

[HV HIS Partner Seminar 2026]

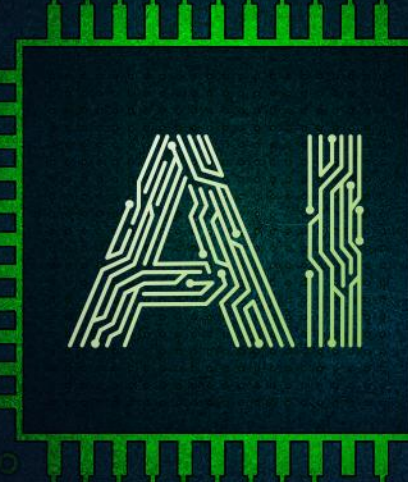
서버 패러다임의 전환

Arm 서버가 만들어가는 새로운 기준

이명관 프로

HS효성인포메이션시스템 데이터사업팀

2026년 3월 4일



Agenda

1. AI 산업 환경 변화
2. Why Arm
3. 제품 라인업 및 레퍼런스

1. AI 산업 환경 변화



AI 워크로드와 기준 변화

- AI 워크로드는 목적과 특성에 따라 학습과 그 이후 두 가지 영역으로 구분
- AI는 이제 PoC 단계를 넘어 업무 프로세스와 현장 시스템에서 상시 개입, 인프라를 평가 기준이 성능에서 운영 안정성으로 변화 중

모델 학습

AI

추론 & 운영



PAST
AI 모델 중심 Model Centric

크기 & 속도

얼마나 큰 모델을 빠르게 학습하는가?

GPU 집약 성능

연산 능력 절대치 중시



NEW
시스템 중심 System Centric

가용성 & 지속성

서비스/현장에서 상시 사용 가능한가?

운영 효율성 관점

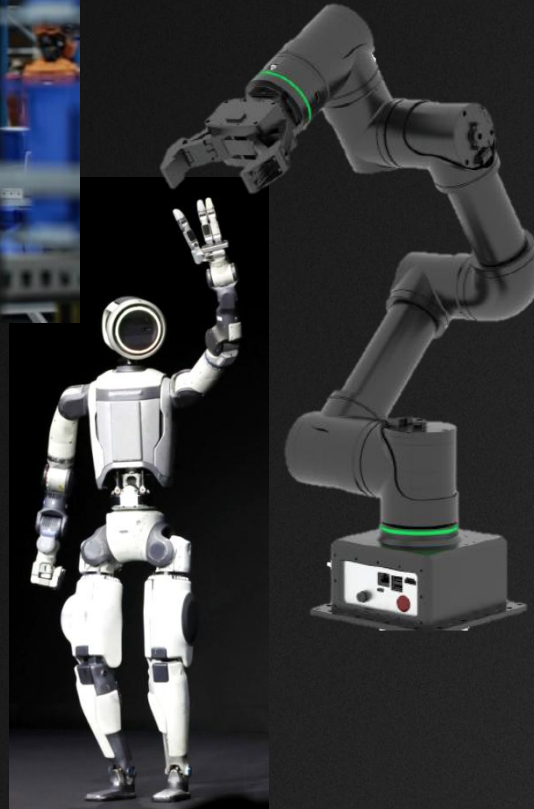
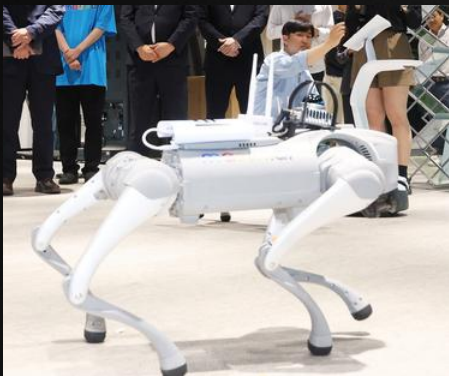
전력 비용, 공간 활용 등 운영 효율성 중시



Physical AI 확산과 Arm의 역할

AI 산업 환경 변화

- 자율주행차, 스마트 공장, 의료 현장 등 Physical AI의 범위 확산 중
- Arm 아키텍처는 낮은 지연 시간, 전력/발열 제약 환경에서의 안정적 운영 등 Physical AI의 요구 조건에 충족



CHALLENGE
현장의 요구 조건

- ✓ 낮은 지연 시간
- ✓ 전력/발열 제약 환경에서의 운용
- ✓ 장시간 무중단 운용



ARM GreenCore

Arm 아키텍처란?

- X86 대비 적은 수의 단순한 명령 집합을 빠르게 수행하도록 설계된 프로세서
- 단순한 구조로 인해 저전력, 저발열, 내열성 등의 장점을 보유

