

배포 즉시 보도 가능합니다.

대전(본원): 대외협력실 이종성 042-869-0976 / 이해준 0676 / 손영주 0997
문의: 슈퍼컴퓨팅인프라센터 홍태영 센터장(042-869-0667)

배포번호 : 2019-74

매수 : 보도자료 5매
(첨부자료 포함)

배포처 : 대외협력실

배포일자 : 2019.10.25.(금)

KISTI/GIST, 바람직한 고성능 AI 컴퓨팅 인프라 구축과 운용 논의를 위한 워크숍 공동개최

KISTI(한국과학기술정보연구원, 원장 최희운) 국가슈퍼컴퓨팅본부와 GIST(광주과학기술원, 총장 김기선) AI연구소가 2019년 10월 25일 금요일 서울 스카이뷰 섬유센터에서 국내환경에 알맞은 바람직한 고성능 AI 컴퓨팅 인프라의 구축과 운용에 관한 논의를 위한 워크숍을 공동으로 개최했다.

GIST AI연구소(소장 임혁)와 KISTI 국가슈퍼컴퓨팅본부(본부장 황순욱)가 공동으로 준비한 본 워크숍에서는 바람직한 AI 컴퓨팅 인프라 구축과 활용의 방향성, 효율적인 AI 컴퓨팅 인프라 운용과 활용 플랫폼, 차세대를 위한 AI 컴퓨팅 칩, 그리고 AI 컴퓨팅 장비 HW 및 SW 주제들에 관한 14인의 국내 컴퓨팅 연구자와 전문가들의 집중적인 발표와 논의가 진행됐다.

인간이 관여하는 모든 활동 영역으로 급속히 파고드는 AI를 유연하게 지원하기 위해서는 엔진 역할을 하는 고성능 컴퓨터에 질 좋은 방대한 데이터 연료를 상황에 맞춰 수시로 투입하면서 최적의 출력으로 운영해야 한다. openAI 자료*에 따르면

알파고처럼 AI 기술 혁신을 선도하는 연구의 학습 소요 시간(계산량)은 3.5개월마다 2배씩 증가하는 경향을 보이고 있다.

* <https://openai.com/blog/ai-and-compute/>

이에 따라 전 세계적으로 보다 큰 AI 컴퓨팅 엔진과 데이터 저장소를 설비하는 경쟁이 벌어지고 있으며, 현재 추진되는 대부분의 사례에서 AI 특화 또는 HPC/AI 겸용 방식으로 국가 차원의 AI 컴퓨팅 인프라를 구축하는 추세이다.

일례로 일본은 2018년 5월(ABCI AI 전용 슈퍼컴퓨터), 대만도 2018년 후반(Taiwania-2 HPC/AI 겸용 슈퍼컴퓨터)부터 국가적인 AI R&D를 직접 지원하는 컴퓨팅 인프라를 투입하여 활용하고 있는 상태다.

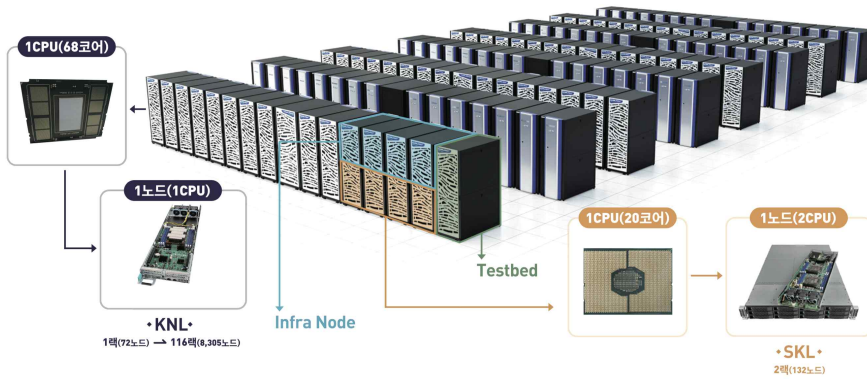
또한 2020~22년 미국, 중국, 유럽의 차세대 슈퍼컴퓨팅 프로젝트로 개발 중인 엑사 스케일급 슈퍼컴퓨터들은 모두 HPC/AI를 겸용으로 활용할 수 있는 형태로 구축될 예정이다.

