

# KISTI 슈퍼컴퓨터 5호기 도입 및 활용

한국과학기술정보연구원 국가슈퍼컴퓨팅본부

강 지 훈, [jhkang@kisti.re.kr](mailto:jhkang@kisti.re.kr)



# KISTI 슈퍼컴퓨팅본부

- 국가초고성능컴퓨팅 활용 및 육성에 관한 법률 및 시행령으로 지정된 국가 초고성능컴퓨팅 센터 (KISTI)
- 8개의 연구센터로 구성

국가슈퍼컴퓨팅본부

슈퍼컴퓨팅서비스센터

과학기술연구망센터

대용량데이터허브센터

과학기술사이버안전센터

계산과학응용센터

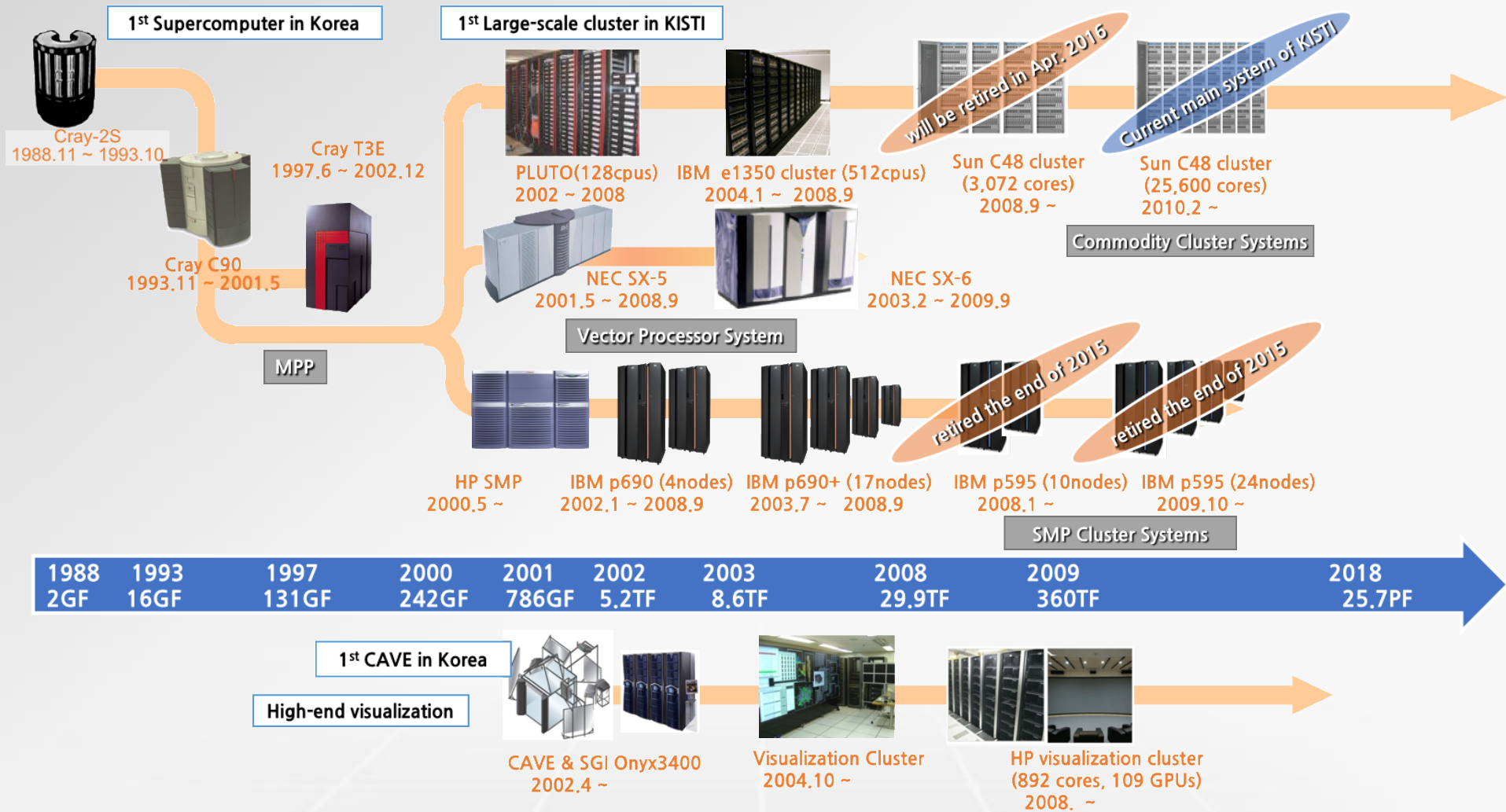
계산과학플랫폼센터

슈퍼컴퓨터개발센터

슈퍼컴퓨팅클라우드센터



# KISTI 슈퍼컴퓨터 역사



# 슈퍼컴퓨터 5호기 구축 배경 (I)

근거 1 「국가초고성능컴퓨터 활용 및 육성에 관한 법률」 제9조, 제11조

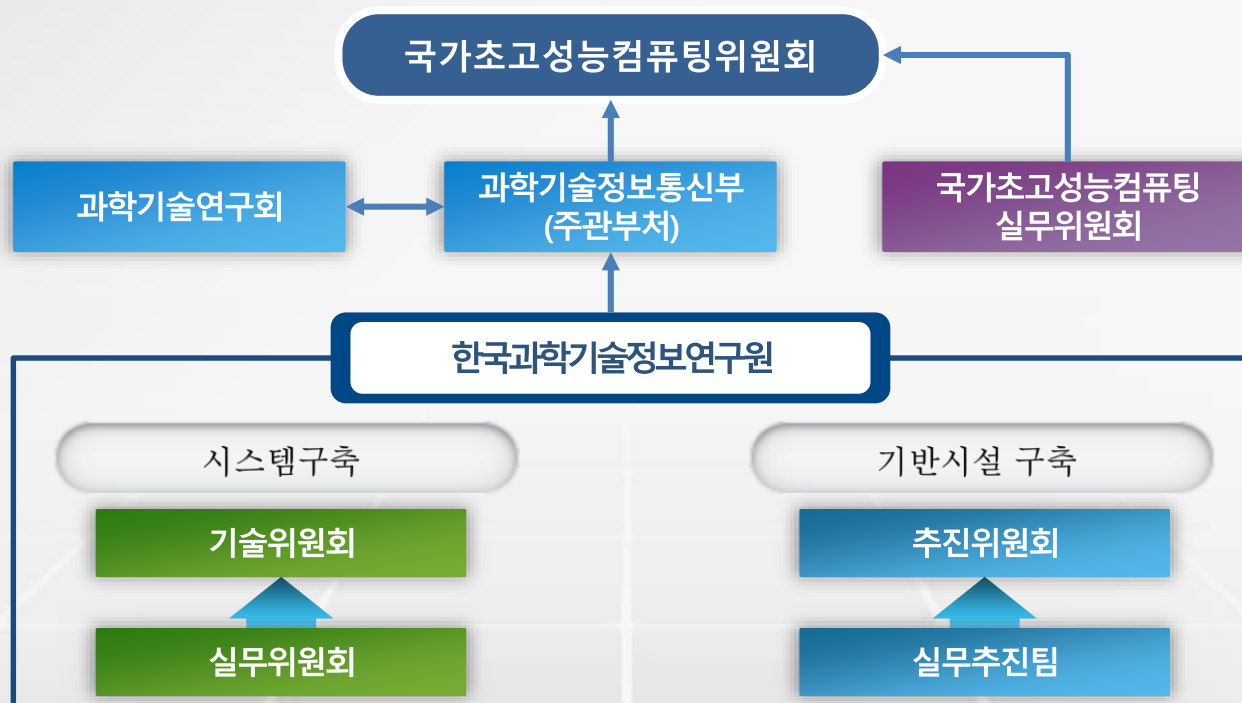


