

에너지 절감을 위한 전원자동관리 시스템 설계

성진우, 이영주, 장지훈
한국과학기술정보연구원
e-mail:jwsung@kisti.re.kr

Design of Power Automatic Management System for Energy Saving

Jin-Woo Sung, Young-Joo Lee, Ji-Hoon Jang
KISTI

요 약

슈퍼컴퓨터 시스템의 성능과 보유 용량이 증가함에 따라 전력소비, 전산실 냉각, 시스템 설치 공간, 유지비용 등의 여러 가지 문제가 대두되고 있다. 이러한 문제의 시작은 슈퍼컴퓨터의 성능증가에 따라 전력밀도가 증가하는 것과 관계가 있다. 그래서 제조사에서는 저전력 서버, 고효율 서버를 위하여 많은 노력을 기울이고 있다. 또한, 시스템 관리측면에서도 시스템 가상화, 통합화를 통하여 서버의 수를 줄여서 전력소비를 줄이는 방안도 나오고 있다. 본 논문에서는 24시간 365일 서비스를 하는 시스템을 위하여 시스템을 사용하지 않는 시간에는 전원을 끄고 작업이 시작되면 다시 서버에 전원을 공급하여 살리는 기능을 자동적으로 하는 전원자동관리 시스템(Power Automatic Management System, PAMS)을 설계하였다. PAMS는 서버 전력소비에서 70%를 차지하는 대기전력을 절감하는 효과를 기대할수 있다.

1. 서론

데이터센터는 IT인프라 운영에 적합한 전용건물에 전기, 공조 등 기반설비를 갖추고, 안전하고 효율적으로 24시간 365일 서비스를 제공하는 곳이다. IT장비의 성능이 계속적으로 증가함에 따라 데이터센터의 전력소비도 비례하여 증가하고 있다. 에너지 자립도가 3%인 우리나라는 국가에서 필요한 에너지의 대부분을 수입하여 전기를 생산하는데 소비하고 있다. 그리고 국제적으로는 에너지.환경문제가 전지구적으로 도전 과제로 부상하여 정부에서도 저탄소 녹색성장 전략을 통해 대응하고 있다.

슈퍼컴퓨터 센터는 넓은 의미로는 데이터센터의 범위에 포함되어 불리어 지지만 일반적인 데이터센터보다는 스토리지 사용영역보다는 계산부분이 많은 영역에 해당되어 전력소비가 더 크다. 전력소비를 절감하기 위한 노력들이 많이 이루어지고 있으나 이것들은 제조사를 위주로 진행되고 있다. 본 논문에서는 슈퍼컴퓨터가 사용하지 않는 시간동안 컴퓨팅 자원의 전원을 자동적으로 관리하는 시스템(PAMS)을 설계하였으며, PAMS의 내부 기능과 구조를 설명하였다. 실험환경은 클러스터 컴퓨팅 시스템 환경으로 대상으로 할 예정이다. 클러스터 컴퓨팅 시스템은 다수의 노드가 병렬로 연결되어 고성능의 컴퓨팅능을 이루는 시스템으로써 슈퍼컴퓨터 시스템의 대부분은 클러스터 형태로 구성되어 있다.

논문의 구성은 1장은 서론, 2장은 관련 연구, 3장은 시스템 구성 및 기능, 4장은 세부 기능 설명, 5장은 결론으로 되어 있다.

2. 기본 개념

2.1 클러스터 시스템

클러스터는 로컬 지역 내의 컴퓨터(워크스테이션, 서버, 슈퍼컴퓨터 등) 여러 대를 이더넷(Ethernet)을 이용해 동일한 시스템(OS) 환경에서 상호 연결하여 규모가 큰 하나의 가상화된 컴퓨터처럼 총체적인 서비스를 제공하는 독립적인 개체(Entity)이다. 시중에 판매되는 장비들을 이용하여 구축할 수 있는 시스템으로, 적은 비용으로 높은 성능을 낼 수 있다.

