

2020 IT 21

Global Conference

Digital New Deal
Technology Essentials
디지털 뉴딜 기술 핵심

Session 6-6

HPC Cloud 서비스 기술 및 동향

정기문 책임 (KISTI)



[요약문]

국내 최대 규모의 슈퍼컴퓨팅 서비스를 제공하는 한국과학기술정보연구원은 슈퍼컴퓨터 “누리온”을 클라우드 환경에서 이용할 수 있는 “KI(KISTI Intelligent) Cloud” 서비스를 제공 중이다. 슈퍼컴퓨터 사용자라면 누구나 사용할 수 있는 KI Cloud는 웹기반으로 손쉽게 접근하여 사용할 수 있으며 사용자 맞춤형 가상 서버, 클라우드 스토리지 서비스 등 오픈스택 기반의 IaaS 환경을 제공한다. 또한 컨테이너 기반의 가상 클러스터 환경에서 대화형 프로그래밍을 수행할 수 있는 Jupyter Notebook 서비스, R 기반 데이터 분석 처리를 지원하는 Rstudio 서비스 등 연구자들이 웹을 통해 실시간으로 이용할 수 있는 다양한 데이터 분석 도구 및 컴퓨팅 환경을 제공한다. 기존의 슈퍼컴퓨팅 사용자라면 고전적인 터미널 기반의 작업 스케줄러 인터페이스를 통해 배치 작업 형태로 슈퍼컴퓨터를 이용해야만 했지만 이제는 다양한 대화형 클라우드 서비스를 통해 연구자가 원하는 컴퓨팅 환경을 쉽게 구성할 수 있게 되었다. 본 발표에서는 KI Cloud의 구조 및 기능 등을 소개하고 HPC 클라우드의 새로운 방향을 제시하고자 한다.

[발표자 약력]

1999년 2월 : 전남대학교 전산학과 학사
2001년 8월 : 전남대학교 전산통계학과 석사
2009년 8월 : 전남대학교 정보보안협동과정 이학박사
2001년 7월~2004년 12월 : 한국정보보호진흥원
2004년 12월~2005년 6월 : 국가사이버안전센터
2005년 6월~현재 : 한국과학기술정보연구원 책임연구원
관심분야 : 클라우드 컴퓨팅, Openstack, Kubernetes, 클라우드 보안, 네트워크 보안, 딥러닝





목 차

1. HPC 소개
2. HPC Cloud 동향
3. KISTI HPC Cloud





HPC (High Performance Computing) 소개



☑ 초대용량 정보를 초고속으로 생산, 처리 활용하게 하는 컴퓨팅 시스템(상대적 개념)

☑ 슈퍼컴퓨터는 통상적으로 세계 성능 순위 500위 내의 컴퓨터를 의미함

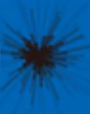
HPC의 정의

한국정보통신기술협회 (TTA)	<ul style="list-style-type: none"> 대용량의 정보를 고속으로 처리 고도로 복잡한 문제 해결, 비즈니스 분석 작업 수행, 연산 집약형 작업 부하 처리 등 복잡한 작업을 수행 	
타 기관의 정의	Gartner	<ul style="list-style-type: none"> 클러스터는 다량의 서버를 포함할 수 있으며 전용통신을 위한 조건을 가지고 있고, 솔루션 유형에 따라 다양한 서버 디자인을 조합, 클러스터링을 할 수 있음
	IBM	<ul style="list-style-type: none"> 복잡한 문제점을 해결하기 위해 슈퍼컴퓨터 또는 컴퓨터 클러스터와 같은 대용량 컴퓨팅 성능이 필요한 기술
	NIST	<ul style="list-style-type: none"> 컴퓨터 자원의 용량과 역량을 초과하는 어려운 문제를 풀며, 매우 복잡한 문제 혹은 컴퓨터 자원이 다량 필요한 어려운 문제를 해결함
도입 효과	고속 연산 지원	복잡한 연산 지원

HPC 구성 요소

수요 영역	산업계	과학기술계
컴퓨팅 서비스 유형	High Performance Computing	High Throughput Computing
주요 특징	병렬 분산 처리 시스템	
	고속연산능력	시간 의존성 문제
	복잡한 데이터 처리	빅데이터
	모델링, 시뮬레이션, 렌더링	

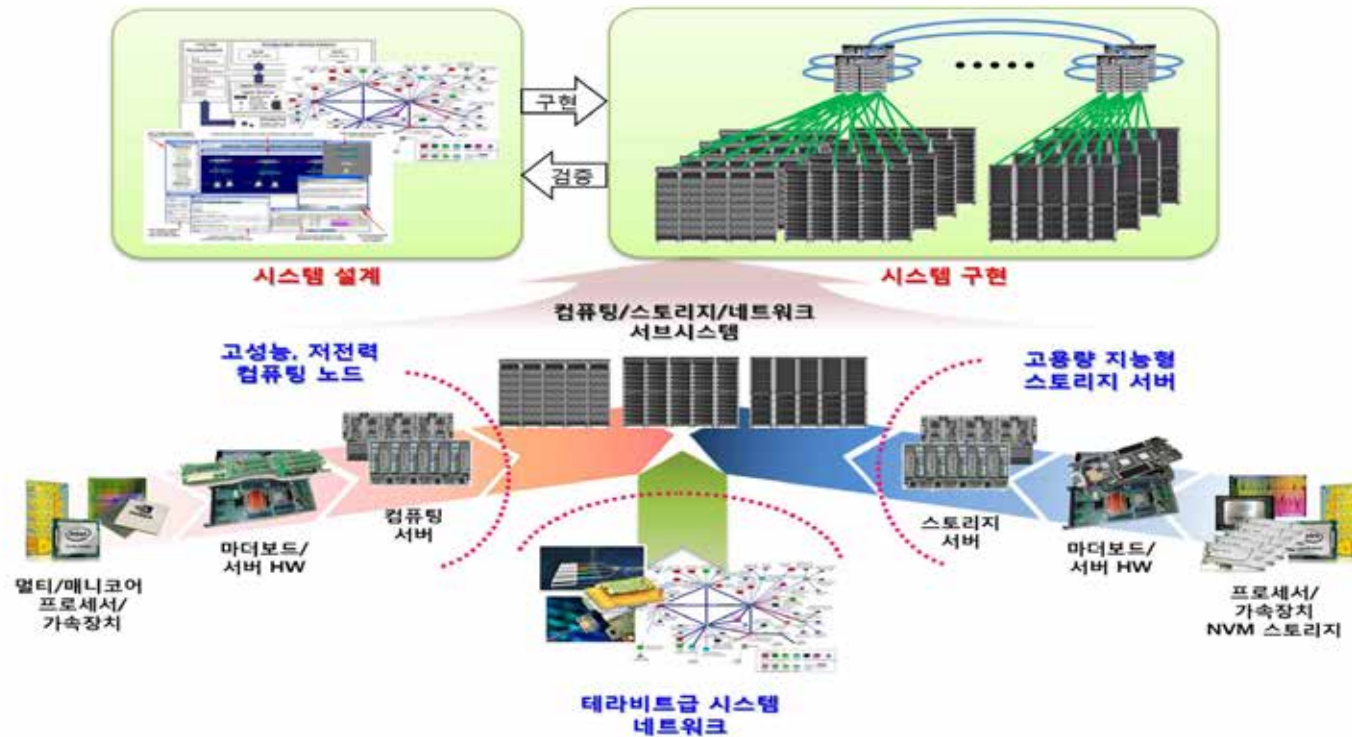
변화하고 있는 고성능 컴퓨팅 수요에 대응하여
영역별 사용자의 니즈를 반영한 서비스 강화 필요

