

01 KISTI 슈퍼컴퓨터 4호기 시스템 소개

글 _ 김성호 선임연구원 · 슈퍼컴퓨팅사업팀
sungho@kisti.re.kr



1. 슈퍼컴퓨터 4호기의 도입 및 선정 과정

지난 1988년 CRAY-2S가 도입되어 국가 슈퍼컴퓨터 1호기로 운영되기 시작한 이후 이번 4호기 슈퍼컴퓨터는 꼭 20년 만에 도입되어 운영될 예정이다. 지난 20년간 한국 과학 기술의 발전에 새로운 연구 수단으로 도입된 슈퍼컴퓨터는 다른 글들에서 밝히겠지만 매우 다양한 연구 분야에서 활용되어 왔다.

슈퍼컴퓨터 4호기는 3호기 IBM시스템과 NEC 시스템의 뒤를 이은 국내 최고 성능의 슈퍼컴퓨터이며, 범용자원으로 국내 모든 연구자에게 제공될 예정으로 지난 2005년 7월 계획이 수립되어 도입이 추진되어 왔다. 이 도입 계획 하에 2005년 12월 1차 입찰을 추진하여 국내외 주요 슈퍼컴퓨터 생산업체들에게 제안서 제출을 요청하였으나 성능, 서비스, 기타 기술지원 등에서 적합한 제안 업체를 찾을 수 없어 유찰을 하고 2차 입찰을 추진하기로 하였다.

슈퍼컴퓨터의 도입 절차는 일반 국가를 상대로 하는 계약법에 따라 정해지는 일반 조달과 유사하나 슈퍼컴퓨터라는 기술적 특수성, 대규모 예산, 초대형 시스템 규모, 슈퍼컴이 가지는 국가적 의미 등으로 인하여 각 중앙부처 주요 요직에 계신 분들과 슈퍼컴퓨터와 관련하여 오랜 경험이 있으신 분, 현재 슈퍼컴퓨터를 사용하여 많은 프로젝트를 수행 중이신 사용자들을 중심으로 선정위원회가 구성되어 제안된 시스템의 기술적, 정책적 사안에 대한 검토 후 업체를 선정하게 된다.

간략히 도입과 관련된 위원회를 소개하면 제안서 및

현행 기술의 문제점을 파악하고 이를 도입해 시스템에 적용할 수 있는지 판단하는 기술위원회, 제안 시스템의 최종성능 및 제안 성능을 검토하는 BMT위원회, 이미 설명한 바 있는 선정위원회가 있다. 각각의 위원회는 위원장과 간사를 두어 회의를 진행한다. 참고로 이번 4호기 선정위원회는 총 8차가 개최되었으며, 기술위원회 및 BMT 위원회는 이보다 훨씬 많은 횟수의 정기, 비정기 회의를 개최하였다.

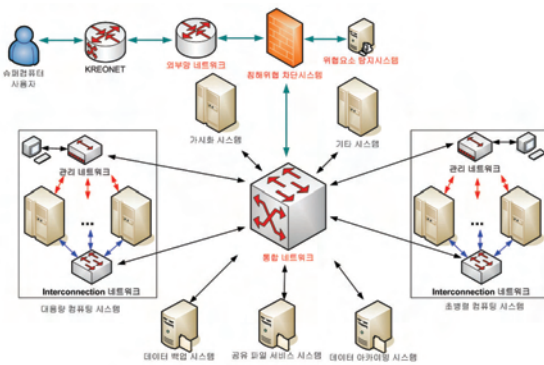
입찰의 방식은 제안 경쟁을 통한 2단계 협상에 의한 계약방식을 택하여 기술 제안서 검토 후 협상대상자를 선정하고 그 중에서 낙찰자를 선정하는 방식으로 하여 최종 도입 설치될 슈퍼컴퓨터의 구성상 미비점을 사전 검토하여 걸러내고, 국내 최고성능 및 전 세계 10위권 이내의 성능을 유지하기 위한 제안업체의 협력을 최대한 이끌어내게 된다.

