

윈도우즈 기반의 병렬컴퓨팅 환경 구축 및 성능평가

Construction and Performance Evaluation of Windows-based Parallel Computing Environment

신재렬^{*1)}, 김명호¹⁾, 최정열²⁾

J.-R. Shin, M.-H. Kim and J.-Y. Choi

A parallel computing environment was constructed based on Windows 2000 operating system. This cluster was configured using Fast-Ethernet system to hold up together the clients within a network domain. For the parallel computation, MPI implements for Windows such as MPICH.NT.1.2.2 and MP-MPICHNT.1.2 were used with Compaq Visual Fortran compiler which produce a well optimized executives for x86 systems. The evaluation of this cluster performance was carried out using a preconditioned Navier-Stokes code for the 2D analysis of a compressible and viscous flow around a compressor blade. The parallel performance was examined in comparison with those of Linux clusters studied previously by changing a number of processors, problem size and MPI libraries. The result from the test problems presents that parallel performance of the low cost Fast-Ethernet Windows cluster is superior to that of a Linux cluster of similar configuration and is comparable to that of a Myrinet cluster.

Key Words: 윈도우즈2000(Windows 2000), 리눅스(Linux), 클러스터(Cluster)

1. 서론

마이크로프로세서의 발전은 대형 슈퍼컴퓨터에서나 가능하던 대형의 공학 문제를 개인용 컴퓨터(PC)에서 해결할 수 있게 하였다. 더욱이 병렬 처리기법의 개발로 인해 슈퍼컴퓨팅의 양상과 규모는 현저한 변화를 가져왔으며, 인터넷의 확산으로 인하여 저가의 고성능 네트워크 장비가 보급됨으로써, 슈퍼컴퓨터에 비해 월등한 가격 대 성능 비를 보이는 PC 클러스터의 구축이 실험실 수준에서 가능하게 되었으며, 이를 이용한 병렬 슈퍼컴퓨팅이 널리 확산되고 있다.

클러스터는 네트워크로 묶여져 하나의 자원처럼 사용되어지는 PC들의 집합으로서, 최근의 다양한 운영시스템을 가지는 PC클러스터들이 선보이고 있다.[1] 이중 가장 널리 이용되고 있는 운영체제는 표준화된 네트워크 구성에 강점을 가지는 유닉스 기반의 리눅스 PC 클러스터로써, 운영체제가 공개되어 있어 저렴하게 PC클러스터를 구축할 수 있는 강점이 있으나, 일반적으로 PC 운

영체제의 주류인 Windows 환경에 익숙한 사용자들이 이를 이용하기 위해서는 새로운 환경을 익혀야 한다는 단점을 가지고 있으며, 다양한 응용프로그램의 부재와 서로 다른 운영 체제환경을 병행 운용하여야하는 번거로움이 있다.

그러나, 네트워크 운영체제인 Windows NT 계열의 운영체제가 개발되면서, 개인용 컴퓨터가 아닌 네트워크 서버 등의 대형 시스템으로 Windows 운영체제의 이용 영역을 넓히고 있다. 슈퍼컴퓨팅 부문에는 Windows 환경을 이용하고자 하는 시도가 과거 수년간 있었으나, 현재로써는 거의 이용되고 있지 않다. 그러나 네트워크 성능이 보강된 Windows 2000의 개발에 따라 병렬 슈퍼 컴퓨팅에도 유용하게 이용될 수 있는 여지를 보이고 있다. 따라서 본 연구에서는 Windows 2000 기반의 PC 클러스터를 구축하여, 기존의 리눅스 기반 클러스터와 성능을 비교 평가함으로써, 병렬 슈퍼 컴퓨팅 부문에서의 Windows 운영체제의 이용가능성을 모색하고자 한다.

