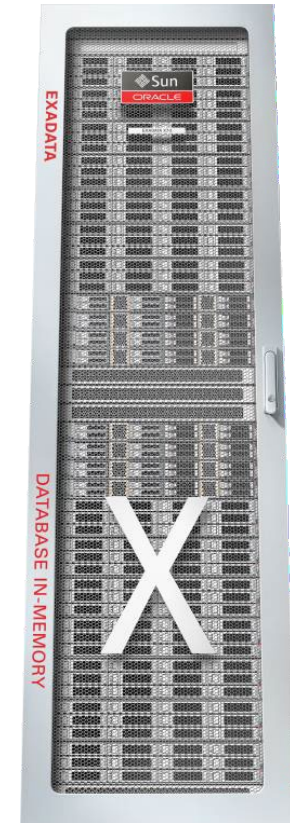


# Exadata 기술 소개

## - Exadata Hardware 기초



경북산업직업전문학교 [1-SEP-2022]  
- 무단배포를 금지합니다 -


---

## Exadata 개요



최고의 성능 / 무한의 확장성 / 준비된 엔터프라이즈 사용 환경

# Runs Oracle Data Warehouses 10x Faster



*Announcing*  
**The World's Fastest  
Database Machine**

- Hardware by HP
- Software by Oracle

**ORACLE®**


oracle.com/exadata  
or call 1.800.ORACLE.1

The performance improvement based on customer reports comparing average performance of Oracle data warehouses on existing customer systems versus Oracle Database Machines. Actual results may vary.

Copyright © 2009, Oracle. All rights reserved. Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

DB\_Step\_Arch090404\_2100

# *Announcing* The World's Fastest Database Machine



**Oracle Database  
Server Grid**

**10x  
Faster  
Query  
Results**

**Exadata Storage Grid  
Fast Query Processing**

**ORACLE®**

**Hardware by HP      Software by Oracle**

The performance improvement based on customer reports comparing average performance of Oracle data warehouses on existing customer systems versus Oracle Database Machines. Actual results may vary.

Copyright © 2009, Oracle. All rights reserved. Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.












DB\_Step\_Arch090404\_2100



# Exadata – The Game Changer

혁신에서 출발, 진화와 발전을 거듭한 혁신적인 기능 탑재



Exadata V1	Exadata V2	Exadata X2	X2(계속)	Exadata X3	Exadata X4	Exadata X5	Exadata X6	Exadata X7	Exadata X8	Exadata X8M
Oracle Database Machine	Sun Oracle Database Machine	Extreme Performance for OLTP, Analytics, Consolidation		In-Memory Database Machine	Best Platform for OLTP, Analytics, Consolidation		Exadata Cloud Service Exadata Cloud at Customer		Autonomous Infrastructure and Database	World's Fastest Database Machine
										
2008	2009	2010	2011	2012	2013	2014	2016	2017	2019	2019
<ul style="list-style-type: none"> <li>HP서버 기반</li> <li>Smart Scan</li> </ul>	<ul style="list-style-type: none"> <li>Sun 서버 기반</li> <li>Flash Cache 탑재</li> <li>Hybrid Columnar 압축</li> </ul>	<ul style="list-style-type: none"> <li>Oracle DB11gR2</li> <li>5.3TB Flash Cache</li> <li>Automatic Service Request</li> </ul>	<ul style="list-style-type: none"> <li>Oracle Solaris 지원</li> <li>Exadata Storage Expansion Rack</li> </ul>	<ul style="list-style-type: none"> <li>Sand-Bridge (8core)</li> <li>메모리 최대2TB</li> <li>Flash Cache Write-back 기능</li> <li>1/8 Rack 출시</li> </ul>	<ul style="list-style-type: none"> <li>Ivy-Bridge(12core)</li> <li>디스크용량 증가 (HC 4T, HP 1.2T)</li> <li>InfiniBand All Ports Active-Active 지원</li> </ul>	<ul style="list-style-type: none"> <li>Haswell(18core)</li> <li>VM Support</li> <li>Flash NVMe 지원</li> <li>Columnar Flash Cache</li> <li>Elastic Configuration</li> </ul>	<ul style="list-style-type: none"> <li>Broadwell(22core)</li> <li>3D V-NAND NVMe Flash cards</li> <li>DDR4 DRAM (13%속도향상)</li> </ul>	<ul style="list-style-type: none"> <li>Skylake(24core)</li> <li>스토리지 DRAM Cache In-Memory OLTP 가속</li> <li>스토리지단의 In-Memory Analytics</li> <li>Hot Swap Flash Card</li> <li>25Gb Network</li> </ul>	<ul style="list-style-type: none"> <li>Cascade Lake(24core)</li> <li>스토리지 서버 - 최신 16 코어 인텔 캐스케이드 레이크 CPU</li> <li>14TB 헬륨 디스크 드라이브</li> <li>새로운 스토리지 서버 확장 (XT)</li> <li>Autonomous Database Service run on Exadata</li> </ul>	<ul style="list-style-type: none"> <li>In-Memory급 스토리지 성능</li> <li>- PMEM Cache/log</li> <li>- 1600만 OLTP Read IOPS</li> <li>- IO Latency 0.019ms</li> <li>100Gb RoCE Switch</li> <li>KVM 가상화</li> </ul>

# Exadata의 Loadmap

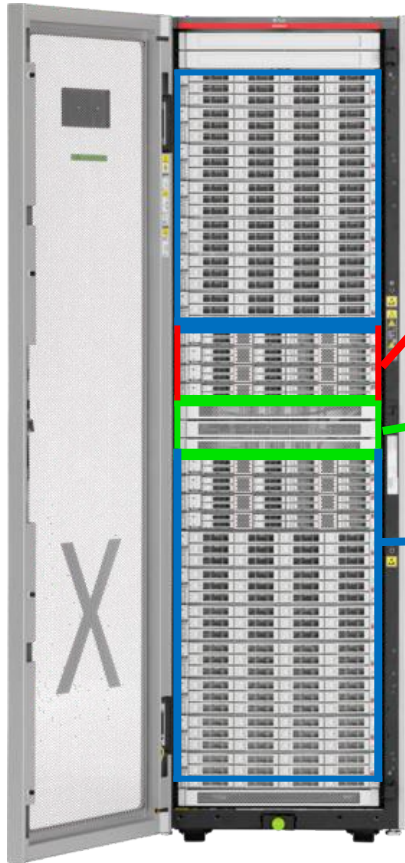
Exadata는 Intel의 최신 Xeon Chip을 사용하기 때문에 Intel의 CPU 출시와 Version을 같이함

V1/ V2	X2	X3	X4	X5	X6	X7	X8/X8M
네할렘 4 core	웨스트미어 6 core	샌드브릿지 8 core	아이비브릿지 12 core	하스웰 18 core	브로드웰 22 core	스카이레이크 24 core	케스케이드레이크 24 core



최근 PC는 커피 레이크, 코멧레이크 ?

# Exadata Database Machine X8-2M 하드웨어 구성



- 확장 가능한 데이터베이스 서버



2소켓 Xeon 8260 프로세스 (2.4GHz)  
서버 당 48 코어  
384 GB ~ 1.5 TB DRAM

- 고속 내부 네트워크

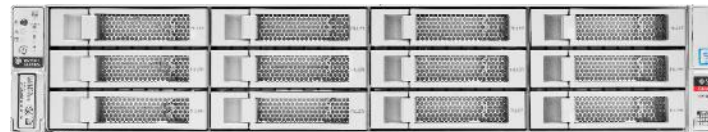
100 Gb/s RoCE  
(RDMA over Converged Ethernet)  
외부접속 25, 10, 1 GbE

