



새로운 슈퍼컴퓨팅 시대 개막: 엑사플롭스, 인공지능, 양자컴퓨터

이 식 · 함재균

대규모 시뮬레이션 연구에 주로 쓰이던 슈퍼컴퓨팅 분야 역시 급격한 변화와 도전을 겪고 있다. 최근 슈퍼컴퓨팅 분야의 변화는 ‘엑사스케일 + 인공지능 + 양자컴퓨터’로 요약될 수 있다.

전 세계에서 가장 빠른 컴퓨터 500대의 순위인 Top500은 매년 2회 발표된다. 2022년 5월 말 발표된 Top500에 최초의 엑사플롭스 컴퓨터가 등장했다. 미국 오크릿지 국립연구소(Oak Ridge National Laboratory, ORNL)의 프론티어(Frontier) 시스템이 그 주인공이다. 프론티어의 실측성능은 1.1엑사플롭스로 1초에 110경 번의 실수연산을 수행할 수 있다. ‘엑사스케일’의 시대가 등장한 것을 기념하기 위해서 10월 18일을 ‘엑사스케일의 날’로 지정했다.

역사적으로 살펴보면, 최초로 테라플롭스(10^{12}) 시대를 연 것은 1997년 6월 미국 샌디아 국립연구소(Sandia National Laboratory, SNL)의 ASCI Red, 페타플롭스(10^{15}) 시대를 연 것은 2008년 6월 로스 알라모스 국립연구소(Los Alamos National Laboratory, LANL)의 로드런너, 엑사플롭스(10^{18})시대를 연 것은 2022년 6월 ORNL의 프론티어이다. 테라 시대에서 페타 시대로 가는 데 9년이 걸렸지만, 페타 시대에서 엑사 시대로 가는 데는 14년이 걸린 셈이다. 엑사 시대가 임박하면서 최근 몇 년간 누가 먼저 엑사 시스템을 개발할 지는 연구계의 큰 관심사였다. 이에 가장 체계적이고 투명하게 엑사 시대를 열어가고 있는 미국과 유럽의 슈퍼컴퓨팅 인프라와 관련 생태계 변화, 최신 동향 등을 살펴본다.

CONTENTS

- | | |
|---|----------------------------------|
| 1. 미국: 정부 차원에서 인프라 구축과 HPC 활용을 체계적으로 지원 | 3. 중국, 일본, 한국의 최신 동향 |
| 2. 유럽: PRACE와 EuroHPC를 통합 범 유럽차원의 생태계 구축 | 4. 슈퍼컴퓨팅 최신 기술 동향 및 특허 분석 |

1. 미국: 정부 차원에서 인프라 구축과 HPC 활용을 체계적으로 지원

- **(HPC법 제정)** 고성능컴퓨팅(High Performance Computing, HPC)의 중요성을 인식하고, 이의 개발·운영·활용에 대한 개별 법¹⁾을 제정하여 정부 차원에서 HPC 활성화를 주도적으로 이끌고 있음
- **(National Strategic Computing Initiative, NSCI)** 슈퍼컴퓨팅을 활용한 사회문제 해결, 기술개발, 산업혁신 등의 전방위적 발전을 위해 정부차원에서 슈퍼컴퓨팅 전략을 수립하여 발표 (2016.7)
 - 5대 전략: 엑사스케일 시스템의 구현, 기술적 일관성, 무어의 법칙을 뛰어넘은 컴퓨팅, 지속가능한 HPC 생태계, 민관협력을 통한 발전
- **(NSCI Update, 2019.9)** 산·학·연 이해관계자가 참여한 논의사항을 중심으로 컴퓨팅 환경의 변화에 따른 NSCI Update 발표
 - 기존 5대 전략을 수정·보완하여 컴퓨팅의 미래활성화, 컴퓨팅을 위한 전략적 기반 제공, 협업 및 조정 접근 방식 보장의 새로운 3대 전략 수립

