

AWS TechCamp Online

Compute in AWS

유재서

Solutions Architect

Amazon Web Services



Agenda

Amazon EC2 개요

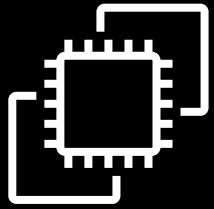
Amazon EC2 인스턴스 유형 및 선택

컴퓨팅 관련 기능 및 서비스

Amazon EC2 개요

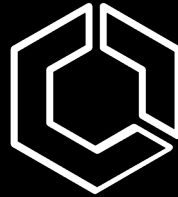


AWS 컴퓨팅 서비스



AMAZON EC2

클라우드에 위치한
가상 서버 인스턴스



AMAZON ECS, EKS, and FARGATE

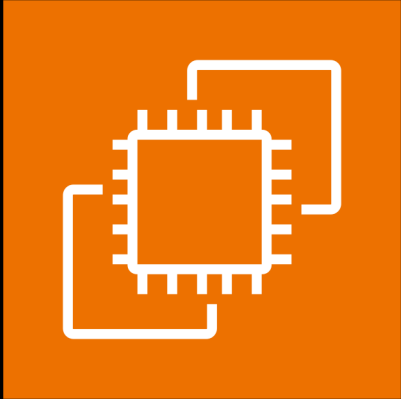
Docker 컨테이너를 대규모로
실행하고 관리할 수 있는
서비스



AWS LAMBDA

이벤트에 대한 응답으로
코드를 실행하는 서버리스
컴퓨팅

Amazon Elastic Compute Cloud (Amazon EC2)



AMAZON EC2

Linux | Windows | Mac

Arm 및 x86 아키텍처

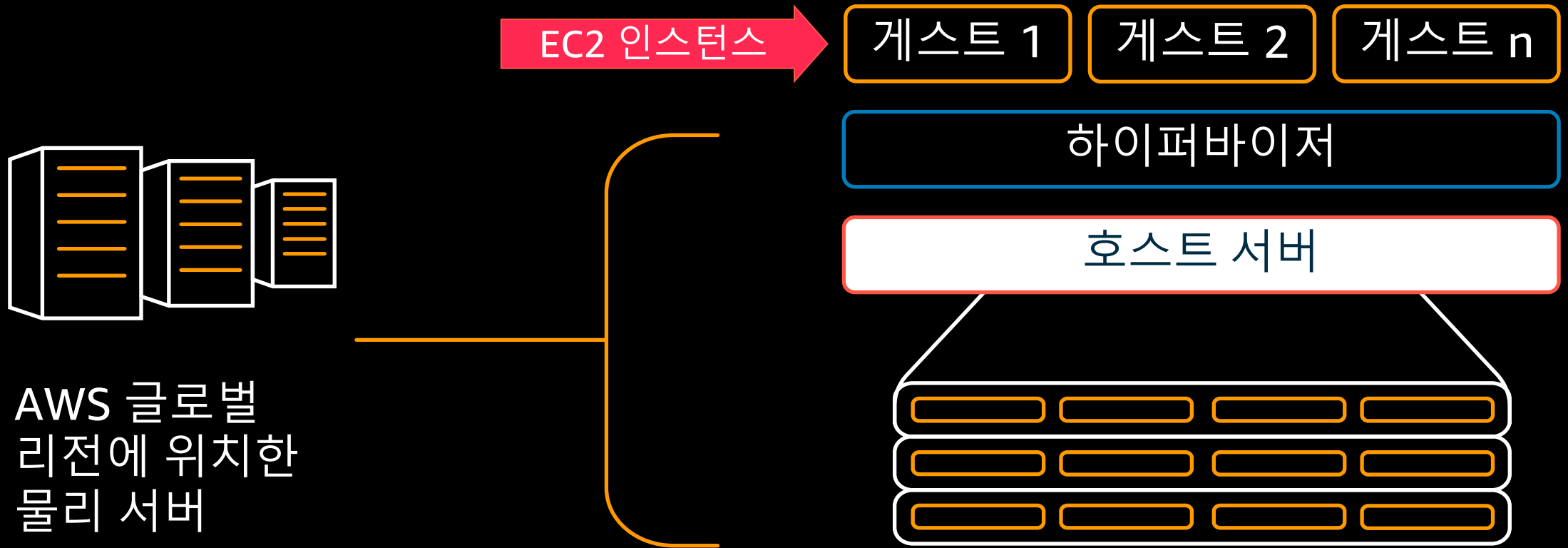
범용 및 특정 워크로드 최적화

베어 메탈, 디스크, 네트워킹 기능

Packaged | Custom | Community AMI

다양한 구매 옵션: 온디맨드 인스턴스, 스팟 인스턴스,
예약 인스턴스(RI), 절감형 플랜(Savings Plans), 전용 호스트

EC2 호스트 가상화



다양한 프로세서와 아키텍처



Intel® Xeon® Scalable
processor



AMD EPYC
processors



AWS Graviton
processor



Latest Tensor
Core GPUs

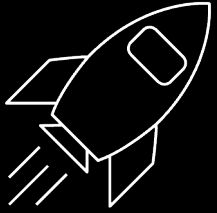


FPGAs for custom
hardware acceleration

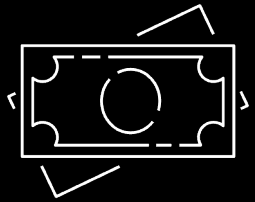


AWS Trainium/Inferentia
chips

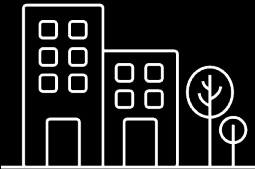
AWS Graviton 프로세서



다양한 워크로드에 대해 **최적의 가격 대비 성능 제공**



동급 x86 기반 인스턴스 대비 **비용 최대 20% 절감***



동급 x86 기반 인스턴스 대비 **에너지 소비량 최대 60% 절감 가능**

* Based on public on-demand pricing

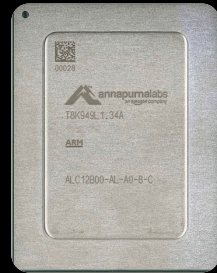
AWS Graviton 프로세서

AWS Graviton



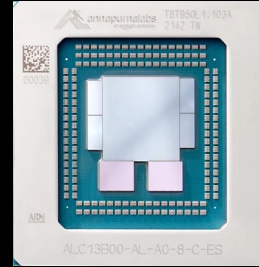
스케일 아웃(Scale-out)
워크로드에서 유사 인스턴스
대비 **최대 45% 비용 절감**