1차 입찰의 유찰 이후, 선정위원회에서 원활한 입찰참여와 경쟁을 유도하기 위하여 1차 입찰 시와는 달리, 초병렬 컴퓨팅 시스템, 대용량 컴퓨팅 시스템, 가시화 시스템, 통합 네트워크 시스템, 기반 시설 등으로 구분하여 각각을 추진하는 것으로 하였다.

초병렬 컴퓨팅 시스템은 많은 수의 CPU를 필요로 하는 응용 분야를 위해, 대용량 컴퓨팅 시스템은 대량의 메모리를 필요로 하는 응용 분야를 위해, 가시화 시스템은 응용 결과물의 과학적 가시화를 위해 그리고 통합 네트워크 시스템은 이상의 다양한 시스템 간 입출력 데이터의 공유 및 관리를 위하여 도입이 결정되었다. 백업 및 아카이빙을 위한 시스템은 초병렬 컴퓨팅 시스템 제안



업체가 제안하도록 하여 통합 시스템으로 활용 가능하도록 구성하였다.



〈그림 1〉 KISTI 슈퍼컴퓨터 4호기 시스템 구성도

2차 입찰을 통하여 2006년 11월 초병렬 컴퓨팅 시스템은 SUN, IBM, HP를 동시협상자로, 대용량 컴퓨팅 시스템은 IBM을 우선협상자로 선정하여 협상을 진행하기로 의결하여 2007년 3월 SUN과 IBM을 최종 낙찰자로 선정하였다.

낙찰자가 선정된 이후 실제 계약이 있기까지 계약 협상을 위한 별도의 팀이 구성되어 시스템 구성을 보완하고 협력 방안, 서비스 계획, 도입 설치 계획, 검수 조건, 계약서 검토 등의 실무를 진행하였다.

2. 슈퍼컴퓨터 4호기의 구성 및 설치

도입될 시스템은 1차분과 2차분으로 나누어 설치가 진행된다. 이러한 분할 설치 일정은 현재 메인 기계실이 포화 상태에 있고 UPS, 전력 등 신규 시스템이 들어옴에 따라 제반 기반 시설이 준비되어야 하는 부분들이 있어 3호기의 사용자를 일부 수용할 수 있는 선에서 1차분을 구축하고 2009년 상반기에 나머지 시스템을 구축할 예정이다.

이에 따라 SUN의 초병렬 컴퓨팅 시스템은 24TFlops의 성능을 가지도록 구성되어 설치되고, IBM의 대용량 컴퓨팅 시스템은 6TFlops의 성능을 가지도록 구성되어 설치된다. 최종 협상을 통하여 확정된 각 시스템의 규모는 SUN의 초병렬 컴퓨팅 시스템이 286TFlops, IBM의 대용량 컴퓨팅 시스템은 36TFlops의 성능을 가지게 된다. 초병렬 컴퓨팅 시스템과 같이 설치될 예정인 외장 스토리지는 1.2페타바이트(1백만 기가바이트)의 용량을 가지게 되며, 이를 백업하기 위한 테이프 라이브러리는 2페타바이트의 용량을 가진다. 각각의 시스템은 Infiniband라고 하는 고성능 내부 네트워크를 통하여 서로 통신하며 계산을 수행할 수 있도록 되어 있다.

3. 초병렬 컴퓨팅의 주요 특성

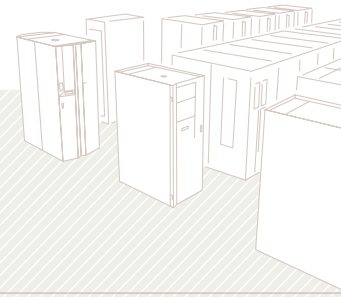
제안된 시스템의 각 특징을 시스템 별로 간략히 설명하겠다.

먼저 <표 1>은 SUN에서 구축하게 될 초병렬 컴퓨팅 시스템의 주요 사양들이다.

