

고성능 컴퓨팅 클라우드의 산업 동향 및 이슈

Industry Trends and Issues of High Performance Computing on Clouds

최정란 (J.R. Choi) 고성능컴퓨팅시스템연구팀 선임연구원
김형환 (H.H. Kim) 고성능컴퓨팅시스템연구팀 책임연구원
우영춘 (Y.C. Woo) 고성능컴퓨팅시스템연구팀 팀장
최 완 (W. Choi) 클라우드컴퓨팅연구부 부장

* 본 연구는 지식경제부 및 한국산업기술평가관리원의 IT산업원천기술개발사업의 일환으로 수행
하였음(10038768, 유전체 분석용 슈퍼컴퓨팅 시스템 개발).

포레스터 리서치(Forrester Research)에 따르면 클라우드 시장은 2010년 400억 달러를 기점으로 2020년 2,400억 달러로 성장할 것이라고 예측하고 있다. 급격하게 성장하고 있는 클라우드의 pay-per-use는 그동안 시스템 활용의 한계를 느끼던 HPC(High Performance Computing) 사용자에게 상당한 매력으로 작용하고 있다. 그러나 현재 대부분의 클라우드는 HPC와 접목하기에는 한계들이 존재한다. 본고는 HPC가 클라우드로 이동하기 위해 해결해야 하는 도전과제와 장점을 알아보고, 현재 HPC 클라우드의 산업 현황에 대해 살펴본다. HPC와 클라우드 접목에 따른 HPC 위주의 성능 측정의 문제와 새로운 측정 지수 도입의 필요성 및 HPC가 클라우드로 서비스되기 위해 클라우드가 지원해야 할 요소들에 대해 기술하였다. 제기된 이슈들에 대한 접근이 해결된다면 HPC의 클라우드로의 전환도 머지않을 것이다.

2013
Electronics and
Telecommunications
Trends

빅데이터 처리 및
분석 기술 특집

- I. 서론
- II. HPC 클라우드 이슈
- III. HPC 클라우드 산업 동향
- IV. HPC 클라우드에서 성능
관련 이슈
- V. 클라우드에서의 HPC
지원
- VI. 결론

I. 서론

HPC(High Performance Computing) 사용자는 복잡한 병렬 시스템을 직접 관리하고 애플리케이션을 대상 하드웨어에 맞춰 개발·사용하는 데 익숙하다. 문제는 실행 환경에 대한 거의 모든 제어를 ‘자발적으로’ 넘겨주는 클라우드에 애플리케이션을 배치하는 것이 HPC 사용자에게 유리한가에 대한 고민이다.

HPC에서 클라우드를 이용하는 대부분 솔루션은 pay-per-use 인프라를 얻는 것이다. 그러나 일반 클라우드 사용자와는 달리 HPC 사용자는 병렬 하드웨어를 사용하고, 이를 이용해 명시적으로 병렬 애플리케이션을 실행하고자 한다. 그들은 고성능 컴퓨팅 클러스터와 유사한 빠른 연결이 가능한 수많은 강력한 장비를 사용하기를 원한다. 또 하나는 클라우드로부터 소프트웨어 요구사항에 따라 직접 형상화할 수 있는 클러스터를 요청(cloud on demand)하고자 한다. 이는 실제로 클라우드 사용자에게 관리가 완전히 위임되는 가상화 클러스터를 통해 가능하다.

클라우드 컴퓨팅은 대규모 HPC 사이트에서 최소한의 통신 의존성으로 워크로드 처리와 내부 HPC 자원 추가 없이 오버로드를 처리하기 위한 측면에서 제안되고 있다. 또한 HPC 전문가 고용이나, HPC 자원의 구매 없이 HPC 기능을 활용하고자 하는 SMB(Small & Medium-sized Business)를 위해 클라우드 컴퓨팅이 제안되고 있다.

대부분의 클라우드 서비스는 소규모의 기업과 소규모 작업을 위한 모델들이 판매되고, HPC와 같은 복잡하고 방대한 작업을 위한 모델은 현재는 제한적으로 제공되고 있다.

프라이빗 클라우드에서는 HPC 클라우드를 위한 실질적인 요구사항이 늘어나기 시작했다. 2010년부터 CERN은 지금은 IBM에 인수된 Platform Computing과 함께 세계에서 가장 큰 입자가속기를 위한 HPC의 프라이빗 클라우드를 개발하고 있다. 유럽과 전 세계에 걸쳐있는

과학자들을 위한 컴퓨팅 자원, 데이터와 애플리케이션을 배포하기 위해 세계에서 가장 큰 과학 컴퓨팅 클라우드를 개발하는 것이다[1]. 2010년 2월 NASA는 원격으로 기후 모델을 실행하기 위해 웹 포털로 클라우드 환경을 구축한다고 발표하였다[2]. NSF(National Science Foundation)는 광범위한 연구자 집단을 위한 HPC 사이클을 제공하기 위해 MS의 클라우드 제품을 사용하고 있다[3].

퍼블릭 클라우드 영역에서는 보잉사와 관련 회사들은 여러 해 동안 인도 Pune의 Tasa CRL에 거대한 슈퍼컴퓨터를 두고 원격으로 접근하여 활용하고 있다[4]. 다수의 석유가스 회사들은 저지연 네트워크를 요구하지 않는 사이즈믹 프로세싱(seismic processing)과 같은 워크로드의 일부 처리를 위해 퍼블릭 클라우드를 이용하고 있다[5]. 그 외 생명과학, 금융 서비스, 빅데이터 처리 등에서 이미 퍼블릭 클라우드가 사용되고 있다. 최근 국내의 KT는 연구소의 연구개발이나 바이오 분석 작업 시 고성능의 분석 작업이 가능하도록 유클라우드 비즈 HPC 서비스를 출시한다고 밝혔다[6].

특히, 현재 전 세계 HPC 워크로드의 80%는 온프레미스 환경에서 돌아가고 있다. 즉, 이 80%는 기존 HPC 인프라에 클라우드 콘셉트를 도입하는 형태가 될 것이고, 그 외에 나머지 20%는 퍼블릭 클라우드 서비스를 활용하게 될 것이라고 전망하고 있다[7].

