

대규모 서비스를 위한 **AWS**의 대표적인 **NoSQL Database** 서비스 알아보기

이 혁

솔루션즈 아키텍트, AWS

Agenda

- 모던 애플리케이션과 데이터베이스
- SQL or NoSQL?
- AWS의 대표적인 NoSQL 데이터베이스 서비스

모든 애플리케이션과 데이터베이스

관계형과 NoSQL 데이터베이스

다양한 문제를 해결하기 위해 그 어느 때보다 많은 데이터베이스가 새롭게 등장

381 systems in ranking, December 2021

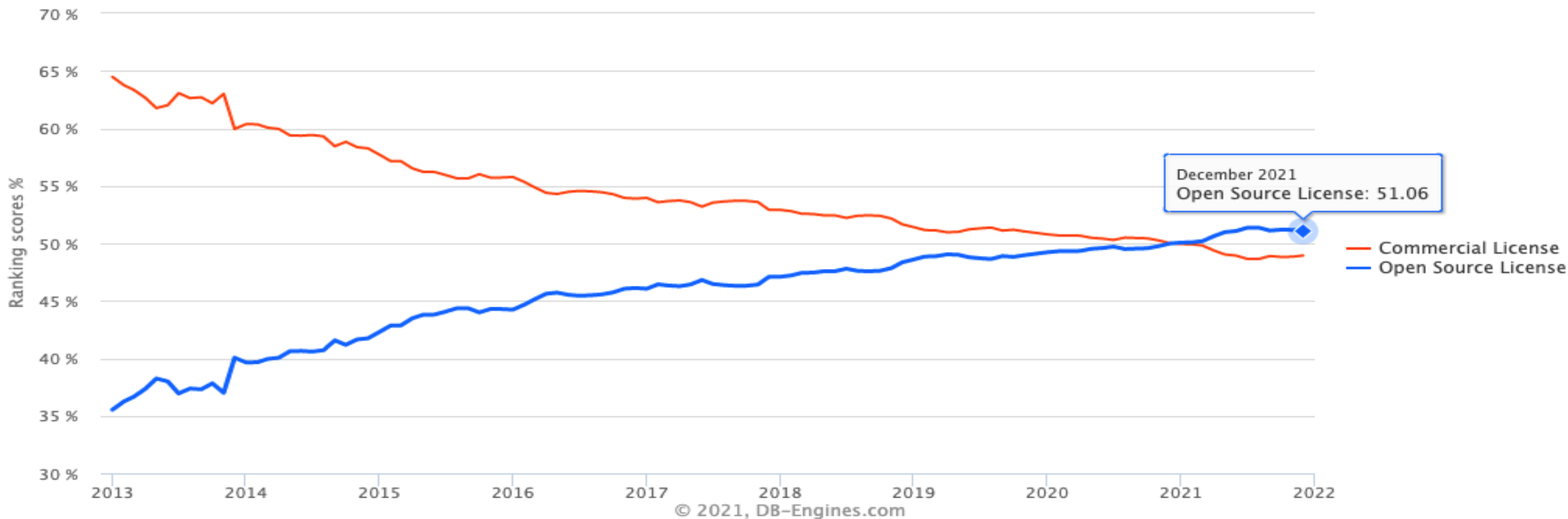
Rank			DBMS	Database Model	Score		
Dec 2021	Nov 2021	Dec 2020			Dec 2021	Nov 2021	Dec 2020
1.	1.	1.	Oracle +	Relational, Multi-model i	1281.74	+9.01	-43.86
2.	2.	2.	MySQL +	Relational, Multi-model i	1206.04	-5.48	-49.41
3.	3.	3.	Microsoft SQL Server +	Relational, Multi-model i	954.02	-0.27	-84.07
4.	4.	4.	PostgreSQL +	Relational, Multi-model i	608.21	+10.94	+60.64
5.	5.	5.	MongoDB +	Document, Multi-model i	484.67	-2.67	+26.95
6.	6.	↑ 7.	Redis +	Key-value, Multi-model i	173.54	+2.04	+19.91
7.	7.	↓ 6.	IBM Db2	Relational, Multi-model i	167.18	-0.34	+6.74
8.	8.	8.	Elasticsearch	Search engine, Multi-model i	157.72	-1.36	+5.23
9.	9.	9.	SQLite +	Relational	128.68	-1.12	+7.00
10.	↑ 11.	↑ 11.	Microsoft Access	Relational	125.99	+6.75	+9.25

<https://db-engines.com>

오픈소스 데이터베이스

기업들은 왜 오픈소스 데이터베이스에 주목하는가?

Popularity trend



<https://db-engines.com>

모던 애플리케이션

사용자	100만 이상
데이터 양	테라바이트 ~ 엑사바이트
성능	1/1000 초 ~ 1/1,000,000 초
요청 량	초당 수백만
접근	다양한 디바이스
확장규모	무제한
지불 방식	사용한 만큼
개발자 접근	관리형 API



인터넷 스케일의 모던 애플리케이션

Amazon Web Services 한국 블로그

Amazon Prime Day 2021 - 이틀간의 클라우드 기반 신기록

by Jeff Barr | on 03 JUL 2021 | in Amazon EC2, News | Permalink | [Share](#)

매년 해왔던 것처럼(2016년, 2017년, 2019년, 2020년 게시물 참고) 올해 Prime Day의 지표를 공유하고 AWS가 어떻게 이런 결과를 달성하는데 도움이 되었는지 알려드리고자 합니다.

올해 저는 개인적으로 MagicMirror로 사용할 도시바 43인치 스마트 TV와 물 뿌리개 몇 개, 작업장용 드레멜 로터리툴 세트 등 각종 유용한 제품을 구입했습니다.

Amazon Aurora — “Amazon Fulfillment Technologies(AFT)는 Amazon에서 이뤄진 구매에 대해 실제 주문 처리를 지원합니다. Prime Day에 AFT의 PostgreSQL 호환 버전 Amazon Aurora 인스턴스 3,715개는 2,330억 건의 트랜잭션을 처리하고 1,595테라바이트의 데이터를 저장하며, 615테라바이트의 데이터를 전송했습니다.”

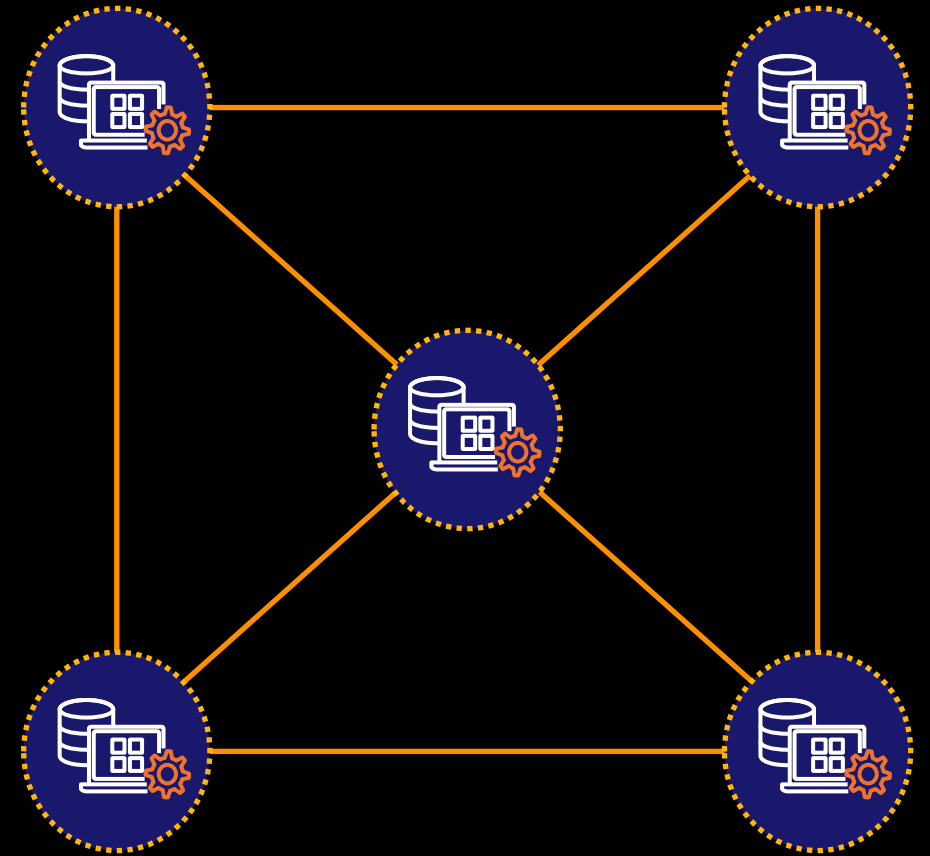
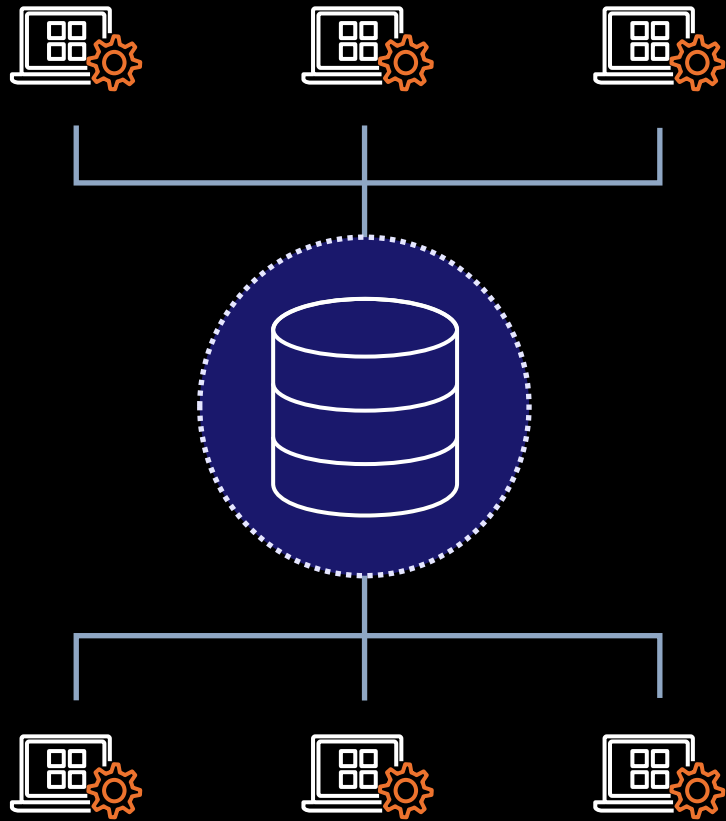
Amazon DynamoDB – “DynamoDB는 Alexa, Amazon.com 사이트 및 모든 Amazon 물류 처리 센터 등 트래픽이 많은 다수의 Amazon 시설과 시스템을 지원합니다. 66시간의 Prime Day 기간 동안 이러한 소스는 수조 건의 API 호출을 수행하면서 한 자릿수 밀리초 성능을 통해 고가용성을 유지하고 초당 최대 8,920만 건의 요청을 소화했습니다.”