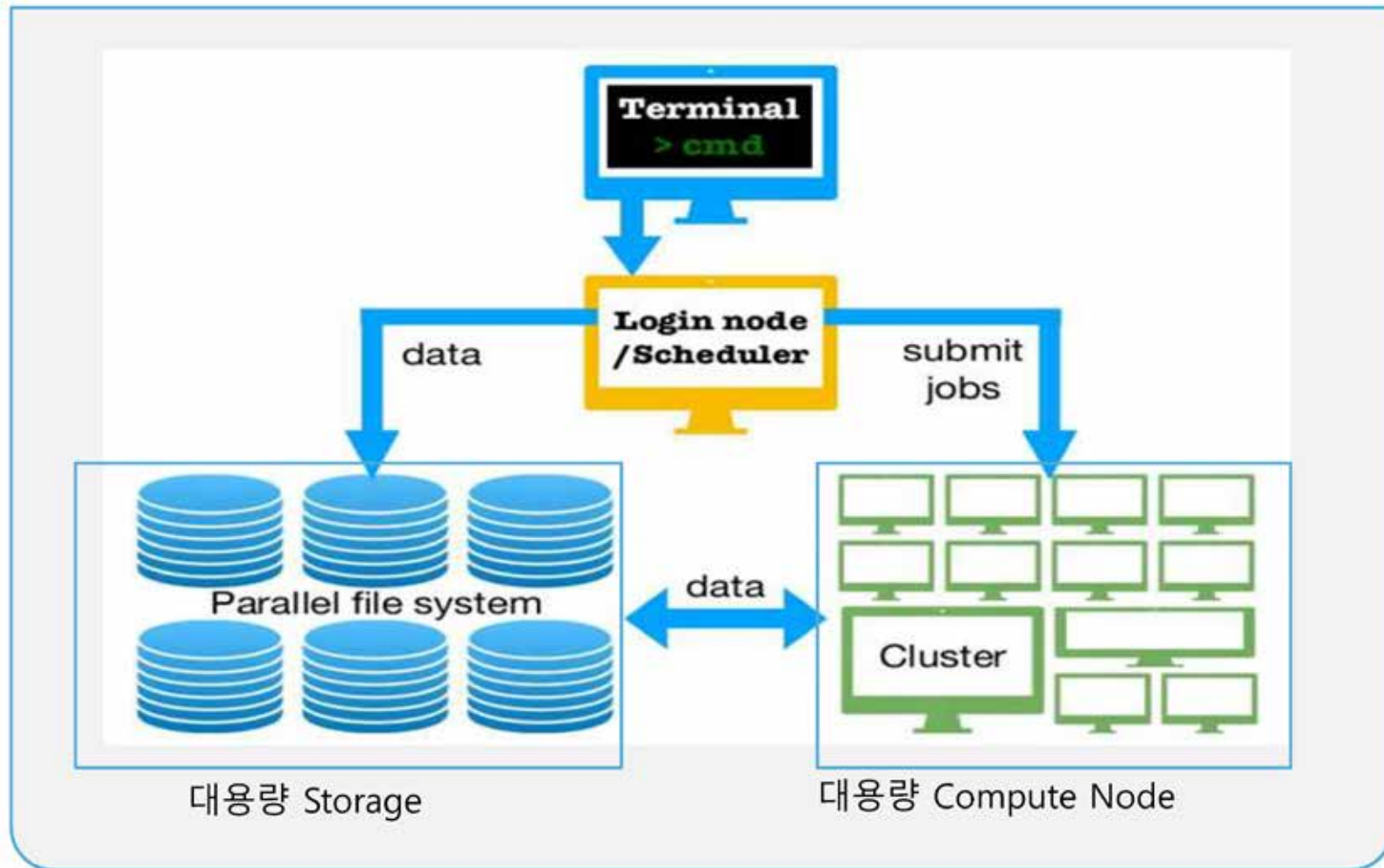


- ☑ HPC의 수요 영역은 산업계 및 과학기술계로 구분할 수 있음
- ☑ 영역별 필요 서비스 등의 다양한 사용자 니즈를 반영한 서비스 제공이 요구됨

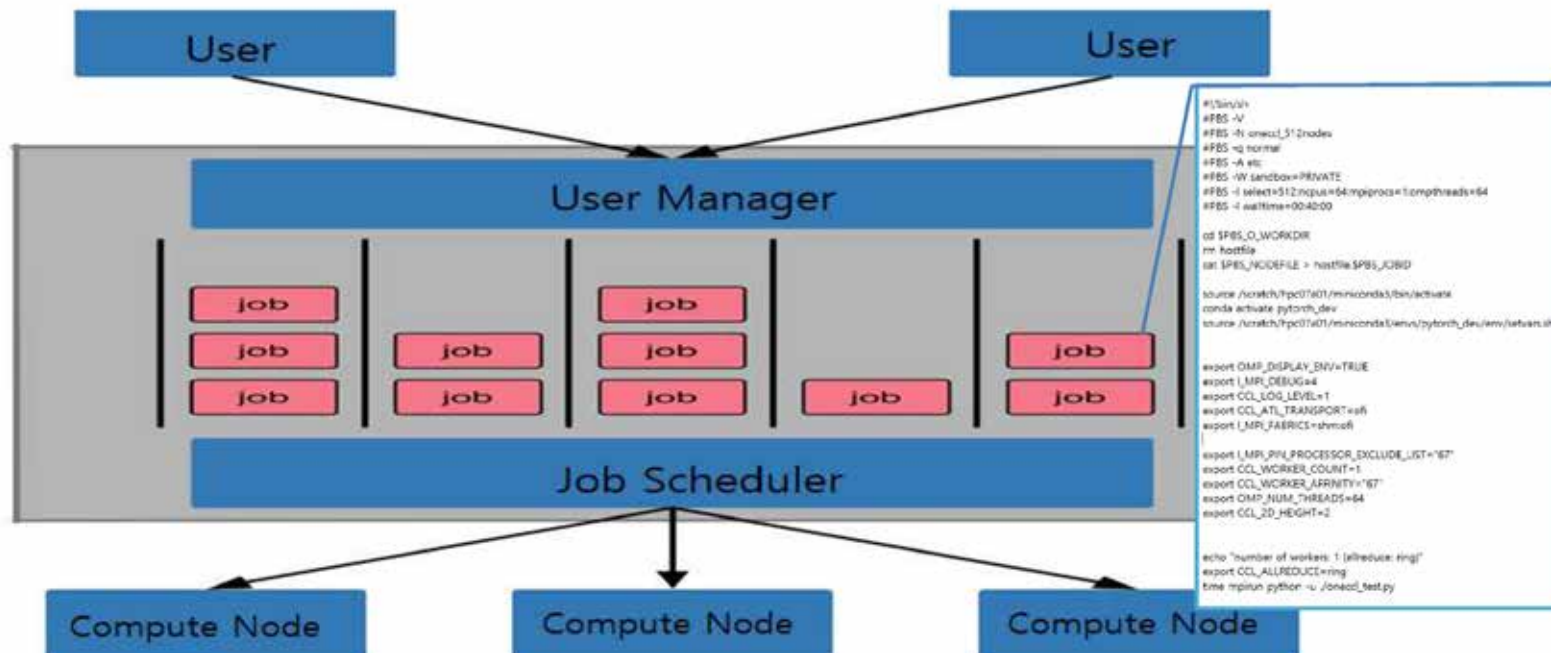
HPC 수요 영역			
	수요 영역별 필요 서비스		주요 내용
산업계	M&S 기술 적용 방법 연구	데이터 기반 기술사업화 지원	<ul style="list-style-type: none">▪ M&S 기술 적용 방법 개발 및 활용 기반 제품 설계를 통한 아이디어 사업화/기술사업화 지원▪ 다양한 산업별 니즈, 수요, 시장 변화 분석 기반 신수요 발굴 및 신규 비즈니스 모델 연구 지원
	M&S 기반 제품 설계	신규 비즈니스 모델 연구	
	데이터 분석 기반 수요 발굴	산업·시장분석 모델 지능화	
	아이디어 사업화 지원	기술사업화 시스템 고도화	
	⋮	⋮	
과학기술계	데이터 중심 협업 연구망 구축	연구 데이터 수집 및 공유	<ul style="list-style-type: none">▪ 논문 및 대용량 연구 데이터의 전송, 처리, 공유 및 아카이빙 지원▪ 오픈 데이터 융합 분석이 가능한 데이터 중심 협업 연구망 구축 및 연구 지원▪ 지능형 빅데이터 분석 및 연구 결과의 시각화·렌더링 지원
	첨단 연구 대상 네트워크 구축	오픈데이터 융합 분석	
	지능형 빅데이터 분석·처리	R&D 논문 아카이빙	
	연구 결과 시각화·렌더링	대용량 데이터 전송	
	⋮	⋮	
다양한 HPC 수요 영역별 사용자의 니즈를 반영한 KISTI HPC 서비스 도출 필요			

- ☑ 초고성능 하드웨어(고성능 CPU, Memory, Accelerator, 고속 Network, 대용량 스토리지 등) 구축
- ☑ 시스템 설계, 클러스터 구축 기술, Parallel Program 소프트웨어 등





- ☑ 다수의 사용자는 Login Node에서 스케줄러에 작업(Job)을 제출하고 순차적으로 Compute Node 활용
- ☑ 사용자는 작업을 임의로 처리할 수 없어 정책에 따라 순차적으로만 사용할 수 있고 점유하여 사용할 수 없음
- ☑ 모든 Compute Node는 동일한 OS 및 소프트웨어가 설치되어 활용됨