ARM

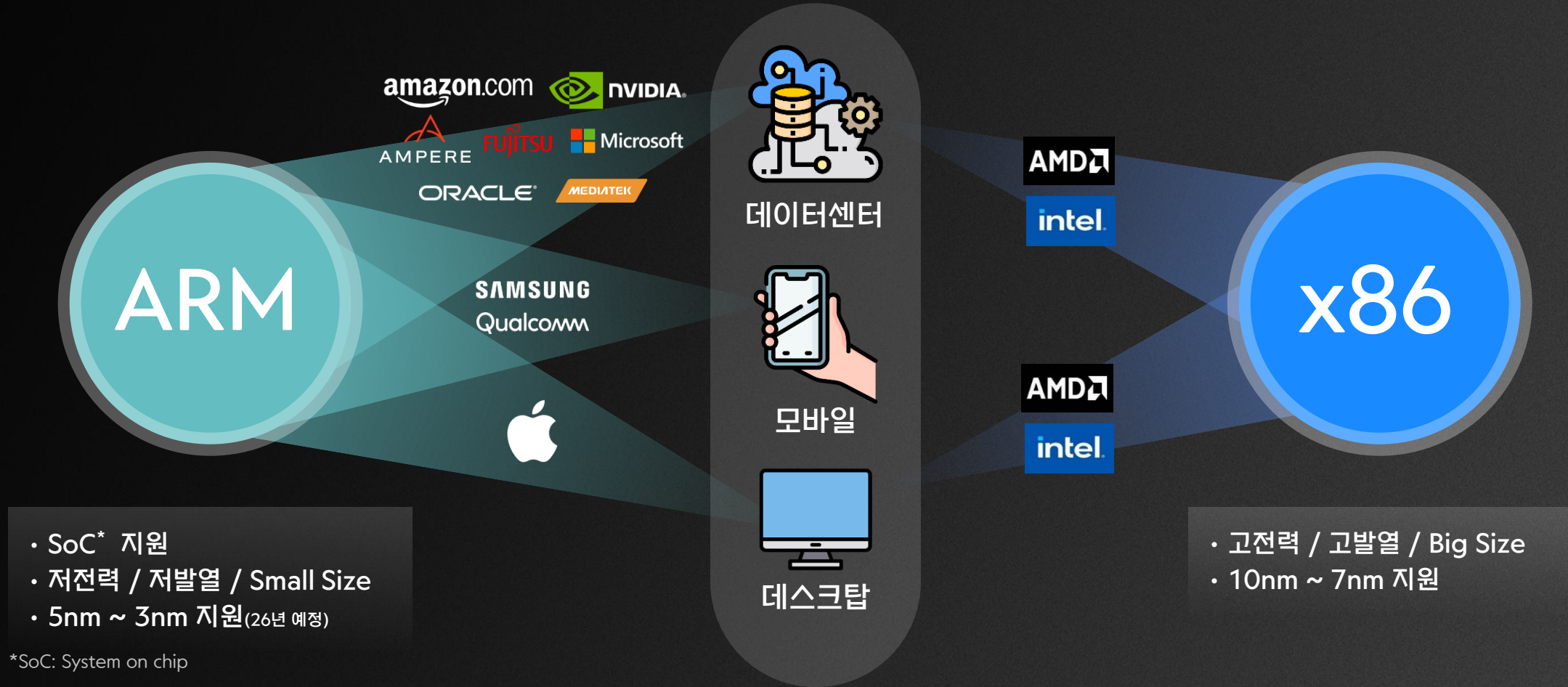
VS



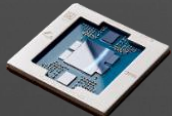

Reduced Instruction Set Computing (RISC)	아키텍처	Complex Instruction Set Computing (CISC)
단순한 작업을 빠르게 수행하여 효율성 증대	개발 목표	다양한 작업을 수행 가능하여 범용성 중시
멀티코어 병렬 구조 단순, 트랜지스터 적음	하드웨어 구조	고성능 단일코어 중심 복잡, 트랜지스터 많음
높음	전력 효율	낮음
적음	발열	많음