AWS Graviton2



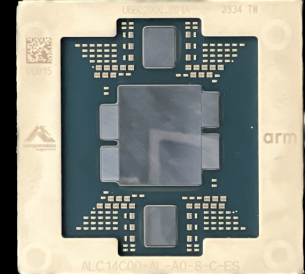
동급 인스턴스 대비 **최대 40%**
향상된 가격 대비 성능 제공

AWS Graviton3



Graviton2 대비 **최대 25% 향상된**
컴퓨팅 성능 및 최대 2배의 부동
소수점 연산 성능 제공

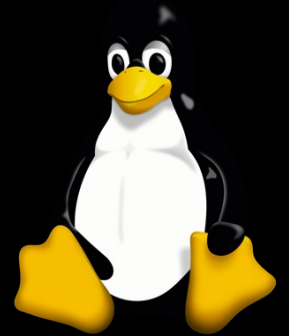
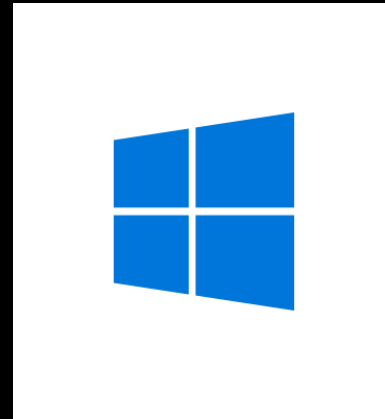
AWS Graviton4



Graviton3 대비 **최대 30% 향상된 컴퓨팅**
성능, 50% 증가된 코어 수, 75% 더 높은
메모리 대역폭 제공

EC2 운영체제

- Windows Server
- Amazon Linux
- Debian
- SUSE
- CentOS
- Red Hat Enterprise Linux (RHEL)
- Ubuntu
- Mac



더 많은 운영 체제를 확인하려면 AWS Marketplace를 방문해 주세요.

EC2 인스턴스 유형 및 선택



다양한 인스턴스 유형

카테고리	기능	옵션
범용 컴퓨팅 최적화 메모리 최적화 가속 컴퓨팅 스토리지 최적화 HPC 최적화	프로세서 선택 (AWS Graviton, Intel, AMD) 빠른 프로세서 (최대 4.5 GHz) 높은 메모리 공간 (최대 24 TiB) 인스턴스 스토리지 (HDD 및 SSD) 가속 컴퓨팅 (GPU, FPGA & ASIC) 네트워킹 (최대 800 Gbps) 베어 메탈 크기 (Nano에서 48xlarge)	Elastic Block Store (EBS) Elastic Fabric Adapter Linux, Windows, macOS

850+

인스턴스 유형

거의 모든 워크로드 및
비즈니스 요구 사항에
적합

인스턴스 표기법



카테고리	기능	옵션
<div>범용</div> <div>C</div> <div>컴퓨팅 최적화</div> <div>메모리 최적화</div> <div>가속 컴퓨팅</div> <div>스토리지 최적화</div> <div>HPC 최적화</div>	<div>g</div> <div>프로세서 선택 (AWS Graviton, Intel, AMD)</div> <div>빠른 프로세서 (최대 4.5 GHz)</div> <div>높은 메모리 공간 (최대 24 TiB)</div> <div>인스턴스 스토리지 (HDD 및 SSD)</div> <div>가속 컴퓨팅 (GPU, FPGA & ASIC)</div> <div>네트워킹 (최대 800 Gbps)</div> <div>베어 메탈</div> <div>크기 (Nano에서 48xlarge)</div> <div>n</div>	<div>Elastic Block Store (EBS)</div> <div>Elastic Fabric Adapter</div> <div>Linux, Windows, macOS</div>

인스턴스 종류

	범용		컴퓨팅 최적화		메모리 최적화				가속화된 컴퓨팅			스토리지 최적화			HPC 최적화
	버스트 가능	범용	컴퓨팅 집약적	최대 200Gbps 네트워크 대역폭	메모리 최적화	인메모리	메모리 집약적	컴퓨팅 및 메모리 집약적	그래픽 집약적	범용 GPU	FPGA	고성능 I/O	고밀도 스토리지	빅데이터 최적화	고성능 컴퓨팅
intel	T3	M7i	C7i	C6 in	R5	X2 idn	X1e		G4 dn	P4		I7 ie	D3	H1	Hpc 6id
로컬 스토리지 (NVMe SSD)		M5d	C7 gd		R7 gd			Z1d				I4g			
AMD	T3a	M7 gd	C5a		R7a				G6e	P5	F2				Hpc 7a
aws	T4g	M7g	C8g	C7 gn	R8g	X8g			G5g			I8g			Hpc 7g

인스턴스 추가 기능

g

AWS Graviton
프로세서

a

AMD 프로세서

i

인텔 프로세서

d

빠른 로컬
NVMe 스토리지

n

최대 100 Gbps
네트워크 최적화

b

블록 스토리지 최적화

e

추가 스토리지 또는 메모리

z

고주파

인스턴스는 하나 이상의 추가 기능을 가질 수 있습니다.