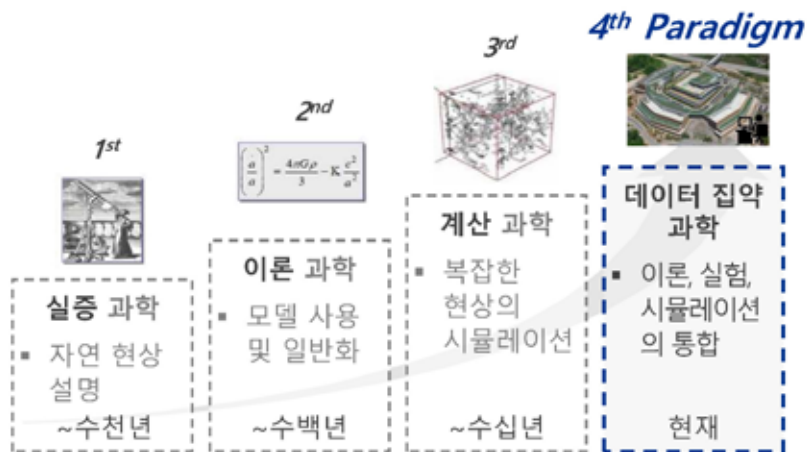


- ☑ 연구 패러다임의 변화로 계산과학 및 HPC 자원에 대한 중요성은 증대되고 있음
- ☑ 전통적인 HPC 사용환경의 개선을 통해서 변화하는 수요에 대응하고자 하는 필요성 커지고 있음

데이터 집약 과학으로의 연구 패러다임 변화

The Fourth Paradigm: Data-intensive Scientific Discovery

-Jim Gray



HPC 기반 연구영역의 확산

	이론적 계산 가능	이론적 계산 불가능
실험 가능	이론과 실험이 모두 가능하여 상호보완적 관계의 전통적 영역	<ul style="list-style-type: none"> ✓ 실험은 가능하나 이론적 계산이 어려운 영역 - 예: 전자구조 계산, 단백질 접힘 → 대규모 계산 실험
실험 불가능	<ul style="list-style-type: none"> ✓ 이론적 계산만이 가능한 영역 - 예: 빅뱅, 블랙홀 연구 → 시뮬레이션 	<ul style="list-style-type: none"> ✓ 이론 및 실험적 접근이 어려운 영역 - 예: 은하계 생성, 금융물리 → 인공지능, 빅데이터

연구 패러다임이 데이터 집약 과학으로 변화됨에 따라
초고속 연산의 중요성이 강화되고 있어, 과학기술계의 HPC 중요성은 증대되는 추세



HPC Cloud 동향



☑ 전통적인 HPC Computing 환경에서 사용자의 편의 요구사항 증가

☑ 사용자 중심의 컴퓨팅 환경에 대한 요구사항 증가

Flexible Computing

```

/apps/Modules/modulefiles/compilers
-----
cce/8.6.3(default) gcc/8.3.0 intel/19.0.5
gcc/8.5.0 intel/17.0.5 pgc/18.10
gcc/7.2.0(default) intel/18.0.3(default) pgc/18.1

/apps/Modules/modulefiles/languages
-----
impi/17.0.5 impi/19.0.5
impi/18.0.3(default) impi/19.1.0

/apps/Modules/modulefiles/libraries_using_mpi
-----
fftw_mpi/2.3.5 fftw-parallel/1.10.2 parallel-netcdf/1.10.0
fftw_mpi/3.3.7 netcdf-netcdf4/4.6.1 psm2/2.3.1

/apps/Modules/modulefiles/libraries
-----
OpenCL/2.2.0 hpccl/1.10.2 nccl/2.5.0 netcdf/4.6.1
netcdf/4.3.10 lapack/3.7.0 mkl/2.1.4 mkl-rt/2.1.7

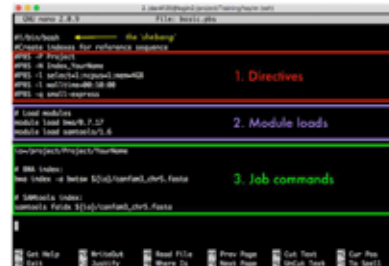
/apps/Modules/modulefiles/applications
-----
conda/ptorch_1.0 gromacs/2018.6 namd/2.10 R/3.5.0
conda/tensorflow_1.13 gromacs/2018.6 PETSc/3.9.4 R/3.6.2
cp2k/5.1.0 gromacs/5.4.6 python/2.7.13 scikit-learn/0.20
ferret/7.4.3 lammps/3.10.0 python/3.7 scikit-learn/0.20
ferret/8.1.2 lammps/3.10.0 qt/5.12 singular/17.0.2
gromacs/5.4.6 namd/2.10 qt/5.12 singular/17.0.2

```

- Ubuntu 환경에서 개발해 왔는데, HPC는 Centos 환경
- HPC에 설치된 패키지 외에 추가적인 패키지 설치 필요
- 준비된 HPC 환경에 맞게 설치 및 포팅

“사용자 요구사항에 맞게 자원을 제공하고 OS를 제공할 수 있는 HPC 환경 필요”

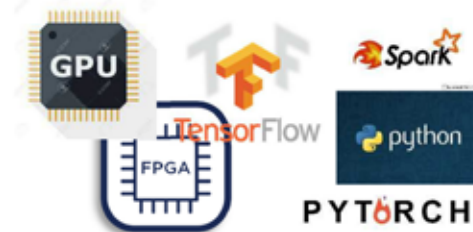
Interactive Computing



- 텍스트 기반 작업 처리
- CLI로 한정된 UI 환경
- Job을 scheduler에 실행한 이후 작업 완료시까지 대기하는 Batch 형태

“사용자 진화적인 환경에서 실행한 결과를 확인하면서 프로그래밍할 수 있는 HPC 환경 필요”

New Computing



- 기존 HPC에 적합한 소프트웨어만 주로 사용 가능한 환경
- 새로운 가속기 디바이스, 데이터분석, 머신러닝, 딥러닝 등을 수행할 수 있는 환경 필요

“4차산업혁명 시대에 새롭게 등장하는 컴퓨팅 디바이스 및 프로그래밍 지원한 HPC 환경 필요”

HPC on Cloud Computing

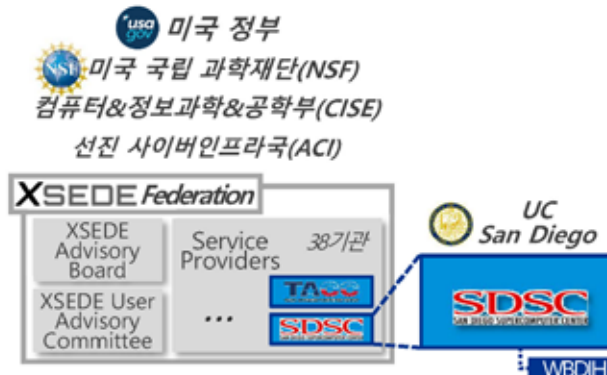
- ☑ SDSC(San Diego Supercomputer Center)는 HPC 기반의 데이터 집약 컴퓨팅 등 Big Data 관련 기능을 중심으로 산학연 및 정부를 포함한 국가 커뮤니티의 지식자원 서비스 활용을 지원함



개요

기관명	SDSC (샌디에고 슈퍼컴퓨터 센터)
설립	<ul style="list-style-type: none"> 1985년, 미국 샌디에고, USD 170M 예산 NPAC 소속 40개의 대학과 파트너
역할	데이터 집약 컴퓨팅과 HPC 기반의 스토리지, 모델링, 데이터 마이닝 및 예측 분석 등의 Big Data 관련 기능을 중심으로 하여, 학계, 산업계 및 정부를 포함한 국가 연구 커뮤니티에 자원, 서비스 및 전문지식 제공

구조



Cloud 서비스

- Openstack Swift 플랫폼을 통해 Cloud Computing 및 Storage 서비스를 제공

구분	세부기능	세부내용
IaaS	Storage	최소 할당량 1TB
	Compute	CPU 1개와 4GB메모리의 가장 기본적인 인스턴스부터 CPU 8개와 64GB메모리의 최대 인스턴스까지 10단계로 구분하여 제공



Cloud Storage & Compute

What is SDSC Cloud Storage and Compute?

The SDSC Research Data Service (RDS) team administers one of the first large-scale academic deployments of cloud storage in the world. UC San Diego campus users, members of the UC community, and UC affiliates are eligible to join the hundreds of users who already benefit from the 3 penalties of tape space, which is organized into object-based cloud storage by OpenStack's Swift platform.

SDSC Cloud Storage is the perfect storage choice for researchers with fixed budgets because unlike other cloud providers, SDSC features a simplified recharge plan that eliminates secondary fees such as bandwidth costs, charges assessed per request, and regional migration fees.

Advantages of SDSC Cloud include:

- Collaboration** - effortlessly share cloud data with colleagues or the public
- Redundancy** - with three copies of each file distributed across multiple internal racks, rest assured that your data is immediately safe from the perils of normal hardware exhaustion
- Technical Support** - communicate directly with SDSC Cloud Storage engineers, who enjoy using their extensive experience to assist you

In addition to storage, SDSC Cloud now includes compute options. Users can spin up instances on demand. Email sdsc@ucsd.edu for more information.

Note: SDSC offers a separate cloud storage system for data that needs to meet FERPA or HIPAA compliance. Inquire sdsc@ucsd.edu for secure solutions.

How Much Does SDSC Cloud Storage and Compute Cost?

