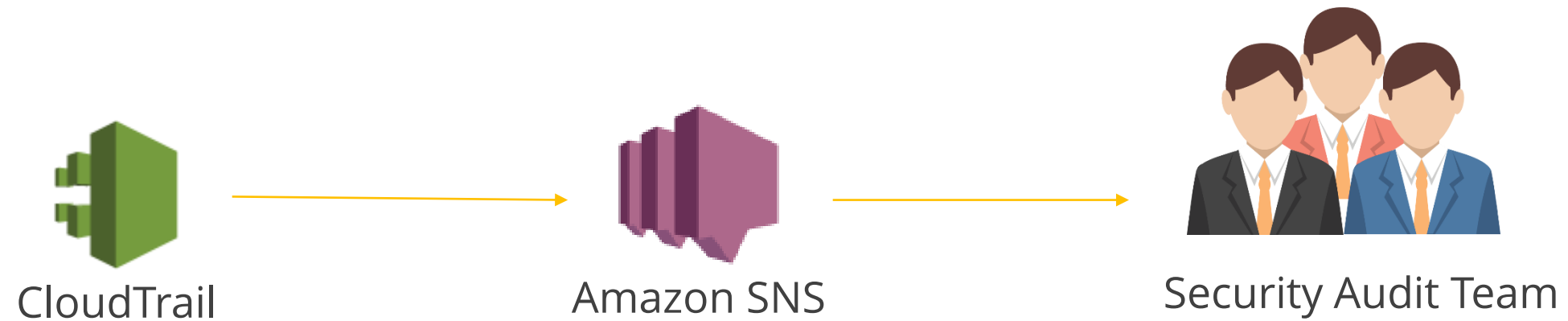
3rd Party Integration

AlertLogic, Boundary, Loggly, Splunk, and Sumologic are some of the companies that offer integrated solutions to analyze CloudTrail log files.