인스턴스 크기



8 – c7gn.xlarge
(4 vCPU, 8 GB mem)

≈



4 – c7gn.2xlarge
(8 vCPU, 16 GB mem)

≈



2 – c7gn.4xlarge
(16 vCPU, 32 GB mem)

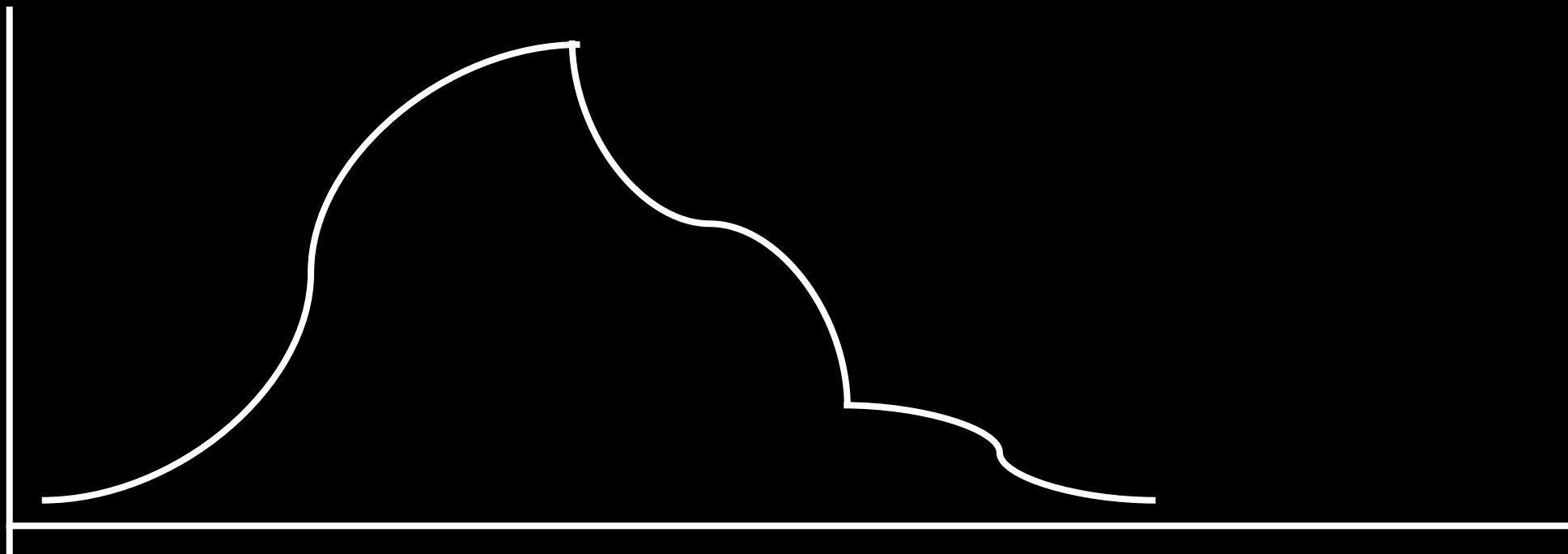
≈



c7gn.8xlarge
(32 vCPU, 64 GB mem)

