

AI와 머신러닝을 통한 지능형 네트워크

AI 인프라 개요
AI를 위한 인프라 서비스
AI를 이용한 서비스

김병진(sirmd@outlook.com)

AI 인프라 개요

AI 이해하기

- 우리가 알고 있는 AI : Siri / 헤이 카카오 / 알파고 ?
- AI : Artificial Intelligence
- 인공지능은 1950년대 컴퓨터가 생각할 수 있는가? 에서 출발
- 사람이 수행하는 지능적인 작업을 자동화 하기 위한 연구 활동
- <https://ko.wikipedia.org/wiki/%EC%9D%B8%EA%B3%B5%EC%A7%80%EB%8A%A5>

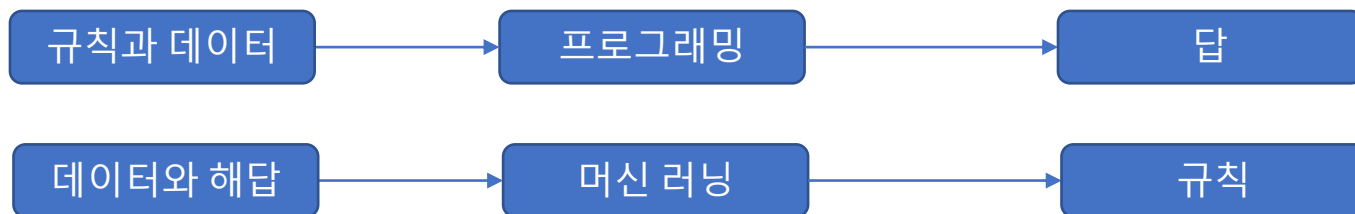
머신 러닝

- 명시적인 규칙을 만들어 지식을 다루면 인간 수준의 인공 지능을 만들 수 있도록 하는 것 : 체스 프로그램
- 심볼릭 AI = 체스 프로그램
- 이미지 분류, 음성 인식, 언어 번역 등의 복잡한 것은 해결하기 어려움 => 머신 러닝의 탄생
- AI 선구자 : 앨런 튜링(인터넷에서 참조하길 권장 : 튜링 테스트등)
- 우리가 어떤 것을 작동시키기 위해 “어떻게 명령할지 알고 있는 것 ” 이상을 컴퓨터가 처리하는 것이 가능한가?
- 특정 작업을 수행하는 방법을 스스로 학습할 수 있는가?
- 컴퓨터가 우리를 놀라게 할 수 있을까?
- 프로그래머가 직접 만든 데이터 처리 규칙 대신 컴퓨터가 데이터 확인하고 자동으로 규칙을 학습할 수 있을까?

AI 인프라 개요

프로그래밍과 머신 러닝

- 프로그램과 머신 러닝



- 머신 러닝은 훈련(Training)이 중요

머신 러닝 주요 사용하는 곳(1990년대 부터)

