NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Principles in Cloud Architecture Design - Cost Optimization

Cloud Infrastructure Engineering

Nanyang Technological University
& Skills Union - 2022/2023

Course Content

- Quick Check-In
- Dive into what Cost Optimization on the cloud means and the best practices in terms of Cost Optimization
- Build architectures that are well-thought and well-considered from a cost point-of-view.

Time	What	How or Why
7:15pm - 7:40pm	Part 1 - Presentation	Cost Optimization Principles
7:40pm - 8:00pm	Part 2 - Activity	
8:00pm - 8:10pm	Break	
8:10pm - 8:30pm	Part 3 - Presentation	Best Practices for Cost Optimization
8:30pm - 8:50pm	Part 4 - Activity	
8:50pm - 10:00pm	Summary & Assignments	

Recap

- Performance Design Principles
  - Democratize advanced technologies
  - Go global in minutes
  - Use serverless architectures
  - Experiment more often
  - Consider mechanical sympathy
- Best Practices
  - Selection
  - Review
  - Monitoring
  - Trade-offs

Self Study Check-In

Q1) What approach will you take to reduce your company's monthly costs?

Overview

Overview

This module focuses on the **cost optimization** pillar, and how to architect workloads with the **most effective use of services and resources**, to **achieve business outcomes at the lowest price point.**

Cost Optimization Principles

Cost Optimization Principles Summary

- Implement Cloud Financial Management
- Adopt a Consumption Model
- Measure Overall Efficiency
- Stop Spending Money on Undifferentiated Heavy Lifting
- Analyze & Attribute Expenditure

Implement Cloud Financial Management

To achieve financial success and accelerate business value realization in the cloud, you need to **invest in Cloud Financial Management /Cost Optimization.**

Your organization needs to **dedicate time and resources** to build capability in this new domain of technology and usage management.

Similar to your Security or Operational Excellence capability, you need to **build capability through knowledge building, programs, resources, and processes** to become a cost-efficient organization.

