컴퓨터 모의실험
기술동향 보고서

슈퍼컴퓨팅 응용실
오광진

1. 서론

1950 년대 최초의 컴퓨터 모의실험이 행해진 이후로 컴퓨터 모의실험은 수립된 이론을 검증하거나 또는 얻어진 실험 결과를 재현하고 분석하는데 많이 이용되어 왔다. 하지만 컴퓨터가 갖고 있는 계산 능력의 한계로 연구할 수 있는 주제는 한정적일 수 밖에 없었다. 그런데 최근에는 슈퍼컴퓨터의 발전과 새로운 컴퓨터 모의실험 기법의 발전에 힘입어 질병의 원인에 대한 분자적 이해나 신약 그리고 새로운 재료의 설계에 이르기 까지 다룰 수 있는 주제가 매우 다양해 졌다. 그 결과 현재는 물리/화학/생명공학/재료공학/전기/전자 등을 망라한 여러 분야에 걸쳐 활발히 이용되고 있다. 한 걸음 더 나아가 앞으로는 컴퓨터 모의실험이 제품개발의 핵심 기술로서 자리잡을 것으로 기대된다.

본 보고서에서는 우선 컴퓨터 모의실험의 전반적인 기술 동향을 살펴보고 컴퓨터 모의실험을 이용한 구체적인 연구 사례를 살펴 보도록 하고자 한다.

2. 전반적인 컴퓨터 모의실험 기술동향

컴퓨터 모의실험은 가상의 모의실험 상자(simulation box) 안에 분자들을 넣고 물리법칙을 이용하여 분자들의 위치를 변화시키는 것이다. 컴퓨터 모의실험의 궁극적인 목적은 어떤 현상이나 물질을 되도록이면 실제와 비슷하게 모사하여 실험을 거치지 않고 그 현상이나 물질의 성질을 이해하고자 하는 것이다.

컴퓨터 모의실험은 보통 분자들 간에 작용하는 에너지를 계산하고 그 에너지를 바탕으로 분자들의 위치를 갱신한다. 그런 다음 축적된 정보를 가지고 후처리 과정을 거쳐 원하는 현상이나 물질의 성질을 얻는다.

2.1 상호작용

상호작용은 보통 경험적 모델에 의해 주어진다. 생체분자 컴퓨터 모의실험에 많이 사용되는 CHARMM 이나 AMBER 가 대표적인 모델이다. 상호작용은 크게 두 가지로 구분된다. 하나는 분자내(intramolecular) 상호작용이고 다른 하나는 분자간(intermolecular) 상호작용이다. 분자내 상호작용은 분자구조의 변화에 따른 에너지 변화를 고려하기 위한 것으로서 결합길이의 변화, 결합각의 변화, 이면각의 변화, 그리고 각 성분간의 상호작용을 고려한 항으로 구성된다. 분자간 상호작용은 분자를 이루는 원자들 사이의 반 데어 발스(van der Waals) 상호작용과 정전기적(electrostatic) 상호작용을 고려한 항으로 이루어 지고 보통 거리의 함수로 주어진다.

여기에 더욱 정확한 분자간 상호작용 에너지의 계산을 위하여 편극률(polarizability)이나 다체 효과가 고려되기도 한다. 현재 많은 상호작용 모델들이 존재하나 앞서 언급한 틀에서 크게 벗어나지는 않는다.

상호작용 모델은 보통 작용점(interaction site)이 원자이다. 물론 대상이 되는 분자가 크다면 작용점도 많아져 상호작용 에너지를 계산하는 데 시간이 많이 소요된다. 이에 여러 작용점을 하나의 작용점으로 통합시켜 전체 작용점의 숫자를 줄이려는 노력도 최근 많이 하고 있다. 그러면 계산 시간에 있어 많은 효과를 얻을 수 있게 된다. 이러한 상호작용 모델은 coarse -grained 모델이라 불리는 데 앞서 언급했듯이 계산 시간을 단축시킬 수 있다는 장점이 있지만 모델에서 세세한 분자 구조에 대한 정보를 생략했기 때문에 원자 수준 에서의 미시적 정보는 제공하지 못한다는 단점이 있다.

상호작용 모델의 성공여부는 그 모델이 실험 결과를 얼마나 잘 재현하는가에 따라 결정된다. 만약 어느 특정한 상호작용 모델을 사용하여 얻어진 결과가 실험 결과와 잘 일치하지 않는 다면 그 상호작용 모델은 모델로서 가치를 상실하게 된다. 그래서 제안된 많은 상호작용 모델들은 양자역학적 계산이나 실험으로 측정된 정보와의 비교를 통해 꾸준하게 개선되어 왔다. 예를 들면 CHARMM 모델의 경우 버전 27 까지 개발되었다.

보통 분자내 상호작용 에너지 계산과는 달리 분자간 상호작용 에너지 계산은 많은 시간을 필요로 한다. 왜냐하면 모의실험 상자 안에 N 개의 원자가 있다면 분자내 상호작용 에너지 계산시간은 N 에 비례하지만 분자간 상호작용 에너지의 계산시간은 N^2 에 비례하기 때문이다. 더군다나 분자간 상호작용의 경우 거리를 계산할 때 sqrt 연산을 해야 하기 때문에 더욱 더 계산시간을 요하게 된다. 이에 분자간 상호작용 에너지를 효율적으로 계산할 수 있는 알고리즘을 개발하려 많은 연구자들이 노력하고 있다. 현재 만히 사용하고 있는 알고리즘에는 potential truncation 방법과 neighbor list 방법이 있다.

potential truncation 방법은 만약 두 원자 사이의 거리가 정해진 작용점 제한거리 (cutoff distance)거리보다 멀면 상호작용을 하지 않는 다고 가정한다. 따라서 작용점 제한 거리이상 떨어진 원자들 사이의 상호작용은 계산할 필요가 없게 되어 계산 시간을 많이 단축할 수 있다. 하지만 작용점 제한 거리를 경계로 해서 상호작용의 단절이 생기기 때문에 에너지 보존에

문제가 발생하게 된다. 이에 그 작용점 제한 거리를 경계로 해서 상호작용이 서서히 변하게 하도록 switching 함수나 shifting 함수를 사용하게 된다. 현재 switching 이나 shifting 에 관한 여러 기법이 개발되어 많이 사용되고 있다.

neighbor list 방법에는 Verlet neighbor list 방법과 cell-linked list 방법 그리고 그 둘을 혼합한 방법이 있다. Verlet neighbor list 방법은 매번 모든 원자 쌍에 대해 거리를 계산하지 않고 작용점 제한거리 내에 있는 이웃한 원자들의 목록을 기억해 두었다가 이용하기 때문에 potential truncation 방법보다도 빠르다. 그리고 cell-linked list 방법은 모의 실험 상자를 균등하게 분할하고 각 분할된 영역에 속해 있는 원자들을 기억해 두었다가 이웃한 영역에 있는 원자들과의 상호작용 에너지만 계산한다. 따라서 cell-linked list 방법을 통해 상당한 계산시간 단축효과를 얻을 수 있다.