- 고성능의 컴퓨터의 증가
- 병렬 처리 기술의 발달
- 대량의 데이터 처리 : 데이터 셋
- 학습하기

AI 인프라 개요

머신 러닝을 넘어 딥 러닝

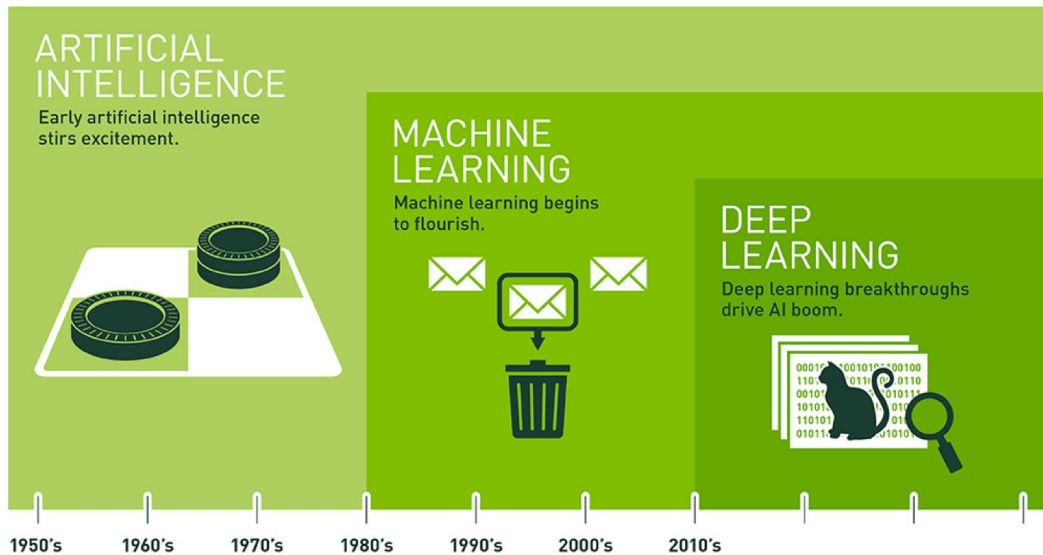
- 머신 러닝의 특정한 한 분야의 연속된 층(Layer)에서 점진적으로 의미 있는 표현을 배우는데 강점을 가지는 데이터 표현을 학습하는 새로운 방식
- Deep = 딥
- 연속된 층을 사용했는지 그리고 깊이
- Layered Representations Learning : 층 기반 표현 학습
- Hierarchical representations Learning : 계층적 표현 학습

딥 러닝과 신경망

- 기본층을 겹겹이 쌓아 올려 구성한 것 : 신경망 = Neural Network
- 신경 생물학의 용어를 사용한 것
- 뇌를 모델링 한 것이 아니라 뇌 구조를 이해하는 것에서부터 영감을 얻어 개발된 부분