- 확장 가능한 지능형 스토리지



High-Capacity Storage Server

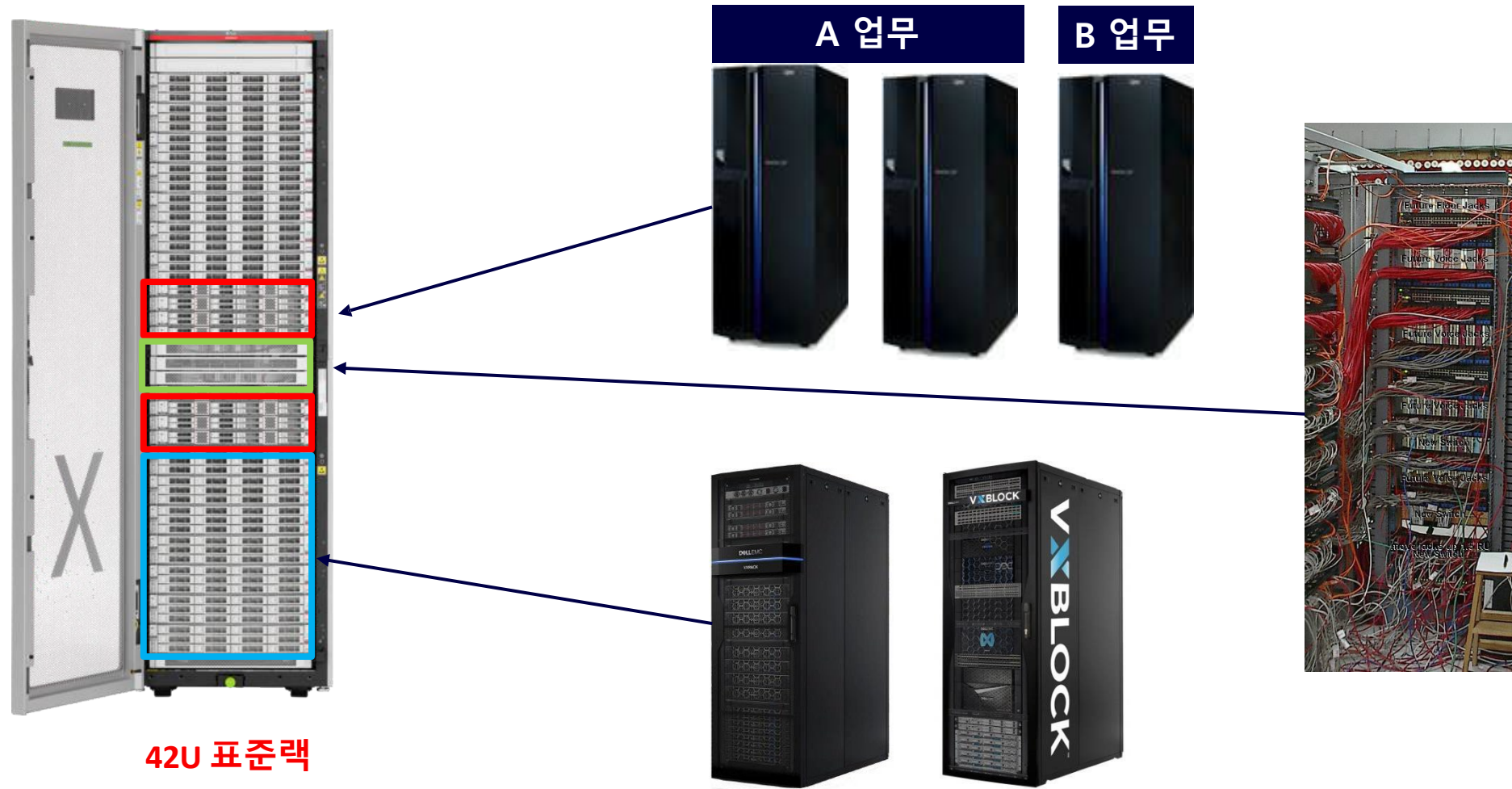
168 TB disk (14 TB 헬륨 디스크)  
25.6 TB NVMe Flash  
1.5 TB Persistent Memory  
32 cores for SQL offload



Extreme Flash Storage Server

51.2 TB NVMe Flash  
1.5 TB Persistent Memory  
32 cores for SQL offload

# Exadata와 타사 Solution H/W 비교



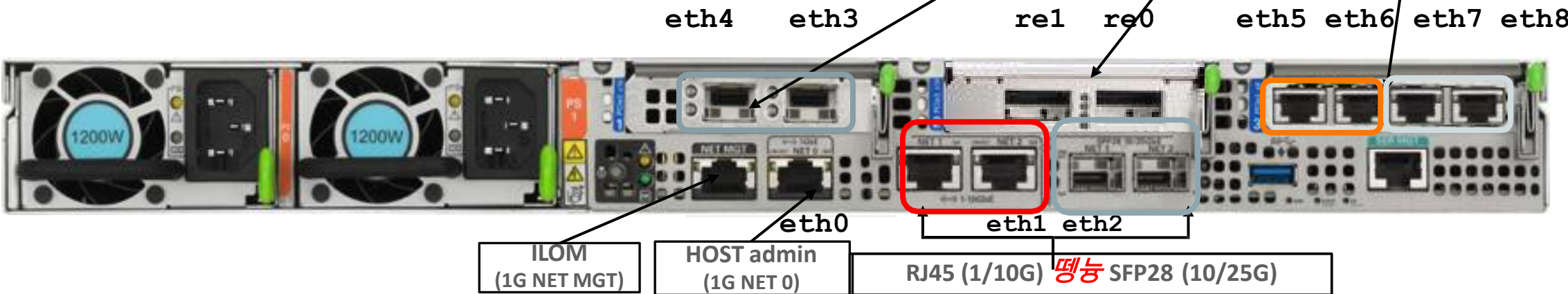
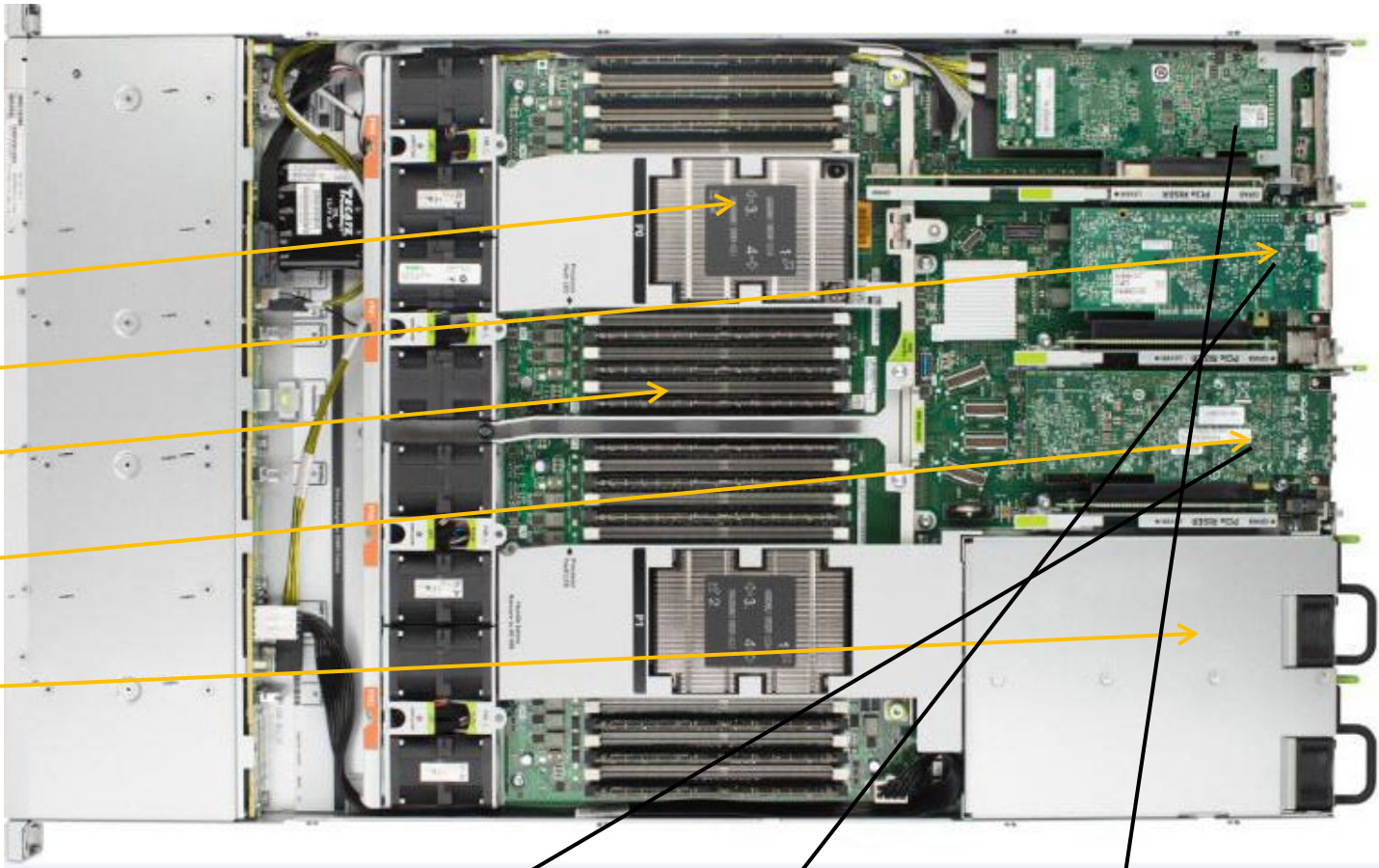


# Exadata Top View

## Database Server

- Intel Xeon CPU
- 2x QSFP28 HCA RoCE
- 384GB DRAM (to 1.5TB)
- 10/25GbE SFP+ PCIe Ethernet Card
- Dual Power Supplies

