# AWS re:Invent

ARC204-R

# Cost optimizing a workload

**Nathan Besh**

Cost Lead, Well-Architected
Amazon Web Services

aws

# Agenda

Introduction

Setup

Baseline

Cost optimization cycles (analyze & optimize)

Bonus: Well-Architected

# Introduction

- **Workshop goal**

- **Simulate the improvements**

- **Present, then implement**

  - A few slides on the topic

  - Hands-on labs

- **Labs are available online**

  - Stay together (don't worry if you don't finish)

  - Do them anytime

# Related Breakouts

ARC219 – AWS Cost Management tools for cost & usage optimization (200-300)

ENT206 – Optimize AWS costs and utilization with AWS management tools (100-200)

CMP410 – Save up to 90 percent on CI/CD and test workloads (spot)

ARC209 – Running lean architectures: How to be cost-effective on AWS (Presentation)

CMP207 – Manage, control, and optimize costs with native AWS products, ft. Intuit (Chalk talk)

CMP323 – Optimize performance and cost for your AWS compute (Presentation)

ENT204 – Managing your cloud financials as you scale on AWS (Presentation)
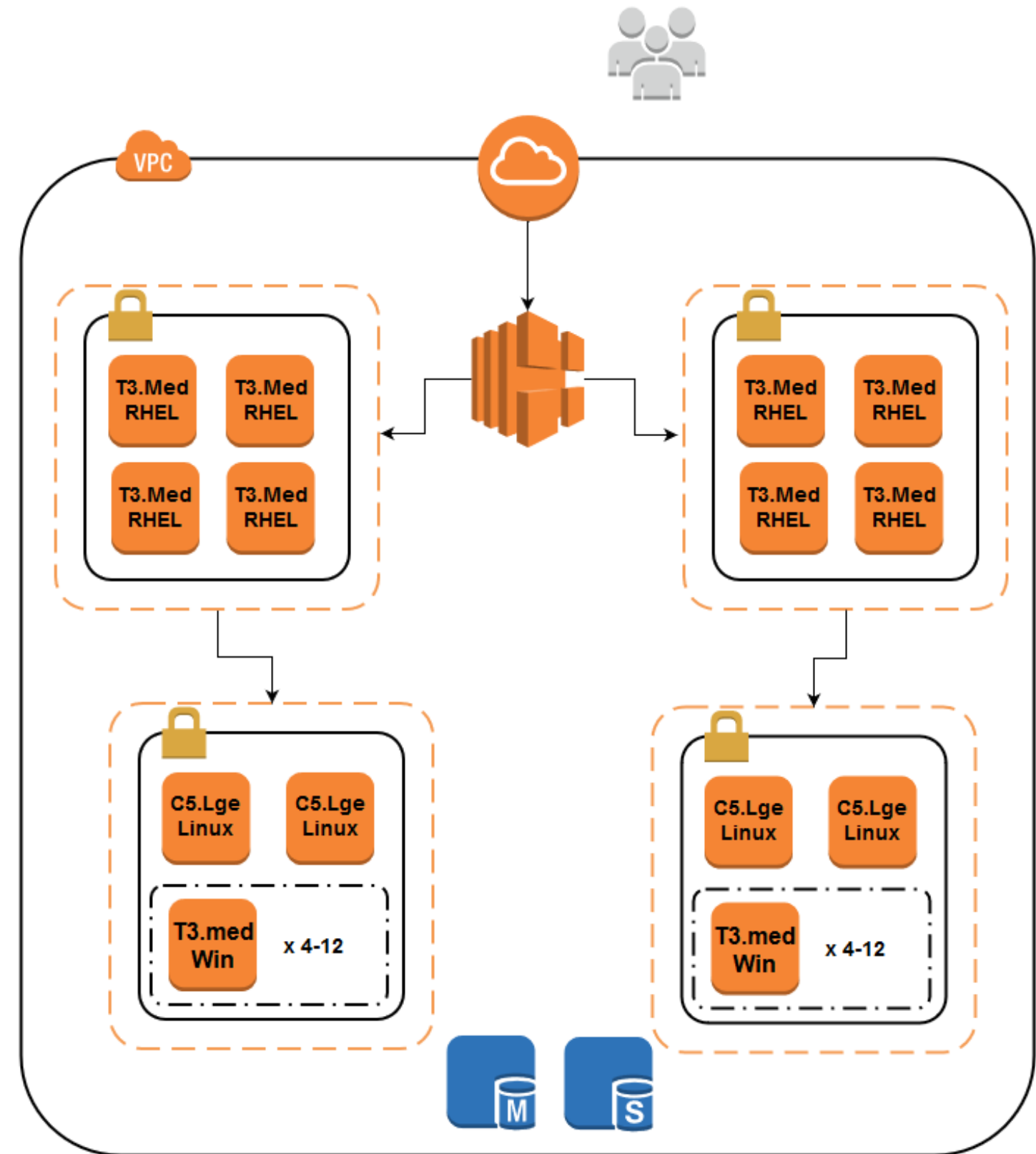
# Environment Setup