<https://aws.amazon.com/ko/blogs/korea/prime-day-2021-two-chart-topping-days/>

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

마이크로 서비스 아키텍처

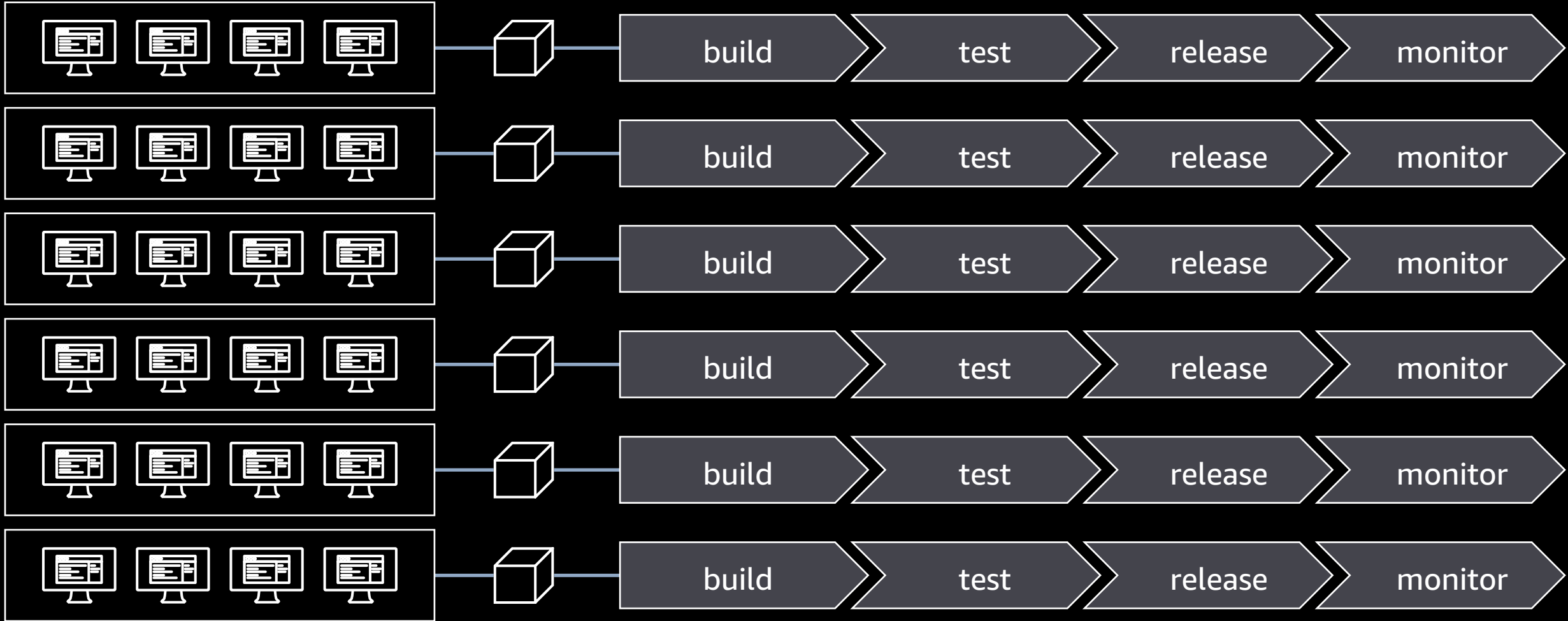


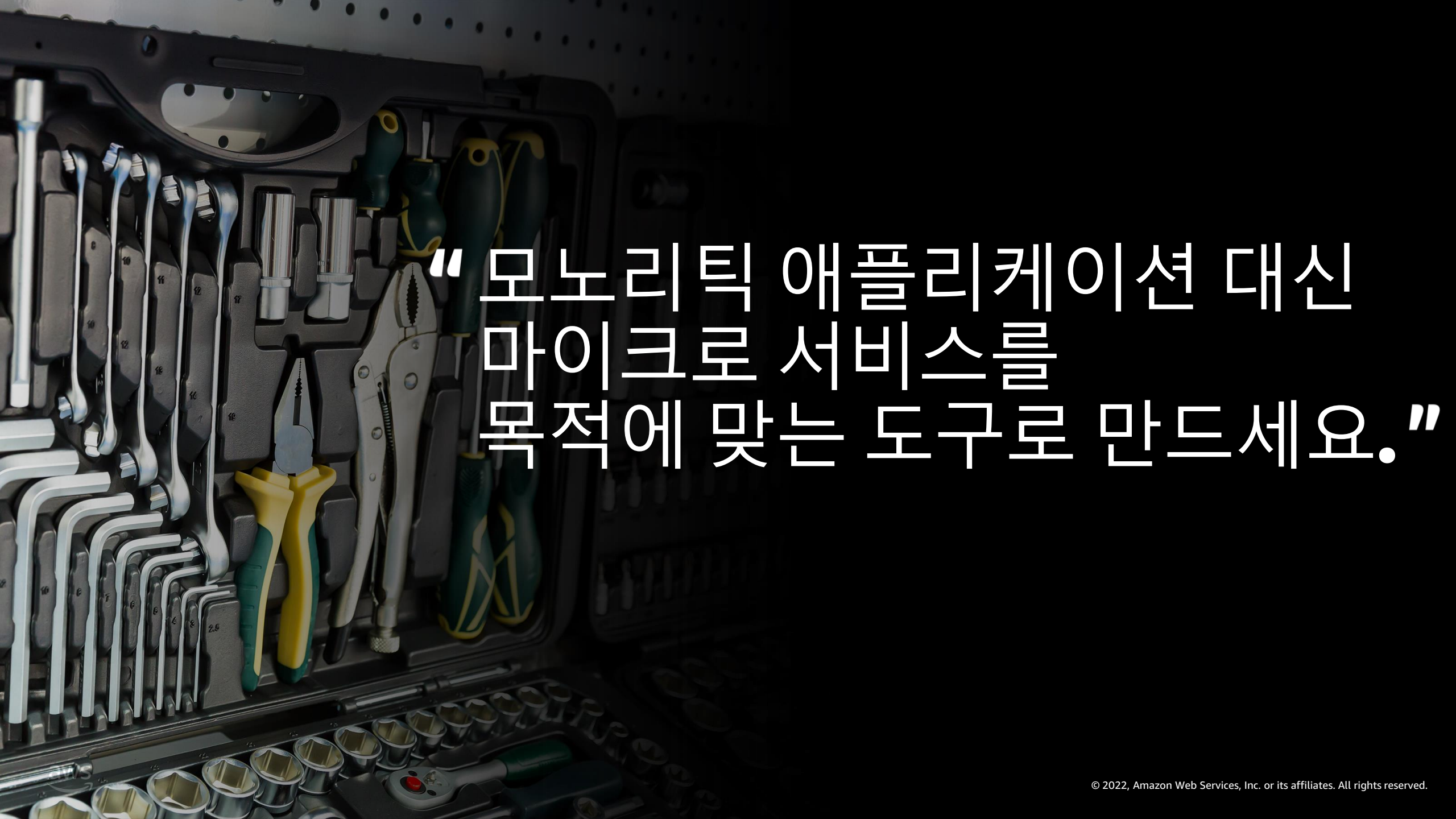
DevOps

DevOps 팀

서비스

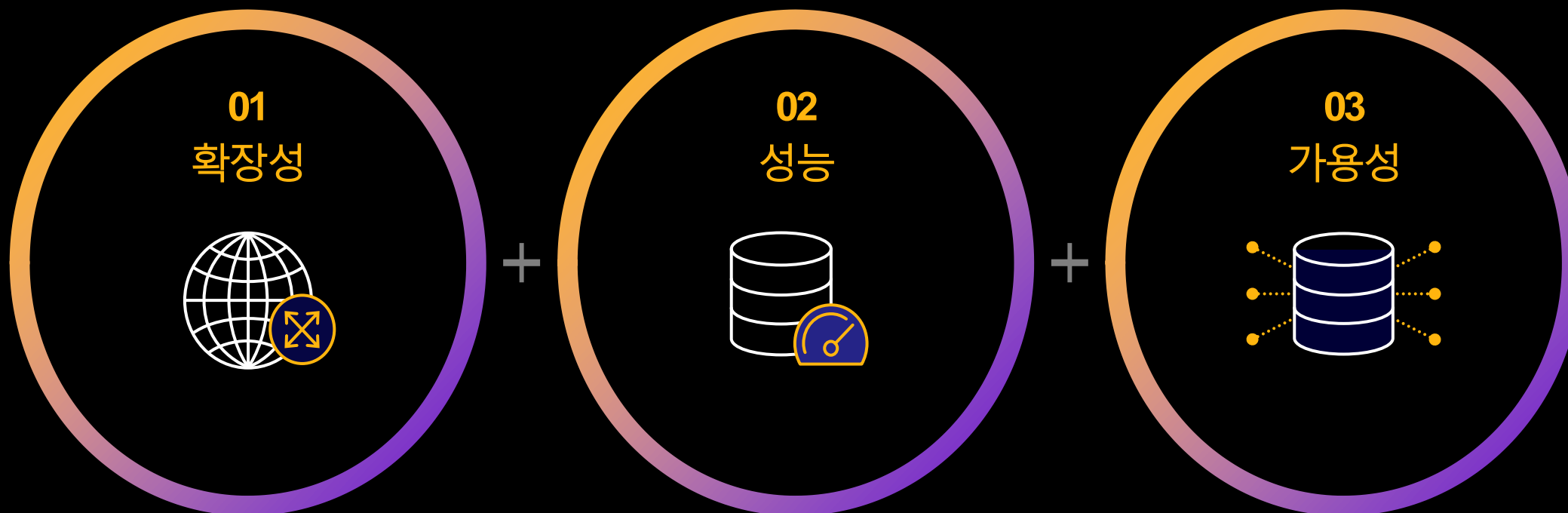
배포 파이프라인





“모노리틱 애플리케이션 대신
마이크로 서비스를
목적에 맞는 도구로 만드세요.”

마이크로 서비스에 왜 목적에 맞는 데이터베이스를 사용해야 할까?



SQL or NoSQL?