Service	Unit	UC Rate	External Rate
Cloud Storage Storage allocation: 1 TB, billed monthly storage cost is averaged by institution's total storage usage	TB/month	\$10.15/TB/month	\$19.50/TB/month
Cloud Compute 100% regional resource usage per hour	hour	\$0.0001	\$0.0001

Recharge rates are subject to change with approval from the UC San Diego Recharge Rate Committee. UC is waived for UC San Diego users of SDSC Cloud, effective July 2017. Please see <https://cloud.sdsc.edu/technology/cloud/understand-how-it-works> for more details.



☑ TACC(Texas Advanced Computing Center)는 HPC 기반의 사이버인프라 활용 생태계 구축을 통해 국가 과학 및 사회 발전에 기여하고자 하며, HPC 및 응용 SW 개발을 비롯한 다양한 연구를 지원함



개요

기관명	TACC (텍사스 첨단 컴퓨팅 센터)
설립	<ul style="list-style-type: none"> 2001년, 미국 텍사스 7개의 컴퓨터 기관이 참여
역할	텍사스와 미국 내 교육, 산업, 연구기관 등을 대상으로 HPC 서비스, 시각화, 데이터분석, 스토리지, 클라우드 등을 포괄하는 사이버 인프라 생태계를 구축하며 매년 3,000개 이상의 프로젝트 진행
구조	<p>미국 정부 미국 국립 과학재단(NSF) 컴퓨터&정보과학&공학부(CISE) 선진 사이버인프라국(ACI)</p> <p>XSEDE Federation XSEDE Advisory Board XSEDE User Advisory Committee Service Providers 38기관 TAOS SDSC University of Texas at Austin CIC¹⁾</p>

Cloud 서비스

- 국가 과학 커뮤니티를 위한 산업용 클라우드 리소스, 애플리케이션, 서비스 및 도구를 개발, 배포 및 관리하고 있음

구분	세부기능	세부내용
IaaS	VM	물리적 서버(가용레온)의 CPU, 메모리, 스토리지 등 하드웨어 자원을 소프트웨어로 적으로 나누어 사용자에게 제공
	Bare Metal	소프트웨어가 설치되어 있지 않은 컴퓨터 하드웨어 별도의 가상화 환경 없이 단독으로 물리 서버 제공
PaaS	AGAVE	생명 과학 연구원을 위한 개인용 스토리지 클라우드 및 분산 컴퓨팅 애플리케이션 및 워크 플로우 최적화에 대한 전문 지식을 갖춘 웹 서비스 API
	Jupyter Hub	Python, R 등 데이터 과학 분야에서 인기있는 다양한 프로그래밍 언어 지원
	Abaco	Linux 컨테이너 기술과 동시 계산을 위한 액터 모델을 기반으로 하는 서비스 형 가능 플랫폼



- ☑ NeCTAR(National eResearch Collaboration Tools and Resources project)는 IT 기반의 연구 네트워크와 전국 8개 노드를 연결하는 지식자원 공유 및 활용 생태계를 구축하여 분야와 기관의 경계를 넘는 공동연구 플랫폼을 제공함



개요

기관명 NeCTAR
(국립 전자 협업도구 및 자원연구 클라우드)

설립

- 2009년, 호주 멜버른, AUD 47M
- 8개의 교육/연구기관, 만명의 사용자 보유

역할 기관 및 분야의 경계를 넘는 공동 연구를 위한 데이터, 모델, 분석 툴 및 워크플로우를 통합하는 온라인 연구 환경인 Virtual Labs와 연구자들의 자유로운 데이터 공유, 저장 및 분석을 지원하는 Cloud 서비스 제공

구조

호주 정부
교육훈련부

Super Science Initiative (2009) National Collaborative Research Infrastructure Strategy (2013 ~)

nectar (8 nodes)

ardc

UTAS, NCI, MONASH University, Qcif, INTERACT, PAWSEY, RDS

Cloud 서비스

- 2015년 Research Cloud 지원 서비스를 시작으로 현재에는 Research Cloud 지원 서비스와 RDS(Research Data Service)지원을 통합하여 운영

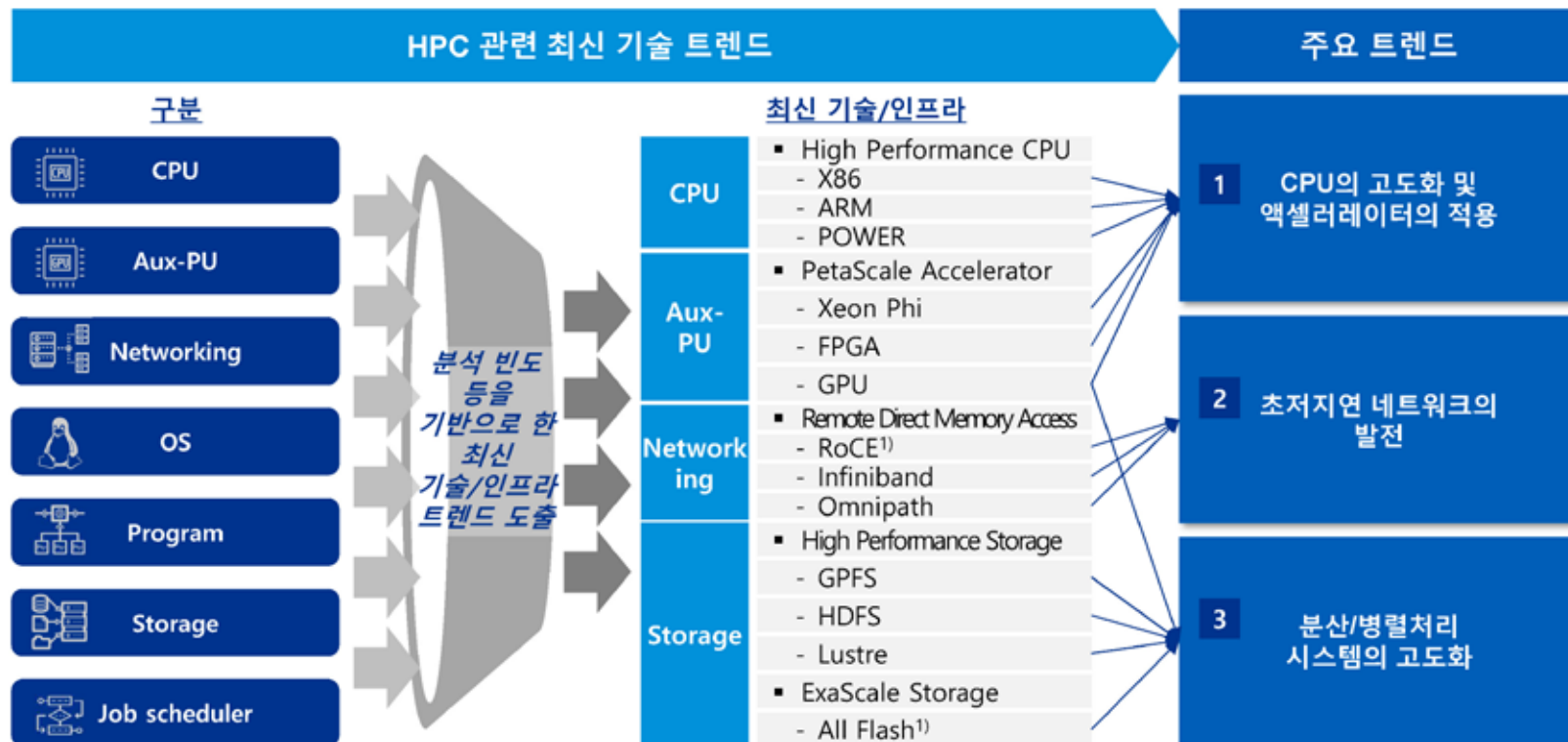
구분	서비스기능	서비스내용
IaaS	VM	윈도우즈용 VM 제공하고 있으며, CPU 1개와 2GB메모리의 인스턴스부터 CPU 16개와 32GB메모리의 최대 인스턴스까지 6단계로 구분하여 제공
	Block Storage	PC USB드라이브와 유사한 방식의 블록 스토리지 서비스 제공
	Object Storage	가상 시스템 이미지, 사진 스토리지, e-메일 스토리지, 백업 및 아카이브와 같은 정적 데이터를 위한 분산 스토리지 시스템 제공
	Shared Filesystem	여러 가상 시스템에 마운트 할 수 있는 공유 파일 시스템을 프로비저닝하고 관리하기 위한 간단한 인터페이스 제공
	GPUs	데이터 병렬 처리를 위한 GPU 서비스 제공
PaaS	Data Transfer	데이터 수집 및 전송을 위한 도구로 네트워크를 통한 데이터 이동 및 물리적 미디어(HDD, USB 등) 전송을 포함
	Workflow	단순하고 반복 가능하여 공유 가능한 연구 데이터 관리 프로세스 지원
	Database	백업 생성 및 복원, 사용자 추가 및 제거 등의 작업에 API 또는 웹 기반 액세스 제공
	Elastic Data Processing	공동 데이터 분석 클러스터를 배포, 관리 및 운영하는 데 사용할 수 있는 EDP (Elastic Data Processing) 서비스
	Application Service	사전 패키징된 클라우드 지원 애플리케이션을 배포하기 위한 간단한 인터페이스 제공
	Virtual Lab	연구 영역별 실험실 및 도구 제공



