

응용의 데드라인에 기반을 둔 하이브리드 클라우드 자원 스케줄링 서비스

고정인, 김서영, 김윤희
숙명여자대학교 컴퓨터과학부
e-mail : {jungin, sssyyy77, yulan}@sookmyung.ac.kr

Hybrid Cloud Resource Scheduling Service Based on Deadline of an Application

Jung-in Koh, Seoyoung Kim, Yoonhee Kim
Dept. of Computer Science, Sookmyung Women's University

요 약

수년간 그리드 자원은 고성능 컴퓨팅을 사용하는 여러 과학 연구 분야의 발전에 크게 이바지해왔다. 최근 클라우드의 등장으로 컴퓨팅 자원 확장이 유연해지면서 과학 분야에서는 하이브리드 클라우드를 활용하려는 시도가 증가하고 있다. 본 논문에서는 응용별 데드라인을 기준으로 그리드 및 하이브리드 클라우드의 효율적인 활용을 극대화하는 자원 스케줄링 서비스를 제안한다. 각 응용에 적합한 데드라인을 기준으로 자원들의 활용도를 높일 수 있고, 응용 수행의 실시간 모니터링 통해 필요에 따른 추가 자원 요구를 공용 클라우드의 자원 활용으로 응용 수행시간을 최적화하는 자원 스케줄링을 제공한다. 또한 제안하는 스케줄링은 공용 클라우드 이용 시 발생하는 비용에 대한 사전 고려를 통해 경제적인 자원 서비스를 제공한다.

1. 서 론

계산과학연구에서 주로 활용된 그리드 자원은 여러 컴퓨팅 자원을 단일 환경으로 통합·운영함으로써 우수한 컴퓨팅 성능을 제공한다. 하지만 이러한 고성능 컴퓨팅은 통합된 물리 자원의 개수에 의해 성능 자체가 국한되므로, 물리 자원이 갖는 컴퓨팅 그 이상의 성능을 기대하기 어렵다. 즉, 그리드가 갖는 컴퓨팅 이상의 성능이 요구될 경우, 물리 자원의 확장이 필요하다. 그러나 일시적으로 증가하는 전력 수요에 따른 자원 확장 및 유지보수 비용에 비해 추가한 자원의 효율성이 높지 않다는 점을 고려해서 필요에 따라 동적으로 자원 확장 및 축소가 가능한 클라우드 컴퓨팅 기술이 다변적 컴퓨팅 수요에 많이 이용되는 추세이다. 최근 공용 및 사설 클라우드를 통합하여 활용하는 하이브리드 클라우드 환경이 주목을 받으면서 하이브리드 클라우드에 대한 스케줄링의 필요성이 증대되었다.

본 논문에서는 응용의 데드라인을 고려하여 그리드 및 하이브리드 클라우드의 효율적으로 활용을 돕는 자원 스케줄링 서비스를 제안한다. 각 응용에 적합한 데드라인을 기준으로 자원 스케줄링을 제공하여 그 활용을 극대화하고, 실시간 모니터링을 통해 추가적 요구가 감지되면, 공용 클라우드의 자원을 사용해 데드라인을 만족하는 자원 스케줄링을 제공한다. 또한 제안하는 스케줄링은 공용 클라우드 이용 시 비용 최소화를 위한 고려를 통해 사용자에게 경제적인 자원 서비스 제공이 가능하다.

본 논문은 2장에서 관련 연구를 소개하고, 3장에서 자원 스케줄링 서비스에 대한 전체적인 구조 설명을 하고, 4장에서는 서비스의 진행 과정을 시나리오 형식으로 구체화한다. 5장에서는 결론 및 향후 연구 방향에 대한 정리로 마무리한다.