베오울프(Beowulf)라고도 불리는 DIY(do it yourself)형 클러스터 슈퍼컴퓨터는 이미 세계 500위 슈퍼컴퓨터 리스트 중 과반수 이상을 차지하는 기업을 토하고 있으며 국내외에서 다양한 분야에서 활용되고 있다. NASA와 Los Alamos 연구원들에 의하여 처음 시도된 Beowulf system은 작게는 수개의 PC에서 많게는 몇 천개의 PC를 연결한 테라급 슈퍼컴퓨터 등으로 운영되고 있다. 베오울프 시스템의 획기적인 성능향상에 기여한 공로로 이들은 1997년과 1998년 Golden Bell Prize를 받게 되었다. 이후 많은 리눅스 클러스터형 슈퍼컴퓨터가 구축되었다.

2.2 큐잉시스템과 스케줄러

한정된 자원을 사용하기 위해 사용자가 대기하는 모든 시스템을 큐잉 시스템이라고 한다. 예를 들어 식당이나 은행에서 고객이 주문하기 위하여 기다리는 과정을 예로 들 수 있다. 은행에서 은행원이 자원이라고 할 때 고객이 서비스받기 위하여 차례로 줄을 서서 서비스 받는 것도 큐

잉시스템의 예라고 할 수 있다.

일반적으로 큐잉 시스템을 구성하는 기본 요소는 다음과 같이 분류해 볼 수 있다.

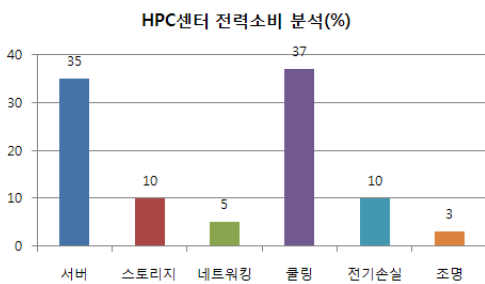
- 고객(customers):시스템에서 자원을 기다리는 객체
- 자원(resources):고객에게 서비스를 제공하는 것
- 큐(queue):고객이 서비스를 기다리는 공간 혹은 구조

슈퍼컴퓨터 시스템에서 큐잉시스템을 사용하는 이유는, 제한된 자원을 효과적으로 사용하여 대기 시간을 줄이기 위해서이다.[1] 큐잉 시스템에 사용하는 소프트웨어를 스케줄러라 부르며 대표적인 소프트웨어로는 PBS, LSF, NQS, Loadlevler, Condor 등이 있다.

2.3 데이터센터 현안

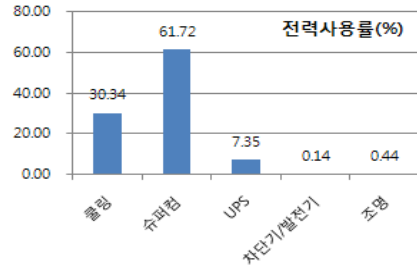
데이터센터가 직면하고 있는 현안 문제는 컴퓨터 시스템의 성능이 증가함에 따라 증가하는 전력 소비와 인프라 스트럭처(전력, 쿨링, 공간)의 증가이다. 컴퓨터 시스템의 성능 증가는 H/W의 소비전력 증가가 필연적이며, 발열량도 증가하여 전산실 온도와의 밀접한 관계가 있다. 전산실에 발생한 열을 제거하기 위해서는 항온항습 설비들의 용량증가가 뒤따르며 또한 설비들의 전력소비도 늘어나게 되어 있다. 그러므로, 컴퓨터의 성능증가는 컴퓨터의 소비전력 증가뿐만이 아니라 설비증설이 필요하며, 설비의 전력소비 증가도 더욱 늘어나게 된다. 전력소비증가는 데이터센터 운영비용과도 직결되므로 에너지 절약 등 효율적인 데이터센터가 유지되도록 노력하여야 한다.