정전기적 상호작용 에너지 계산에는 potential truncation 방법을 사용할 수 없다. 왜냐하면 정전기적 상호작용은 보통 사용하는 작용점 제한거리보다도 먼 거리에 걸쳐 작용하기 때문이다. 그래서 정전기적 상호작용 계산을 위한 특별한 방법이 고안되었는데 그것이 ewald 방법이다. 이 방법은 아직까지도 널리 사용되고 있는 고전적인 방법이다. 하지만 최대로 최적화된 ewald 방법도 계산속도가 $N^{3/2}$ 에 비례하기 때문에 원자들의 수가 증가함에 따라 계산시간이 기하급수적으로 늘어난다. 이러한 난점을 해결하기 위해 많은 연구가 진행되어 왔고 최근 계산속도가 NlogN 에 비례하는 더욱 개선된 방법들이 고안되었다. 그 중 하나가 SPME(smooth particle mesh ewald)와 같은 고속 푸리에 변환 (fast Fourier transform)을 이용한 방법이고 나머지가 FMM (fast multipole method)와 같은 다중극자(multipole)에 근거한 방법이다. 현재 이러한 방법들에 대한 병렬화 작업이 활발히 진행되고 있다.

2.2 수치 적분

일단 상호작용 모델이 주어지면 어떠한 규칙에 의해서 분자들의 궤적을 갱신하여야 한다. 그 대표적인 것이 분자동력학 모의실험에서 사용하는 뉴우튼 방정식이다. 분자동력학 모의실험은 초기 위치와 속도로부터 시작하여 작은 시간 간격(time step) 단위로 뉴우튼 방정식을 수치적으로 적분하여 분자들의 위치와 속도를 갱신한다. 시간 간격을 작게 할수록 더욱더 정확한 분자의 궤적을 얻을 수 있다. 그런데 분자가 나타내는 운동이 느리게 일어나면 어려움이 발생한다. 왜냐하면 작은 시간 간격으로 긴 시간

동안 모의실험을 수행해야 하기 때문이다. 그렇지 않으면 그 분자의 운동을 컴퓨터 상에서 관측할 수 없게 된다. 그래서 분자의 궤적을 정확하게 얻으면서 동시에 시간 간격을 늘릴 수 있는 방법을 많이 연구해 오고 있다.

고전적으로 많이 사용하는 방법에는 SHAKE/RATTLE 방법이 있다. SHAKE/RATTLE 방법에서는 화학결합의 진동과 같이 빠르게 일어나는 운동을 고정시켜 버림으로써, 다시 말해서, 화학 결합의 길이를 고정시킴으로써 좀 더 큰 시간 간격을 사용할 수 있게 한다. 분자동력학 모의실험에서 보통 사용되는 시간 간격이 1 펨토 초인데 SHAKE/RATTLE 방법을 사용하면 약 2 펨토 초까지 시간간격을 늘리 수 있다.

SHAKE/RATTLE 방법과는 달리 최근 다른 빠르기로 일어나는 운동에 대해 다른 시간 간격을 사용할 수 있게 하는 RESPA 와 같은 다중 시간 간격(multiple time step) 방법이 개발되었다. 예를 들면 진동운동과 같은 빠른 운동은 짧은 시간 간격을 사용하고 분자의 병진운동과 같은 상대적으로 느린 운동은 좀 더 긴 시간 간격을 사용하게 하는 것이다. RESPA 를 사용하면 최대 10 펨토 초 이상까지 시간 간격을 늘리 수 있다는 보고가 있다. 다시 말하면 RESPA 에 의해 약 10 배정도의 계산속도 상승효과를 얻을 수 있다.

분자들의 궤적을 갱신하는 또 하나의 방법은 몬테카를로 모의실험에서 사용하는 Metropolis 알고리즘과 같은 확률론적인 방법이다. 분자동력학 모의실험이 시간에 따른 분자의 위치와 속도를 제공하고 시간에 의존하는 정보들을 제공해 주는 것과는 달리 몬테카를로 모의실험은 시간에 의존하는 정보들을 제공해주지 못한다. 그리고 분자동력학 모의실험은 병렬처리가 용이한 반면 몬테카를로 모의실험은 병렬처리가 효율적이지 못하다. 그래서 슈퍼컴퓨팅과 관련하여 분자동력학 모의실험이 몬테카를로 모의실험보다 많이 이용되고 있다.

아무런 조건이 가해지지 않는다면 뉴우튼 방정식에 의해 산출된 위치와 속도는 항상 계의 총 에너지를 유지 시킨다. 다시 말해서 그 위치와 속도는 일정한 에너지 조건 하에서 산출된 것이다. 하지만 종종 일정한 온도나 또는 일정한 온도와 압력 하에서 위치와 속도를 구할 필요가 있는 경우가 있다. 이런 경우 특별한 알고리즘이 필요하게 된다. 그 대표적인 것이 constraint 방법과 extended system 방법이다. constraint 방법은 역학적으로 앞서 언급한 조건들을 만족하도록 뉴우튼 방정식에 제한을 가하는 방법이다.

extended system 방법은 온도와 압력 각각에 대해 독립적인 운동 자유도를 부여하고 그에 해당하는 운동방정식을 세우는 방법이다. 최근 extended system 방법과 관련하여 일반적으로 적용될 수 있는 이론적인 체계가 정립되었다. 이밖에도 온도와 화학퍼텐셜(chemical potential)이 일정한 조건 하에서 분자동력학 모의실험을 수행하게 하는 알고리즘도 개발되어 있다. 이러한 새로운 알고리즘은 컴퓨터 모의실험을 좀 더 유연하게 한다. 예를 들면 구조의 변화를 수반하는 고체의 상전이 연구가 가능해진다. 그리고 물질의 상평형 연구도 수월하게 된다.

몬테카를로 모의실험도 위에서 언급한 조건 들에서 수행되도록 하는 알고리즘이 개발되어 있다.

2.3 자유에너지 계산

컴퓨터 모의실험으로 얻으려고 하는 정보 중 하나가 자유에너지이다. 왜냐하면 자유에너지는 열역학적으로 어떤 현상이 일어날 지 판단해주는 매우 중요한 양이기 때문이다. 현재 많은 연구자들이 컴퓨터 모의실험을 통해 자유에너지를 효과적으로 계산할 수 있는 방법을 모색하고 있다.

널리 알려진 방법으로는 thermodynamic integration 방법, FEP(free energy perturbation) 방법, slow growth 방법 등이 있다. Particle insertion 방법이나 Umbrella sampling 방법도 많이 사용되는 방법 들이다. 여기서 언급한 방법들은 모두 계가 평형 상태에 있다는 가정 하에서 사용될 수 있는 방법들이다. 그리고 여러 번의 모의실험을 필요로 한다. 예를 들면 열역학적인 상태 1 과 2 의 자유에너지 차이를 계산한다고 할 때 상태 1 과 2, 뿐만 아니라 그 사이의 여러 중간 단계에서도 모의실험을 수행하여야 한다. 중간 단계를 많이 거치면 거칠수록 얻어진 자유에너지 값은 정확해 진다. 하지만 많은 단계를 거치면 그만큼 오차가 전파되어 최종 상태에서의 오차가 커질 수 있다는 단점이 있다. 그래서 두 열역학적 상태 사이의 거리가 너무 먼 경우에는 자유에너지 계산에 주의를 필요로 한다.