- ☑ 해외 일부 슈퍼컴퓨팅센터에서 Cloud 서비스 기술을 점진적으로 제공해나가고 있음
- ☑ IaaS 중심의 컴퓨팅 서버와 스토리지 제공을 기본으로 하고 있으며, Cloud 연구환경 제공하는 기관은 다양한 서비스 제공

기관명	서비스명	IaaS 기능	PaaS 기능
SDSC(미국)	Cloud Storage & Compute	종량제 VM 인스턴스 및 Object Storage 서비스	
CSCS(스위스)	OpenStack	OpenStack 기반 IaaS 서비스	
ORNL(미국)	CADES Cloud	OpenStack 기반 IaaS 서비스	
TACC(미국)	Chameleon, Jetstream, RODEO	BareMetal 및 VM 서비스	Agave, Abaco, TACC JupyterHub
NERSC(미국)	Spin	Container as a Service	
JSC(독일)	HDF Cloud	OpenStack 기반 IaaS 서비스	
NCHC(대만)	TWCC	VM 및 컨테이너 서비스	
NSCC-GZ(중국)	Tianhe-2 Cloud Platform	VM 임대 서비스	
NCI(호주)	NeCTAR Research Cloud	VM, Storage 등 IaaS	Virtual Lab

☑ 고성능 프로세서, 엑셀러레이터, 고속 네트워킹, 분산/병렬처리 시스템 등의 기술이 엑사스케일 속도의 HPC 구축을 위한 HPC 주요 기술 트렌드임



Note: 1) RDMA Over Converged Ethernet, 2) 하드 디스크(HDD)를 배제하고, 솔리드 스테이트 드라이브(SSD)만을 저장장치로 채택하는 스토리지 지칭

1 CPU의 고도화 및 액셀러레이터의 적용



정의

- CPU는 연산을 실행하는 가장 핵심적인 부품을 지칭
- 액셀러레이터는 연산 속도의 가속을 지원하는 부품을 지칭

기대
효과

- 고성능 CPU, GPU, 액셀러레이터를 적용하여 연산속도의 비약적 상승 도모

고성능 CPU, GPU, 액셀러레이터
기반 대규모 연산처리 역량 고도화

2 초저지연 네트워크의 발전



정의

- 최소 지연으로 대용량 데이터를 처리하도록 최적화된 네트워크를 지칭

기대
효과

- 시스템간의 연결 네트워크 성능 제고
- 대형 계산 문제의 Input/Output 전송시간 단축

초저지연 Interconnect 등의 활용을
통한 시스템간 네트워크의 성능 제고

3 분산/병렬처리 시스템의 고도화



정의

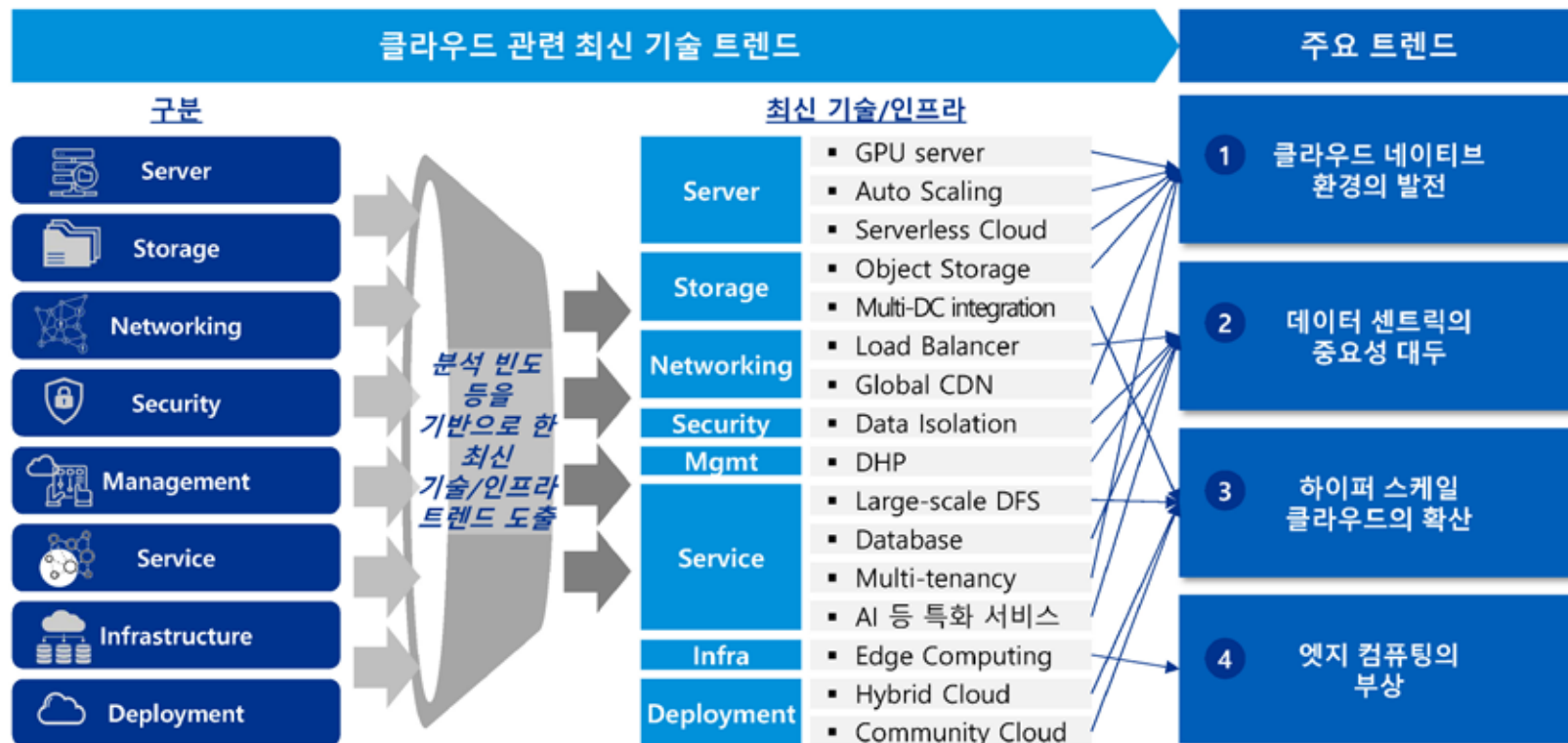
- 병렬로 연결된 컴퓨터 또는 여러 개의 분산된 기기들을 네트워크로 연결하여 동시에 처리하는 시스템

