

규 격 서

품명 및 규격 Description	단위 Unit	수량 Quantity
AI 학습용 GPU 서버	set	3

◇ 개요

- 대용량 스토리지에 분석이 필요한 자료들을 저장하고 분석할 수 있는 안정적인 운용이 가능
- 리눅스 병렬 프로그램을 구동하여 병렬 계산 목적으로 모델 계산에 활용이 가능
- 고성능 GPU 사용이 가능 병렬 처리로 활용 가능

◇ 특징

- System의 냉각 방식은 수냉 쿨링 방식
- 자체 설계된 도면도(구성도) 포함
- GPU를 병렬로 연결하여 Deep Learning의 연산 처리가 가능
- GPU가 PCI-E 3.0 16배속으로 동작 되어야 하며, GPU는 별도 라이저 카드 없이 마더보드의 PCIe 슬롯에 바로 장착
- 서버당 GPU Mem는 Total 96GB이상이어야 하며 GPU 구동 시 Throttling 없이 365일(1일 24시간) 이상 무정지 상태로 정상 작동
- GPU Burnnig 코드 구동 시에 GPU Throttling이 없음
- NVIDIA GPU Mem는 Total 96GB이상 장착하고 Deep-learning Full load 운전 시 Nvidia-Smi로 측정하였을 때 각각의 GPU 온도가 섭씨 60°C 미만으로 7일 이상 유지
- Workstation 규격으로 CPU는 Total 20core 이상, 40Thread 이상 지원하고, RAM은 Total 512GB 이상, Expansion slot은 PCI-E 3.0 16배속으로 4ea 이상, SYSTEM FAN은 23ea 이상으로 구성

◇ 상세 규격

- CPU는 Dual로 구성되며 Total 20 Core, 40 Thread 이상
- 각 CPU당 PCI Express Revision은 3.0, Max number of PCI Express Lanes 48이상 지원
- RAM의 총용량은 512GB 이상
- 메인보드는 메모리는 표준 DDR 모듈을 기준으로 12개 이상, Max 768GB 이상을 지원
- 저장장치는 SSD Capacity 2TB이고 규격은 Interface PCIe 3.0 4배속(32GT/s), TLC, Read 600K IOPS, Write 550K IOPS, MTBF 1.5M hour, TBW 600TB, DRAM이며 수량은 1ea를 필요
- Lane에 대한 정확한 이해를 바탕으로 병목현상 없이 최적화된 셋팅을 지원
- 네트워크는 Intel series이고 접속 단자는 RJ45 2port 이상
- 파워서플라이의 공급전력은 2050W이상 이고 한 개 이상의 제품의 에너지 효율은 80 Plus 이상의 Titanium Level이어야 한다.

No	상 세 규 격		수 량
1	품명	AI 학습용 GPU 서버	1 set
	CPU	Dual CPU, Total 20 Core, 40Thread 이상	
	GPU	Total GPU Mem 96GB 이상	
	RAM	Total Mem 512GB 이상	
	SSD	저장장치는 SSD Capacity 2TB이고 규격은 Interface PCIe 3.0 4배속(32GT/s), TLC, Read 600K IOPS, Write 550K IOPS, MTBF 1.5M hour, TBW 600TB, DRAM+SSD 4TB	
	POWER	Dual parallel Power Supply 2030W이상	
	O/S	Linux Ubuntu and Window dual Install or Cent OS + KVM 등 (환경에 맞는 커스텀 환경 제공 필요)	
	Cooling System	BARO Custom Made Special Water Cooling System (GPU and CPU 전부 커스텀 수냉 방식으로 진행 필요, Flow 14이상)	

◇ A/S 기간등 기타사항

- 하자보증조건

- ▶ 하드웨어와 소프트웨어의 무상하자보증기간은 검사완료일로부터 3년으로 함
- ▶ 하드웨어와 소프트웨어에 대해 장애 발생 시 계약대상자는 24시간 연중무휴로 A/S에 응대하여야 하며 익일 이내의 현장방문으로 장애 해결작업을 시작해야 함
- ▶ Warranty Service기간 내 분기별 1회 방문 After Service를 반드시 행해야함

- 설치 및 검수 조건

- ▶ 설치현장을 확인하고 지정하는 시간 및 장소에 납품이 가능해야 하며 정상 작동을 위해 요구되는 네트워크와 전력 연결환경을 구축
- ▶ 운영을 위한 구축 및 운영체제와 소프트웨어를 설치 후 이상이 없음을 확인할 수 있는 소프트웨어 테스트를 진행

- 교육조건

- ▶ AI & Future Technology 강좌
- ▶ 딥러닝 강좌(딥러닝 이론 및 실습)
- ▶ Docker 사용법 교육 (1일 3시간 / 총 8시간)
- ▶ CUDA 사용법 교육 (1일 2시간 / 총 6시간)
- ▶ Cloud 구조 및 사용법 교육 (1일 2시간 / 총 20시간)
- ▶ Workstation 사용법 교육 (1일 1시간 / 총 10시간)

- 소프트웨어 설치 조건

- ▶ 운영체제는 Ubuntu 18.04 LTS를 설치. (필요에 따라 듀얼부팅 또는 다른 리눅스 운영체제를 설치 및 지원이 가능해야함.)
- ▶ Nvidia drivers (latest Cuda-compatible drivers), CUDA, cuDNN, NCCCL, Docker, Nvidia-dock)
- ▶ Nvidia CUDA repository, Nvidia machine learning repository, Docker CE repository
- ▶ 차후 CLOUD를 이용하여 Scale up, Scale out 가능해야 함
- ▶ 복구솔루션은 납품 시 최초 설정된 운영체제 및 딥러닝 소프트웨어들을 재설치할 경우 1시간 이내 복구할 수 있어야 하며 납품 후 3년 동안 요청할 경우 2영업일 내에 방문하여 복구해야 함
- ▶ S/W Customizing 기술 지원이 가능해야 함
- ▶ 시스템 보안설정 (계정, 파일, 접근권한 등)
- ▶ 시스템 자동화 구축 지원

- 기타조건

- ▶ Deep Learning, Cloud 솔루션의 구축경력이 5년 이상인 엔지니어를 구성원으로 보유하여야 하며 요청 시 즉각적인 방문 지원이 가능해야 함
- ▶ 기업부설 연구소 설립이 되어있는 곳의 기술이 바탕이 되어야 하며 회사 내부에서 모든 문제에 대해 처리가 가능해야 함
- ▶ 입찰에 참여하는 업체는 제조업체로 제한함