간략히 설명하면 1차분은 이론 성능 24TFlops로, AMD사의 64bits CPU를 총 3,008개 사용하게 되며 188노드로 구성이 된다. 각 노드는 4개의 CPU 소켓을 가지고 있으며, 각 CPU는 4개의 계산 코어를 가지고 있어 한 노드 당 16대의 컴퓨터를 사용하는 것과 같은 효과를 가지고 있다. 각 노드는 32GB를 가지게 되므로 각 코어가 이를 동시에 공유하여 활용하게 된다.

2차분의 경우 2009년 상반기 구축 예정이며, 구축될 시점에서 최대 성능을 가지는 CPU를 활용하여 시스템을 구성하게 된다. 현재 예상으로는 대략 21,056의 CPU를 사용할 것으로 예상된다.

01 KISTI 슈퍼컴퓨터 4호기 시스템 소개



〈표 1〉 슈퍼컴퓨터 4호기 초병렬 컴퓨팅 시스템 개요

시스템명(Compute tier rack)	C48	Blade System	
아키텍처(Chip architecture)	x64	x64	
프로세서명(Chip model)	AMD (Barcelona)	Intel or AMD	
이론 성능 (GFlops) Rpeak (GFlops)	CPU(core) 당	8/core	13.32/core
	노드(node) 당	128	106.56
	시스템 전체(total)	24,064	280
노드 수(개)	케비넷(랙) 당(per rack)	48	192
	전체(total)	188	2,632
노드 당 소켓수(개) (sockets per node)	소켓 당(per socket)	4	4
	노드 당(per node)	16	8
CPU[Core]수(개)	전체(total)	3,008	21,056
	노드 당(per node)	32	24
메모리 용량 (GB)	전체(total)	6,016	63,168
	Infiniband Interconnect	포트 당 성능[단방향] (GB/sec)	4X DDR, 2.5GB/sec(20 Gbps) Uni-direction
통합네트워크 포트수	10 GbE	15	89
	1 GbE	10	20
대용량 스토리지	디스크 스토리지 용량(TB)	207	1,230
	테이프용량[비압축](TB)	410	2,050
컴퓨팅 노드 운영체제		Cent OS	Cent OS
운영환경	총 전력소비량(kW/h)	372	1,941
	열 용량(kcal)	179,414	1,573,740
	시스템 표준배치면적(m2)	12.88	34.62
	총하중(kg)	12,457	32,563
	컴퓨팅시스템 냉각장치	냉수식 쿨링 모듈	냉수식 쿨링 모듈

4. 대용량 컴퓨팅의 주요 특성

대용량 컴퓨팅 시스템은 초병렬 컴퓨팅 시스템과는 달리 대량의 메모리를 활용할 수 있도록 구성되어 있으며 IBM에서 생산하는 Power CPU를 사용하게 된다. 1차분은 총 10노드를 운영하게 되며 POWER5+CPU를 사

용하는 p5 595 모델을, 2차분은 p6 H모델을 사용하며 POWER6 CPU를 사용하게 된다.

5. 기타 부가 시스템

대피컴퓨터 4호기는 현재 IBM의 경우 구축을 완료하고



〈표 2〉 슈퍼컴퓨터 4호기 대용량 컴퓨팅 시스템 개요

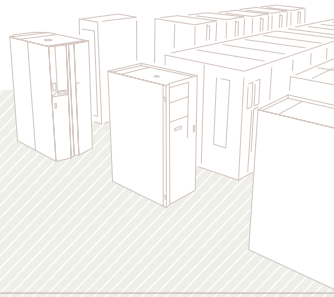
노 드		
구 분	1차	2차
모델명	p5 595	p6 H
프로세서 명	POWER5+	POWER6
프로세서	Clock Speed(GHz)	2.3GHz
	이론최고성능(GFlops)	9.2
노드당 이론 최고성능(GFlops)	588.8	1,280.0
노드당 소켓수(개)	32	32
소켓당 Core수(개)	2	2
노드당 Core수(개)	64	64
캐시 메모리 종류 및 크기	L1: 64/32KB(l/D)/core L2: 0.95MB/core L3: 18MB/core	L1: 64/64KB(l/D)/core L2: 8MB/core L3: 32MB/core
I/O 방식/슬롯수(개)	PCI-X/20 GX/4	PCIX-X/20 GX/4
메모리	노드 당 용량(GB)	256 (9 노드) 512 (1 노드)
	방식	DDR-2 533MHz
	노드 당 대역폭(GB/sec)	799.5
Form Factor[WHD](mm)	785×2,025×1,326	785×2,025×1,326
Interconnection 네트워크 스위치		
구 분	1차	2차
방 식	HPS	Infiniband 4X DDR
총 스위치 수량(대)	3	4
전체 포트 개수(개)	96	96
포트별 대역폭(GB/sec)	2(단방향)	2.5(단방향)
디스크 스토리지		
구 분	1차	2차
물리 디스크 용량(GB)	86,400	374,400
가용 디스크 용량(GB)	63,000	273,000
RAID 구성 방식	RAID 5	