국가 초고성능컴퓨팅센터가  
세계적 수준의 국가초고성능컴퓨팅 자원 확보 및 운용

근거 2 「제1차 국가초고성능컴퓨팅 육성 기본계획(’13년~’17년)」 수립



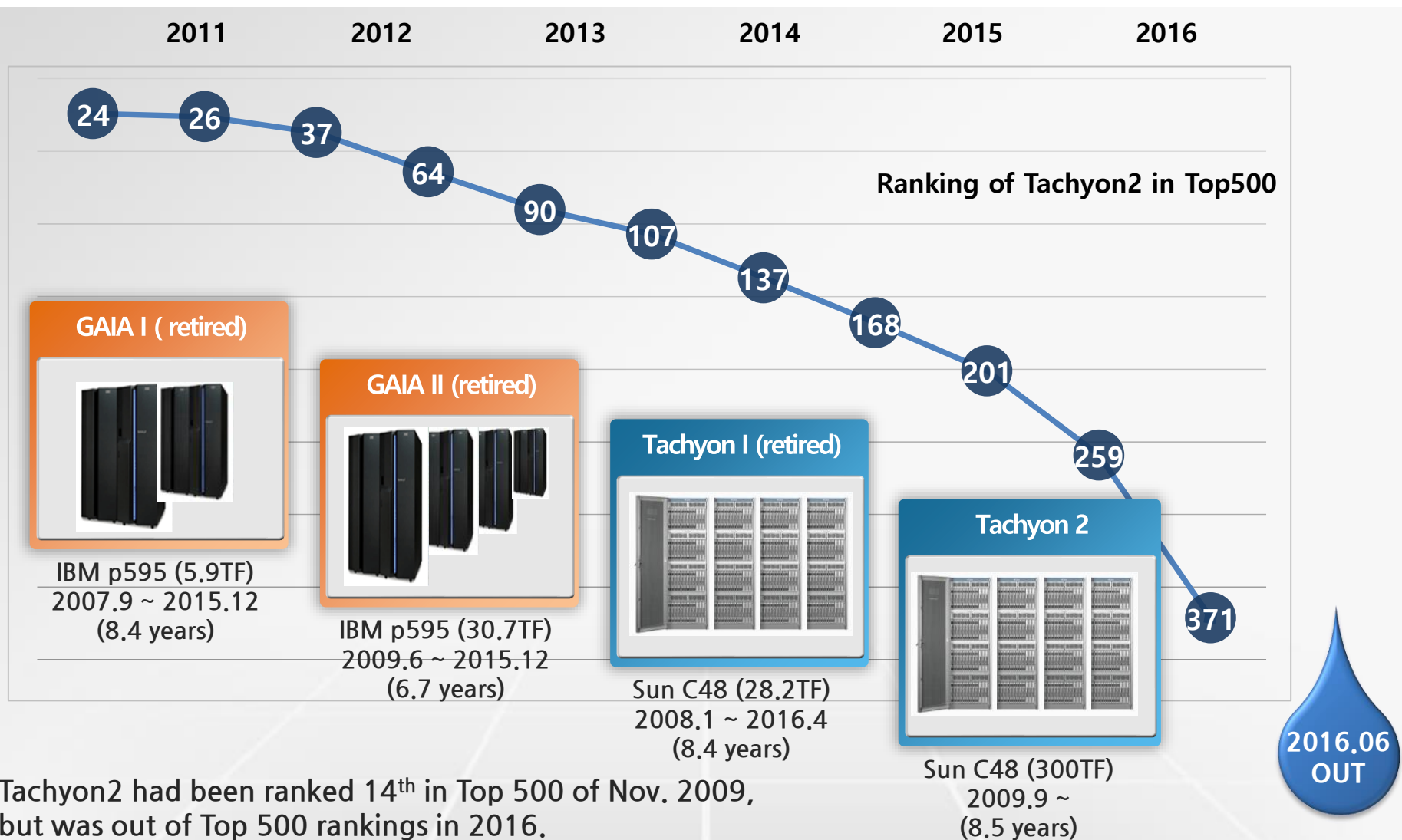
‘초고성능컴퓨팅 서비스 기반 구축’을 3대 전략 중 하나로 채택하고  
관련 정책과제를 제시



# 슈퍼컴퓨터 5호기 구축 배경(II)



Current KISTI-4 system, Tachyon II, was procured in 2009 and served for 8 years.



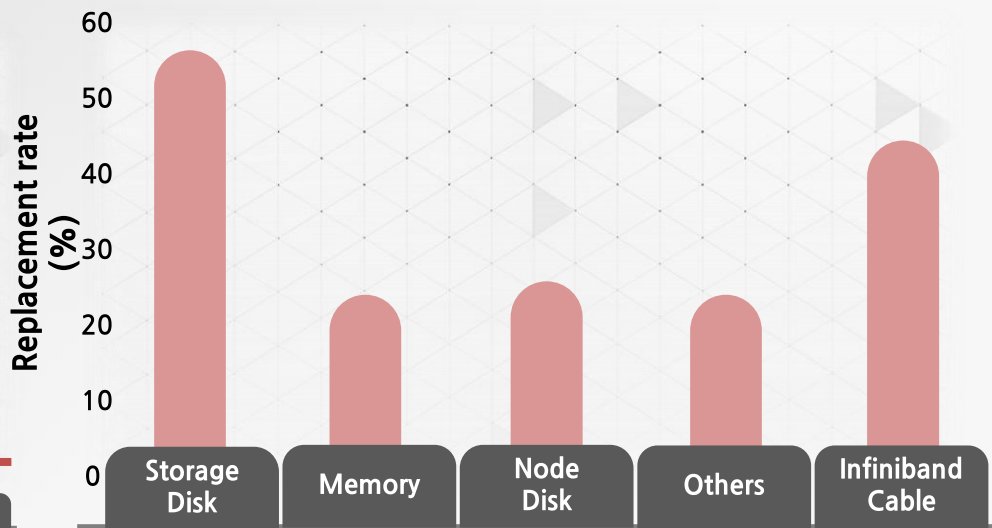


# 슈퍼컴퓨터 5호기 구축 배경(III)

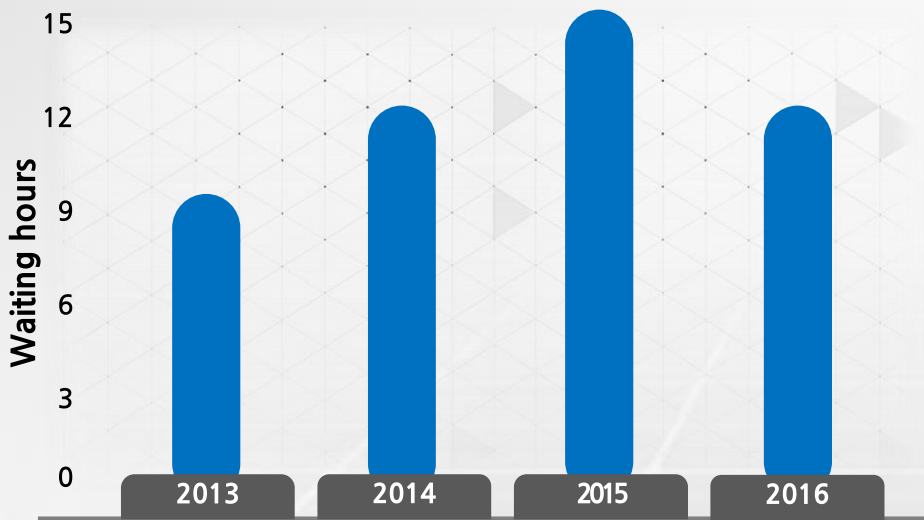
## Operating cost beyond the residual value