기대
효과

- 크고 복잡한 문제를 분해하여 병렬적으로 해결
- 대형 계산 문제 처리 시간 단축

분산/병렬처리 시스템의 고도화를
통한 4차 산업혁명 관련 대응 역량 제고

☑ Cloud 네이티브 환경의 발전, 데이터 센트릭 트렌드의 중요성 대두, 하이퍼스케일 클라우드의 확산 등이 클라우드 주요 기술 트렌드



1 클라우드 네이티브 환경의 발전



정의

- 어플리케이션 또는 서비스 시스템 관리를 컨테이너 기반 기술로 자동화한 환경

배경

- 시스템 관리 방법이 서버 기반의 정적 인프라에서 가상화 기반의 동적 인프라로 변화

특성

- 기존 환경 대비 시스템의 높은 안정성, 운영/관리 업무효율, 비즈니스 요구에 대한 빠른 응대, 시스템의 확장 용이성 등 제고

클라우드 네이티브 환경을 통해 유연성 확보 및
고객의 니즈에 빠르게 대응 가능한 경쟁력 확보

2 데이터 센트릭의 중요성 대두



정의

- 데이터 집적지 구상의 역할을 하며 데이터 중심으로 서비스 되는 클라우드

배경

- 가상 머신의 확장성 중심에서, 원활한 데이터 연계 여부 중심으로 변화

특성

- 하나의 대형 스토리지를 기반으로 데이터에 어디서나 접근 가능토록 하며, 저지연 네트워크를 통해 서비스를 그물처럼 연결

데이터 중심의 클라우드 환경으로
원활한 데이터 연계 제공

3 하이퍼스케일 클라우드의 확산



정의

- 분산된 컴퓨팅 환경을 최대 수천 개의 서버로 확장할 수 있는 클라우드를 지칭

배경

- 폭발적으로 증가하는 워크로드에 대한 수요를 경제적으로 수용하기 위해 탄생

특성

- 다양한 비즈니스 니즈에서 요구되는 규모 및 컴퓨팅 집적도를 오토스케일링/ 오토프로비저닝 기반의 최적화된 경제성으로 해결 가능

하이퍼스케일 클라우드를 통한
폭증하는 데이터 수요 대응

4 엣지 컴퓨팅의 부상



정의

- 데이터가 생성되는 단말과 가까운 곳에서 데이터를 처리하는 방식

배경

- 대용량 데이터의 실시간 처리를 필요로 하는 분야에 대응하기 위해 탄생

특성

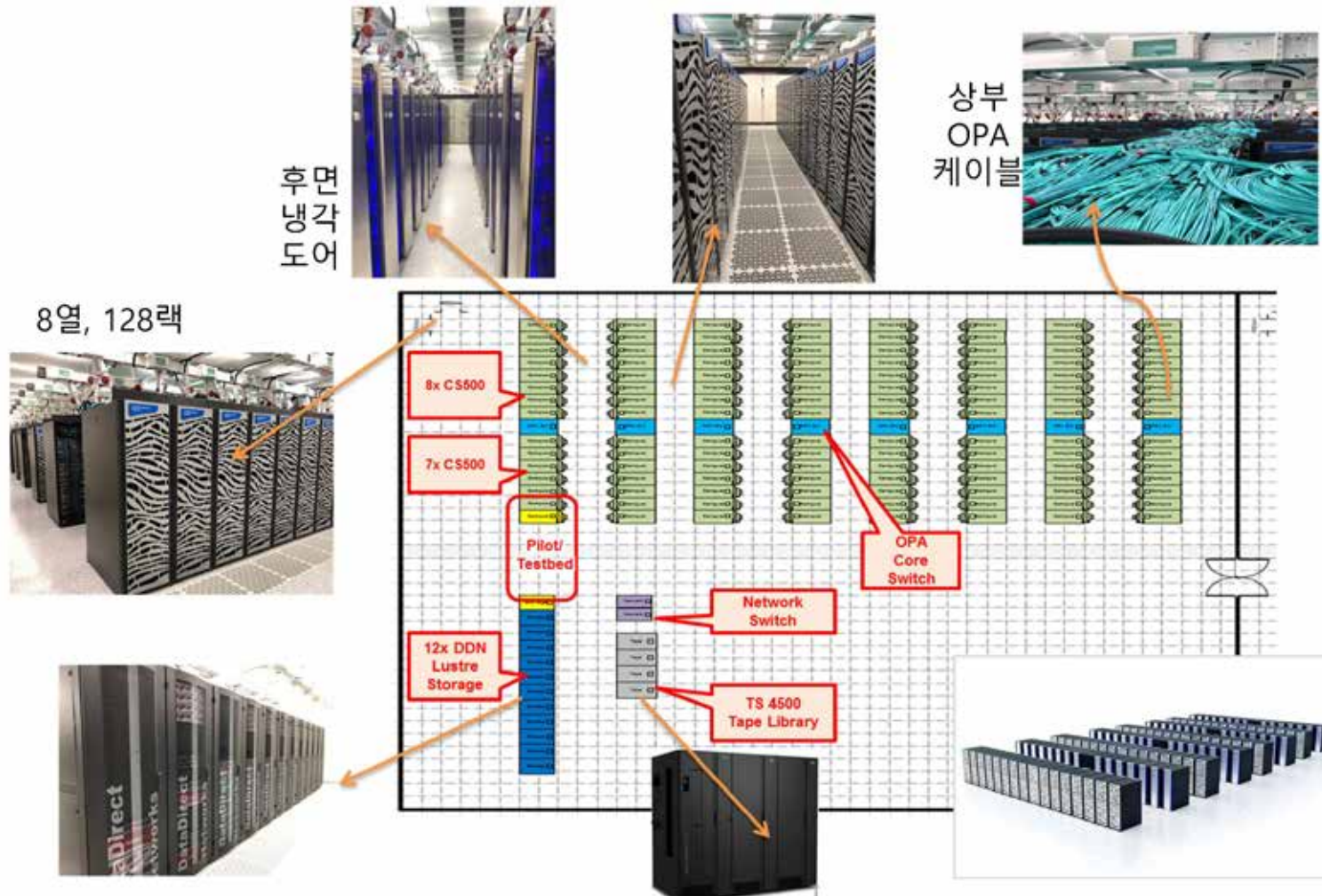
- 엣지의 활용으로 대용량 데이터의 효율적 처리 및 대역폭 사용량 감소
- 5G의 낮은 지연시간의 이점을 활용하는 엣지 컴퓨팅 시장 성장 중

엣지 플랫폼 등 구축을 통해
실시간 데이터 처리가 필요한 분야에 대응

KISTI HPC Cloud

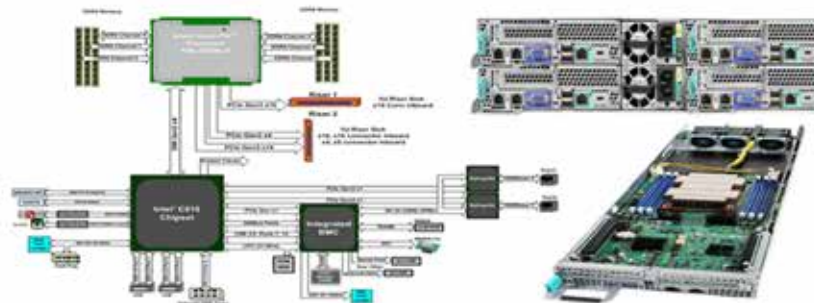






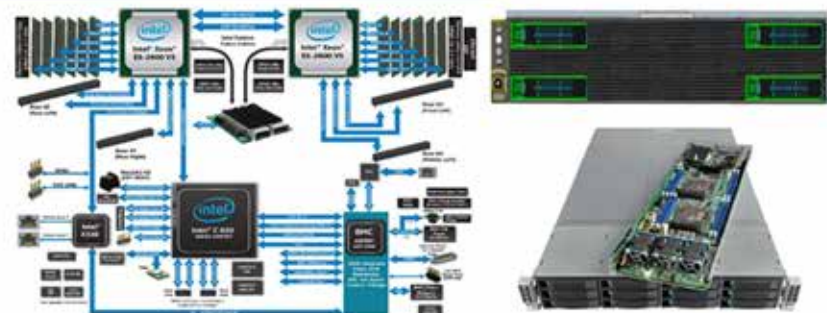
계산노드 Cray 3112-AA000T(2U enclosure), KNL 기반 계산 노드 (8,305개)

- 1x Intel Xeon Phi KNL 7250 processor
- 68 Cores per Processor
- 96GB (6x 16GB) DDR4-2400 RAM
- 1x Single-port 100Gbps OPA HFI card
- 1x On-board GigE (RJ45) port



CPU-only Cray 3111-BA000T(2U enclosure), Skylake 기반 계산 노드 (132개)

- 2x Intel Xeon SKL 6148 (2.4 GHz) processors
- 20 Cores per processor (total 40 Cores)
- 192GB (12x 16GB) DDR4-2666 RAM-6CH
- 1x Single-port 100Gbps OPA HFI card
- 1x On-board GigE (RJ45) port



- ☑ (초고성능컴퓨팅 기본계획) 초고성능컴퓨팅 활용 확대를 위해 다양한 기술 분야의 요구사항에 대응할 수 있는 동적인 서비스를 클라우드 기반 컴퓨팅 기술을 이용하여 구현
- ☑ (빅데이터·AI 서비스 수요) 대용량 데이터를 이용한 분석, 딥러닝 기반 예측 등의 데이터 기반 R&D 증가
- ☑ (5호기 서비스 전략) 유연한 인프라·플랫폼 서비스로 경쟁력 강화
- ☑ (목적) 산학연 연구자들이 슈퍼컴퓨팅 환경에 쉽게 접근하여 HPC·AI·빅데이터 연구를 수행할 수 있는 환경을 제공



< 슈퍼컴퓨터 5호기 서비스 전략 >



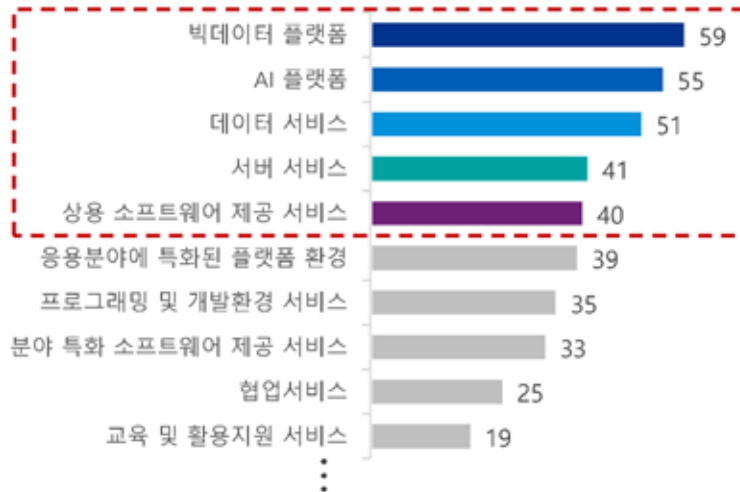
- ☑ 정책, 시장, 기술 관점의 메가트랜드 도출을 위한 환경분석 연구
- ☑ 국내외 주요 기관의 서비스 분석 및 수요조사 분석을 토대로 KISTI HPC 클라우드 핵심 서비스 도출
- ☑ 연구자 대상 설문조사 분석을 토해 서비스 방향성 도출을 위한 서비스 분석연구