Implement Cloud Financial Management

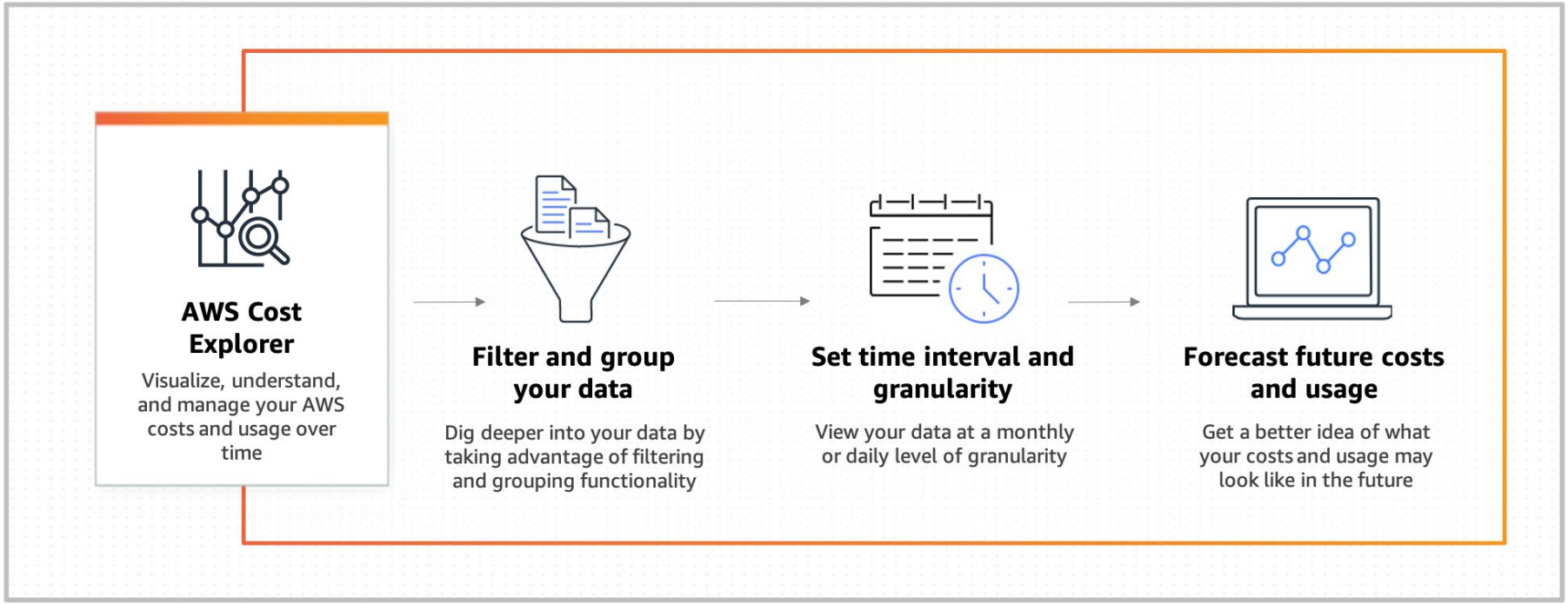
Think - Cloud FinOps

https://www.finops.org/introduction/what-is-finops/

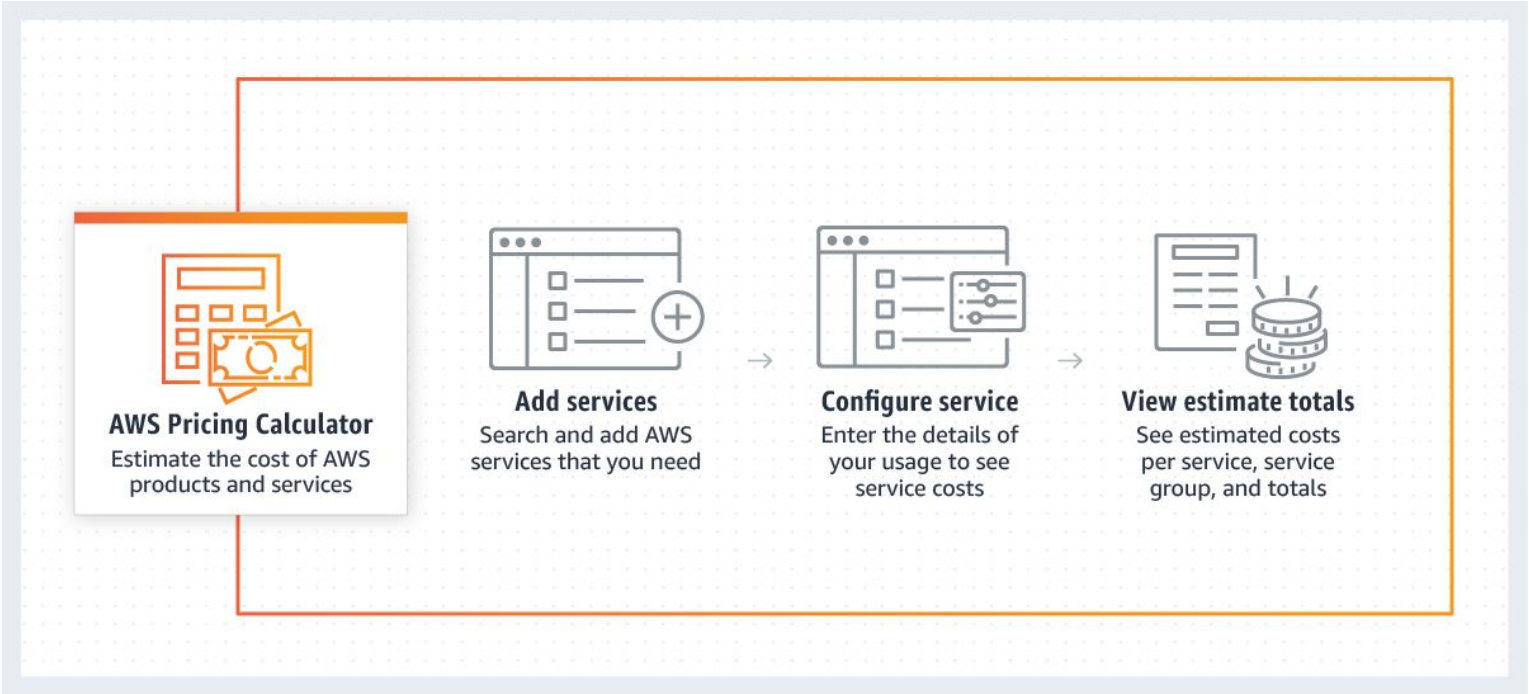
Implement Cloud Financial Management

https://medium.com/contino-engineering/becoming-a-finops-certified-practitioner-69c462d7842e