인스턴스 크기

작은 인스턴스 vs 큰 인스턴스

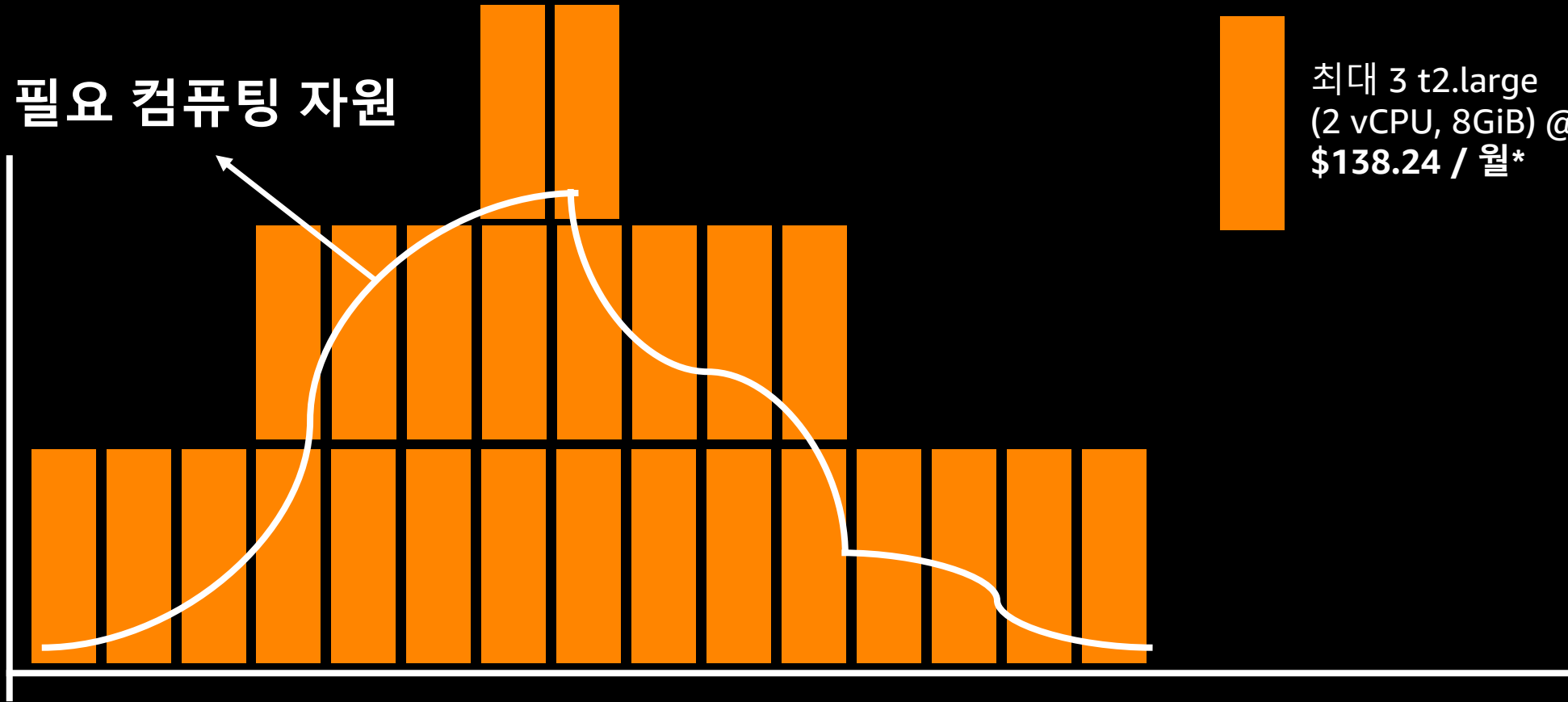


*서울 리전 리눅스 인스턴스 월 720시간(30일) 사용 사례

인스턴스 크기

작은 인스턴스 vs 큰 인스턴스

필요 컴퓨팅 자원

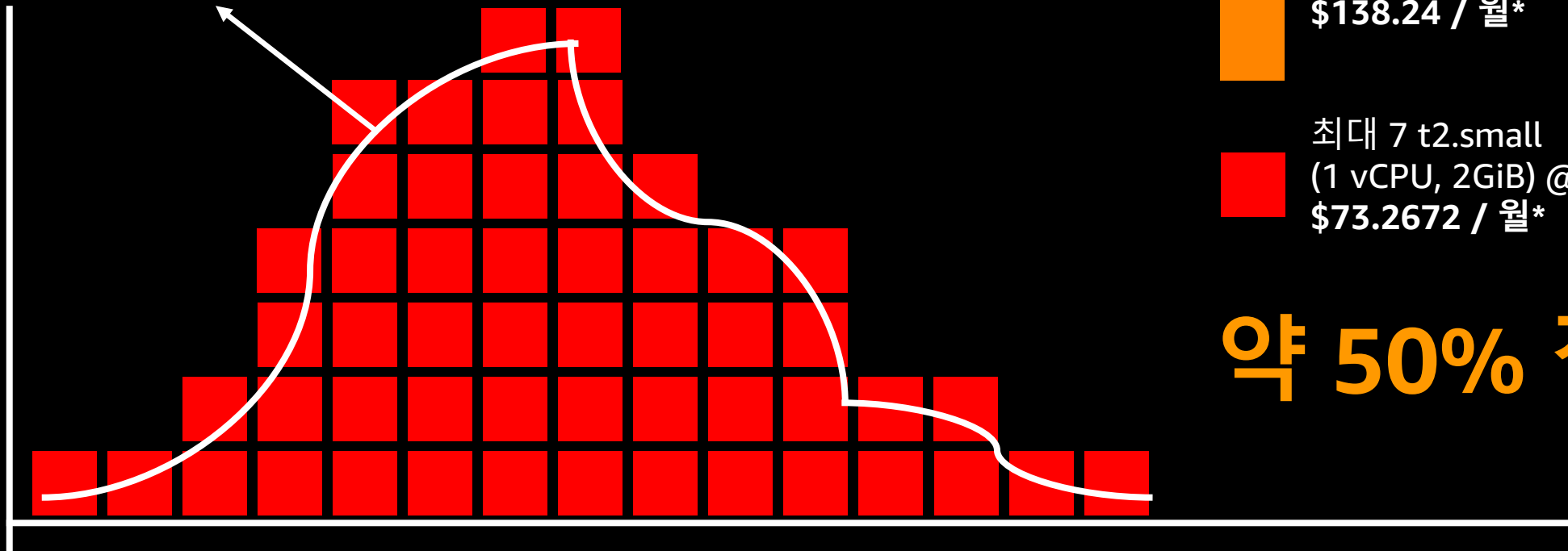


*서울 리전 리눅스 인스턴스 월 720시간(30일) 사용 사례

인스턴스 크기

작은 인스턴스 vs 큰 인스턴스

필요 컴퓨팅 자원



최대 3 t2.large
(2 vCPU, 8GiB) @ \$0.1152 /시간
\$138.24 / 월*

최대 7 t2.small
(1 vCPU, 2GiB) @ \$0.0288/시간
\$73.2672 / 월*

약 50% 절약

*서울 리전 리눅스 인스턴스 월 720시간(30일) 사용 사례

최적의 인스턴스 선택 (신규 워크로드)

▼ Instance type [Info](#) [Get advice](#)

Instance type

t2.micro

Family: t2 1 vCPU 1 GiB Memory Current generation: true

On-Demand Windows base pricing: 0.0162 USD per Hour

On-Demand Ubuntu Pro base pricing: 0.0134 USD per Hour

On-Demand SUSE base pricing: 0.0116 USD per Hour

On-Demand RHEL base pricing: 0.026 USD per Hour

On-Demand Linux base pricing: 0.0116 USD per Hour

Free tier eligible

☐ All generations

[Compare instance types](#)

Additional costs apply for AMIs with pre-installed software

Get advice on instance type selection

Tell us about your instance type requirements and we'll provide suggestions for you.

Workload type

Web/App Server

Use case

Web Hosting

Priority

Price/Performance

CPU manufacturer

No preference

▼ Advanced parameters

Add parameters to narrow down the list of recommended instance types

Select a parameter

Add

Minimum number of vCPUs

Specify a value e.g. 1

Maximum number of vCPUs

Specify a value e.g. 32

Minimum instance storage (GB)

Specify a value e.g. 100

Maximum instance storage (GB)

Specify a value e.g. 950

Cancel

Get instance type advice

Get advice on instance type selection

EC2 instance type finder recommends instance families: M8g, C8g, C7gn, M7gd, C7gd

View recommended instance family details

▼ Additional information

1. The recommended instances are based on the AWS Graviton processor platform, suitable for Application hosting use case with Web/App Server workload type, prioritizing Performance as per the user preference.
2. M8g and C8g instances, powered by the latest-generation AWS Graviton4 processors, offer the highest performance among the recommended options, with up to 30% better performance compared to their Graviton3-based predecessors.
3. M8g instances are general-purpose and provide the best price performance for a wide range of workloads, making them well-suited for application hosting scenarios.
4. C8g instances are compute-optimized, offering high performance for compute-intensive web/app server workloads that require sustained CPU performance.
5. C7gn instances are network-optimized, delivering up to 200 Gbps of network bandwidth and up to 3x higher packet-processing performance per vCPU compared to x86-based network-optimized instances, which can be beneficial for high-traffic web applications.
6. M7gd and C7gd instances, powered by AWS Graviton3 processors, include local NVMe-based SSDs, offering up to 45% better real-time NVMe storage performance than comparable Graviton2-based instances. This can be advantageous for applications requiring fast local storage access.
7. The M7gd instances provide a balance of compute, memory, and storage resources, while C7gd instances are more compute-optimized with local storage, both suitable for different web/app server workload profiles.
8. For applications requiring high network bandwidth, C7gn instances are available, offering enhanced network performance for data-intensive or high-traffic web applications.
9. M7gd and C7gd instances are available for workloads that benefit from local storage, providing fast access to data stored on the instance.