1) 부산대학교 기계공학과 대학원

2) 부산대학교 기계공학과



| | Fast-Ethernet Windows Cluster (8node-16CPU) | Fast-Ethernet Linux Cluster (8node-16CPU) | Myrinet Linux Cluster (16node-32CPU) |
|----------------|--|---|--|
| CPU | Intel Pentium III 1GHz×2 | Intel Pentium III 933MHz×2 | Intel Pentium III 550MHz×2 |
| Memory | 512MBytes SDRAM | 512MBytes SDRAM | 512MBytes SDRAM |
| HDD | IDE 40GB (ATA100) | IDE 30GB | SCSI-II 9.1GB |
| NIC | Realtek 8139A (100Mbps) | Fast-Ethernet 100Mbps x2 | Netgear FA310TX + Myrinet Lan-PCI 1Gbps |
| Network Switch | 3Com Office Connect Dual Speed Switch (3C16735B) | Intel Express 410T Standalone Switch | 3Com Superstack II 3900 + Myrinet-LAN switch |
| OS | Windows 2000 (Sp2) Server/Professional | WOW Linux 6.2 (kernel 2.4.0 smp) | Alzza Linux 6.1 (kernel 2.2.12-20krmp) |
| MPI Lib. | MPICH.NT.1.2.2 MP-MPICH.1.2 | MPICH-1.2.1 | MPICH-1.2.1 |
| Compiler | Compaq Visual Fortran 6.1A | G77 (egcs 1.1.2) | Absoft Pro Fortran 6.2 |

Table. 1 Specification of Each Node of Clusters

2. Windows 2000 PC Cluster의 구성

2.1 Hardware 구성

Table 1은 본 연구에서 이용한 Windows 클러스터와 기존의 리눅스 클러스터의 하드웨어 사양을 비교 정리한 것이다. 이러한 하드웨어의 구성은 일반적인 리눅스 클러스터와 크게 다르지 않으며, 구체적인 하드웨어의 사양은 개발 시점에서 전체 시스템 차원에서 최적의 가격 대 성능비와 신뢰성을 고려하여 결정하였다. 프로세서는 일반적으로 높은 신뢰성을 보이는 것으로 알려진 [2] Intel사의 제품을 이용하였으며, 주기판과 기타 부품 역시 같은 기준에서 선택하였다. 네트워크 장비로는 각 노드에 100Mbps 급 NIC를 설치하였으며, 외부에서의 접근을 위해 서버에 한 개의 NIC 추가하였다. 각 노드는 통신에 충분한 대역폭을 확보해 줄 수 있는 동급의 Switch 장비를 이용하여 연결되었다.

2.2 Windows 2000 설치 및 네트워크 구성

Windows 2000의 설치와 다른 운영체제의 설치와 다르지 않고, 단지 Cluster를 구성하고 운영하기 위한 서버의 특수 기능만이 다르다. 따라서 서버에는 Windows 2000 Server를 설치하였으며, 각 노드에는 Windows 2000 Professional을 설치하였다. 각 사용자 계정 및 네트워크 자원에 대한 정보를 관리하기 위하여

서버에는 액티브 디렉토리를 설치하였으며, 이를 통해서 클라이언트에 접속한 사용자나 수행하고 있는 작업을 확인하고 관리할 수 있다. 한편 외부에서는 서버에 사용자계정을 통해 원격으로 클러스터를 이용할 수 있도록 Terminal Server의 기능을 수행하도록 설정하였으며, MS 네트워크 클라이언트를 통하여 공유자원에 대한 접근이 가능하며, Telnet이나 FTP와 같은 원격지 서비스도 지원하도록 하였다.

한편 네트워크 구성의 경우 병렬처리 도구들은 주로 TCP (Transport Control Protocol), SMP (Symmetric MultiProcessor) 그리고 VIA (Virtual Interface Architecture) 같은 통신 매체를 이용하므로, 본 연구에서는 TCP 와 DNS(Domain Name

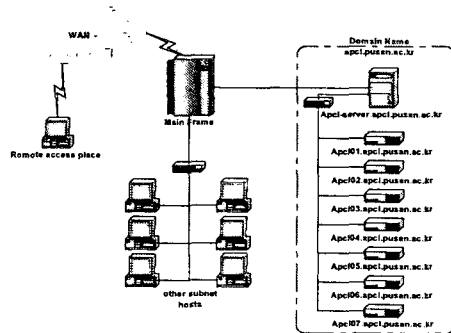


Fig. 1 Network Design

Service)를 이용하여 네트워크를 구축하였다. Fig. 1은 클러스터의 네트워크 구성을 나타낸다. 여기서 서버는 외부 네트워크에 접속하기 위한 공인 IP를 가지고 있으며, 내부 네트워크는 사설 IP를 이용하여 구성하였다. 따라서 서버에는 두 개의 NIC가 필요하며, 네트워크는 한 개의 도메인에 클라이언트들이 존재하는 형태이다. 따라서 외부에서는 서버 이외에 다른 클라이언트에는 직접 접속할 수 없으며 외부에서는 하나의 도메인 네임을 가지는 단일 시스템으로 인식된다.