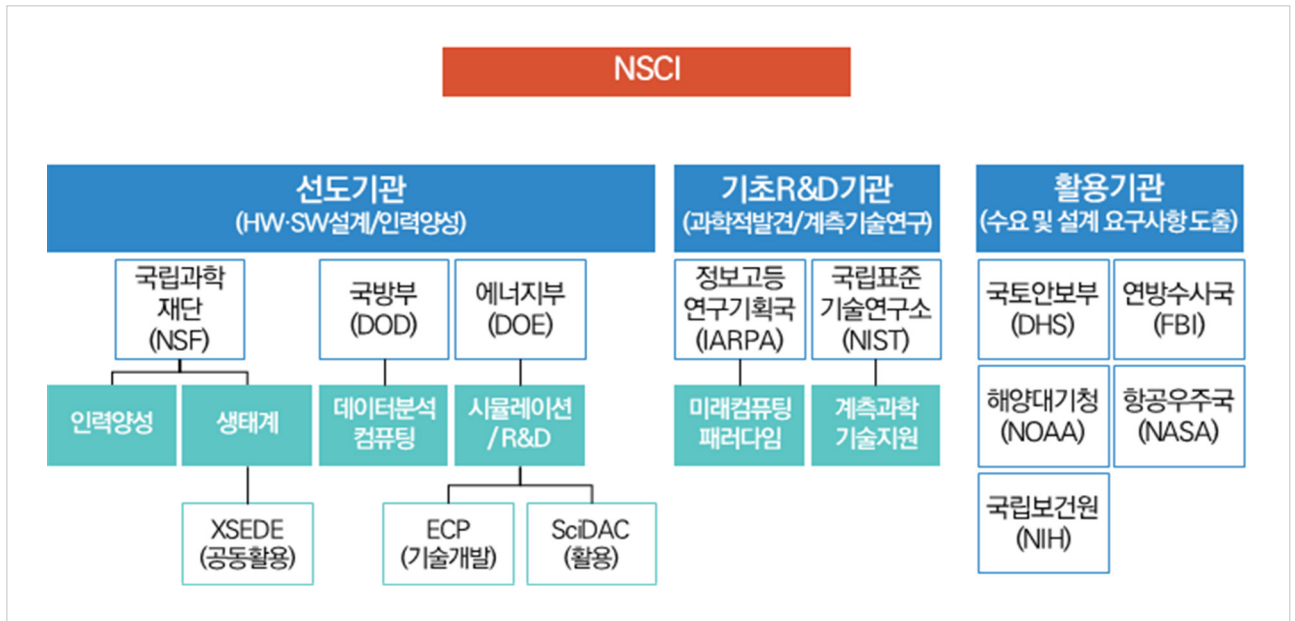
<그림 1> NSCI update 추진전략 및 세부내용

추진전략	전략목표	권고안 (Recommendations)
컴퓨팅의 미래 활성화	디지털 및 비디지털 계산의 새로운 영역 개척	옛지, 엑사스케일 및 그 이상까지의 효율적 활용을 위한 생태계 확장 단대단(end-to-end) 애플리케이션 워크플로우 및 통합시스템 개발 컴퓨팅에 대한 중요한 과학적, 기술적 한계 탐구
컴퓨팅을 위한 전략적 기반 제공	컴퓨팅 인프라와 생태계의 개발, 확장 및 발전	강력한 하드웨어 및 소프트웨어 기반 제공 컴퓨팅 생태계에서 사이버 보안의 중요성과 사이버 보안 강화 및 가속화 컴퓨팅을 위한 데이터 사용 및 관리 지원 광범위한 미국 혁신 생태계를 지원하는데 필요한 인력 양성
협업 및 조정 접근방식 보장	미래 컴퓨팅 파트너십 구축 및 확장	연방기관, 학계 및 산업계 협업 촉진 미래 컴퓨팅 이니셔티브에 초점을 맞춘 기관 건 거버넌스 구조 구현

1) High Performance Computing Act of 1991

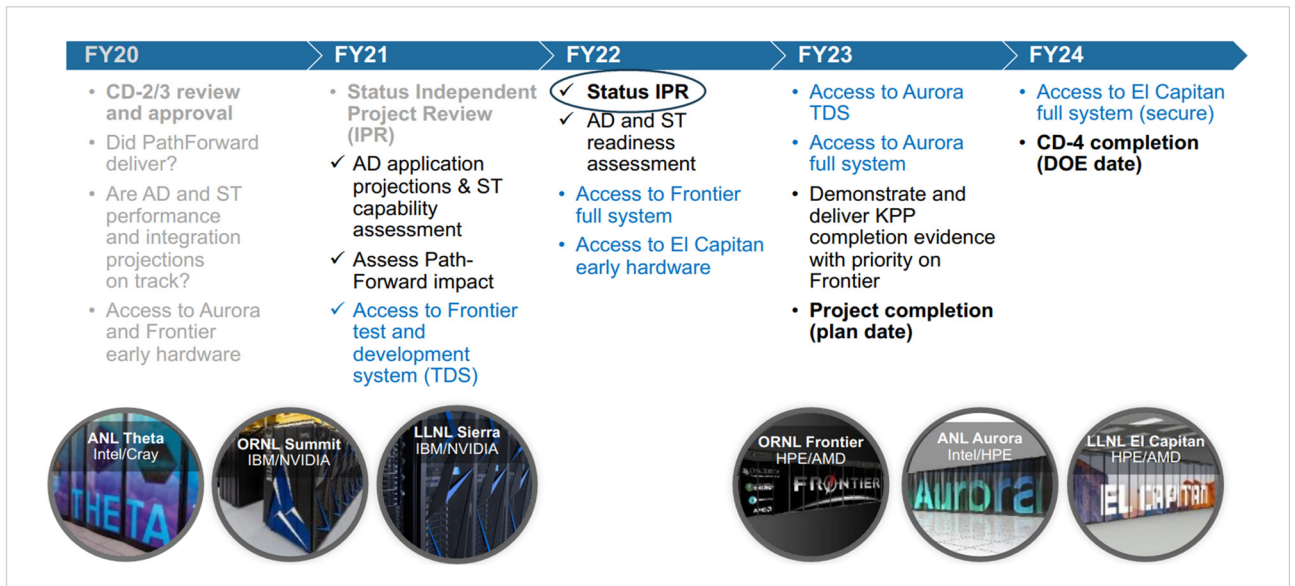
- **(NSCI 참여기관)** 전략적 목적 달성을 위해 선도기관(Lead), 기초R&D기관, 활용기관 선정하여 기관별 역할 명시