2.4 데이터센터 전력소비 분석



(그림 1) 해외 데이터센터의 전력 소비 분석

그림 1은 해외 데이터센터의 전력 소비 분석자료로서 서버와 쿨링 부분이 각각 35%와 37%를 차지한다는 것을 알 수 있다. 쿨링부분은 서버에서 발생하는 열을 제거하여 냉각시키는 것을 말하며, 항온항습기, 공기조화기, 가습기, 냉동기, 냉각탑, 모터, 팬 등이 여기에 해당된다. 전기손실은 교류-직류 변환 손실이 있는 UPS와 고압차단기, 변압기에서의 손실을 말한다. 쿨링에 소비되는 전력이 서버와 거의 맞먹는 것을 알 수 있다.



(그림 2) KISTI 슈퍼컴센터의 전력소비 분석

그림 2는 KISTI 슈퍼컴센터의 전력소비 특성을 나타내며, 해외의 경우와 유사하게 서버(슈퍼컴)과 쿨링부분이 대부분을 차지하고 있다.

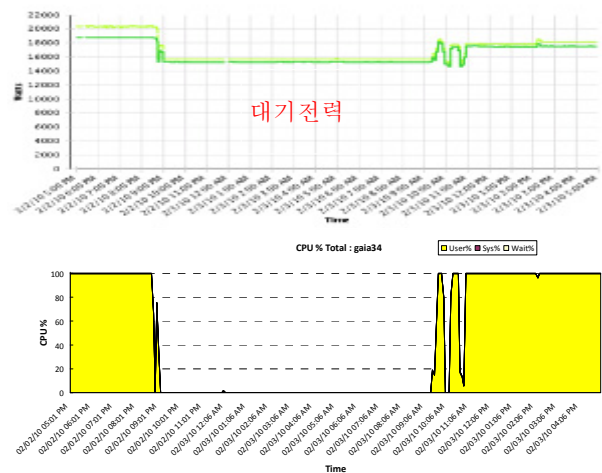


(그림 3) 슈퍼컴퓨터의 전력소비 분석

그림 3은 슈퍼컴부분의 전력소비를 분석한 자료로 서버부분이 대부분을 차지하는 것을 알 수 있다. 이러한 통계분석 자료에 의하여 에너지 효율화를 하기 위한 노력은 서버와 쿨링 부분에서 많이 이루어져야 한다는 것을 알 수 있다.

2.5 에너지 절감 부분

슈퍼컴퓨터 시스템은 일반적으로 랙당 10kW이상의 전력을 소비하고 있다. 그림 4는 시간대별 슈퍼컴퓨터 시스템의 사용량과 전력소비 통계자료의 실제 자료의 예이다.



(그림 4) 시스템 사용율과 전력소비량 관계

시스템의 사용율은 약 45%정도이다. 위 그래프는 시간대별 전력 소비량을 나타내는 표이고, 아래는 시스템 사용률

을 나타내는 표이다. 2개의 표를 비교하면서 살펴보면 2가지 사실을 알 수 있다. 첫 번째는 시스템에 작업이 없을 때에도 전력소비가 15kW가 일정하게 소비되고 있으며, 두 번째는 작업이 있어서 시스템 사용율이 100%이더라도 전력소비 증가량은 많지 않으며, 증가율은 20~30%정도임을 알 수 있다. 슈퍼컴의 이러한 전력소비 특성을 고려하면 에너지를 절감하기 위해서는 항상 70%~80%를 차지하는 대기전력(Waiting Power)을 절감하는 게 에너지 절감에 효과적이다.

대기전력을 절감하기 위해서 전력을 소비하는 부분들(전원부, 냉각팬, 스토리지, 드라이버 등)을 고효율 장치로 개선하는 작업이 필요하나 이 일은 제조사들의 몫이다. 시스템 관리자 입장에서 대기전력을 절감하기 위하여 취할 수 있는 가장 단순한 방법은 사용하지 않을 때에는 전원을 끄는 것이다. 그리고 필요할 때에는 다시 전원을 투입하고 시스템을 살리는(부팅) 작업을 하는 것이다.