2.3 병렬처리 도구 및 설치

병렬처리 클러스터는 분산 메모리 체계이므로 기본적으로 메시지 패싱을 통하여 병렬처리를 구현한다. 메시지 패싱 규약에는 PVM, MPI 등이 있지만 본 연구에서는 가장 널리 이용되는 MPI를 이용하여 구성하였다. Windows 시스템에 이용될 수 있는 공개된 MPI 패키지로는 MPICH.NT, MP-MPICH 등이 있으며, 상용으로는 WMPI, MPI-Pro 그리고 PaTENT이 있다.[3-7] 이들 중에 여러 패키지들이 MPICH.NT를 기본으로 개발되었으므로 본 연구에서는 MPICH.NT.1.2.2를 이용하였으며, 이를 좀 더 개량한 MP-MPICH.1.2를 비교하여 보았다. 한편 컴파일러로는 Windows 환경에서 널리 이용되고 있으며, 좋은 성능을 보이는 것으로 알려진 Compaq Visual Fortran 6.1A를 이용하였으며, 필요에 따라 MS Visual C++ 등을 이용할 수 있다.

3. 병렬처리 성능평가

3.1 Windows Cluster와 Linux Cluster의 사양

병렬처리 성능을 비교 평가하기 위하여 Table 1에 제시된 Windows Cluster와 두 가지 Linux Cluster의 처리 성능을 비교하였다. 모두 Dual 프로세서 시스템으로 유사한 사양을 가지지만, 구성 시점에 따라 프로세서 및 네트워크 사양에는 약간씩 차이가 있다.

모두 동일 계열의 프로세서이므로 성능은 클럭 속도에 비례한다고 할 수 있지만, 최적의 성능을 이끌어 내기 위해서는 컴파일러의 성능이 중요하다. Windows 환경에서 많이 이용되는 Visual Fortran은 충분히 좋은 성능을 보이고 있는 것으로 알려져 있으며, 리눅스 환경에서는 Abssoft사

의 것이 좋은 평가를 받고 있다. 그러나 공개 소프트웨어인 G77은 최신 프로세서에 최적화된 실행 코드를 생성하지 못하는 것으로 알려져 있다.

한편 네트워크의 경우 이들의 이론상의 대역폭은 Fast-Ethernet이 100Mbps이고 Myrinet은 1Gbps지만 실험실상에 네트워크를 구축했을 때는 이보다 작은 대역폭을 갖는다. 지체시간에서는 Fast-Ethernet은 수십밀리초, Myrinet은 수십 마이크로초의 지체시간을 가진다. Fig. 2는 Fast-Ethernet과 Myrinet에서 노드간의 데이터

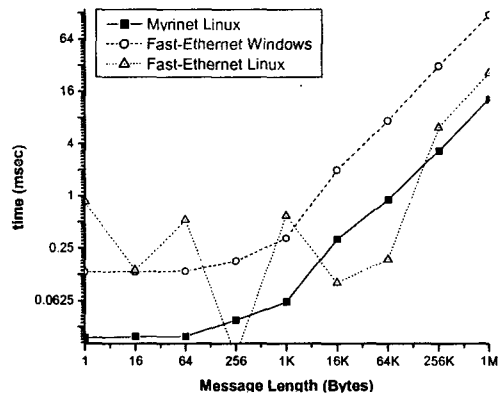


Fig. 2 Ping-Pong between nodes

전달에 걸리는 시간을 알아보기 위해 데이터의 크기에 따른 전송에 걸리는 시간을 보여준다. Myrinet을 이용한 클러스터와 Fast-Ethernet으로 구축한 Windows Cluster의 경우 근본적인 네

| CPU No. | Fast-Win | Fast-Linux | Myri-Linux |
|---------|----------|------------|------------|
| 1 | 98.986 | 160.735 | 176.405 |
| 4 | 23.592 | 41.975 | 46.839 |
| 8 | 12.893 | 25.121 | 23.522 |
| 16 | 7.545 | 15.505 | 11.982 |