2. 관련 연구

최근 하이브리드 클라우드의 활용이 급증함에 따라, 이를 효율적으로 활용하는 동시에 사용자에게 경제적인 자원 서비스를 제공하는 정책을 위한 필요성이 증가하고 있다. 이를 위해 하이브리드 클라우드를 대상으로 더욱 효율적인 자원 활용 및 사용자 요구 만족 극대화를 위한 몇몇 스케줄링 연구가 진행되고 있다[1], [2].

[1]은 하이브리드 클라우드 상에서 비용과 수행 QoS 모두를 고려하는 스케줄링 제공을 통해 사용자 요구에 따른 경제적인 클라우드 서비스 모델을 제시하였다. 각 작업량은 사용자가 제시한 데드라인, 메모리, CPU, 데이터 전송량과 같은 요구로 표현하여, 이를 통해 최적화 문제를 구성 후 선형계획법을 통해 분석한다. 적절한 VM 인스턴스 선택을 위해서는 이진정수계획법을 활용한다. 하지만 단지 사용자가 제시한 속성 위주의 스케줄링을 제공하므로, 실행 작업 특성이 다양한 경우, 오히려 전반적인 성능 저하 및 부하의 결과를 낳을 수 있다.

Comet Cloud[2]는 대규모 그리드, 사설 및 공용 클라우드를 대상으로 데드라인 정책에 따라 제공할 자원의 종류를 달리하여 제공한다. 응용 계층을 정의하여 자원 활용도를 높이며 동시에 비용 절감이 가능한 자원 서비스를 제공한다. 데드라인과 비용에 따른 정책을 제시하지만, 두

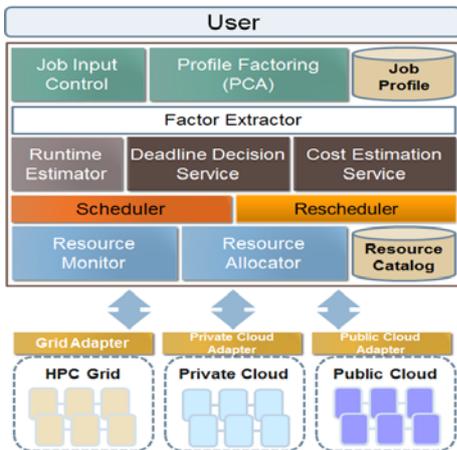
"본 연구는 한국과학기술정보연구원의 응용연구서비스개발과제의 연구결과로 수행되었음" (K-12-L06-C02-S05)

가지를 동시에 고려한 자원 스케줄링은 제공하지 않는다는 점에서 한계가 존재한다.

3. 자원 스케줄링 서비스

본 연구에서 제안하는 하이브리드 클라우드 스케줄링은 크게 응용의 데드라인 결정, 자원 스케줄링, 자원 할당이 3가지 역할로 나뉘지고, 전체적인 스케줄링 서비스의 구조는 그림 1에 자세히 나타나있다.

응용의 데드라인은 작업이력을 바탕으로 PCA(Principle Component Analysis)[4]를 통해 해당 응용에 맞는 데드라인을 결정하게 된다. Profile Factoring을 통해서 각 인자가 응용에 미치는 영향력을 수치화하고, 이를 통해서 얻은 높은 수치의 인자로 인해 추출된 작업이력 중에서 Dead-line Decision Service가 응용에 적합한 데드라인을 제공한다. 자원스케줄링은 Scheduler, Rescheduler, Run time Estimator, Resource Monitor로 구성된다. Runtime Estimator를 통해 작업의 수행시간을 예측하고, Resource Monitor에서 자원의 상태를 주기적으로 모니터링한다. Scheduler는 예측된 작업 수행시간과 현재 자원 상태를 고려해, 해당 응용의 작업을 어떤 자원에 할당할지를 결정하고, Rescheduler는 수행 중인 작업의 지연으로 데드라인을 맞출 수 없다고 판단되면, 대기 중인 작업들에 대해 다시 스케줄링을 수행한다. 공용 클라우드 자원 할당 시엔, Cost Estimation Service에서 최소비용의 인스턴스 선택 서비스를 제공한다. Resource Allocator는 Scheduler에서 정한 Resource Adapter로 수행할 작업을 전달해서 실질적인 작업 수행을 제공하는 역할을 한다.