본고는 II장에서 HPC가 클라우드로 이동할 때 제기되는 이슈를 다루고, III장에서는 현재 HPC 클라우드 산업 동향, IV장에서는 II장에서 언급된 이슈 중 성능 관점에서 기술하고, V장은 HPC를 클라우드 서비스하기 위해 지원해야 할 두 가지 요소에 대하여 기술하고, VI장 결론으로 마무리 지었다.

II. HPC 클라우드 이슈

본 장은 HPC가 클라우드로 이동할 경우 발생할 긍정

적인 요인들과 도전과제에 대해 기술한다[8].

긍정적인 요인들로는 비용, 유연성, 빠른 배치, 확장성, 탄력성, 이식성 등을 들 수 있다.

- Cost: 비용은 용량 비용(CAPEX)이라기보다 운용 비용(OPEX)이라는 측면에서 클라우드를 이용하여 컴퓨팅 구축 비용 절감은 비즈니스 구미에 맞음.
- Flexibility: 클라우드는 특정 워크로드를 위해 빠른 프로비저닝과 특정 OS, SW 및 HW 형상들의 동적 롤아웃 기능을 지원함.
- Speed of deployment: 온프레미스 환경에서 새로운 HW를 프로비저닝하는 데 수개월 걸리는 것과 달리 온디맨드 컴퓨팅은 수분 안에 새로운 환경/클러스터를 빠르게 프로비저닝함.
- Scalability: 요구 용량에 맞추어 탄력적으로 규모를 늘리는 것(scale-out)으로 Windows Azure나 AWS(Amazon Web Service)와 같은 대형 프로바이더들이 제공하는 성능 레벨은 자체 하드웨어로는 헤아릴 수 없는 정도임.
- Resiliency: 실행 중인 워크로드의 스냅샷을 찍어 MPI 워크로드의 체크포인트 저장 기능 등을 활용하여 다수의 가상화 솔루션들은 동적으로 물리 호스트에서 다른 호스트로 게스트 VM(Virtual Machine)을 탄력적으로 이동하는 기능을 제공함.
- Portability: 특정 클라우드 플랫폼에서 다른 클라우드 플랫폼으로 애플리케이션 수정 없이 바로 워크로드를 옮기는 기능은 동시에 다수 클라우드에서 실행 가능한 고가용성을 지원함.

해결해야 할 도전과제들은 보안, 성능, 라이선스 관리, 통합관리, 데이터 처리, 비용 관리, 클라우드 전환 등이다.

- Security: 클라우드 플랫폼 내에서 데이터는 정교함이나 위협성에 비례하여 데이터 암호화, 데이터 상주시간 제한, 데이터 익명화 등 보안에 대한 노력이 더욱 필요함.
- Performance: 클라우드에서 성능에 대한 질문에

는 명확한 답은 없지만, HPC에서 병렬/IO 집중형 워크로드, 네트워크 성능 저하에 따른 지연, 가상화 적용 등이 저하의 원인으로 대두됨.

- License management: 클라우드에서 다수의 SW를 제한된 시간 동안 활용하는 모델로 트랜잭션에 따른 책정 등 별도의 정책 정의가 필요함.
- Management: 한 개 이상의 플랫폼에 걸쳐있는 워크로드 및 자원의 관리는 HPC를 위한 클라우드의 효율적 활용을 위해 하나의 도구로 통합 관리하는 것이 중요함.
- Data: 애플리케이션의 특성에 따라 전체를 클라우드 서버로 옮길 것인지, 필요에 따라 적용할 것인지 선택 후 적용함. 현재는 네트워크 문제로 FedEx 등을 통해 직접 배송하는 사례도 있음.
- Managing costs: 미래 엔터프라이즈 고객들을 위한 더욱 복잡한 청구 관계 또는 부분 비용 최소 등을 지원하기 위한 상업적 청구 모델이 필요함.
- Cloud lock-in: 클라우드를 적용하고자 하는 기업들은 벤더들이 기존 시스템 대체 비용이 어마어마하여 기술 전환을 하지 못하는 상태임.

위에서 언급한 것처럼 대부분의 시나리오에 대해 장점이 도전과제보다 훨씬 크다. 클라우드는 이미 HPC 워크로드에 이상적인 플랫폼으로 인식되고 있다. 언급한 대부분의 도전과제는 IaaS 플랫폼 제공자나 SW에 벤더에 의해 지속적으로 줄여나가고 있다. 그러나 HPC가 복잡하다는 인식이 클라우드로 전환했을 때 이점들과 비즈니스 가치에 다가가기 어렵게 하고 있다. IV장에서는 제기된 도전과제들 중 성능에 관한 이슈를 자세히 기술하였다.

III. HPC 클라우드 산업 동향

1. Amazon의 EC2 HPC 서비스

Amazon은 Cluster Compute와 Cluster GPU 두 종류

의 HPC 서비스를 제공하고 있다. Xen 하이퍼바이저 3.4를 활용하여 HVM 기반의 10G 네트워크를 제공하는 Linux 3.x 커널의 EC2 인스턴스이다. I/O 중심의 워크로드를 처리하기 위해서 양방향 고대역 네트워크에서 작동하는 Cluster Compute 또는 Cluster GPU 인스턴스를 선택하거나, 처리량 지향 애플리케이션을 위해 수천 개의 코어로 확장할 수도 있다. 현재 바이오 제약, 석유가스, 금융 서비스, 제조 등 산업 및 학술연구단체에 걸쳐 AWS 인스턴스에서 CAE(Computer Aided Engineering), 분자 모델링, 계층 분석, 수치 모델링 같은 분야에서 HPC 애플리케이션을 운영하고 있다.

Amazon EC2상의 HPC 서비스[9]는 클러스터 인스턴스 제품군인 Cluster Compute와 Cluster GPU 인스턴스는 다른 Amazon EC2 인스턴스와 동일한 방식으로 사용할 수 있지만, 다음과 같이 HPC 애플리케이션에 최적화된 기능을 갖추고 있다.

클러스터 인스턴스는 Placement Group 내에서 시작하는데, 모든 인스턴스는 지연시간이 짧고, 인스턴스 간 full bisection 10Gbps 대역폭이 확보된다. 자원의 동적인 유연성을 제공하여 필요에 따라 조절 가능하고 여러 Placement Group을 연결해 대규모 병렬처리 작업을 위한 대용량 클러스터를 구축할 수도 있다.