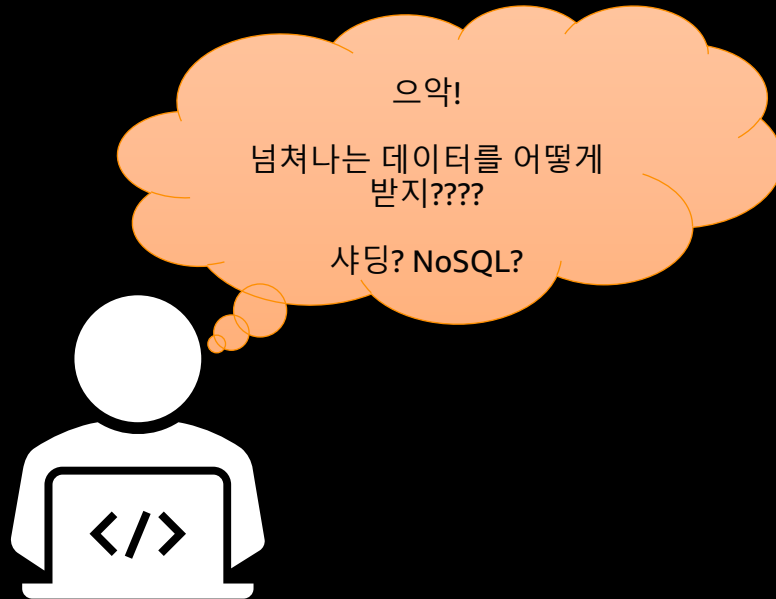
SQL? NoSQL?

SQL ~~vs~~ NoSQL

SQL & NoSQL

Why NoSQL?

ACID, MVCC, Normalized Table 등 RDBMS가 제공하는 이점보다, 다른 형태(최근 아키텍처)의 요구사항이 커질 경우, RDBMS를 이용한 아키텍처가 도리어 부담으로 작용할 수 있습니다.

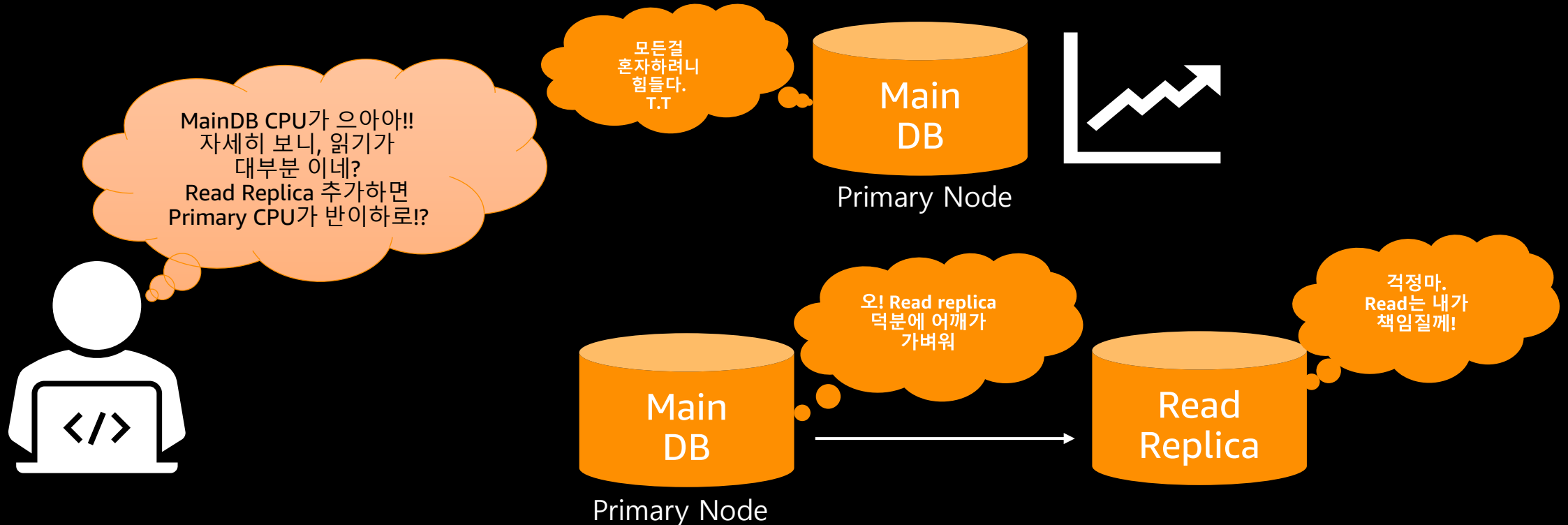


최근 아키텍처 요구사항

- 대용량 데이터 처리
- 시계열 처리
- 그래프 처리
- 개발자들의 가변적 스키마 요구
- Document 처리

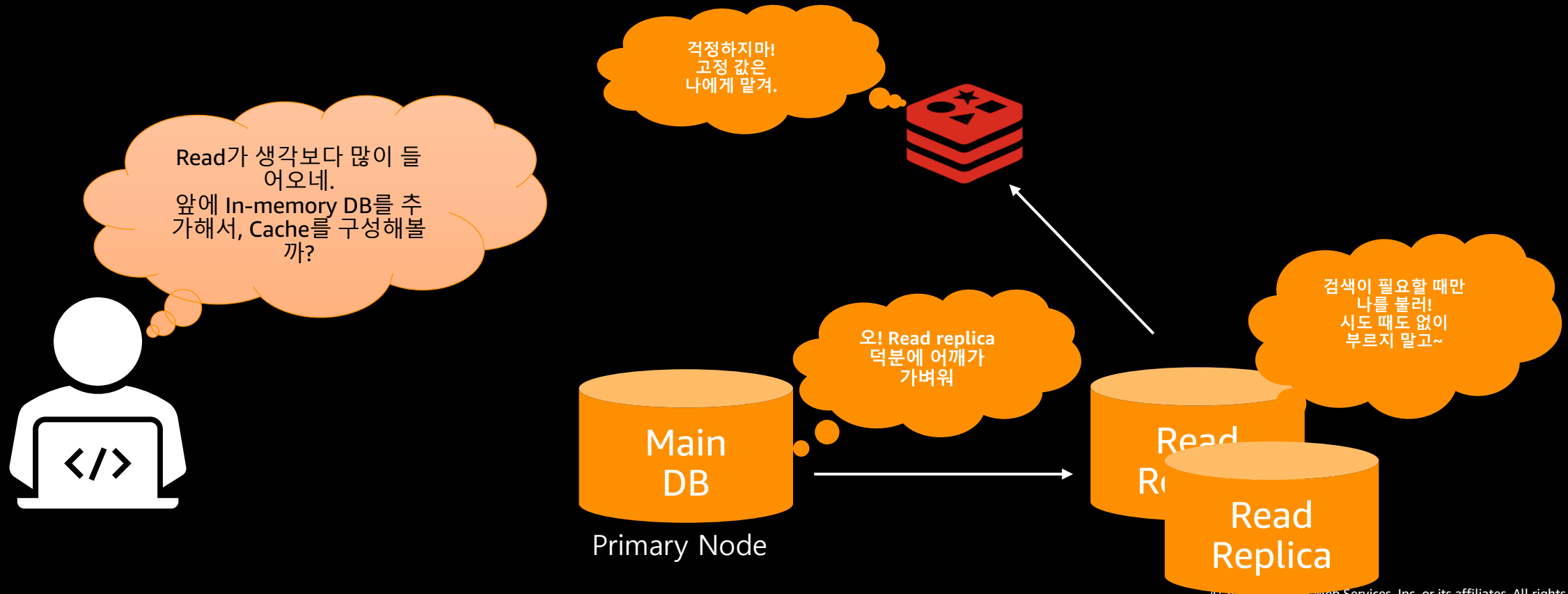
대용량 데이터 처리 - 1

OLTP 형태의 데이터 처리에서는 보통 Write:Read 비율이, 1:9~2:8 수준으로 발생합니다. 이럴 경우 일단 접근할 수 있는 전략은 기존의 RDBMS에 Read Replica를 추가하여, Read 확장성을 이용하는 방법입니다.



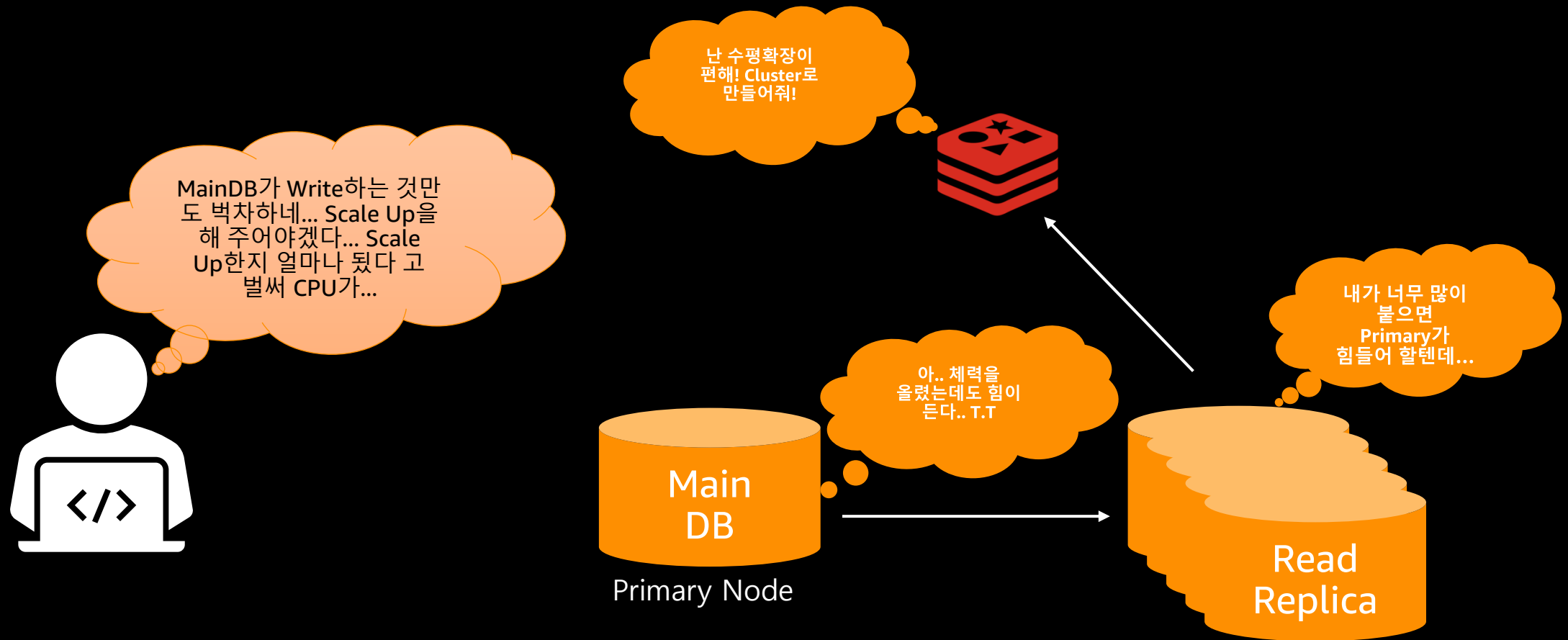
대용량 데이터 처리 - 2

극단적으로, Read가 많이 발생하는 경우, 아니면 동일한 값을 재활용하는 Cache전략을 적용할 수 있는 경우, In-memory Database를 추가로 앞에 위치시킬 수 있습니다.