CloudTrail Log Files

121
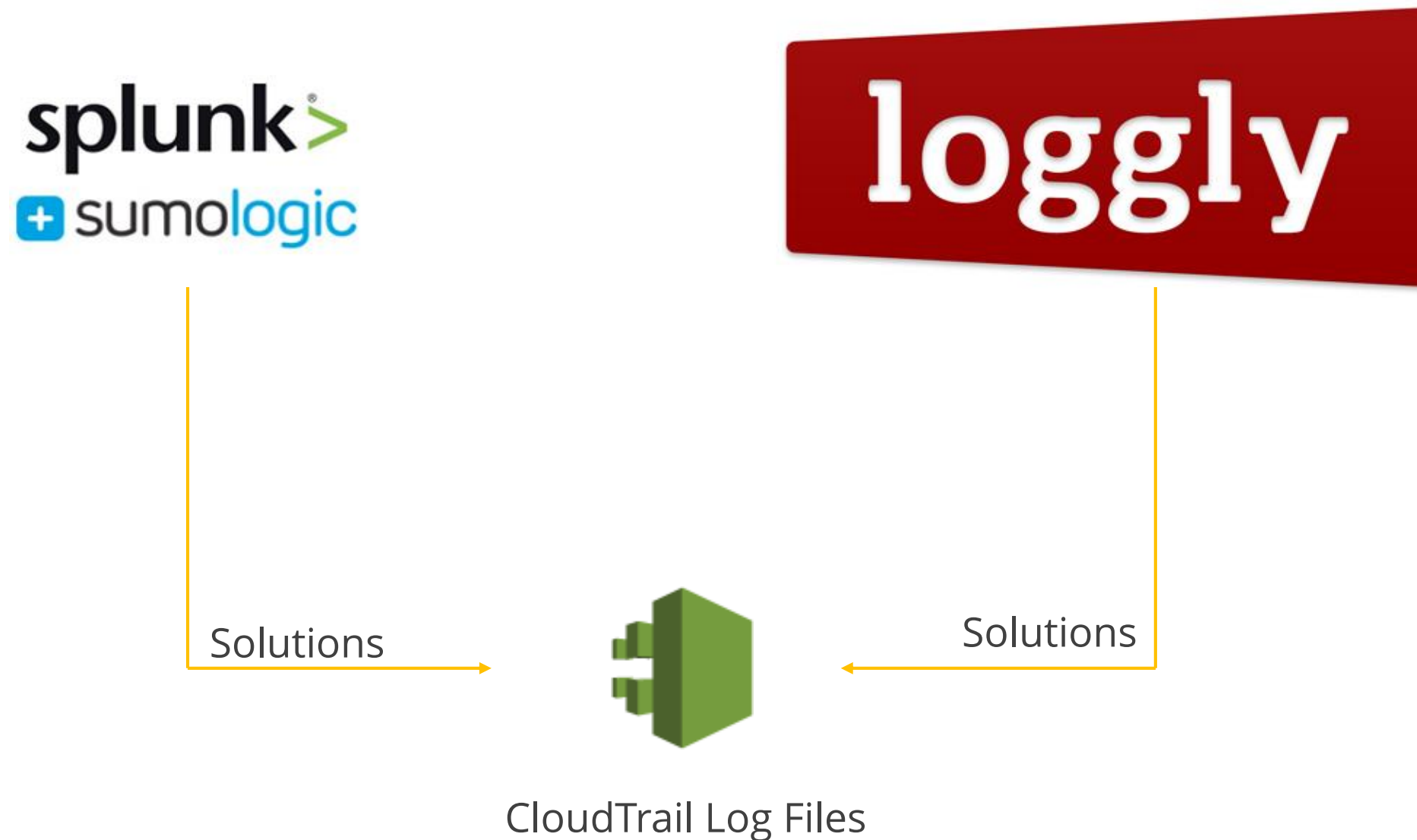
Log File Aggregation

CloudTrail can be configured to aggregate log files across multiple accounts and regions so that log files are delivered to a single bucket.
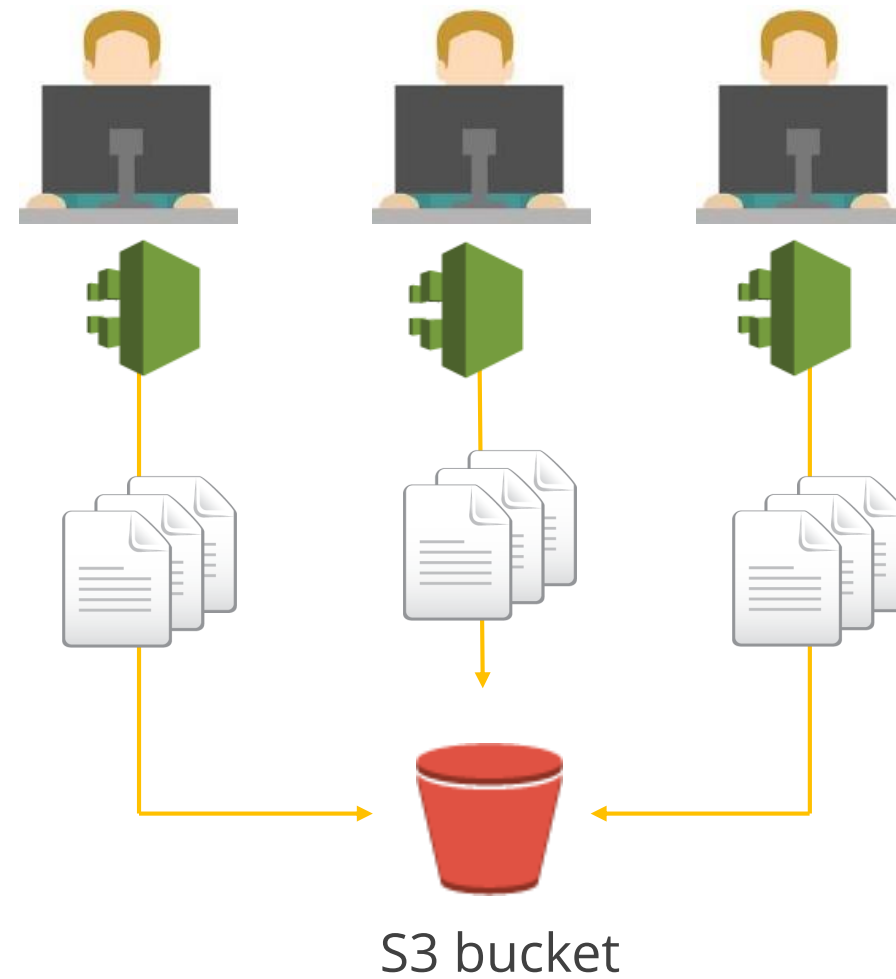
S3 bucket

Encrypted Log Files

CloudTrail encrypts all log files delivered to the specified Amazon S3 bucket using Amazon S3 server-side encryption (SSE). Log files can be further secured by using the AWS Key Management Service (KMS) keys.

CloudTrail

Log File

Amazon S3 bucket

Log File Integrity Validation

You can validate the integrity of CloudTrail log files stored in your Amazon S3 bucket and detect whether the log files were unchanged, modified, or deleted since CloudTrail delivered them.

CloudTrail Log File
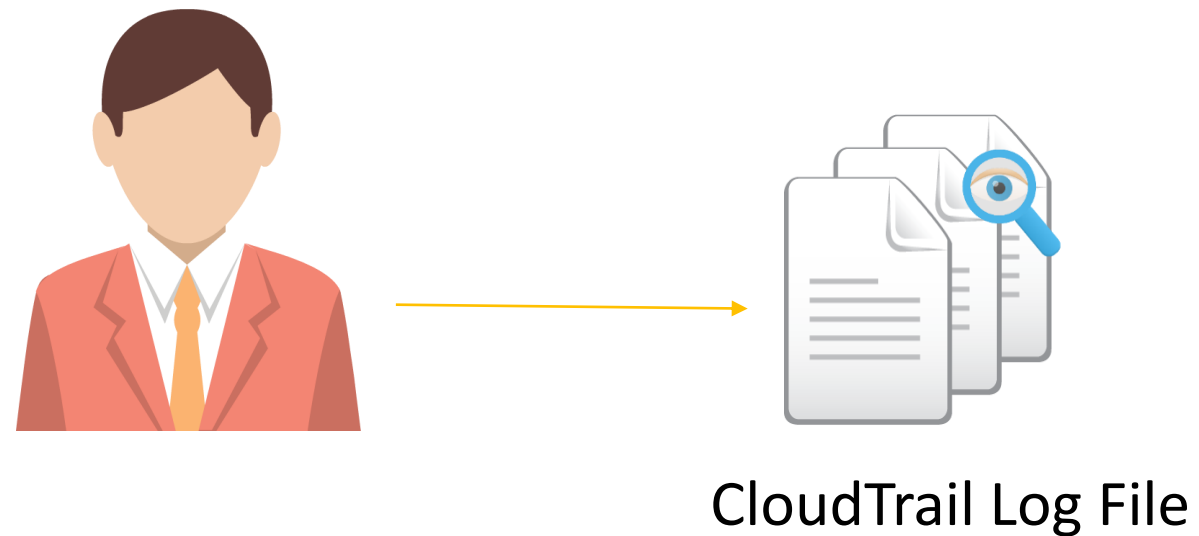
AWS Config

AWS Config is a fully managed service that provides an AWS:

