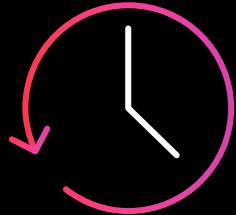
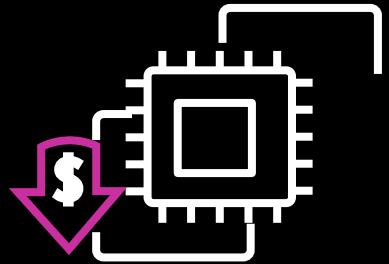


Cost Optimization

AWS
Technical Account Manager 박양수



The Big Five



Rightsizing: Choose the optimal sized instance

Scheduling: Turn them off if you don't use them

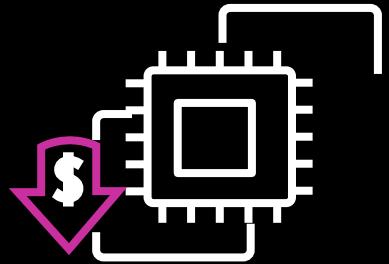
Pricing: On-demand is easy but not the best



Storage: Leverage different storage classes

Monitoring: Set goals and take actions

The Big Five



Rightsizing: Choose the optimal sized instance

Scheduling: Turn them off if you don't use them

Pricing: On-demand is easy but not the best



Storage: Leverage different storage classes

Monitoring: Set goals and take actions



\$732

↓ 75%

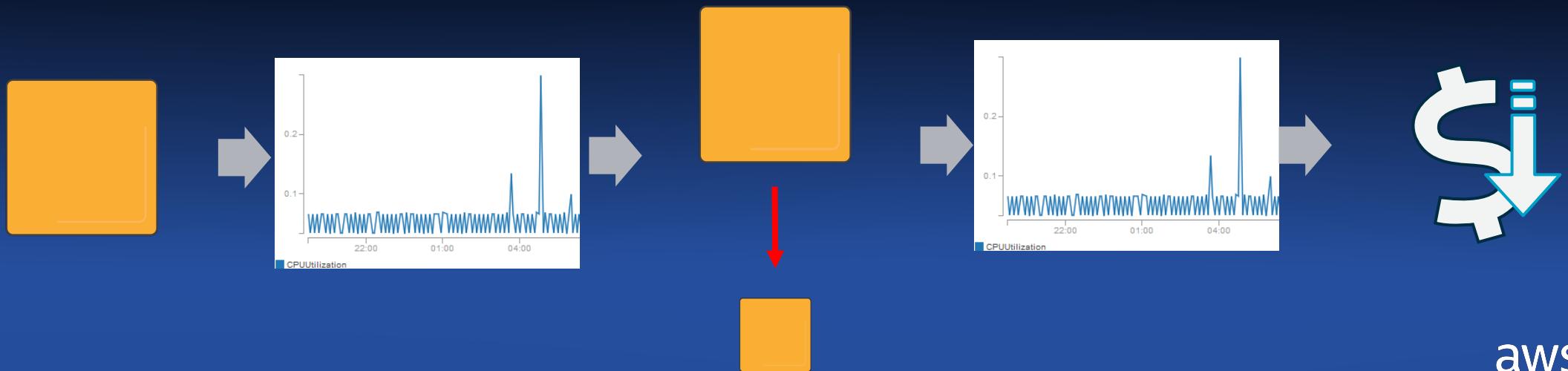
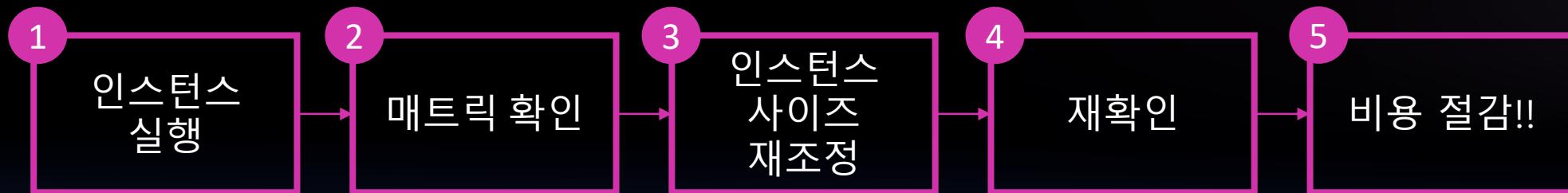
\$183

M4.4xlarge
\$0.984/hr

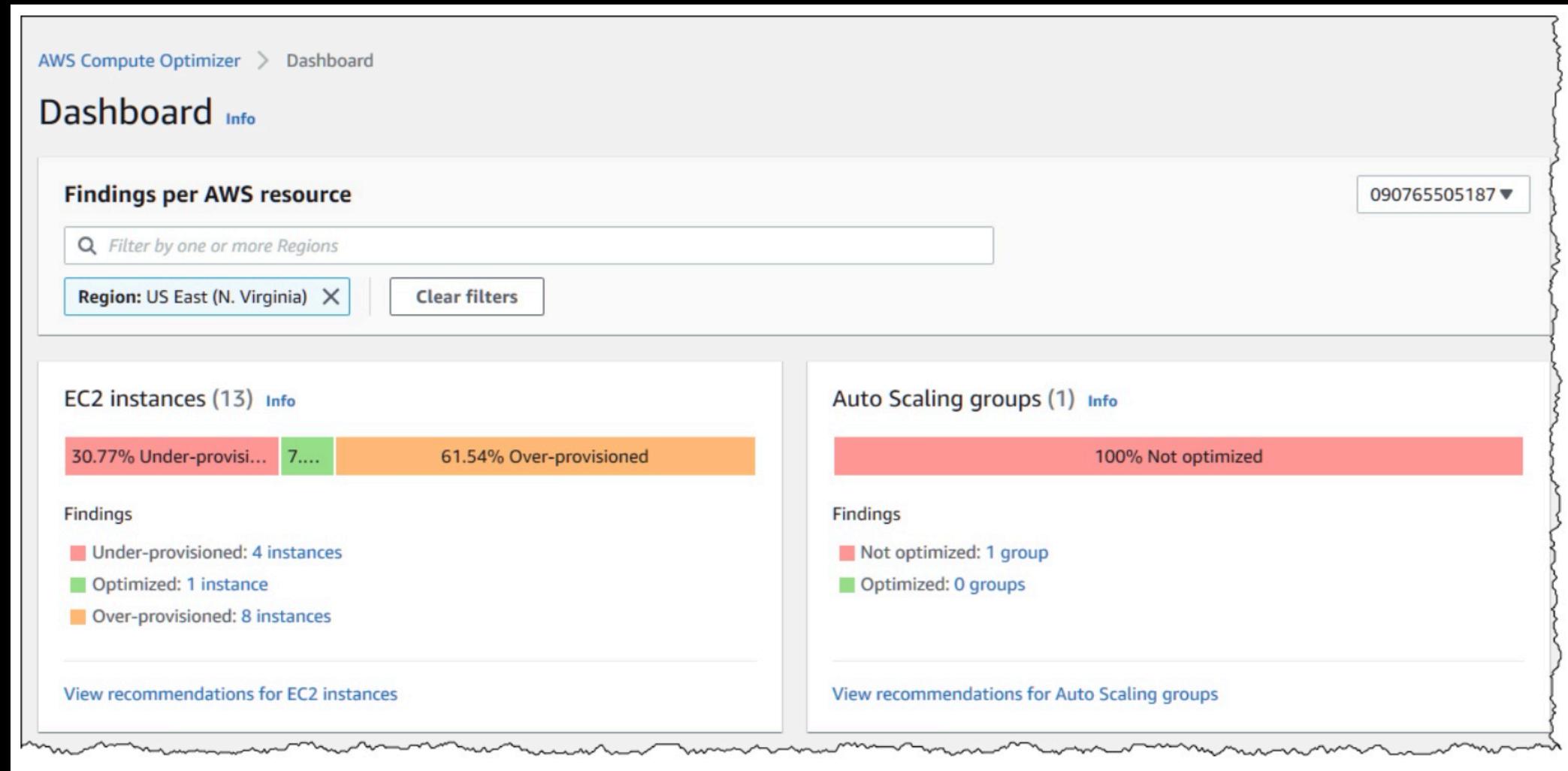
M4.xlarge
\$0.246/hr



Rightsizing, 어떻게 검사할까?