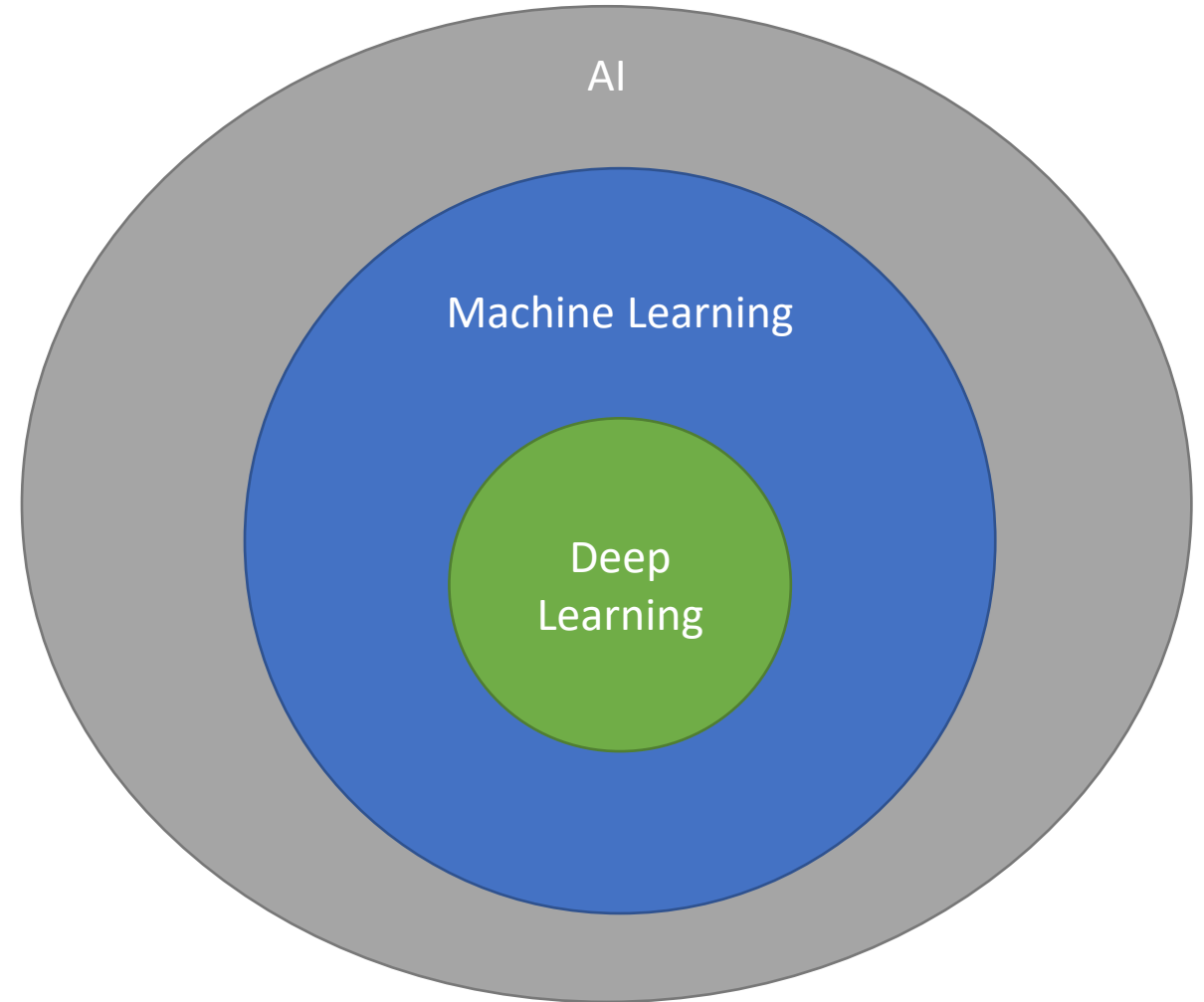
AI 인프라 개요

인터넷에 있는 그림 참조