Cluster Compute와 Cluster GPU는 기본 프로세서 아키텍처가 포함되어 있어 개발자는 특정 프로세서 아키텍처를 컴파일하여 애플리케이션을 세밀하게 조정해 최적의 성능을 얻을 수 있다.

Cluster GPU 인스턴스를 통해 GPGPU 컴퓨팅용 CUDA와 OpenCL 프로그래밍 모델을 사용하여 Nvidia Tesla GPU 병렬 성능을 활용할 수 있다.

Amazon EC2 클러스터 인스턴스의 한 예로, cc2.8xlarge 인스턴스로 구성된 1,064개 인스턴스 클러스터(17,024개 코어)는 High Performance LINPACK 벤치마크의 240.09TFLOPS(Tera Floating Operations per Second)를 달성하고, 2011년 11월 top500 리스트에서 42위를

〈표 1〉 Amazon HPC 클라우드 구현 기술

기술	설명
HVM	최신의 Linux Kernel(3.x) 및 DomU에 물리 장비(GPU)를 제공하기 위해서 full 가상화(HVM) 방식 사용
PV on HVM	HVM에서 network/block device 에뮬레이션으로 인한 성능 저하를 막기 위해 network/block device는 Xen PVHVM driver 적용
GPU Passthrough	Cluster GPU 인스턴스는 PCI에 붙은 GPU 카드를 GPU Passthrough 기술을 통해서 Attach함.
CUDA SDK	Nvidia의 GPU computing 프로그래밍을 지원하는 소프트웨어 툴

〈표 2〉 Amazon 서비스 이용 사례

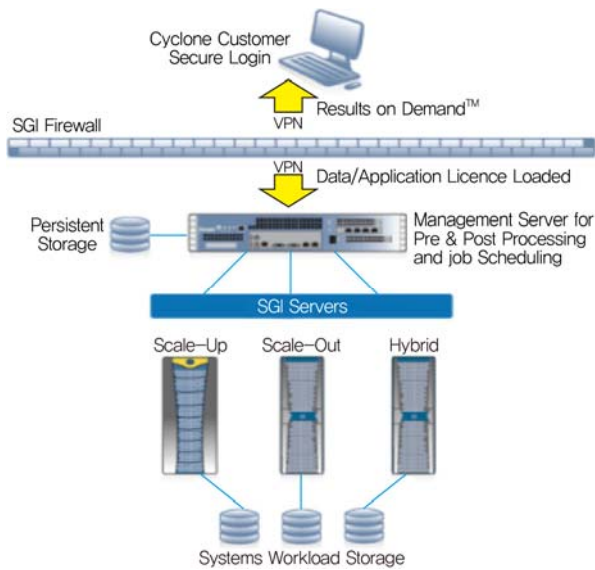
고객	설명
Bankinter	신용 위험 확인 시뮬레이션, 평균 23시간 → 20분
PSR	천연가스와 전기 시설에 기술 솔루션과 컨설팅
AeroDynamics Solutions	미공군 제트엔진 시뮬레이션
Cyclopic energy	풍력 발전 시뮬레이션
국립타이완대학교	암호화 알고리즘 병렬 구조 실험 등

차지했다.

〈표 1〉은 Amazon의 클라우드에서 HPC 서비스를 위해 사용한 구현 기술이고, 〈표 2〉는 서비스 이용 사례를 나타내고 있다[10].

2. SGI의 Cyclone

SGI의 Cyclone[11],[12]은 “Result on Demand”로, 바이오, 화학물질, 유체역학, 구조역학, 전자자기학, 온톨로지 분야 등의 과학기술 컴퓨팅에 특화된 HPC 클라우드이다. 주로 IT보다 과학과 연구 분야에 초점을 맞춰 지원하고 있다. Cyclone은 Itanium에 기반한 Altix 시스템의 SMP(Symmetric Multiprocessing)를 지원하는 HPC 시스템이 접근할 수 있는 클라우드 컴퓨팅 사이클과 서비스를 제공한다. (그림 1)은 Cyclone의 활용 모델이다.



(그림 1) SGI의 Cyclone 활용 모델

Cyclone은 플랫폼, 운영체제, 네트워크 등에서 다음과 같은 유연한 선택을 제공한다.

- 플랫폼: Scale-up(Altix 4700, Altix UV)과 Scale-out(Altix ICE, Altix XE)을 지원하고, Nvidia Tesla, ATI FireStream과 Tiler 가속기를 가진 하이브리드 시스템을 제공
- 운영체제: SLES, RHEL
- 네트워크: NUMalink, InfiniBand, GigE
- 물리적 접근 및 가상화 모두 지원

SGI의 클라우드는 클러스터가 설치되고 실행될 때까지 임시 가교로써 SGI 클라우드 자원을 활용할 수 있다.

SGI Cyclone은 네트워크상에서 사용자에게 HPC 애플리케이션과 SGI 소스를 제공하는 SaaS 모델과 프로세싱, 스토리지, 네트워크 용량 등과 자신의 애플리케이션을 실행하도록 하는 IaaS 모델을 제공한다. IaaS 모델은 이미 자체 개발되고 라이선스가 있는 애플리케이션의 클라우드 전환과 관련이 있다.

비용 정책은 유틸리티 모델(utility model)을 사용하여 필요한 클러스터를 획득하고, 관리하는 비용은 지불하지 않고, 고객이 자원을 사용한 만큼 지불하는 방식(per-use)을 채택하고 있다. 필요한 클러스터를 주 단위