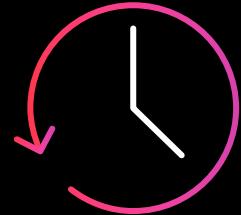
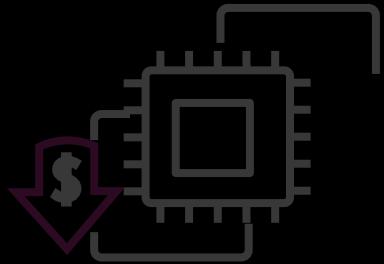
Compute Optimizer



인스턴스 타입, 비용절감에 중요한 요소 중 하나입니다.



The Big Five



Rightsizing: Choose the optimal sized instance

Scheduling: Turn them off if you don't use them

Pricing: On-demand is easy but not the best



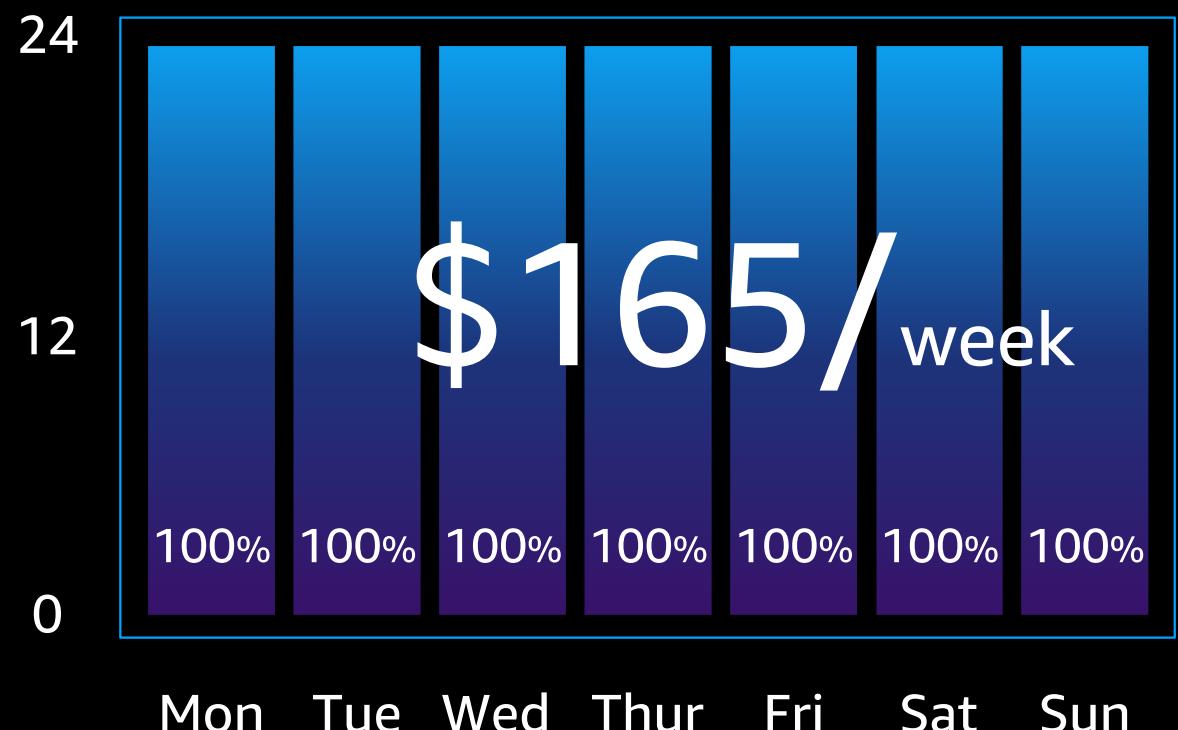
Storage: Leverage different storage classes