Implement Cloud Financial Management



Implement Cloud Financial Management



Adopt a Consumption Model

Pay only for the **computing resources that you require** and **increase/ decrease usage depending on business requirements**, not by using elaborate forecasting.

For example, development and test environments are typically only used for eight hours a day during the work week.

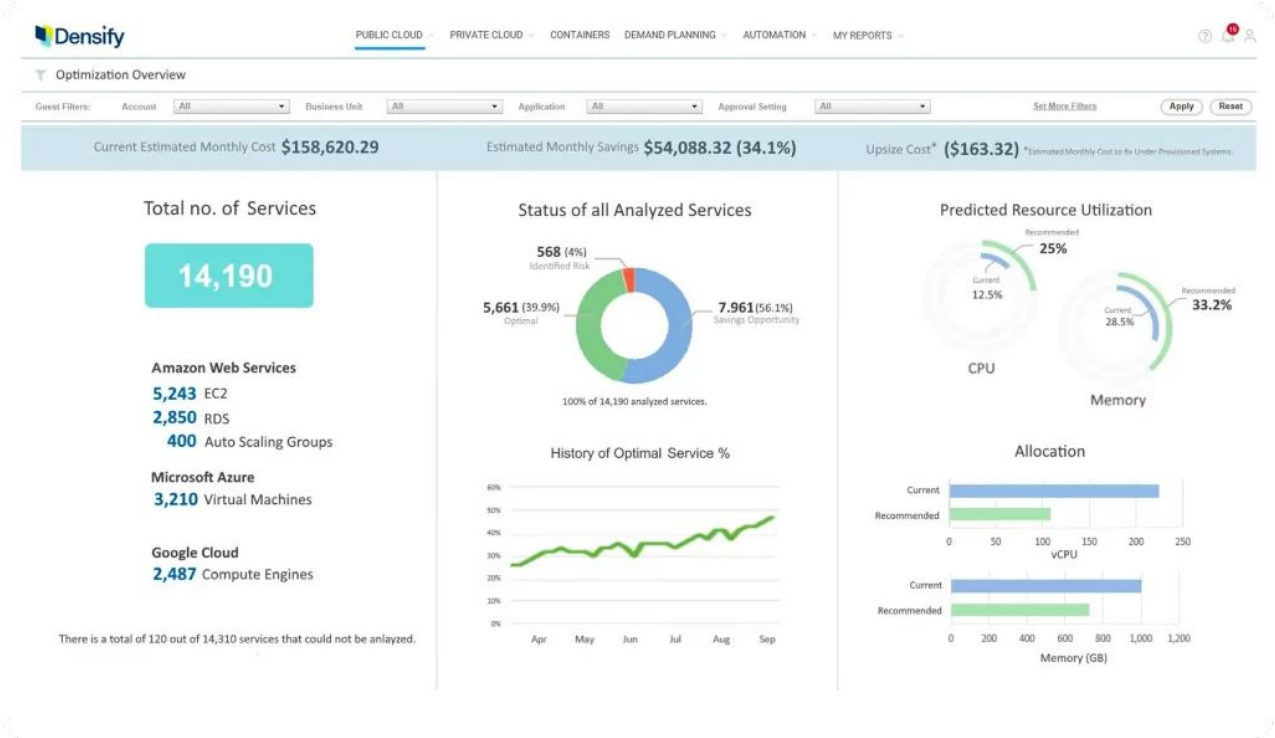
You can **stop these resources when they are not in use** for a potential **cost savings of 75%** (40 hours versus 168 hours).

Measure Overall Efficiency

Measure the **business output of the workload** and the **costs associated with delivering it**.

Use this measure to **know the gains** you make from increasing output and reducing costs.

