슈퍼컴퓨터의 기술발전추세와 미래

유 여 백
(몬타나주립대 전자계산학과 조교수)

1. 서 론

한동안 상아탑에 머물러있던 슈퍼컴퓨터는 이제 모든 과학기술 community로 급속히 보편화되고 있다. 현재의 슈퍼컴퓨터 성능은 Giga FLOPS(billion floating-point operations per second) 단위로 말해지나 머지않은 장래에는 Tera FLOPS(trillion floating-point operations per second) 단위로 말해질 것이다. (1 Tera FLOPS 컴퓨터 한대의 power는 약 1000만개 personal computer의 힘을 모은것에 해당한다) 이 목표를 향하여 미국, 일본을 비롯한 여러 나라에서 각 방향의 연구가 활발히 진행되고 있다. 예를들어 90년대초에 나올 cray-4는 64개의 processor, GaAs(gallium arsenide) chip 등을 활용하여 cray-1의 1000배에 해당하는 throughput을 낼 것으로 예상한다. 현재의 컴퓨터에서 내일의 Tera FLOPS 컴퓨터로 가기 위해서는 1000배 이상의 performance 인상이 필요하다. 새로운 소재와 기술의 개발, 그리고 parallel processihg을 이용하면 Tera FLOPS 컴퓨터가 이론적으로 가능하다. 한편 이제까지의 electronic computer가 쓰는 computing method와 근본적으로 다른 computing method가 등장하여 Supercomputing의 영역을 넓히고 있다. 현재의 발전추세로 보아 다음 세기초에는 neural network, optical computer, molecular computer같은 이색적인 computer를 보게될 것이다. 다음 장에서는 현재 진행되고 있고, 미래에 일어날 것으로 예상되는 기술발전추세를 여러가지 슈퍼컴퓨터와 미니슈퍼컴퓨터별로 나누어서 살펴보고 이어서 미래에 등장할 것으로 예상되는 새로운 슈퍼컴퓨터들을 소개한다.

2. 슈퍼컴퓨터 기술발전추세

2.1 슈퍼컴퓨터

2.1.1 반도체기술

앞으로 10년안에 4~16 megabit Dynamic RAM이 등장하게 된다. Logic gate 밀도가 높아져서 현재 ECL array당 10,000개 gate이던 밀도가 머지않아 15,000~20,000개 gate로 늘어나게 된다. 이 과정에서 optical interconnection 기술이 널리 쓰이게 될 것이다. 또한 TAB(tape-automated bonding)이라는 새로운 packaging 기술이 개발되고 있는데 이 방법을 쓰면 printed-circuit board에 더많은 IC dice를 넣을수 있게 된다. Chip이 이렇게 complex하게 됨에 따라 2000년까지는 1000개의 vax 780의 power에 해당하는 chip를 만들 수 있게 될 것이다. 현재 cray는 GaAs를 통한 기술개발을 계속하고 있다. (전자는 GaAs를 통하면 Silicon에서보다 10배까지 빠르게 움직인다) GaAs chip으로 만들어질 Cray-3은 cycle time이 2 ns(nanosecond), memory cycle

time이 20 ns가 될것이다.(현재 Cray-2의 cycle time은 4.1 ns, memory cycle time은 80 ns이다).

2.1.2 parallel processing

현재 슈퍼컴퓨터성능의 가장 중요한 원동력은 Extended Fortran과 vector의 고속연산에 알맞게 고안된 pipelined vector processor이다. 앞으로 슈퍼컴퓨터성능향상은 개량된 vector processor에서보다 parallelism을 높이는데서 나올 것이다. 앞으로 10년간 기술발전및 parallel processing을 통하여 1,000배의 performance gain이 가능할 것으로 예상된다. (parallel processing에 잘 맞는 문제에 한함) Cray-3은 16개 processor에 최고 16 GFLOPS의 연산속도를 가지고, Cray-3에 이어 등장할 Cray-4는 64개 processor에 최고 128 GFLOPS의 연산 속도를 가지게 된다. 일본 NEC의 SX-3은 4개 processor에 20 GFLOPS를 예상하고, Fujitsu는 1990년을 목표로 4 GFLOPS짜리 VP 2600/10을 개발중이다.

parallel processing을 극대화한 것으로 소위 "massively parallel computer"가 있는데 Thinking Machine, Inc.의 Connection Machine, CM-2가 대표적이다. CM-2는 SIMD(single-instruction, multiple-data stream) machine으로 한 instruction을 동시에 65만개 processor에 전달하면 각 processor는 각기 private memory에 있는 data에 그 instruction을 적용하게 된다. CM-2의 현재 최고속도는 2,500MIPS (million instruction per second)이며 더 parallelism을 짜낼 수 있는 software 개발이 진행중이다.