Monitoring: Set goals and take actions

Instance A: Production



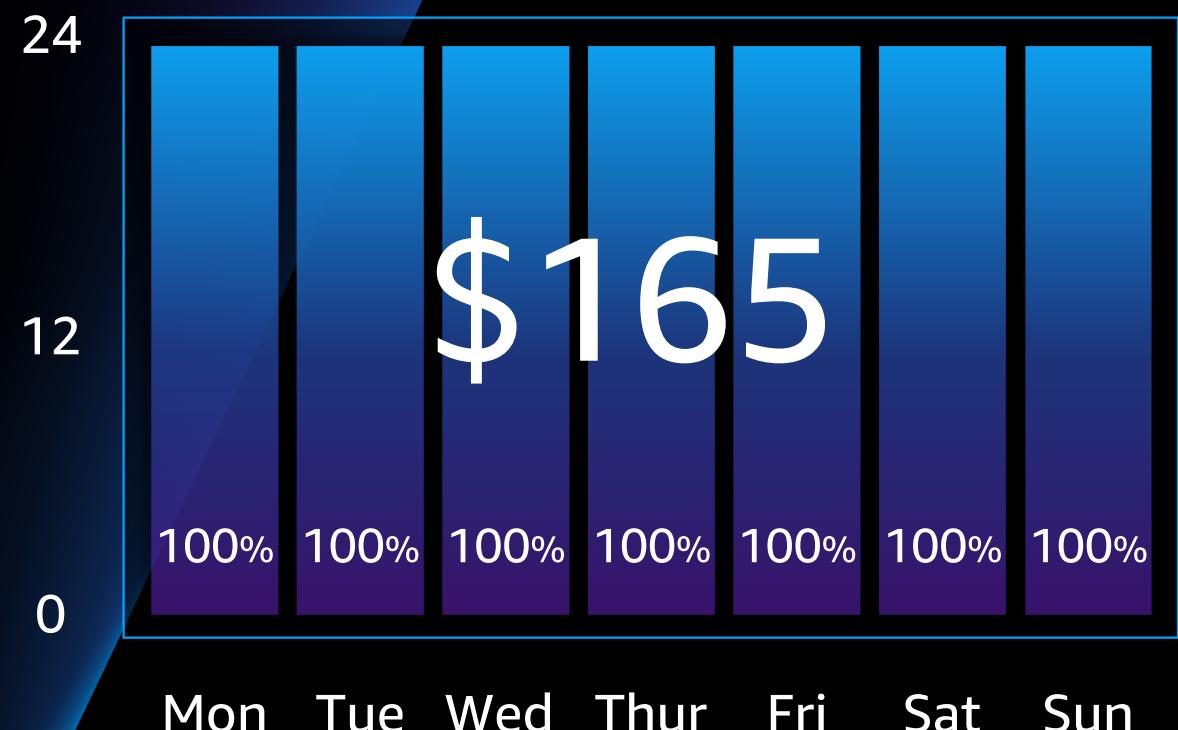
Instance B: Test



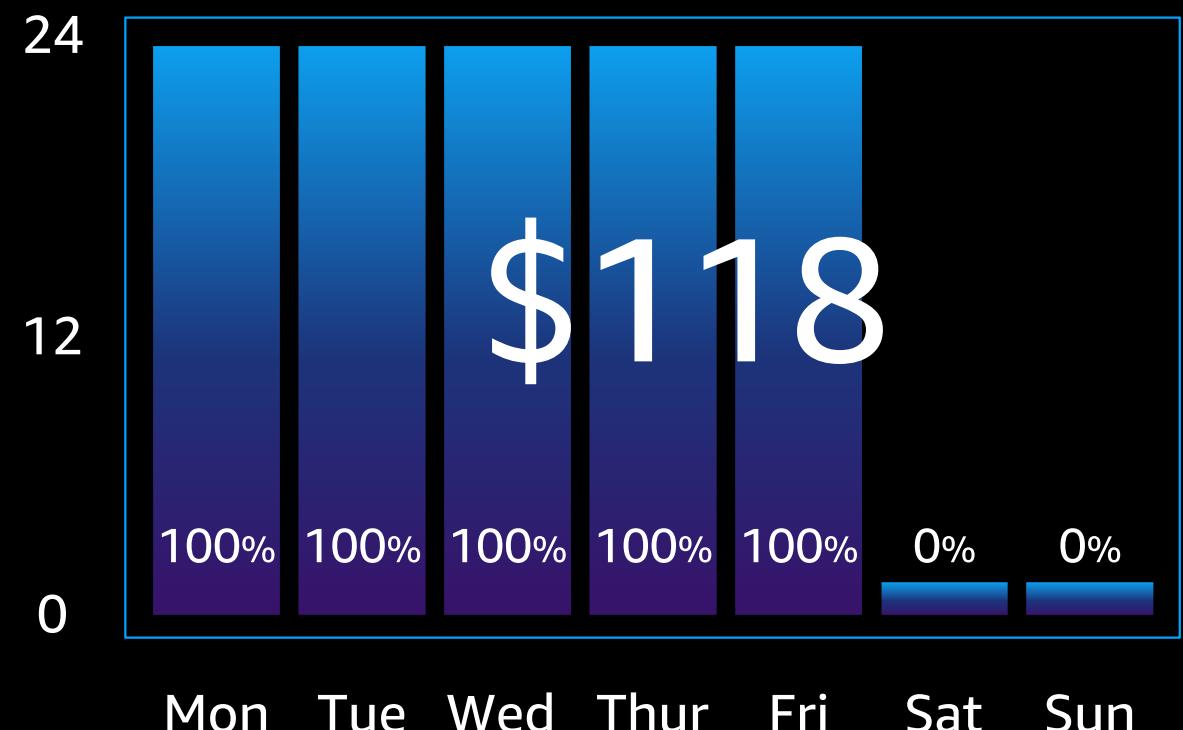
* Seoul Region, m4.4xlarge ₩/조



Instance A: Production



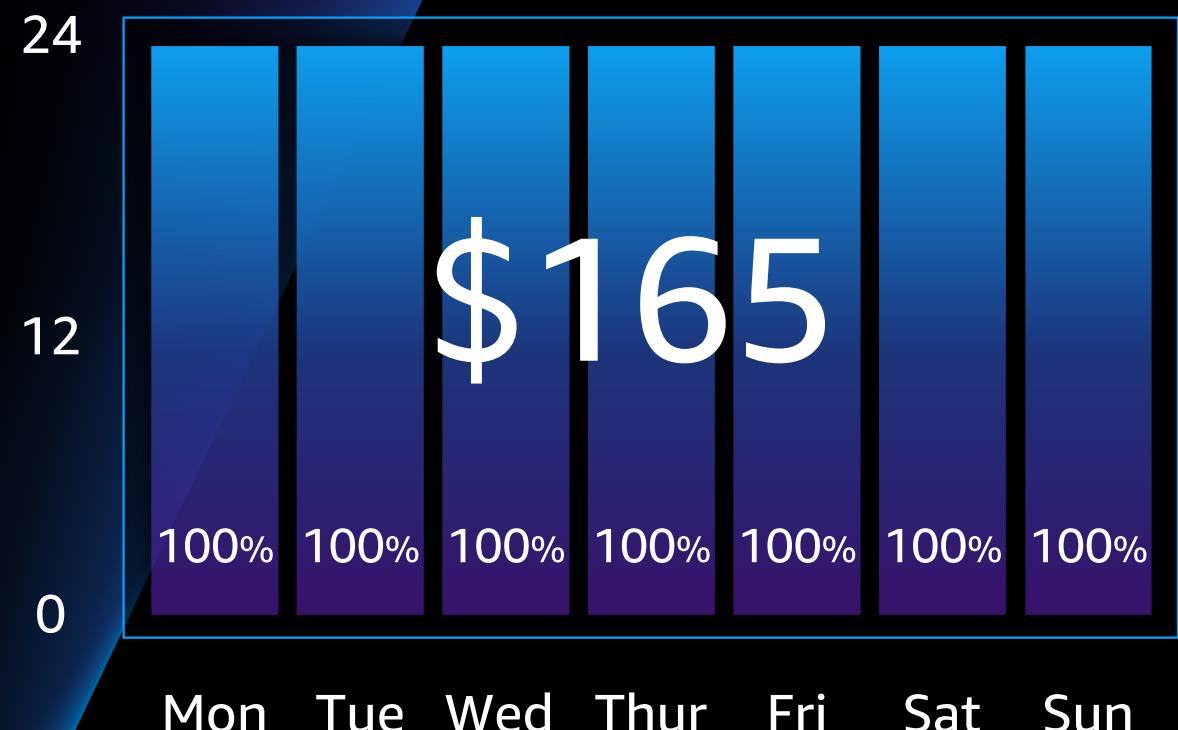
Instance B: Test



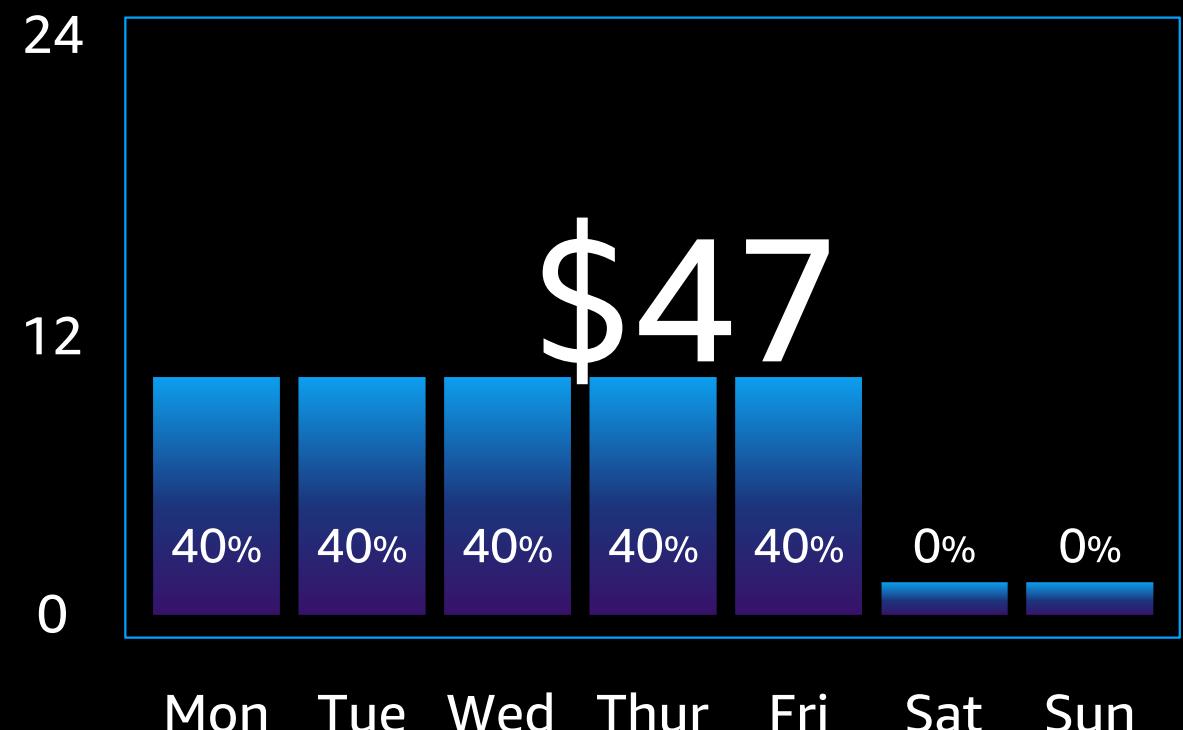
* Seoul Region, m4.4xlarge ₩/조



Instance A: Production



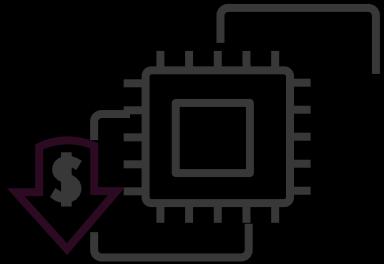
Instance B: Test



* Seoul Region, m4.4xlarge ₩/조



The Big Five



Rightsizing: Choose the optimal sized instance

Scheduling: Turn them off if you don't use them

Pricing: On-demand is easy but not the best

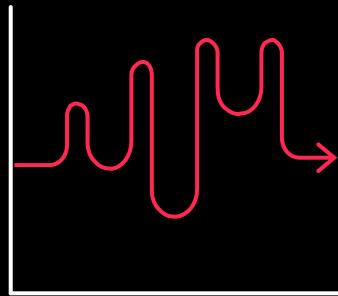


Storage: Leverage different storage classes

Monitoring: Set goals and take actions

온디맨드

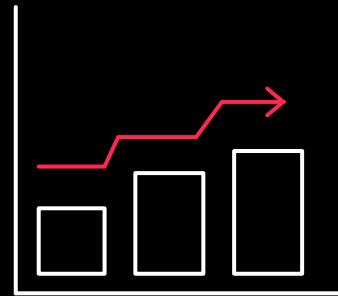
약정 없이 **사용한 EC2**
인스턴스에 대한 비용만 지불



- 트래픽이 예측 불가할 시
- 비즈니스 환경 자주 변경시

RI/SP

1년과 3년 약정으로 온디맨드
대비 **최대 75%** 절약



일정한 워크로드 및 항시
켜두어야 하는 워크로드

스팟 인스턴스

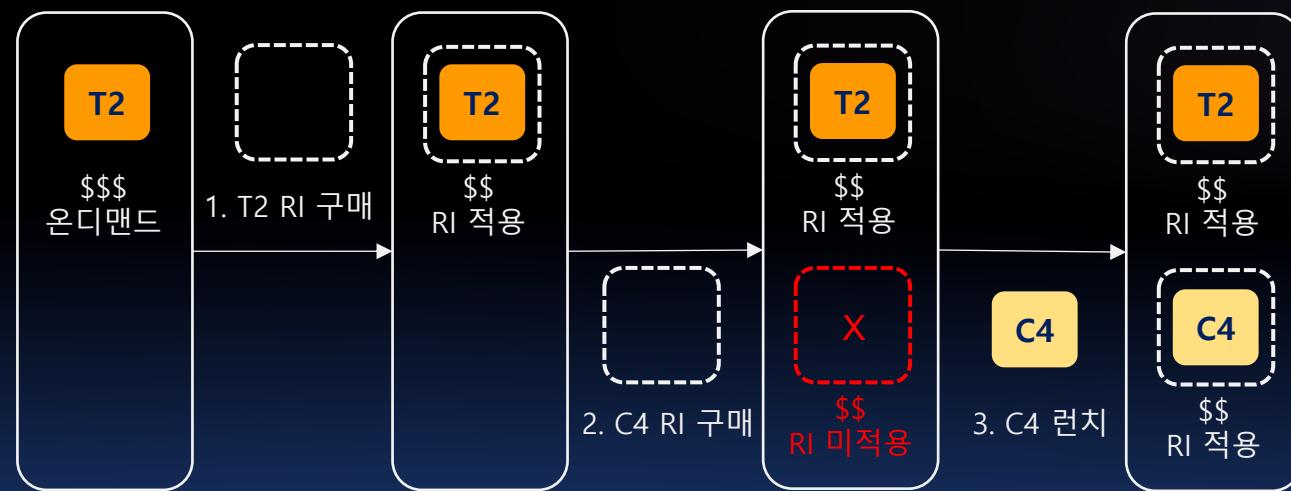
예비 컴퓨팅 용량을 통해
온디맨드 대비 **최대 90%** 절약



시간 제한이 없는 배치성
워크로드



예약 인스턴스 (Reserved Instance)



* 예약 인스턴스 작동 원리

가격 비교