- ☑ 설문조사 : 출연연, 산업계, 학계 대상 181명 응답(1600명 대상) (2018.9.10~9.20)
- ☑ HPC 클라우드 기반의 빅데이터, AI 등 미래 핵심 과학기술 관련 서비스 및 플랫폼 서비스에 대한 수요 높음
- ☑ 산업별 여러 분야(정보통신/ 기계, 물리학 등)에서도 4차 산업혁명 관련 서비스에 대한 수요가 높음

HPC 클라우드 서비스 미래 수요분야

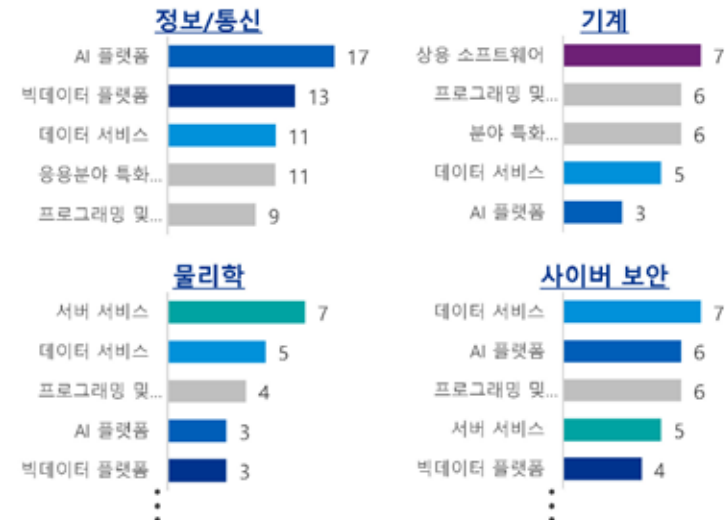
(단위: 응답 수)



- ✓ 빅데이터, AI 등 미래 핵심 과학기술 관련 서비스 수요 증가
- ✓ 미래 수요 상위 5개 중 2개 항목이 플랫폼 서비스에 해당

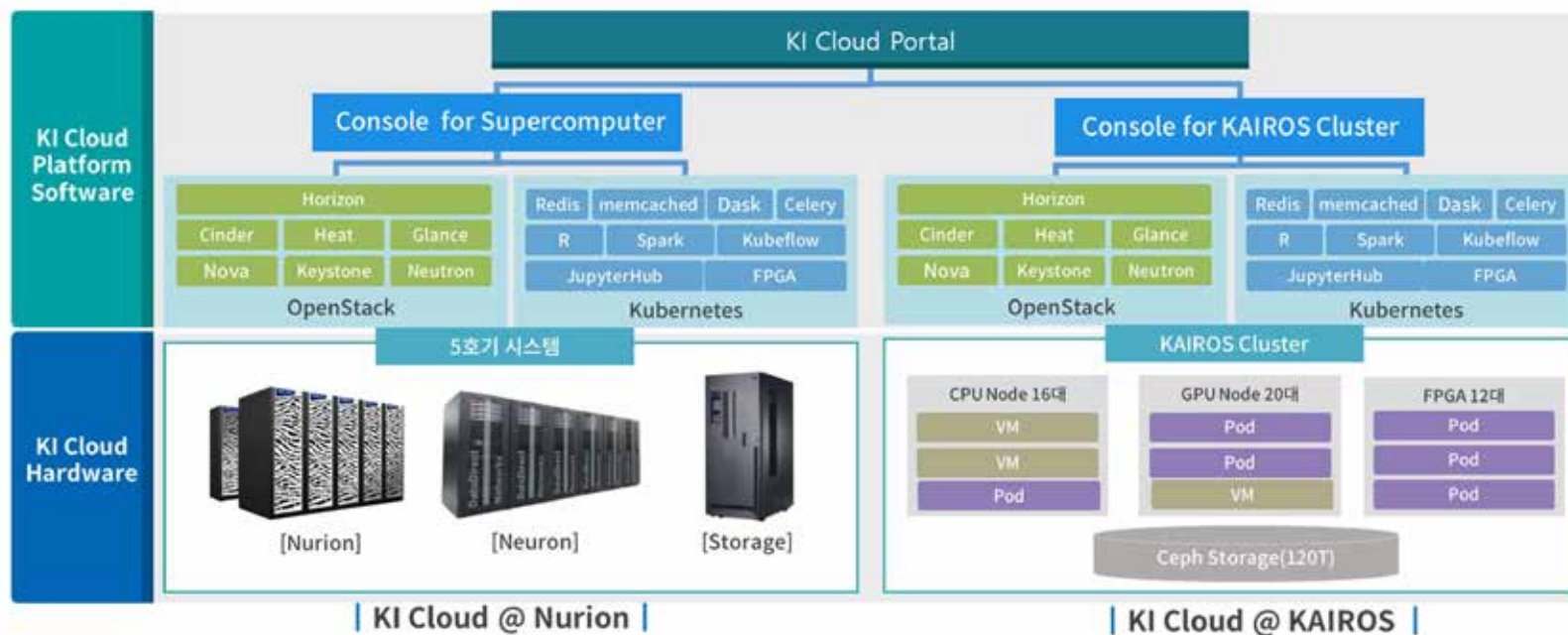
산업별 HPC 클라우드 미래 수요 분야

(단위: 응답 수)

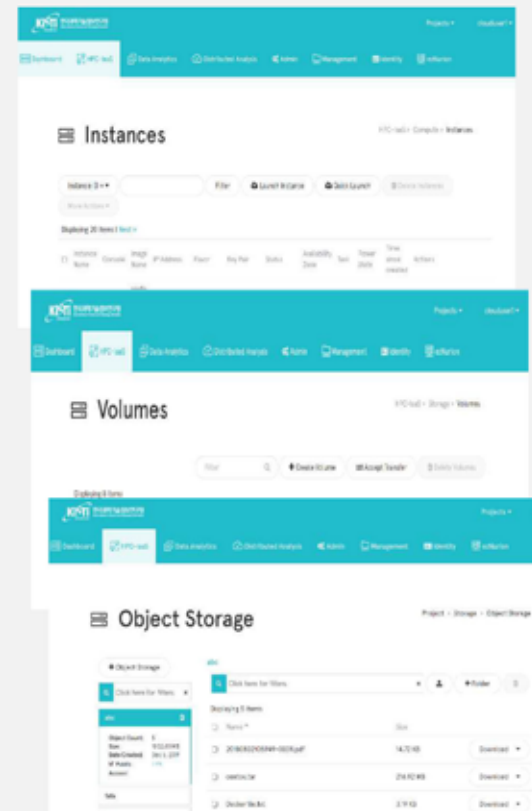


- ✓ 산업별 주요 서비스 또한 수요층 공통의 주요 수요와 유사
- ✓ 분야별 특성에 따라 데이터 서비스 및 플랫폼 관련 수요 존재

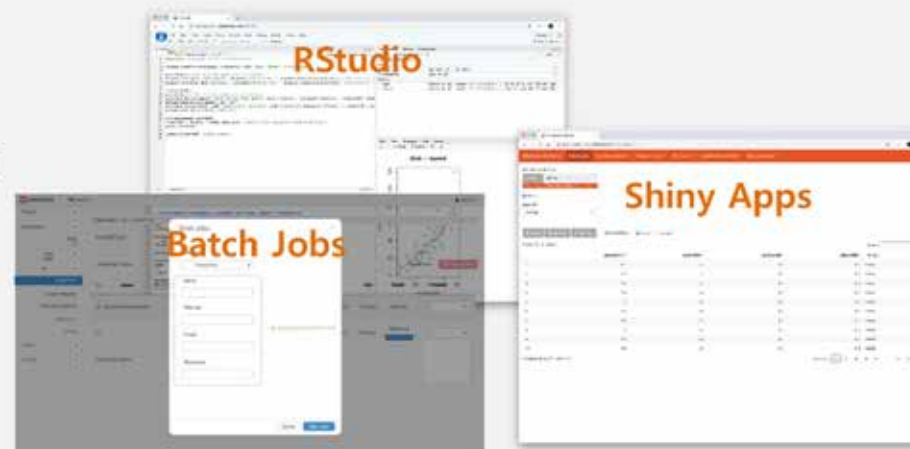
- ☑ 오픈소스 기술을 통합 슈퍼컴퓨터 가상화 환경 및 클라우드 서비스 기능 구현
- ☑ 5호기를 구성하는 SKX와 KNL 자원 특성에 적합한 가상화를 지원할 수 있도록 코어 계층 구축
 - SKX 노드는 Openstack VM 기반의 IaaS 서비스, KNL 노드는 Container 기반의 PaaS 서비스 구축
- ☑ 범용 환경에서 개발된 오픈소스 플랫폼을 KISTI 슈퍼컴퓨터에 적용하기 위한 Customization을 통해 구축