*주기억용량 : 1990년초에는 주기억용량이 1 giga word로 늘어나게 될 것으로 예상된다.

2.1.3 Network

workstation, minisupercomputer, supercomputer를 함께 묶을 수 있는 고속 network의 필요가 증대하고 있다. 현재 NSFNet의 경우 초당 1.5megabit (T1의 경우) 또는 45megabit(T3의 경우) 정도인데 앞으로 초당 3 gigabit의 정보를 전달할 수 있는 초고속 Network이 계획되어 있다.

2.2 Hypercube

Mini supercomputer 분야에서 가장 popular한 Architecture는 Hypercube topology이다. Hypercube은 각 corner마다 한 node(processing element)씩 자리잡고 있는 고차원의 입방체로 생각할 수 있다. 현재 10차원, 즉 1024개의 32-bit processing element로 구성된 Hypercube이 나와있으며 (NCUBE 10), 수년내에 16차원, 즉 16000개 이상의 node를 가지는 Hypercube이 등장할 계획이다. 앞으로 memory chip의 비용이 떨어짐에 따라 1000만 byte의 memory를 가지는 node도 가능해질 것이다. 이렇게 memory가 커지면 node당 저장할 수 있는 data가 많아지고 따라서 더 많은 계산이 node내에서 끝나게 되므로 node간의 communication이 줄어들게 되어 efficiency가 커지게 된다.

각 분야의 기술이 빠르게 발전하고 있어서, 앞으로 Hypercube이 Large-Scale computing을 위한 표준 Architecture가 될지는 아직 미지수이다. 그리고 주어진 문제를 작은 단위로 나눠서 균형있게 각 node에 분배하는 방법에 대한 연구가 이제 겨우 시작되었다. 앞으로 Hypercube이 Supercomputing의 주류에 들어서기 위해서는 위에 언급한 문제가 해결되어야 할것이다.

2.3 Graphics Supercomputer

초기의 Supercomputing에서는 복잡한 simulation 결과가 긴 숫자로만 표시되어서 결과를 읽고 분석하기가 어려웠다. Graphics Supercomputer가 나옴에 따라 이제는 계산결과를 Graphics로 볼 수 있게 되었다. Graphics Supercomputer를 쓰면 engineer는 계산과정을 시각적으로 표현하고 design을 쉽게 바꾸고 또 결과를 금새 볼 수 있게 된다. 예를 들면, 자동차 위의 공기의 흐름이 designer가 차체의 모양을 바꿈에 따라 어떻게 다르게 움직이는가를 눈으로 볼 수 있게 된다. Graphics Supercomputer는 최근에 발달하였으며 아직은 전통적인 Supercomputer에 비하면 성능이 상대적으로 약하나, 현재의 빠른 발전속도로 보아 지난날의 Workstation같이, 곧 Supercomputer의 주류에 들어오게 될 것으로 전망된다. 현재 graphics supercomputer가 쓰이고 있는 분야는 animation, computer-aided design, image processing, real-time simulation, scientific visualization 등이다. 최근에 나온 Alliant/Raster의 VFX82

는 Vector processor를 갖춘 3-D graphics workstation이다. Peak performance는 377 MFLOPS에 달한다. 이제 Visualization은 minisuper급을 지나서 supercomputer급으로 진입하고 있다. 앞으로 얼마 안가서 graphics supercomputer의 성능은 Gray X-MP, 크기는 책상정도, 가격은 10만불 이내가 될것으로 예상된다. 이렇게되면 graphics supercomputer가 널리 보급되어 전기 및 기계 Design, 유체역학, 지진학, 분자 modeling 등에도 쓰이게 될 것이다.