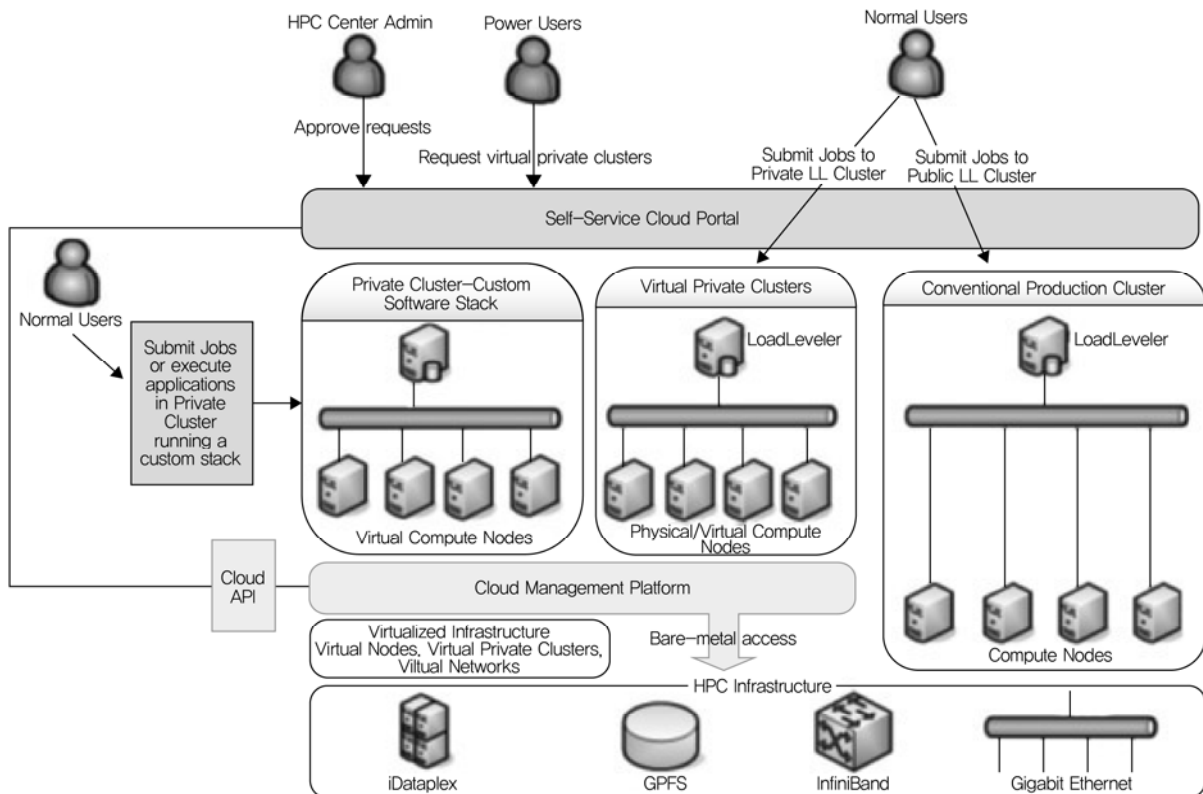
로 예약하고, 고객의 데이터는 초고속 인터넷 네트워크 백본을 통하여 업로드하거나 FedEx를 통해 스토리지를 배달한다.

3. IBM의 HPC 클라우드 솔루션

IBM은 1992년부터 Platform Computing사의 솔루션을 통해 클러스터, 그리드, 클라우드 관리 소프트웨어 분야를 지속적으로 개발 및 서비스하고 있다. 최근에는 Platform Computing사를 인수하여 HPC 클라우드 서비스를 위한 자원관리 및 워크로드 지원(플랫폼 LSF, 플랫폼 심포니 등)을 위한 기술 기반을 강화하고 있다. IBM의 HPC 클라우드[13]는 속도와 유연성에 대한 요구사항을 맞추기 위해 최적화되었다. 핵심 기술은 경량(light-weight), 고성능의 작업 스케줄러와 최소 4분 이내 VM/물리적 이미지를 모두 지원하는 자원 관리자, 그리고 표준 작업큐 인터페이스에 따라 사용자 편의의 인터넷 기반 HPC 포털을 제공하는 것이다.

IBM HPC Management Suite for Cloud 제품군은 동적으로 bare-metal HPC 클러스터의 프로비저닝과, 코어, 저장공간, 물리/가상 메모리로 이루어진 VM들을 통해 클라우드에서 HPC 애플리케이션을 서비스할 수 있도록 돕는다. 제공되는 기본 컴포넌트는 통합 분산환경 관리 및 프로비저닝 지원 도구, 워크로드 스케줄러, 클라우드 포털, 병렬 파일 시스템, 확장성 및 고성능을 제공하는 스케일아웃 NAS(Network Attached Storage) 솔루션, 연구자를 위한 HPC 오픈소스 SW 스택 등이다.

(그림 2)는 자동차, 우주항공, 디지털 제조업 등을 포함하는 기계 관련(mechanical) 클라우드 아키텍처 사례이다. 설계, 시뮬레이션과 분석을 위한 인터랙티브 모드와 배치 모드의 2D/3D 워크로드를 지원하도록 최적화된 HPC 클라우드 인프라와 엔지니어링 포털을 제공한다. 주요 기능은 고성능컴퓨팅 인프라와 유연한 시스템과 워크로드 관리, 파트너들과 원격의 3D 그래픽 처리 가속화 지원, ISV 파트너 통합 등을 제공한다.



(그림 2) IBM의 Mechanical 클라우드 아키텍처[13]

IBM은 프라이빗 클라우드 형태로 HPC 인프라를 구축할 수 있도록 지원하며 현재 ‘스마트 클라우드 엔터프라이즈(SCE)’와 ‘스마트 클라우드 엔터프라이즈 플러스(SCE+)’라는 퍼블릭 클라우드 서비스를 통해 HPC 클라우드 환경을 제공한다.