HCA : Host Channel Adapter  
HBA : Host Bus Adapter

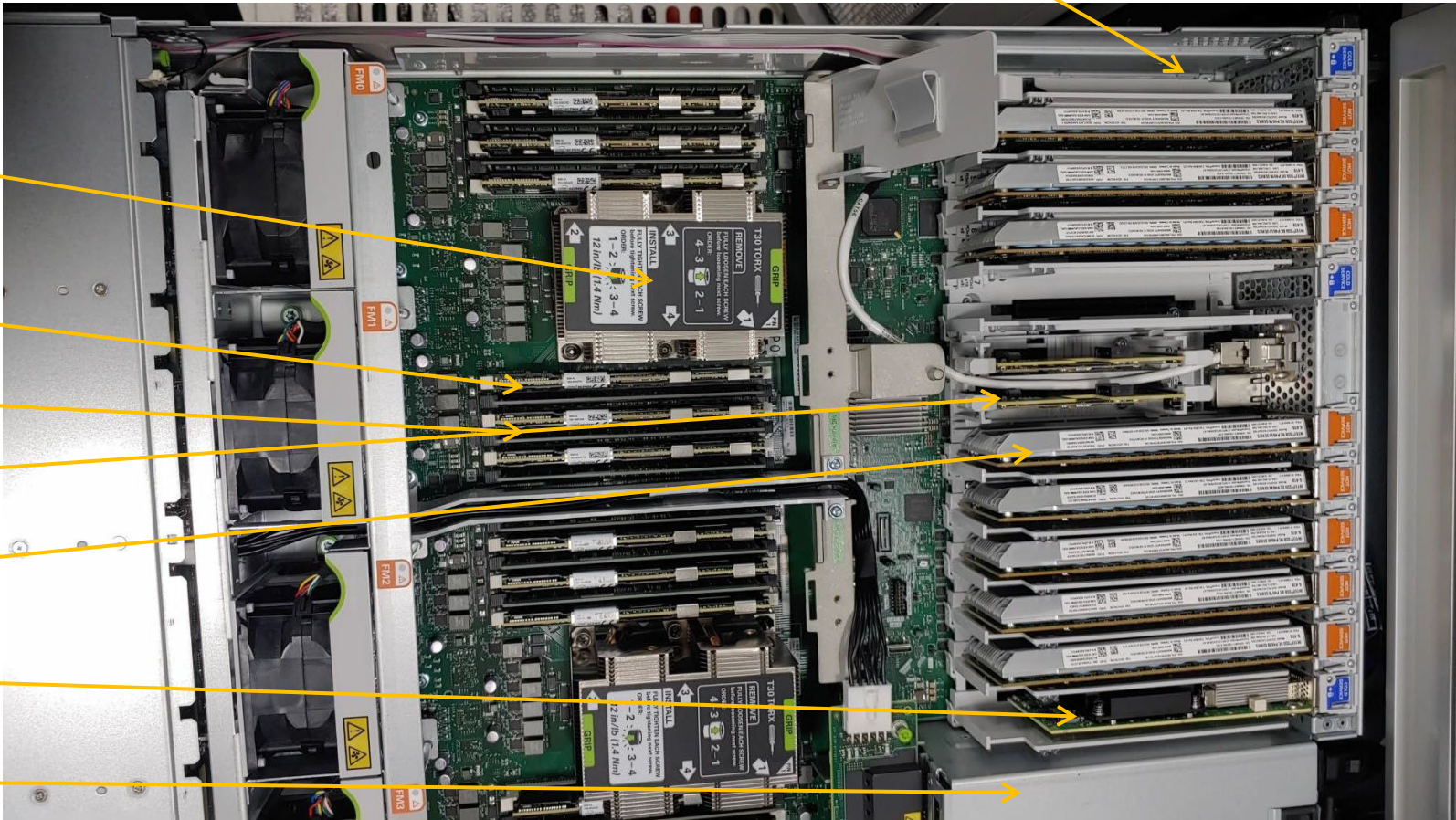




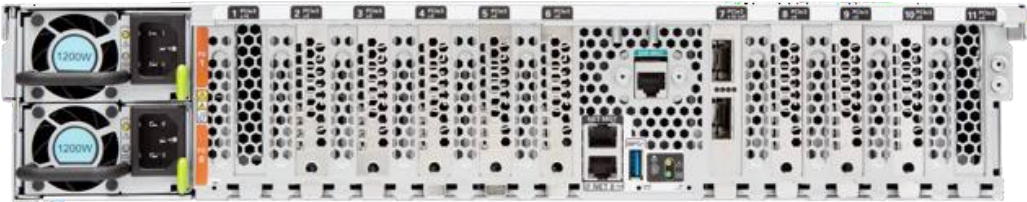
# Exadata Top View

## Storage Server

- Intel Xeon CPU
- DRAM
- PMEM
- 2 x 150GB M.2 Drives
- 6400GB F640 Flash Card
- 2x QSFP28 HCA RoCE
- Dual Power Supplies