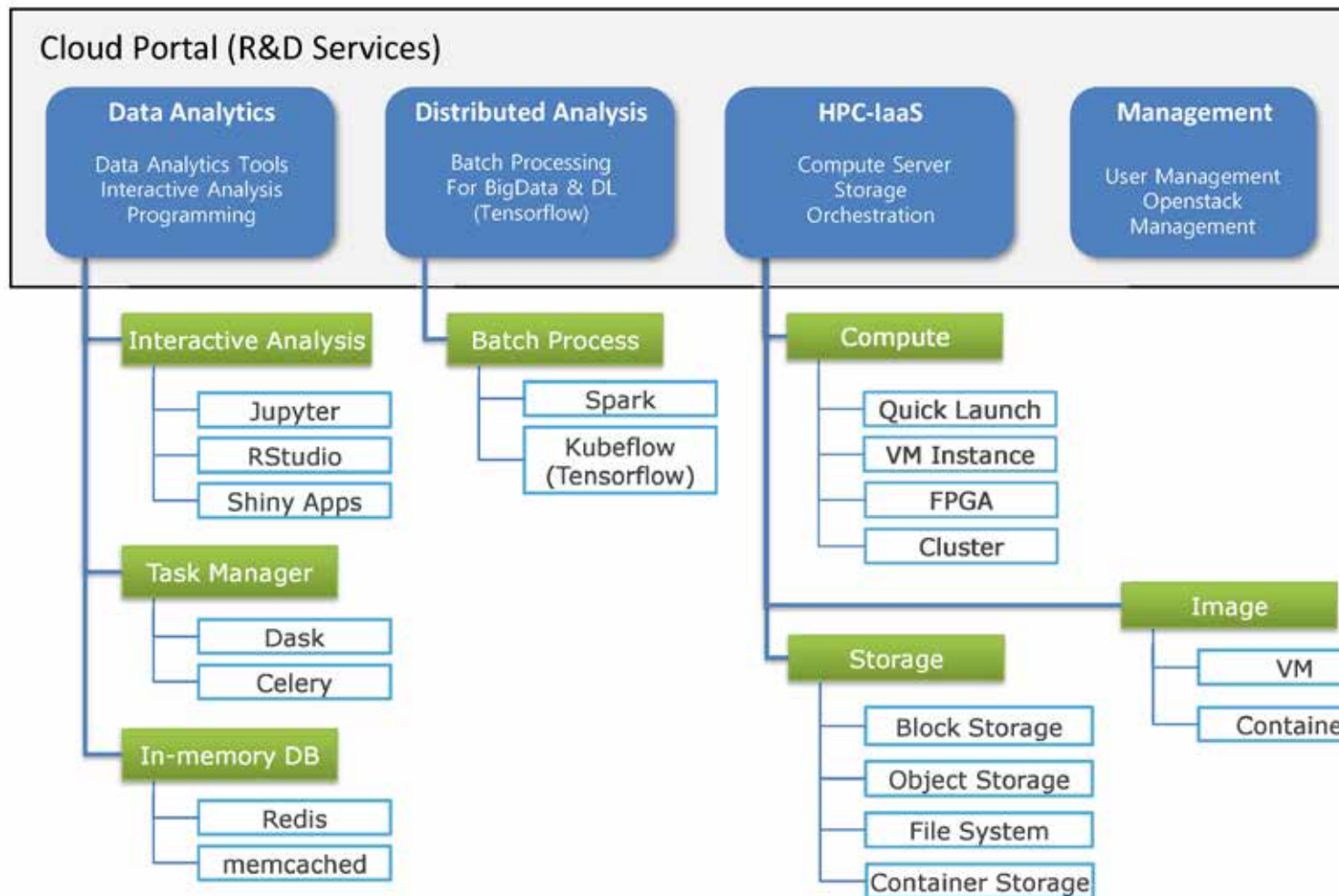


- ☑ 웹 기반 인터페이스를 통해 VM 생성 및 관리 기능 제공
 - VM 생성/삭제/관리
- ☑ VM에 접속 가능한 Web 터미널 기능
- ☑ 스토리지 관리 및 서비스
- ☑ 네트워크 관리 기능
 - 네트워크 생성/삭제/수정/조회
 - 라우터 생성/삭제/수정/조회
 - 보안 그룹 생성/조회/삭제/수정
 - Floating IP 할당
- ☑ 이미지 관리 기능
 - 이미지 생성/삭제/조회/이미지로부터 Volume 생성
- ☑ 자원 사용량 모니터링 및 과금 관리 기능
 - Cloud 기반 자원에 대한 과금을 부과할 사용량 제공

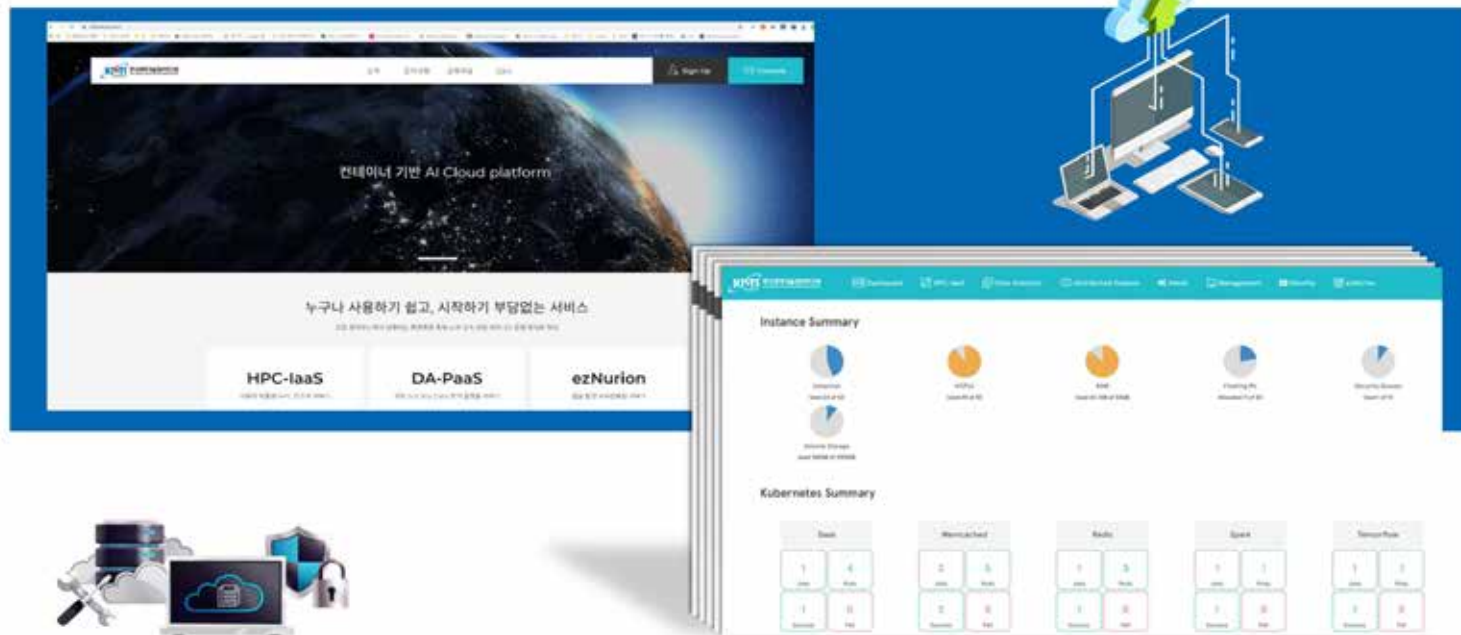


- ☑ 웹 브라우저를 통한 Interactive Analytics 환경 제공
 - Jupyter 기반의 Python, R, Spark등 다양한 개발 환경 제공
 - 데이터분석 및 머신러닝/딥러닝 환경 제공
 - Lustre Home Mount
- ☑ In-memory DB, Task Manager 등 데이터 분석 도구 제공
 - In-memory DB : Redis, Memcached
 - Task Manager : Dask, Celery
- ☑ 분산 딥러닝 등 ML/DL 실행 환경
 - Spark, Kubeflow 기반 딥러닝 지원
- ☑ R 기반 웹 어플리케이션 플랫폼 지원
 - RStudio Server
 - R Shiny 플랫폼





- ☑ KISTI HPC cloud : KI Cloud 서비스
 - 접속 URL : <https://kicloud.ksc.re.kr>
- ☑ 품질 및 기능에 대한 인증
 - NIPA 클라우드 품질 성능 확인제 인증 완료
 - TTA의 소프트웨어 확인 및 검증(V&V) 인증 완료







<https://kicloud.ksc.re.kr>



THANK
YOU