Measure Overall Efficiency



Stop Spending Money on Undifferentiated Heavy Lifting

AWS does the heavy lifting of data center operations like racking, stacking, and powering servers.

It also **removes the operational burden** of managing operating systems and applications with managed services.

This allows you to **focus on your customers and business projects** rather than on IT infrastructure.

Analyze & Attribute Expenditure

The cloud makes it easier to accurately identify the usage and cost of systems, which then allows transparent attribution of IT costs to individual workload owners.

This helps **measure return on investment (ROI)** and gives workload owners an opportunity to **optimize their resources and reduce costs**.

Cost Optimization Principles Summary

- Implement Cloud Financial Management
- Adopt a Consumption Model
- Measure Overall Efficiency
- Stop Spending Money on Undifferentiated Heavy Lifting
- Analyze & Attribute Expenditure

Activity - Understanding Cost Optimization

In this activity, gather into your own group and each group should take on one or two research problem.

Ensure all research problems are taken and presented by the end of this section.

Activity - Understanding Cost Optimization

- Implement Cloud Financial Management
- Adopt a Consumption Model
- Measure Overall Efficiency
- Stop Spending Money on Undifferentiated Heavy Lifting
- Analyze & Attribute Expenditure

Best Practices For Cost Optimization

Summary

There are five best practice areas for cost optimization in the cloud:

- Practice Cloud Financial Management
- Expenditure and usage awareness
- Cost-effective resources
- Manage demand and supply resources
- Optimize over time

Practice Cloud Financial Management

With the adoption of cloud, technology teams **innovate faster** due to shortened approval, procurement, and infrastructure deployment cycles.

A new approach to financial management in the cloud is required to **realize business value and financial success**.

This approach is Cloud Financial Management, and builds capability across your organization by **implementing organizational wide knowledge building, programs, resources, and processes**.

Practice Cloud Financial Management

Many organizations are composed of many different units with different priorities.

The ability to **align your organization to an agreed set of financial objectives**, and provide your organization the mechanisms to meet them, will create a more efficient organization.

A capable organization will innovate and build faster, be more agile and adjust to any internal or external factors.