최근 자유에너지 계산은 물리적인 상태가 변할 뿐 아니라 계의 화학적 조성이 변하는 경우에도 적용되고 있다. 예를 들면, 단백질 내에 있는 한 아미노산을 다른 아미노산으로 치환했을 경우, 다시 말해서, 단백질에 돌연변이가 일어 났을 경우, 자유에너지의 변화가 그 한 예이다. 계산 중 화학적 조성을 변화시키는 방법에는 현재 single topology 방법과 dual topology 방법이 있다. 둘 다 널리 사용되고 있는 방법이다.

최근 앞서 언급한 방법들 보다 효과적인 새로운 방법이 고안되었는데 이 방법은 평형상태 보다는 비평형 상태에서의 모의실험으로부터 자유에너지를 구한다. 뿐만 아니라 이론적으로 한 번의 모의실험으로도 자유에너지를 계산할 수 있게 한다. 따라서 계산 시간 단축이라는 측면에서 보면 매우 효율적인 방법이라고 할 수 있겠다. 이 방법은 모의실험 동안 외부에서 힘을 가함으로써 현상을 빨리 진행시키게 한다. 예를 들면 두 분자가 결합하고 있는 경우 한 분자는 고정시키고 다른 한 분자는 핀셋으로 움직이는 것과 같이 외부의 힘으로 두 분자의 분리를 가속시킨다. 그런 다음 후처리 과정을 거쳐 자유에너지를 구하게 한다. 현재 이 새로운 방법을 이용한 연구가 많이 보고되고 있다.

최근 WHAM(weighted histogram analysis method)이라고 하는 방법이 고안되었는데 자유에너지 계산 후처리 과정에 사용하면 좀 더 좋을 결과를 얻을 수 있다. 왜냐하면 WHAM 은 수행된 모든 모의실험으로부터의 정보를 활용하기 때문이다.

2.4 양자역학적인 모의실험

지금까지 언급한 방법들은 고전역학적 방법이라는 큰 범주에 속한다. 다시 말해서 양자역학으로 기술되는 현상을 고전적 모의실험으로는 재현할 수 없다는 뜻이다. 예를 들면 화학결합의 생성과 소멸을 수반하는 화학반응 같은 현상을 고전적 모의실험으로는 연구하기 힘들다. 이러한 어려움을 극복하기 위하여 최근 밀도함수론(density functional theory)을 이용한 CPMD(Car-Parrinello molecular dynamics)와 같은 양자역학적인 컴퓨터 모의실험 방법이 고안되었다. 이 외에도 양자역학적인 모의실험에는 파인만의 path integral 이론에 근거한 path integral 모의실험도 있다.

그런데 양자역학 모의실험은 슈뢰딩거 방정식을 풀어야 하기 때문에 고적적인 모의실험에 비해 확연하게 계산속도가 느리다는 단점이 있다. 따라서 최근 많은 관심을 받고 있는 단백질 같은 생체분자에 적용하기에는 아직 어려움이 있다. 대신 최근 연구되고 있는 방법이 QM/MM 방법이다. QM/MM 방법은 고전역학적으로 다룰 수 있는 부분은 고적역학으로 다루고 양자역학으로 다루어야 하는 부분만 양자역학으로 다룬다. 따라서 많은 계산 시간을 절약할 수 있을 뿐만 아니라 비교적 큰 생체분자도 다룰 수 있게 되었다. 최근 효소와 같은 생체촉매 활성부위에서 일어나는 화학반응의 연구 등 QM/MM 을 이용한 연구가 많이 보고되고 있다.

이와 같이 고전적 컴퓨터 모의실험 기법의 발전과 더불어 양자역학적인 모의실험 기법의 발전은 컴퓨터 모의실험의 응용 범위를 한층 확대시키는 전기를 마련하였다.

2.5 병렬처리

컴퓨터 모의실험의 병렬처리와 관련된 알고리즘에는 대표적인 것이 현재 원자분할(atom decomposition) 기법과 영역분할(spacial decomposition) 기법이 있다. 그림 1 에서 보는 바와 같이 원자분할 기법에서는 각 프로세서마다 일정한 수의 원자가 할당되고 그 수는 컴퓨터 모의실험 동안 유지된다. 이 방법은 매번 모든 프로세서가 서로 정보를 교환해야 하기 때문에 프로세서의 숫자가 증가하면 그 병렬처리 효율성이 떨어지는 것으로 알려져 있다. 한편 그림 2 이 보여주는 영역분할 기법에서는 모의실험 상자가 프로세서의 숫자만큼 균일하게 분할되고 프로세서마다 각 영역에 포함되어 있는 원자들이 할당된다. 여기서 각 영역의 크기는 상호작용 제한 거리보다 크게 설정한다. 그러면 이웃한 영역과의 프로세서 들과 정보를 교환하면 되므로 원자분할 기법보다 프로세서간의 정보교환에 따른 계산 시간 손실이 적게 된다.

한편 고전적인 방법으로 원자분할 기법과 비슷한 replicated data strategy 방법이 있다. 예를 들면 프로그램이 루프를 계산할 때 루프를 프로세서 마다 균등하게 배분하는 것이다. 이 경우 원자나 영역에 구애 받지 않고 루프 안에 있는 어떠한 자료에도 적용할 수 있다.