AWS Config

Resource Inventory

Configuration History

Configuration Change Notifications

125

simpli|learn
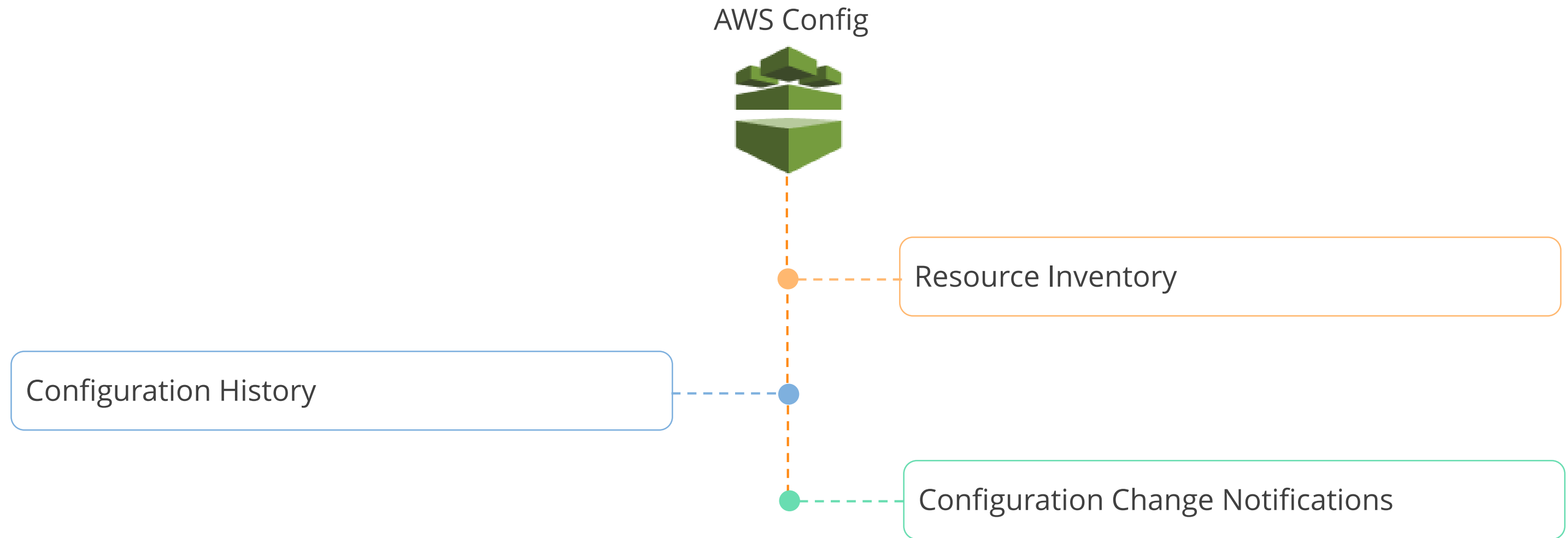
AWS Config

Using Config Rules, an IT Administrator can quickly determine when and how a resource went out of compliance. For example, it ensures EBS volumes are encrypted, EC2 instances are tagged, and Elastic IP addresses (EIPs) are attached to instances.

Admin

Encrypted EBS Volumes

Tagged EC2 Instances

EIP attached Instances

Tagged
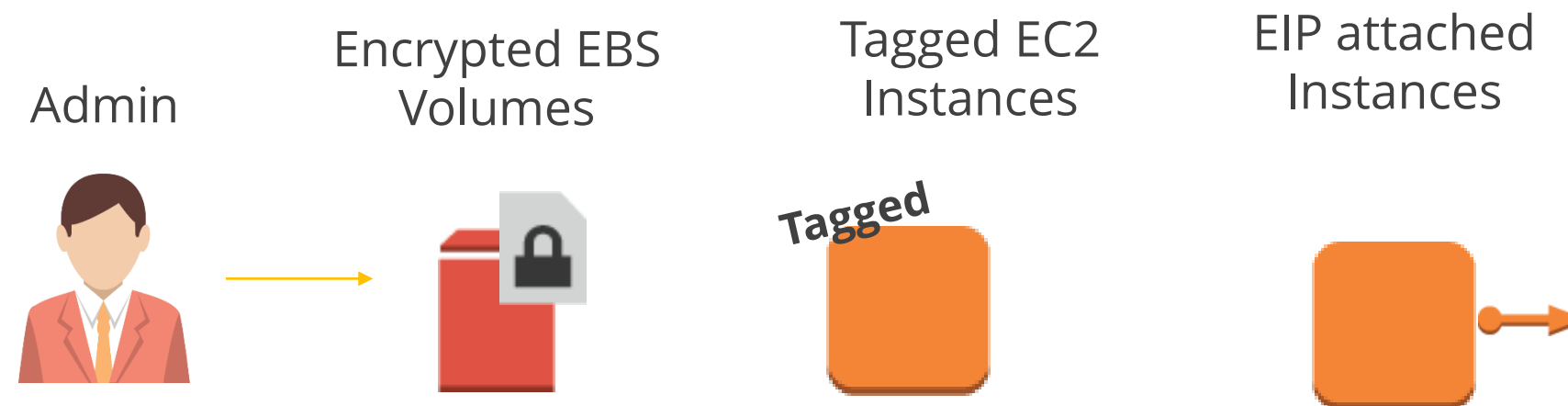
Configuration Visibility

View all the configuration attributes of your AWS resources in real time. Amazon Simple Notification Service (SNS) will notify you of any updated configuration or specific changes from the previous state.