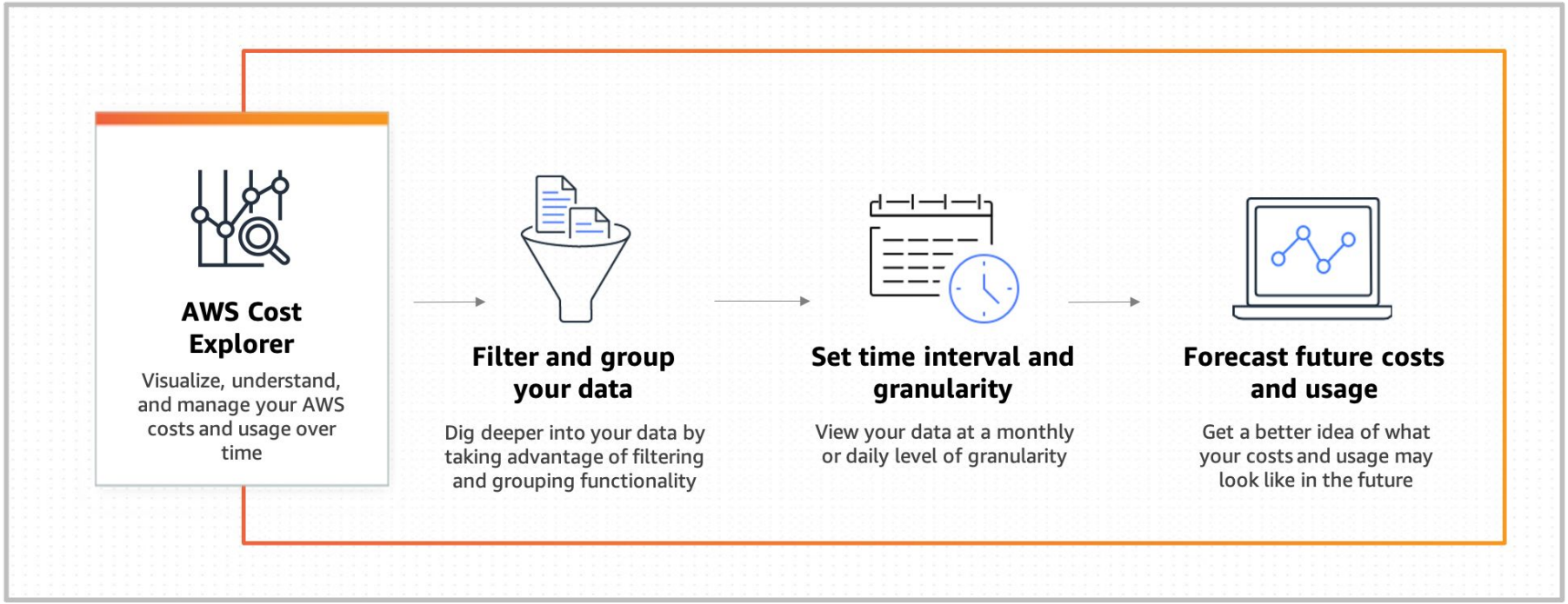
Practice Cloud Financial Management

In AWS you can use **Cost Explorer**, and optionally **Amazon Athena and Amazon QuickSight with the Cost and Usage Report (CUR)**, to provide cost and usage awareness throughout your organization.

AWS Budgets provides proactive notifications for cost and usage.

The AWS blogs provide information on new services and features to ensure you keep up to date with new service releases.

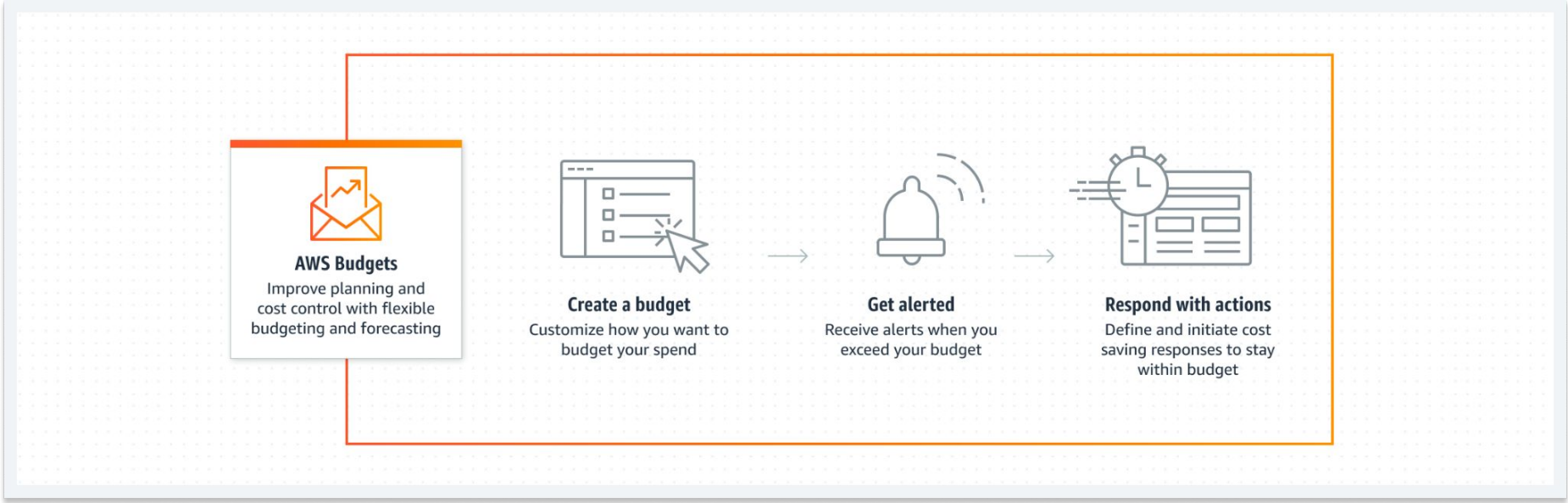
Practice Cloud Financial Management



Practice Cloud Financial Management

QuickSight powers millions of dashboard views weekly, so your end users can make better data-driven decisions.

Practice Cloud Financial Management



Expenditure & Usage Awareness

The increased flexibility and agility that the cloud enables encourages innovation and fast-paced development and deployment.