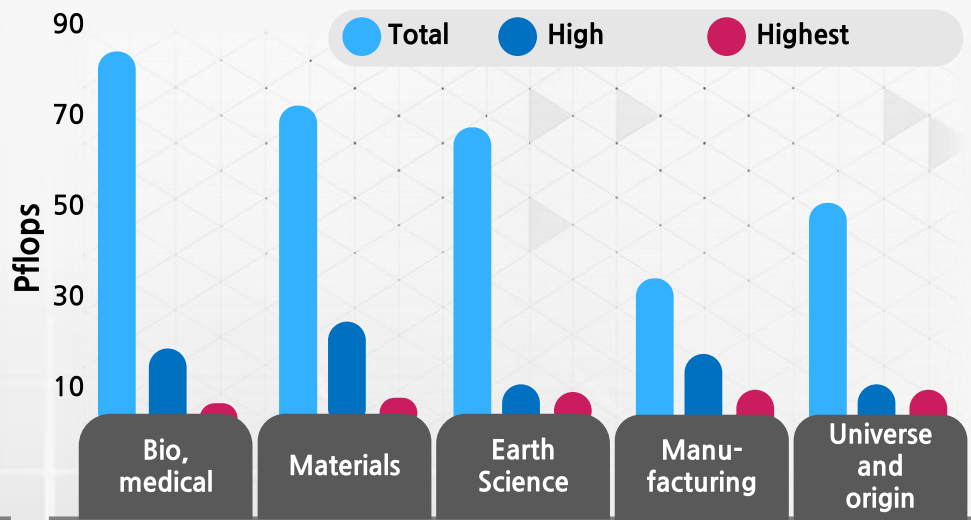
## Increase of replacement due to hardware failure



## Increased waiting time due to limited resources



## Insufficient infrastructure supply to demand



# 슈퍼컴퓨터 5호기 구축 개요

## 사업목적

국가 초고성능컴퓨팅 인프라 선진화를 위한 리더십 시스템 도입

총 이론성능 25.7PF (세계 10위권)

## 기간/예산

2016년 ~ 2020년 (5년) / 총 908억원

[백만원]

단위사업	세부과제	세세부과제	2016	2017	2018	2019	2020	합계
국가 리더십 초고성능 컴퓨터 구축 활용 사업	국가 리더십 초고성능컴퓨터 시스템 구축 [사업 1]		-	14,675	14,675	14,675	14,675	58,700
	초고성능컴퓨터 운영환경 구축 [사업 2]	안정적 운영 환경을 위한 기반시설 구축	8,000	17,839	-	-	-	25,839
		초고성능컴퓨터 응용소프트웨어 및 실행환경 구축	-	2,235	1,345	1,345	1,345	6,270
	소계 (단위사업 1)		8,000	34,749	16,020	16,020	16,020	90,809

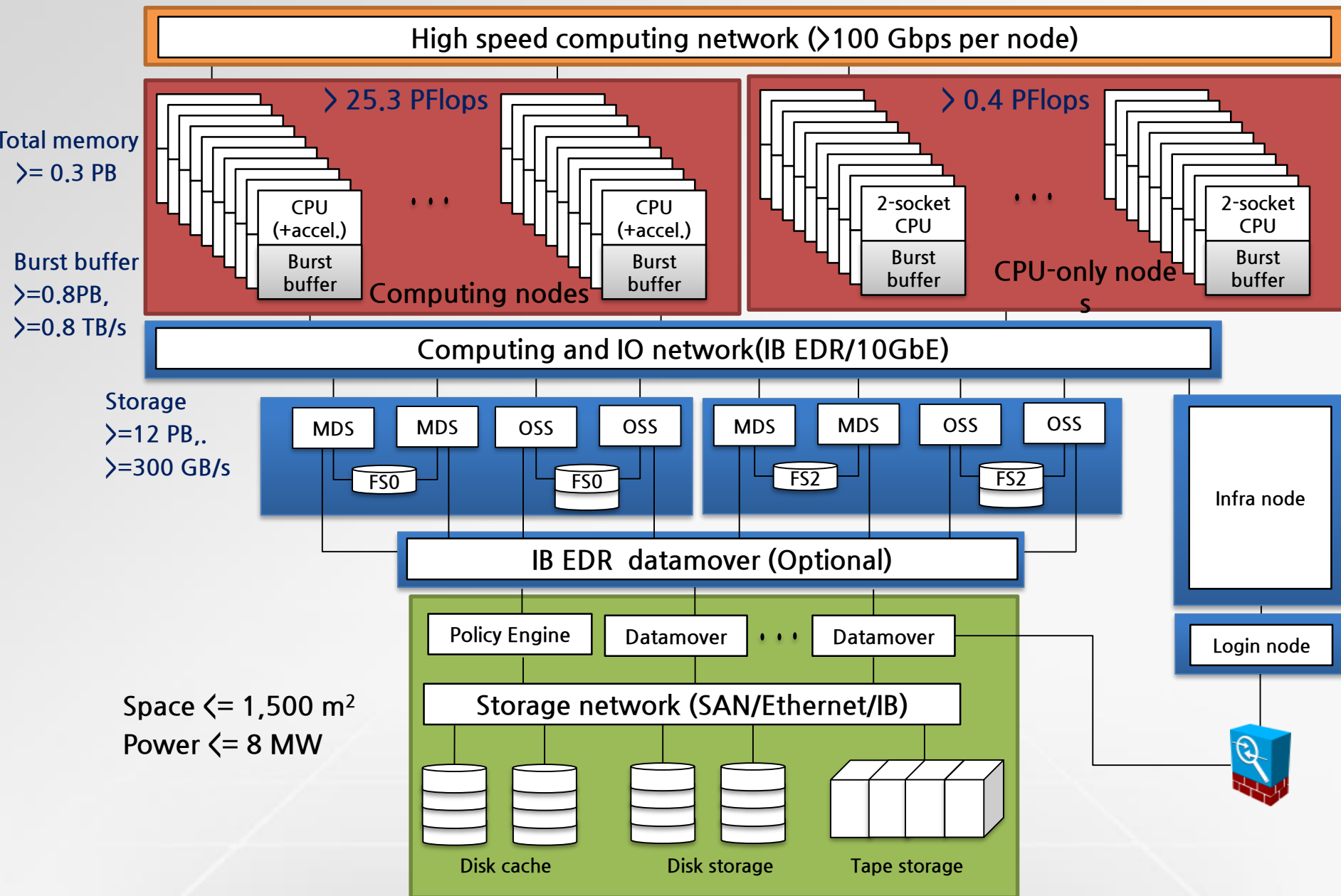
## 사업내용

국가 리더십 초고성능컴퓨팅 시스템 구축

초고성능컴퓨터 운영환경 구축

- 안정적 운영 환경을 위한 기반시설 구축
- 초고성능컴퓨터 응용 소프트웨어 및 실행환경 구축

# 슈퍼컴퓨터 5호기 시스템 요구 사항





# 슈퍼컴퓨터 구축 추진 경과

- '15.06 ○ ▶ Building construction completed
- '15.07.07 ○ ▶ Preliminary feasibility study approved
- '16.03 ○ ▶ RFI and BMT announcement
- '16.07 ○ ▶ RFP and transfer to PPS(調達廳)
- '16.10~12 ○ ▶ Bidding
- '17.01 ○ ▶ Failure in bidding and RFP modification
- '17.02~05 ○ ▶ Bidding
- '17.06~07 ○ ▶ Cray winning bid and negotiation
- '17.08 ○ ▶ **Contract (49M USD)**
- '17.11 ○ ▶ Pilot system(16nodes) delivery completed
- '17.12 ○ ▶ Main system delivery started
- '18.01~02 ○ ▶ Main system construction
- '18.03~04 ○ ▶ Software installation and system setup
- '18.05~06 ○ ▶ BMT verification and system inspection
- '18.07~09 ○ ▶ NOA and early access