AWS Config

Resources

SNS
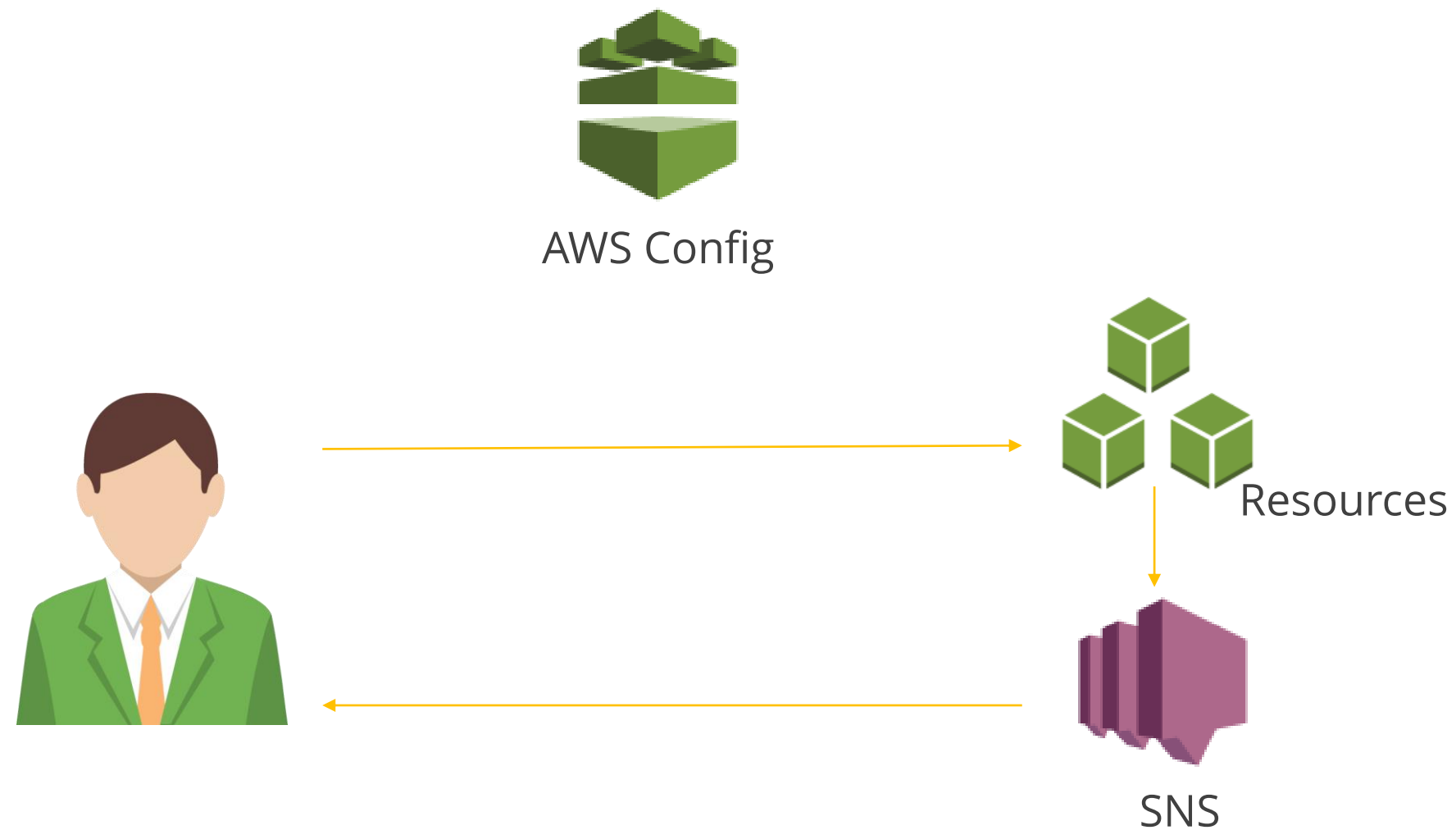
Continuous Assessment

Assess the overall compliance of your AWS resource configurations based on your organization's policies and guidelines.

Cloud Governance Dashboard

AWS Config Rules give you a visual dashboard with lists, charts, and graphs to help you quickly spot non-compliant resources and take appropriate action.

Knowledge Check

What services assist you with the monitoring and logging of your cloud environment?

a. Amazon CloudFront, Amazon CloudFormation, and Amazon Trusted Advisor

b. Amazon CloudWatch, Amazon CloudFormation, and Amazon CloudTrail

c. Amazon CloudWatch, Amazon CloudTrail, and Amazon Config

d. Amazon CloudFront, Amazon CloudFormation, and Amazon Trusted Advisor

What services assist you with the monitoring and logging of your cloud environment?

a. Amazon CloudFront, Amazon CloudFormation, and Amazon Trusted Advisor

b. Amazon CloudWatch, Amazon CloudFormation, and Amazon CloudTrail

c. Amazon CloudWatch, Amazon CloudTrail, and Amazon Config

d. Amazon CloudFront, Amazon CloudFormation, and Amazon Trusted Advisor

The correct answer is **c)**

Amazon CloudWatch, Amazon CloudTrail, and Amazon Config are the managed services that provide monitoring and logging of your cloud environment.

KNOWLEDGE CHECK	Which tool would you use to monitor AWS resource and performance utilization?

a.   Amazon CloudTrail

b.   Amazon CloudWatch

c.   Amazon Config

d.   Amazon CloudFront

Which tool would you use to monitor AWS resource and performance utilization?

a. Amazon CloudTrail

b. Amazon CloudWatch

c. Amazon Config

d. Amazon CloudFront

The correct answer is **b)**

Amazon CloudWatch monitors your Amazon Web Service (AWS) resources and the applications in real-time in a particular region. Amazon CloudTrail records AWS API calls for your account. Amazon Config reports on configuration changes made to your AWS resources. Amazon CloudFront is the Amazon CDN service.