Leave Feedback

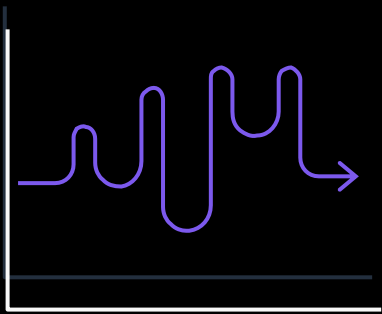
Try again

Close

Amazon EC2 구매 옵션

온디맨드 (On-Demand)

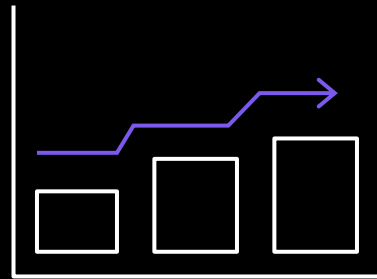
장기 약정 없이 **초 단위로**
컴퓨팅 용량 비용 지불



요구 사항을 정의하기
위한 급증하는 워크로드

예약형 인스턴스 (Reserved Instances, RI)

1년 또는 3년 약정으로
온디맨드 요금에서
상당한 할인



약정 및
안정적인 상태의 사용

절감형 플랜 (Savings Plans)

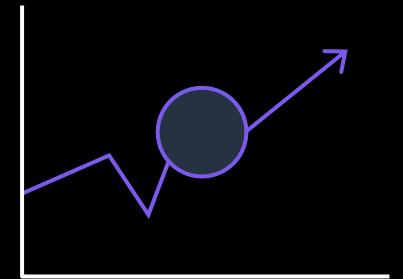
더 많은 유연성을 갖춘
Amazon EC2 RI와 동일한
대폭 할인



약정 및
유연한 요금 모델

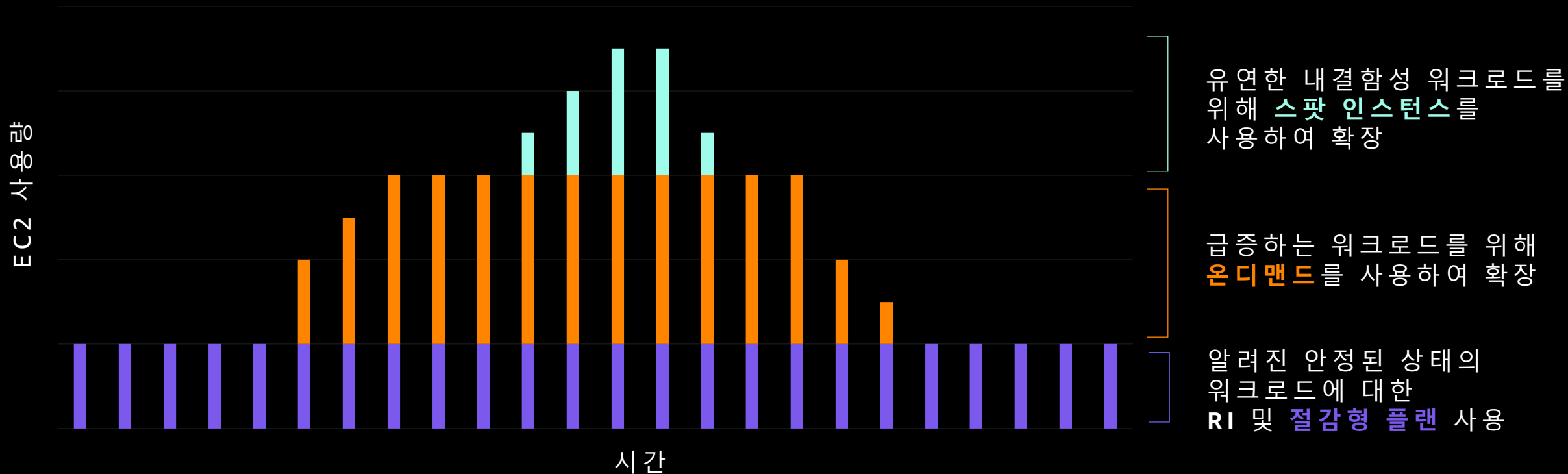
스팟 인스턴스 (Spot Instances)

예비 Amazon EC2 컴퓨팅 용량으로
온디맨드 요금에 비해
최대 90% 할인된 가격으로 사용



유연한 상태 비저장,
내결함성 워크로드

비용 및 용량 최적화



컴퓨팅 관련 기능 및 서비스

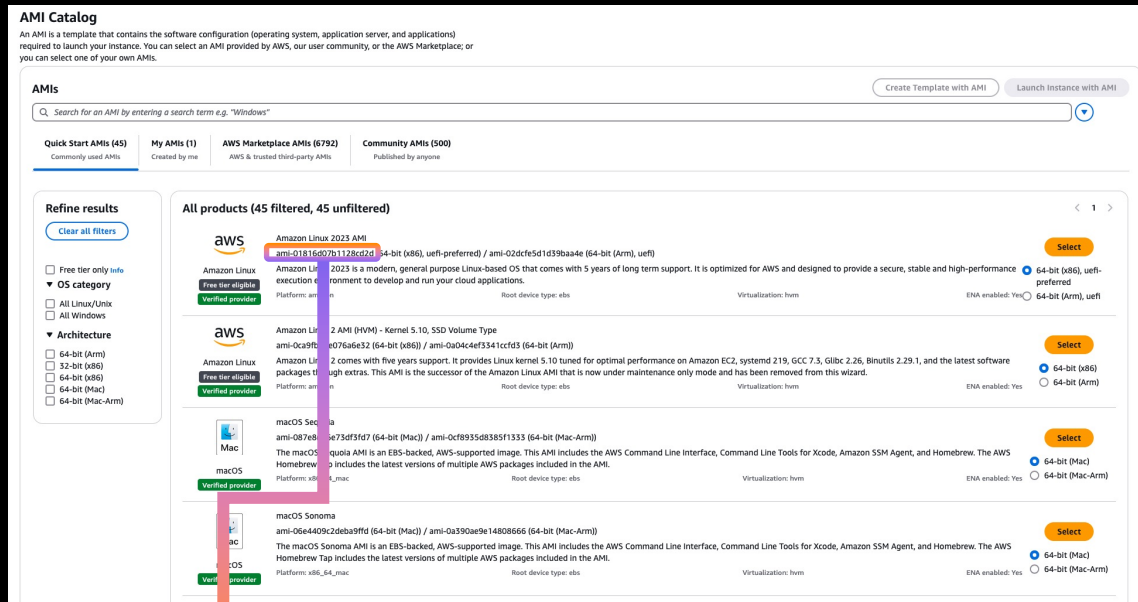


Amazon Machine Image(AMI)