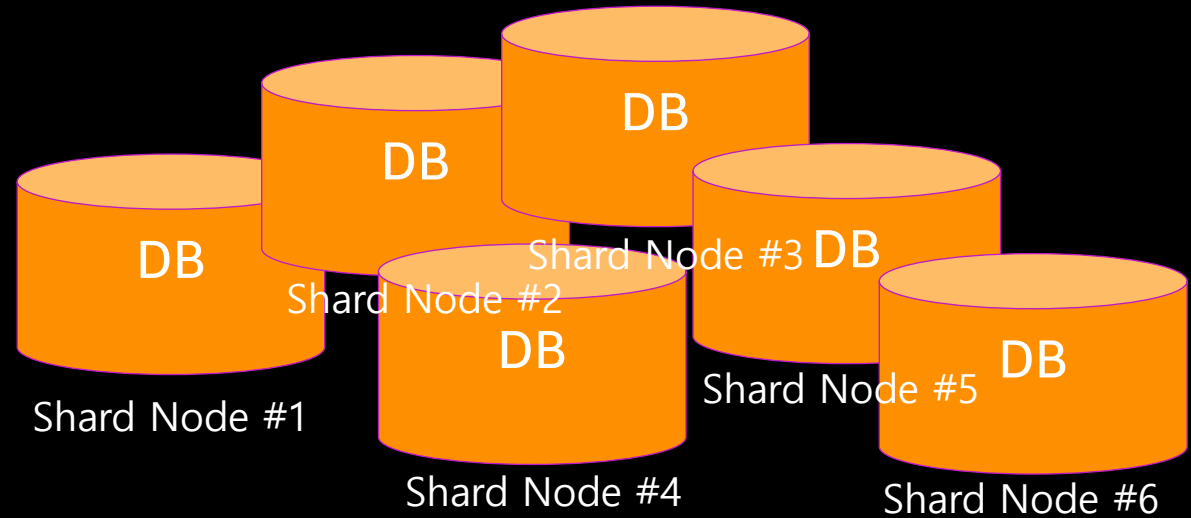
대용량 데이터 처리 - 3

비즈니스가 확장될 경우, Primary node에 대한 Write를 초기에는 Scale Up 으로 대처하지만 얼마 지나지 않아, 그 한계점에 도달하게 됩니다.



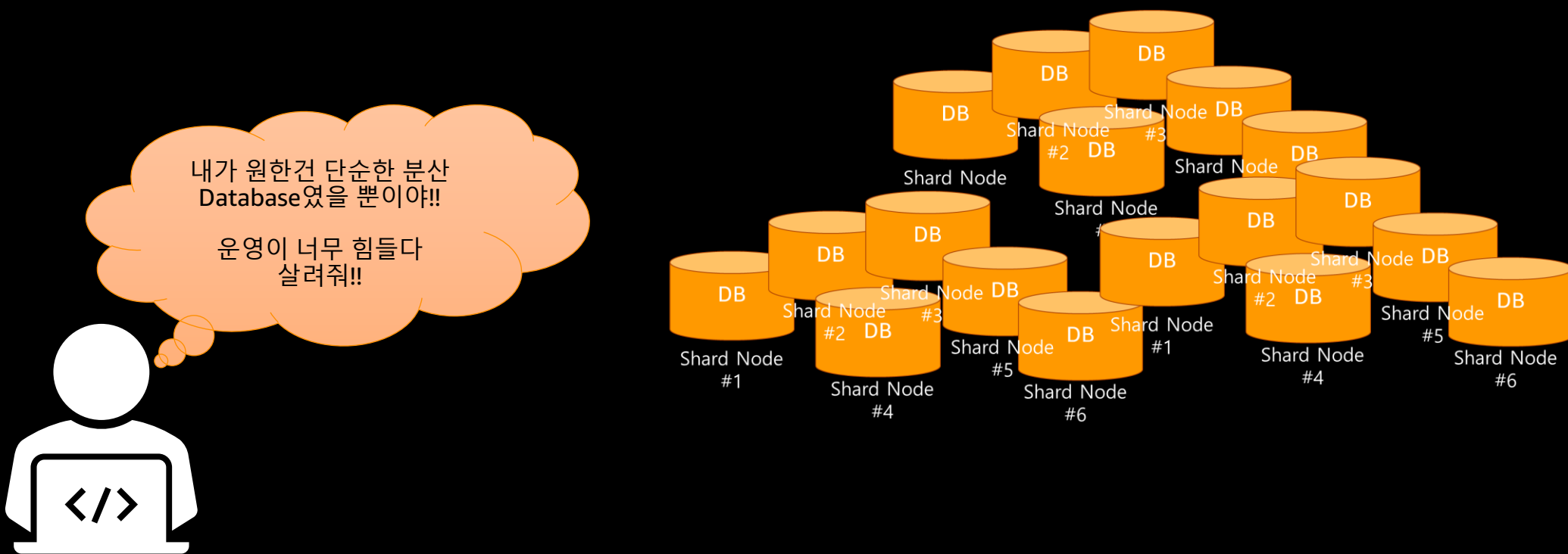
대용량 데이터 처리 - 4

하나의 물리적인 서버가 처리할 수 있는 용량에는 한계가 있음을 인정하고, 여러 개의 물리적인 서버에 분산 저장할 수 있는 Database나 Framework을 도입하게 됩니다.



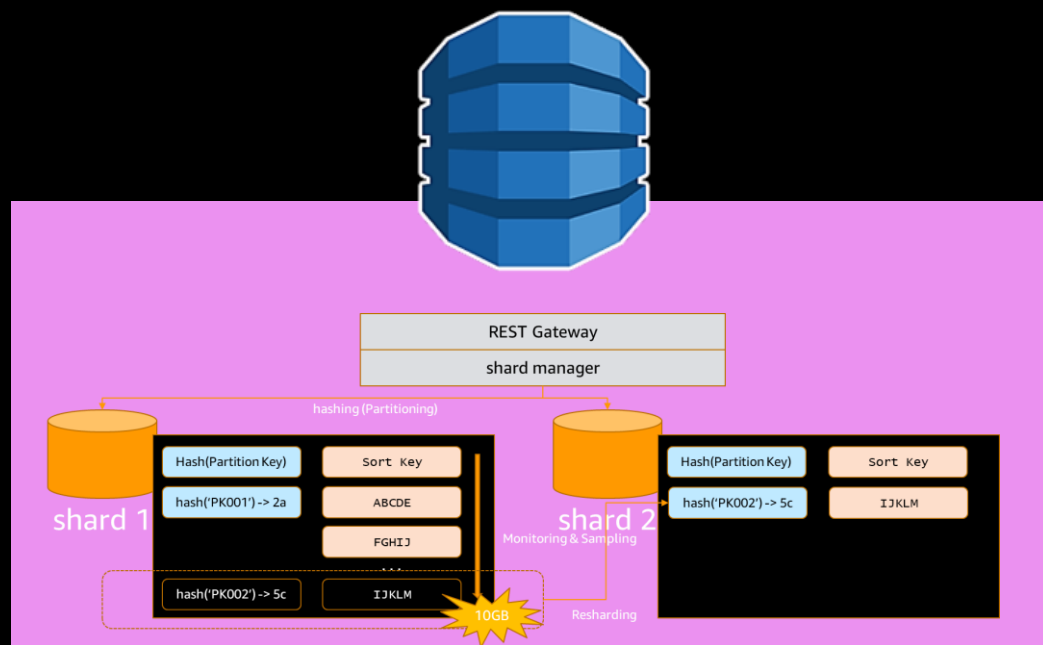
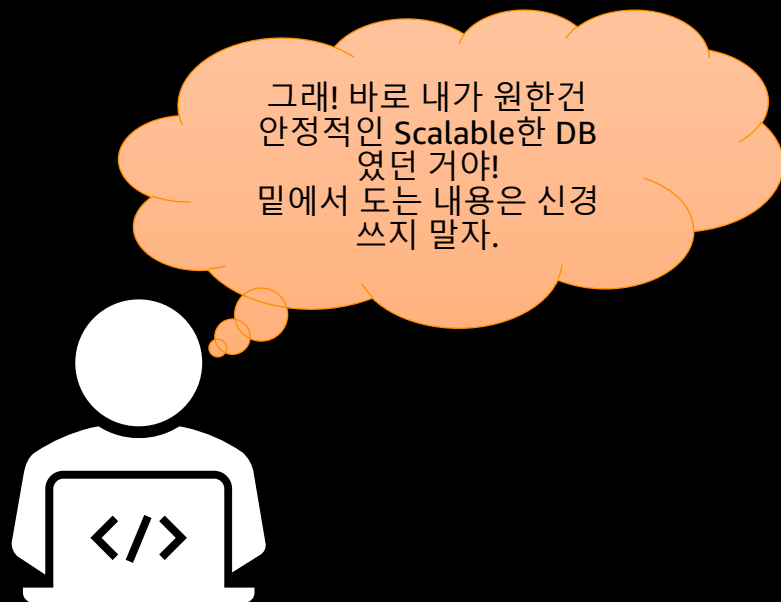
대용량 데이터 처리 - 5

- 분산 Database에서 Node의 증가는 바로 운영 비용 및 운영상의 어려움을 마주하게 됩니다.