Hybrid IT architectures

Overview of the tools and functionality available to run hybrid cloud architectures

simplilearn

Hybrid IT Architectures

In this section you'll learn the following areas of hybrid architectures:

Networking and how to link onsite infrastructure with AWS

Common controls for security and access

Integrate data and control lifecycle management

Available tools for integrated resources and deployment management

136
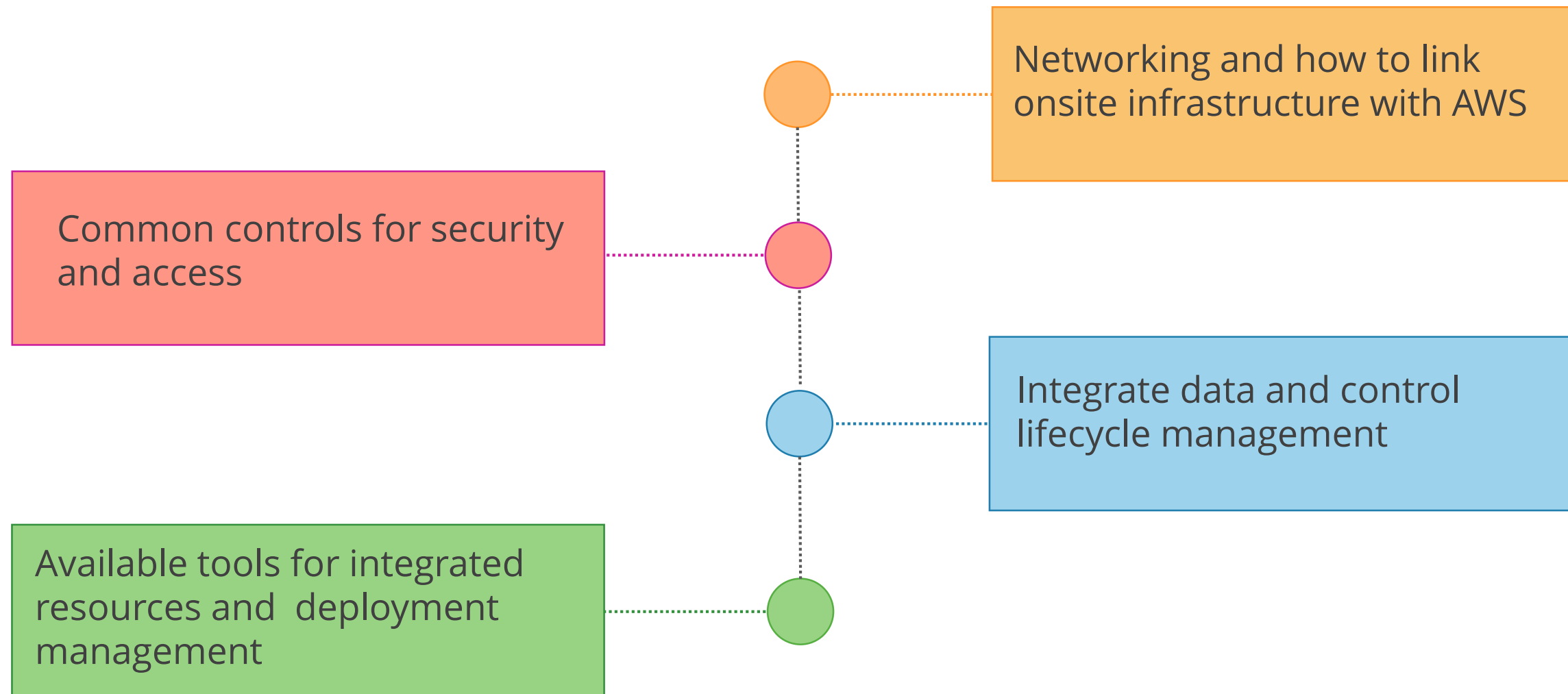
Network

Extending the existing on-premises network configuration onto your AWS virtual private cloud ensures your AWS resources operate as a part of your existing network. Amazon VPC is a logistically isolated network in the Amazon cloud that gives you complete control over a virtual networking environment.
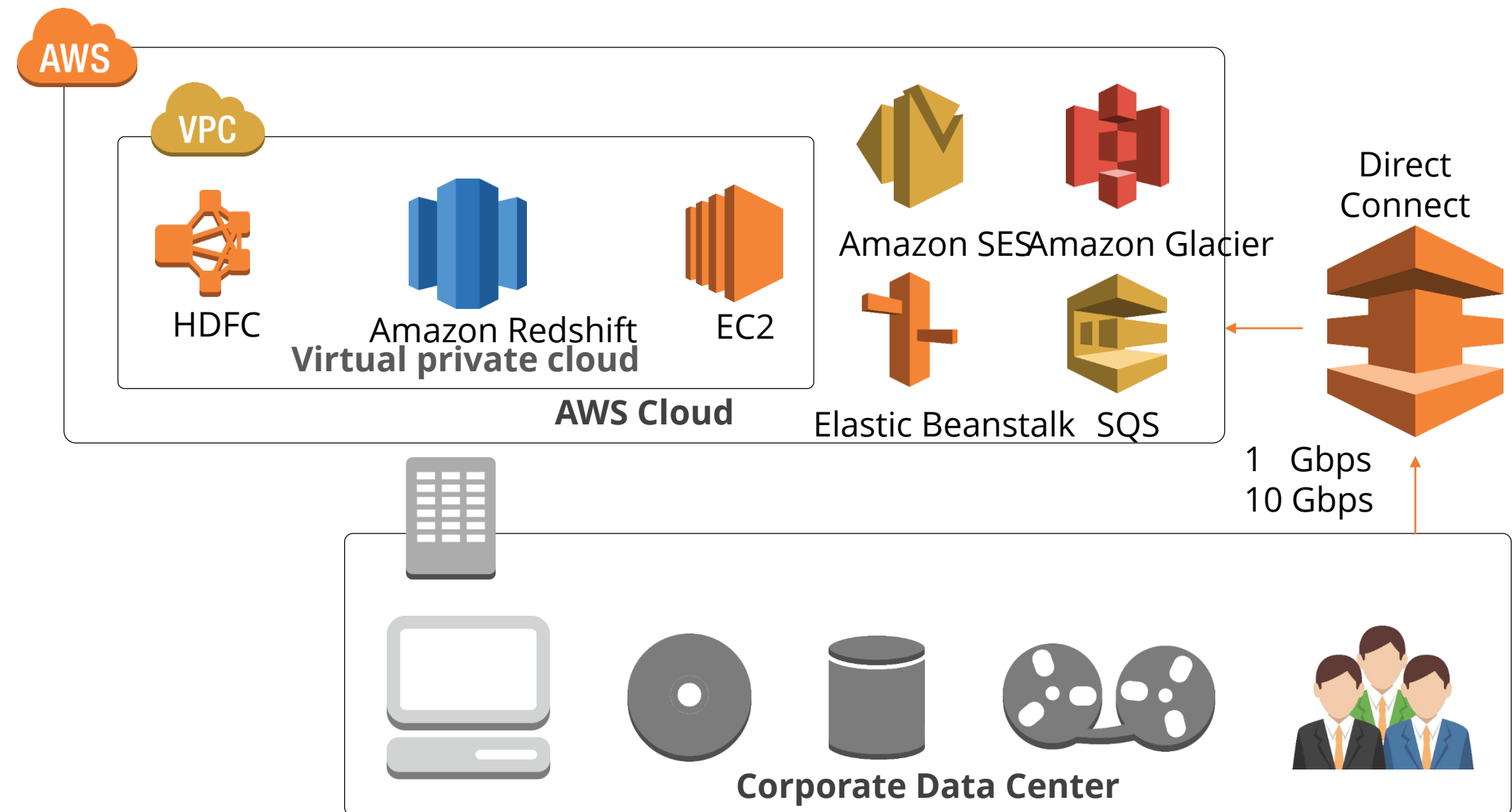
Data Integration and Lifecycle Management