시스템을 끄고 켜는 일을 수작업으로 한다면 상당히 고통스러운 일이며, 더욱이 온라인상에서 다수의 사용자가 스케줄러를 통하여 작업을 제출하는 체계라면 전원관리의 자동화는 필수사항이 될 것이다.

2.6 관련 연구 조사

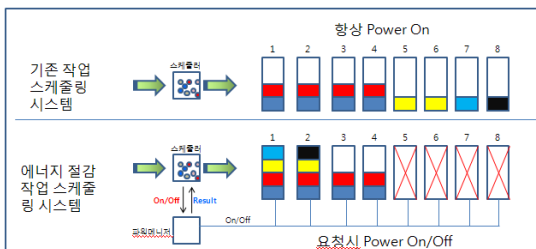
슈퍼컴퓨터의 전원관리는 통상적으로 인간(시스템 관리자)이 하여 왔으며, 컴퓨터의 전원관리를 인간이 아닌 다른 무엇(프로그램 등)에게 맡긴다는 것은 찾기 힘들다.

3. 시스템 설계

전원자동관리 시스템의 구성은 자원관리자, 작업관리자, 전원관리자로 구성되며, 각각에 맡겨진 기능은 서로 독립적이며 차별되어 있다.

3.1 전원관리 시스템 구조 및 기능

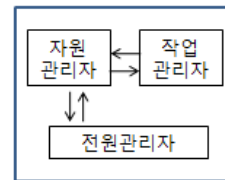
그림 3은 기존의 작업스케줄링 방법과 개발하고자하는 에너지절감 작업스케줄링 시스템의 비교를 나타내는 그림이다. 기존의 클러스터 시스템에서 스케줄러는 사용자 작업을 계산 노드에 분배하는 방식은 부하균등법이거나 순차배분법이다. 그러나 전원자동관리 시스템에서는 사용자 작업을 계산노드의 사용률이 일정수준(상한 한계 설정: 약 70%~80%)이 될 때까지 집중적으로 작업을 배분하는 방식



(그림 3) 작업 관리 방법 비교

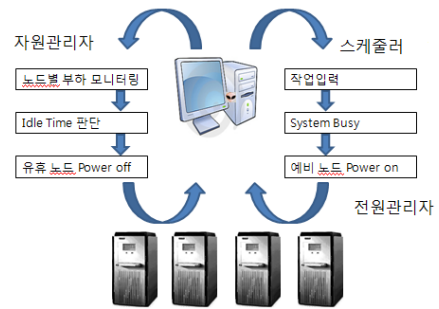
을 채택한다. 이렇게 되면 작업이 없는 계산노드들은 전원 관리자에 의하여 전원이 off되고 다시 작업이 늘어나면 꺼져있던 노드를 다시 살리는 것이다.

그림 4는 전원자동관리 시스템의 구조를 나타낸다. 시스템 내부적으로는 자원관리자, 작업관리자, 전원관리자가 존재하며, 자원관리자는 작업관리자로부터 작업의 정보를 가져오고 노드들의 작업상태를 모니터링하여 계산노드들의 유휴정도를 판단하여 필요시 전원을 off하는 명령을 내린다. 그리고 수행해야 될 사용자 작업이 대기중일 때에는 power off되어 있던 계산 노드들을 다시 살려서 작업이 수행될 수 있도록 조치를 한다. 작업관리자는 스케줄러로서 사용자 작업을 계산노드에 배분하여 작업이 수행되도록 하는 역할을 한다. 전원관리자는 자원관리자의 명령에 따라 계산노드의 전원을 관리한다.



(그림 4) 전원 자동 관리 시스템의 구성

그림 5는 전원관리 시스템의 전체적인 기능을 설명해주는 그림이다. 사용자가 작업을 제출하면 작업관리자는 작업을 계산노드에 배정을 하고 계산 노드가 부족하면 전원관리자가 계산노드를 추가해주어 작업이 수행된다. 그리고 자원관리자는 작업을 수행하지 않고 있는 노드가 있으면 전원관리자에게 명령을 내려 전원을 off한다.