2.4 Dataflow computer

Dataflow architecture는 근본적으로 control-flow Architecture(Von Neumann computer)와 다르며, 많은 종류의 computation에 내재되어 있는 parallelism을 효과적으로 활용할 수 있다. Dataflow computer에서는 operand가 준비되는대로 Instruction이 수행되게 된다. Dataflow computer는 소위 "single-assignment language"라는 high-level language로 program을 짜게 된다. (VAL이 대표적) "single-assignment language"의 기본 rule은 각 variable이 program segment의 어느 한 statement의 왼쪽에만 나타난다는 것이다. 그래야만 program에서의 data dependence가 쉽게 나타나서 program내 statement의 순서에 관계없이, data만 준비되면 해당 statement가 수행되는 것이다(data-driven). 이와같이 intuitive한 appeal에도 불구하고, dataflow computer의 개발속도가 늦은데, 다른 parallel computer architecture와 경쟁하려면 더 많은 연구가 필요하다. 현재 대표적인 dataflow computer인 Cydro 5의 성능은 Cray X-MP의 절반정도이다.

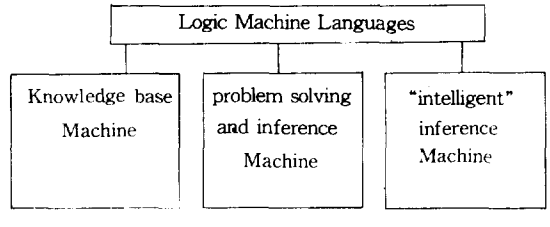
2.5 Systolic Array

Systolic array는 regularity와 modularity가 area-efficient layout에 꼭 필요한 VLSI/WSI architecture의 generation에 속한다. Systolic array는 pipeline과 같이 작동한다 : data가 한 operation을 수행하는 processor에 보내진다. 그 processor가 처리한 결과는 다음 processor에 전달되어 다른 operation이 수행되며, 이런 seguence가 systolic array의 끝까지 되풀이된다. 현재의 VLSI기술로는 feature size가 1에서 3 micron인 circuit의 제조가 가능하다. 이것은 50만개 transistor를 가진 chip에 해당한다. 2000년까지는 advanced optical, 전자 빔, 이온 빔, 또는 X-ray lithography 등을 써서 한개 chip에 10억개의 transistor를 올려 놓을 수 있는 기술이 개발될 것이다. 세련된 design 및 programming tool의 개발과 systolic array에 맞는 application과 algorithm이 개발되면, signal processing, 고속연산, symbolic processing, intelligent database분야에서 systolic array가 널리 쓰이게 될것이다.

2.6 차세대 Computer(제 5 세대 Computer)

차세대 computer는 knowledge-based expert system을 support하기 위해 design된 knowledge information processing system이다. Expert system은 human expertise의 전문분야에 대한 잘 짜여진 지식의 module을 포함하고 있다. Expert system의 구조는 Logic programming의 구조와 잘 일치된다. 1981년 일본은 "intelligent"한 computer를 위한 기초를 쌓고 일본을 전자계산연구분야의 leader로 끌어올리기위한 10개년계획의 제 5 세대 computer project를 발표하였다. 그들이 추구한 computer는 large-scale knowledge-base를 처리하고 인간과 interact함으로서 인간이 수행하고 생각하는 방식을 바꿀 수 있게 하는 computer였다. 제 5 세대 computer는 세개의 전문화된 machine과 이것을 한데묶어 연결해주는 logic machine language와 architecture로 볼 수 있다.

일본의 제 5 세대 computer project는 미국, 영국, EEC 등에 비슷한 project를 낳게 하는 원인이 되었다. 1981년 이후 5억불이상이 차세대 computer연구에 쓰여졌다. 이런 노력은 간단한 expert system이나 소규모 multiprocessing에 이미 알려진 Technigue을 적용하여 많은 진보를 가져왔다. 예컨대 AI



언어가 parallel computer에 맞게 확장되었다. symbolic processing의 궁극적인 언어는 logic이며, combinational logic은 원래 parallel processing에 부합된다. 현재 순수하게 discriptive라고 highly parallel한 parallel programming paradigm이 순수한 logic 언어에 기초하여 개발되고 있다. 그러나 아직도 일본이 제5세대 computer project를 시작하면서 제시한 breakthrough를 가져온 project는 없는 실정이다. 궁극적으로 차세대 computer는 종래의 computer design 방식과 다른 logic machine이 될것이나 1990년대에는 decentralised된 control flow architecture가 될 것으로 예상된다.