On Demand



On-demand c4.xlarge
Seoul region (Linux)
720 Hours per Month

\$0.227 per Hour x 720 hours

= \$163.44 per month

= \$1961.28 per year

1 year All Upfront RI(Standard)



All Upfront RI c4.xlarge
Seoul region (Linux)
720 Hours per Month

All Upfront cost: \$1262

= \$105.16 per month ($1262/12$)

= \$1262 per year total cost (1262)

35% Savings per RI per Year

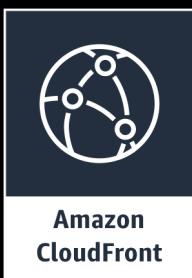
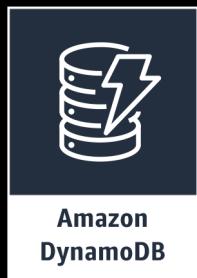
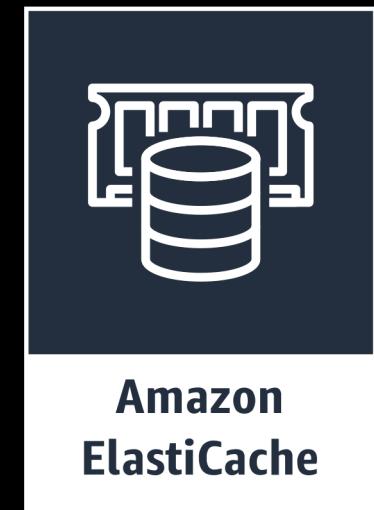
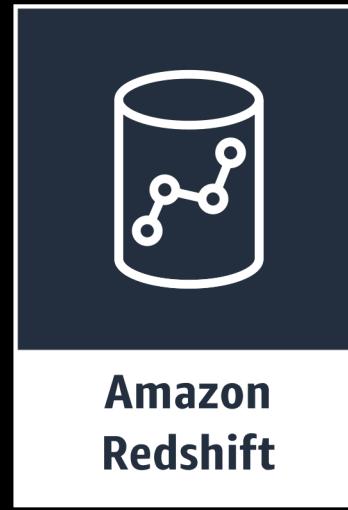


예약 인스턴스 (Reserved Instance)

제공 클래스	기간	결제 옵션
표준 (Standard)	1년	전체 선결제 (All upfront)
전환 (Convertible)	3년	부분 선결제 (Partial upfront)
인스턴스 세대 (M4 → M5) 인스턴스 패밀리 (M4 → C5) OS (Linux → Windows) 테넌시 (shared → dedicated)		선결제 없음 (No upfront)



예약 인스턴스 제공 서비스



* 예약 인스턴스는 아니지만, 약정을 통해 할인받으실 수 있는 서비스

Instance Size Flexibility

- All Regional Linux/UNIX RIs with shared tenancy now apply to all sizes of instances within an instance family and AWS region.

E.g. 1 m4.xlarge RI can apply to:

2 m4.large instances or
 $\frac{1}{2}$ an m4.2xlarge instance

- Family 타입 중에 작은 타입부터 큰 타입으로 순서대로 적용된다.