(그림 5) 작업 흐름

3.3 전원관리 시스템 요구기능

클러스터 시스템을 위한 전원관리 시스템의 주요기능 단계는 아래와 같다.

- 자원 관리
 - 사용자 작업의 상태에 따라 다수의 계산노드들 계속 사용하여야 하는지 아니면 전원을 꺼야하는지를 결정할수 있어야 한다.
 - 전원 관리자에게 전원 on-off 작업을 내리고 그 결

과에 따라 후속 작업을 할 줄 알아야 한다.

- 전원 관리
 - 1개의 관리노드와 다수의 계산노드가 네트워크로 구성된 클러스터 시스템에서 관리노드는 계산노드의 전원을 off하고 on할 수 있어야 한다.
- 작업 관리
 - 사용자의 작업은 스케줄러를 사용하며, 사용자의 작업 정보를 자원관리자에게 제공해줄 수 있어야 한다.
- 클라이언트 노드 모니터링
 - 사용자의 작업이 수행되는 각 클라이언트 시스템의 상태도 주기적으로 모니터링 한다.
- 파일 관리
 - 통계를 위하여 관련 작업 정보를 기록한다.

3.4 시험 환경

시험 환경은 1대의 서버와 8대의 계산 노드로 구성되어 있다. 시험에 사용될 서버와 노드의 사양은 Intel Xeon Processor Quad-core X3430이며, 2.4GHz이다. O.S는 Linux이며, 스케줄러는 OpenPBS이다. 자세한 정보는 표 1에 나타나 있다.

<표 1> 시스템 개발 환경

서버명	규격	비고
O.S	Linux	
구조	N+1 cluster	
Processor	Xeon Quad Core X3430 2.4GHz	
메모리	8MB	
Scheduler	openPBS	
스토리지	4TB	
네트워크	Netgear 10/100/1000 Gigabit 스위치	
기타	랙(전원 모니터링 PDU 포함)	

4. 결론

컴퓨터의 성능이 증가함에 따라 많은 전력을 소비하게 되었으며, 발열량도 증가하게 되었다. 이러한 변화는 기반 시설의 증가로 이어지게 되어 데이터센터 관리자 또는 경영자에게는 시스템 구입비용뿐만 아니라 전체 운영경비도 증가하게 되었다. 이러한 변화에서 에너지 절감은 선택이 아니라 필수사항이 된 것이다.

KISTI 슈퍼컴퓨터 시스템은 국내외 연구자에게 컴퓨팅파워를 제공하는 서비스를 수행하고 있으며, 24시간 365일 슈퍼컴퓨터에 작업이 수행되고 있다. 슈퍼컴퓨터의 전력소비분석 자료에 의하면 대기전력이 70%~80%이고 사용전력이 20%~30%로 나타났다. 이러한 분석을 바탕으로 대기전력을 절감하기 위한 노력의 일환으로 작업유무에

따라 전원자동관리 시스템의 개발을 시작하게 되었다.

전원자동관리 시스템의 기본적인 구성은 자원관리자, 작업관리자, 전원관리자이다. 전원자동관리 시스템의 주요 특징은 작업관리 기능과 전원관리 기능을 서로 연계시켰다는 점이 그 큰 특징이다.

전원자동관리 시스템을 통하여 유휴자원의 대기전력을 절감하게 되어 에너지 효율적인 슈퍼컴퓨터 데이터센터가 되리라 생각한다.

향후의 연구는 시스템 개발 및 시험을 거쳐 절감되는 전력량을 측정하는 일과 부팅으로 인한 사용자의 작업수행시간의 변화를 측정하는 작업이다.

참고문헌

[1] 리눅스 클러스터로 만드는 슈퍼컴퓨터, 이정훈, p299
 [2] 성진우, 이상동, 김성준, 이영주, 김중권, “통합컴퓨팅 환경을 위한 스케줄러 설계 및 Prototype 구현”, 2005. 정보처리학회