# Setup

3-Tier Web App

Front:
  8 x t3.medium RHEL
  Load Balancer

Application:
  4 x c5.large Linux
  4-12 (ASG) x t3.medium Windows

Database: Amazon Aurora

# Tools

- Application log files
- Cost and Usage Report (billing files)
- Amazon Simple Storage Service (Amazon S3)
- Amazon Athena
- Amazon QuickSight

# Process

- Copy application & billing files
- Perform analysis
- Decide action
- Simulate implementation
- Confirm optimization

# Lab

https://wellarchitectedlabs.com/Cost/arc204.html

Step 1: Setup

# Baseline

aws

# Baseline

- How does my workload demand behave?
- How do my workload resourcing and costs behave?
- What is my workload efficiency?
- Business focus & high level

- What does your workload do? (business outcomes)
- What shows these outcomes? (log file)
- Workload resourcing/costs = Cost & Usage Report (CUR)

# Baseline

- Demand profile
- Workload profile
- Workload efficiency (outcomes/$)

# Lab

https://wellarchitectedlabs.com/Cost/arc204.html

## Step 2: Baseline

# Cycle #1: Licensing

aws

# Licensing

- What is the cost of a software license?

- What is the cost of running licensed software?

- If software mandates additional resources, those resources are the cost of running the software & should be associated

# Lab

https://wellarchitectedlabs.com/Cost/arc204.html

**Step 3: Licensing**

# M5.xlarge – US-East-1 (N. Virginia) – Shared tenancy

| | Linux | RHEL | Windows |
|---|---|---|---|
| Cost/hr | $0.192 | $0.252 | $0.376 |

# T3.medium – US-East-1 (N. Virginia) – Shared tenancy

|  | Linux | RHEL | Windows |
|---|---|---|---|
| Cost/hr | $0.0416 | $0.1016 | $0.06 |

# M5.xlarge – US-East-1 (N. Virginia) – Shared tenancy

- What's different when you launch?
  - Storage
  - Network traffic (updates, chatty OS)
- Console defaults, all free-tier eligible

|  | Linux + 8Gb GP2 | RHEL + 10Gb GP2 | Windows + 30Gb GP2 |
|---|---|---|---|
| Cost/hr | $0.193096 ($0.192) | $0.253370 ($0.252) | $0.380110 ($0.376) |

# M5.xlarge – US-East-1 (N. Virginia) – Shared tenancy

- Discounted Pricing Model
- 1-year, no upfront cost, regional/standard

| | Linux + 8Gb GP2 | RHEL + 10Gb GP2 | Windows + 30Gb GP2 |
|---|---|---|---|
| Cost/hr | $0.124096 ($0.192) | $0.184370 ($0.252) | $0.311110 ($0.376) |

# T3.medium – US-East-1 (N. Virginia) – Shared tenancy

- Reserved Instances & Storage

| | Linux | RHEL | Windows |
|---|---|---|---|
| Cost/hr | $0.027096 | $0.087370 | $0.049110 |
| License + Storage Cost/hr | | $0.060274 | $0.022014 |
| % extra cost | | 3.2247 x | 1.8124 x |

# Cycle #2: Storage

aws

# Storage

- Plan for scale

- Find unused storage efficiently

- Find what storage is used for, efficiently

# Storage

- Unused storage = remove the waste

- What storage is used for = another service?