Disk Controller  
SAS HBA





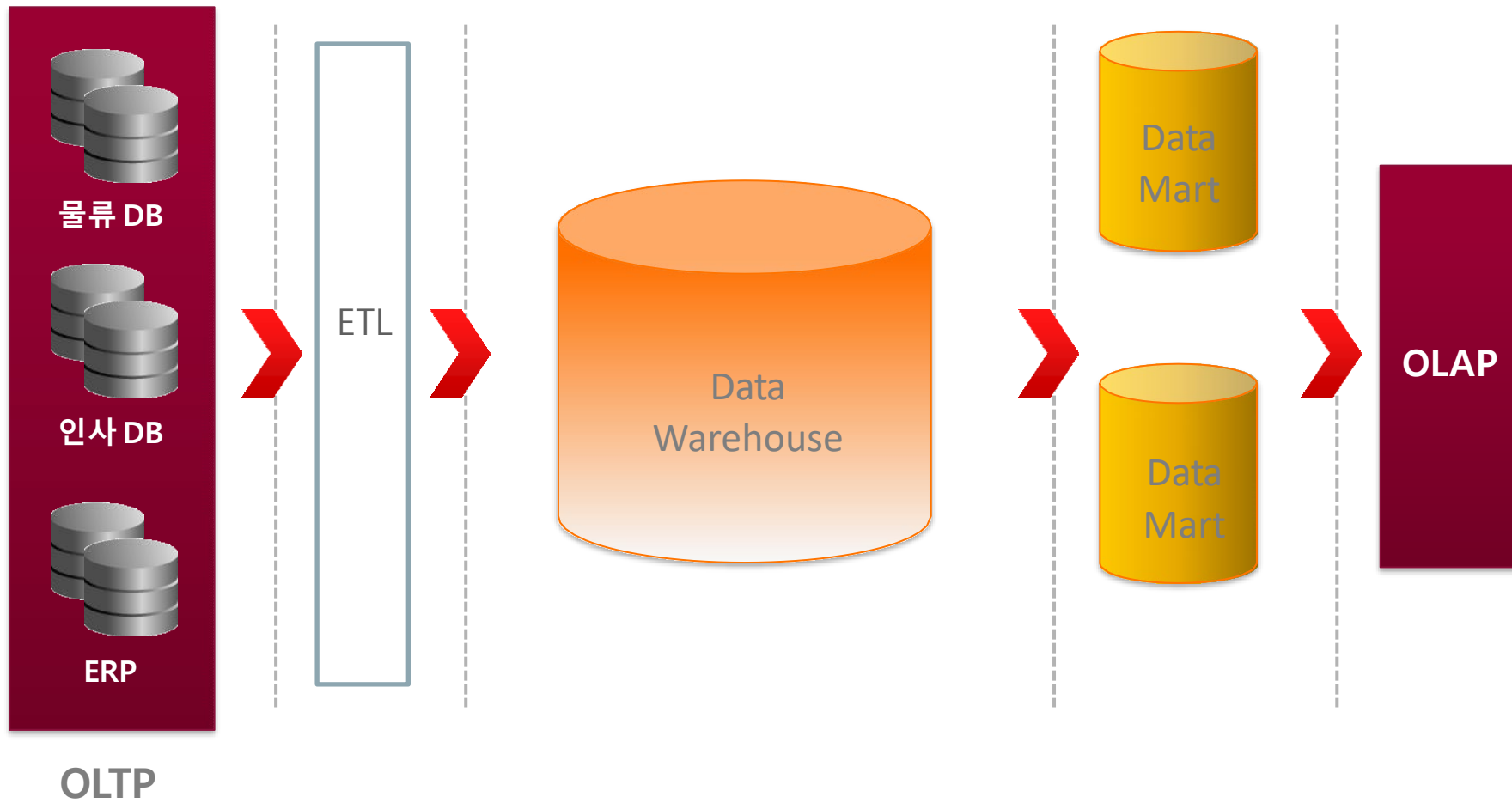
---

## Exadata 출현 배경



# 데이터베이스 용도별 DBMS

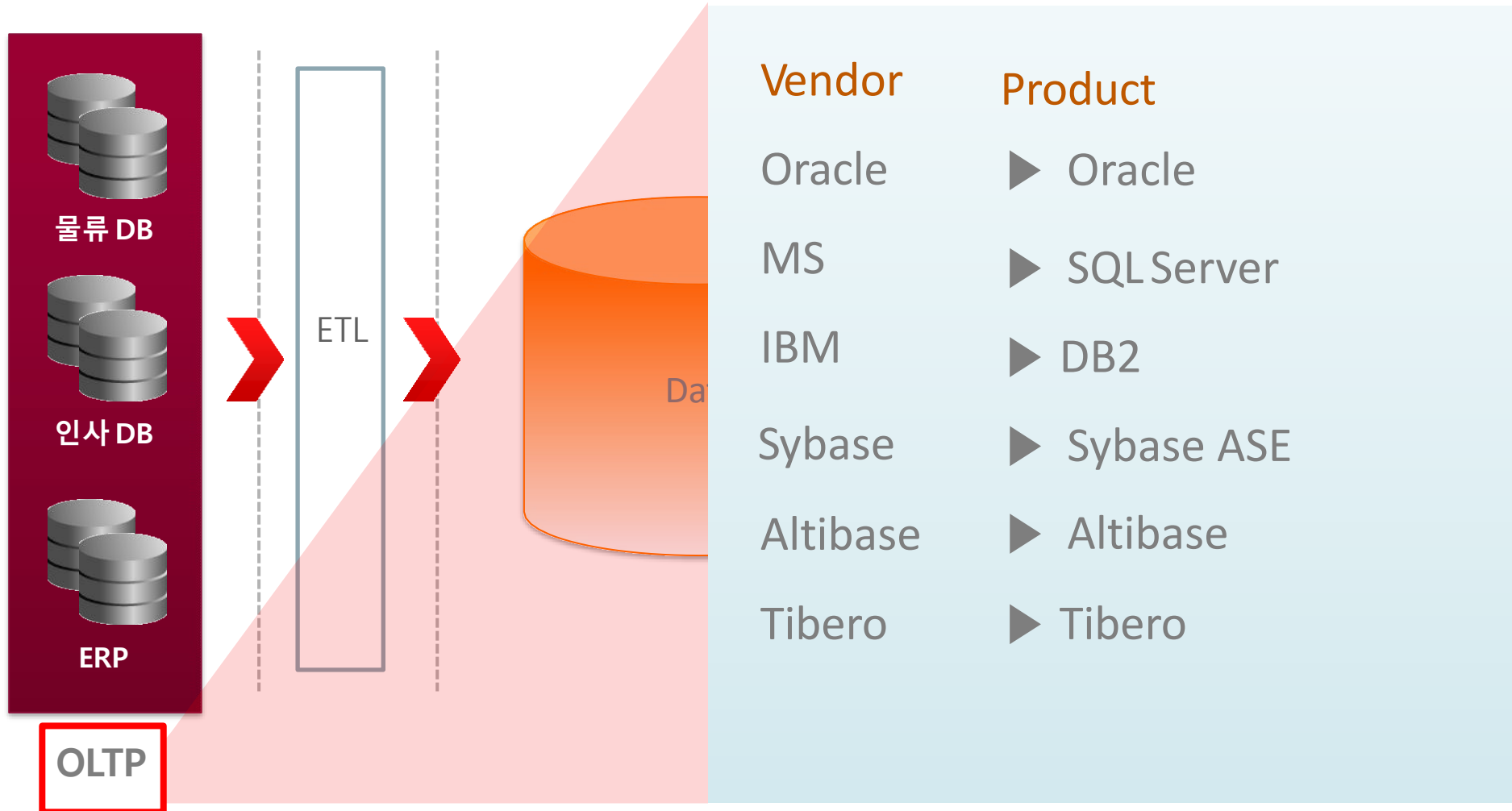
- DW 기본 Architecture





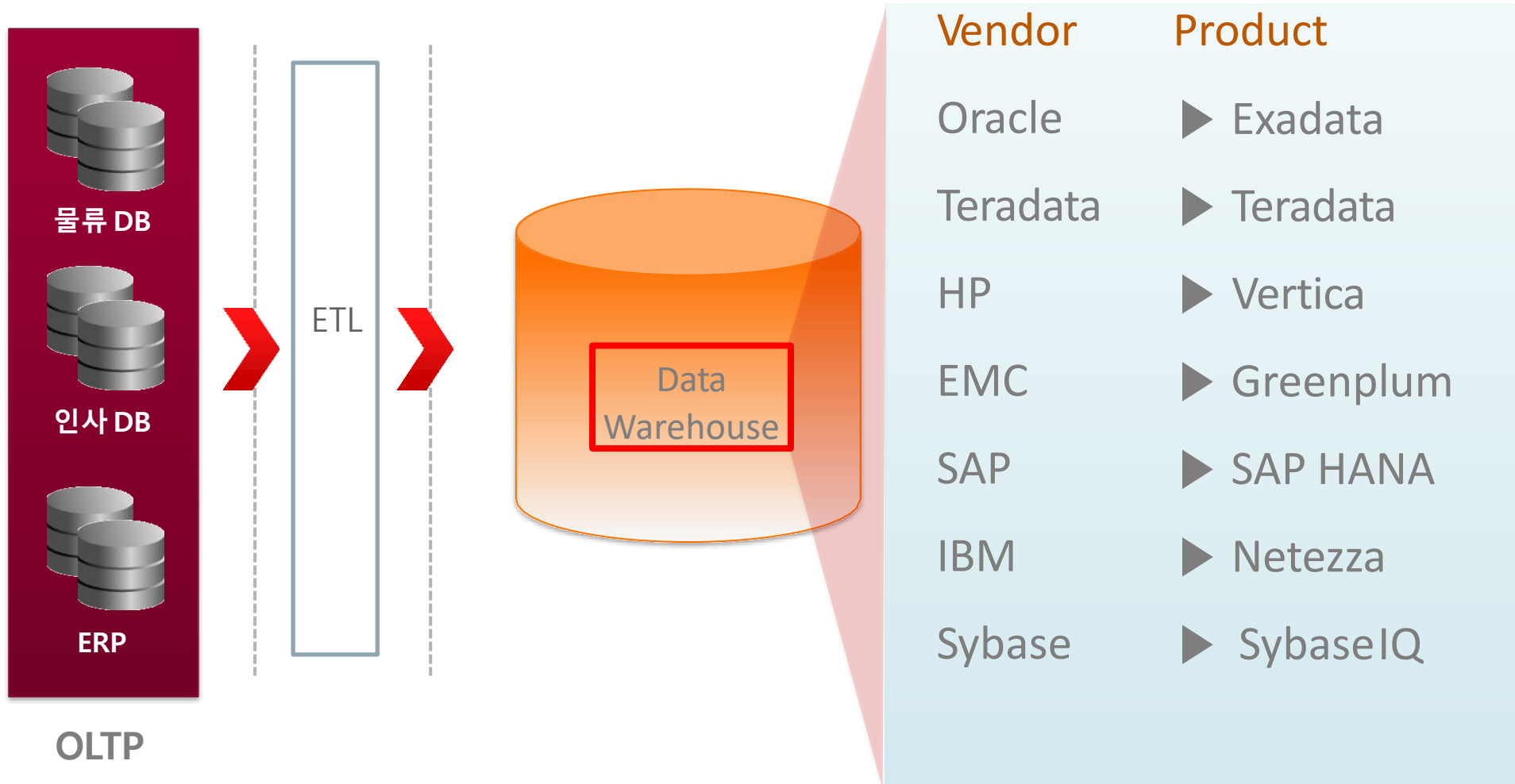
# 데이터베이스 용도별 DBMS

- DW 기본 Architecture

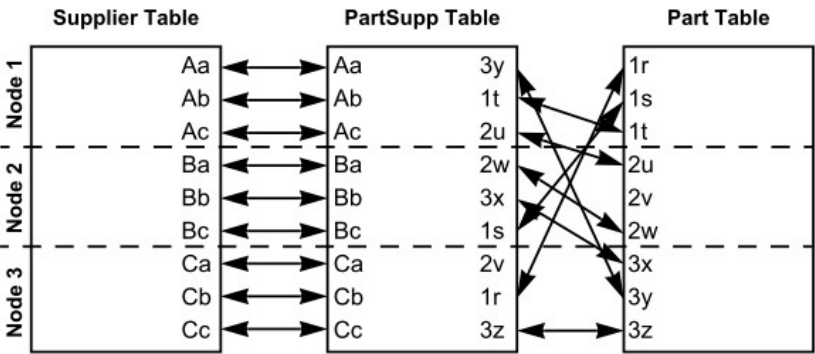
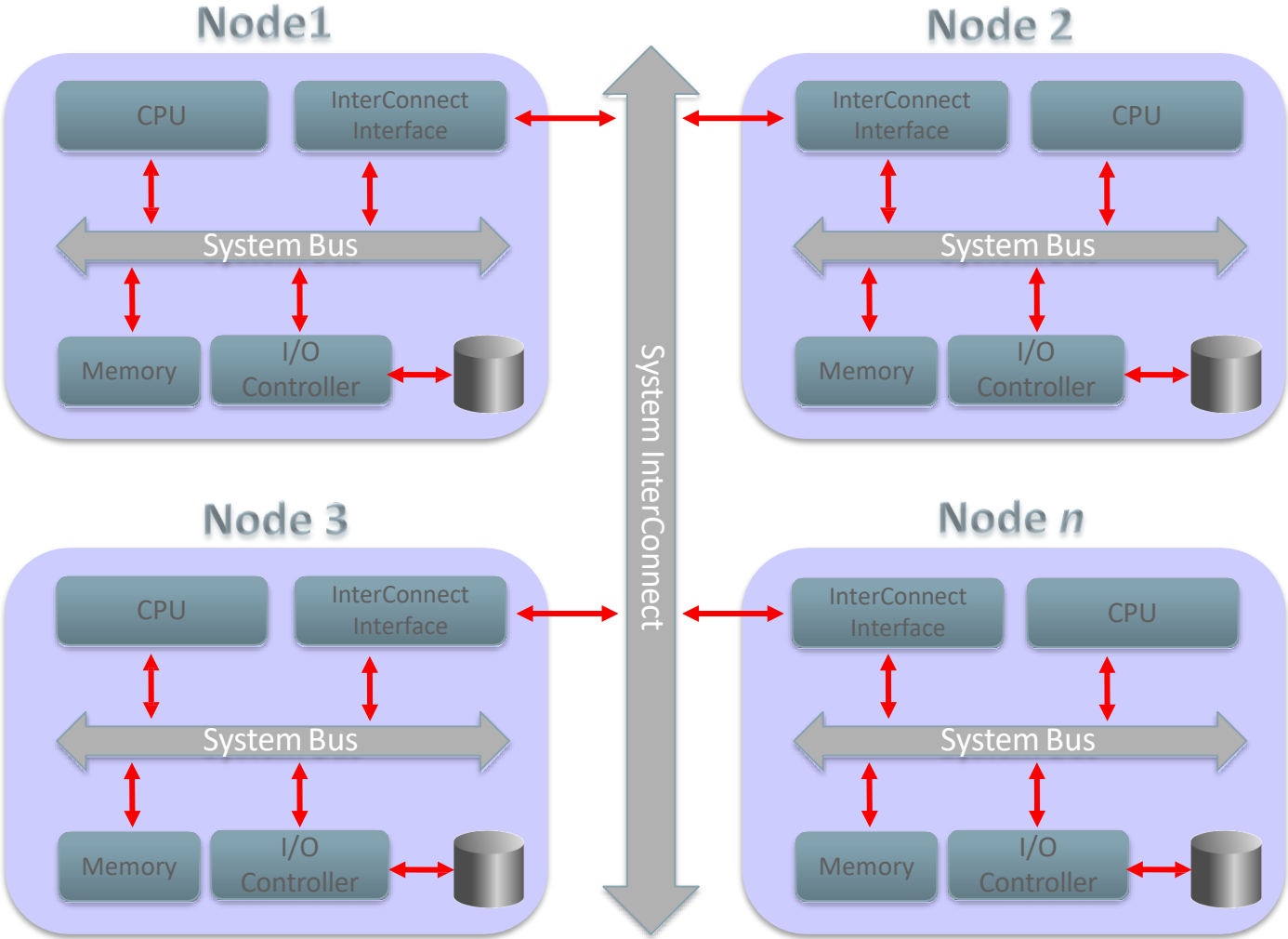