글로벌 ARM 시장 확대

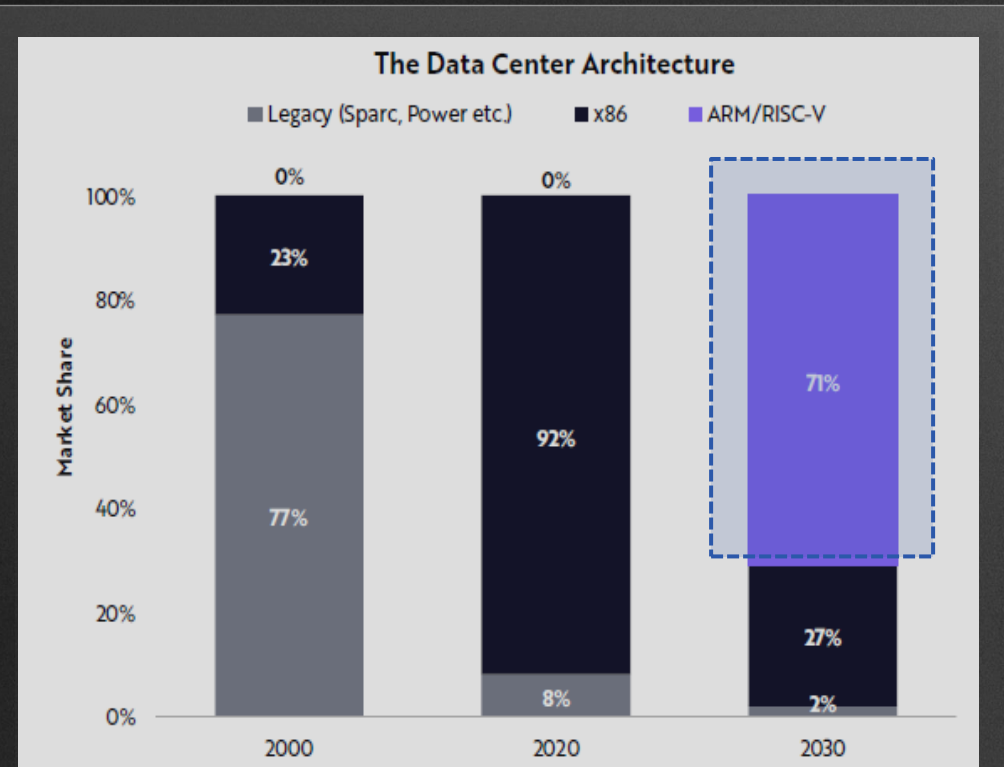
AI 산업 환경 변화



글로벌 대기업 ARM 시장 확대

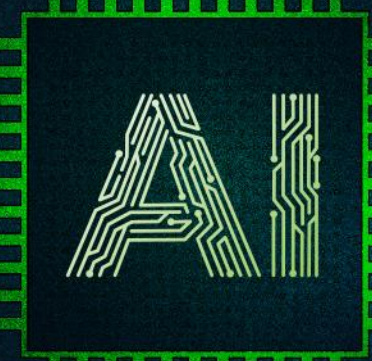
<p>AWS / 2018</p>  <ul style="list-style-type: none"> 2018년 클라우드 서버용 ARM '그래비톤' 발표 ARM 칩 50% 이상 사용 20%~70% 비용 저감 	<p>Microsoft / 2022</p>  <ul style="list-style-type: none"> Azure 클라우드 ARM칩 기반 '코발트 100' 발표 인텔 대비 최대 50% 효율성(성능) 향상 발표
<p>NVIDIA / 2023</p>  <ul style="list-style-type: none"> ARM 기반 데이터센터 CPU 'Grace' 발표 GPU 'Hopper'를 연결한 GH200 슈퍼칩 발표 	<p>SKT / 2023</p>  <ul style="list-style-type: none"> AWS '그래비톤' 클라우드 인프라 환경 국내 구축 (40% 비용↓, 120% 성능↑) ARM 기반 클라우드 상품 출시 (x86 대비 40% 절감)

ARM 기반 아키텍처 시장 전망



* 출처: ARK Investment Management LLC

2. Arm 서버 도입 효과



도입 원가 절감

- X86(Intel & AMD) 대비 CPU 칩 자체 비용 저렴
- 서버 팬, 전원 등 x86 대비 저렴한 부품 구성 가능



실시간 병렬 처리 지원

- 멀티코어 기반의 실시간 병렬 처리 지원, 데이터 병목현상을 줄여 빠른 응답속도 기대 가능
- 32/64/80/128/192 Core 등 고객 요구에 맞춘 코어 대응 가능



저전력/저발열 지원

- RISC의 단순한 구조로 인해 x86 대비 적은 전력으로 원활한 운영 가능
- 저발열로 인해 냉방비 및 공간 활용 측면에서 이점이 있음



환경적 제약 극복

- X86 대비 전원/온도/습도 제약이 적어 열악한 환경 속에서도 운영 가능
- 전장 장비(배, 트럭 등)/사무실/야외 환경 등 상온에서 운영 가능



최고 사양 CPU 칩 가격 비교

CPU	모델	Core	가격	코어당 금액
ARM	Altra MAX 128-3.0GB	128 Core	\$5,800	\$45.3/Core
Intel	Xeon Platinum 8380 2.3GB	40 Core	\$8,409	\$210.2/Core
AMD	EPYC 7763 2.4GB	64 Core	\$7,890	\$123.3/Core

최저 사양 CPU 칩 가격 비교

CPU	모델	Core	가격	코어당 금액
ARM	Altra Q32-1.7GB	32 Core	\$800	\$27/Core
Intel	Xeon Silver 'IceLake-SP'	8 Core	\$500	\$62.5/Core
AMD	EPYC 7302	16 Core	\$978	\$61.1/Core

