

# 클러스터 컴퓨팅

- 홍정우 ■ KISTI 슈퍼컴퓨팅센터, 연구원 / jwhong@hpcnet.ne.kr
- 이지수 ■ KISTI 슈퍼컴퓨팅센터, 선임연구원 / jysoo@hpcnet.ne.kr

이 글에서는 클러스터 컴퓨팅의 필요성과 문제점들을 이야기하고 또 국내외의 클러스터 컴퓨팅 현황을 소개하고자 한다.

## 클러스터 컴퓨팅의 필요성

초기에 등장하였던 슈퍼컴퓨터들은 초고가의 자원이었다. 이러한 슈퍼컴퓨터들은 실제 실험이 불가능한 경우나 결과를 얻기 위해서는 반복적인 실험이 필요한데 그에 필요한 비용이 현실적이지 않은 등의 경우에 실험을 대신하는 도구로 사용되어 왔다. 슈퍼컴퓨터를 이용한 연구들이 늘어나고 또 연구의 규모도 커져감에 따라 지속적으로 보다 큰 용량의 슈퍼컴퓨팅 자원이 요구되어져 왔고, 기존의 자원으로 이러한 요구를 만족시키기에는 한계가 있는 것으로 여겨지고 있다. 특히 BT, NT, ET, MT분야의 문제들은 그 계산의 복잡함이라든지 다루어야 하는 대규모의 데이터 등의 문제 때문에 기존 슈퍼컴퓨터에 적용되던 단일 CPU 혹은 물리적인 SMP(Shared Memory Processor) 컴퓨터 구조에 적합하지 않은 경우가 종종 있다. 이

러한 상황에서 뛰어난 가격대비 성능을 가진 클러스터가 슈퍼컴퓨팅 자원으로 주목을 받게 된 것은 자연스럽다고 생각된다.

현존하는 많은 클러스터들이 리눅스를 사용한다는 사실은 드러나지는 않지만 매우 중요한 장점을 클러스터 컴퓨팅 방법에 추가한다. 컴퓨터를 사용하는 과학기술분야의 연구들은 물리적인 현상들을 수식으로 표현하고 이러한 수식의 계산을 통하여 자연현상을 예측하려 한다. 따라서, 자연현상들의 보다 정확한 기술을 위하여 수치적인 모델들이 지속적으로 개발되어 왔다. 또한 매우 오래 전에 개발되었던 코드들이 수십 년 동안 새로운 기능을 축적해 가며, 계속 사용되어지고 있는 경우도 흔하다. 그러나 컴퓨터 하드웨어와 소프트웨어의 빠른 발전은 컴퓨터를 활용하는 과학기술자들에게 도움과 함께 그에 따른 부담도 같이 전가해 왔다. 즉, 신기술로 만들어진 새로

운 컴퓨터를 쓰기 위해서 과거에 작성하여 놓은 코드를 새로운 구조의 컴퓨터에서 적절히 수행될 수 있도록 재작성 혹은 이식해야 한다. 언제부터인가 리눅스라든가 CopyLeft의 주장 등이 이러한 문제를 해결할 수 있는 방안으로 제안되었다. 즉, 프로그램 소스코드의 공유와 과거에 비해서는 이식성이 뛰어나게 좋은 환경의 등장이 그 이유이다. UNIX 분야의 표준인 Posix를 근간으로 하는 리눅스는 무료인 데다가, 소스가 공개되어 있고, 다양한 사용자층을 확보하고 있다. 과학기술자들에게는 코드를 구하기가 쉽고, 직접 소스코드를 조작할 수 있고, 무료라는 사실들이 매력적으로 여겨지고 있다. 이러한 환경속에서 각 연구그룹에서 개발된 코드들이 쉽게 공유될 수 있고, 지속적인 사용도 보장될 수도 있다는 장점들이 재조명되어지고 있다.