3. 미래의 Supercomputer

3.1 Neural Network

현재 Neural Network이라는, 종래의 Algorithmic 방식과는 다른, 정보처리 모델이 연구중이다. Neurobiology의 발전과 massively parallel VLSI implementation 기술이 발달함에 따라 Neural Network에 관한 관심이 높아지게 되었다. Neural network은 인간두뇌의 구조를 model로 하고 있다. 뇌에는 neuron이라는 100억개의 뇌세포가 있는데, 각각의 뇌세포는 만개의 다른 뇌세포와 연결되어 있어서 서로 자극을 주거나 받는다. 어떤 연결은 강하고, 어떤 연결은 약하다. 뇌가 input을 받으면 미리 program되어진 구조에 따라서 반응하기도 하고 어떤 때는 스스로 그 반응에 맞게 자기 자신을 organize하기도 하는데 이것이 바로 Learning process인 셈이다.

현재 쓰이고 있는 거의 대부분의 computer는 "stored program" 방식에 따른 Von Neumann형 computer이다. 따라서 computer가 어떤 정보처리 기능을 수행하려면 누군가가 그 기능을 이해하고 그 기능을 수행할 수 있는 program을 짜야 한다. 그런데 어떤 종류의 복잡한 문제는 이러한 계산처리과정을 쓴다는 것이 거의 불가능한 경우가 있다. 예컨대 아직까지 복잡한 pattern recognition 문제—handwritten-character reader, 적과 아군의 비행기를 식별해내는 system, continuous speech를 recognize하는 system—를 해결할 수 있는 Algorithmic software가 없다. 이런 문제들은 다음과 같은 공통점이 있다.

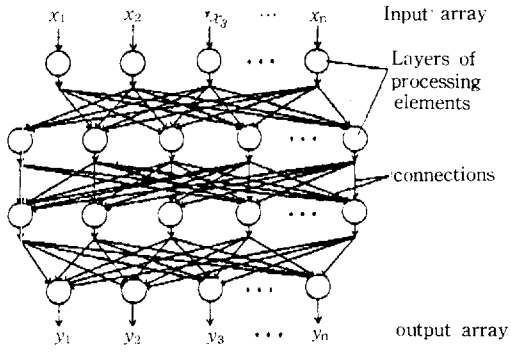


그림 1. Neural Network[5]

① 인간에게는 이런 문제가 복잡한 계산문제보다 오히려 쉽다.

② 수많은 종류의 example이 있을 수 있다.

③ 각 task는 결국 한쪽의 object를 다른쪽의 object에 associate하는 일이다. 예를 들어 computer가 문장을 읽을 수 있게 되려면, 주어진 문자, 빈칸, 구두점 등을 각각에 해당된 소리, 쉼, 억양에 associate(mapping)할 수 있어야 한다는 말이다. Neural network은 이러한 association을 지배하는 내규를 만들어서 그 결과를 실제 Example과 비교하면서 향상해 나간다. 이렇게 시행 착오를 통해서 neural network 스스로 자기를 가르칠 수 있게 된다.

그림 1에서 보는 바와 같이 neural network에 input data가 들어오면 첫번째 Layer에 있는 processor들이 transfer function에 따라서 병렬방식으로 처리하고 그 결과를 다음 Layer에 위치한 processor들에 보내게 된다. 결국 output array에 나타난 결과는 input에 associate된 characteristic인 셈이다.

neurocomputing은 algorithmic programming을 대체할 수 없다. Neurocomputing과 종래의 computing방식은 서로 compatible하지 않다. 따라서 Neural network을 과학계산용으로 써서는 안된다. 현재 10여개의 neural network이 나와 있으나 neuro computing 자체가 아직 초창기여서 사용범위는 극히 제한되어 있다. 한가지 어려움은 neural network을 어떻게 효과적으로 program해야 할지 잘 모른다는 것이다. 그러나 앞으로 neurology와 VLSI

technique가 더 발달하고 neurocomputing을 위한 software가 개발되면 실용적인 real-time neural network system이 등장하게 될것이다.

3.2 Optical Computer

Neurocomputing과 더불어 발전하고 있는 또하나의 기술이 optical computing이다. 종래의 전자와는 다른 광자의 성질을 최대한 이용하여 인간의 뇌와 같이 생각하고 행동하는, 이제까지의 computer와 전혀 다른 새로운 computer architecture의 연구가 활발히 진행되고 있다. optical computer의 장점은 electronic computer를 제작할 때 극복해야 하는 connectivity와 bandwidth 문제를 빛이 가지고 있는 parallelism이라는 속성을 이용하면 쉽게 극복할 수 있다는 것이다. Lens나 거울은 수백만개의 input과 output을 아무런 핀이나 줄, 또는 금속 joint없이 연결시킬 수 있기 때문이다. Hologram을 이용하면 1조개의 연결이 가능하다. 이러한 Hologram덕분에 소위 "random problem"을 풀 수 있다는 것이 확인되었다. ("Random problem"이란, algorithm으로 나타내기에는 너무 복잡한 pattern recognition같은 문제를 말한다) pattern recognition같은 종류의 문제를 푸는 가장 좋은 방법은 사진에서 각 pixel을 동시에 처리하는 것이다.