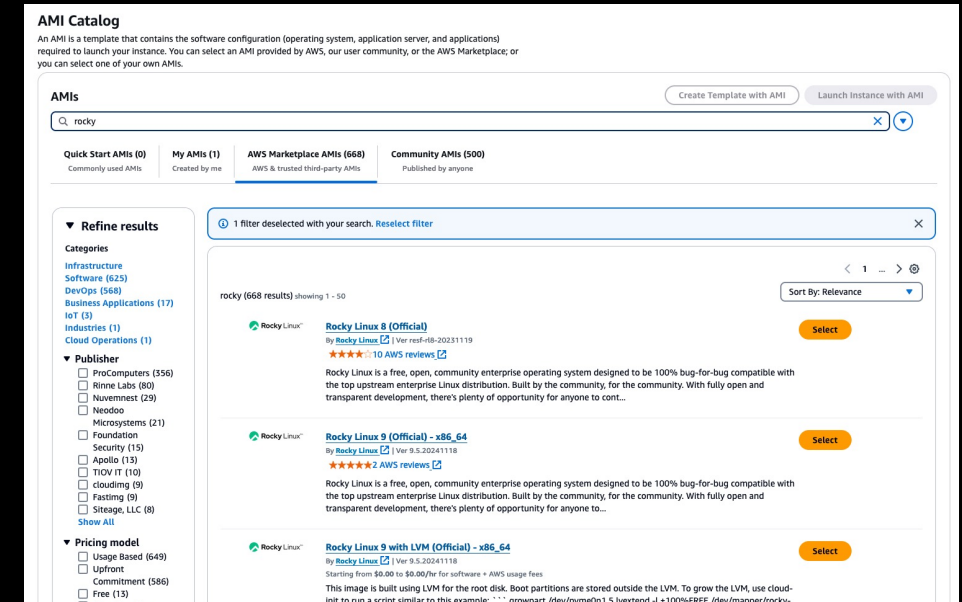
- 인스턴스 시작에 필요한 정보 제공
- 동일한 구성으로 한 AMI에서 여러 인스턴스 시작 가능
- AMI는 다음을 포함합니다.
 - 1개 이상의 Amazon Elastic Block Store (Amazon EBS) 스냅샷 또는 루트 볼륨에 대한 템플릿(운영 체제, 애플리케이션)
 - AMI를 사용하여 인스턴스를 시작할 수 있는 AWS 계정을 제어하는 시작 권한
 - 인스턴스에 연결할 볼륨을 지정하는 블록 디바이스 매핑

AMI 선택

AWS Packaged AMI



AWS Marketplace AMI



AMI ID를 사용하여 API 또는 AWS 명령줄 인터페이스(AWS CLI)를 통해 시작
`aws ec2 run-instances --image-id ami-01816d07b1128cd2d --instance-type c4.xlarge --count 10 --key-name MyKey`



사용자 데이터 (User data)

일반적인 구성 작업을 자동으로 수행

Shell 스크립트 방식

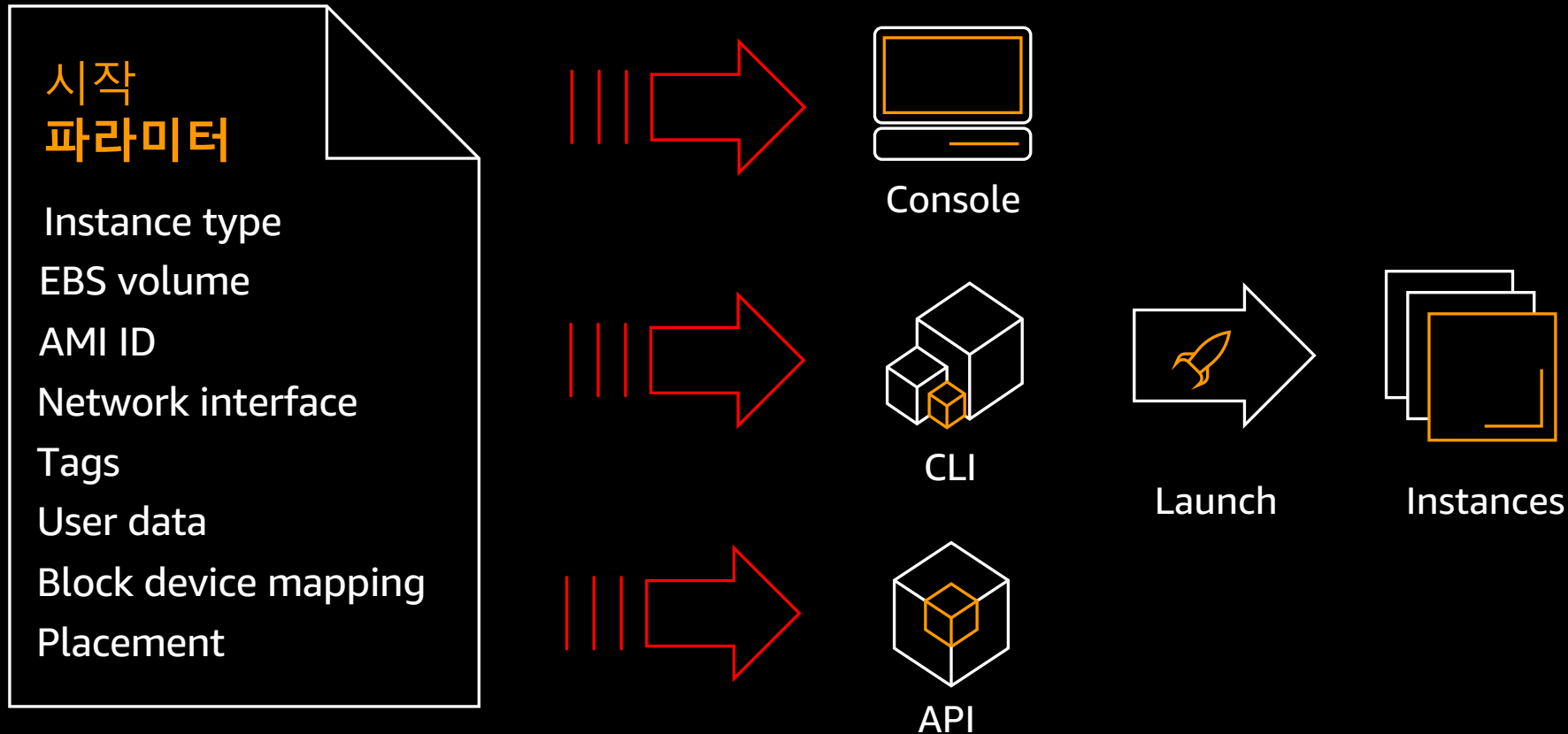
```
#!/bin/bash
yum update -y
amazon-linux-extras install -y lamp-mariadb10.2-php7.2 php7.2
yum install -y httpd mariadb-server
systemctl start httpd
systemctl enable httpd
```