### 클러스터 컴퓨팅의 문제점

앞 절에서는 클러스터가 등장하게 된 배경과 그 이용이 점차 활성화 되어가고 있는 이유들에 대하여 이야기하였다. 그러면 클러스터 컴퓨팅의 문제점들은 어떠한 것들이 있을 수 있을까? 우선 안정된 활용을 하기 위하여 필요한 전문지식을 첫째로 들 수 있다. 리눅스의 일반 활성화에 가장 큰 장애는 설치나 사용도중에 문제가 발생할 경우 컴퓨터 전문가 수준의 지식을 필요로 하는 어려움이라고 한다. 이 문제는 과학기술 분야에 활용되는 클러스터 컴퓨팅에도 똑같이 적용된다. 즉, 클러스터에 주로 설치되는 운영체계가 리눅스이고, 이를 이용하는 사람들이 컴퓨터 전문가가 아니고 물리, 화학, 생명공학, 기계, 기상 등 각 응용분야의 전문가들이라는 사실에 기인한 문제들이다. 또 다른 문제는 대규모 클러스터 컴퓨팅을 위해서는 동시에 매우 많은 수의 컴퓨팅 노드들이 동시에 작동되어야 하고, 그중 한 개의 노드가 고장이 나더라도 전체 기계가 작동을 하지 않을 수도 있다는 점이다. 소요 전력 또한 쉽지 않은 문제이다. 추정방법에 문제가 있을 수 있으나 예를 들어 최대 전력소모가 300W/ hour인 PC를 100 대쯤 사용한다면 시간당 소비전력이 30 kW이고 하루에 720 kW가 소요되므로 연간 262,800 kW이며 2001년 12월 현재 학교나 관공서에 적용되는 전력의 가격인 110 원으로 계산하면



생명공학회사 LOCUS discovery INC에서 운영 중인 dual 1GHz Pentium III 시스템 클러스터

28,908,000 원이 소요되게 된다. 이 문제는 기술발전 등으로 인해 점차 해결될 것으로 예상되거나 지금 당장 일반 대학의 연구실 수준에서 부담하기에는 적지 않은 비용이다. 마지막 문제는 클러스터 컴퓨팅 방법이 모든 문제에 다 적합하지는 않다는 점이다. 클러스터 컴퓨팅은 병렬처리를 기본으로 하는데 다루는 문제 자체가 근본적으로 병렬형태의 계산이 효율적이지 않은 경우도 많다. 따라서 클러스터형 컴퓨터가 기존의 슈퍼컴퓨터들을 완전히 대체하지는 못할 것이며, 가까운 미래에는 같이 공존할 것으로 예상된다.

### 클러스터컴퓨팅의 현황

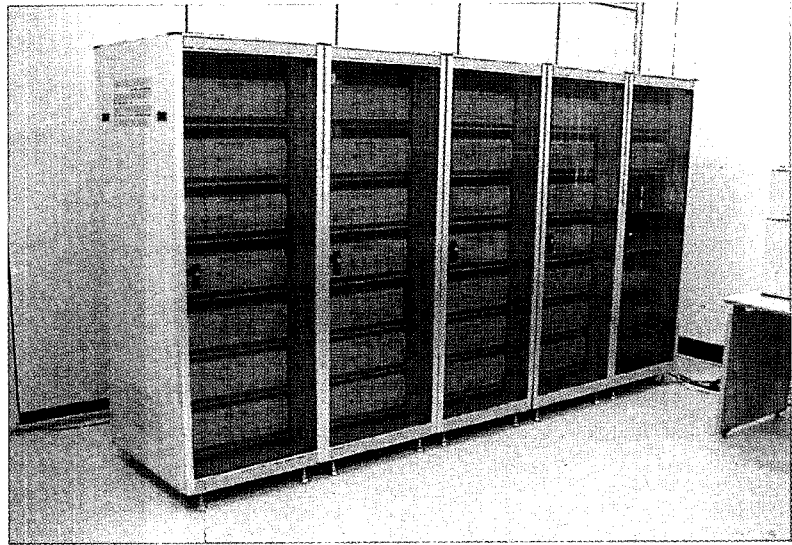
선진국의 경우 학교 및 연구기관들에서 클러스터의 사용이 매우 보편화 되어 있다. 대표적인 예로는 미국의 Sandia National Lab에서 3단계에 거쳐 구축한

Computational Plant (Compaq DS10L 1,536 대가 Myricom 사의 Myrinet gigabit network으로 연결), 생명공학 회사인 Locus Discovery에 설치된 Locus Supercluster(708 노드를 Fast Ethernet으로 연결, 각 노드는 dual 1GHz Pentium III), NCSA에 설치된 Tera32(516 노드를 Myrinet으로 연결, 각 노드는 dual 1GHz Pentium III) 등의 주목할만한 시스템들이 있다. 현재 클러스터 현황을 파악하는 데는 Top 500 클러스터 리스트를 보는 것이 도움이 될 것이다 (<http://www.top500clusters.org>).