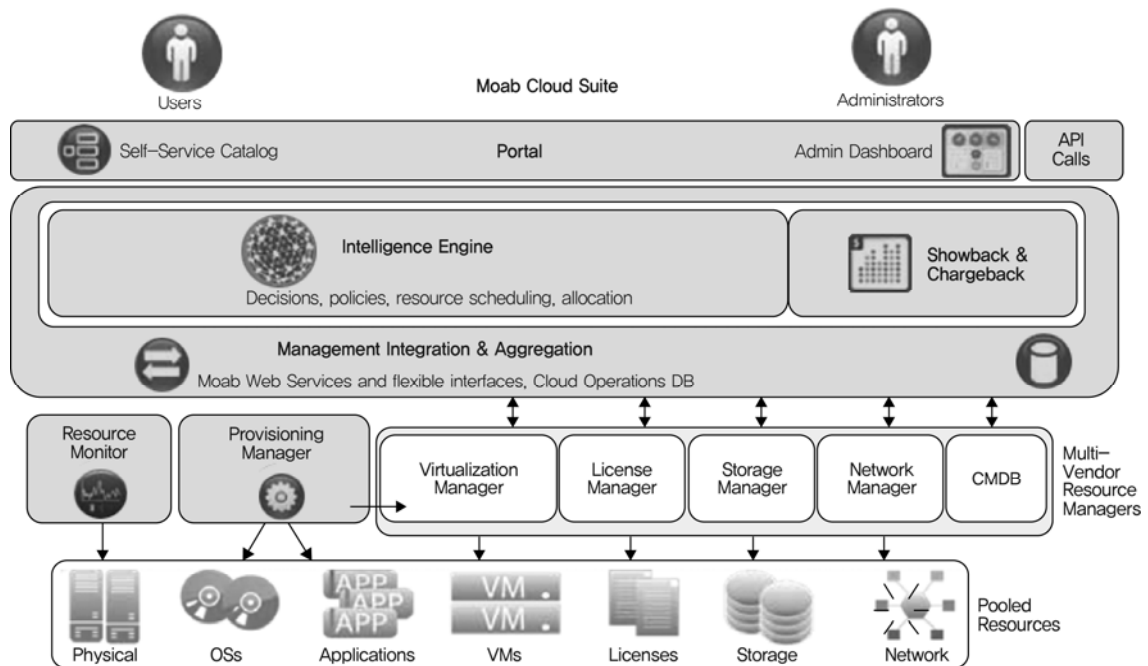
4. Adaptive Computing의 Moab HPC Suite

Adaptive Computing의 Moab[14]은 스스로 최적화 기능을 제공하는 동적 워크로드 관리 소프트웨어와 HPC 클라우드 컴퓨팅 환경을 관리하는 솔루션이다. Moab 인텔리전스 엔진은 엔터프라이즈급 HPC 시스템의 복잡하고, 임무-위급한 워크로드를 우선순위에 따라 자동으로 처리한다. 서비스가 요청되면, Moab의 인텔리전스 엔진은 사용자의 요구사항을 맞추기 위해 사용 가능한 자원의 할당 및 재프로비전(re-provision)을 수행

한다. 이 자동화 계층은 시스템 관리자의 오버로딩 없이 복잡한 환경을 셋업하기 위해 중요한 역할을 한다.

(그림 3)은 Moab의 구성도이며, 다음과 같은 기능을 제공한다.

- Simplified user experience: 작업 할당 및 관리를 위한 Moab Viewpoint라 불리는 웹 기반 포털
- Usage accounting/Charge management: 요금 부과 체계로 작업을 서밋하기 전에 얼마나 많은 비용을 소모할지 측정(개별 컴포넌트 프로세서, GPU, 라이선스 등, 또는 총 사용 비용)
- Dynamic scaling for HPC resource: 작업 환경 내에서 요청에 따라 사용 가능한 자원으로 대체하고 바로 활용할 자원이 없을 경우 노드 재인스톨
- HPC cloud bursting: 요청한 자원이 로컬 클라우드에 없을 경우, 외부 클러스터 또는 퍼블릭 클라우드에서 새로운 자원이 프로비전됨.



(그림 3) Moab Cloud Suite 구성도[14]

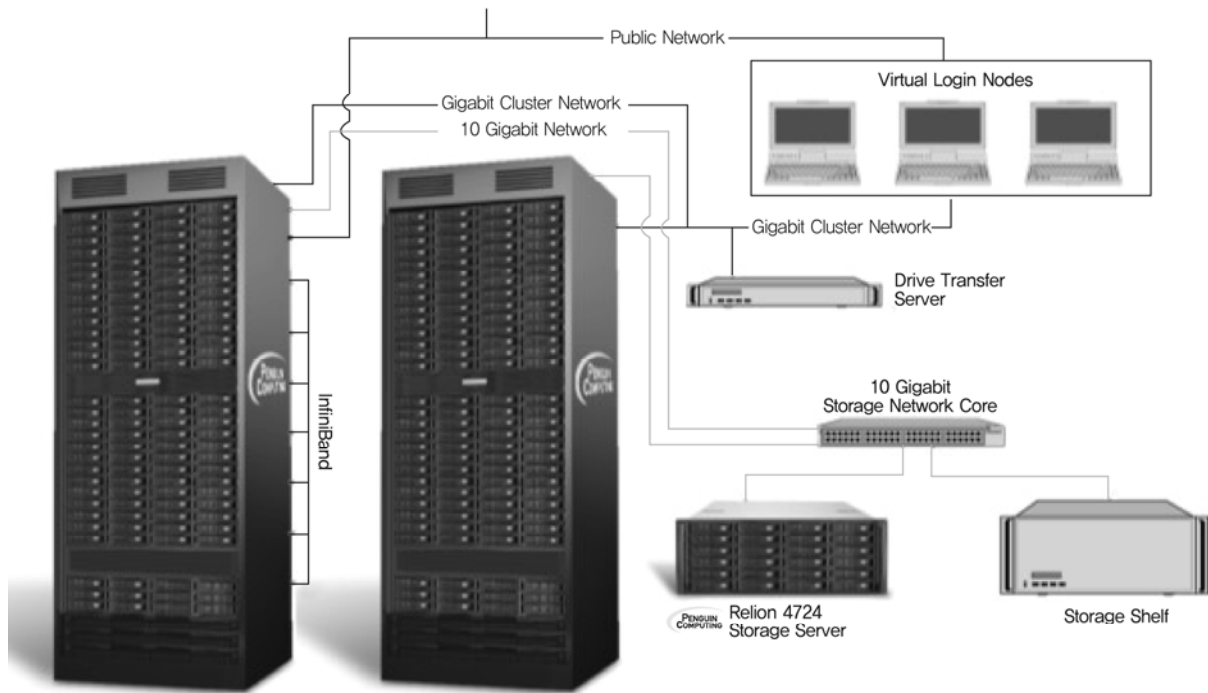
- Service catalog: 포털을 통해 여러 사용자에게 그에 맞는 여러 서비스를 제공하기 위해 테일러링할 수 있는 기능 제공
- Diverse environments: 실행 중인 작업에 물리/가상의 자원을 지원하기 위해 VMware의 ESX 플랫폼과 같은 하이퍼바이저가 들어오는 작업에 맞춰 재-프로비전되어, 클러스터 내 모든 자원을 허가하고 있음.
- Migrating workloads: 물리/가상 워크로드의 균형을 유지하기 위해 자동으로 워크로드를 효과적인 노드로 이동시켜 성능을 유지시킴.