optical computing에서 system component를 외부적으로 연결하는 것은 단순한 문제이나 cpu내에서는 전기 energy를 optical energy로 바꾸고, 다시 전기 energy로 바꾸는 과정에 소요되는 시간때문에 심각한 latency 문제가 생긴다. 현재 진행되고 있는 여러 research중 Holmdel의 Lawrence West는 non bistable optical logic gate를 연구중이다. west의 device는 새로운 양자현상인 QWEST(quantum-well envelope state transition)을 최대한 이용하도록 되어있다. 소재는 GaAs로 된 여러겹의 quantum-well 구조로 구성되어 있어서 전자가 GaAs conduction band의 두 최저 energy level 사이에서만 움직일 수 있도록 design되어 있다. 이렇게해서 얻어지는 공명효과는 자연에서 발견되는 어떤 물질보다 적어도 1000배 이상 빛의 흡수력의 변화를 감지할 수 있게 된다. 이 Transition는 매우 좁고, 온도에 영향을 받지 않는다. QWEST효과로 Switching speed가 겨우 수pico second까지 달할 수 있게 된다.

West의 최종 목표는 현재의 cray computer보다 10배 빠른 "optical cray"를 만드는 것이다. 그러나 IBM의 Keyes같은 과학자는 optical computing에 관해 회의적이다. Keyes는 광자의 독특한 성질과 optical technique의 한계때문에 범용 digital optical computer를 만드는 것이 불가능하다고 생각한다. 다른 과학자들은 optics와 electron은 서로 보완적이라고 생각한다. 그래서 궁극적으로 electronic computer와 optical computer를 합친 hybrid형 computing으로 가게 될것으로 생각한다.

현재 미국, 일본, 유럽, 소련 등지에서 optical computing에 관한 연구가 활발히 진행되고 있는데, 미국에서는 SDI(Strategic Defence Initrative)로부터 optical computing에 관한 연구가 크게 support되고 있다. SDI에서 이분야에 관심을 쏟는 이유는 광자는 전자와는 달리 핵폭발시에도 EMP(electro-magnetic pulse)효과의 영향을 받지 않기 때문이다.

3.3 Molecular Computer

Molecular computer는 silicon IC대신 단백질과 여러 분자들로 구성된다. Molecular computing의 기본적인 개념은 생물분자정보처리 시스템이 개발된 1970년대에 정립되었으나, 본격적인 연구는 1980년대에 들어와서 시작되었다. 1980년대에 들어 많은 과학자들이 biosensor, 단백질 공학, DNA재조합기술, polymer화학, 인공막 등 molecular computer제조에 필요한 기술을 발전시켰다. Molecular computing이라는 새로운 computing방식을 가져오게 된 중요한 feature는 그것이 context-dependent하다는 것이다. 따라서 입력을 bit-by-bit으로 처리하지 않고 dynamic physical structure로 한번에 처리할 수 있게 된다. molecular computer는 단백질 효소의 형상인지능력을 이용하여 광자 또는 electric pulse 등 nontactile input signal을 처리하는데 이것이 가능하게 되기 위해서는 input signal이 효소가 인지할 수 있는 분자적, 생리화학적인 형태로 바뀌어져야 한다. 효소가 기질(substrate)을 인지하는 과정은 key와 lock이 들어 맞는 것과 비슷하다(그림 2).

효소 자체는 다른 모양 상태로 변함으로서 분자 level에서의 control과 memory가 가능하게 된다.