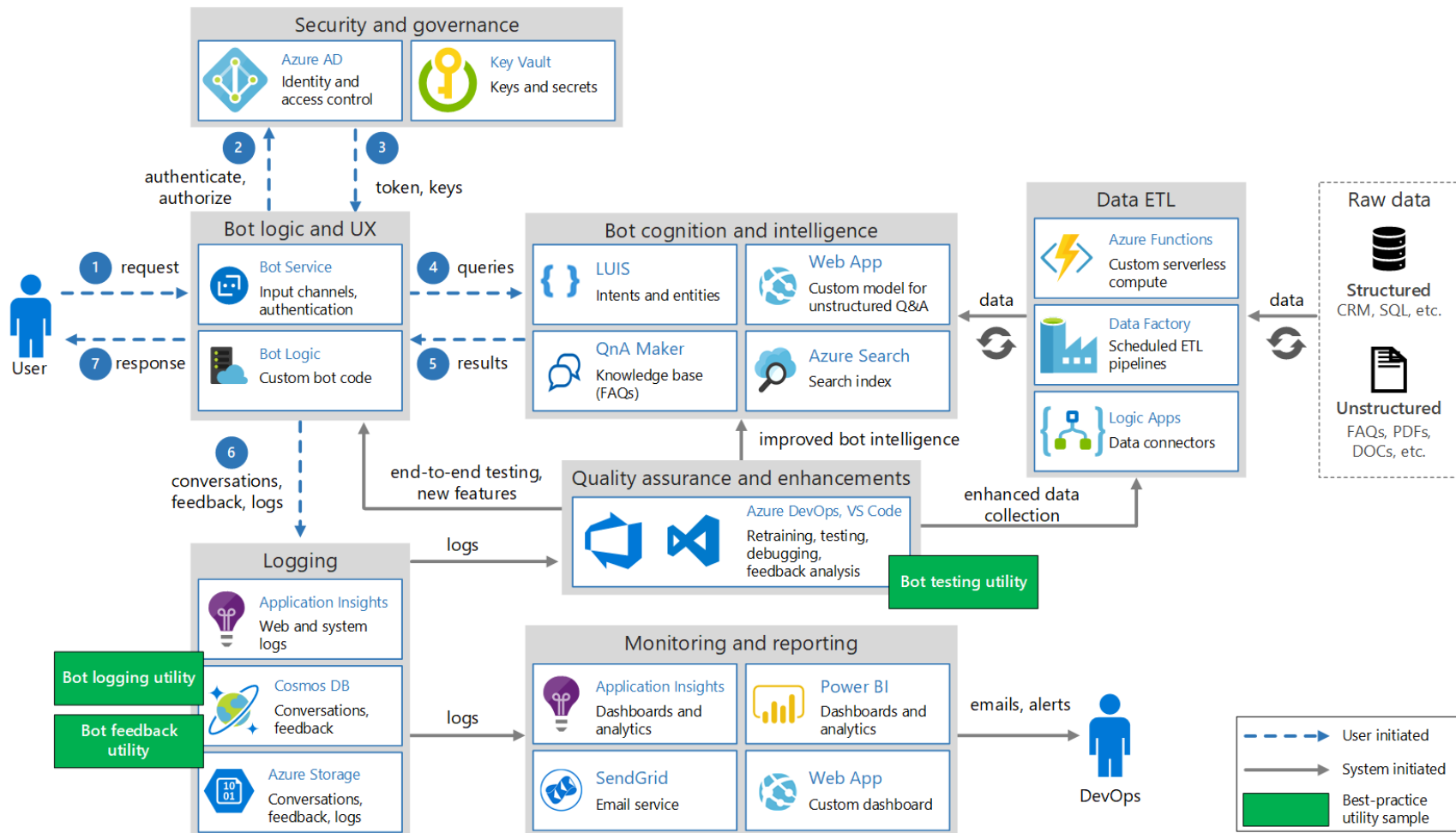
Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