<그림 2> NSCI 참여 기관 및 역할



- 중심적 역할을 할 선도기관에는 국립과학재단과(NSF), 국방부(DoD), 에너지부(DoE)가 포함
- NSF는 XSEDE와 후속인 ACCESS 프로그램을 통해 생태계 육성과 HPC 자원의 공동활용을 이끌고 있고, 에너지부는 ECP(Exascale Computing Project, <http://exascaleproject.org>)와 SciDAC(Scientific Discovery through Advanced Computing, <http://www.scidac.gov>) 프로그램을 통해 기술개발과 거대컴퓨팅 관련 R&D를 지원
- **(DoE의 엑사급 슈퍼컴퓨터 개발)** 2016~2020s까지 엑사스케일 컴퓨팅에 22억 달러를 투자. 2024년이면 세계에서 가장 강력한 슈퍼컴퓨터 3대를 보유하게 됨
 - 에너지부 산하 세 개의 국립연구소에서 서로 다른 아키텍처의 엑사스케일 슈퍼컴퓨터를 개발 중. 절대 강자인 Nvidia 대신 Intel과 AMD를 지원하는 형태로 개발 사업이 추진되고 있음

<그림 3> 미국 엑사스케일 슈퍼컴퓨터 구축 로드맵



출처) ECP Application Development Update 2022

- ORNL은 9,472개의 AMD의 “Trento” 64코어 CPU(총 606,208 코어)와 37,888개의 MI250X 220코어 GPU 가속기(총 8,335,360 코어)를 이용하여 최초의 엑사스케일 컴퓨터인 프론티어를 2022년 완료(<http://www.top500.org> 참고)

* 이론 성능 1.68EFlops, 실측성능 1.1EFlops의 이기종 시스템

- 아곤 국립연구소(Argonne National Laboratory, ANL)는 인텔의 사파이어 래피즈(Sapphire Rapids) CPU와 폰테 베키오(Ponte Vecchio) GPU 기반으로 “오로라(Aurora)” 시스템을 구축하기로 했으나 일정이 지연되고 있음

- 로렌스 리버모어 국립연구소(Lawrence Livermore National Laboratory, LLNL)는 AMD “Genoa” CPU와 MI300 APU²⁾로 2.0 엑사플롭스³⁾의 슈퍼컴퓨터인 “엘 캐피탄(El Capitan)”을 설치 중

● (NSF의 자원 활용 프로그램) 5개년 사업인 XSEDE를 성공적으로 완수하고 후속사업인 ACCESS 프로그램을 2022년 시작

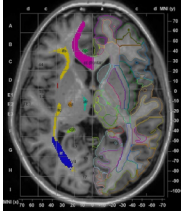
- (XSEDE, 2016~2021) 전문가들이 연구를 위해 사용하는 슈퍼컴퓨터, 시각화 및 저장 시스템, 데이터 모음, 소프트웨어, 네트워크, 전문가 지원 등의 강력한 통합 디지털 리소스 및 서비스의 모음. XSEDE는 이러한 리소스와 서비스를 통합하여 더 쉽게 사용할 수 있도록 하고 더 많은 사람들이 이를 사용할 수 있도록 지원함. 분야별 특정 과학 커뮤니티에 최적화된 Science Gateway 기능도 제공

* XSEDE: eXtreme Science and Engineering Discovery Environment

2) AMD에서 출시된 GPU 통합형 CPU로 단일 칩 안에 CPU와 GPU가 함께 있음

3) 엑사(Exa)는 10¹⁸, 플롭스(FLOPS)는 Floating-point Operation Per Second)이므로 엑사플롭스는 1초에 10¹⁸번(100경 번)의 실수연산이 가능함을 의미

<표 1> XSEDE의 Science Gateway

이름	분야	내용	
Atomic and Molecular Physics Science Gateway	원자, 분자 및 광학 물리학	원자 또는 분자에서 일어나는 전자 산란 및 방사선과의 상호 작용을 연구할 수 있는 충돌 코드 제공	
Computational Anatomy	뇌 과학 (생물학)	Computational Anatomy의 해부학적 이미지를 토대로 형태학적 변화를 직접 비교하고 디지털화 할 수 있음	
DesignSafe: Natural Hazards Engineering Research Infrastructure	엔지니어링	자연 재해 연구를 위한 포괄적인 사이버 인프라 환경으로 데이터 관리, 분석 및 이해하는 데 필요한 계산 도구를 제공하는 연구 플랫폼	
Asteroseismic Modeling Portal	항성 천문학 및 천체 물리학	Aarhus Stellar Evolution Code를 병렬 유전 알고리즘과 함께 사용하여 맥동 주파수에서 태양과 같은 별의 속성을 도출할 수 있도록 웹 기반 인터페이스를 제공	