2008년 2월부터 서비스를 진행할 예정이며, SUN의 경우 4월부터 서비스가 가능할 것으로 예상하고 있다.

슈퍼컴퓨터 응용 연구를 통하여 생성되는 사용자 데이

터는 가시화 시스템을 통하여 스크린 상에 보여 질 수 있다. 가시화 시스템은 2007년 11월부터 구축을 시작하여 1월 구축을 완료하고 2월부터 서비스가 될 예정이다. 가시화 시스템의 구성개요는 〈표 3〉과 같다.

01 KISTI 슈퍼컴퓨터 4호기 시스템 소개



〈표 3〉 슈퍼컴퓨터 4호기 가시화 시스템 개요

항 목	신규시스템	슈퍼컴퓨터3호기(nobel)	Onyx3400	
CPU	코어수	784개이상	672개	20개
	이론성능	9.4TFlops이상	4.32TFLOPS	16GFlops
GPU	모델	NVIDIA Quadro FX5600	그래픽처리 능력부재	InfiniteReality3
	GPU수	98개이상		5개
	이론성능	49TFlops이상		-
메인메모리	3.6TB이상	4.4TB	6GB	
네트워크	내부	Infiniband, 20Gbps	-	-
	외부	10Gbpsx26	1Gbps	1Gbps
디스크	용량	360+TB	100TB	400GB전후
	I/O성능	12GB/s이상	-	-

기존 3호기 시스템과 비슷한 정도의 가시화 전용 백엔드 시스템과 최고성능의 그래픽 카드를 탑재하여 사용자 데이터를 처리하도록 구성되어 있으며, 소니 4K 해상도의 프로젝터 4개를 사용하여 입체영상까지 구현이 가능하다.

고 있으며 초고속망과 같은 지식 사회 구현의 중요한 인프라로 인식되어 도입되고 활용되었다. 새로이 도입되는 KISTI 슈퍼컴퓨터 4호기가 앞으로 국가 과학 기술 발전의 중요한 밑거름으로 활용되어 우리나라가 보다 성장하는데 중요한 역할을 할 것으로 믿어 의심치 않는다.

6. 결론

슈퍼컴퓨터 4호기는 기존 3호기에 비하여 성능 면에서 약 100배 이상 빠른 시스템이며, 현재 시점에서 구축이 된다면 전 세계 10위권에 위치하는 대형 시스템이다. 이를 수주한 SUN과 IBM에게는 중요한 참고 사이트로 우리 기관이 알려져 있다. 이러한 시스템은 도입하여 구축하는 것 못지않게 그 성능을 충분히 활용할 수 있는 사용자 지원 및 서비스 체계가 중요하다.

이미 KISTI 슈퍼컴퓨팅센터에서는 이와 관련된 활성화 계획 및 공동 활용 체계 구축계획 등을 세워 4호기의 성능을 십분 발휘할 수 있도록 많은 준비를 하고 있다. 선진국에서는 이미 오래전부터 슈퍼컴퓨터를 핵무기 등과 같이 국가 차원에서 준비되어야 하는 전략 물품으로 인식하



〈그림 2〉 가시화실 및 화면



〈그림 3〉 소니 4K 해상도 프로젝터