정리함



AI의 일반적인 아키텍처

- Microsoft Azure 아키텍처 참조 : <https://docs.microsoft.com/ko-kr/azure/architecture/reference-architectures/ai/conversational-bot>



AI의 일반적인 아키텍처

복잡한 컴퓨팅 환경 : HPC 컴퓨팅 환경부터 이해

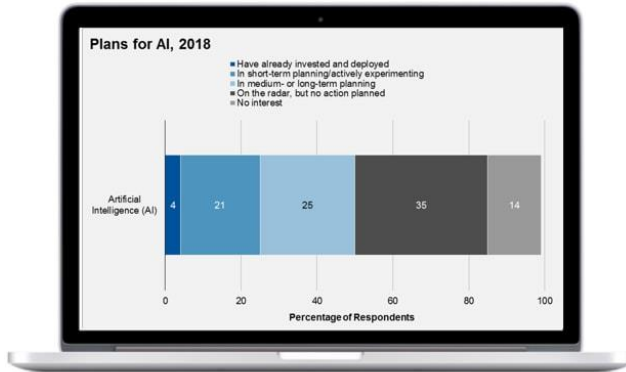
- 고성능 컴퓨팅 : High-Performance Computing = HPC
- 고급 연산 문제를 풀기 위하여 슈퍼 컴퓨터 및 컴퓨터 클러스터를 사용한 것
- 최신은 테라플롭스 이상의 규모를 HPC 컴퓨터로 구분하느 기준으로 사용

AI를 해야 하는데 고성능 컴퓨터

- 초기에는 고성능 컴퓨터에 의존하여 처리
- 지금은 GPU? CPU
- AI를 위한 칩셋의 개발 : NVIDIA
- IBM의 서버 : <https://www.ibm.com/kr-ko/it-infrastructure/solutions/ai>
- 핵심은 NVIDIA의 GPU