국내에서는 2018년 11월에 개통된 CPU 기반의 KISTI 슈퍼컴퓨터 5호기 ‘누리온’ 과 GPU 기반 시스템 ‘뉴론’ 을 활용한 AI 컴퓨팅 지원에 덧붙여서 국내 산학연 AI에 특화된 GPU 기반의 초고성능 컴퓨팅을 지원하기 위해 ‘광주 AI산업융합집적단지’ 사업의 일환으로 AI 클라우드 데이터센터 구축을 추진하는 사업이 준비 중인 것으로 알려져 있다.

KISTI 황순욱 본부장은 본 행사의 취지에 대해 “지능정보사회에 진입하는 데 있어 핵심적인 역할을 수행하는 인공지능(AI)의 기술 혁신을 더욱 가속하기 위해서는, 고성능컴퓨팅을 비롯한 컴퓨팅 인프라, 관련 하드웨어 및 플랫폼 기술에 대한

최신 현황 공유 및 효율적인 활용 방안에 대한 고민이 지속적으로 필요하며, 본 행사를 통해 산학연 각계의 전문가들이 모여 함께 논의할 기회를 갖는다는 점에서 의의를 찾을 수 있다” 고 설명했다.

첨부 1.



주요 장치 및 구성도

메인시스템	스토리지	테이프 라이브러리	기반시설
<ul style="list-style-type: none"> 계산노드 8,305대 CPU-only노드 132대 	<ul style="list-style-type: none"> 12랙 (총 23.88PB) 8TB NL-SAS 3,200개 1.2TB SSD 768개 	<ul style="list-style-type: none"> 4랙 (총10PB) LTO7 미디어 1,700개 	<ul style="list-style-type: none"> 냉동기 400RT 4EA 프리쿨링 250RT 2EA 냉각탑 450CRT 2EA 향온환습기 30RT 11EA

<그림 1> KISTI 슈퍼컴퓨터 5호기 ‘뉴리온’

뉴론 NEURON

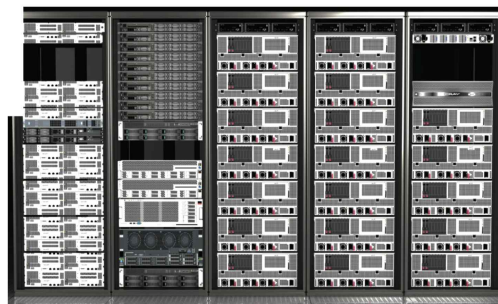
뉴럴 네트워크(신경망)에서 신경세포를 나타내는 뉴런을 의미함과 동시에 ‘뉴(NEW, 새로운)+라온(은 우리말, 전부의)’ 새로운 아키텍처를 지속적으로 확장하여 사용자에게 제공한다는 의미

Tesla V100

- DOUBLE-PRECISION
7.5 TeraFLOPS
- SINGLE-PRECISION
15 TeraFLOPS
- DEEP LEARNING
120 TeraFLOPS



구분	Lenovo nx360	HP Apollo 6500
운영체제	CentOS 7.4(Linux, 64-bit)	
GPU 종류	Tesla K40	Tesla V100
	HBM2 16GB	HBM2 32GB
노드당 GPU수	2	1
노드 수	8	15
GPU	Intel Xeon Ivy Bridge(E5-2670) /2.5GHz (10-core)/2 Socket	Intel Xeon skylake(Gold6230) /2.1GHz (20-core)/2 Socket
메인 메모리	노드당 128GB DDR3	노드당 384GB DDR4



신규 사용자
회원가입
<http://neuron.kisti.ac.kr>
회원 가입

기본 사용자
기밀정보(연구비) 접근
이전 시, 뉴론 계정으로
Newton 사용 가능

가속기 기반 시스템
AI, 빅데이터, HPC 등
다양한 분야의 주요 대용

로그인
뉴론 계정 정보 로그인
neuron.kisti.ac.kr

<그림 2> KISTI 슈퍼컴퓨터 5호기 보조 시스템 ‘뉴론’

첨부 2.



[인사말을 하는 KISTI 국가슈퍼컴퓨팅본부 황순욱 본부장]



[AI컴퓨팅 인프라에 대해 발표하는 GIST AI대학원 김종원 교수]