ARM

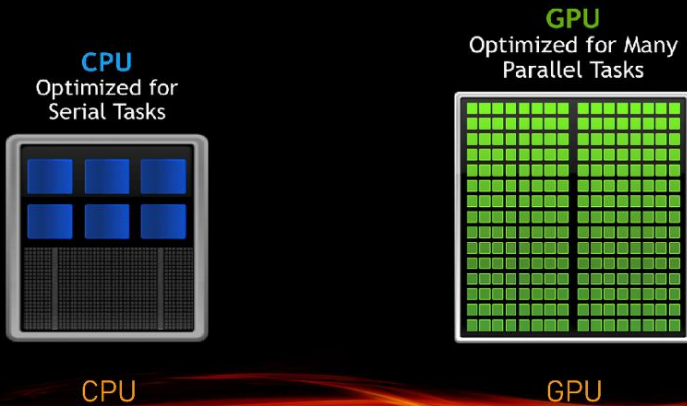


VS

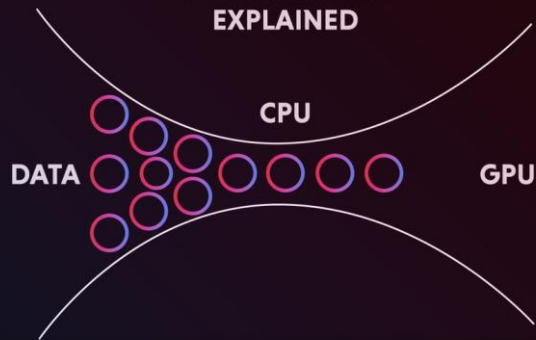


(Arm) 막강한 병렬 처리 능력을 통한 '데이터 전처리' 가속

Nvidia GPU architecture



PC BOTTLENECK EXPLAINED

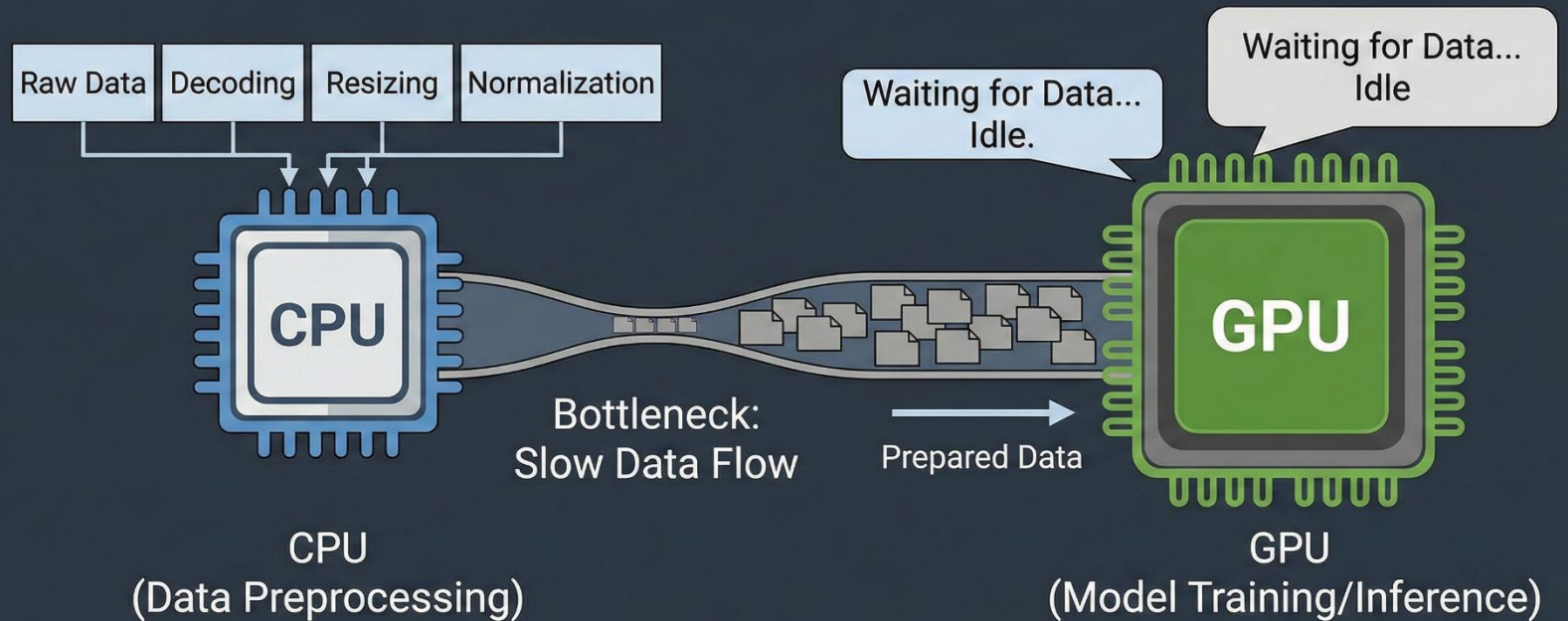


Bottleneck occurs when there is a limit on how much data can be processed by a computer. If there is differences in processing speed, bottleneck will happen. CPU bottleneck happens when the processor isn't fast enough to process/transfer the data received. The same concept applies to GPUs - only this happens if you pair a fast processor with a low level GPU.

#4

데이터 기아 (Data Starvation) 현상

CPU Bottleneck in AI Server



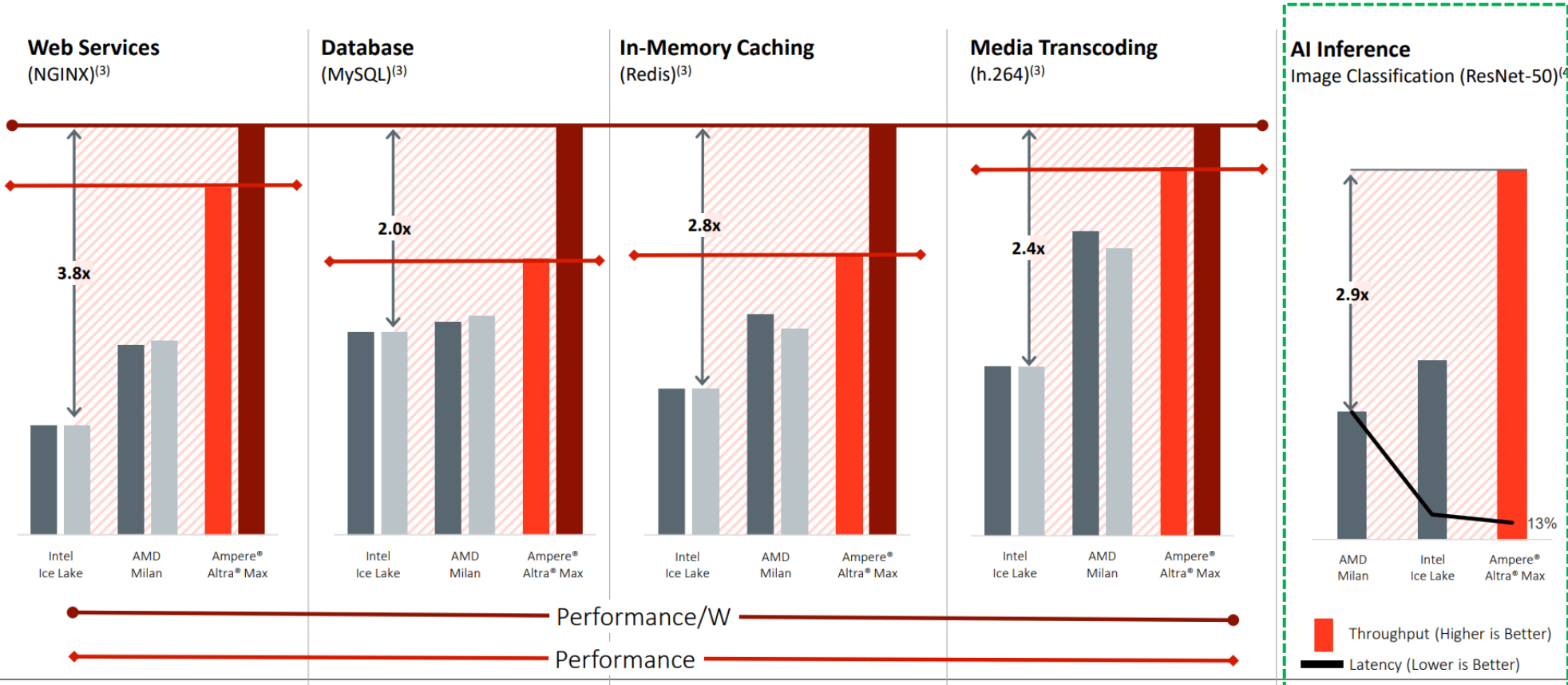
CPU struggles to prepare data fast enough, causing the powerful GPU to wait, limiting overall performance.