Efficiency: PUE < 1.3



Performance: 25.7PFlops



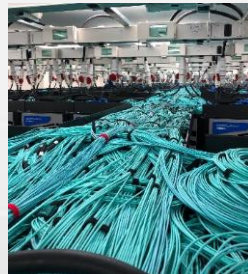
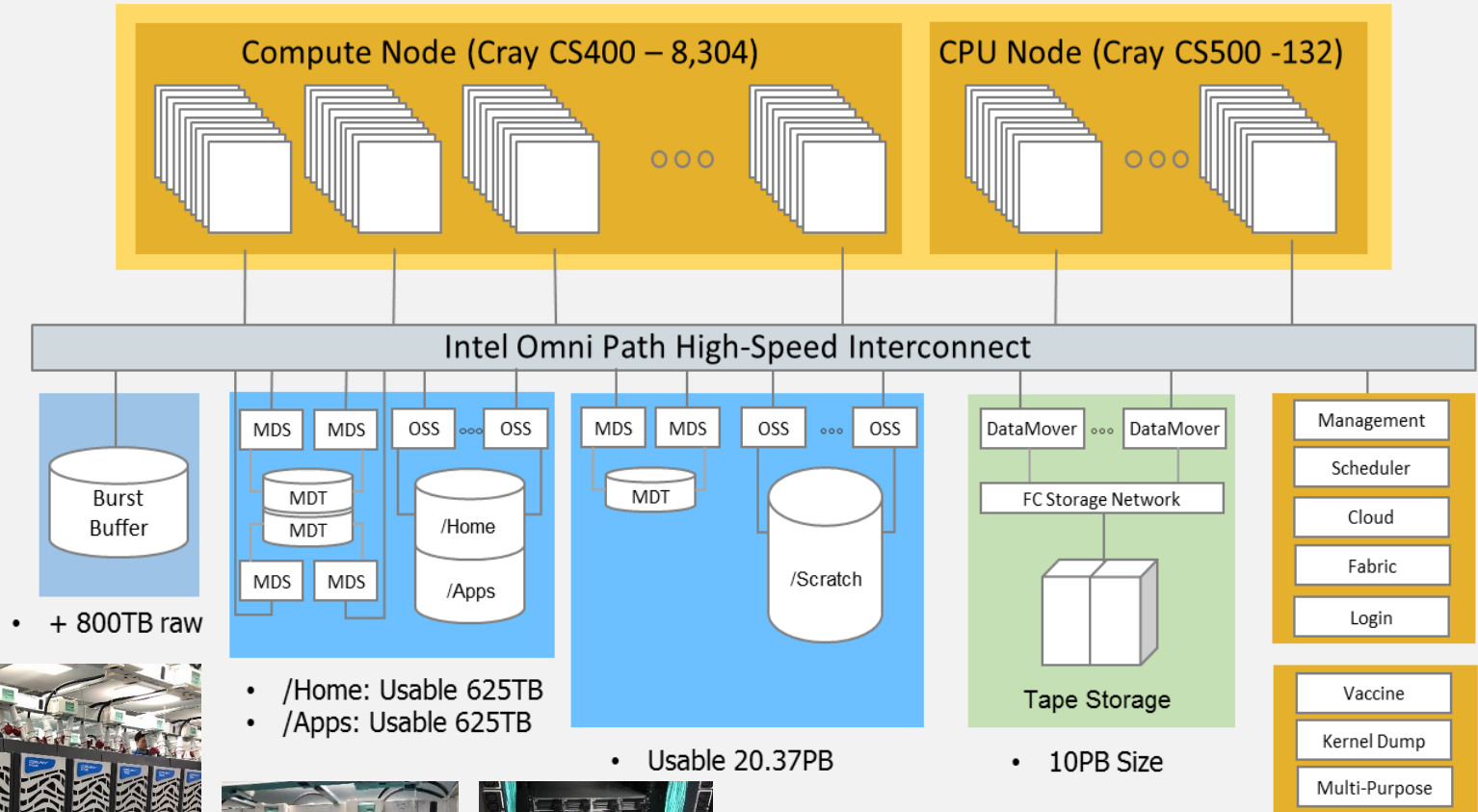
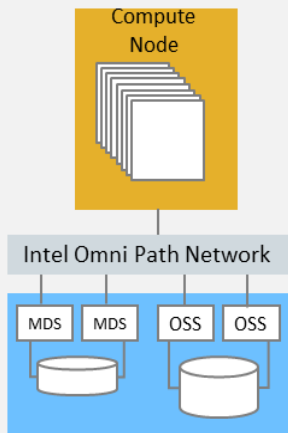
# 슈퍼컴퓨터 5호기 개요

## Main System

- 1 x Intel KNL (3.06TFLOPS/node peak)
- 6 x 16GiB DDR4-2400 DIMM(96GiB)
- 1 x Intel OPA 100Gbps HFI

- 2 x Intel SKL 6148 (3.072TFLOPS/node peak)
- 12 x 8GiB DDR4-2666 DIMM (96GiB)
- 1 x Intel OPA 100Gbps HFI

## Pilot -> Testbed





# 슈퍼컴퓨터 5호기 세부 구성(I)

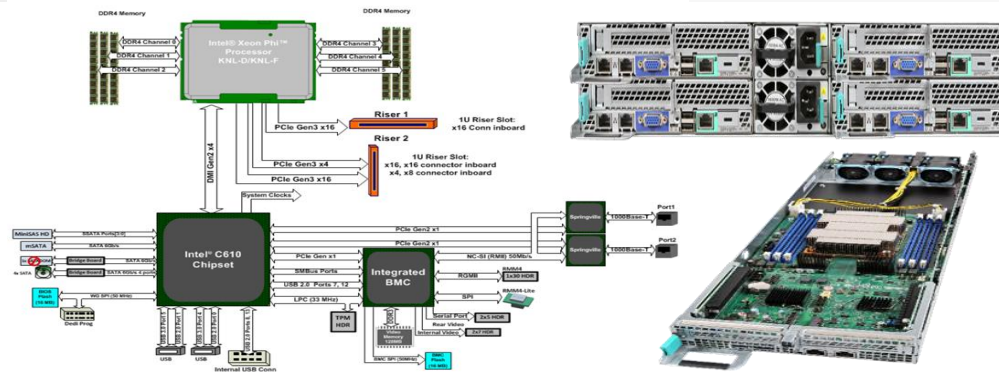


Theoretical performance 25.7PF = 25.3PF CS400 (KNL) + 0.4PF CS500 (SKL)

## Computing nodes

Cray 3112-AA000T(2U enclosure), 8,305 KNL Computing modules

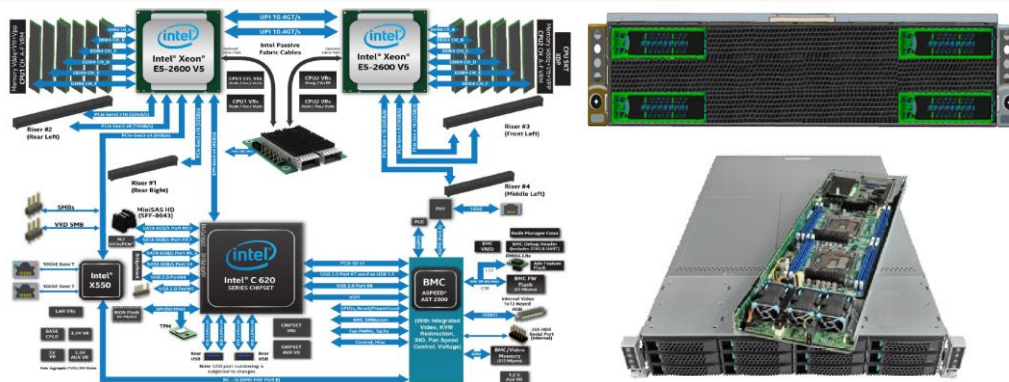
- 1x Intel Xeon Phi KNL 7250 processor
- 96GB (6x 16GB) DDR4-2400 RAM
- 1x Single-port 100Gbps OPA HFI card
- 1x On-board GigE (RJ45) port