# Lab

[https://wellarchitectedlabs.com/Cost/arc204.html](https://wellarchitectedlabs.com/Cost/arc204.html)

Step 4: Storage

# Cycle #3: Data Transfer

# Data Transfer

- AWS = managed network provider by default
- Networking is a resource (like compute)
- Accurate & transparent allocation drives efficient workloads
- Efficient compute, efficient storage, efficient network
- Ensure data transfer delivers value

# Data Transfer

- Pricing tips
  - Start with the source service
  - Look at VPC pricing (services inside a VPC)
  - **Gateway** VPC endpoints are free (Amazon S3 & Amazon DynamoDB)
  - **Interface** VPC endpoint / AWS PrivateLink is a private connectivity service

- How far from **your** resource is data traveling?
  - LAN = free
  - Remote/Router = paid

# Lab

https://wellarchitectedlabs.com/Cost/
arc204.html

Step 5: Data Transfer

# Cycle #4: Pricing Models

aws re:Invent

# Pricing Models

- On Demand (OD)
- Spot: Instances, defined duration, fleet
- Reserved Instances (RI)
- Savings Plan (SP)

# Pricing Models: Flexibility & Discounts

Flexibility →

| Standard RI | Regional RI (AZ) | Size Flex (AWS Linux) (AZ, Size) | Convertible RI (AZ, Size, Family, OS, tenancy) |
|---|---|---|---|

| Instance Saving Plan (AZ, Size, OS, tenancy) | Compute Savings Plan (AZ, Size, Family, OS, tenancy, Region, Service) |
|---|---|

Highest Discount
Up to 72%

High Discount
Up to 66%

# Coverage vs. Utilization (RI)

- 100 instances of Linux, m5.large, Linux, 1-year commit, no upfront cost
- On Demand: $84,096
- RI Discount = 36%

- 50% coverage (50 RI/50 OD): $68,766
- 25% coverage (25 RI/75 OD): $76,431

- Purchase 128 RIs (need 100!): $68,398
- Purchase 143 RIs (need 100!): $76,413

# Coverage vs. Utilization (SP)

- 100 instances of Linux, m5.large, Linux, 1-year commit, no upfront cost
- On-Demand: $84,096
- 100% coverage instance SP @ 36%: $53,436
- 100% coverage compute SP @ 26%: $62,196

| 74 Instances running: $62,231 On-Demand | |
|---|---|
| | |

| 64 Instances running: $53,821 On-Demand | |
|---|---|
| Instance SP: $385 (SAVING) | Compute SP: $8,375 (LOSS) |

# Pricing Models = Commitment = Investment

- ## What is the risk?

  - Risk = investment over the full commitment

  - How quickly can I pay off the **full** commitment with savings?

  - Formula = (Monthly discounted amount x 12) / (Monthly On Demand)

  - Breakeven = wrong

- ## What is the return?

  - Savings from discount

  - Variable!

- ## What if it is unused?

  - DON'T Worry!

# Lab

https://wellarchitectedlabs.com/Cost/
arc204.html

Step 6: Pricing Models

# Bonus: Well-Architected Tool

aws

# Cost Optimization: Where should I focus?

- Well-Architected Tool

- 9 Questions

  - Different focus areas of cost optimization

  - Best practices in each area

- Open discussion within your organization

  - No right or wrong

  - How you are achieving the outcomes

- Free learning tool for your organization

- Track your progress and improvement

# Lab

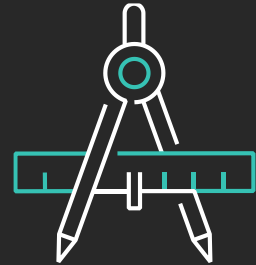[https://wellarchitectedlabs.com/Cost/arc204.html](https://wellarchitectedlabs.com/Cost/arc204.html)
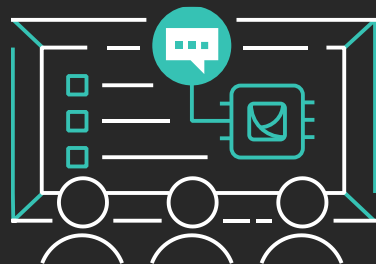
Step 7: Well-Architected Tool

# Learn to architect with AWS Training and Certification

Resources created by the experts at AWS to propel your organization and career forward

Free foundational to advanced digital courses cover AWS services and teach architecting best practices

Classroom offerings, including Architecting on AWS, feature AWS expert instructors and hands-on labs

Validate expertise with the **AWS Certified Solutions Architect - Associate** or **AWS Certification Solutions Architect - Professional** exams

Visit aws.amazon.com/training/path-architecting/

aws training and certification

# Thank you!

**Nathan Besh**

natbesh@amazon.com

Please complete the session survey in the mobile app.