(현재) TCO & ESG 달성을 위한 Green Computing

Ampere: Leadership Performance for Cloud Workloads

Highest Performance and Power Efficiency Across Key Cloud Workloads⁽¹⁾⁽²⁾



워크로드 별 와트 당 성능

- 조건 : 멀티 코어 운용
- AI 서비스(추론) 운용에 최적화 된 솔루션

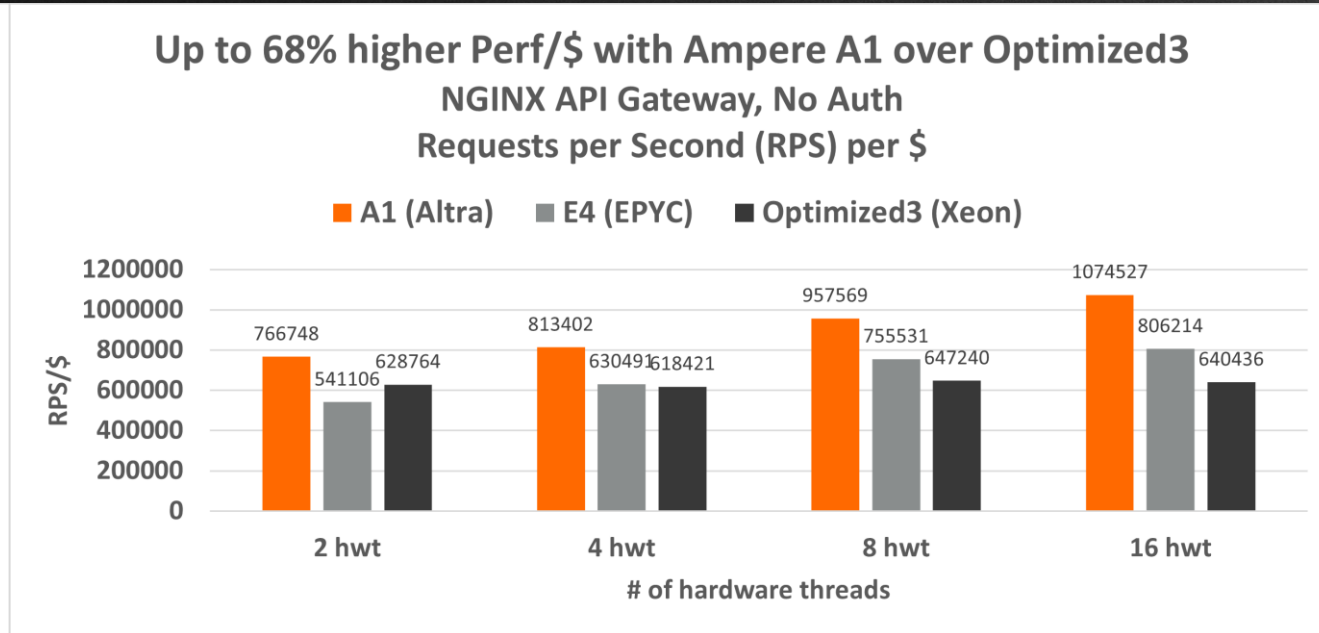
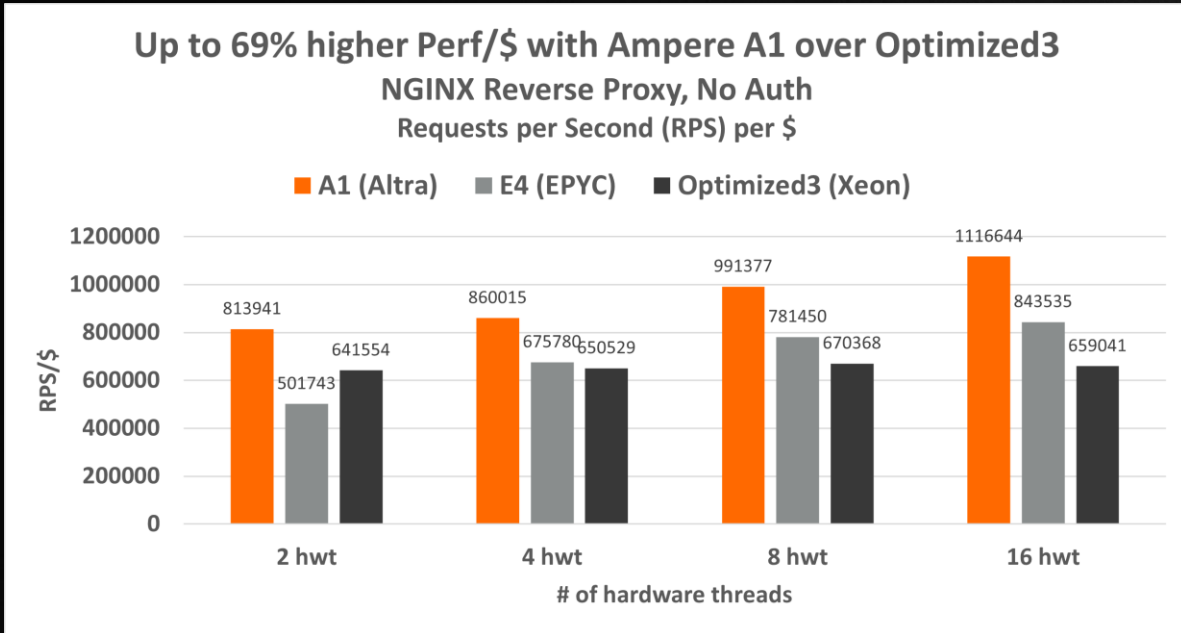