## CPU-only nodes

Cray 3111-BA000T(2U enclosure), 132 Skylake Computing modules

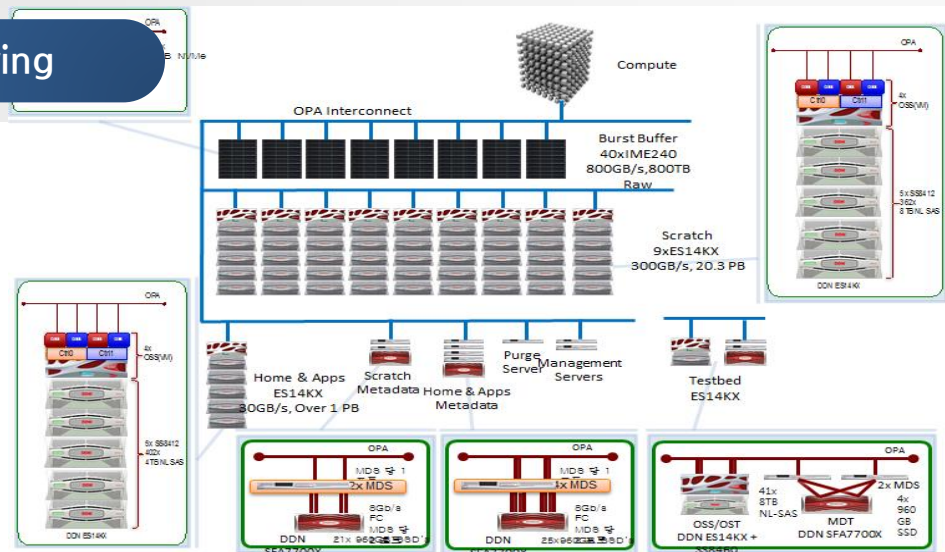
- 2x Intel Xeon SKL 6148 processors
- 192GB (12x 16GB) DDR4-2666 RAM
- 1x Single-port 100Gbps OPA HFI card
- 1x On-board GigE (RJ45) port



# 슈퍼컴퓨터 5호기 세부 구성(II)

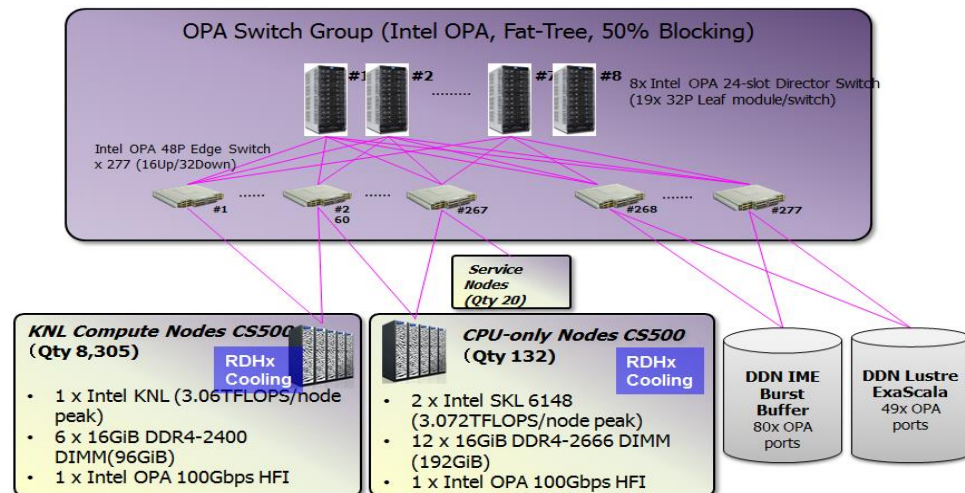
## Storage 20PB SFS@300GB/s, 10PB Archiving

- Global scratch: 20PB, 0.3TB/s  
(DDN ES14KX 9ea, 360 x 8TB disk each)
- Home and application directory 1PB
- NVMe Burst Buffer: 0.8PB, 0.8TB/s  
(IME240 40ea 19 NVMe SSD each)
- Cray TSMSF and IBM TS4500



## Interconnect OPA(Omni-Path Architecture), Fat-Tree, 50% Blocking

- Intel OPA High-speed interconnect switch  
274x 48-port OPA edge switches  
8x 768-port OPA core switches
- Bandwidth: 12.3 GB/sec
- Bisectonal Bandwidth : 27 TB/sec
- $10^{-16}$  BER(Bit Error Rate), Adaptive routing



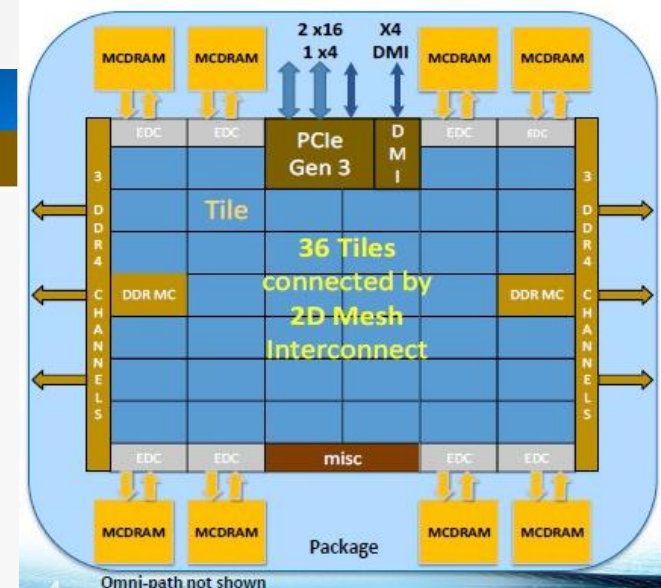
# Knights Landing 매니코어 프로세서



Enabling cost-effective high-performance computing

## Unique features of Knights Landing

- Many-core with low-clock speed
  - 68 cores with 1.4GHz (7250 processor)
- High-bandwidth memory(HBM)
  - 16GB on-chip memory with ~400GB/s
- 512bit vector unit
  - Processing 8 double-precision calculations simultaneously
- High theoretical peak performance
  - 3TFlops per processor, twice of 20-core SKL processor



## But...

- Smaller size of L2 cache per core
  - 0.5GB/core, half of SKL processor, 1GB/core
- No L3 cache
- Insufficient amount of HBM, 0.23GB/core



# Knights Landing 성능 평가

## KNL testbed

Components	Specification
Number of nodes	16
Processor model	Xeon Phi KNL 7250, 1.4 GHz, 68 cores/node
On-package HBM	16 GB/node
Main memory	96GB/node, DDR4 (16 GB×6 channels)
Interconnect	Intel Omni-Path Architecture (OPA)
Theoretical performance	3TFlops / node