# 데이터베이스 용도별 DBMS

- DW 기본 Architecture



# MPP란 무엇인가?

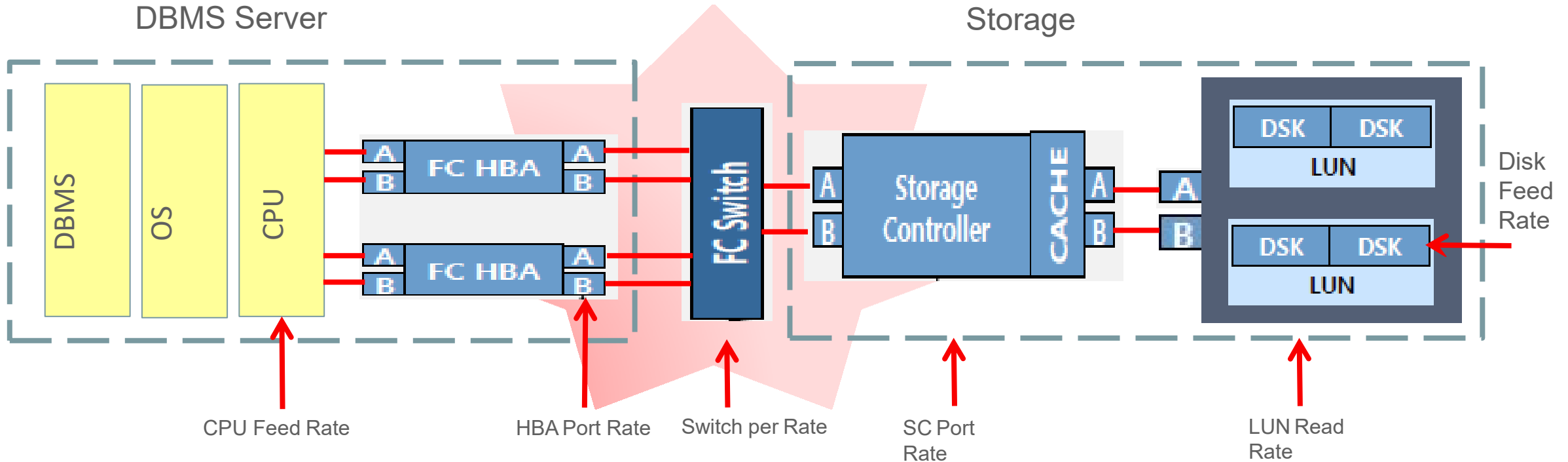


MPP: costly inter-node communication required



# Exadata 스토리지의 출현 배경은 SMP의 한계

- 일반 스토리지 구성 시 스토리지 데이터 대역폭(Bandwidth)한계로 병목 현상 발생



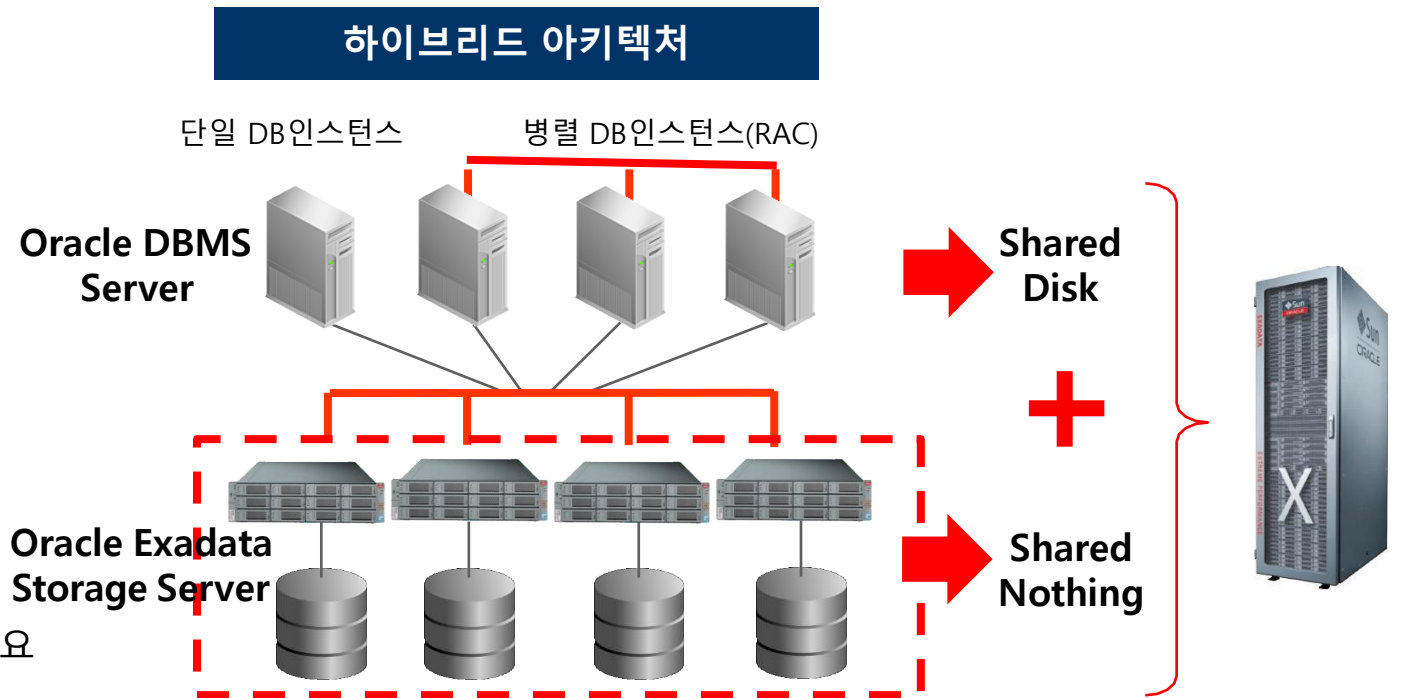
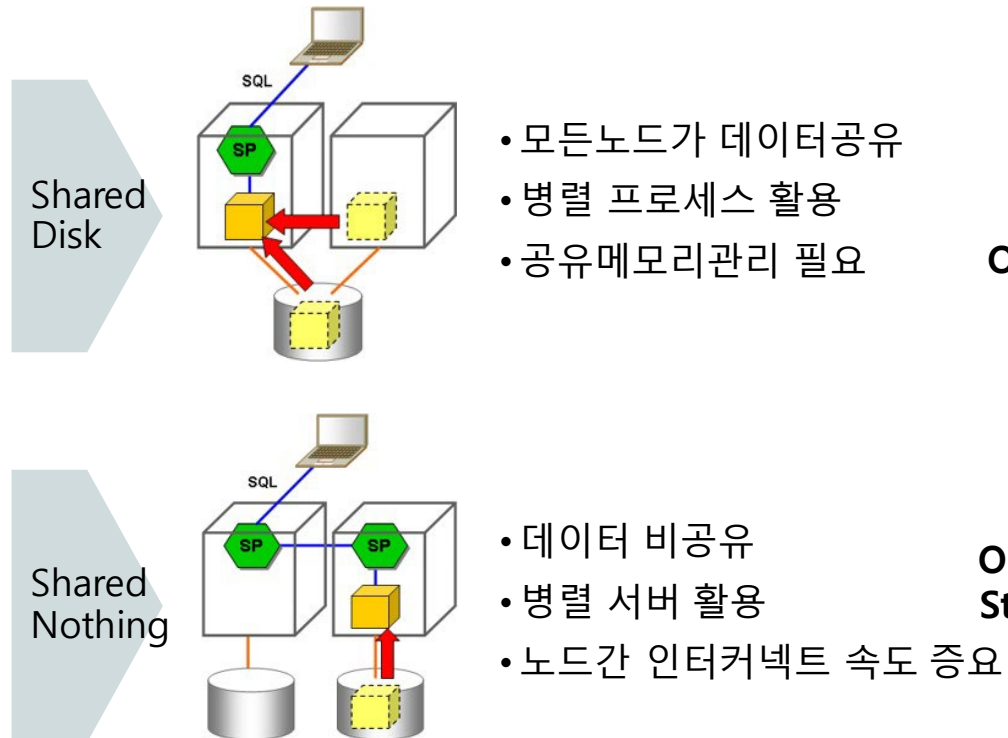
- 데이터베이스 성능은 Storage에 의해 제한되는 상황
  - 데이터는 기하급수적으로 늘어나고 있는 상황에서 기존의 SAN(Storage Array Network) 기반의 구조에서는 Database의 성능이 제한되어 대용량 데이터 처리에 한계가 존재