Shell 스크립트

- `#!` 문자 및 스크립트를 읽을 인터프리터의 경로(일반적으로 `/bin/bash`)로 시작
- root 사용자 권한으로 실행
 - **sudo** 명령 불필요
 - 생성하는 모든 파일의 소유권자는 root
 - root 이외의 사용자에게 파일 액세스를 허용하려면 권한 수정 필요
- 대화형으로 실행되지 않으므로 사용자의 입력이 필요한 명령은 포함할 수 없음
 - 예: `-y` 플래그 없는 **yum update**

시작 템플릿 (Launch Template)

시작(launch)을 간소화하고 단순화하기 위해 시작 요청을 템플릿화



일관된 경험

단순한 권한

거버넌스 및
모범 사례

생산성 향상

Amazon EC2 관련 자격 증명

EC2 키 페어

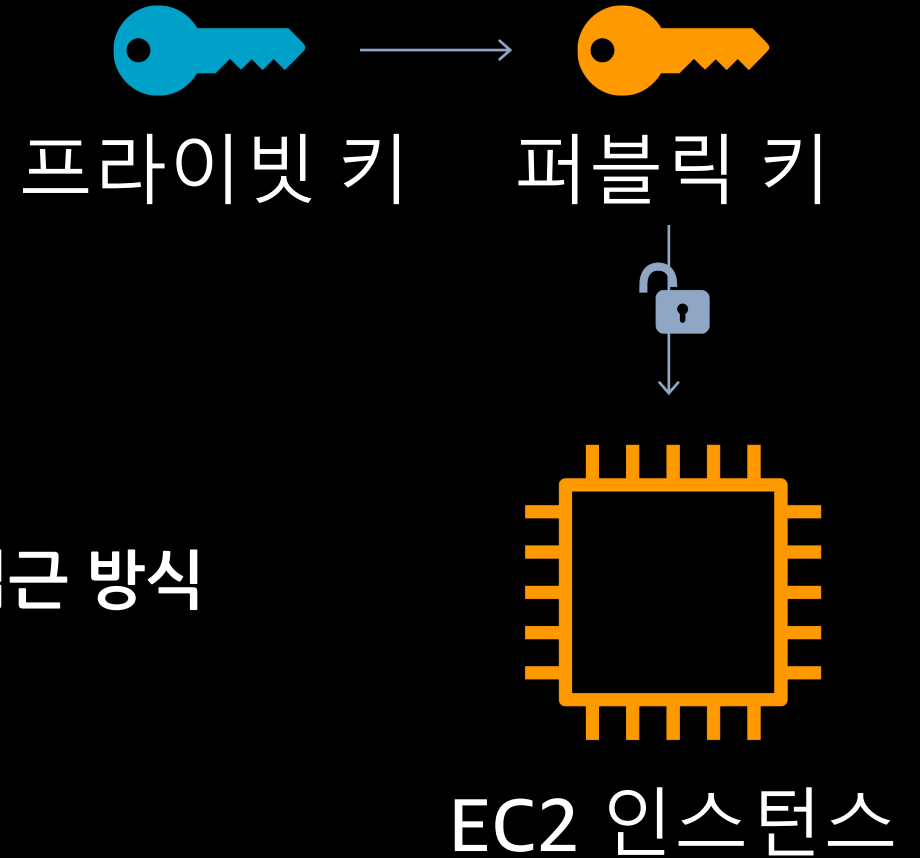
- Linux – 최초 호스트 로그인을 위한 SSH 키 페어
- Windows – Windows 관리자 암호 검색

표준 SSH RSA 키 페어

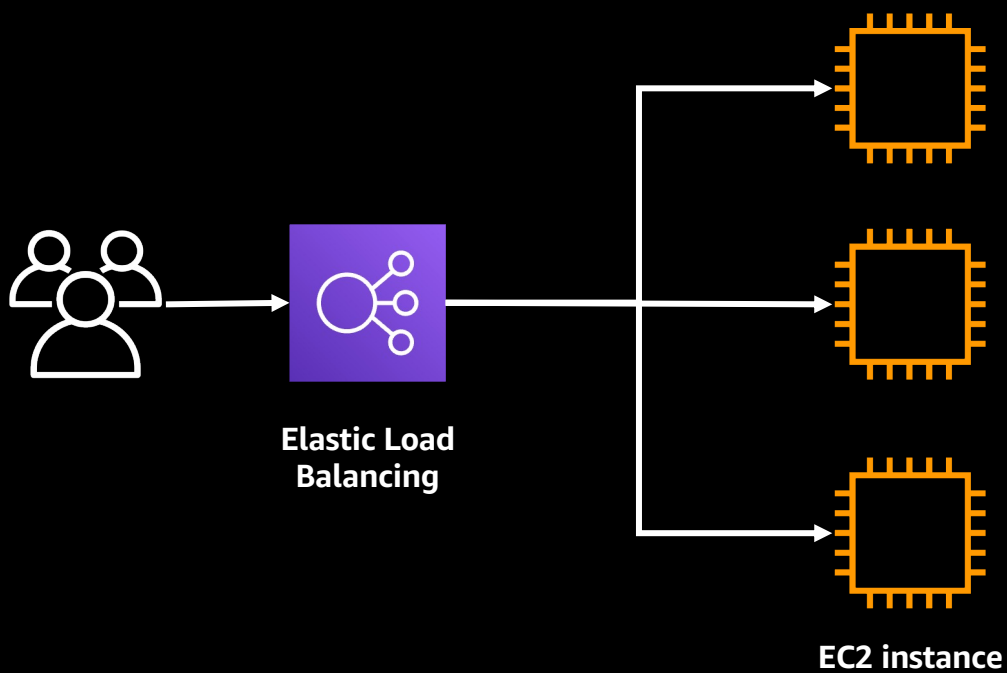
- 퍼블릭 키, 프라이빗 키
- AWS에는 프라이빗 키가 보관되지 않음

일반 OS에 대한 초기 액세스를 제공하기 위한 AWS 접근 방식

- 안전한
- 개인화된
- 비제네릭 (NIST, PCI DSS)