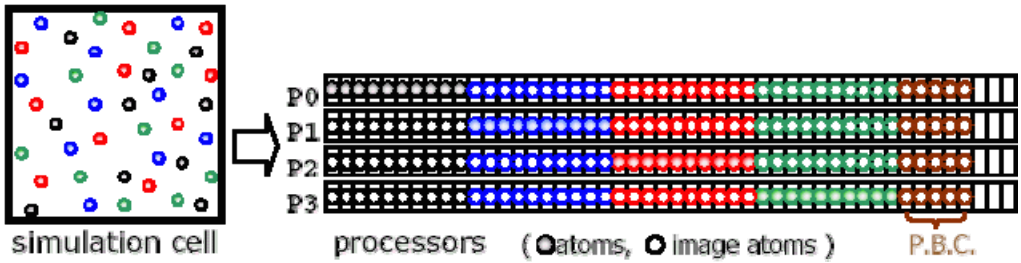


그림 1. 원자분할 기법

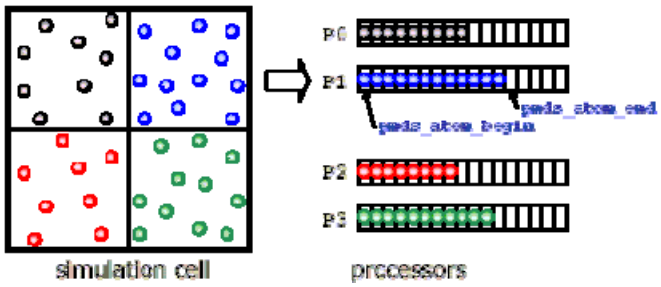


그림 2. 영역분할 기법

그럼 현재 많이 사용되고 있는 분자동력학 모의실험 프로그램들 중 CHARMM, AMBER, NAMD, GROMACS 의 병렬처리 효율성을 살펴보도록 하겠다. 표 1 은 샌디에고 슈퍼컴퓨팅 센터의 Blue Horizon 컴퓨터에서 시험한 것으로서 세 개의 프로그램을 비교해 놓았다. 여기서 사용된 계는 총 23558 개의 원자로 구성돼 있다. 표에서 숫자가 크면 클수록 병렬처리 효율이 좋은 것이다. 표 1 이 나타내는 바와 같이 NAMD 가 전반적으로 좋은 효율성을 보여준다. NAMD 는 다음 절에서도 다시 언급될 것이다.

processors	CHARMM c29b1	AMBER 7	NAMD 2.4	Specifications:
1	2643	2538	2986	
2	1371 (1.9)	1229 (2.1)	1809 (1.7)	
4	724 (3.7)	658 (3.9)	930 (3.2)	
8	414 (6.4)	375 (6.8)	516 (5.8)	
16	255 (10.4)	228 (11.1)	308 (9.7)	
32	195 (13.6)	162 (15.7)	202 (14.8)	
64	182 (14.5)	124 (20.5)	124 (24.1)	
128			126 (23.7)	

- IBM Power3
- 375 Mhz
- MPI vendor
- Mpxlf90

표 1 CHARMM, AMBER, NAMD 병렬처리 비교
출처: http://www.scripps.edu/brooks/Benchmarks/

표 2 는 GROMACS 의 병렬처리 효율성을 보여주는데 총 121856 개의 원자로 이루어진 계에 대해서 여러 컴퓨터 기종에서 시험한 결과를 보여 준다.

Machine	Processors								
	1	2	4	8	12	16	20	24	32
IBM SP2	100%	107%	104%	99%	92%	89%	81%	75%	65%
Linux	100%	114%	94%	76%	60%	53%	-	-	-
Scali	100%	111%	108%	104%	99%	94%	89%	86%	72

표 2 CHARMM, AMBER, NAMD 병력처리 비교
출처: http://www.gromacs.org/benchmarks/scaling.php

병렬처리와 더불어 최근 각광을 받고 있는 것이 분산처리이다. 현재 단백질 접힘 현상이나 단백질과 리간드의 도킹(docking) 등 몇몇 분야에 실제 적용되고 있다.

3. 컴퓨터 모의실험을 이용한 연구 사례

3.1 일리노이 주립대학교 Klaus Schulten 그룹 (www.ks.uius.edu/Research/namd/)

Schulten 그룹은 객체지향 프로그래밍 기법을 사용하여 NAMD 라는 분자동력학 모의실험 프로그램을 개발하였다. 그림 3 은 F1-ATPase 에서 확인한 NAMD 의 병렬처리 효율을 보여준다. 그림이 보여주는 바와 같이 NAMD 는 약 2000 개 의 프로세서까지도 좋은 효율을 기록하고 있다.

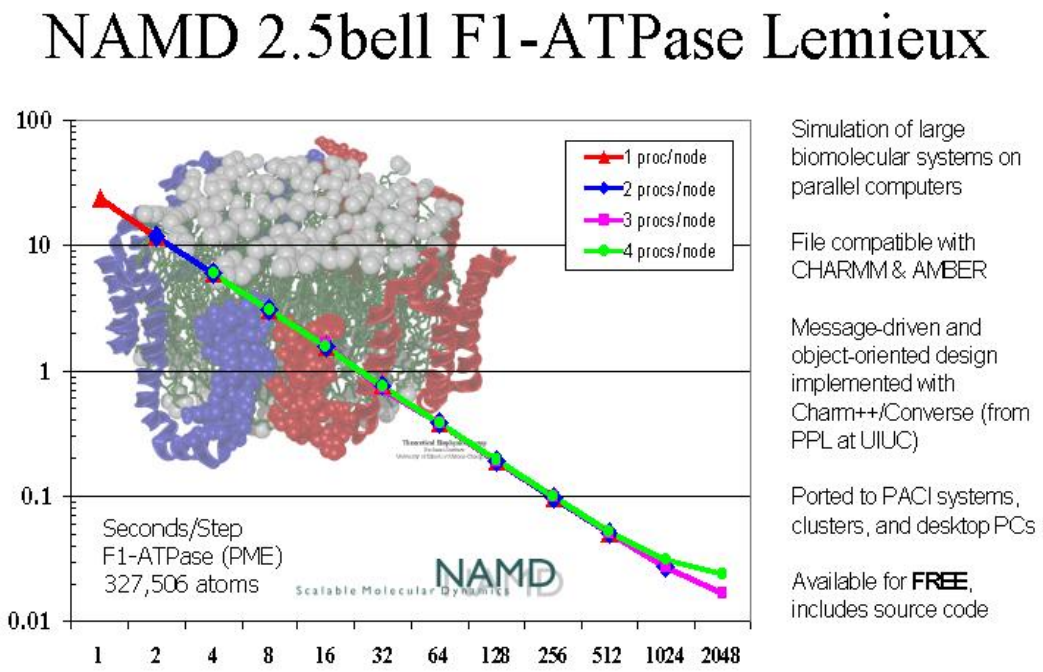
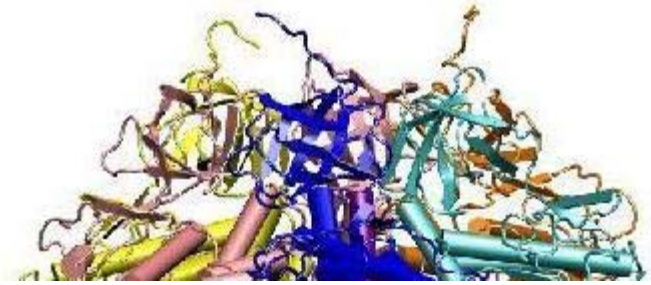
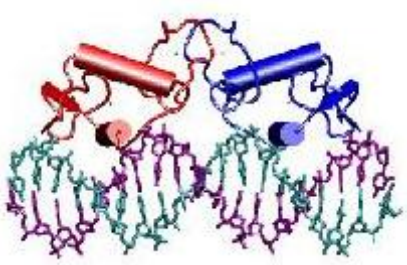


그림 3. NAMD 병렬처리 효율성
그림출처: http://www.psc.edu/~oneal/charts/es45/namd/NAMD -2.5bell -f1atpase -pme -lemieux -pe -node.jpg

NAMD 는 2002 년 미국 볼티모어에서 열린 SC2002 에서 Gordon Bell 상을 수상하였으며 피츠버그 슈퍼컴퓨팅 센터의 LeMieux 에서 초당 1 테라플롭스라는 경이적인 계산속도를 얻었다. Schulten 그룹은 그림 4 에서 보여주는 바와 같이 여러 생체분자들을 연구하였으며 최근에는 단백질을 포함한 총 320,000 개의 원자로 이루어진 계도 연구하였다.