It eliminates the manual processes and time associated with provisioning on-premises infrastructure, including identifying hardware specifications, negotiating price quotations, managing purchase orders, scheduling shipments, and then deploying the resources.

However, the ease of use and virtually unlimited on-demand capacity requires a new way of thinking about expenditures.

Expenditure & Usage Awareness

Many businesses are composed of multiple systems run by various teams.

The **capability to attribute resource costs** to the individual organization or product owners **drives efficient usage behavior and helps reduce waste.**

Accurate cost attribution allows you to know **which products are truly profitable**, and **allows you to make more informed decisions** about where to **allocate budget.**

Expenditure & Usage Awareness

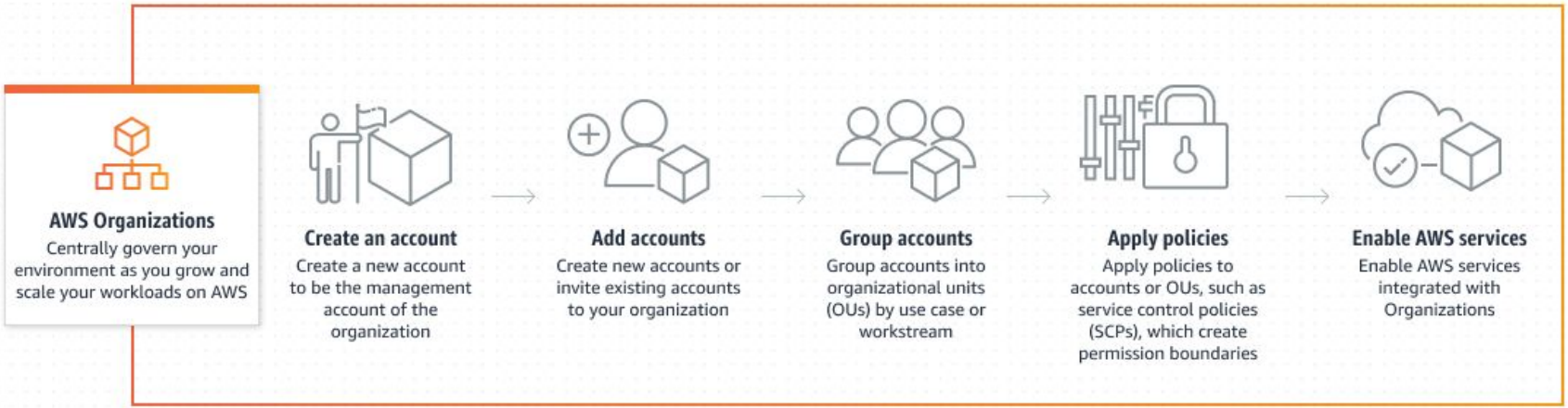
In AWS, you create an account structure with **AWS Organizations or AWS Control Tower**, which provides separation and assists in allocation of your costs and usage.

You can also use resource tagging to apply business and organization information to your usage and cost.