국내에서의 클러스터 연구도 활기를 띄어가고 있다. 서울대 기계항공공학부에서 여러 종류의 노드를 Fast Ethernet으로 연결한 클러스터를(2000년 말 기준 총 90 노드) 구축하여 활용하고 있고 부산대 기계공학부에서 32 노드 Fast Ethernet 시스템 및 32 노드 Myrinet 시스템을 구축하여

활용하고 있다. 또한, 고등과학원(KIAS)에서 76개의 노드를 Fast Ethernet으로 연결한 클러스터를 구축하여 단백질 구조해석 등의 분야에 사용하고 있고, 포항공대 화학과에서도 클러스터 구축이 진행되고 있다. 이외에도 자체 생산한 Alpha EV6프로세서 및 시스템보드와 Myrinet장비로 128 노드의 시스템을 구축하여 내·외부 서비스를 하고 있는 삼성종합기술원과 공공 서비스와 슈퍼컴퓨팅 자원의 자체 확보능력 배양을 목적으로 관련 기술을 개발중인 KISTI 슈퍼컴퓨팅센터 등이 있다. 이 곳에서 언급한 활동들은 국내 클러스터 활용의 일부에 불과하며, 클러스터 시스템의 구축 및 활용은 많은 대학 및 연구기관에서 현재 계획 또는 수행되고 있다.

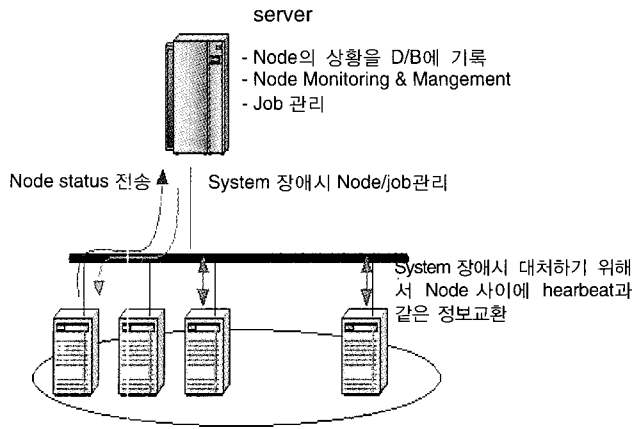
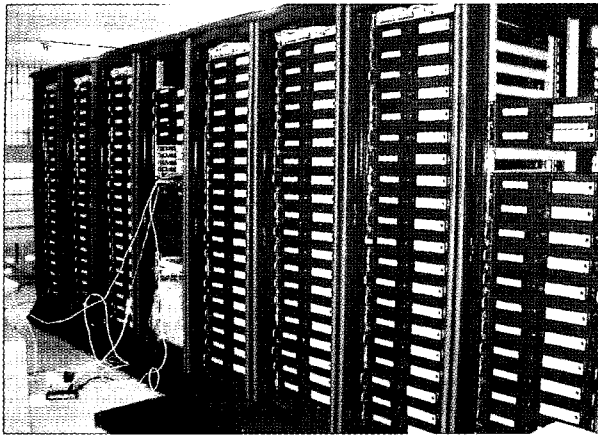
과학기술자들에게 슈퍼컴퓨팅을 자원을 지속적으로 제공해오고 있는 KISTI 슈퍼컴퓨팅센터에서는 BT, NT, ET, MT 등의 신기술 분야에서 대규모의 컴퓨팅을 필요로 하는 과제들을 지원하기 위한 TeraFlop급 클러스터 시스템의 개발과 활용에 관한 연구를 수행하고 있다. 이 연구는 클러스터가 가지는 가격 대비 성능, 응용분야에 따른 맞춤형 하드웨어 구성이 가능함 등과 같은 장점을 극대화하면서 위에서 언급한 문제들에 기인한 단점들을 보완하여 신뢰도가 높은 컴퓨팅 자원으로 활용할 수 있도록 하는 것이 목적이다. 이를 위하여 1999년 초기시스템 구축을 필두로 2000년에는 Intel, Alpha, AMD 등의



부산대학교에서 구축 운영중인 클러스터 : 1999년 5월부터 2000년 6월까지 구축 및 보안을 거쳐 현재 운영 중. 총 128 CPU 규모로 Fast Ethernet으로 구성된 Dual Pentium II 450Mhz/512MB Memory 64CPU 시스템과 Myrinet 장비를 근간으로 구축한 Dual Pentium III 550MHz/512MB Memory 64CPU 시스템으로 구성된다.

다양한 종류의 프로세서와 Fast Ethernet, SCI, Myrinet 등의 다양한 네트워크의 장비들의 조합에 대하여 물리, 화학, 기상, 열유체, 구조해석 등의 분야의 응용프로그램들의 성능을 소규모의 클러스터에 대하여 벤치마킹하였다. 또한 이러한 실험의 결과들을 바탕으로 특정 응용프로그램에 최적화된 클러스터 구성에 대한 연구와 구현을 수행하여왔다. 현재는 128 노드로 구성되는 중간 규모의 클러스터를 센터 내에 설치하고 있으며, 여러 분야의 대표적 응용프로그램들의 성능의 벤치마킹과 최적화에 관한 연구를 수행할 계획이다. 또한 COTS (Commodity Off The Shelf)장비를 사용하여 구축하는 클러스터 시스템의 성능한계를 극복하고 upgrade 요구를 충족시킬 방