BPTI
3000 atoms





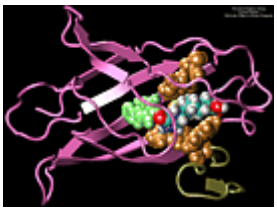
ATP Synthase
327,000 atoms (2001)

Estrogen Receptor
36,000 atoms (1996)

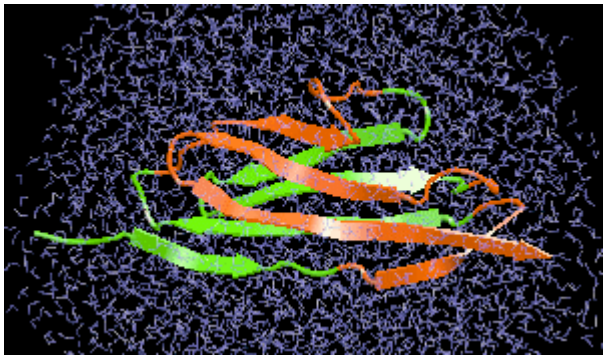
그림 4. NAMD 로 연구한 생체분자들
그림출처: www.ks.uiuc.edu/Research/namd/motivation

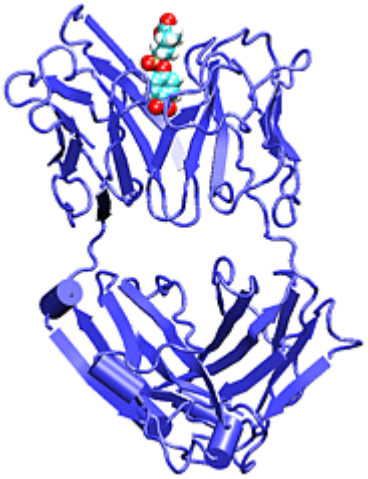
뿐만 아니라 Schulten 그룹은 steered molecular dynamics simulation(SMD)이라는 방법을 이용하여 분자동력학 모의실험을 제어함으로써 고전적인 방법으로는 힘든 생체분자와 관련된 여러 중요한 현상들을 연구하고 있다. 예를 들면, 그림 5 에서 보이는 바와 같이 단백질과 리간드의 상호작용, 근육단백질의 수축 또는 이완, 항원과 항체의 상호작용, 그리고 이온채널에서 이온의 통과 등이 있다.

리간드와 단백질간의 상호작용

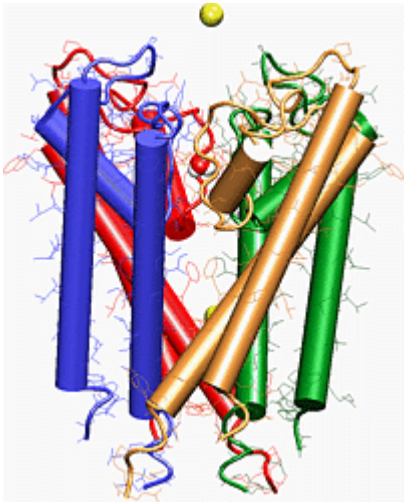


근육의 이완과 수축





항원과 항체의 상호작용



이온채널에서 이온의 통과

그림 5. SMD 응용 사례
그림출처: www.ks.uiuc.edu/Research/smd_imd

SMD 가 가지고 있는 장점은 외부에서 시간에 의존하는 섭동을 가함으로써 우리가 원하는 방향대로 현상을 빠르게 진행시킬 수 있다는 점이다. 최근에는 VMD 라는 가시화 프로그램을 개발하였는데 SMD 와 VMD 를 연동시켜 실시간으로 분자동력학 모의실험을 제어하고자 하는 프로젝트를 추진하고 있다.

3.2 스탠포드 대학교의 Vijay Pande 그룹 (www.stanford.edu/group/pandegroup/folding/)

Pande 그룹은 최근 분산처리 방법을 이용하여 단백질의 접힘 현상을 연구하고자 하는 프로젝트 (folding@home)를 추진하고 있다. 2000 년 10 월 1 일 이후 지금 현재 전 세계적으로 400,000 개 이상의 컴퓨터들이 folding@home 프로젝트에 참가하고 있다. (그림 6 참조)

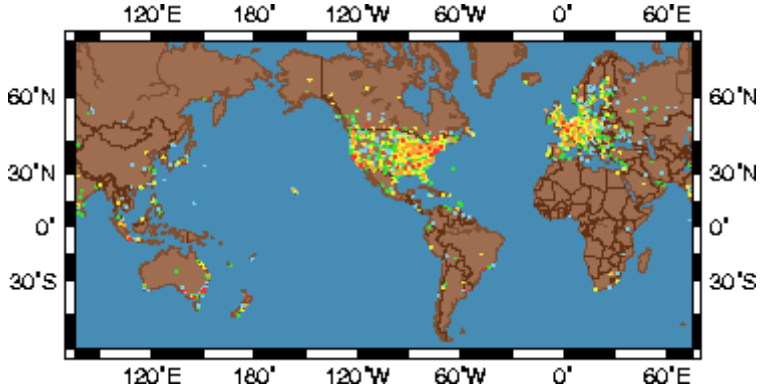


그림 6. folding@home 에 참가한 컴퓨터 현황
그림출처: www.stanford.edu/group/pandegroup/folding

Pande 그룹은 현재 분산처리 방법을 이용하여 여러 단백질의 접힘 현상에 대해 컴퓨터 모의실험을 수행하고 있다. 그림 7 는 컴퓨터 모의실험을 통해 얻어진 단백질 접힘 시간과 실험적으로 측정된 시간을 비교해 놓은 그림이다. 그림에서 보는 바와 같이 컴퓨터 모의실험을 통해 얻어진 결과는 실험으로 얻어진 결과와 근사하게 일치함을 알 수 있다.