Notes:
 1. Based on Company benchmarking
 2. Intel Ice Lake represents Intel 8380 SKU; AMD Milan represents AMD 7763 SKU.
 3. Percentages represent AMD Milan and Ampere Altra Max indexed against Intel Ice Lake
 4. Percentages represent Intel Ice Lake and Ampere Altra Max indexed against AMD Milan



(미래) 클라우드 네이티브와 AI 시대를 위한 인프라

OCI 기반 Altra (Arm 서버) 워크로드 성능: No Auth 상태의 NGINX 가격 대비 성능 (가성비)



성능 평가 조건

- A1.Flex : Ampere® Altra® CPUs based on Arm Neoverse N1 cores, 3.0GHz
- E4.Flex: AMD EPYC third generation processors, 2.55GHz base, 3.5GHz single-core turbo
- Optimized3.Flex: latest third Gen Intel Xeon (Ice Lake) scalable processors, 3.0GHz base, 3.6GHz single-core turbo

CPU 종류별 클라우드 서비스 가격 요금 비교

OCI virtual machine	Per OCPU per hour(\$)	Per GB memory(\$)
Intel.VM.Optimized3	0.054	0.0015
AMD.VM.E4	0.025	0.0015
Ampere.VM.A1	0.01	0.0015

AI가 불러온 '전력 위기'...대비 나선 샘 알트먼, '이것'에 276억 '베팅'

MS 나델라 "AI산업 병목은 GPU 공급 아닌 전력 부족"



ARM 칩 기반 서버

VS

X86 칩 기반 서버

1,142,892원/년

가상 운용 비용

2,469,058원/년



ARM
상면 효율화

115.6W

전력량
70%

265.0W

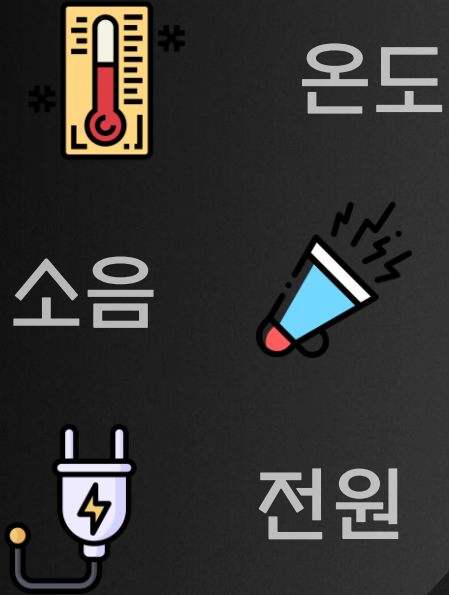
124.8W

전력량
100%

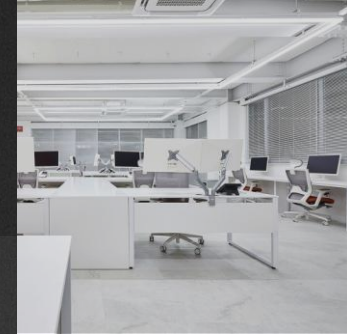
269.9W



x86
상면 효율화



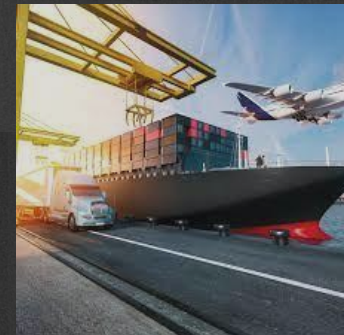
일반 사무실



공장 설비실



운항 장비
(자동차/배/로봇)



3. 제품 및 레퍼런스



HS 효성인포메이션시스템

ARM 서버

서버 보드 직접 설계/개발



- 국내 유일 ARM 서버 보드 일체 설계/개발 협력 지원
- 고객 중심 커스텀 서버 제작 지원

맞춤형 관리 환경



- 국내 ARM 서버 전용
- BMC 제공 및 관리 GUI 화면 제공

All in One 통합 지원



- H/W 와 S/W의 통합 지원
- ARM 기반 오픈소스 대한 가이드 지원

ARM 기반 고급 인력 확보



- ARM 관련 S/W 전체 설치 지원
- ARM 기반의 고급 S/W 엔지니어 확보 및 기술력 보증

GreenCore 제품 라인업

제품 및 레퍼런스

모델	GreenCore SQ mini	GreenCore LX	GreenCore Altra 32/64	GreenCore Altra 80/128
외형				
사이즈	1U	2U	1U/2U/4U	
CPU/Mem	24(1.0)core ~64Gb	16(2.2)core ~128Gb	32(1.7)/64(2.2)Core/ ~2TB	80(3.0)/128(3.0)Core ~4TB
Disk	HDD/SSD 지원	HDD/SSD 지원	HDD/SSD 지원	
포지션	엣지 서버 시장 (엔트리 환경)	데이터센터 서버 (미드레인지)	클라우드 서비스 / 데이터센터 고성능 서버 시장(하이엔드)	
주요 업무 (타겟)	저가용 서버/ 엣지 환경	데이터센터 저장 중심 서버	고성능 서버 클라우드 서버용 (가상화, 오픈스택, K8s 등)	