# Exadata 아키텍처

Exadata는 SMP의 장점과 MPP의 장점을 수용한 하이브리드 아키텍처 방식

- SMP와 MPP는 수십년 동안 논란이 이어져 왔음
  - SMP(Symmetric Multi Processing : Shared Disk) → OLTP
  - MPP(Massive Parallel Processing : Shared Nothing) → DW

Exadata는 일관된 네트워킹, 임팩트, 실행용, 고품질의 것



# Exadata 스토리지의 출현 배경

- 대역폭의 한계를 해결하는 방안



고속도로 정체가  
심한 경우  
해결 방안은?

해결 방안 1

통행 차량을 줄인다

해결 방안 2

고속도로의 차선을 넓힌다

해결 방안 3

고속도로를 추가로 건설한다



# Exadata 스토리지의 출현 배경

- 대역폭의 한계를 해결하는 방안

해결 방안 1

통행 차량을 줄인다



Cell-offload 을 통해 Storage에서  
필요한 데이터만 추출 / 전송

해결 방안 2

고속도로의 차선을 넓힌다



Network로 기존 대역폭을 6배 향상  
(8/16Gbps -> 100Gbps)

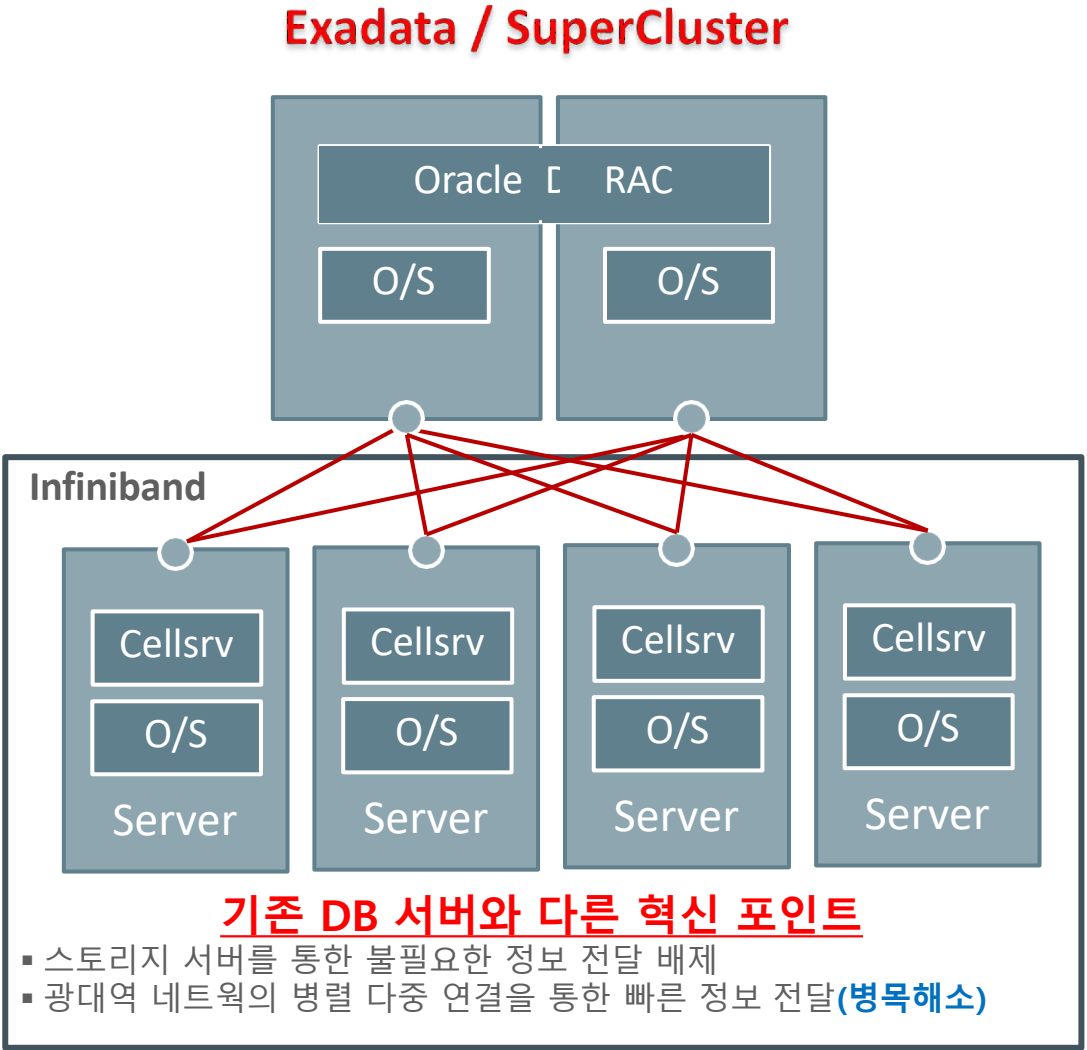
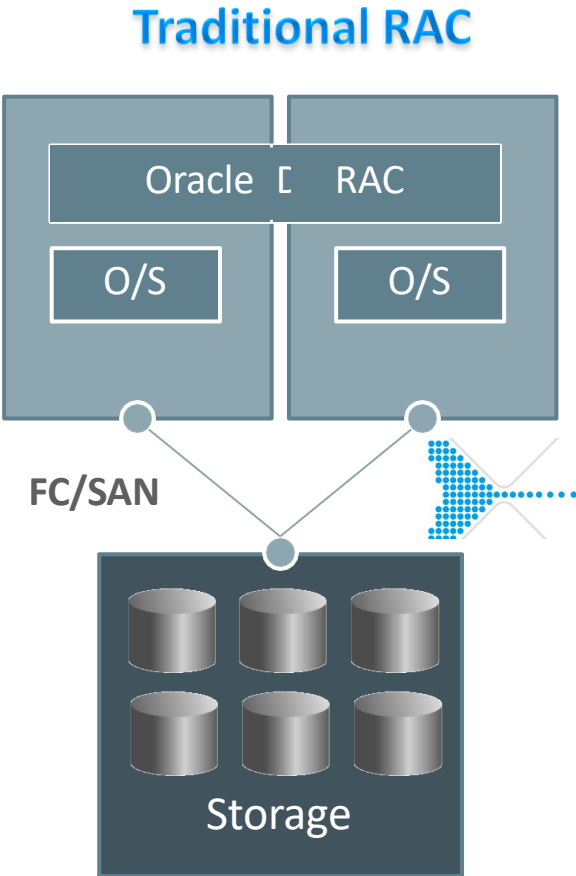
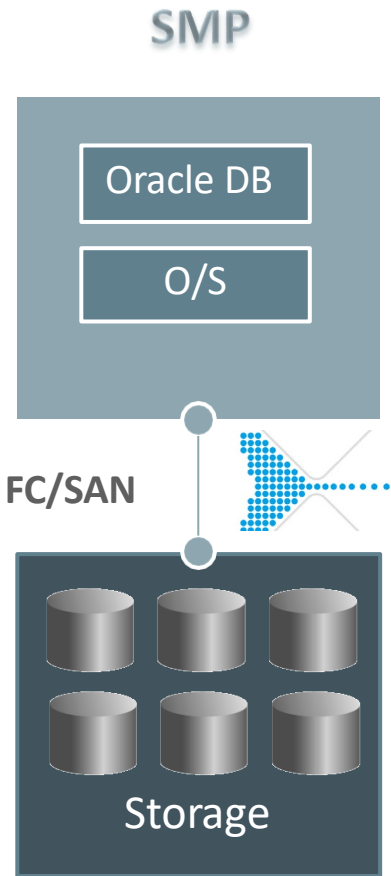
해결 방안 3