Table. 2 Computing Time[s] for 631x211 Grid

트위크의 크기의 차이를 빼고는 같은 경향을 띤다. Fast-Ethernet으로 구축한 Linux Cluster는 두 개의 NIC를 가지고 있지만 네트워크의 불안정한 면을 보여준다. 또한 운영체제에 따라, Linux Cluster와 Windows Cluster는 많은 데이터를 주고받을 때는 거의 비슷한 지체시간을 보이지만, 데이터의 양이 10KB 보다 작은 경우 Windows가 Linux보다 2-4배 정도 작은 지체시간을 가진

다.[8] 한편 MPI Library의 성능도 네트워크의 성능에 영향을 미치므로 병렬처리 시스템의 성능은 종합적으로 평가하여야 한다.

3.2 병렬성능 평가 모델

병렬처리 성능 평가는 이전의 연구에서와 일관성을 유지하기 위하여 이차원 압축성/점성 유동의 예조건화 해석 코드를 이용하였다.[8] 성능 평가 문제는 낮은 마하수의 압축기의 익셀 유동해석을 문제로써, 계산 및 통신 부하에 따른 영향을 살펴보기 위하여 작은 크기의 격자인 211x71 격자와 큰 크기의 격자인 631x211 격자에 대한 해석을 각각 100 번 및 10 번의 반복수행에 따른 계산 시간을 평가하였다. 해석에 이용된 수치기법으로는 공간이산화에 Roe의 풍상 차분법과 TVD-MUSCL을 이용한 고차정확도 기법이, 시

를 적용하였다.[9] 한편 병렬처리는 영역분할을 통하여 구현하였으며, 각 분할 영역에 동일한 수의 격자수를 배분함으로써 연산 부하를 동일하게 배분하였다.

3.3 병렬 처리 성능 비교

Fig. 3과 Fig. 4는 문제의 크기에 따른 각 시스템에서의 속도향상 비를 프로세서 수에 따라 도시한 것이다. 이 결과로부터 Windows 클러스터가 1개의 Fast-Ethernet NIC만을 이용함에도 불구하고 Myrinet 클러스터에 버금가는 속도 향상 비를 보여 주고 있음을 알 수 있으며, 문제의 크기가 작은 경우에도 비교적 안정된 속도 향상 비를 확인할 수 있다. 더욱이 Table. 2 에 요약된 절대 계산시간을 비교한다면 Windows 환경의 우수한 컴파일러를 이용할 수 있음에 기인하여 매우 우월한 성능을 보임을 알 수 있다.

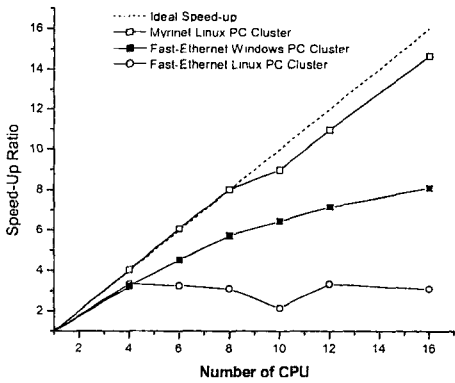


Fig. 3 Speed-Up Ratio for 211x71 Grid

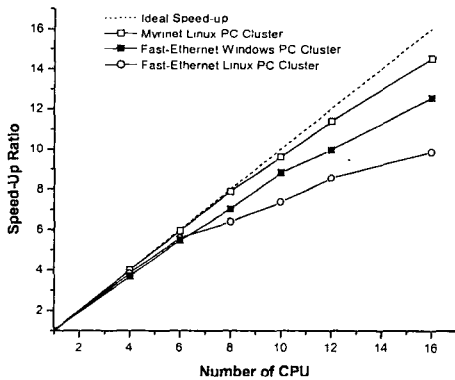


Fig. 4 Speed-Up Ratio for 631x211 Grid

3.4 MPI Library 및 Topology에 따른 성능

Windows 환경에서 이용 가능한 MPICH.NT. 1.2.2 및 MP-MPICH1.2 의 성능을 비교하고, 프로세서 지정 방식에 따른 영향을 살펴보기 위하여 각 경우에 따른 성능 향상 비를 Fig. 4에 도시하였다. 여기서 Case_I는 노드 순서로 프로세스를 지정(최대의 노드수 사용)한 경우이며 Case_II는 각 노드 당 2개의 프로세스를 지정(최소의 노드수 사용)한 경우이다. 이 결과로부터 Windows에 기반 한 병렬처리 환경은 프로세스