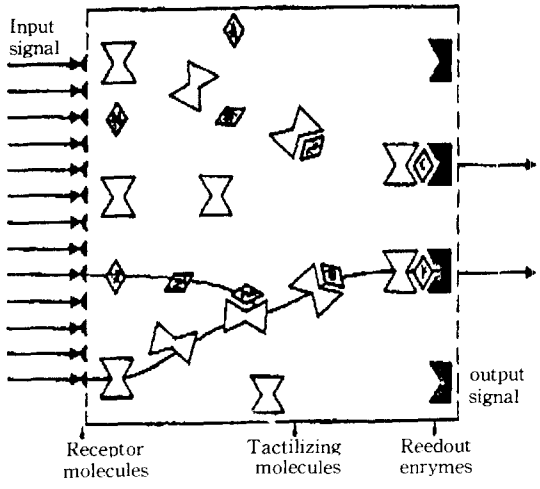


그림 2. Molecular Computing[9]

molecular computer에서 말하는 "계산"이란 이와같이 단백질분자의 움직임과 그 생리화학적인 환경의 역학을 말한다. 따라서 program 자체가 단백질 구조에 내포되어 있는 것이다. 현재 molecular computing연구는 형태를 보고, 표면을 느끼며, 화학적인 전위차를 감지하고 나서 얻어지는 정보를 bit-by-bit process를 위해서 전통적인 computer에 공급하는 intelligent한 preprocessor개발에 주력하고 있다. Molecular computing은 digital machine에 생물학적 능력을 갖게 해준다. 현재는 생물공학 intrastructure가 아직 molecular computer제작에 필요한 물질을 제공할 정도로 발전되어 있지 않지만 몇년후에는 간단한 정보처리를 할 수 있는 molecular device가 제작될 것으로 기대된다. 그러나 완전한 능력을 갖춘 molecular computer는 다음세기 전에는 제작되기 힘들 것으로 전망된다.

4. 결 론

지금까지 Vector supercomputer를 비롯한 여러종류의 supercomputer의 기술발전 추세를 간단히 살펴보았다. 앞으로의 Supercomputer는 VLSI기술의 발달, GaAs같은 새로운 소재의 chip, optical connection을 이용한 더 나은 Package방식, 보다 큰 memory, 그리고 parallel processing을 최대한 이용하여 현재의 supercomputer 성능보다 엄청나게 강력한 Tera FLOPS급의 성능을 발휘할 것으로 기대된다. 또한 전문분야별(Graphics, signal processing, AI 등) Supercomputer들도 발전을 거듭하면서 성능은 크게 증가하고 값은 떨어져서 과학기술 분야를 포함한 각분야에 일상적으로 쓰이게 될 것이다. 21세기는 이제까지 보던 electronic computer가 아닌 neural network, optical computer, molecular computer와 같은 새로운 supercomputer까지 등장하여 바야흐로 supercomputer시대가 될 것이다.

참 고 문 헌

[1] Bell et al., "Supercomputing for one," IEEE, *Spectrum* 25(4), Apr. 1988, pp.46-50.

[2] Bell, "Optical computing: a field in flux," IEEE, *Spectrum* 23(8), Aug. 1986, pp.34-57.

[3] Batchman and parrish, "Introduction : integrated optical computing," IEEE, *computer* 20(12), Dec1987, pp.7-8. (Also, six other related articles.)

[4] Fortes and Wah, "Systolic arrays-from concept to implementation," IEEE, *computer* 20(7), Jul 1987, pp. 12-17. (Also seven other related articles.)

[5] Hecht-Nielson, "Neurocomputing : picking the human brain," IEEE, *spectrum* 25(3), Mar 1988, pp.36-41.

[6] Hill, "Supercomputing '88," *The Spang Robinson Report on Supercomputing and Parallel processing* 3(1), John Wiley and Sons, Jan 1989, pp.23.

[7] Hutcheson et al., "optical interconnects replace hardwire," IEEE, *Spectrum* 24(3), Mar 1987, pp. 30-35.

[8] Fitzgerald and Wallich, "Next-generation race bogs down," IEEE, *spectrum* 24(6), Jun 1987, pp. 28-33.

[9] Conrad, "The lure of molecular computing," IEEE, *spectrum* 23(10), Oct 1986, pp.56-60.

[10] Milutinovic, "GaAs micro processor technology," IEEE, *Computer* 19(10), Oct 1986, pp.10-13. (Also, six other related articles)

[11] Moriarty, "parallel processing of large-scale applications on powerful multiple processors,"

MIT Press, *The International Journal of Supercomputer Applications* 3(1), spring 1989, pp. 82-87

[12] Perry and Zorpette, "Supercomputer experts predict expansive growth," IEEE, *Spectrum* 26 (2), Feb 1989, pp.26-33.

[13] Shriver. "Artificial neural systems," IEEE, *Computer* 21(3), Mar 1988, pp.8-9(Also, eight other related articles)

[14] Williamson, "Neural networks : glamour and glitches." *Computerworld*, Feb 15, 1988, pp. 89-92.