고속도로를 추가로 건설



스토리지 단위별로 Channel을  
추가하는 새로운 Architecture 도입

# 스토리지 별로 Channel을 추가하여 무한 Channel 확장



---

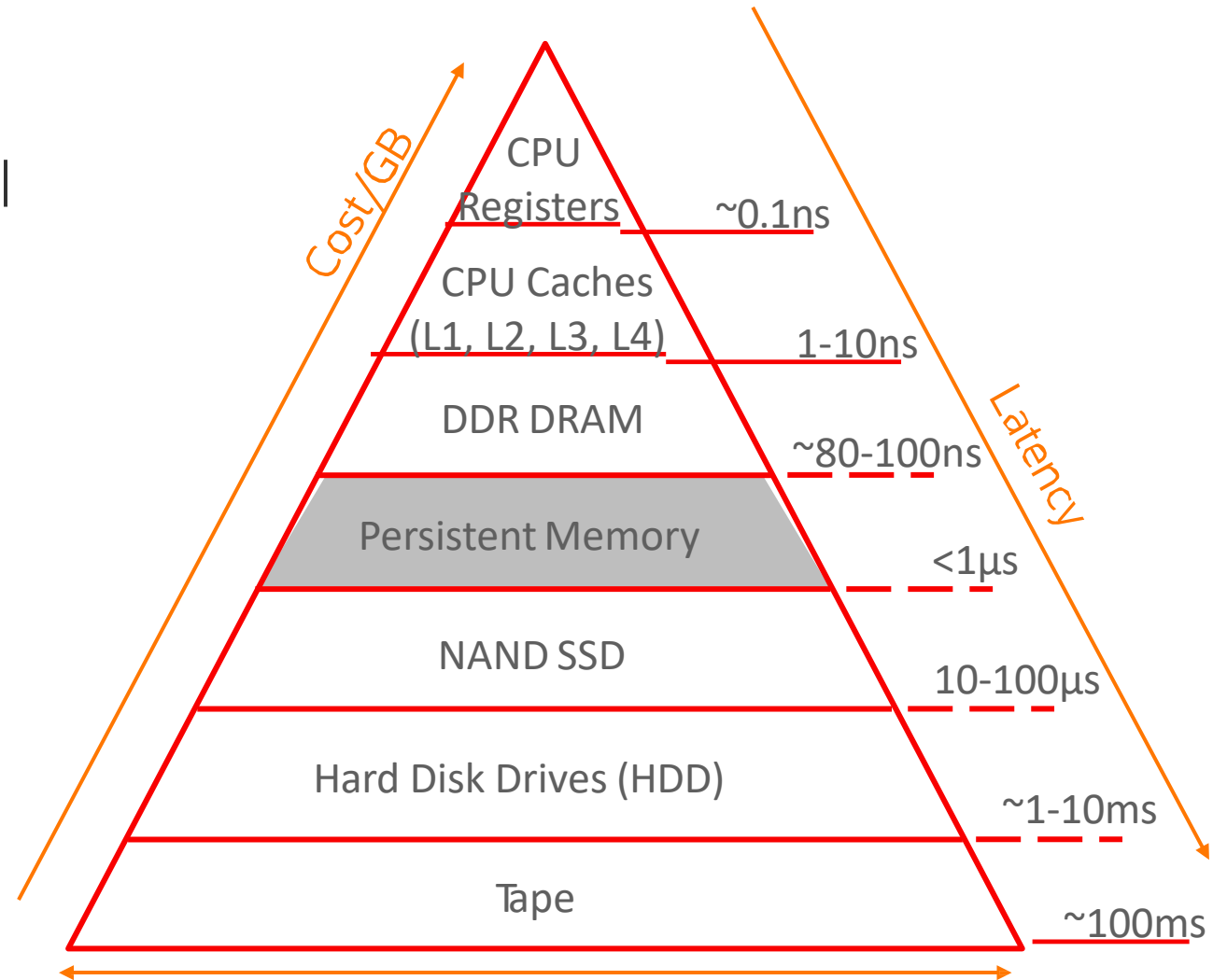
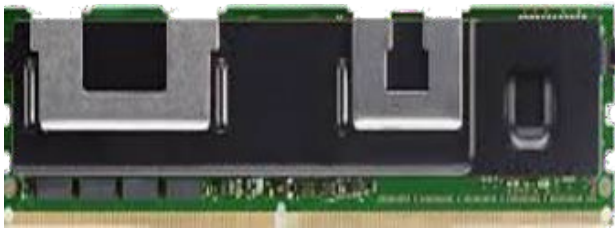
## Exadata OLTP로의 도약 X8M의 새로운 기능





# 새로운 Persistent Memory의 Exadata 적용

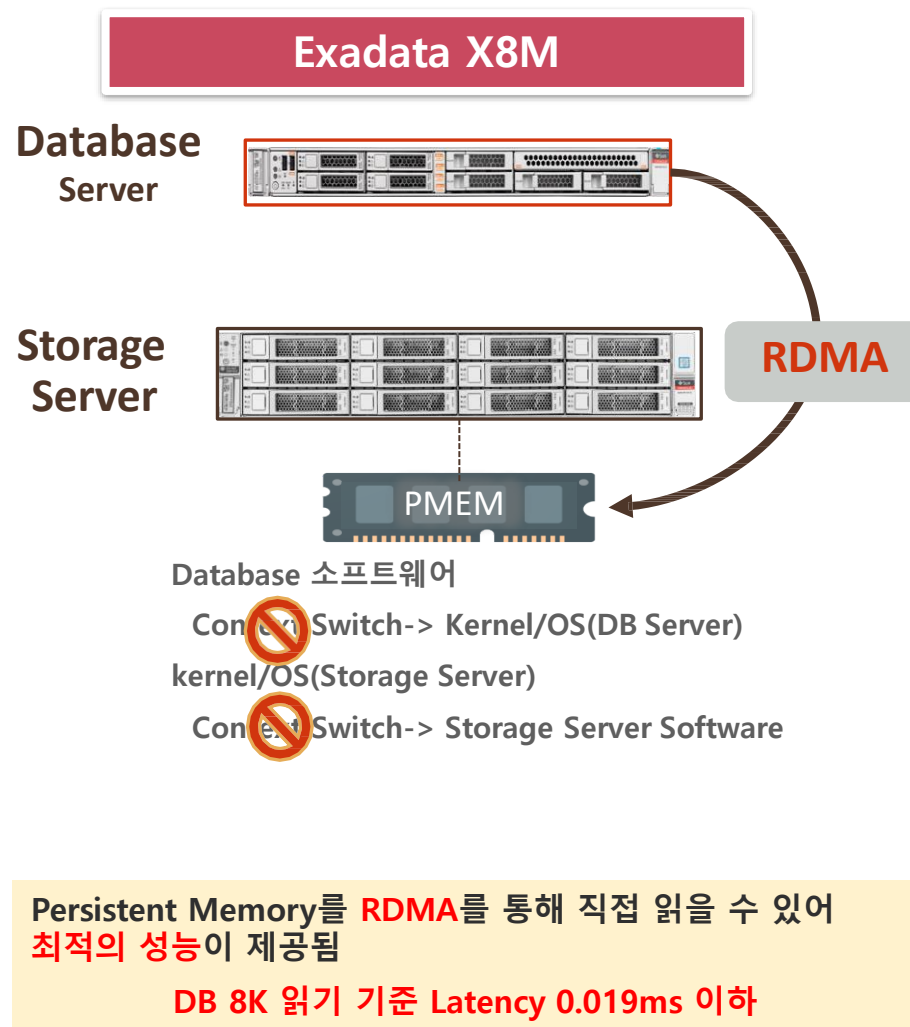
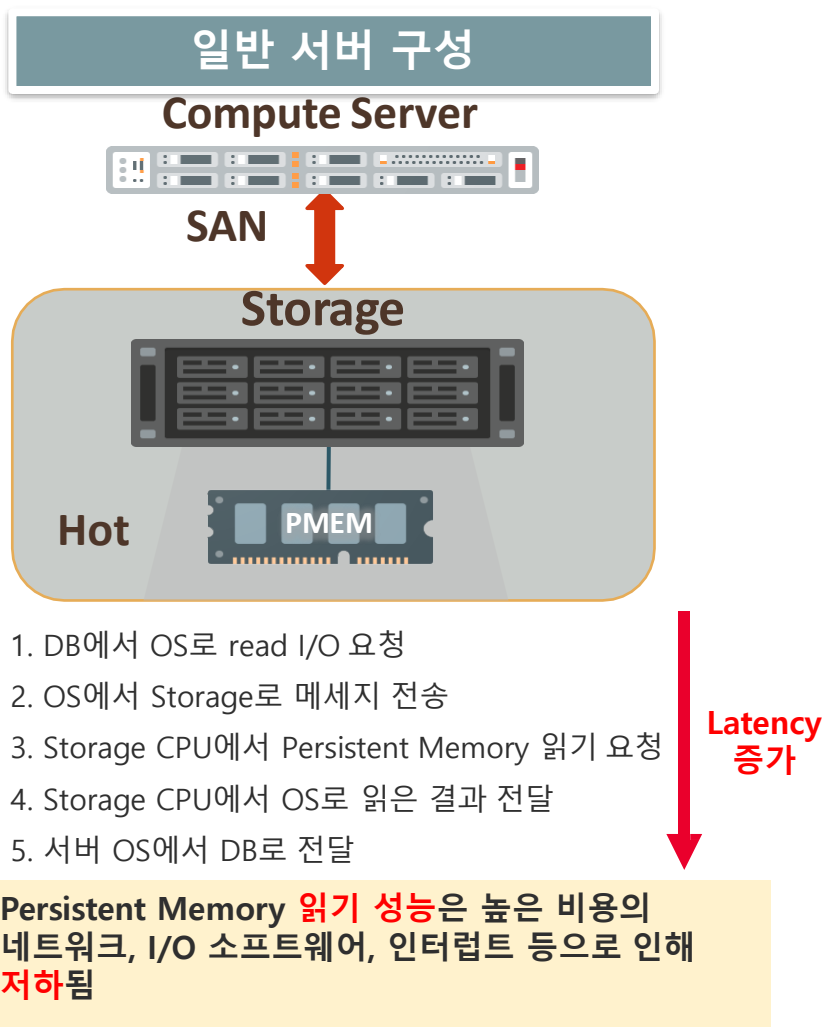
- Persistent memory는 새로운 실리콘 기술
  - 용량, 성능 및 가격은 DRAM과 플래시 사이
- Intel® Optane™ DC Persistent Memory:
  - 메모리 속도 읽기 - 플래시보다 훨씬 빠름
  - DRAM과 달리 정전시에도 쓰기 보존
- Exadata는 PMEM을 Read는 Cache로 사용하여 사용량을 최대로 유지하고 Re-log write에 PMEM을 사용함.