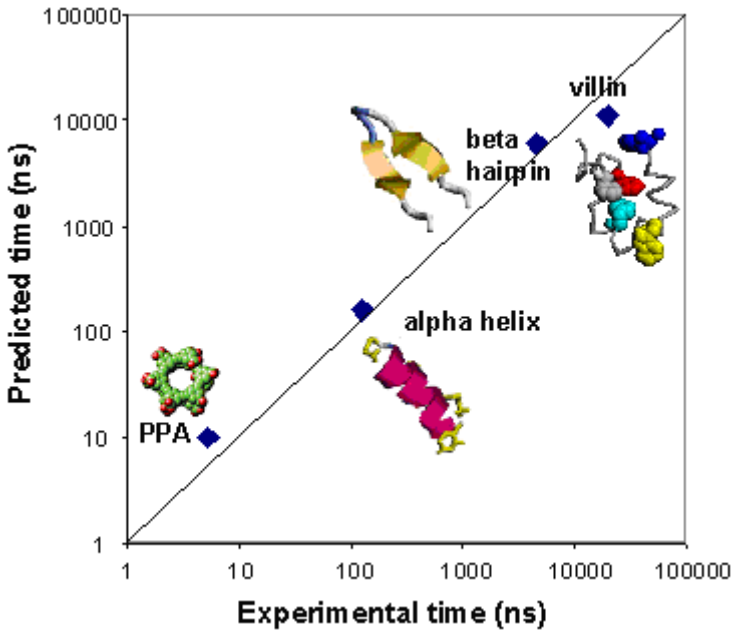


그림 7. 컴퓨터 모의실험으로 얻어진 단백질 접힘 시간과 실험적으로 측정된 단백질 접힘 시간과의 비교
그림출처: www.stanford.edu/group/pandegroup/folding/results.html

하지만 비교적 큰 단백질 (50 개 이상의 아미노산을 포함하는 단백질)에 대한 결과는 아직 보고되고 있지 않다.

3.3 CPMD consortium (www.cpmd.org)

스위스 공과 대학의 Parrinello 그룹과 IBM 의 Andreoni 그룹에 의해서 개발된 CPMD 프로그램은 현재 많은 사람들에 의해서 사용되고 있다. 고전적인 컴퓨터 모의실험이 고적역학에 의해 기술되는 현상만 다룰 수 있는 것과는 달리 CPMD 는 양자역학적인 현상까지도 다룰 수 있다는 장점이 있다. CPMD 프로그램은 많은 개선을 거듭하여 현재 버전 3.7 까지 발전하였다. 그림 8 은 IBM SP 기종에서 CPMD 프로그램의 병렬처리 효율을 나타낸다. 그림이 나타내는 바와 같이 프로세서 약 50 개 까지 좋은 효율성을 보여주고 있음을 알 수 있다.

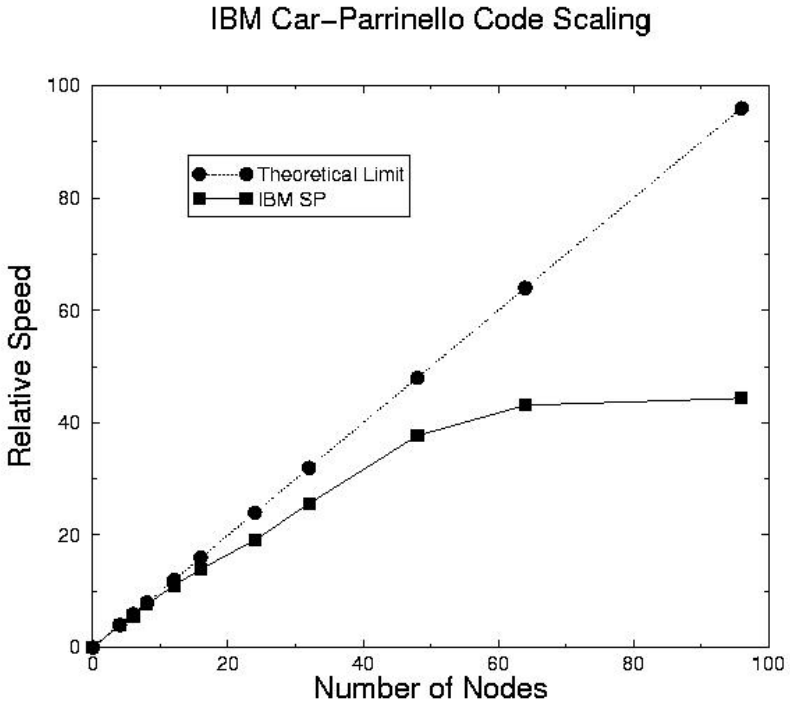


그림 8. CPMD 의 병렬처리 효율성
그림출처: www.hec.utah.edu/~rbevans/scaling.html

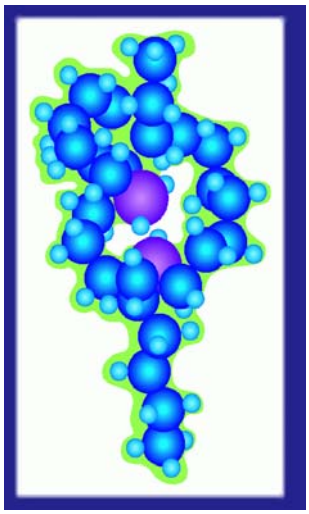


그림 9. CPMD 을 이용한 고분자 매듭의 연구
그림출처: http://www.cmm.upenn.edu/research/index.html

그림 9 는 라디칼 상태에 있는 고분자가 형성할 수 있는 매듭을 보여준다. 여기서 고분자는 28 개의 탄소 원자와 58 개의 수소 원자로 구성돼 있다. 이외에도 CPMD 는 응집상태에서의 화학반응이나 반도체 같은 재료의 연구에도 유용하게 쓰이고 있다. 이와 더불어 CPMD 는 생체분자 QM/MM 의 계산에도 이용된다. 그림 10 는 Alcohol dehydrogenase 의 활성부위에서 아연의 화학에 대한 결과를 보여주고 그림 11 은 리간드의 결합에 따른 HIV -1 Protease 의 활성부위 구조의 변화를 보여준다.

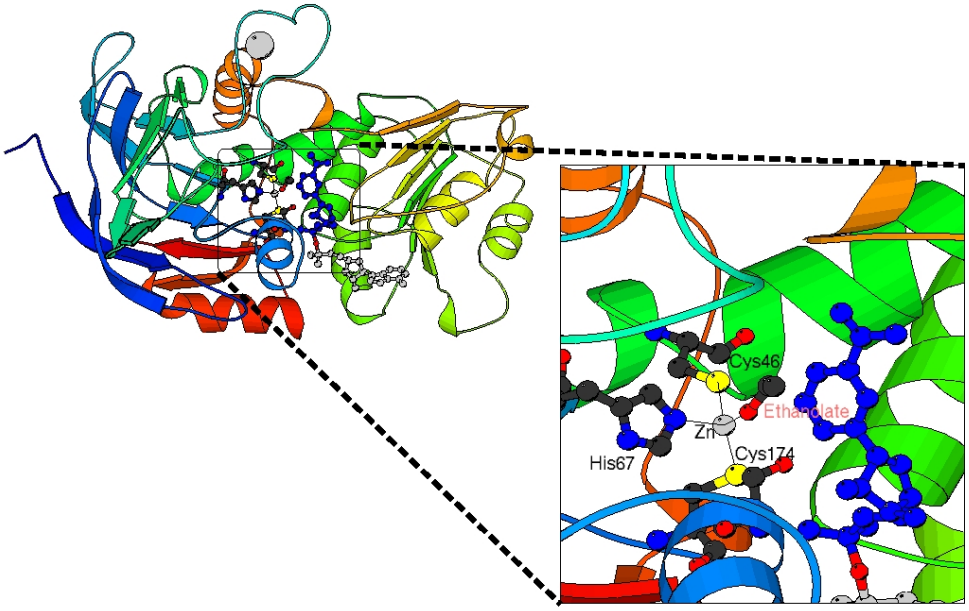


그림 10. CPMD 을 이용한 Alcohol dehydrogenase 의 활성부위 연구
그림출처: www.sissa.it/cm/bc/research/research_main.html