## KNL mode

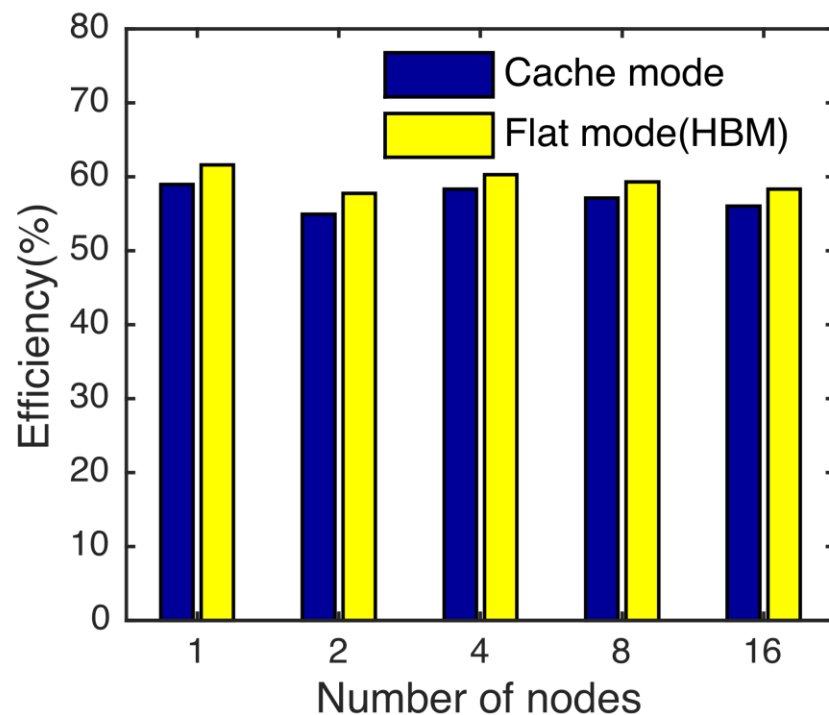
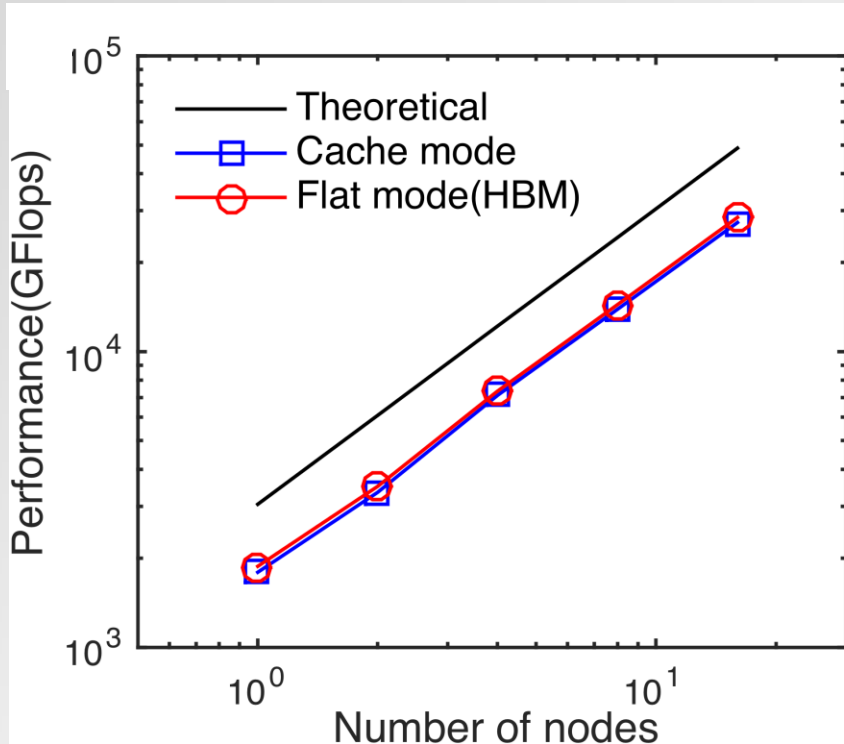
- Cluster mode : Quadrant mode
- Memory mode: Cache mode / Flat mode (HBM and DRAM)

## Tachyon2

Components	Specification
Number of nodes	3200
Processor model	Xeon Nehalem X5570, 2.93GHz, 4*2 cores/node
Main memory	24 GB/node, DDR3
Interconnect	Infiniband 4x QDR
Theoretical performance	100GFlops / node

# 표준 벤치마크 결과 (I)

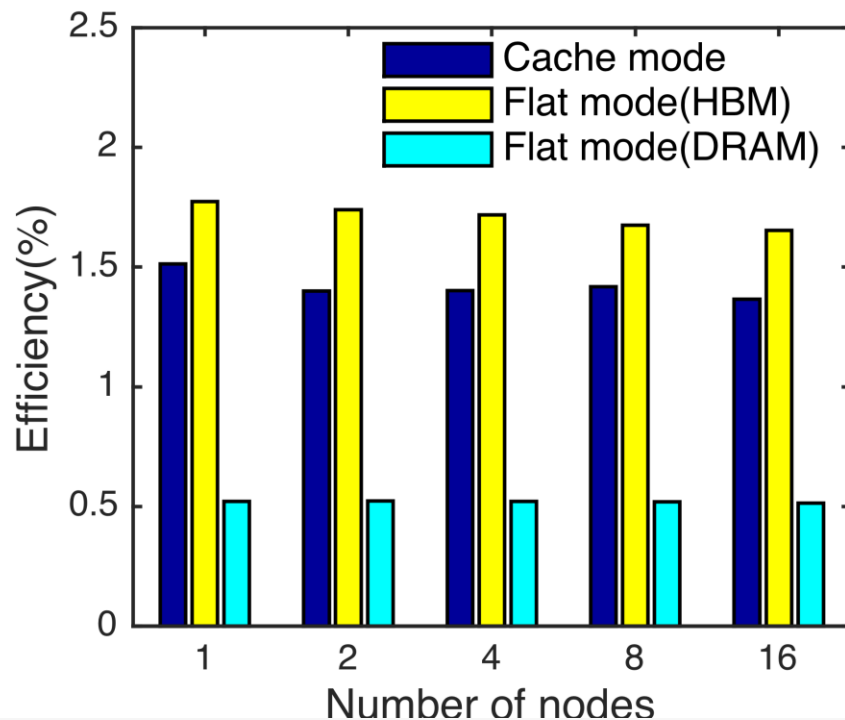
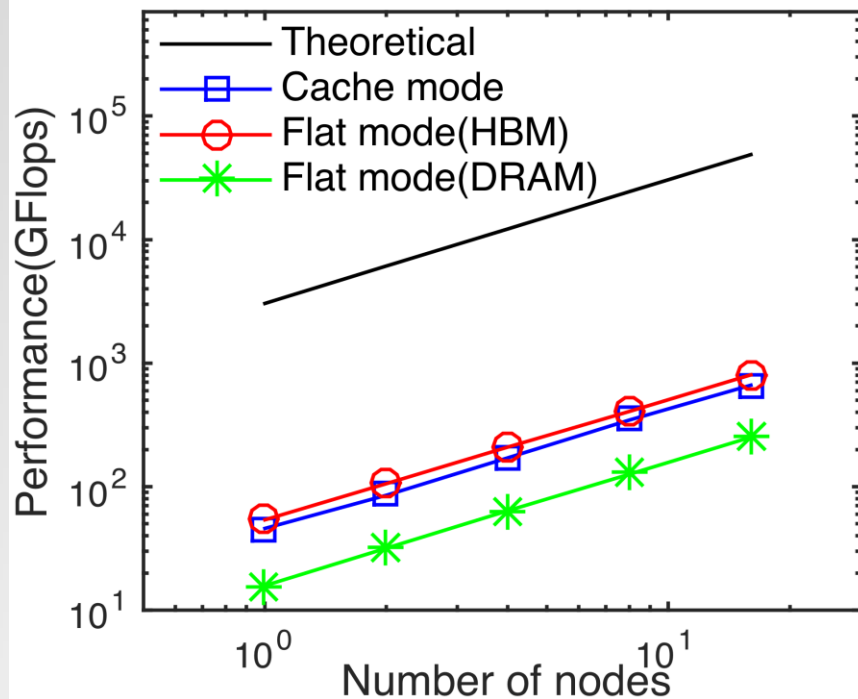
## High Performance Linpack



- Cache mode and flat mode with HBM show almost same performance
- 1.8TFLOps / node is obtained for both memory modes
- Efficiency, the ratio of measured performance to theoretical performance is  $\sim 60\%$