- (ACCESS, 2022~2027) XSEDE의 후속 프로그램. 플랫폼 서비스로 HPC, HTC, 가시화 및 데이터 분석, 대용량 메모리, 스토리지 및 클라우드 서비스를 총괄적으로 지원. 미국내 13개 기관의 시스템을 연결하여 Compute(CPU, GPU, regular memory, large memory), Storage 및 Testbeds 서비스 제공

* ACCESS: Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support

<표 2> ACCESS 제공 서비스

구분	프로젝트	관련 활동
ACCESS Service Tracks	할당 서비스	할당 서비스, 혁신적인 파일럿, 서비스 모델
	최종 사용자 지원 서비스	일반 사용자 지원, 할당 및 활용 지원, 최종 사용자 교육, 전산과학 지원 네트워크
	운영 및 통합 서비스	운영 지원, 데이터 및 네트워킹 지원, 사이버 보안 지원
	모니터링 및 측정 서비스	모니터링 및 측정 작업, 서비스 모델, 데이터 분석 프레임워크
	기술 번역 서비스	파일럿/개발 단계 운영, 생산 단계 운영, 번역 파이프라인
ACCESS Coordination Office Service	ACCESS 조정 사무소(ACO)	집행위원회, 외부 자문 위원회, 커뮤니케이션 및 홍보

출처) <https://www.xsede.org/advancetoaccess>

2. 유럽: PRACE와 EuroHPC를 통합 범 유럽차원의 생태계 구축

- 초고성능컴퓨팅 공동활용 및 커뮤니티를 활성화 지원을 위한 PRACE와 인프라 투자를 지원하는 EuroHPC를 통해 HPC 생태계 조성. 서비스 되는 컴퓨팅 자원의 규모는 EuroHPC가 훨씬 더 큼

<그림 4> PRACE와 EuroHPC와 파트너십 관계



- (PRACE) 범 유럽 차원의 과학 및 산업계 발전을 위해 범 유럽차원의 거대자원인 Tier-0과 국가차원의 대형시스템인 Tier-1 시스템의 HPC 서비스를 제공하고, 대학 및 사용자 커뮤니티의 활성화를 위한 생태계를 육성하기 위한 Bottom-up 접근 방식. 2008년 시작되어 현재 유럽 25개국에 참여하고 있으며, 프랑스, 독일, 이탈리아, 스페인, 스위스가 핵심역할을 수행 중

* Partnership for Advanced Computing in Europe

<표 3> PRACE Tier-0 슈퍼컴퓨팅 시스템 현황

국가	기관명	시스템명	RPeak (PF)	
프랑스	CEA (GENCI)	Joliot-Curie	KNL	2.52
			Rome	11.75
			SKL	6.86
독일	FZJ (GCS)	JUWELS	Booster	70
			Cluster	12
	LRZ (GCS)	SuperMUC-NG	26.9	
	HLRS (GCS)	HAWK	26	
이탈리아	CINECA	Marconi100	32	
스페인	BSC	MareNostrum 4	11.14	
스위스	CSCS & ETH	Piz Daint	7.787	
합계			201 PF (최대 77 PF 허용)	

* 독일은 국가센터가 3곳이어서 PRACE에는 GCS(Gauss Centre for Supercomputing)란 대표이름으로 참여하고 있음.

<표 4>PRACE 자원의 액세스 유형

유형	세부내용	추천 사용자
PROJECT ACCESS	높은 확장성과 최적화가 입증된(사전 테스트가 완료된) 코드를 사용하는 프로젝트에 대한 PRACE Tier-0 HPC 시스템에 대한 액세스 권한	Tier-0 사용자
PREPARATORY ACCESS	RACE 프로젝트 접근에 대한 제안요청에 적용하기 전에 PRACE Tier-0 시스템에서 최적화, 확장 및 테스트 코드를 사용할 수 있도록 준비하는 과정	프로젝트 액세스를 준비하는 사용자
SHAPE ACCESS	중소기업이 PRACE RI 내에서 연구한 전문성과 지식으로부터 이익을 얻을 수 있도록 지원	HPC 사용하는 중소기업
DECI ACCESS	소규모 프로젝트를 위해 Tier-1 아키텍처 접근 지원	소규모 프로젝트

● **(EuroHPC)** 세계 최고 수준의 HPC 기술을 개발하고 인프라를 구축하여 유럽의 연구자들에게 제공하는 Top-down 방식의 프로그램으로 Horizon 2020 및 Horizon Europe을 통해 자금을 지원