# X8M 특징점 - 고성능

## RoCE 네트워크 - Remote Direct Memory Access

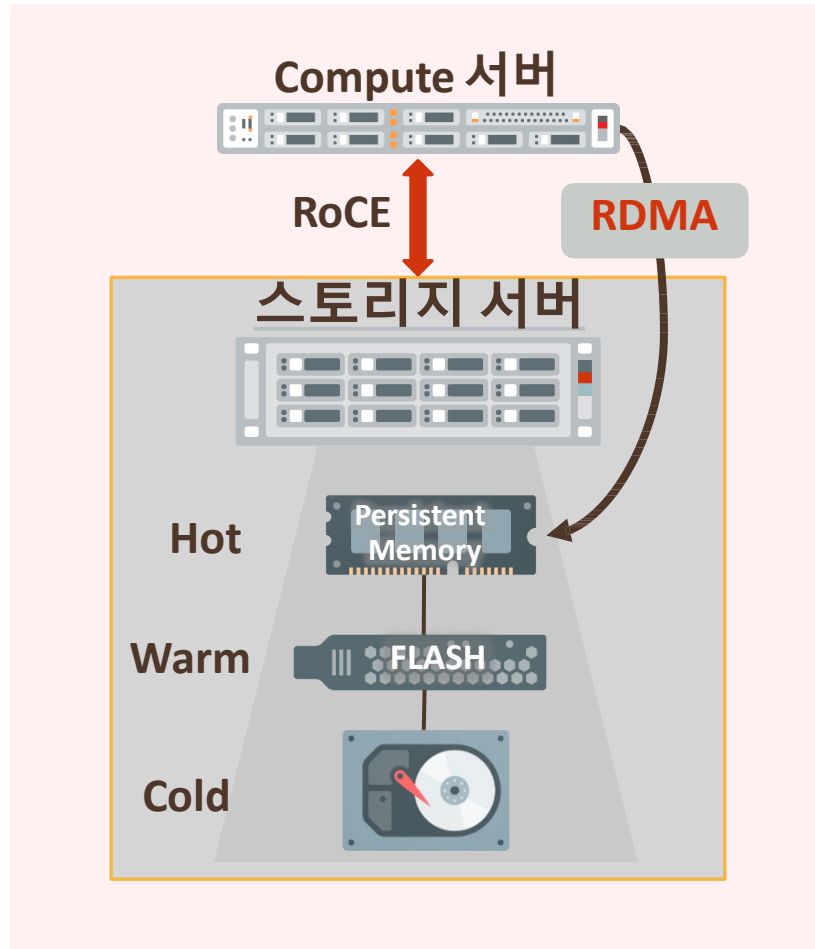
Exadata RDMA 기술은 RoCE 스위치를 통해 100Gbps의 속도로 Ethernet fabric을 통해 안정적으로 제공됨



# X8M 특징점 - 고성능

## PMEM(Persistent Memory) – PMEM Cache 데이터 가속

Exadata X8M 장비는 데이터베이스에 최적화된 세계 최초, 유일한 공유 방식의 Persistent Memory가 탑재 되어 IO 성능을 극대화 시킴



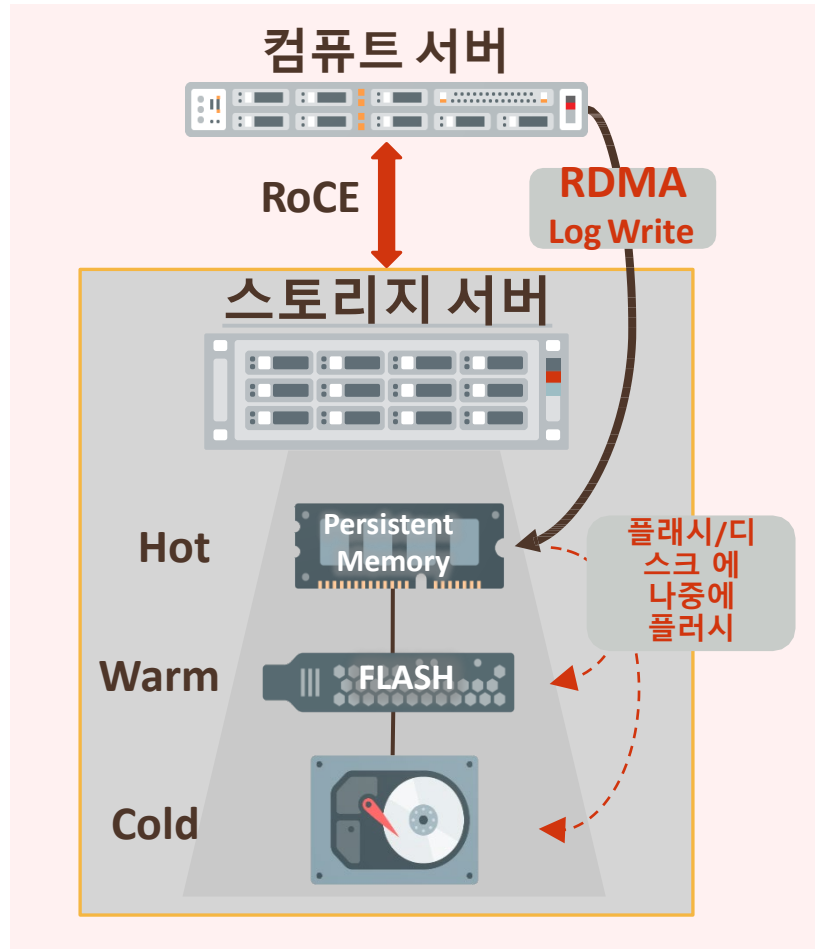
- Exadata 스토리지 서버는 플래시 메모리 앞에 Persistent Memory Accelerator를 투명하게 추가
  - ✓ 기존 X8 대비 2.5X 높은 IOPS – 1600 만 IOPS(FR)
- 데이터베이스는 IO 대신 RDMA를 사용하여 원격 PMEM 읽음
  - ✓ 네트워크 및 IO 소프트웨어, 인터럽트, 컨텍스트 스위치 우회
  - ✓ 기존 X8대비 10X 빠른 latency - <19  $\mu$ sec (8K database read)
- PMEM이 자동으로 계층화 되고 DB간 모두 공유
  - ✓ 가장 많이 사용되는 데이터를 캐시 하여 유효 용량 10X 증가
- Fault-tolerance를 위해 스토리지 서버에서 자동으로 Persistent Memory 미러링

Exadata System Software 19.3 및 Database Software 19c 에서 가능

# X8M 특징점 - 고성능

## PMEM(Persistent Memory) – PMEM Log Commit 가속

Exadata X8M 장비는 DB에 최적화된 세계 최초의 유일한 공유 방식의 Persistent Memory가 탑재 되어 Log write 성능을 극대화 시킴



### ▪ Log Write latency는 OLTP 성능에 매우

#### 중요

- ✓ 빠른 Redo Log write IO는 빠른 트랜잭션 Commit 시간을 의미함
- ✓ Log write 속도 저하는 전체 데이터베이스 성능에 영향을 줌

### ▪ 자동 Commit 가속 기능

- ✓ 데이터베이스는 여러 스토리지 서버의 PMEM에 단방향으로 RDMA 쓰기를 수행함
- ✓ 네트워크 및 IO 소프트웨어, 인터럽트, 컨텍스트 스위치 등을 건너뛸
- ✓ 기존 X8대비 최대 8x 빠른 Log Writes

Exadata System Software 19.3 및 Database Software 19c 에서 가능