5. Penguin Computing의 POD

Penguin Computing사는 HPC 솔루션 및 엔프라이즈 솔루션 업체로 필요할 때 바로 사용 가능한 퍼블릭 HPC 클라우드 POD(Penguin Computing on Demand)[15]를 서비스한다. 2012년 수요의 증가로 30% 이상의 컴퓨팅 용량을 확장하였고, AOL, Yelp, Catepilla, Life Technologies, Dolby, Lockheed Martin 및 US Air Force 등

의 업체들과 협력하고 있다.

POD는 고밀도의 계산 노드와 직접 추가할 수 있는 초고속 스토리지를 가진 고성능, 확장성의 온디맨드 HPC 환경이다. POD는 간소화된 로그인 절차를 제공한다. 사용자가 로그인 후 가상 게이트웨이 노드의 OS 이미지를 선택하고, 프로비전되어 바로 HPC 인프라를 접근할 수 있다. (그림 4)와 같이 POD 사용자는 지속적인 보안 컴퓨팅 환경을 제공하는 자신의 로그인 노드를 가진다. 로그인 노드로부터 시작된 작업은 가상화되지 않은 계산 노드의 물리적 코어에서 바로 실행할 수 있고, direct attached storage의 사용자 요구사항을 맞추기 위해 여러 계층의 스토리지 옵션을 제공하여 스토리지가 병목되는 것을 막는다. I/O 집약 애플리케이션을 처리하기 위한 초고속 병렬 파일 시스템을 제공하기 위해 Ceph 분산 오픈소스 스토리지 시스템을 기반으로 분산 파일 기반/블록 기반/객체 기반의 스토리지를 지원한다. 클라우드 환경의 남용과 오용을 막기 위한 VPN 장비 제공을 통하여 보안 연결(virtual private cluster)과 스토



(그림 4) POD 구성[15]

리지의 암호와 옵션을 제공한다.

POD 컴퓨팅 인프라는 최적의 성능을 보장하기 위해 물리적 컴퓨팅 노드에 작업을 바로 실행시키고, InfiniBand와 GPU를 사용할 수 있다. 웹 기반의 관리 대시보드를 제공하여 관리자에게 사용자의 계정 생성 및 관리, 게이트웨이 노드의 계정 할당, 가상 게이트웨이 이미지 관리, 스토리지 자원 할당 등의 기능을 제공한다. 가상 게이트웨이 서버도 오픈소스의 OpenStack을 이용하고 있다.

6. Sabalcore의 HPC On-Demand와 Univa의 UniCloud

Sabalcore[16]는 인터넷상으로 주문형/전용 솔루션을 위한 HPC On-Demand 서비스를 제공한다. 서비스 측면에서는 관리자가 사용자, 프로젝트, 자원의 사용에 따른 할당 및 추적을 지원하는 accounting 도구를 제공한다. 특정 작업의 계산 노드 주변에 방화벽을 구축하여

실행하는 작업을 논리적으로 분리하는 기능(dynamic cluster isolation)을 제공하여 미국 국방과 관련한 까다로운 규정도 만족시키고 있다. 빠른 서비스 시작을 지원하기 위해 OS가 RAM에 미리 설치되어 있고 실행할 준비가 되어있는 다수의 오픈소스 애플리케이션을 제공한다. 또한 고객의 SW 설치 및 자신의 SW를 구축하고 설치하도록 허가하고 있다. HW는 노드당 8~32 코어 2.33~3.0 GHz의 Intel Xeon과 AMD Opteron CPU, 32~128GB의 RAM으로 구성되어 있다. 테스트용의 Nvidia의 GPU도 일부 가지고 있다. 이들 자원들은 Ethernet과 InfiniBand의 가상화를 지원하지 않는 네트워크로 연결되어 있다. 기본 파일 시스템은 특정 애플리케이션을 위한 병렬 파일 시스템을 지원하는 NFS (Network File System)를 기반으로 하고 있다. 기본 스케줄러는 Torque이나 사용자의 스케줄러도 사용할 수 있다. 현재는 클라우드 버스팅을 지원하지 않고 있으며, 비용은 용량에 의존하여 RAM, 스토리지, 지원현황, 대역폭 등을 포함하여 0.6~0.29달러/core hour로 책정되

고 있다. 그 외 설치, 애플리케이션 조정, 워크플로우 최적화, 스크립팅, 일반 컨설팅, 디버깅, 문제 해결 등의 포괄적 기술지원 서비스도 하고 있다.

Univa는 Grid Engine의 개발 및 지원뿐 아니라, 파트너 RightScale과 함께 UniCloud[17]라는 “one-click” HPC를 제공하고 있다. 클라우드를 구축하기 위한 두 가지 명시적 방법을 제공하는데, 첫째는 프라이빗 클라우드 인프라이다. 이 방법은 VMware, KVM, Xen, 그 외 OpenStack이나 VMware vSphere 같은 더욱 세련되고 다양한 가상화 기법 중 하나를 사용하여 ‘raw’ VM을 제공한다. 두 번째는 Amazon EC2 AWS와 같은 퍼블릭 클라우드 인프라를 제공하는 것이다. Univa는 바로 CPU 클라우드 가상화를 제공하고, GPU-기반 클라우드는 제공하지 않는다. 10GigE의 네트워크를 지원하며 모두 가상화되고, 성능 향상을 위해 paththrough를 사용하고 있다. 스토리지 옵션은 NFS를 지원하고, Gluster 등의 다른 파일 시스템도 사용할 수 있다. Univa Grid Engine을 사용하여 클라우드 노드의 요청 및 추가/삭제 등을 자동으로 할 수 있다. 클라우드 버스트 기능을 제공하여 내부에서 퍼블릭 클라우드로, 퍼블릭에서 다른 퍼블릭 클라우드로 버스팅이 가능하다.

IV. HPC 클라우드에서 성능 관련 이슈

HPC를 클라우드 컴퓨팅에 적용할 경우 주요 이슈는 물리적 하드웨어와 비교하여 발생 가능한 성능 손실 평가에 관한 것이다. 클라우드에서 성능 불이익은 다음의 두 계층에서 나타날 수 있다[18].