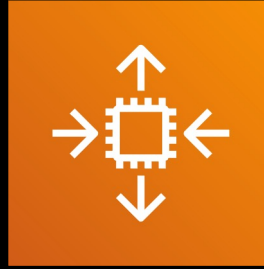


Elastic Load Balancing (ELB)



- 네트워크 트래픽 분산을 통한 애플리케이션 확장성 개선
- 여러 가용 영역을 기반으로 고가용성 제공
- 트래픽에 따라 자동 조정
- 종류
 - Application Load Balancer
 - Network Load Balancer
 - Gateway Load Balancer

Amazon EC2 Auto Scaling



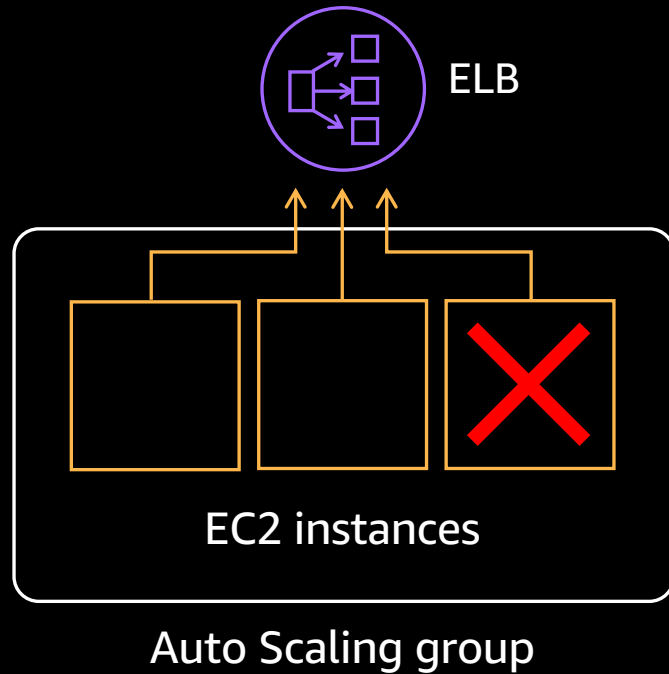
Amazon EC2 Auto Scaling은 정의된 조건에 따라 애플리케이션의 부하를 처리하기 위해 **EC2 인스턴스를 자동으로 추가하거나 제거**하여 애플리케이션의 **가용성을 유지하는 완전 관리형 서비스**입니다.

- **Groups**
 - 수요에 맞게 확장할 수 있는 논리적으로 관리되는 EC2 인스턴스 컬렉션
- **Configuration templates**
 - EC2 인스턴스를 시작하는 데 사용되는 템플릿
- **Scaling options**
 - Auto Scaling 확장 방법

Amazon EC2 Auto Scaling

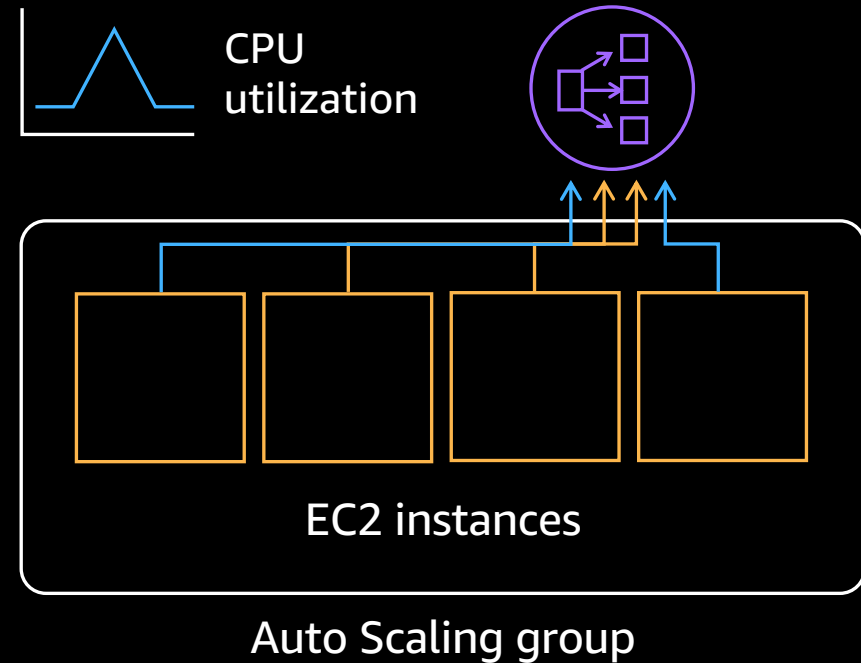
Fleet management

비정상 인스턴스 교체



Dynamic scaling

수요에 맞게 확장



Wrap up

- Amazon EC2란 손쉽게 사용 가능한 컴퓨팅 파워를 제공하는 가상머신 서비스
- 모든 워크로드를 위한 고성능의, 비용 효율적인 다양한 컴퓨팅 인스턴스 타입을 제공
- Linux, window, Mac 등 다양한 유저 환경에 맞는 AMI 제공
- On-demand는 물론, RI, Savings Plan 등 다양한 구매 옵션을 제공
- User data와 시작 템플릿을 통한 편리하고 일관성 있는 환경 구축
- Auto Scaling과 ELB를 활용해 가용성, 확장성, 비용 효율적 아키텍처 구성

Thank you!