- '20년 미드레인지 시스템 구축을 시작으로 '23년까지 엑사스케일 시스템 개발을 추진

(0단계) 불가리아 외 4개 국가에 4PF ~ 14PF 규모의 미드레인지 시스템

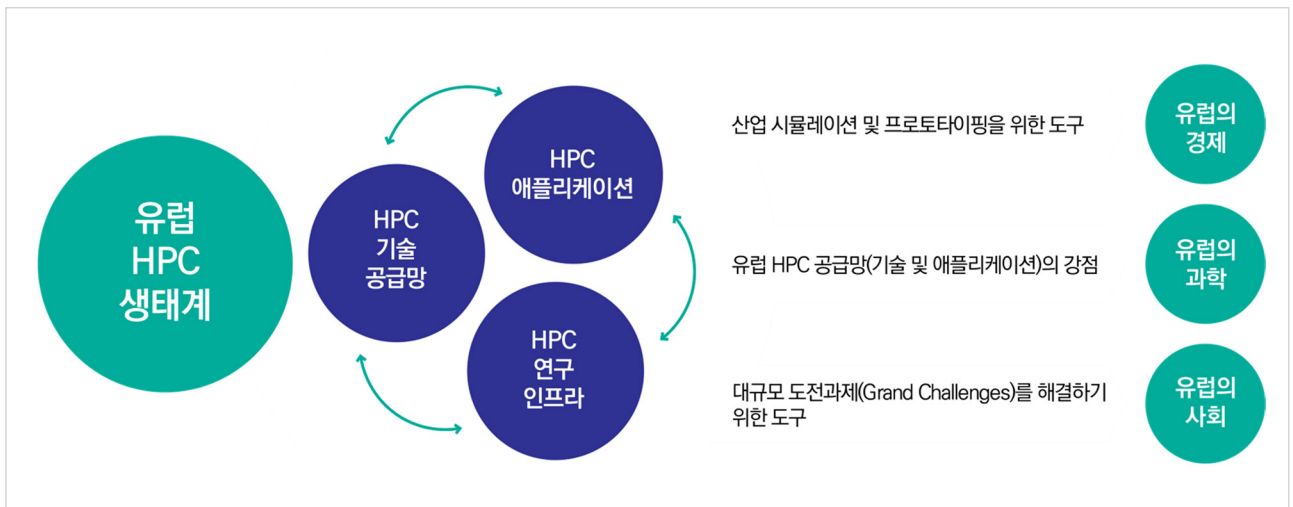
(1단계) 스페인 외 2개 국가에 200PF 급 pre-엑사스케일 시스템 구축

(2단계) EPI 프로젝트를 통해 개발되는 프로세서 기반의 엑사스케일 시스템

- 연구 인프라, HPC 기술 및 애플리케이션, 전문성의 세 가지 중심으로 추진

* European Processor Initiative

<그림 5>EuroHPC 생태계 모식도



출처) <https://www.etp4hpc.eu/>, 미래비전그룹 수정

● **European High-Performance Computing Joint Undertaking(EUROHPC JU)을 통해 2023년까지 엑사급 슈퍼컴퓨터 개발하는 프로젝트 추진**

- 2018년 10월, 유럽 HPC 생태계의 개발·촉진, 엑사스케일 시대에 도달하고 퀀텀 기술 개발, 고성능 컴퓨팅 및 데이터의 배포·운영 인프라를 위해 설립
- JU의 구성원은 EU회원국으로서 Horizon Europe, Digital Europe Program 또는 Connecting Europe Facility와 관련된 고성능 컴퓨팅 및 양자컴퓨팅에 대한 유럽 공동 이니셔티브를 따라야 함
- EuroHPC JU 활동은 5개의 핵심 의제(Pillar)로 구성됨

<표 5> EuroHPC JU사업의 핵심의제

핵심의제	세부내용
인프라	글로벌 리더 수준의 슈퍼컴퓨팅 및 양자컴퓨팅 인프라 확보·활용·운영
슈퍼컴퓨팅 통합 서비스	클라우드 기반으로 고성능컴퓨팅, 양자컴퓨팅, 데이터 인프라의 통합 서비스를 제공하며, 안전한 클라우드 기반의 서비스 플랫폼 운영
기술	세계 최고 수준의 혁신적인 슈퍼컴퓨팅·양자 컴퓨팅 생태계를 위한 컴퓨팅 HW 및 SW 기술과 관련 생태계 가치사슬 확보
애플리케이션	HPC 우수센터(CoE) 지원을 포함하여 과학계 및 산업계(특히, 중소기업)에서 활용하는 어플리케이션 지원
저변확대	회원국의 HPC CoE 구축 및 연계로 슈퍼컴퓨팅, 양자컴퓨팅, 데이터 활용 및 기술 교육, 훈련 제공

● **(EuroHPC JU의 10년 계획)**

- 중급 고성능컴퓨터와 최소 2개의 최상급 엑사스케일 및 2개의 최상급 포스트 엑사스케일 시스템이 있는 HPC 및 데이터 인프라 구축. 특히, 각 범주 중 하나 이상은 유럽 기술로 구축

<표 6> EuroHPC JU pre-엑사스케일 시스템 구축 현황

구분	LUMI	Leonardo	MareNostrum5
구축 국가와 기관	핀란드 CSC	이탈리아 Cineca	스페인 BSC
구축 년도	2022년 상반기	2022년 하반기	진행 중
실측 성능	375 PF	249.47 PF	205 PF
이론 성능	550 PF	323.40 PF	314 PF
Top500 순위	세계 3위, 유럽 1위	세계 4위, 유럽 2위	N/A
CPU	AMD Milan	Intel IceLake / Sapphire Rapid	Sapphire Rapids / Emerald Rapids
GPU	AMD MI250X	Nvidia A100	Nvidia H100