Two Types of Savings Plans



Compute Savings Plans

Offer greatest flexibility & up to 66% discount



EC2 Instance Savings Plans

Up to 72% off on the instance family within AWS region

FLEXIBLE ACROSS

- ✓ Instance family
- ✓ Region
- ✓ OS(Family 변경 필요 없음)
- ✓ Tenancy
- ✓ Compute options (EC2, Fargate)

FLEXIBLE ACROSS

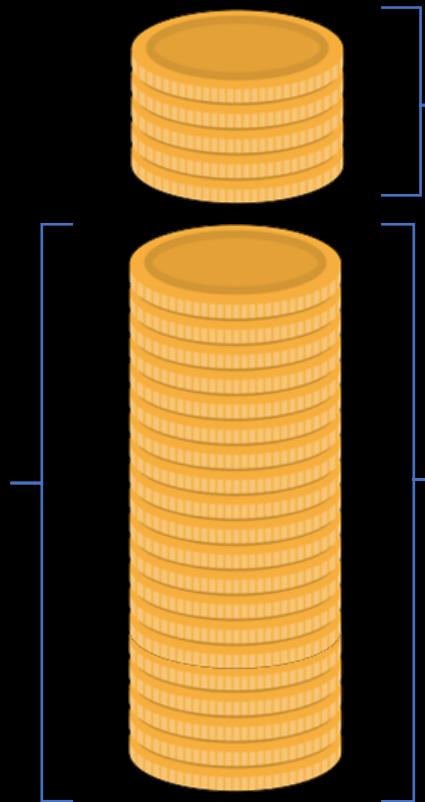
- ✓ Size
- ✓ OS
- ✓ Tenancy



Savings Based on Hourly Monetary Commitment



Commitment
of \$10.00 per
hour for
qualified
usage



On-Demand
Rate

Savings Plan
Rate

RI, SP 비교



스팟 인스턴스 (Spot Instance)

~50-90%
Savings

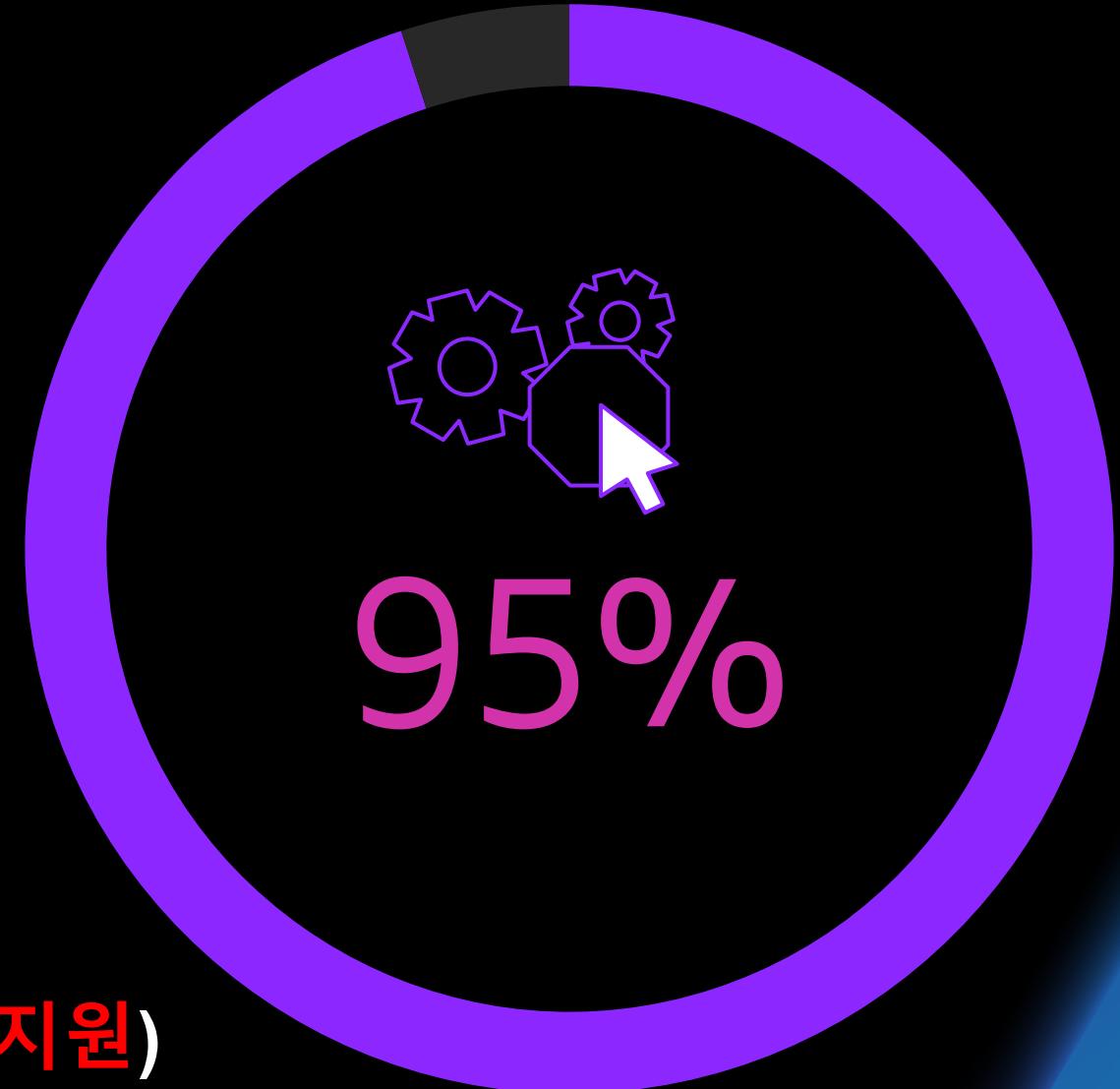
스팟 사용 Use case:

- 시간에 구애받지 않는 워크로드
- 내부 리서치 작업
- 테스트환경
- 실제 서비스에 영향없는 배치 작업

**주의: 스팟은 꺼질 수
있습니다!**

하지만...

**2분의 준비 시간을 드립니다.
(rebalance recommendation 지원)**



스팟 인스턴스 어드바이저

리전: 아시아 태평양(서울) OS: Linux/UNIX

인스턴스 유형 필터:

vCPU(분): 4 메모리 GiB(분): 3 EMR에서 지원하는 인스턴스 유형

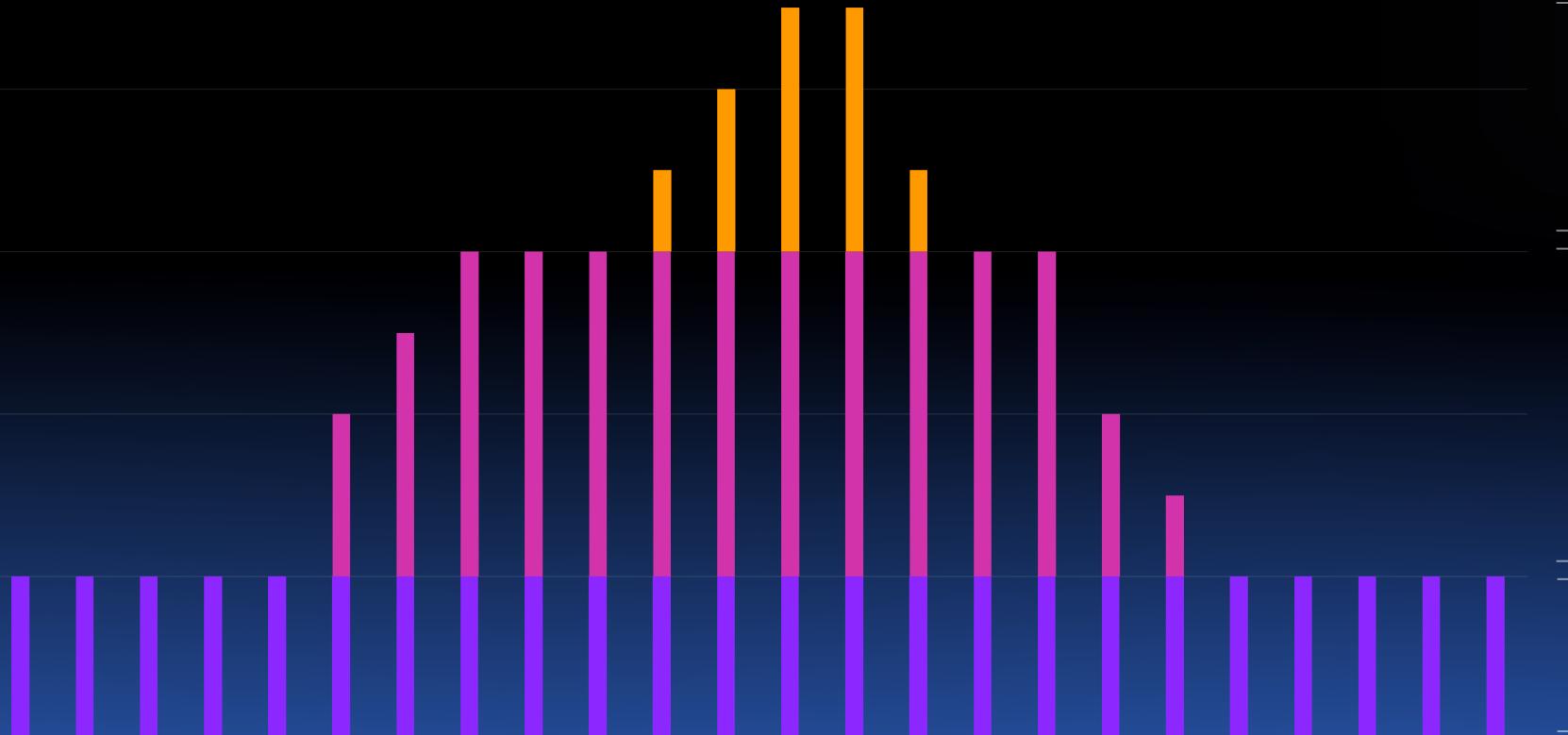
인스턴스 유형	vCPU	메모리 GiB	온디맨드 대비 절감액*	중단 빈도 ▾
c5.xlarge	4	8	69%	<5% □□□□□
r5d.12xlarge	48	384	0%	<5% □□□□□
p3.8xlarge	32	244	70%	<5% □□□□□
r5.12xlarge	48	384	0%	<5% □□□□□
i3.xlarge	4	30.5	70%	<5% □□□□□
r4.4xlarge	16	122	80%	<5% □□□□□
c5.4xlarge	16	32	69%	<5% □□□□□
x1.32xlarge	128	1952	70%	<5% □□□□□
m5d.12xlarge	48	192	0%	<5% □□□□□
m4.xlarge	4	16	76%	<5% □□□□□
p2.xlarge	4	61	70%	<5% □□□□□

중단 빈도를 AWS Website에서 확인하실 수 있습니다.

<https://aws.amazon.com/ko/ec2/spot/instance-advisor/>



온디맨드뿐만 아니라, 다양한 요금옵션을 활용하십시오



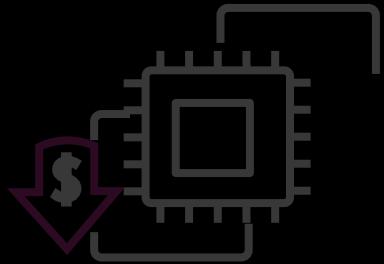
배치성이나 렌더링 워크로드는
스팟으로 운영하시길 바랍니다

예측이 불가한 워크로드에는
온디멘드를

예측이 가능하고 안정적인
워크로드는 **RI** 적용을



The Big Five



Rightsizing: Choose the optimal sized instance

Scheduling: Turn them off if you don't use them

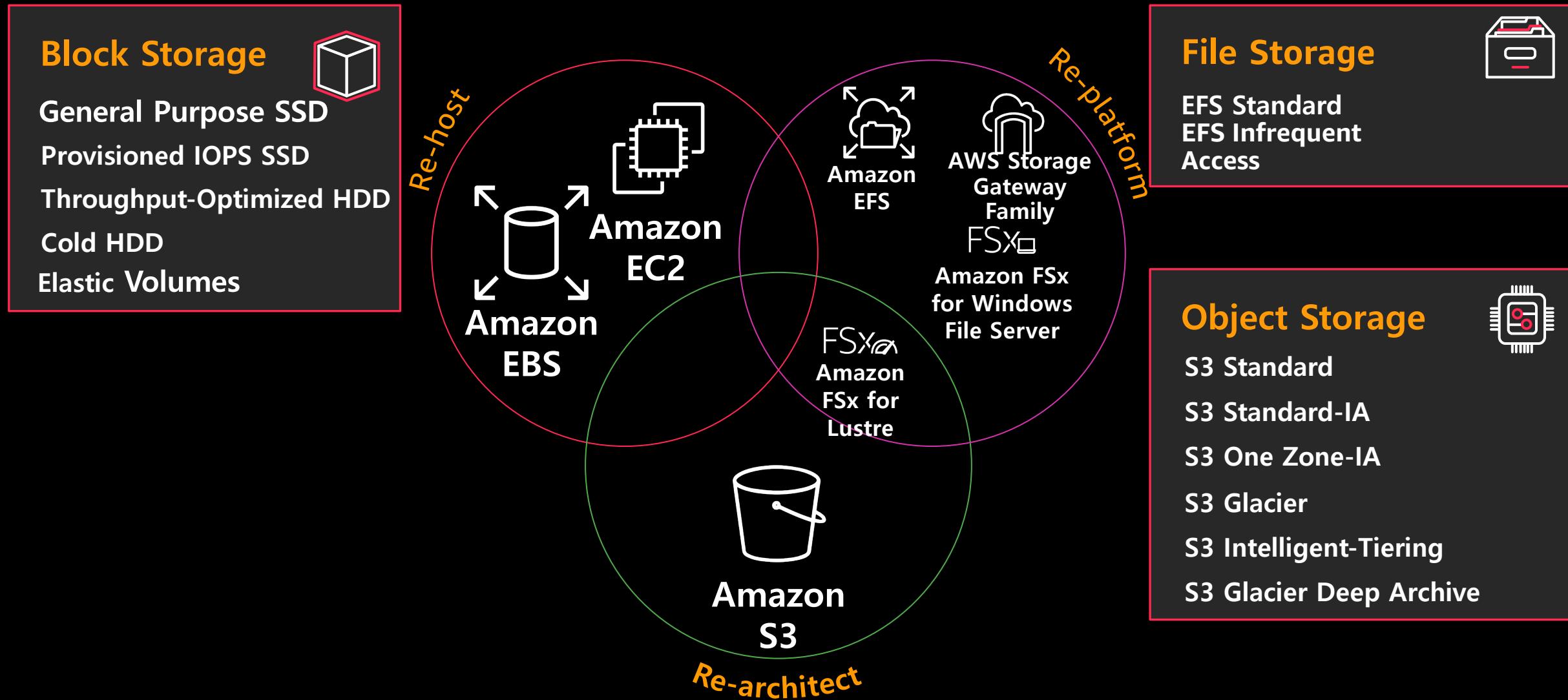
Pricing: On-demand is easy but not the best