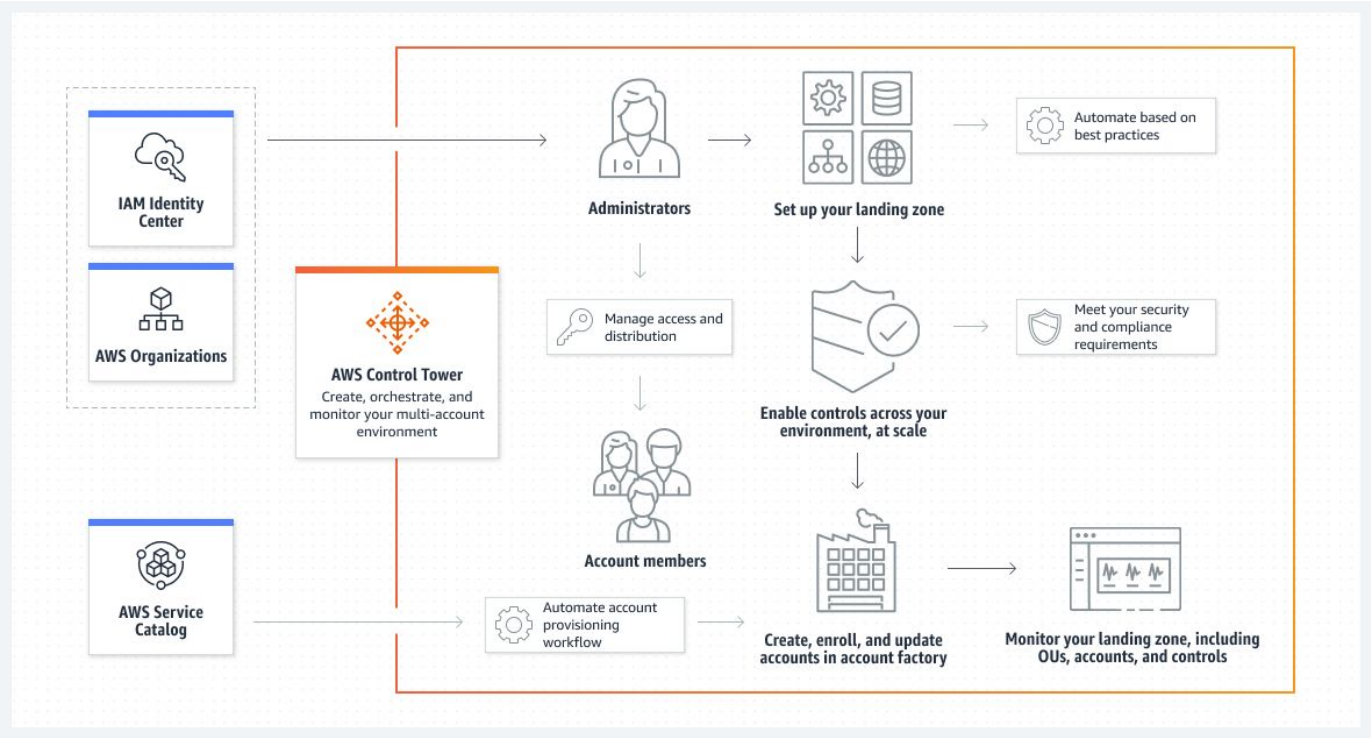
Use **AWS Cost Explorer** for visibility into your cost and usage, or create customized dashboards and analytics with **Amazon Athena and Amazon QuickSight**.

Controlling your cost and usage is done by notifications through **AWS Budgets**, and controls using AWS Identity and Access Management (IAM), and Service Quotas.

Expenditure & Usage Awareness



Expenditure & Usage Awareness



Cost-Effective Resource

Using the **appropriate instances and resources for your workload is key to cost savings**.

For example, a reporting process might take five hours to run on a smaller server but one hour to run on a larger server that is twice as expensive.

Both servers give you the same outcome, but the smaller server incurs more cost over time.

Cost-Effective Resource

A well-architected workload uses the **most cost-effective resources**, which can have a **significant and positive economic impact.**

You also have the opportunity to use managed services to **reduce costs.**

For example, rather than maintaining servers to deliver email, you can **use a service that charges on a per-message basis**.

Cost-Effective Resource

AWS EC2 Usage

On-Demand Instances allow you to pay for compute capacity by the hour, with no minimum commitments required.

Savings Plans and Reserved Instances offer **savings of up to 75% off On-Demand pricing**.

With Spot Instances, you can leverage unused Amazon EC2 capacity and offer **savings of up to 90% off On-Demand pricing**. Spot Instances are appropriate where the system can tolerate using a fleet of servers where individual servers can come and go dynamically, such as stateless web servers, batch processing, or when using HPC and big data.

Cost-Effective Resource

https://aws.amazon.com/savingsplans/compute-pricing/

Manage Demand & Supply Resources

When you move to the cloud, you **pay only for what you need.**

You can **supply resources to match the workload demand** at the time they're needed, this **eliminates the need for costly and wasteful over provisioning**.

You can also **modify the demand**, using a throttle, buffer, or queue to smooth the demand and serve it with less resources resulting in a lower cost, or process it at a later time with a batch service.