## 표준 벤치마크 결과 (II)

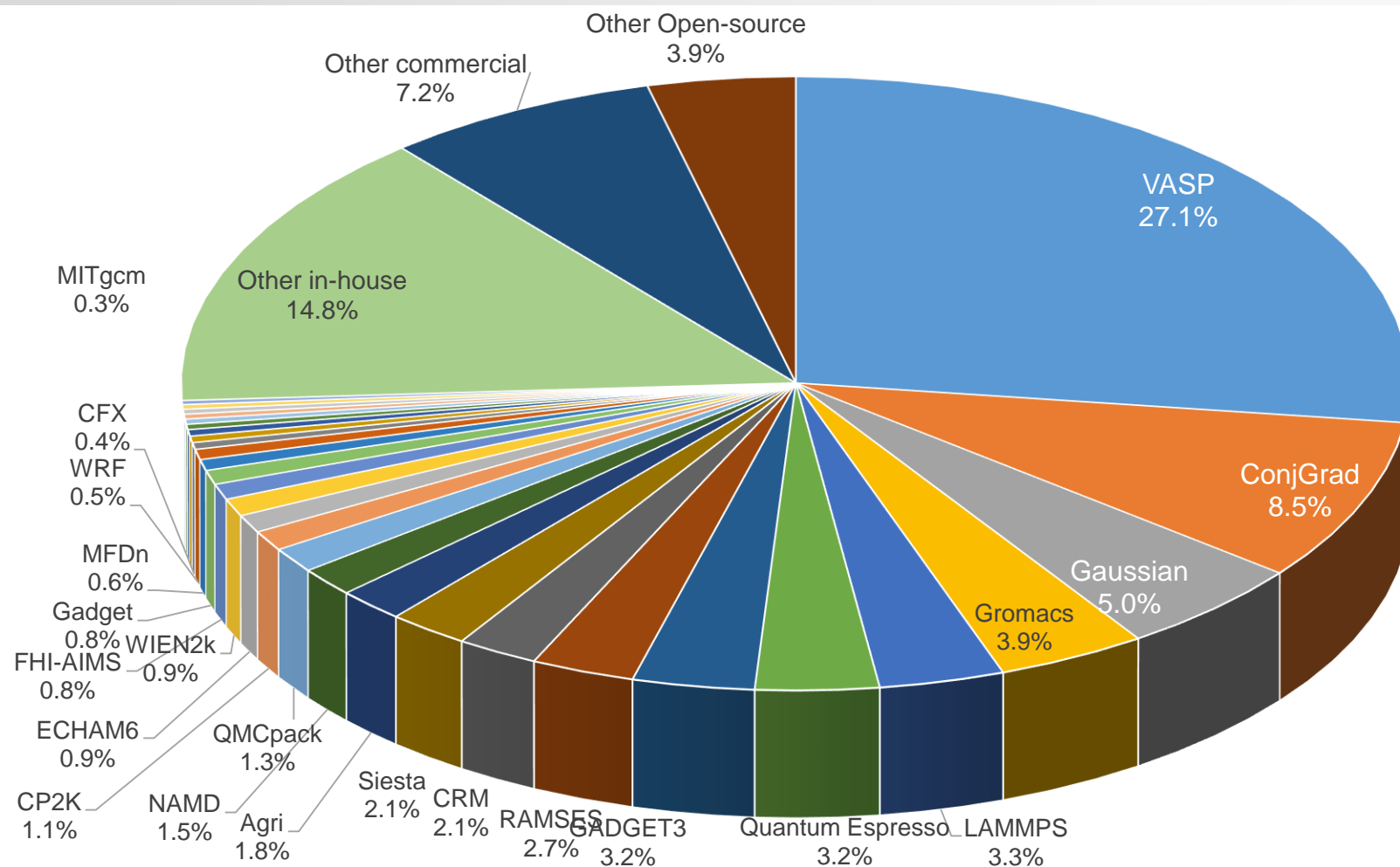
### High Performance Conjugate Gradient



- HPCG performance is much lower than theoretical performance:  $\sim 51$  GFlops / node
- Efficiency, the ratio of measured performance to theoretical performance is  $\sim 1.7\%$
- Cache mode and flat mode show quite large performance difference  $\rightarrow$  HPCG highly depends memory bandwidth.

# 응용소프트웨어 성능 비교

## 슈퍼컴퓨터 4호기 워크로드 분석 및 응용소프트웨어 선정



- Top10 applications occupy ~61% of workload
- Top20 applications occupy ~72% of workload

# 응용소프트웨어 성능 평가 항목

Applications	Algorithm	Domain	Problems
Gaussian	Sparse linear algebra	Quantum mechanics	Electron structure of $[\text{PdCl}_4]^{2-}$ , 88 electrons and def2-TZVP basis set
Quantum Espresso	Sparse linear algebra	Quantum mechanics	Electron structure of gold supercell, 112 Au atoms and 800 bands
Gromacs	N-Body	Molecular dynamics	Solvated Alcohol dehydrogenase(ADH), 134K ADH and 1,563K water atoms
LAMMPS	N-Body	Molecular dynamics	Rhodopsin protein in lipid bilayer, 616K rhodopsin and 407K water atoms
WRF	Structured grids	Weather modeling	Regional model for east-asia, $1501 \times 1201$ grids and 25 vertical layers
OpenFOAM	Structured grids	Fluid dynamics	Lid-driven cavity flow, $80 \times 32 \times 32$ grids

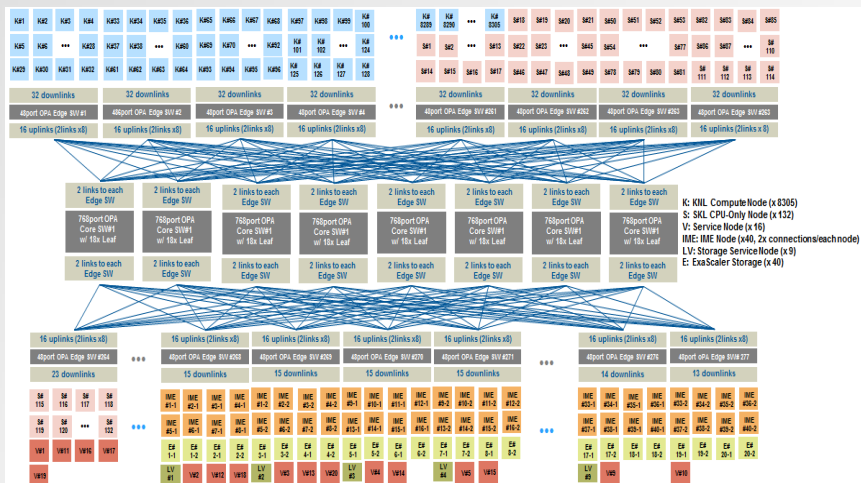
# 슈퍼컴퓨터 운영방향

## 안정성

## 슈퍼컴퓨팅 인프라의 안정적 서비스

거대 문제를 비롯한 HPC의 수요에 슈퍼컴퓨팅 인프라를 조기에 안정화하여 최적의 연구 환경 제공

- 8,300노드 규모의 초대형 클러스터 시스템,
- OPA, Burst buffer, MCDRAM 등의 최신 기술,
- 20PB급 단일 대형 병렬 파일시스템



## 유연성

## 신규 서비스 지원을 위한 유연한 환경

가상화/클라우드 기술을 활용하여 지능 정보 등 다양한 사용자 요구를 지원하는 환경제공

- 계산노드 일부 베어메탈 방식의 클라우드 기술 및 OS레벨 가상화 지원
- CPU-only노드의 일부는 하드웨어 기반 가상화 서비스 환경 제공

HPC

HTC

AI & Big Data

가상화  
(Container, BareMetal)

스케줄러/Provisioning

병렬파일시스템(Lustre)