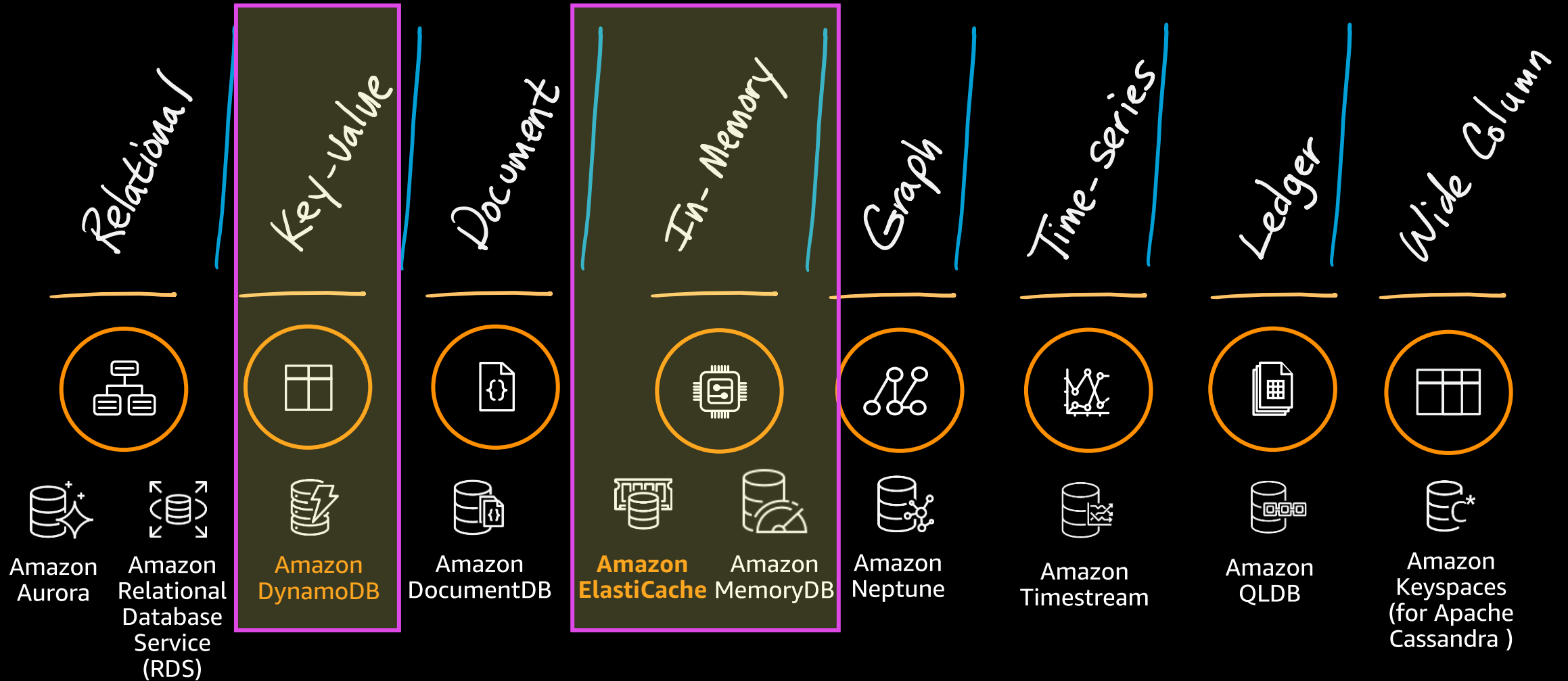
대용량 데이터 처리 - 6

태생적으로 분산 처리를 고려하여 만든 Database이면서, 대규모 Cluster를 구성해도 안정적인 운영 환경을 제공하는 서비스를 채택하게 됩니다.



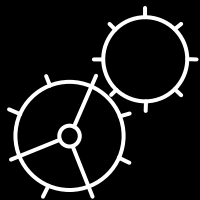
AWS의 대표적인 NoSQL 데이터베이스 서비스

AWS 데이터베이스 서비스



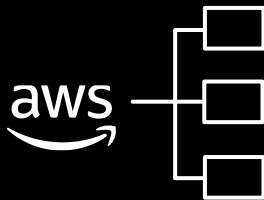
Amazon DynamoDB

어떤 규모에서든 빠르고 유연한 NoSQL 데이터베이스 서비스



규모에 따른 성능

- 초당 수백만 이상의 요청 처리하도록 확장 가능
- 일관된 10밀리초 미만의 성능
- 다중 리전 복제 자동화
- Kinesis Data Streams for DynamoDB로 고급 스트리밍 애플리케이션 개발