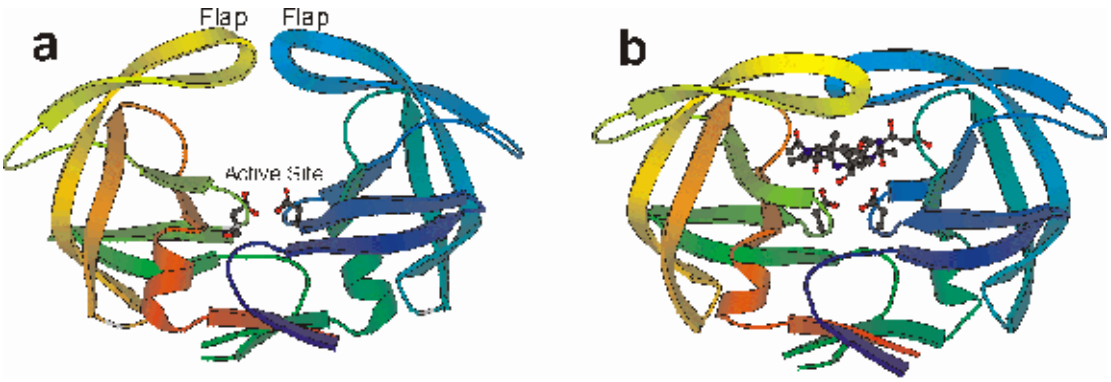


그림 11. CPMD 을 이용한 HIV -1 Protease 의 활성부위 연구

그림출처: www.sissa.it/cm/bc/research/research_main.html

3.4 Pacific Northwest National Laboratory(PNNL)의 MS^3 프로젝트(http://www.emsl.pnl.gov/docs/nwchem/nwchem.html)

PNNL 은 강력한 슈퍼컴퓨팅 자원을 효율적으로 이용하여 사용자가 손쉽게 계산을 할 수 있는 환경을 구축하는 MS^3 프로젝트를 추진하고 있다. MS^3 는 크게 세가지 프로젝트로 구성되어 있다. 첫째가 강력한 계산화학 프로그램(NWCHEM)의 개발이고, 둘째가 슈퍼컴퓨팅 자원을 효율적으로 할 수 있게 하는 병렬처리 도구(Parsoft)의 개발이고, 셋째가 NWCHEM 을 손쉽게 사용할 수 있게 하는 사용자 인터페이스의 개발(Ecce)이다. 그림 12 는 MS^3 프로젝트가 추구하는 distributed client-server 모델을 도식적으로 보여주는 그림이다. 자신의 컴퓨터로 원거리에 있는 프로그램을 원격으로 구동시키고 그 결과를 얻은 후 그래픽 사용자 인터페이스를 통해 분석하는 전 과정을 보여준다.

NWCHEM 은 양자화학의 전자구조 계산에서부터 앞 절에서 설명한 생체분자의 컴퓨터 모의실험에 이르기 까지 계산 화학 전반에 걸쳐 일반적으로 사용될 목적으로 개발되고 있다. 여기서 병렬처리는 Parsoft 를 통해 구현되었는데 그림 13, 그림 14, 그림 15 가 보여주는 바와 같이 양자화학 계산이나 컴퓨터 모의실험에 대해 모두 좋은 효율성을 보여주고 있다.

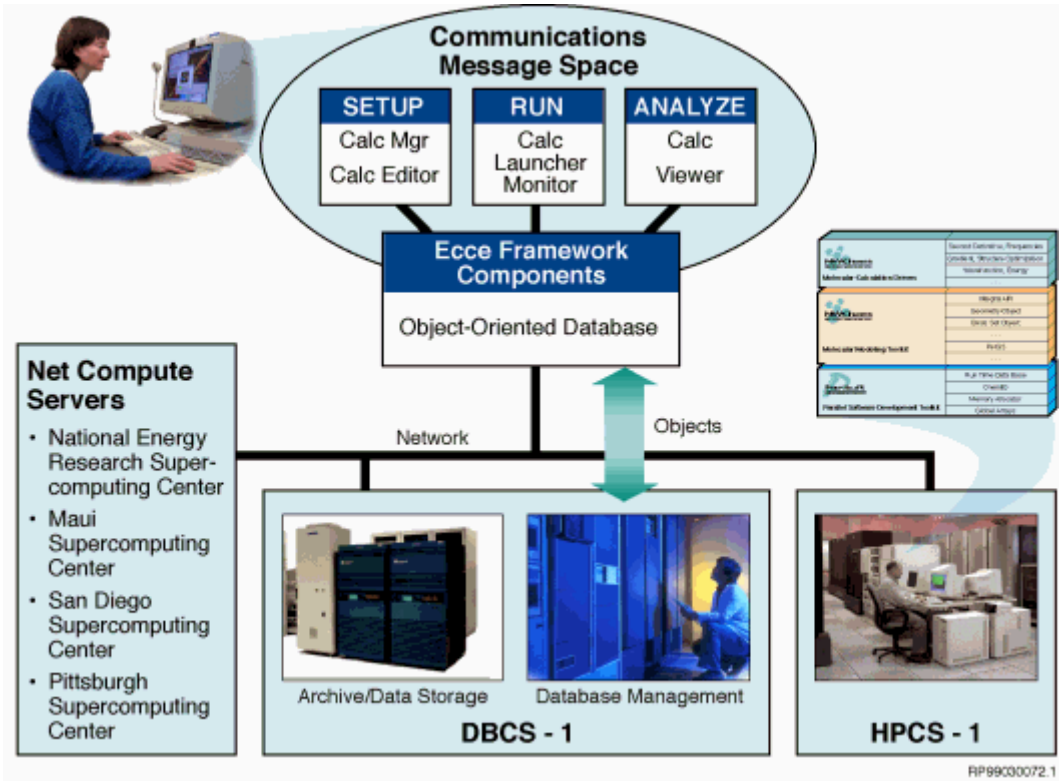


그림 12. MS[3] client -server model

그림출처:

http://www.emsl.pnl.gov/capabs/mscf/?/capabs/mscf/software/ms3.html

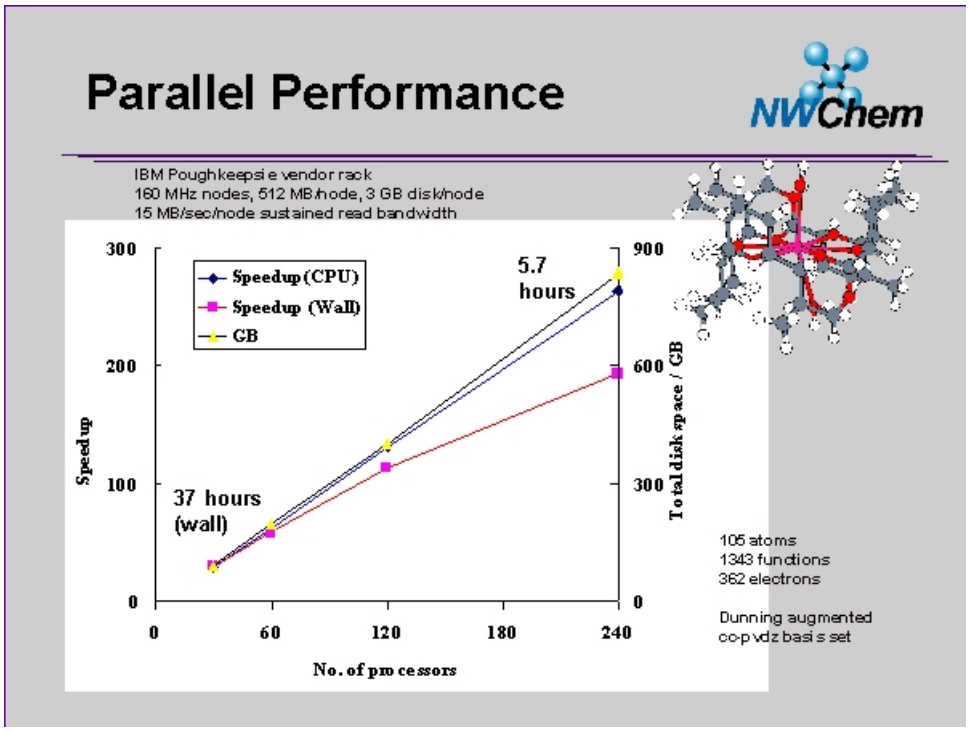


그림 13. SCF 계산에 대한 병렬처리 효율성
그림출처: www.emsl.pnl.gov/docs/nwchem/nwchem.html

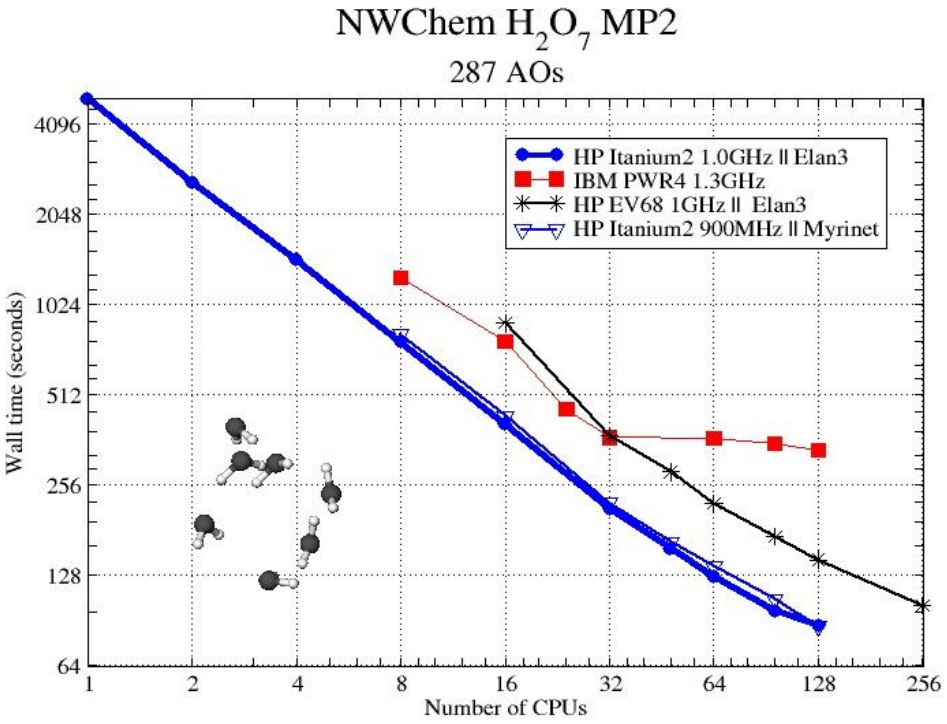


그림 14. MP2 계산에 대한 병렬처리 효율성
그림출처: www.emsl.pnl.gov/docs/nwchem/nwchem.html

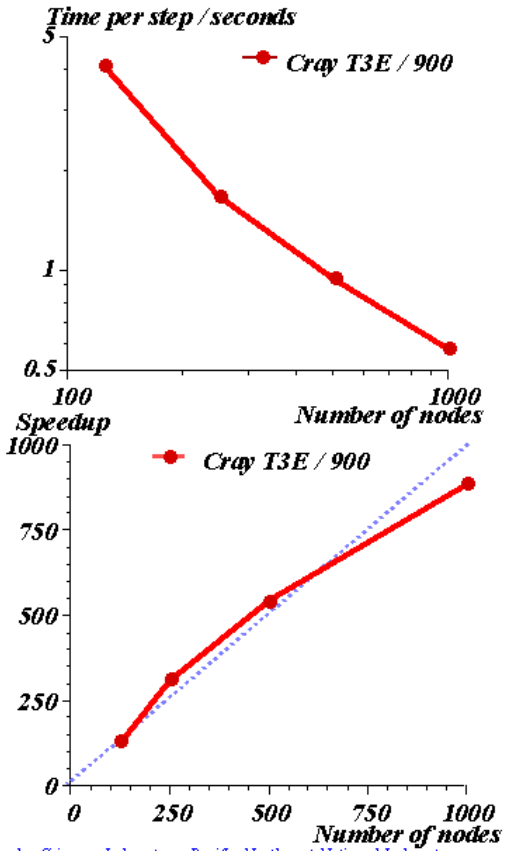


그림 15. 옥탄올에 대한 컴퓨터 모의실험 병렬처리 효율성
그림출처: www.emsl.pnl.gov/docs/nwchem/nwchem.html

이러한 노력의 결과 MS^3 프로젝트는 1999 년 R&D Magazine 이 수여하는 R&D 100 Award 를 받았다.

4. 결론

지금까지 살펴본 바와 같이 컴퓨터 모의실험은 많은 발전을 거듭하였으며 슈퍼컴퓨팅의 발전과 더불어 그 응용범위가 다양해졌다. 새로운 컴퓨터 모의실험 기법의 개발과 슈퍼컴퓨팅의 발전을 통해 앞으로 현재의 한계를 넘어 컴퓨터 모의실험의 응용범위를 더욱 확대하고자 많은 연구자들이 노력할 것으로 기대된다.