Storage: Leverage different storage classes

Monitoring: Set goals and take actions

Broad portfolio of storage services



Your choice of Amazon S3 storage classes

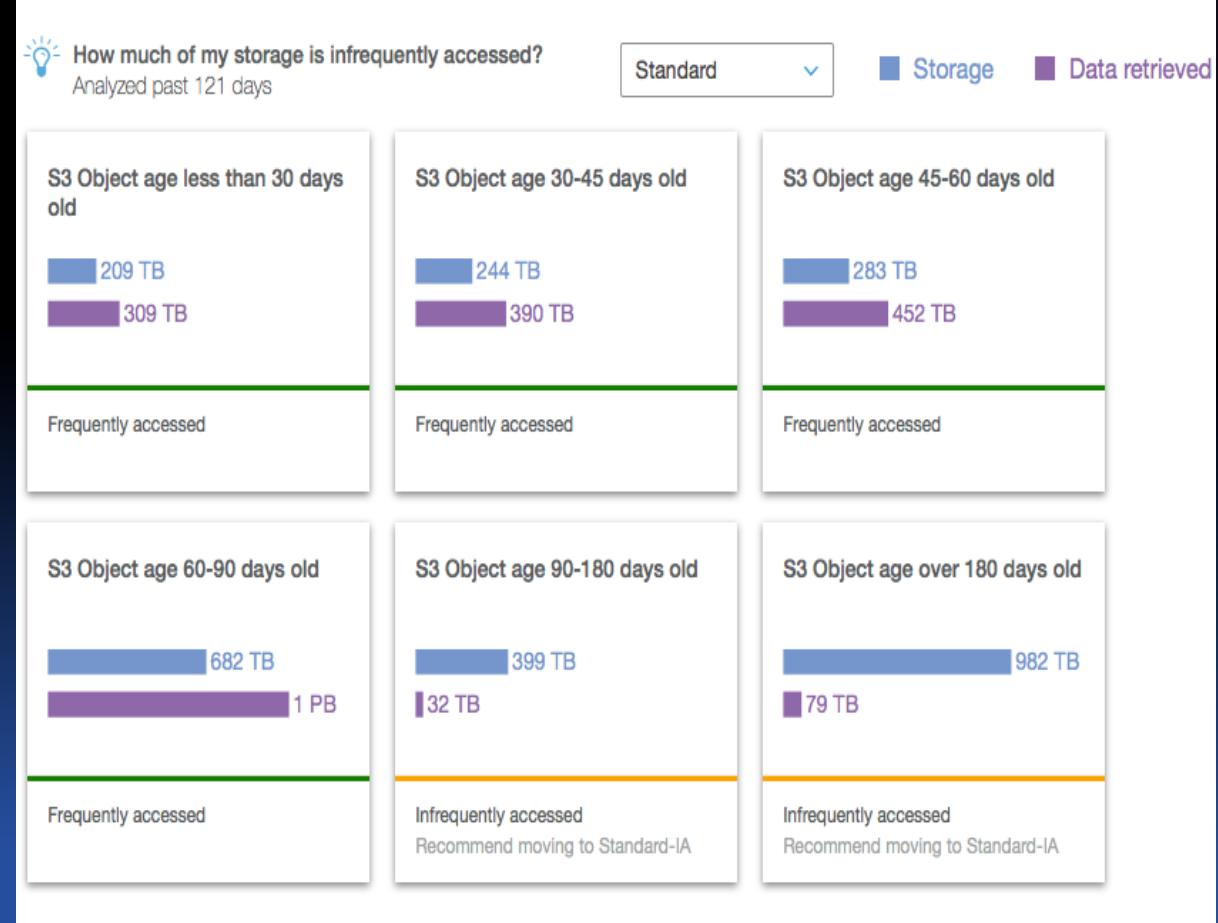
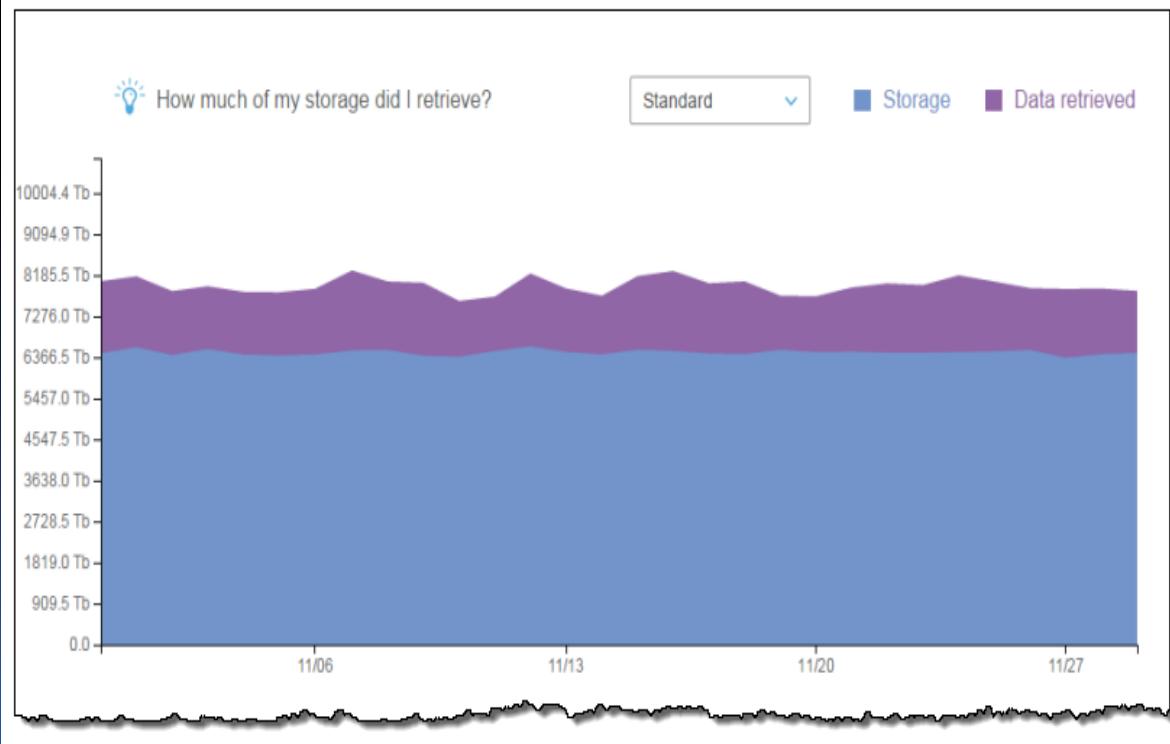


S3 Standard	S3 Intelligent-Tiering	S3 Standard-IA	S3 One Zone-IA	S3 Glacier	S3 Glacier Deep Archive
<ul style="list-style-type: none">• Active, frequently accessed data• Milliseconds access• ≥ 3 AZ• From: \$0.0210/GB	<ul style="list-style-type: none">• Data with changing access pattern• Milliseconds access• ≥ 3 AZ• From: \$0.0210 to \$0.0125/GB• Monitoring fee per obj.• Min storage duration	<ul style="list-style-type: none">• Infrequently accessed data• Milliseconds access• ≥ 3 AZ• From: \$0.0125/GB• Retrieval fee per GB• Min storage duration• Min object size	<ul style="list-style-type: none">• Re-creatable less accessed data• Milliseconds access• 1 AZ• From: \$0.0100/GB• Retrieval fee per GB• Min storage duration• Min object size	<ul style="list-style-type: none">• Archive data• Minutes to hours access• ≥ 3 AZ• From: \$0.0040/GB• Retrieval fee per GB• Min storage duration• Min object size	<ul style="list-style-type: none">• Archive data• Hours access• ≥ 3 AZ• From: \$0.00099/GB• Retrieval fee per GB• Min storage duration• Min object size

S3 Analytics provides additional functionality over standard S3 lifecycle policies

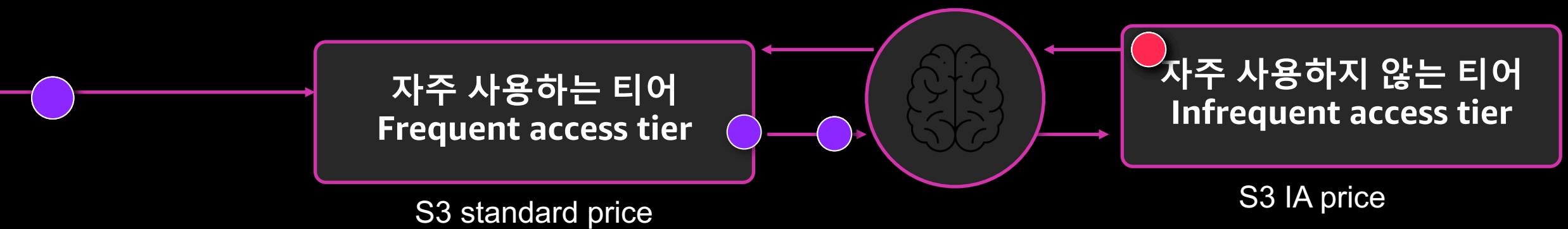
How Much of My Storage did I Retrieve?