- 가상 엔진(virtual engine): 가상화 메커니즘에 따라 성능 손실이 발생하고, 적용된 가상화 기술의 영향을 많이 받음.
- 클라우드 환경(cloud environment): 클라우드 환경에 따른 상위 계층에서 주로 컴퓨팅과 통신 자원의 공유 및 오버헤드에 의해 발생됨.

경제적 측면에서 HPC와 클라우드 환경에서의 차이점은 HPC는 buy-and-maintain, 클라우드는 pay-per-use이다. 성능 문제는 경제적 문제가 될 수 있다. 클라우드에서 낮은 성능은 긴 대기 시간(long waiting time)과 높은 가격(high cost)을 의미한다. 다양한 가격 책정의 기준들(RAM 메모리 할당량, VM 동작 시간 등)은 다양한 고려사항과 최종 실행 비용을 줄이기 위한 다양한 애플리케이션 최적화가 요구될 것이다.

〈표 3〉은 고전적인 환경과 클라우드 환경에서의 HPC의 차이점을 요약 기술한 것이다[18]. 성능 데이터의 가용성은 애플리케이션이 얼마나 빨리 실행되고, 얼마를 지불할지를 아는 데 중요한 요소가 된다. 그러나 무엇을 어떻게 측정할지에 관해서는 여전히 불확실성이 존재하고, 측정된 데이터의 의미를 해석하는 데에도 어려움이 따른다.

일반적으로 HPC 시스템은 LINPACK, NPB(NAS Parallel Benchmarks), mpttest 등의 벤치마크 실행을 통해 성능을 측정한다. HPC 사용자들을 HW의 사양과 벤치마크 측정 지수를 통해 시스템의 주요 특징과 잠재성을 설명한다. 최근에는 이러한 방법들이 클라우드 환경의 가상 노드에 적용된 HPC에서도 측정 지수로 활용되고, HPC가 클라우드로 서비스되기 위한 장애 요소로 언급되고 있다. Napper와 Beintinesi[19], Ostermann [20]은 Amazon EC2에서 LINPACK 결과가 178GFLOPS (Giga Floating Operations per Second)로 측정되었는데, 이는 top500의 최하위 순위(약 76.41TFLOPS)와도

〈표 3〉 고전적 HPC와 클라우드에서 HPC 차이점

	HPC	HPC in Clouds
비용	Buy-and-maintain	Pay-per use
성능 최적화	HW에 맞춰 애플리케이션 조율	시스템과 애플리케이션 함께 조율
시스템 규모 (System dimensioning)	시스템 획득 시간에, 시스템 관리자 제어로 전역적인 성능 지수 사용	매 애플리케이션 실행 시, 사용자 제어로 애플리케이션 중심의 성능 지수 사용

비교할 수 없이 낮은 결과이다. Walker[21]도 TeraGrid 장비의 물리적 클러스터와 EC2를 비교한 결과 가상화 계층과 클라우드 환경 계층에서 오버헤드가 상당히 높았다. 대부분의 경우 네트워크 대역폭, 지연 등의 이유들이 주요 원인이었다. 그러나 이는 HPC와 클라우드에서 HPC의 차이점을 전혀 고려하지 않고 측정한 결과라 할 수 있다.

다른 지수를 적용할 경우 결과는 바뀔 수 있다. 우선, 애플리케이션 응답 시간은 클라우드 환경에서 중요한 성능 지수가 될 수 있다. 이안 포스터(Ian Foster)[22]는 전체 애플리케이션 응답 시간을 $RT = \langle \text{job submission time} \rangle + \langle \text{execution time} \rangle$ 으로 정의하고 있다. 일반적인 HPC 환경에서 job submission time은 모든 컴퓨팅 자원을 얻을 때까지 몇 분 혹은 몇 시간까지 꽤 길다. HPC 워크로드를 실행하는 클라우드에서는 큐가 없기 때문에 낮은 계산 능력을 제공하더라도 최종 응답 시간은 물리적인 HPC 시스템과 비교할 만 할 것이다.

V. 클라우드에서의 HPC 지원

클라우드에서 HPC 애플리케이션을 지원하기 위해, 클라우드 제공자로부터 적절한 SLA(Service-Level Agreement)가 승인되어야 한다[18]. HPC 사용자를 위한 SLA는 일반 클라우드의 SLA와는 달리 애플리케이션 성능 예측을 보장하기 위해 관련 파라미터를 정형 명세 형태로 제공해야 한다.

Amazon은 고장허용 파라미터(fault-tolerant parameter, 서비스 가동 시간 등)를 붙여 시스템 가용성을 보장한다. 대신 네트워크 성능 예측은 없고, 물리 머신에서 동일한 메모리를 사용하는 VM이 없음을 보장하고 있다. GoGrid는 Amazon의 가용성 파라미터에 더하여 네트워크 파라미터에 관한 보장(Jitter: 지연 < 0.5msec, latency: 포인트 간 패킷 전송에 걸리는 시간 < 5msec, Packet loss: 10초를 초과한 지연 < 0.1% 등)을 제공하

지만, VM 간의 물리적 컴퓨팅 자원의 공유에 대한 보장은 제공하지 않고 있다.

SLA 적용은 HPC 성능을 위한 솔루션이 될 수 있지만, 현재의 솔루션은 너무 일반적인 SLA 약정만을 제공하거나 파라미터를 제어하기엔 너무 적은 값을 제공하고 있다.

성능 측정 기술 및 도구와 관련하여 클라우드 제공자로부터 서비스가 매우 적다. 현재 Amazon의 CloudWatch, Hyperic(Aamazon에 합병)의 CloudSatus 등의 성능 모니터링 도구들이 있으나, 이는 트랜잭션 애플리케이션에 유용한 도구이고, HPC 코드와 같이 오랜 실행을 동반하는 애플리케이션의 행위를 예측하기 위한 특성들은 제공하지 않고 있다.

최근 실험 차원이지만, C-meter[20], PerfCloud[23]와 같은 솔루션들이 벤치마크-온디맨드 서비스로 클라우드에서 제공되는 VM 또는 가상 클러스터를 동적으로 벤치마크하는 프레임워크를 제공하고 있다. C-meter는 GrenchMark 벤치마크(통합 워크로드 생성)를 지원하고 웹 애플리케이션에 맞춰져 있다. PerfCloud는 HPC 환경을 위한 여러 벤치마크(NPB, mptest, SkaMPI 벤치마크 등)를 지원한다.