* https://eurohpc-ju.europa.eu/about/our-supercomputers_en

<표 7> EuroHPC JU 미드레인지 시스템 구축 현황

구분	Meluxina	Vega	Karolina	Discoverer	Deucalion
구축 국가	룩셈부르크	슬로베니아	체코	불가리아	포르투갈
구축 년도	2021	2021	2021	2021	2023
실측 성능	12.81 PF	6.92 PF	9.59 PF	4.51 PF	7.22 PF
이론 성능	18.29 PF	10.05 PF	15.69 PF	5.94 PF	10 PF
특징	FPGA 포함	-	-	-	ARM 노드 포함
CPU	AMD EPYC	AMD EPYC	AMD EPYC	AMD EPYC	ARM A64FX AMD EPYC
GPU	NVIDIA A100	NVIDIA A100	NVIDIA A100	-	Nvidia A100
스토리지	(플래쉬) 20PB	NVMe Luster(1PB), 대용량 Ceph(23PB)	-	2PB	NVME 430 TB 대용량 10.6 PB
응용분야	계산과학, AI, 빅데이터/ HPDA	계산과학, AI, 빅데이터/ HPDA	계산과학, AI, 빅데이터/ HPDA	바이오, 생태학, 재료공학	계산과학, AI, 빅데이터

- HPC 인프라에서 고급 컴퓨팅 시스템(특히 양자 컴퓨터 및 양자 시뮬레이터)을 통합하는 하이브리드 컴퓨팅 인프라
- 유럽의 개인 사용자를 위한 안전한 클라우드 기반 HPC 및 데이터 인프라
- 과학자, 산업체 및 공공 부문을 위한 유럽 공공 데이터 공간을 기반으로 하는 HPC 기반 저장 서비스
- 차세대 기술 빌딩 블록(하드웨어 및 소프트웨어) 및 엑사스케일 및 포스트-엑사스케일 시스템을 위한 혁신적인 HPC 아키텍처로의 통합
- 차세대 슈퍼컴퓨터에 최적화된 새로운 알고리즘, 코드 및 도구를 갖춘 HPC 응용 프로그램 및 HPC 소프트웨어 산업화의 우수 센터
- 주요 산업 분야의 HPC 및 데이터 애플리케이션 및 서비스를 위한 대규모 산업 파일럿 테스트 베드 및 플랫폼
- 국가 HPC 우수센터(Centers of Excellence), 중소기업을 포함한 산업 혁신을 위한 특화 서비스 및 리소스를 통한 HPC의 사용 범위 확장
- HPC 기술 및 노하우를 보유한 유럽 인력 양성
- 회원국 전체 공익 분야에서 데이터 저장, 처리 능력 및 새로운 서비스 강화

- **(범 유럽차원의 Top500 현황)** 비EU회원국인 노르웨이, 스위스, 영국을 포함하여 총 133대가 포함되어
합계성능 1359.8 PF(점유율 20.0%)임

<표 8> EU 슈퍼컴퓨터 인프라 현황(TOP500, 22.6)

구분	대표 보유기관	대표 활용분야	보유대수	합계성능(실측성능)
산	Eni S.p.A, Total Exploration	에너지(석유, 가스)	36	209.4 PF
학	GENCI, CINECA	범용	47	494.1 PF
연	FZJ, HLRS, DKRZ	에너지, 기상·기후, 우주	41	611.5 PF
벤처/기타	Atos, HPE	인터넷 서비스	9	44.8 PF
합계			133	1359.8 PF

3. 중국, 일본, 한국의 최신 동향

- **(중국과 일본)** Top500에서 1위 시스템은 미,중,일이 교대로 차지하고 있음. 중국은 Top500에 등재하지 않은 두 대의 엑사급 슈퍼컴퓨터를 이미 구축하여 가동 중이다. 일본은 현재 세계 2위인 Fugaku 시스템의 차기 시스템에 대한 논의를 진행 중인데, 시기적으로 볼 때 이 시스템은 수 엑사플롭스 슈퍼컴퓨터가 될 것
- Top500 순위 중 성능 면에서는 미국이 45.8%, 일본이 12.5%, 중국이 8.9%를 차지해 이들 3개 국가가 전체의 67.2%를 차지했으며, 수량에서는 중국이 134대(26.8%), 미국이 120대(30.0%), 독일이 36대(7.2%) 순서임
- **(대한민국)** 국내의 경우 KISTI와 기상청 외에도 삼성전자, SK텔레콤, 광주과기원, MKO의 슈퍼컴퓨터가 최근 Top500 순위에 포함되었다. 빅데이터 분석과 인공지능의 중요성이 커지면서 국내에서도 민간기업들이 슈퍼컴퓨터를 본격적으로 도입하고 있는 것을 확인할 수 있다. 우리나라의 슈퍼컴퓨터는 총 8대가 포함되어 국가별 보유 대수 순위 9위에 랭크되었음
- **(KISTI 6호기)** 과학기술정보통신부와 KISTI를 중심으로 슈퍼컴퓨팅 생태계를 확장하는 노력이 계속되고 있다. 2018년 상반기에 세계 11위였던 누리온 시스템은 서비스 개시 3년 6개월이 지나면서 순위가 49위까지 하락했다. 이에 정부는 2024년 상반기 가동을 목표로 이론성능 600 페타플로스(0.6 엑사플롭스)의 슈퍼컴퓨터 도입에 2929억원을 투자하기로 결정하였다. CPU와 GPU를 모두 가진 하이브리드 방식으로 구축될 차기 시스템은 서비스가 개시될 2024년 하반기 기준으로 세계 Top10 안에 드는 슈퍼컴퓨터이다. 차기 시스템은 기존의 시뮬레이션 연구는 물론이고 GPU 가속기의 장점을 살려 초거대 인공지능 연구도 활용될 예정임