The Amazon S3 console graphs how much of the storage in the filtered data set has been retrieved for the observation period as shown in the following example.



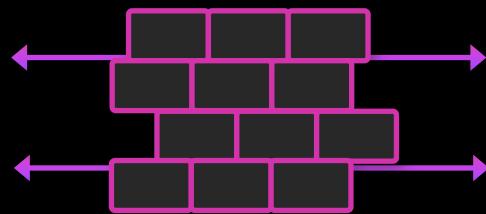
Amazon S3 인텔리전트 티어링

액세스 패턴을 분석해 스토리지 비용을 자동으로 최적화

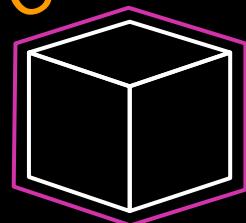


블록 스토리지 클래스 비교

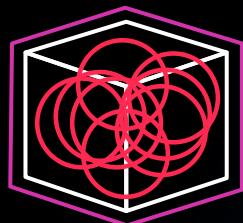
IOPS



GOOD



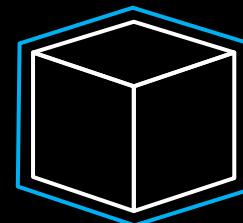
범용 SSD
gp2, gp3



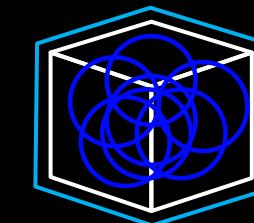
프로비저닝 된 IOPS SSD
io1, io2

VS

Throughput



처리량 최적화 HDD
st1



콜드 HDD
sc1

범용 SSD
vs

50% 절감

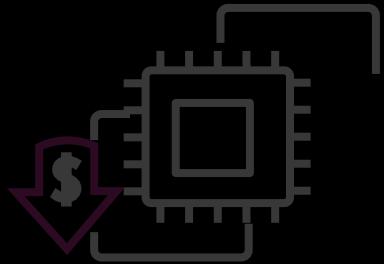
75% 절감



EBS Metrics check from Cloudwatch



The Big Five



Rightsizing: Choose the optimal sized instance



Scheduling: Turn them off if you don't use them



Pricing: On-demand is easy but not the best



Storage: Leverage different storage classes

Monitoring: Set goals and take actions

Trusted Advisor

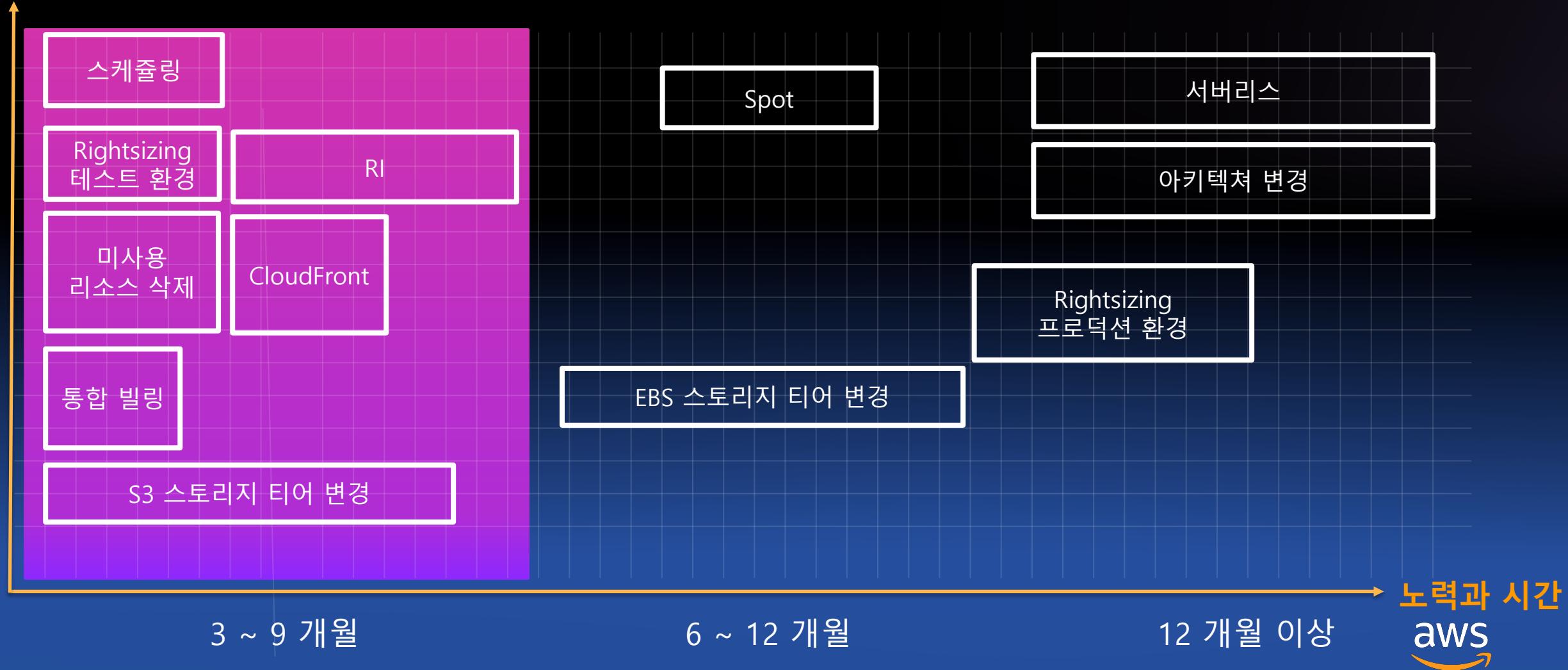
Trusted Advisor Dashboard

[Refresh All Checks](#) [Download All Checks](#)

Cost Optimizing	Performance	Security	Fault Tolerance
 0 ! 2 ! 7 ✓ 0 n/a	 0 ! 2 ! 9 ✓ 0 n/a	 2 ! 1 ! 14 ✓ 0 n/a	 0 ! 2 ! 15 ✓ 7 n/a
<ul style="list-style-type: none">! Amazon EC2 Reserved Instances Optimization! Amazon EC2 Reserved Instance Lease Expiration✓ Low Utilization Amazon EC2 Instances✓ Idle Load Balancers✓ Underutilized Amazon EBS Volumes✓ Unassociated Elastic IP Addresses✓ Amazon RDS Idle DB Instances✓ Amazon Route 53 Latency Resource Record Sets✓ Underutilized Amazon Redshift	<ul style="list-style-type: none">! Service Limits! CloudFront Content Delivery Optimization✓ High Utilization Amazon EC2 Instances✓ Amazon EBS Provisioned IOPS (SSD) Volume Attachment Configuration✓ Large Number of Rules in an EC2 Security Group✓ Large Number of EC2 Security Group Rules Applied to an Instance✓ Amazon Route 53 Alias Resource Record Sets	<ul style="list-style-type: none">! AWS CloudTrail Logging! IAM Access Key Rotation! IAM Password Policy✓ Security Groups - Specific Ports Unrestricted✓ Security Groups - Unrestricted Access✓ IAM Use✓ Amazon S3 Bucket Permissions✓ MFA on Root Account✓ Amazon RDS Security Group Access Risk	<ul style="list-style-type: none">! Amazon S3 Bucket Logging! Amazon S3 Bucket Versioning✓ Amazon EBS Snapshots✓ Amazon EC2 Availability Zone Balance✓ Load Balancer Optimization✓ VPN Tunnel Redundancy✓ Auto Scaling Group Resources✓ Amazon RDS Backups✓ Amazon RDS Multi-AZ✓ Auto Scaling Group Health Check

어려운 것부터 말고, 가능한 것부터 차근차근 시작하세요

비용절감 효과



감사합니다!