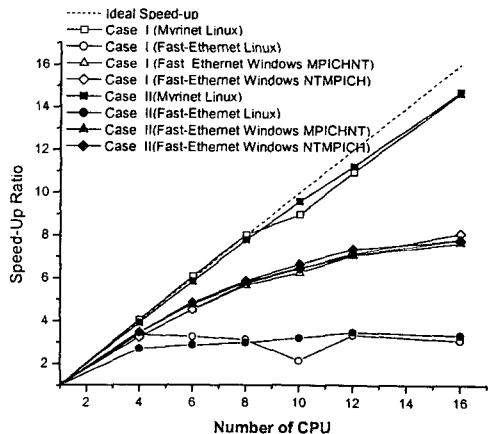


Fig. 5 Dependence of Speed-up Ratio on MPI Library & Processor Topology

간적분에는 Gauss-Seidel 내재적 해법이 이용되었으며, 난류모델로는 Menter의 k- ω SST모델

의 네트워크 구성 방식에 따른 성능 차이가 크지 않으며, 각 라이브러리에 따른 성능 차이도 두드러지지 않아 안정된 병렬 효율 향상을 보임을 확인할 수 있다

4. 결 론

본 연구에서는 Windows 2000 PC 클러스터 환경에 기반한 병렬 슈퍼컴퓨팅 환경을 구성하였으며, 기존의 리눅스 클러스터와 병렬 라이브러리 성능을 비교 평가하였다.

이 결과, Windows 환경은 저가의 보편적인 하드웨어 구성에도 불구하고 기존의 리눅스 클러스터에 비해 매우 우수하고 안정적인 병렬처리 성능을 보여 주었다. 본 연구에서 구성한 시스템은 더 많은 수의 노드로의 확장에 별다른 제한이 없으므로, 이용자의 편의와, 다양한 소프트웨어와 높은 호환성, 단일 운영체제에 의한 컴퓨팅 환경 구성 등의 장점을 고려한다면 슈퍼컴퓨팅 환경의 새로운 대안으로 이용될 수 있을 것이다. 또한 널리 확산된 인터넷과 Windows 운영체제를 기반으로 병렬처리를 위한 네트워크를 구성하는데 이론적인 제한이 없으므로, 향후 연구에서는 인터넷 슈퍼컴퓨팅 환경으로서 Windows 운영체제의 이용가능성을 모색할 것이다.

참고문헌

- [1] Baker.M, "Cluster Computing White Paper", Version2.0, University of Portsmouth, UK, Sept. 2000.
<http://www.dcs.port.ac.uk/%7Emab/tfcc/WhitePaper/>
- [2] Pabst. T and Völkel. F, "Hot Spot - How Modern Processors Cope With Heat Emergencies", <http://www.tomshardware.com/cpu/01q3/010917/index.html>, Spet. 2001.
- [3] "MPICH.NT.1.2.2", <http://www-unix.mcs.anl.gov/mpi/index.html>
- [4] "MP-MPICH.1.2", <http://www.lfbs.rwth-aachen.de/~karsten/projects/nt-mpich/index.html>
- [5] "Wmpi1.5", <http://www.criticalsoftware.com/wmpi/home/index.html>
- [6] "MPI-Pro", <http://www.mpi-softtech.com/default.asp>
- [7] "PaTENT", <http://www.genias.net/geniasde.ht>

ml

[8] Luecke. G.R, Raffin.B and Coyle.J, "Comparing the Communication Performance and Scalability of a Linux and NT Cluster of PCs, a SGI Origin 2000, an IBM SP and a Cray T3E-600", Iowa State University, 2000.

[9] 이기수, 김명호, 최정열 외, "Myrinet과 Fast-Ethernet PC Cluster에서 예조건화 Navier-Stokes 코드의 병렬화 효율 비교," 2001년 항공우주공학회 춘계학술발표회 논문 (1992), p.18.