AI의 일반적인 아키텍처

- IBM 웹 사이트 참조하여 물리적인 구성



AI 서비스 앱1

AI 서비스 앱2

AI 서비스 앱3

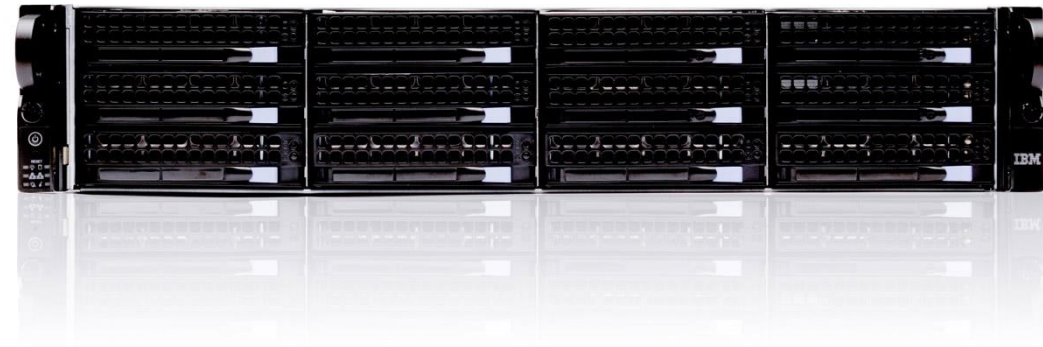
AI 서비스용 API

AI 학습용 서버

AI를 위한 트레이닝 서버
NVIDIA Tesla GPU 등

AI 데이터 저장소

많은 양의 데이터 저장소와 빠른 I/O



AI 의 일반적인 아키텍처

인프라를 도입하기 위한 고려사항

- 성능
- 총 소요비용
- 개발과 기간

AI 인프라 플랫폼 구축

- GPU / CPU / 네트워크 / 저장소 등의 하드웨어
- 운영체제와 머신 / 딥 러닝 관련 소프트웨어 스택
- 개발 환경과 개발 도구
- AI 인프라를 구축하기 위해서 최적의 조합을 찾기 위해 많은 경험과 테스트가 필요
- 컴퓨터 월드 기사 참조 : <http://www.comworld.co.kr/news/articleView.html?idxno=49483>

AI 의 일반적인 아키텍처

인프라 도입을 위한 시스템 구축 또는 클라우드 사용

- 인프라 플랫폼 구축을 위해 하드웨어부터 소프트웨어 까지 선택
- 최적의 플랫폼을 구축할 경우의 비용
- GPU의 사용과 X86 서버 아니면 RISC 계열 서버
- 운영체제는 AIX 인가 Linux 인가 Windows 인가
- 클라우드에는 이미 플랫폼은 구축되어 있다.