AWS is used to reliably, cost effectively backup, and secure your data. You can replicate data across geographical regions, manage the lifecycle of the data, or even synchronously replicate your data to a local AWS data center.

AWS Storage Gateway is a VM that is installed in your data center and configured to be associated with your AWS account.

Gateway Cached lets you use Amazon S3 as your primary data storage while retaining frequently accessed data locally in your storage gateway.

Gateway-stored volumes let you store your primary data locally, while asynchronously backing up that data to AWS.

Amazon Simple Storage Service provides unlimited storage that is available for backing up and archiving your critical data.

Amazon Glacier is an extremely low-cost storage service for infrequently accessed data, for which recovery of several hours are suitable.

simpli learn

Common Controls for Security and Access

A few common security and access controls are:

AWS Identity and Access Management or IAM is the service that enables you to securely control user access to all AWS services and resources.

AWS Directory Service is a managed service that connects your AWS resources with an existing on-premises Microsoft Active Directory or to set up a new, stand-alone directory in the AWS Cloud.

Microsoft AD allows you to manage user accounts and group memberships, create and apply group policies, domain-join Amazon EC2 instances, and provide Kerberos-based single sign-on (SSO).

Common Controls for Security and Access

- Microsoft AD is a full blown managed Microsoft Active Directory service that supports up to 50,000 users.

- Simple AD is a stand-alone, managed directory that is available in two sizes; small and large.

- AD Connector is a directory gateway that allows you to proxy directory requests to your on-premises Microsoft Active Directory. AD Connector comes in two sizes; small and large.

Active Directory
Microsoft

Microsoft Active Directory Services

Microsoft AD	Simple AD	AD Connector
50,000 Users	Small / Large	Small / Large

Simple AD:
- Small → 500 Users
- Large → 5000 Users

AD Connector:
- Small → 500 Users
- Large → 5000 Users

simplilearn

Integrated Resource and Deployment Management

AWS provides monitoring and management tools with robust APIs so you can easily integrate your AWS resources with on-site tools.

AWS
OpsWorks

You can deploy and operate applications in the AWS Cloud or in your own data center using AWS OpsWorks application management service.

Automate code deployments to instances in your existing data center or in the AWS cloud using AWS CodeDeploy.

AWS
CodeDeploy

simpli·learn
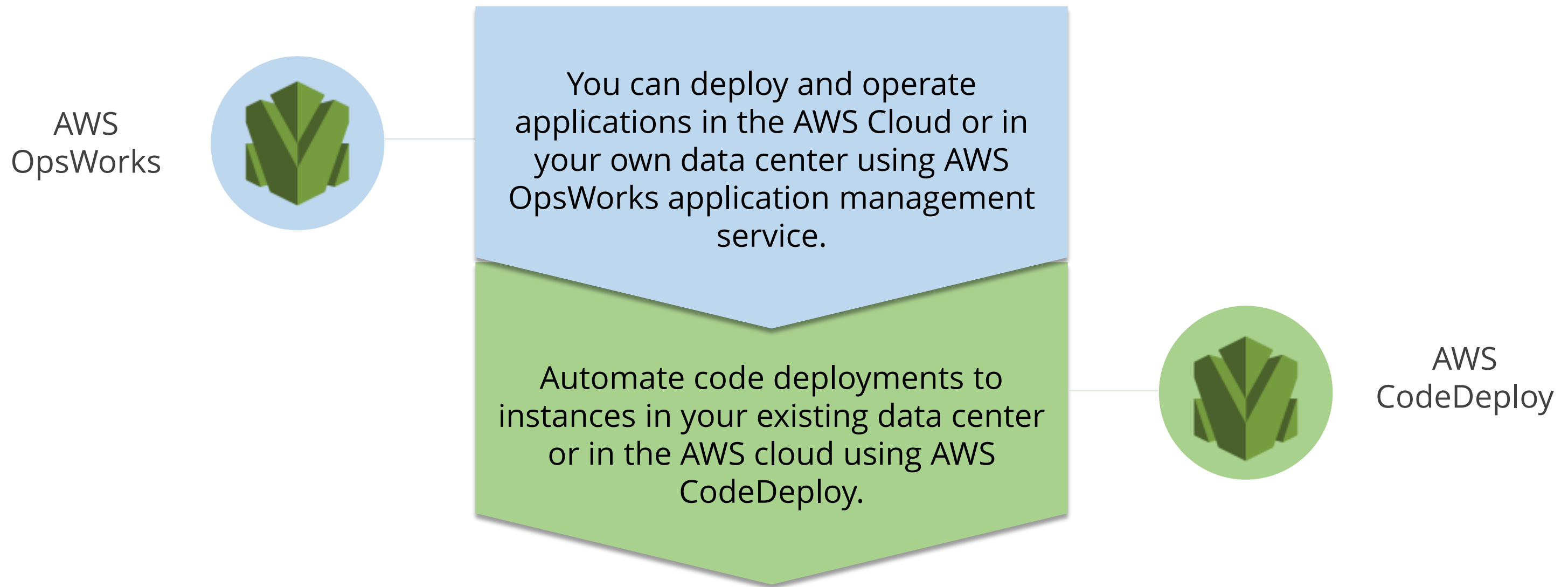
Knowledge Check

KNOWLEDGE CHECK	Which of the following is NOT a service used directly in hybrid architectures?