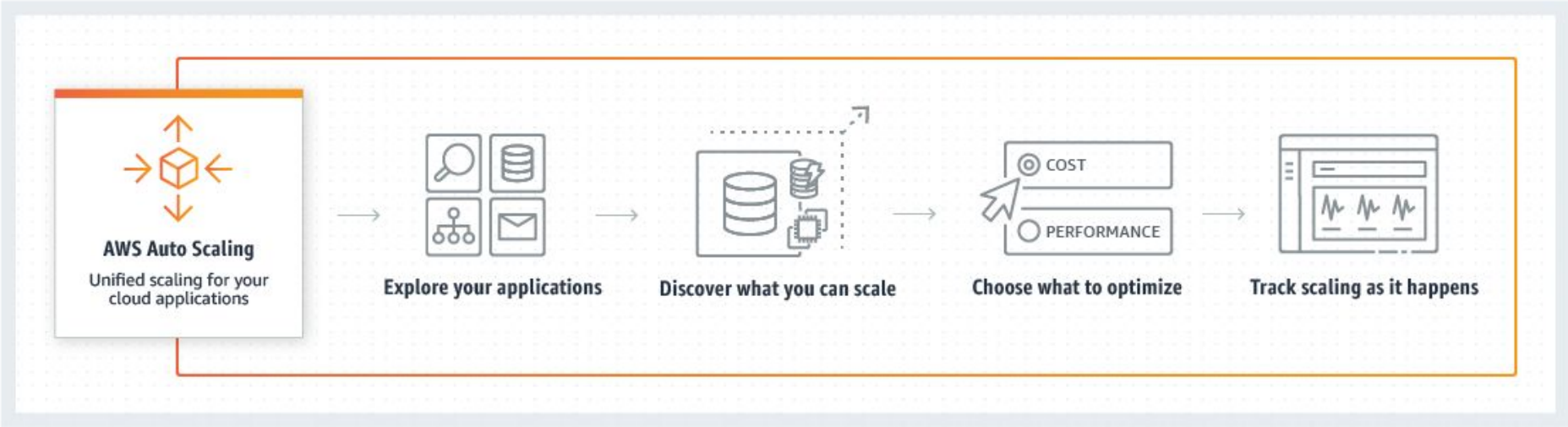
Manage Demand & Supply Resources

In AWS, you can automatically provision resources to match the workload demand.

Auto Scaling using demand or time-based approaches allow you to **add and remove resources as needed**.

If you can anticipate changes in demand, you can save more money and ensure your resources match your workload needs.

Manage Demand & Supply Resources



Manage Demand & Supply Resources

You can use **Amazon API Gateway to implement throttling**, or **Amazon SQS to implementing a queue in your workload**.

These will both allow you to modify the demand on your workload components.

Optimize Over Time

As AWS releases new services and features, it's a best practice to **review your existing architectural decisions** to ensure they continue to be the most cost effective.

As your requirements change, be aggressive in decommissioning resources, entire services, and systems that you no longer require.

Optimize Over Time

Implementing **new features or resource types** can optimize your workload incrementally, while minimizing the effort required to implement the change.

This provides **continual improvements in efficiency** over time and ensures you remain on the most updated technology to **reduce operating costs.**

Optimize Over Time

You can also **replace or add new components to the workload** with new services.

This can provide significant increases in efficiency, so it's essential to regularly review your workload, and implement new services and features.

Summary

There are five best practice areas for cost optimization in the cloud:

- Practice Cloud Financial Management
- Expenditure and usage awareness
- Cost-effective resources
- Manage demand and supply resources
- Optimize over time

Services to take note

- Trusted Advisor
- AWS Cost Explorer(Billing Console)
- AWS Budgets

Understanding Cost Optimization

- How do you implement cloud financial management?
- How do you monitor usage and cost?
- How do you evaluate new services?

Activity - Understanding Cost Optimization

In this activity, gather into your own group and each group should take on one or two research problem.

Ensure all research problems are taken and presented by the end of this section.

Activity

Learner:

- Clean up AWS.
- Remove/delete/terminate all service/ resources that created.

Instructor

- Clean up AWS.
- Remove/delete/terminate all service/ resources that created.
- Check the AWS account after learner clean up.

What’s Next?