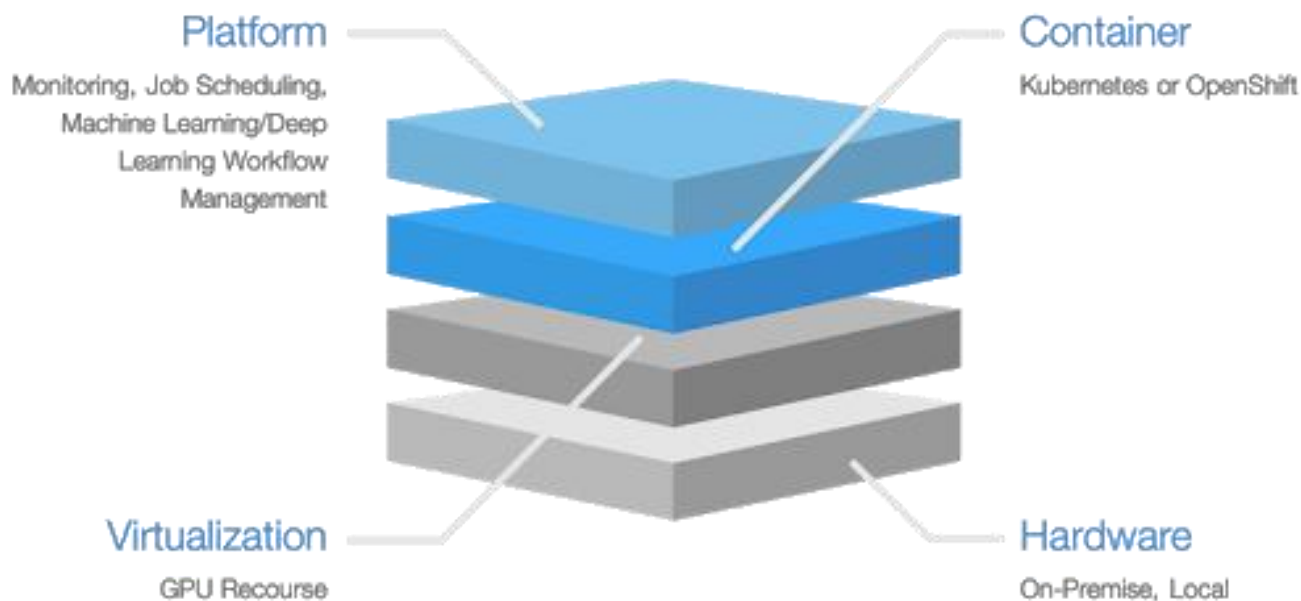
구축된 플랫폼에서 개발언어와 라이브러리만 선택하여 우선 개발

- 플랫폼 구축 경험이 부족할 경우 클라우드에서 시작하여 우선 서비스를 시작
- 내부 플랫폼은 클라우드에 있는 인프라를 하나씩 내부로 이동
- 하이브리드 클라우드 환경
- 마지막에는 내부 데이터 센터로 최종 이관

AI를 위한 인프라 서비스

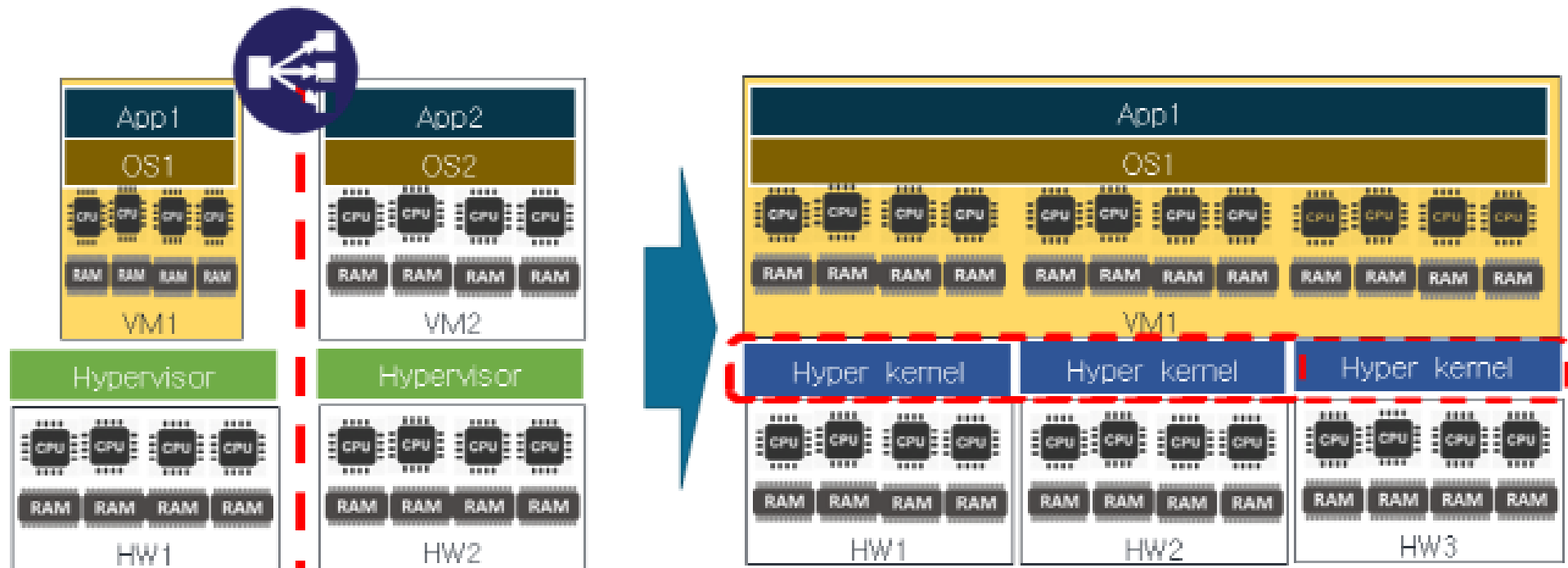
하드웨어 구성과 네트워크 인프라

- (주) 유클릭 웹 사이트 참조
- <https://blog.naver.com/PostView.nhn?blogId=uclick2016&logNo=221847941186&redirect=Dlog&widgetTypeCall=true&directAccess=false>