VI. 결론

2012년 11월 발표된 top500에서 클라우드 컴퓨팅으로 구성된 가상 시스템인 Amazon의 AWS가 102위에 올랐다(2011년 11월 42위). 이처럼 HPC와 클라우드 컴퓨팅을 결합하려는 움직임이 다양하게 시도되고 있다. 본고에서는 이러한 움직임에 맞추어, HPC가 클라우드로 이동할 경우 발생할 수 있는 이슈들을 점검하고, 특히 성능 이슈에 대하여 HPC와 클라우드의 특성을 모두 반영한 성능 측정 지수의 필요성에 대해 언급하였다. 그 외 클라우드로 HPC가 서비스되기 위해 클라우드에서 지원해야 할 기능들과 HPC 클라우드 서비스를 제공하

는 다양한 업체의 현황에 대하여도 기술하였다. 최근에는 퍼블릭 형태보다 프라이빗 클라우드 컴퓨팅 구축 환경을 기반으로 HPC가 확장되는 모습이지만, 앞서 언급된 다양한 업체에서의 HPC 클라우드 서비스들과 제기된 이슈들에 대한 접근이 해결된다면 퍼블릭 클라우드로의 전환도 머지않을 것이다.

용어해설

HPC 고급 연산 문제를 풀기 위하여 슈퍼컴퓨터 및 컴퓨터 클러스터를 사용하는 것을 말하며 최근에는 TFLOPS 이상의 규모를 HPC에 사용되는 컴퓨터를 구분하는 기준으로 활용

HPC 클라우드 HPC 인프라에 클라우드 컴퓨팅을 결합하는 형태로, 기존의 HPC 클러스터와 공유의 그리드 방식과 셸프 서비스, 동적 자원 할당, 광범위 네트워크, 과금 등이 결합된 서비스 형태

약어 정리

AWS	Amazon Web Service
GFLOPS	Giga Floating Operations per Second
HPC	High Performance Computing
NAS	Network Attached Storage
NFS	Network File System
NPB	NAS Parallel Benchmarks
NSF	National Science Foundation
SMB	Small & Medium-sized Business
SMP	Symmetric Multiprocessing
TFLOPS	Tera Floating Operations per Second
VM	Virtual Machine

참고문헌

[1] CERN Cloud Project, 2010. <http://cloud.web.cern.ch/cloud/>

[2] Nebula Cloud Computing Platform, 2010. <http://nebula.nasa.gov/>

[3] D. Klug et al., "Microsoft and NSF Enable Research in the Cloud," NSF, Feb. 4th, 2010. http://www.nsf.gov/news/news_summ.jsp?cntn_id=116336

[4] Boeing, Boeing Takes to the Cloud with a Solution to Market its Best-Selling Airplane. <http://www.cloudconnect-vent.com/santaclara/downloads/whitepapers/Boein>

gAzureFinalrevlocked.pdf

[5] J. Feblowitz, "Oil and Gas: Into the Cloud?," *JPT*, vol. 63, no. 5, 2011. 5. <http://www.spe.org/jpt/print/archives/2011/05/11Management.pdf>

[6] Ucloud HPC, KT. <https://ucloudbiz.olleh.com/portal/ktcloudportal.epc.productintro.hpc.info.html>

[7] 디지털데일리, "고성능컴퓨팅(HPC)과 클라우드의 만남... '효율성 극대화'," 2012. 5. 23. http://www.ddaily.co.kr/news/news_view.php?uid=91166

[8] GreenButton, "Cloud-The Natural Evolution for High Performance Computing," HPC in the Cloud, Nov. 5th, 2012. http://www.hpcinthecloud.com/hpccloud/2012-11-05/cloud_-_the_natural_evolution_for_high_performance_computing.html

[9] Amazon Web Wervice, "AWS에서의 HPC," 2012. <http://aws.amazon.com/hpc-applications/>

[10] 손춘호, "HPC 클라우드 개요," KT 클라우드, 2012. 5.

[11] SGI Cyclone, "CycloneTM-HPC Cloud results on demandTM," 2012. http://www.sgi.com/products/hpc_cloud/cyclone/index.html

[12] E.C. Joseph, S. Conway, and J. Wu, "A New Approach to HPC Public Clouds: The SGI Cyclone HPC Cloud," IDC Technical Computing, Mar. 2010.

[13] S. Chari, "Smart IBM Solutions for High Performance Engineering Clouds," IBM, 2010.

[14] Moab, "HPC Cloud-Moab HPC Suite Enterprise Edition," Adaptive Computing, 2012. <http://www.adaptivecomputing.com/>

[15] J. Bernstein and K. McMahon, "Computing on Demand-HPC as a Service," Penguin Computing, 2012.

[16] "HPC On-Demand," Sabalcore Computing, 2012, http://www.sabalcore.com/service_cluster_on_demand.html

[17] D. Eadine, "Moving HPC to the Cloud," ADMIN HPC, 2012. <http://hpc.admin-magazine.com/Articles/Moving-HPC-to-the-Cloud>

[18] R. Aversa et al., "Performance Prediction for HPC on Clouds," *Cloud Computing: Principles and Paradigms*, R. Buyya, J. Broberg, and A.M. Goscinski, Ed., John Wiley & Sons, 2011, pp. 437-454.

[19] J. Napper and P. Bientinesi, "Can Cloud Computing Reach the top500?" *Proc. Combined Workshops Unconventional High Perform. Comput. Plus Memory Access (UCHPC-MAW)*, May 2009.

- [20] S. Ostermann et al., "An Early Performance Analysis of Cloud Computing Services for Scientific Computing," Technical Report, TU Delft/PDS, PDS-2008-006, Dec. 2008.
- [21] E. Walker, "Benchmarking Amazon ec2," *USENIX ;Login;*, vol. 33, no. 5, Oct. 2008. pp. 18-23.
- [22] I. Foster, "A Critique of 'Using Clouds to Provide Grids...'," Sept. 2008. <http://ianfoster.typepad.com/blog/2008/09/a-critique-of-u.html>
- [23] E.P. Mancini, M. Rak, and V. Villano, "PerfCloud: Grid Services for Performance Oriented Development of Cloud Computing Applications," *Proc. WETICE*, 2009, pp. 201-206.