4. 슈퍼컴퓨팅 최신 기술 동향 및 특허 분석

- **(다양한 가속기의 발전)** CPU 성능 향상의 제약 때문에 GPU, FPGA, TPU, 뉴로모픽 칩, 양자 칩 등 다양한 형태의 가속기를 함께 사용하게 될 것
 - GPU를 채용한 하이브리드 방식의 컴퓨터가 늘고 있는 것은 최근 발표된 Top500 순위에서도 확인할 수 있다. 상위 10개 시스템 중 중국과 일본의 시스템 3개를 제외한 나머지 7개는 모두 GPU를 이용해서 컴퓨터의 성능을 높이고 있다. 필요전력과 요구성능, 가격 등을 고려하면 현 시점에서 유일한 해결 방안
 - GPU 가속기 시장은 Nvidia가 대부분의 시장을 독점하고 있었으나, 최근 들어 경쟁업체들의 도전이 거세다. AMD에서 출시한 MI250X GPU는 Top500에서 1위인 프론티어, 3위인 루미, 12위인 에이다스트라(Adastra) 등 상위권 시스템에서 많이 채용되었음
 - 단일 칩에 CPU와 GPU 모두 넣어서 만든 AMD MI300 APU는 2023년 구축될 로렌스 리버모어 국립연구소의 슈퍼컴퓨터 ‘엘 캐피탄’에 탑재 예정
 - 인텔 역시 GPU 시장에 진출하기 위해 ‘폰테 베키오’란 GPU를 출시하여 아르곤 국립연구소에서 구축 중인 테라급 슈퍼컴퓨터 ‘오로라’에 탑재하여 테스트하고 있다. 폰테 베키오 이을 후속 GPU인 ‘리알토 브리지’를 2023년에 중반에 내놓겠다는 계획을 발표했다. 최대 160개의 코어를 탑재하게 될 리알토 브리지는 전작 대비 30%의 성능 향상이 있을 것이라 함
- **(저전력화·고집적화)** 데이터의 대규모화, 초고속처리 및 집중화 요구로 초고성능화, 고에너지 효율(저전력), 고집적화로 발전되는 추세
- **(인공지능의 비중 확대)** 전문가들의 분석에 따르면 인공지능 분야에 필요한 컴퓨팅 자원은 1.1~1.4년 마다 10배씩 증가하는 추세이다. 인공지능 분야의 혁신을 이끈 AlexNet, 알파고, BERT, GPT-2/3/4 등의 주요 모델은 슈퍼컴퓨터급의 대형 계산 자원을 활용할 수 있었기에 가능했고, 이런 추세는 기술 경쟁에 따라 더 심화될 것이다. 2018년 발표된 구글의 ‘알파고 제로’의 경우 딥러닝의 돌파구가 되었던 AlphaNet 보다 30만배 더 많은 계산을 필요로 한다. 2020년에 등장한 1,750억개의 변수를 갖는 자연어처리 GPT-3 모델은 10,000개 이상의 GPU를 필요로 한다. ChatGPT의 등장과 적용 확대로 인공지능에 필요한 컴퓨팅 자원 수요는 예사보다도 더 빠르게 증가할 것으로 보임
- **(양자컴퓨터)** 최근 들어 유럽의 대표적인 슈퍼컴퓨팅센터들이 경쟁적으로 양자컴퓨터를 도입하고 있음
 - 독일 울리히 슈퍼컴퓨팅센터(Juelich Supercomputing Centre, JSC)는 2022년 1월부터 D-Wave Systems 사의 5천 큐비트 양자 어닐러를 설치하여 서비스 시작. 유럽에 설치된 첫 양자컴퓨터임
 - 울리히 슈퍼컴퓨팅센터와 프랑스의 GENCI는 중성원자 방식인 파스칼(Pascal)의 양자컴퓨터를 도입해 내부적으로 사용하고 있음