매니코어 프로세서 시스템

x86 시스템

고성능 인터커넥트

대용량 스토리지

Burst Buffer (고성능 스토리지)



# 슈퍼컴퓨터 5호기 활용 계획

## ■ 국가 전략자원으로서의 활용

- 재난·재해 대응, 중소기업지원, 지능정보 분야 등 국가가 수행해야 할 전략적 중점 분야에 대한 비중 확대
- 초고성능컴퓨터로만 해결 가능한 대형 및 첨단 R&D 분야에 대한 비중 확대

## ■ 국가 중요 공공재로서의 유·무상 배분

### [무상 지원]

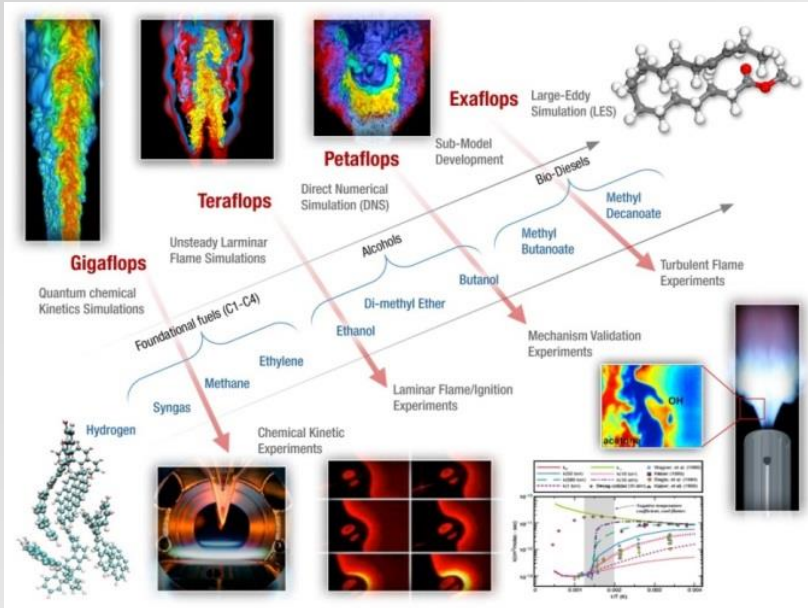
- 국가 차원의 지원이 필요한 기초 및 순수 연구, 도전적인 연구에 대하여 엄격한 동료 평가를 통한 무상 지원
- 초고성능컴퓨팅 분야에서 국가 위상을 제고할 수 있는 관련 기술 개발 등의 분야
- 기타 초고성능컴퓨터 운영과 활성화에 필요한 분야에 대해서는 최소한의 범위 내에서 무상 지원

### [유상 지원]

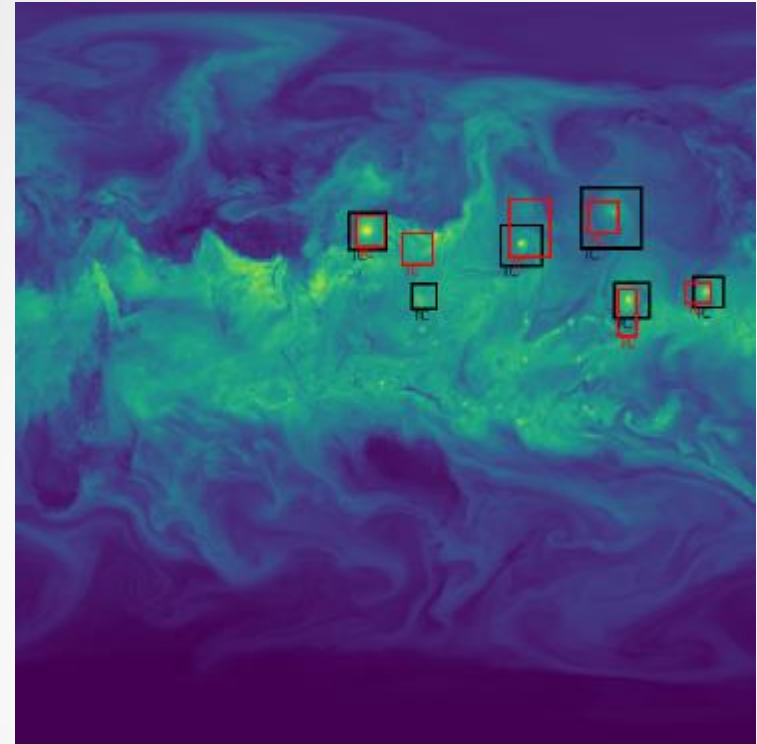
- 정부 지원 국가 R&D과제에 대하여 연구비 내에서 지불 가능한 초고성능컴퓨터 운영비용 수준의 사용료 산정
- 영리 목적 사용에 대하여 초고성능컴퓨터 구축비용과 운영비용을 고려한 사설 서비스 수준의 사용료 산정

# 거대문제 사례 - Capability computing

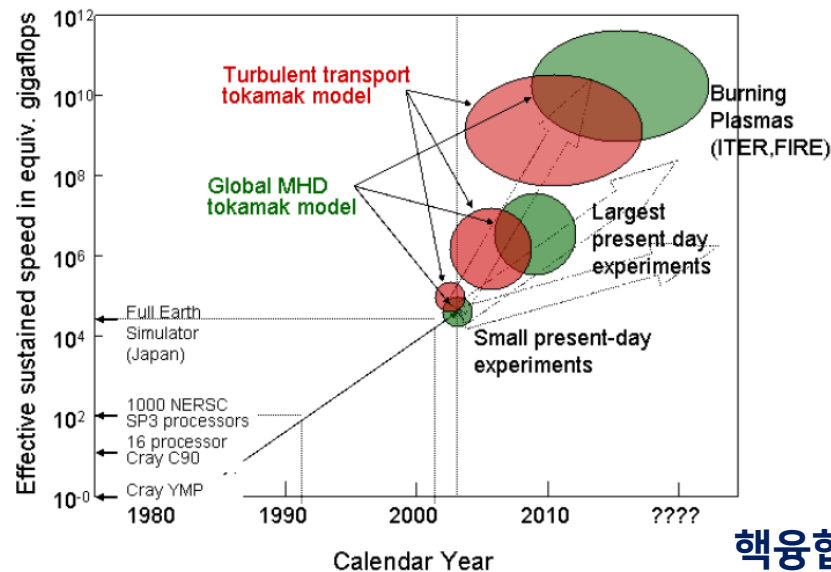
## 연소과학



## 기계학습



CAM5 25-km 해상도 모델 30년 시뮬레이션 단일 작업  
 -> 100 TB 데이터 산출  
 -> 심각한 기상 이벤트 탐색 필요  
 (NERSC CORI 활용 15 PF 성능 달성, 2017)



## 핵융합

## ■ KISTI 슈퍼컴퓨터 5호기 구축

- 총 이론성능 25.7PFlops의 시스템: 고효율계산을 위한 매니코어프로세서 노드와 전통적인 멀티코어프로세서 노드로 구성
- 고성능 스토리지와 대용량 스토리지의 계층적 구성
- 3분기 사전 베타서비스를 거쳐 3/4분기 공식서비스 준비

## ■ KISTI 슈퍼컴퓨터 5호기 사전 성능 평가 및 응용소프트웨어 준비

- 매니코어 프로세서의 활용 고도화를 위한 응용소프트웨어 관점의 성능 평가
- 응용소프트웨어 최적화를 통한 슈퍼컴퓨터 5호기 활용 고도화 수행

## ■ 자원 배분 정책

- 국가 차원의 전략적 자원 활용 고려
- 재난·재해 대응, 중소기업지원, 지능정보 분야 등 국가가 수행해야 할 전략적 중점 분야에 대한 비중 확대
- 초고성능컴퓨터로만 해결 가능한 대형 및 첨단 R&D 분야에 대한 비중 확대

**Thank you**