a.   AWS Storage Gateway

b.   AWS Direct Connect

c.   Amazon Config

d.   AWS Directory Service

Which of the following is NOT a service used directly in hybrid architectures?

a. AWS Storage Gateway

b. AWS Direct Connect

c. Amazon Config

d. AWS Directory Service

The correct answer is **c)**

AWS Storage Gateway is used to store your data on the cloud via your data center. AWS Direct Connect lets you establish a dedicated network connection between your onsite premises and AWS. AWS Directory Service is a managed service that allows you to connect your AWS resources with an existing on-premises Microsoft Active Directory or to set up a new, stand-alone directory in the AWS Cloud.
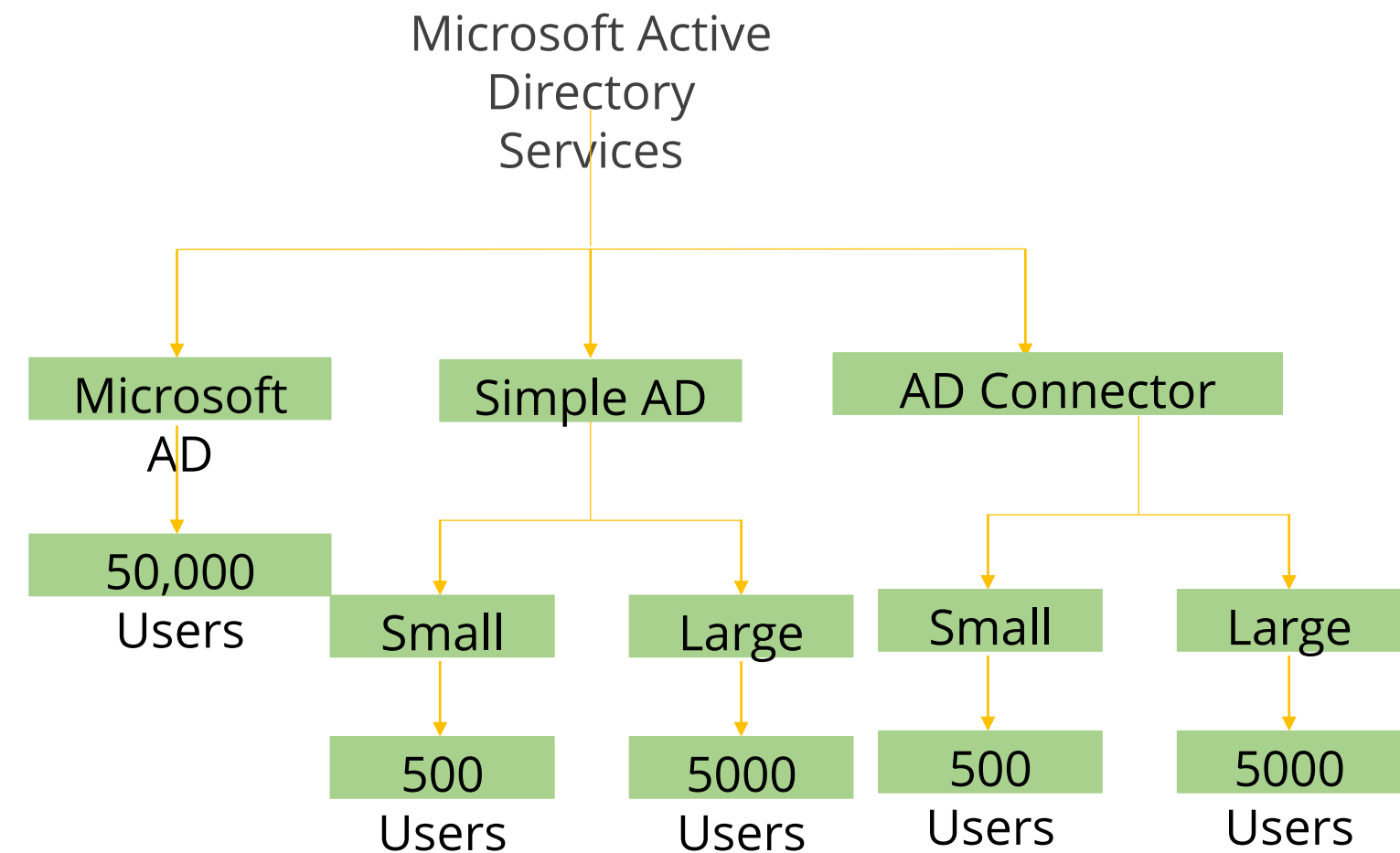
Which Storage Gateway option would you choose if you wanted to use solely AWS storage?

a. Gateway Cached

b. Gateway S3

c. Gateway Stored

d. Gateway Remote

Which Storage Gateway option would you choose if you wanted to use solely AWS storage?

a. Gateway Cached

b. Gateway S3

c. Gateway Stored

d. Gateway Remote

The correct answer is **a)**

Gateway Cached lets you use Amazon Simple Storage Service (Amazon S3) as your primary data storage while retaining frequently accessed data locally in your storage gateway. Gateway Stored volumes let you store your primary data locally, while asynchronously backing up that data to AWS.