고성능 워크로드 대응을 위한
GPU / NPU 추가 구성 가능

온라인 및 실제 장비
테스트 환경 제공

H/W부터 S/W까지
고객 맞춤 구성 및 지원

멀티코어 서비스, 기존 장비 대체 등 오픈소스 기반 클라우드 구축에 최적의 솔루션

Open 가상화

- ✓ OpenStack 기반 오브젝트 스토리지 구축
- ✓ Ceph DR 환경 구축

- H 공공기관

모바일 환경 구축

- ✓ 모바일용 Win OS 환경 구축 테스트
- ✓ 기존 x86 대비 전력량 30% 감소 효과

- S 전자

엣지컴퓨팅

- ✓ 컴퓨터 운영 환경이 터프한 곳
- ✓ 스마트팩토리, 자율주행, 로봇 등

- H 자동차, S 전자
- H 주립대학(해외)

NPU 환경 지원

- ✓ 국내 생산 NPU 솔루션과의 협업 환경 구성
- ✓ 기존 Intel /AMD 대비 다수 Core 지원으로 성능 서포트 지원

- NPU Q사 / NPU M사

GreenCore 사례 (가상화 DR 구축 사례)

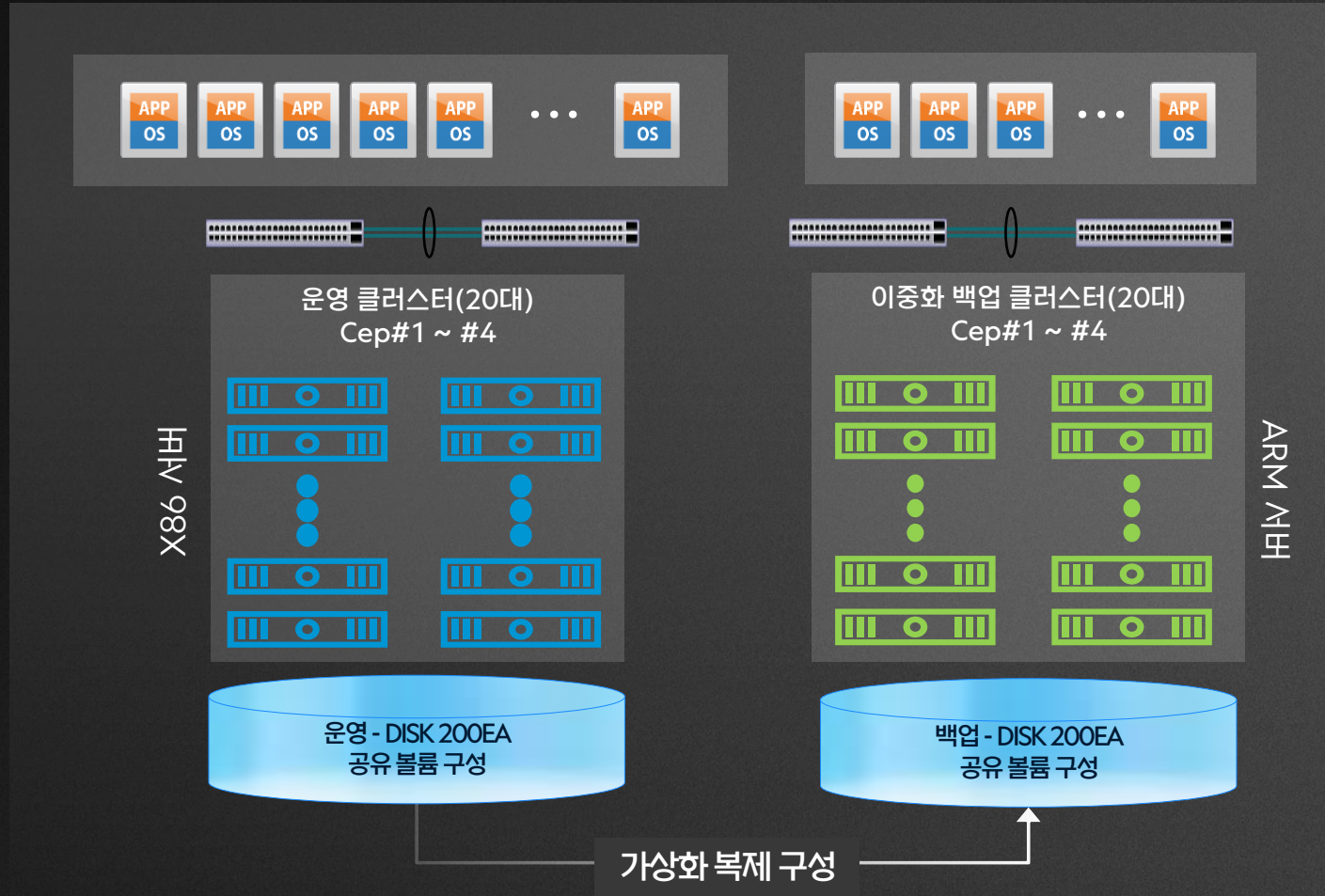
구성 특징

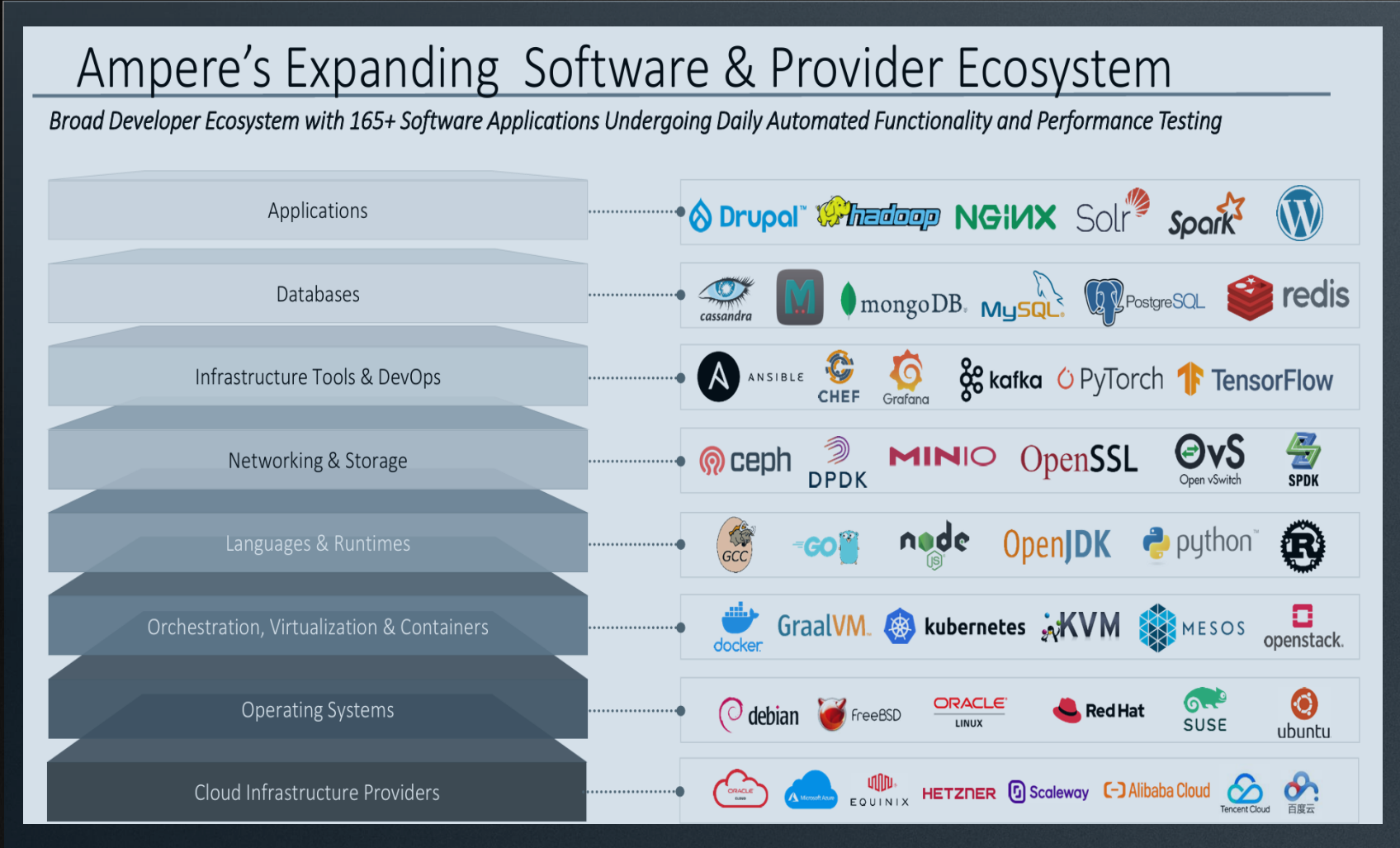