서버리스

- 완전 관리형
- Auto scaling
- 온디맨드 모드
- 트리거를 사용한 변경 추적

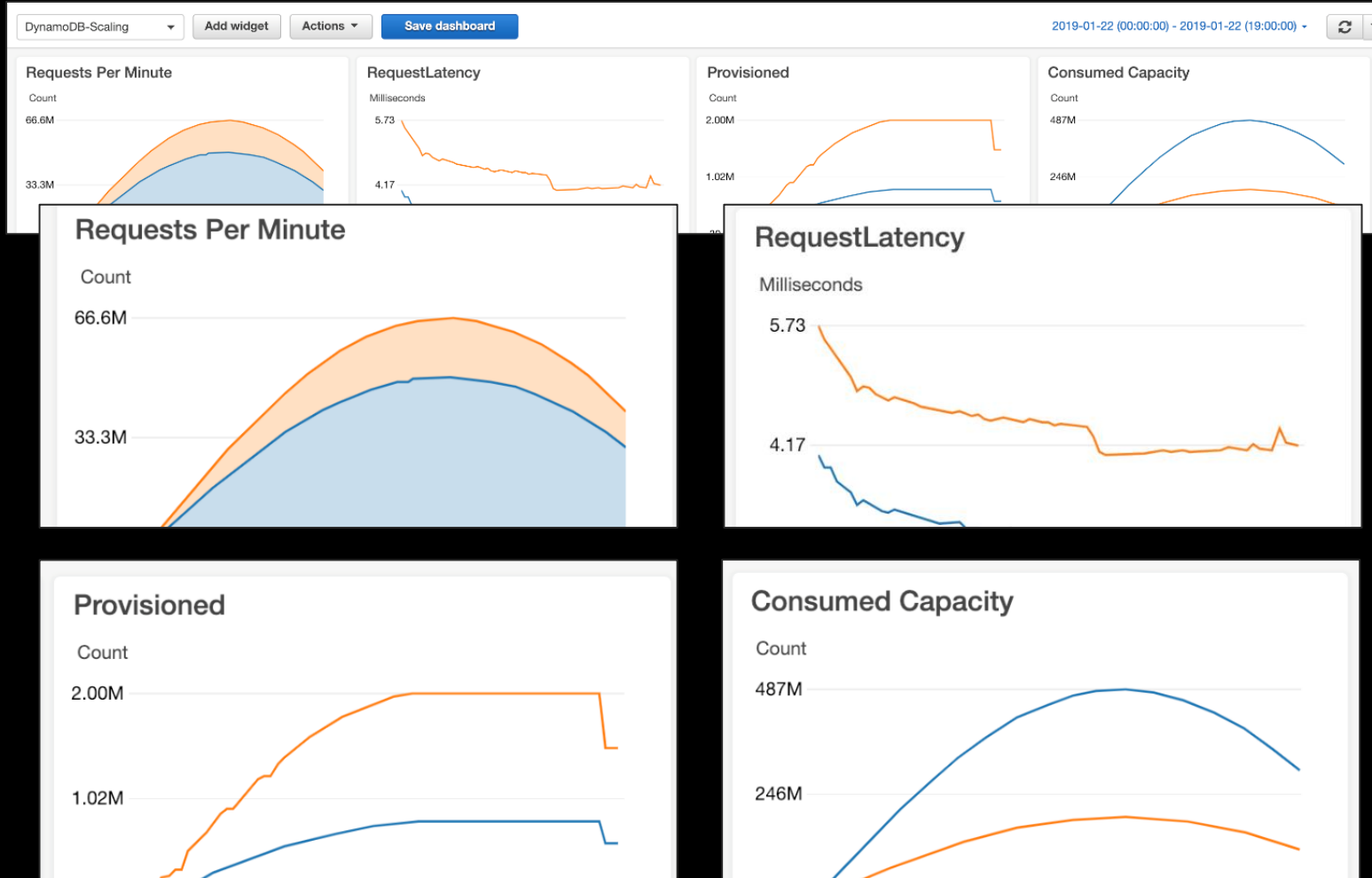


엔터프라이즈에 사용 가능

- ACID 트랜잭션
- 저장 중 암호화
- 특정 시점으로 복구 (PITR)
- 온디맨드 백업 및 복구

페타바이트 규모에서도 10밀리초 미만의 성능

규모에 따른 성능



초당 수백만 이상의 요청

수 조 이상의 아이템

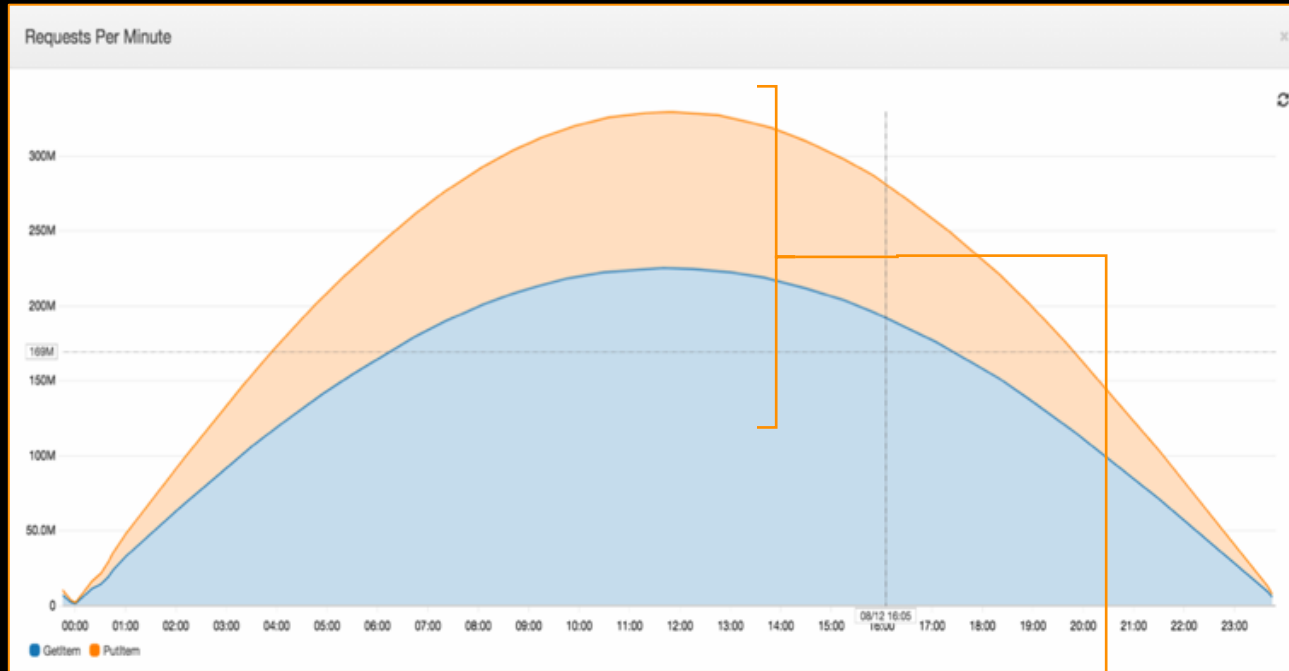
페타바이트 이상의
스토리지

10밀리초 미만의
쓰기/읽기 성능

대규모 트래픽에서도 일관적인 지연 시간

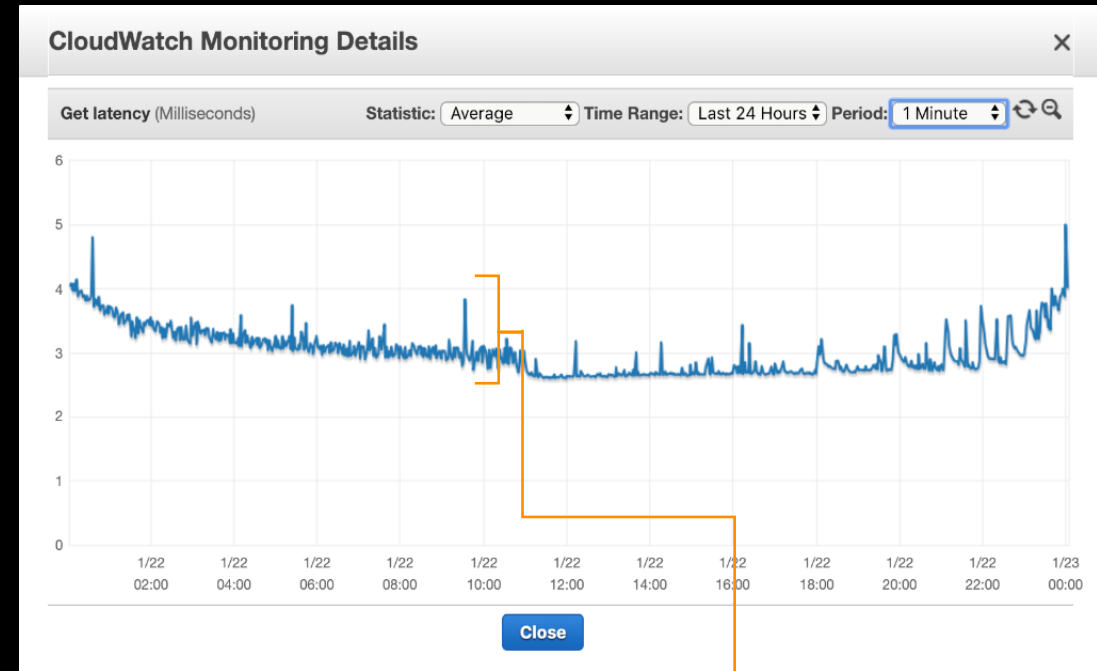
규모에 따른 성능

대규모 트래픽



초당 수백만 요청

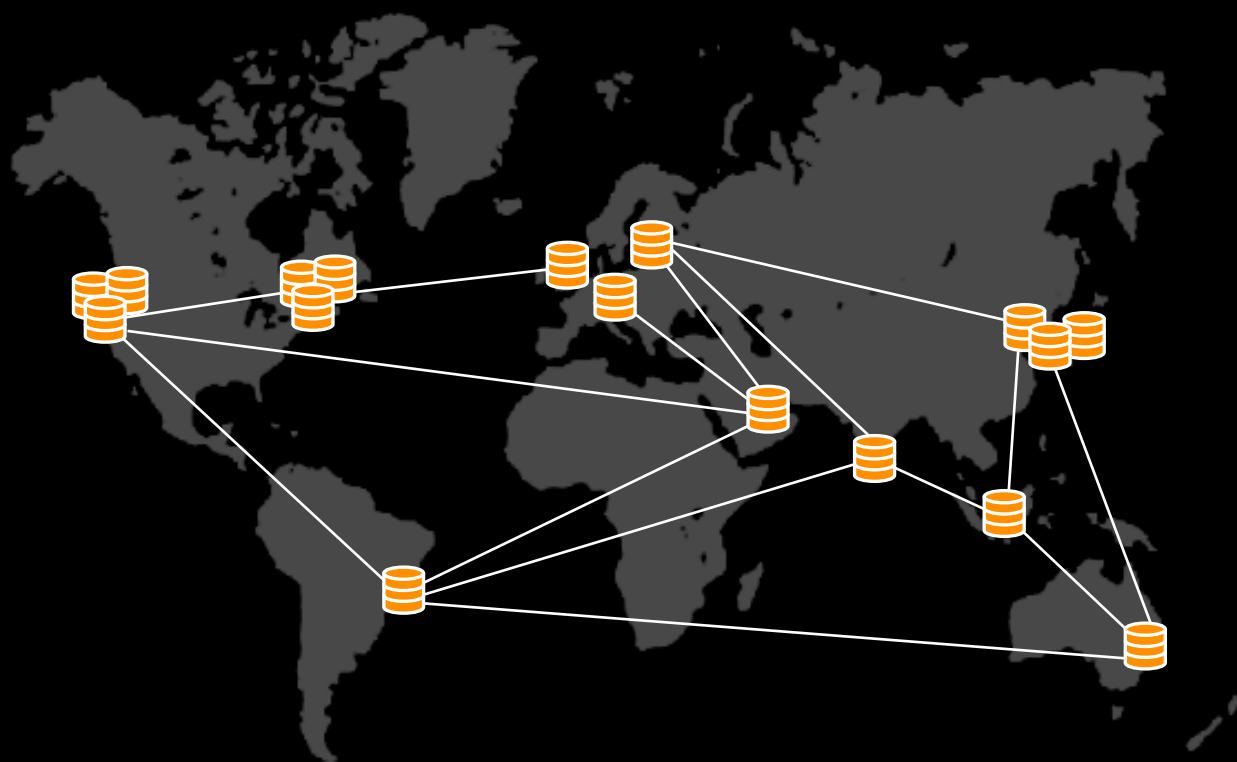
일관적인 저 지연



오직 1ms 차이

전역 테이블을 사용하여 전역 복제 자동화

규모에 따른 성능



글로벌로 분산된 고성능
애플리케이션을 제작

동일 리전 내 가능한 테이블로 저 지연
읽기 및 쓰기

다중 리전 중복을 통한 99.999% 가용성

모든 리전에서 쓰기 가능

간단한 설정과 애플리케이션 변경 없음

고객의 비즈니스를 위한 소중한 시간 돌려드리기

서버리스

보안

- 운영체제 패치
- 데이터베이스 패치
- 액세스 제어
- 심사
- 암호화
- 규정 준수

내구성

- 서버, 랙 및 데이터센터 장애 대비
- 하드웨어 장애 시 신속하게 데이터 재복제
- 백업 및 복원 관리

가용성

- 고가용성 설정
- 모니터링
- 리전 간 복제

성능

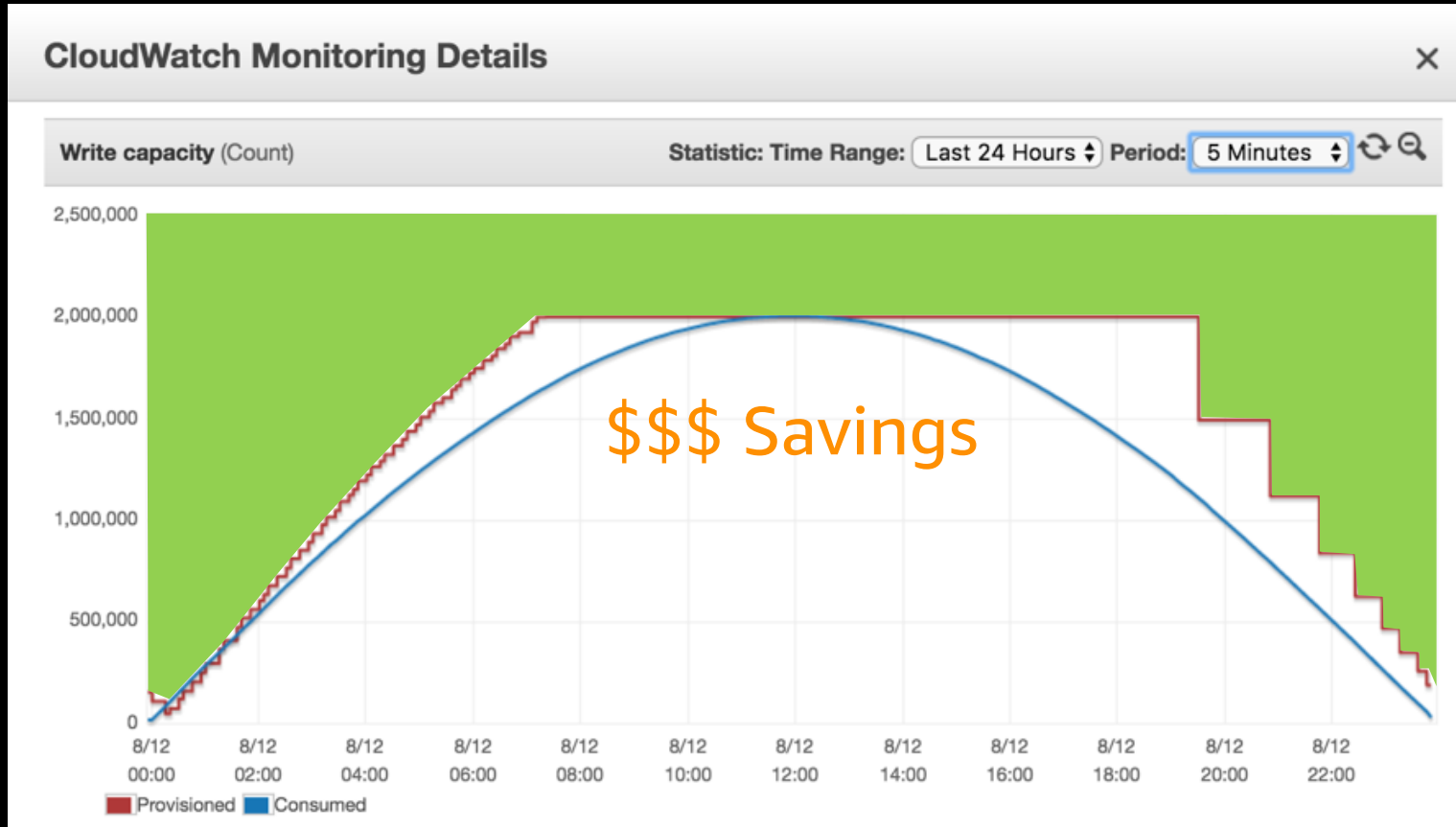
- 성능 튜닝
- 인덱스
- In-memory 캐싱

확장성

- Host 제공
- Host 복구 및 정리

프로비저닝 용량 모드: 자동 크기 조정, 성능 유지

서버리스



자동화된 확장 정책

필요 시 확장

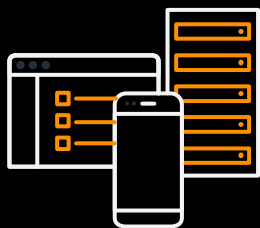
필요 없을 시 축소

스케줄에 따른 자동
크기 조정

프로비저닝 용량 모드

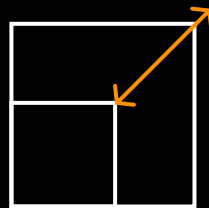
온디맨드 용량 모드: 빠르고 유연한 확장

서버리스



용량 관리 불필요

얼마만큼의 읽기/쓰기 처리량을
사용할지 예측할 필요 없음



예측 불가능한 워크로드에 최적화

요청 시 초당 0에서 수만 요청까지
자동으로 증가



사용한 만큼만 지불

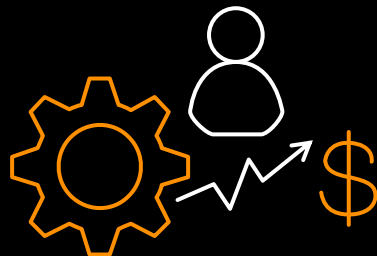
요청당 지불 가격

ACID 트랜잭션에 대한 서버 측 지원

엔터프라이즈에 사용 가능



ACID 보장으로
애플리케이션 코드
간소화



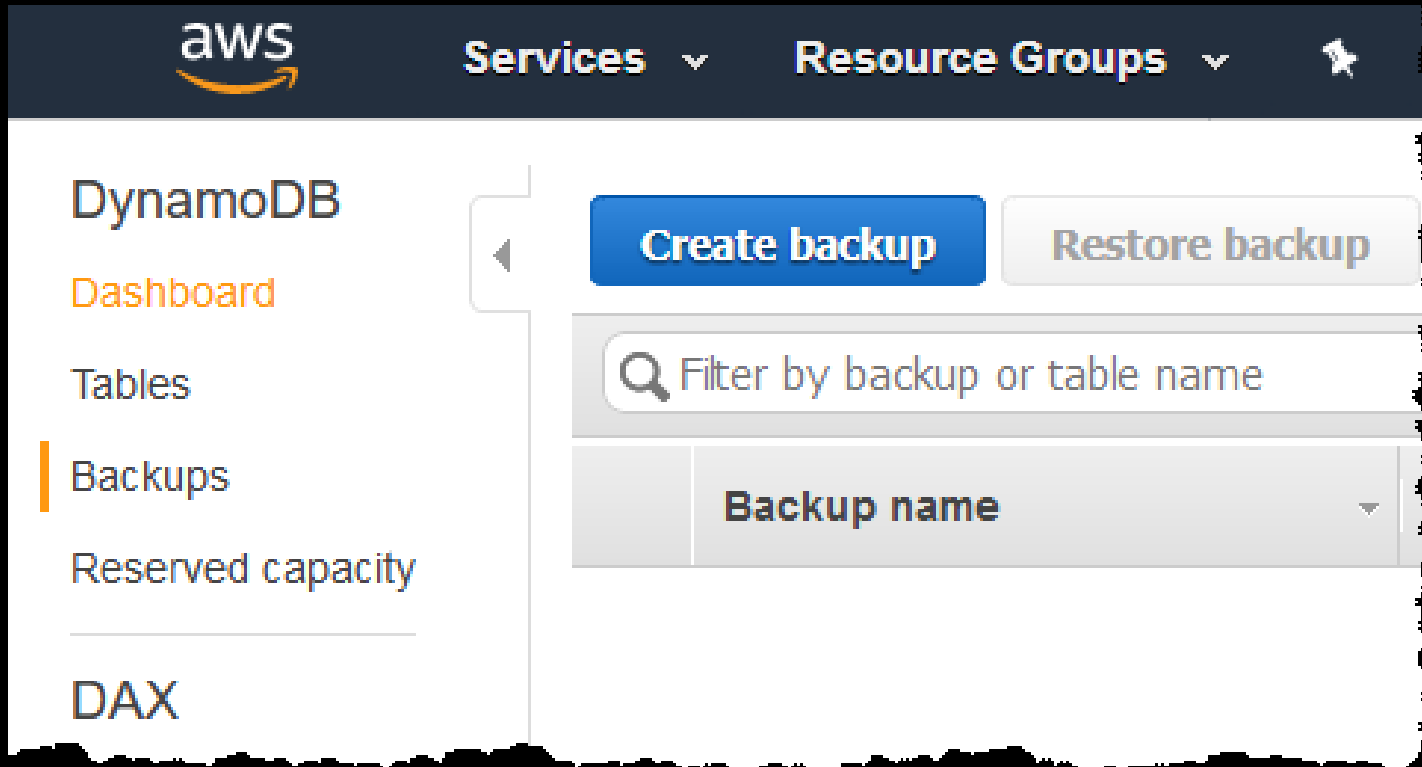
대규모 워크로드에
대한 트랜잭션
실행



레거시 마이그레이션
가속화

백업 및 복원

엔터프라이즈에 사용 가능



장기 데이터 보관 및 규정
준수를 위한 온디맨드
백업

특정 시점 복구(PITR)를
위한 연속 백업

성능에 영향을 미치지
않고 즉시 PB 단위의
데이터 백업

Amazon ElastiCache

REDIS 또는 MEMCACHED와 호환되는 완전관리형 인 메모리 데이터 스토어

Redis &
Memcached 호환



오픈소스 Redis