Practice Assignment: Designing Hybrid Storage

Configure a basic plan to resolve an on-premise storage problem

Designing Hybrid Storage

You have been hired by a medium sized law firm. They have an aging storage solution which they want to replace, but they do not want to purchase any hardware.

They store several terabytes of data, comprising documents and images. A lot of the data is legacy data, which is rarely accessed and can be archived. However, the most recent files need to be available instantly.

You have been asked to provide a basic, high-level plan for a hybrid storage solution using the client's existing data center and AWS.

Detail the products and services you would use and sketch out a basic plan of the infrastructure.

Key Takeaways

Key Takeaways

- AWS Well-Architected Framework helps you to understand the pros and cons of decisions you make while building systems on AWS.

- The AWS Well-Architected Framework is based on four pillars: Security, Reliability, Performance efficiency, and Cost Optimization.

- Cloud computing helps achieve optimal server configuration by providing various features.

- You can configure Amazon CloudWatch, CloudTrail, and AWS Config to provide you with alerts and notifications.

- With hybrid technologies you can link your existing on-premises network configuration onto your AWS virtual private cloud.

Key Takeaways

- AWS Well-Architected Framework helps you to understand the pros and cons of decisions you make while building systems on AWS.

- The AWS Well-Architected Framework is based on four pillars: Security, Reliability, Performance efficiency, and Cost Optimization.

- Cloud computing helps achieve optimal server configuration by providing various features.

- You can configure Amazon CloudWatch, CloudTrail, and AWS Config to provide you with alerts and notifications.

- With hybrid technologies you can link your existing on-premises network configuration onto your AWS virtual private cloud.

simplilearn

Quiz

QUIZ

1

Where should you look to find documentation about AWS architecture?

a. AWS Architecture Center

b. AWS Whitepapers

c. AWS Case Studies

d. AWS Quick Reference Deployments

QUIZ

1

Where should you look to find documentation about AWS architecture?

a. AWS Architecture Center

b. AWS Whitepapers

c. AWS Case Studies

d. AWS Quick Reference Deployments

The correct answer is **a, b, and c**

Explanations: AWS Architecture Center, AWS Whitepapers, and AWS Case Studies will provide you information about AWS architecture.

QUIZ 2

What service does AWS Quick Start Reference Deployments use?

a. EC2 Container Service

b. CloudFront

c. CloudFormation

d. RDS

QUIZ 2

What service does AWS Quick Start Reference Deployments use?

a. EC2 Container Service

b. CloudFront

c. CloudFormation

d. RDS

The correct answer is **c**

Explanations: CloudFormation is used to rapidly deploy a fully functioning environment for a variety of enterprise software applications.

QUIZ

3

Which of these is NOT a benefit of cloud computing?

a. Dynamic scaling

b. Global deployment

c. Fixed Capacity

d. Cost efficiency

QUIZ

3

Which of these is NOT a benefit of cloud computing?

a. Dynamic scaling

b. Global deployment

c. Fixed Capacity

d. Cost efficiency

The correct answer is   **c**

Explanations: Fixed capacity is associated with traditional IT infrastructure. With AWS you don't need to worry about provisioning capacity based on estimates.

QUIZ 4	What does Vertical Scaling mean?

a.   Increasing the monitoring of a resource

b.   Increasing the number of resources

c.   Increasing the specifications of a resource

d.   Increasing the number of applications on a resource

QUIZ

4

What does Vertical Scaling mean?

a. Increasing the monitoring of a resource

b. Increasing the number of resources

c. Increasing the specifications of a resource

d. Increasing the number of applications on a resource

The correct answer is   **c**

Explanations: Vertical Scaling means increasing the specifications of a resource, for example increasing the memory and CPU.

QUIZ

5

What is a stateless application?

a. One that retains all application logs on S3

b. One that is running on AWS EC2

c. One that maintains information based on previous interactions and stores session information

d. One that needs no knowledge of previous interactions and stores no session information

QUIZ 5	What is a stateless application?

a. One that retains all application logs on S3

b. One that is running on AWS EC2

c. One that maintains information based on previous interactions and stores session information

d. One that needs no knowledge of previous interactions and stores no session information

The correct answer is **d**

Explanations: A stateless application is one that needs no knowledge of previous interactions and stores no session information. An example of this would be a webserver that provides the same web page to any end user.

simpl¡learn

| QUIZ 6 | Which one is an example of a Push Model distributing load to multiple nodes? |

a.   AWS Elastic Load Balancer

b.   AWS SQS

c.   Amazon Kinesis

d.   Storage Gateway

QUIZ 6	Which one is an example of a Push Model distributing load to multiple nodes?

a. AWS Elastic Load Balancer

b. AWS SQS

c. Amazon Kinesis

d. Storage Gateway

The correct answer is **a**

Explanations: A load balancer, such as the AWS Elastic Load Balancer, is a popular way to distribute a workload across multiple resources.

| QUIZ

7	Which of these components is not stateful by definition?

a. AWS Lambda

b. Application running on a single server

c. DynamoDB

d. RDS

Which of these components is not stateful by definition?

a. AWS Lambda

b. Application running on a single server

c. DynamoDB

d. RDS

The correct answer is   **a**

Explanations: Databases are stateful by definition as they store and retain data. Lambda uses a stateless programming model.

QUIZ

8

What method ensures Loose Coupling?

a. Hard Failures

b. Hardcoded IP addresses

c. Synchronous Integration

d. Well-Defined Interfaces

What method ensures Loose Coupling?

a. Hard Failures

b. Hardcoded IP addresses

c. Synchronous Integration

d. Well-Defined Interfaces

The correct answer is   **d**

Explanations: Ensure that all components only interact with each other through specific, technology-agnostic interfaces, for example RESTful APIs, will result in being able to modify resources without affecting other components.

This concludes "Designing Highly Available, Cost-efficient, Fault-tolerant Scalable Systems."

The next lesson is "AWS IAM."

simplilearn