- ✓ H공공기관에 가상화 DR 구축
(ARM 서버 20대+ 구축)
- ✓ **OpenStack Ceph 설치를 통해**
대용량 스토리지 운영 (메인: x86 / DR: ARM-25% 절감)
- ✓ 기존 인텔 백업 서버를 자사의 ARM 서버로 교체, Ceph 기반의 분산형 스토리지 구축까지 일괄 지원

[구축 효과]

ARM 서버와 Ceph 기반 스토리지 구축으로 **비용 절감**, **호환성 확보**, **성능 지원** 등, 대용량 데이터 처리 효율 및 유연한 인프라 관리 효과 달성







다양한 테스트
환경 지원

온라인 구성 환경을 통한 체험 환경 지원

직접 장비 대여를 통한 테스트 환경 지원

DX센터 방문을 통한 테스트 환경 체험



스토리지 수준의
유지 보수 지원

H/W 와 S/W가 일체 된 일괄 된 지원 환경 제공

전국 지원을 통한 일괄 지원

스토리지 지원 환경과 동일한 지원 환경 제공

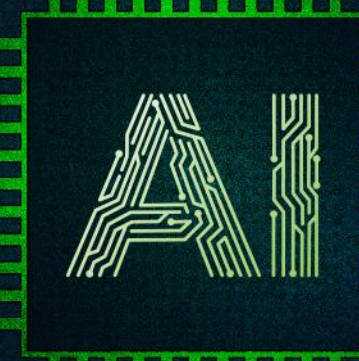
VM 테스트 링크

blueshift.xslab.co.kr



ARM 서버

이제 선택이 아닌
필수입니다.



감사합니다.