and Memcached와 호환

탁월한 성능



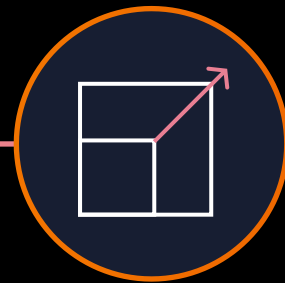
1밀리초 미만의 응답시간이
필요한 애플리케이션을 위한
인 메모리 데이터 스토어

보안 및 안정성



VPC 지원, 전송 및 저장 중
암호화, HIPPA, PCI, FedRAMP,
다중 가용 영역, 자동 장애
조치

간편하게 확장 가능



샤딩과 리플리카로
읽기와 쓰기 확장

Redis란?

#1 인메모리 데이터 스토어, #1 키-값 데이터베이스



redis

Rank			DBMS	Database Model	Score		
Sep 2021	Aug 2021	Sep 2020			Sep 2021	Aug 2021	Sep 2020
1.	1.	1.	Redis +	Key-value, Multi-model i	171.94	+2.05	+20.08
2.	2.	2.	Amazon DynamoDB +	Multi-model i	76.93	+2.03	+10.75
3.	3.	3.	Microsoft Azure Cosmos DB +	Multi-model i	38.52	+1.66	+6.86
4.	4.	4.	Memcached	Key-value	25.66	-0.32	-0.68
5.	5.	↑ 6.	etcd	Key-value	10.32	+0.54	+1.47

* <https://db-engines.com/en/ranking/key-value+store>

개발자들이 Redis를 사랑하는 이유:

1. Blazing fast
2. 다양한 자료 구조
3. Feature rich and easy to use
4. 오픈소스



자체 관리 Redis는 어렵고 시간 소모적입니다.



관리가 어려움

서버 프로비저닝, 소프트웨어 패치, 설정, 구성 및 백업 관리

매력적인 고객 경험을 구축하는데 더 많은 시간을 할애할 수 있음



고가용성 구축의 어려움

빠른 오류 감지 및 수정 구현 필요



확장의 어려움

온라인 확장은 오류가 발생하기 쉬우며 복제 성능을 모니터링해야 함



고 비용

사람, 프로세스, 하드웨어 및 소프트웨어에 투자해야 함

실시간 수요에는 종종 엄청난 스파이크가 발생함

Amazon ElastiCache

AWS의 모든 기능을 기반으로 구축



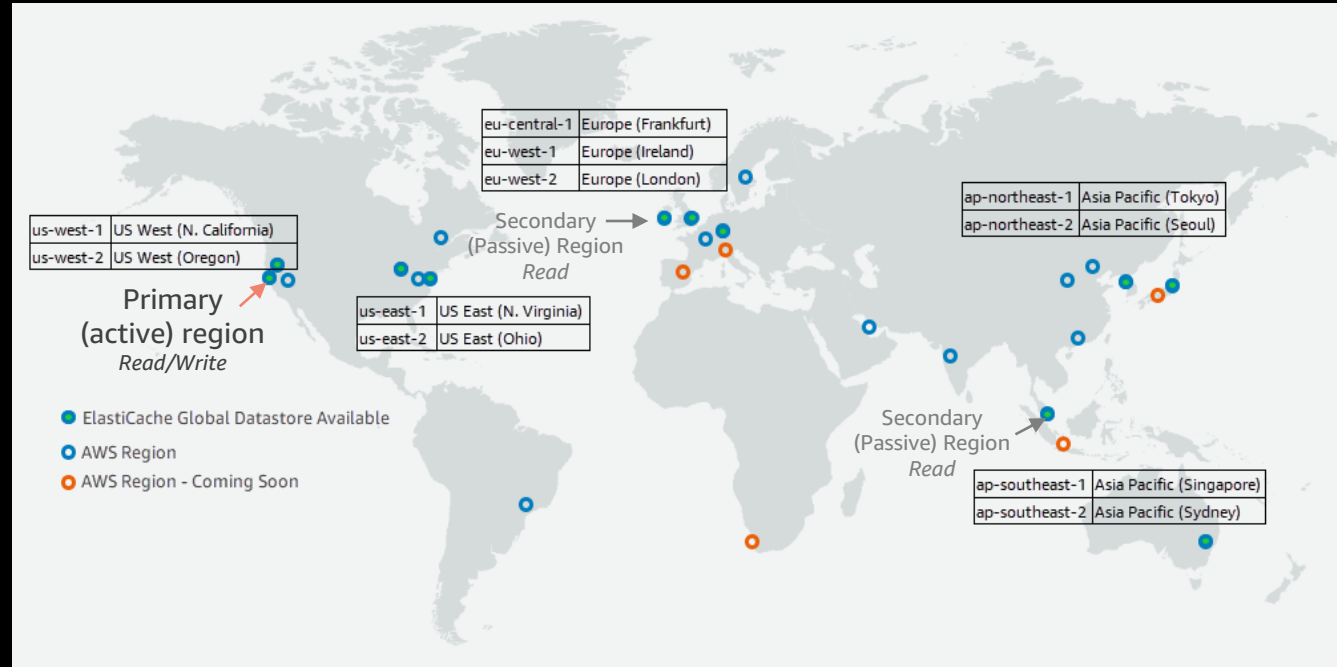
- 모든 AWS 리전에서 사용 가능
- Multi-AZ 자동 장애 조치
- 글로벌 데이터스토어 – 리전 간 복제
- 클러스터 당 최대 500 노드
- 최대 340TB 용량
- T2, M5, R5, M6[G], R6[G] 타입 지원
- 예약된 스냅샷 지원
- 스케일 아웃/인
- 스케일 업/다운



글로벌 데이터스토어(리전 간 복제)

- 기존 클러스터의 원클릭 설정
- 로컬에서 쓰고, 글로벌로 읽기 가능
- 리전 간 재해 복구 활성화
- Redis의 밀리초 미만 지연 시간으로 극한의 성능 활용
- 교차 리전 트래픽에 대한 전송 중 보안 암호화
- AWS Management Console 또는 최신 AWS SDK 또는 CLI와 함께 사용

글로벌 애플리케이션의 예제



사용 예제 - 캐싱



도전과제

디스크 기반 데이터베이스 및 응용 프로그램에서 데이터를 검색하는 것은 느립니다. 특수 하드웨어, 독점 인메모리 데이터베이스 및 데이터 감소와 같은 해결 방법은 비용이 많이 들고 종속성이 높으며 확장되지 않으며 관리하기 어렵습니다.