AI를 위한 인프라 서비스

- 컴퓨터월드 사이트 참조
- <http://www.comworld.co.kr/news/articleView.html?idxno=49407>



- VM1의 크기는 HW1의 사양을 넘지 못한다.
- HW2를 추가하더라도 VM1을 더 크게 만들지 못한다.
- VM은 하나의 HW위에서만 운영된다.

- HW2가 추가되면, 타이달 풀(Tidal Pool)이 확장된다.
- VM도 확장된다.
- HW3를 빼면, VM은 축소된다.











AI를 위한 인프라 서비스

- HP 웹 사이트 참조

HPC and AI Industry Solutions

Financial Services **Academia, Research, Gov't** **Life Sciences, Health** **EDA / CAE Manufacturing** **Oil and Gas, Energy** **Weather and Climate Research**

POINTNEXT Advisory, Professional and Operational Services – HPE Flexible Capacity for HPC, HPE Datacenter Care for Hyperscale

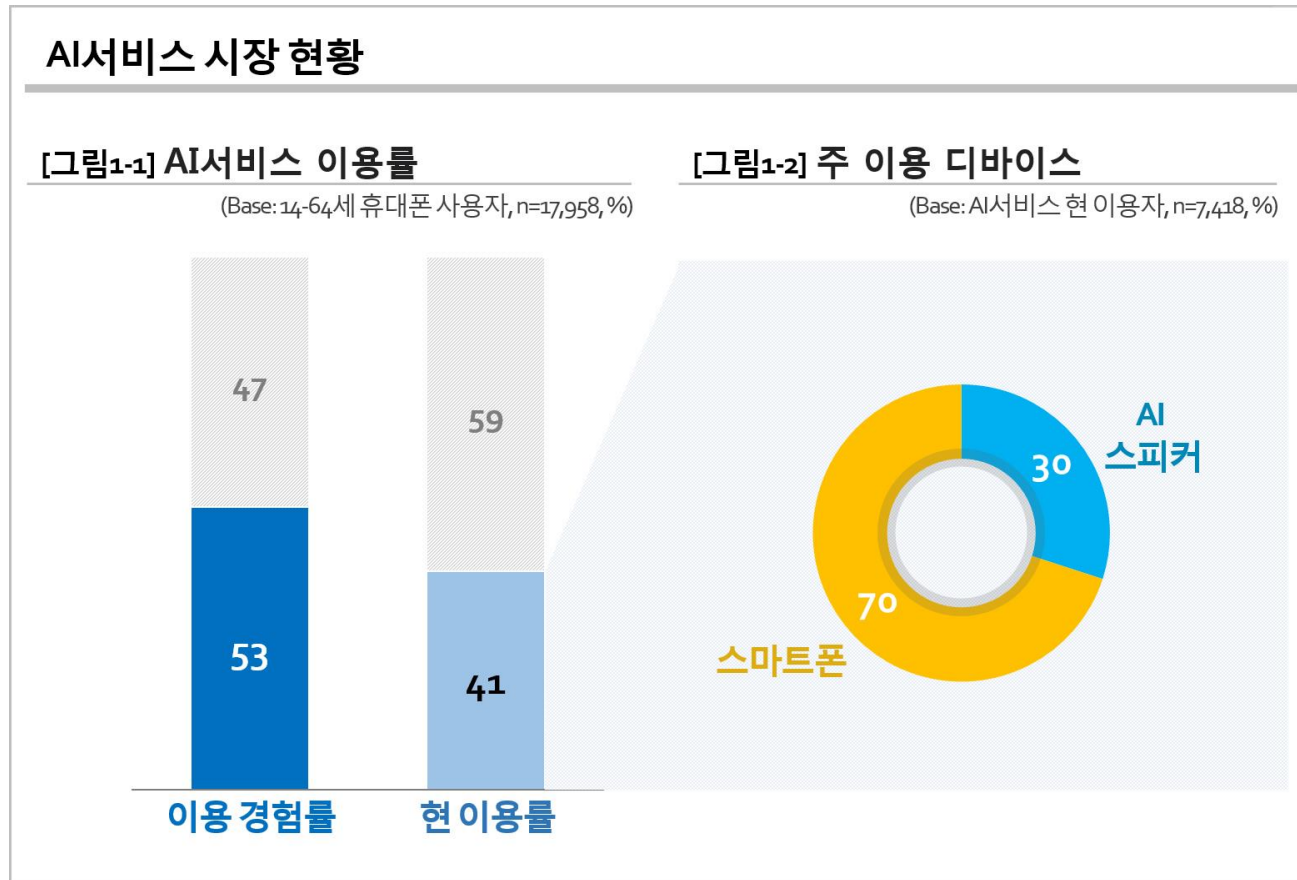
Supercomputing / Enterprise / Commercial HPC			High Performance Computing Software Portfolio		
<p>HPE SGI 8600</p>  <p>Liquid cooled, delivering industry leading performance, density, & efficiency</p>	<p>HPE Apollo 6000</p>  <p>Extreme Compute Performance in High Density</p>	<p>HPE Apollo 2000</p>  <p>The bridge to enterprise scale-out architecture</p>	<p>HPE Performance Software Suite</p> <ul style="list-style-type: none">- HPE Insight Cluster Management Utility- HPE SGI Management Suite- HPE Performance Software – Message Passing Interface		
Emerging HPC			HPC Storage	HPC In-memory Compute	Networking
<p>HPE Apollo 6500</p>  <p>Rack-scale GPU Computing</p>	<p>HPE Apollo sx40</p>  <p>HPC optimized industry standard servers supporting latest NVIDIA GPU technology</p>	<p>HPE Apollo k120</p>  <p>Intel Xeon Phi self hosted servers</p>	<p>HPE Apollo 4520</p>  <p>Additional Storage Options Available</p> <p>HPC Data Management Framework Software</p> <p>Large-scale, storage virtualization & tiered data management platform</p>	<p>HPE Superdome X</p>  <p>HPE MC990 X</p>  <p>Scale Up HPC</p>	<p>Arista 7500E Series</p>  <p>Low latency, high bandwidth</p>

클라우드를 이용한 AI 서비스 및 프로그래밍

Azure 클라우드 서비스 이용

AI를 이용한 서비스

- 컨슈머인사트의 제 28차 이통통신 기획 조사(2018년 10월 실시)
- https://www.consumerinsight.co.kr/voc_view.aspx?PageNo=1&id=ins02_list&no=2940&schFlag=0



Q: 아래 인공지능(AI) 브랜드 중 들어보셨거나 알고 계신 것을 모두 선택해 주십시오.
알고 계신 AI서비스 중에서 사용해 보신 것을 모두 선택해 주십시오./그럼, 현재 어떤 AI서비스를 이용하고 계십니까?

AI를 이용한 서비스

AI 서비스 : 일상 생활

- 스마트 폰에서 사용한 AI 비서 서비스
- 손목으로 이동
- 집에서 이용한 서비스

AI 서비스 : 기업에서

- 보안 관제 서비스 : 침입탐지 활동 (24시간 감시 가능)
- 빅 데이터 분석 후 다음 작업에 대한 방향성 제시
- 자율 주행 자동차 서비스
- 기업에서 제공하는 일상 생활 서비스 데이터와 연계(개인 금융 데이터 분석 후 활용)
- 인사 및 법률 서비스
- 의료 서비스

수고하셨습니다.

질문사항 : sirmd@outlook.com