- 핀란드의 슈퍼컴센터인 CSC는 루미(LUMI) 슈퍼컴퓨터와 5-큐비트의 양자컴퓨터 헬미(HELMI)를 연결한 서비스를 제공하기 시작했다. 일반적인 목적의 슈퍼컴퓨터와 최적화 분야에서 강점을 갖춘 양자컴퓨터의 장점을 결합시킴으로서 더 빠르게 문제를 풀 수 있을 것으로 기대됨
- EuroHPC JU에선 체코, 독일, 스페인, 프랑스, 이태리, 폴란드 6개국에 양자컴퓨터를 설치하여 서비스하겠다는 계획을 발표했다. 각국의 슈퍼컴퓨팅센터에서 향후 수년에 걸쳐 양자컴퓨터를 개발하고 설치하고 서비스하는 대형 프로젝트임
- 양자컴퓨터를 이용하면 기존의 컴퓨터로 풀 수 없었던 많은 문제가 해결될 수 있을 것이라 기대된다. 다만, 양자컴퓨터는 슈퍼컴퓨터와 적대적인 관계 아니라 서로 보완하는 관계이다. 당분간 수치 연산을 많이 필요로 하는 분야에서 계속 슈퍼컴퓨터가 사용되고, 조합최적화와 기계학습 분야에서 양자컴퓨터가 적용하는 방식으로 서로 역할분담을 하게 될 것임
- **(기술특허분석)미국,유럽,PCT,일본,중국,한국을대상으로Wintelips데이터베이스(00.1.1~21.12.10.)를 이용하여 초고성능 컴퓨팅 관련 특허 검색·분석 실시**
 - 특허 검색을 통해 도출된 총 26,168건의 특허 중 기술연관성이 낮은 19,973건을 제외한 6,195건의 유효 특허에서 수적으로는 중국이 압도적으로 1위를 기록
 - 기술 분야로는 클라우드 관련 특허(A1)가 2/3를 차지함

<표 9> 산업분야별 특허 초고성능 컴퓨팅 기술 분야 주요기술 분야별 유효특허 DB

주요기술	한국	미국	일본	유럽	중국	PCT	계
(A1) Cloud Computing (클라우드컴퓨팅)	75	634	26	90	3,209	128	4,162
(A2) HPC 및 Super Computing(고성능 및 슈퍼컴퓨팅)	17	139	12	19	392	41	620
(A3) Distributed Cluster(분산클러스터)	7	14	1	2	212	6	242
(A4) MPI(메시지 전달 인터페이스)	7	23	0	6	115	10	161
(A5) Scientific Computing(계산과학)	6	10	5	3	365	5	394
(A6) Autonomic Computing(자율컴퓨팅)	25	167	30	18	33	44	616
합계	13	987	74	138	4,625	234	6,195

- 미국 등록 특허로 한정하면 기술력지수에서 미국이 압도적 1위

<표 10> 산업분야별 특허 최고성능 컴퓨팅 기술 분야 미국 등록 특허 상위 국가 주요 특허지수 현황

국가	미국 특허 등록 건수	기술 경쟁력 판단 지수		기술 질적 수준 판단 지수	
		기술력 (TS) 지수	시장확보 (PFS) 지수	특허 인용도 (CPP) 지수	특허영향 (PII) 지수
미국	522	550.76	2.63	20.73	1.06
중국	24	14.46	3.50	11.83	0.60
독일	23	32.07	1.74	27.39	1.39
사우디	15	8.70	4.20	11.40	0.58
인도	7	2.65	1.71	7.43	0.38
아일랜드	6	2.49	4.17	8.17	0.42
한국	4	0.20	1.75	1.00	0.05

* 주요 특허지수 분석은 통상 미국 등록 특허를 대상으로 분석함

맺음말

지금 전 세계는 더 빠른 슈퍼컴퓨터를 만들고, 이를 이용한 시뮬레이션 연구와 인공지능 연구를 통해 경쟁력을 키우기 위해 총성 없는 전쟁을 벌이고 있다. 미국은 슈퍼컴퓨터 관련 주요 제품을 중국에 수출하는 것을 금지했고, 일본과 유럽은 미국 의존도를 낮추기 위해 자체 기술로 슈퍼컴퓨터를 개발하기 위해 노력하고 있다. 최근의 Top500 순위를 보면 상위권에 자리 잡고 있는 대부분의 슈퍼컴퓨터들이 GPU를 가속기로 활용하고 있는 것을 볼 수 있다. 소요전력과 발열, 설치 면적 등의 장점 때문에 가속기 이용은 피할 수 없는 흐름으로 보인다. 향후에는 GPU는 물론이고 뉴로모픽 컴퓨터와 양자컴퓨터 등도 가속기 방식으로 슈퍼컴퓨터에 통합될 것이라 예측하는 전문가가 많다.

참고문헌

- 4차 산업혁명 퀀텀점프를 위한 국가초고성능컴퓨팅 혁신전략(2021.5, 비상경제 중앙대책본부)
- <http://www.top500.org>
- 국가초고성능컴퓨팅 서비스 선진화 사업(2023.2, KISTI)
- 미래를 읽다 과학이슈 11 Season13, 이식 외(2023.2, 동아엠앤비)

저 자 이 식
KISTI 국가슈퍼컴퓨팅본부
책임연구원
T. 042-869-1039
E. siklee@kisti.re.kr

함 재 군
KISTI 국가슈퍼컴퓨팅본부
초고성능컴퓨팅정책센터
책임연구원
T. 042-869-0580
E. jaehahm@kisti.re.kr

KISTI ISSUE BRIEF 제55호

발 행 일 2023. 05. 23.

발 행 인 김재수

편 집 위 원 조민수, 최희석, 최장원, 정한민, 함재군,
이준영, 이상환, 곽영

발 행 처 34141 대전광역시 유성구 대학로 245
한국과학기술정보연구원 정책연구실
<https://www.kisti.re.kr>

I S S N 2635-5728