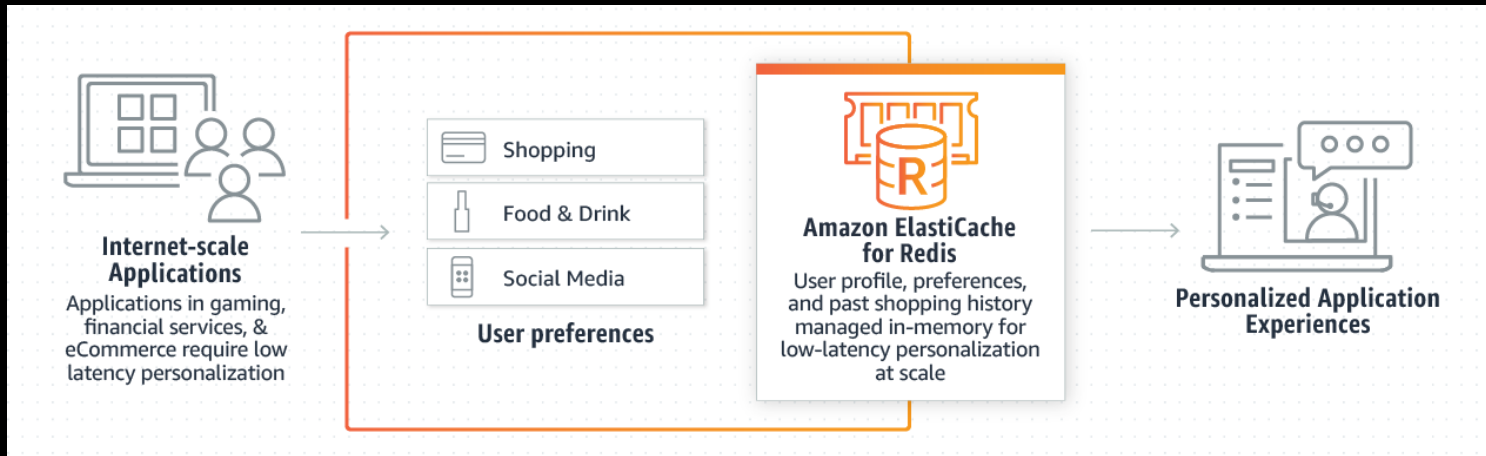
해결책

오픈 소스 인메모리 데이터 저장소 및 캐시인 Redis를 사용하여 마이크로초 대기 시간 및 높은 처리량을 위해 자주 사용하는 데이터를 인메모리에 저장합니다.

사용 사례

데이터베이스 쿼리 캐싱 및 전체 웹 페이지 캐싱

사용 예제 – 세션 스토어



도전과제

인증 토큰, 쿠키 및 세션 상태와 같이 빠르게 움직이는 임시 세션 데이터를 처리하기 위해 짧은 대기 시간과 높은 동시성 데이터 저장소가 필요합니다.

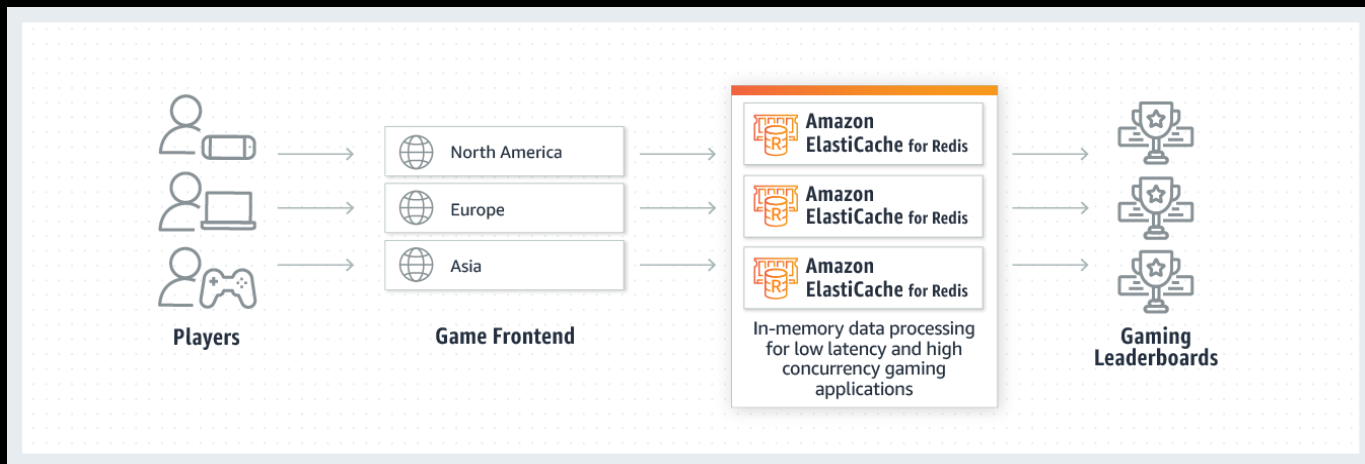
해결책

Redis를 메모리 내 키-값 저장소로 사용하고 적절한 키 만료 전략과 함께 캐시를 사용하여 마이크로초 응답을 위한 임시 세션 데이터를 저장하고 관리합니다.

애플리케이션

웹 / 모바일 앱

사용 예제 – 리더보드



리더보드에 데이터 입력

```
> zadd leaderboard 556 Andy  
> zadd leaderboard 819 Barry  
> zadd leaderboard 105 Carl  
> zadd leaderboard 1312 Derek
```

리더보드 질의

```
> zrevrange leaderboard 0 -1  
1) "Derek"  
2) "Barry"  
3) "Andy"  
4) "Carl"
```

도전과제

리더보드는 변화하는 점수를 처리하고 실시간으로 순위를 업데이트하기 위해 짧은 대기 시간과 높은 동시성 데이터 저장소가 필요합니다.

해결책

Redis를 마이크로초 단위로 점수별로 정렬된 목록을 유지하면서 요소의 고유성을 보장하는 정렬된 세트 및 순위 목록과 같은 데이터 구조와 함께 메모리 내 데이터 저장소로 사용하세요.

애플리케이션

게임 리더보드, 영업 리더보드, 고객 리더보드 등

사용 예제 – 지리정보 앱

도전과제

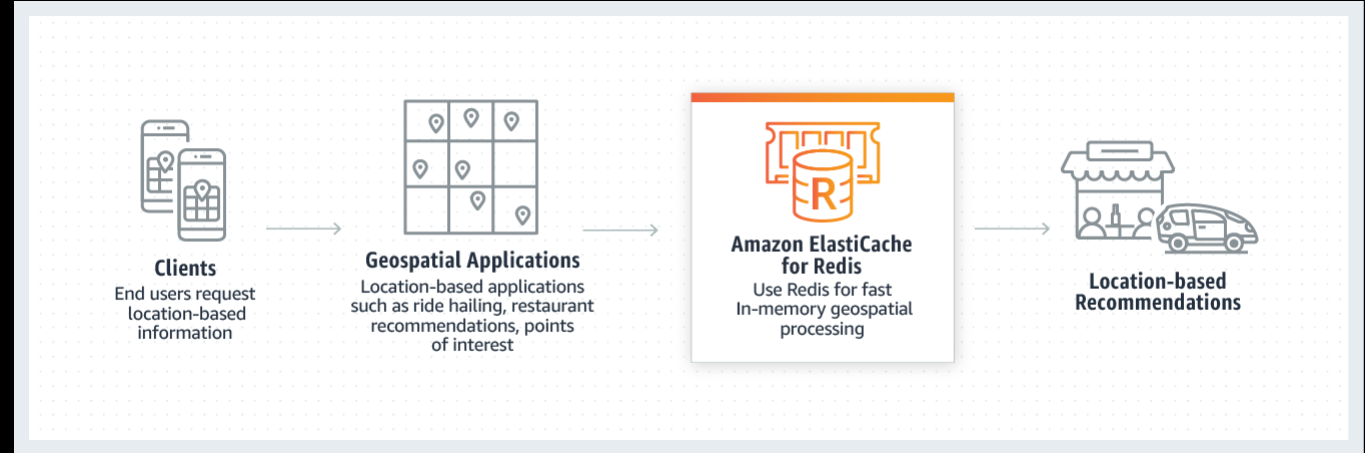
실시간으로 빠르게 이동하는 위치 데이터를 처리하기 위해 짧은 대기 시간과 높은 동시성 데이터 저장소가 필요합니다.

해결책

Redis를 정렬된 세트, 위치 데이터의 인코딩/디코딩, 위치별 인덱싱, 검색 및 정렬을 위한 API와 같은 목적에 맞게 구축된 데이터 구조와 함께 메모리 내 데이터 저장소로 사용하세요.

애플리케이션

차량 공유, 데이팅, 소셜 미디어, 게임, 전자 상거래



> GEORADIUS restaurants 47.6062 122.3321 10 mi

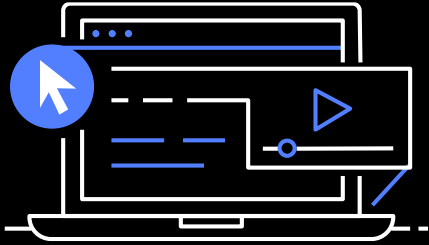
- 1) "Starbucks"
- 2) "McDonalds"
- 3) "The Best Burger Place"

**“There is no compression
algorithm for experience”**

Andy Jassy

CEO, Amazon

AWS 디지털 교육



Flexibility to learn
your way

550개 이상의
무료 디지털 교육 및
심층적 강의실 교육을 통해
클라우드 기술 역량을
업그레이드 하세요!

추천 과정 (한글 자막)

- [AWS Cloud Practitioner Essentials](#)
AWS 클라우드 기초에 대해 학습하고, 기초 자격증인 AWS Certified Cloud Practitioner 시험을 준비할 수 있습니다.
- [Amazon DynamoDB for Serverless Architectures](#)
Amazon DynamoDB의 전반적인 소개와 Amazon DynamoDB가 서버리스 아키텍처 구축에 어떻게 활용되는지 알아봅니다.
- [AWS Security Fundamentals](#)
AWS 액세스 제어 및 관리, 거버넌스, 로깅, 그리고 암호화 방법을 포함한 기본적인 클라우드 컴퓨팅 및 AWS 보안 개념에 대해 알아봅니다.
- [AWS Database Offerings](#)
다양한 데이터베이스 기술 및 아키텍처에 대한 기본 개요와 AWS 데이터베이스 서비스를 소개합니다.

AWS Builders Online Series에 참석해주셔서 대단히 감사합니다.

저희가 준비한 내용, 어떻게 보셨나요?
더 나은 세미나를 위하여 **설문을 꼭 작성해 주시기 바랍니다.**



aws-korea-marketing@amazon.com



twitter.com/AWSKorea



facebook.com/amazonwebservices.ko



youtube.com/user/AWSKorea



linkedin.com/company/amazon-web-services



twitch.tv/aws

Thank you!

이 혁

hyuklee@amazon.com