안으로 각 노드 단위의 연산능력을 향상시키기 위하여 각각의 노드에 장착해서 특정연산을 고속으로 처리할 수 있는 수치가속기 보드도 연구 중이다. 또한 본 센터에서는 대규모의 클러스터의 유지 및 관리에 필요한 관리 소프트웨어들의 연구도 활발히 진행하고 있다. 이러한 연구들로 얻어진 결과는 위에서 언급한 TeraFlop급 클러스터의 설계, 설치 및 운영에 매우 유용하게 사용되리라 예상한다. 이 Tera클러스터는 현재 본 센터의 Cray 사의 T3E을 사용하는 사용자들 중에서 비상용코드를 사용하는 사용자그룹과 정책적으로 지원이 필요한 거대과제 등의 지원에 사용될 계획이다. 또한 슈퍼컴퓨터, 대규모의 저장장치, 고가의 실험 장비등의 자원들을 초고속 네트



KISTI 슈퍼컴퓨팅센터에서 구축 중인 TeraCluster 개발을 위한 기반 클러스터 시스템 : 계산노드 128 대 규모이며, 2001년 12월 초 현재, Linpack 수치로 208Gflops 이상을 보이고 있으며, 최적화 작업 이후에는 230Gflops 수준의 성능이 가능할 것으로 예상된다. 2002년~2003년도에는 512~1024노드 이상 규모의 TeraFlop급 클러스터를 구축하고 운영할 수 있는 기술을 개발하고 실제 적용하는 것을 목표로 하고 있다.

KISTI에서 개발 중인 클러스터 관리 소프트웨어 환경의 시스템감시기능도 : 클러스터 시스템은 다수의 개별시스템이 공존하는 형태이므로 그 단위 계산 노드의 수만큼 감시 및 관리의 대상이 존재하게 된다.

위크로 연계하여 효율적으로 사용하는 방안인 Grid에 이러한 클러스터들을 효과적으로 연계하는 방안에 대하여도 연구하고 있다.

**맺음말**

클러스터 컴퓨팅은 컴퓨팅의

한 방법으로서 중요한 역할을 하게 될 것이다. 클러스터 컴퓨팅 기법이 여러 응용분야로 확산된다는 것은 이러한 방법론이 필요한 분야들에 대해서 컴퓨팅 기법의 고도화가 이루어지고 있다는 것을 의미하며, 주어진 시스템에서 보다 많은 가능성들을 시험해 볼 수 있다거나 경제적인 이유들로 인해 접근이 어려웠던 대규모의 계산문제들에 대해 좀더 쉽게 접근이 가능하게 될 것으로 판단된다. 즉, 각 응용 분야의 과학기술자들이 좀더 크고 복잡하며, 사회, 경제, 과학, 기술 등의 분야에 커다란 파급효과를 지니는 연구를 수행할 수 있는 환경을 제공할 수 있다는 것을 의미한다.

System	Architecture	User%	Application%	User Affiliation%
Cray C90 (Vector)	NEC SX4/5	13%	Mechanics : 60% Atm. Env. : 26% Etc : 14%	Industry : 45% Edu : 32% Research : 7% Etc : 13%
Compaq SMP (Non vector) (Commercial)	Bio. etc Special Purpose	9%	Physics : 59% Chemistry : 16% Mechanics : 13% Etc : 12%	Industry : 2% Edu : 87% Research : 10% Etc : 0%
Cray T3E (Parallel) (Non commercial)	IBM SP (commercial) (Parallel) (Non Vector)	78%	Mechanics : 64% Chemistry : 12% Atm : 9% Etc : 15%	Industry : 5% Edu : 79% Research : 15% Etc : 0%
	Cluster (Non Commercial) (Parallel)	283		

1993 1997 2000 2001 2002 2003 Arch. User% Application% User Affiliation%

KISTI 슈퍼컴퓨팅 자원계획(2001년도 사용현황 분석 포함) : 현재 KISTI 슈퍼컴퓨팅 자원 정책은 다양한 형태의 슈퍼컴퓨팅 자원을 모두 제공하는 것으로 계획하고 있으며, 클러스터 컴퓨팅 자원은 국가적 정책적으로 중요한 파급효과를 지니는 거대 컴퓨팅 과제를 위해 제공하며, 기존 형태의 슈퍼컴퓨팅 자원으로써는 기존의 중·소규모의 계산을 필요로 하는 연구와 병렬처리가 불가능한 계산 및 상용 프로그램을 필요로 하는 분야 등을 지원할 계획으로 있다.