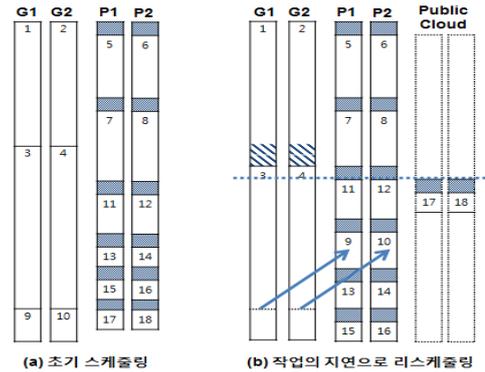


(그림 1) 자원 스케줄링 서비스 구조

4. 자원 스케줄링 시나리오

자원 스케줄링 서비스는 Bag-of-Task 속성[5]을 가진 응용을 대상으로 한다. 우선, 스케줄링을 위해서 한 응용의 데드라인 D를 22시간으로 두고, 한 응용에 속하는 작업이 각각 10시간, 5시간, 2시간, 1시간이 소요되는 작업들 4개씩이라고 한다면 이 작업들에 대해서 수행시간이 긴 순서대로 정렬을 시킨 후, 정렬된 작업은 그리드-사설 클라우드-공용 클라우드 순으로 해당 자원에 작업을 수행할

수 있는지의 여부를 판단한다. 이 때, 자원으로는 그리드(코어 2개: G1, G2), 사설 클라우드(코어 2개: P1, P2), 공용 클라우드가 있다고 하자. 초기 스케줄링 시, 가용 코어가 존재하면 작업이 수행되고, 그렇지 않으면 (대기시간+수행시간)을 고려해서 데드라인 안에 수행이 가능하다고 판단되면 해당 코어에 예약되고, 데드라인에 위배된다면 다음 자원의 가용 코어를 확인하는 형식으로 초기 스케줄링이 된다. 만약 작업의 지연이 발생하면 리스케줄러가 대기 중인 작업을 모아 다시 스케줄링 하는데 그림 2와 같이 데드라인을 넘기는 작업들을 최소화 수행시간이 짧은 작업으로 간추려서 공용 클라우드에 스케줄링하게 된다.



(그림 2) 리스케줄링 시나리오의 예

5. 결론 및 향후 연구 방향

본 논문에서 제안하는 스케줄링 서비스는 주어진 데드라인을 통해서 응용별 수행 시간의 최적화를 가능케 하며, 데드라인을 맞출 수 없는 경우, 상대적으로 수행시간이 짧은 작업들을 간추려서 공용 클라우드 자원에 스케줄링하기 때문에 자원 사용 비용이 절감이 가능하다. 이로써 스케줄링 서비스는 급증하는 전산 수요에 따른 자원 증설 및 관리 비용을 절감시키고, 기존 유휴 자원에 대한 효율적 활용을 증대시키는 효과를 보인다.

향후에는 스케줄링 서비스를 적용한 실험을 통해서 본 논문의 주장을 확고히 하고, 각 응용에 대한 스케줄링 적용결과에 대한 평가 기능을 구현하여 스케줄링의 정확성을 높일 계획이다.

참고 문헌

[1] Ruben Van den Bossche, Kurt Vanmechelen, Jan Broeckhove, "Cost-Efficient Scheduling Heuristics for Deadline Constrained Workloads on Hybrid Clouds", 3rd IEEE International Conference on Cloud Computing technology and Science, 2011.
 [2] Comet Cloud, <http://www.cometcloud.org>
 [3] Amazon Elastic Cloud (EC2), <http://aws.amazon.com/ec2>
 [4] Jolliffe, I.T. "Principal Component Analysis", Springer Verlag, 2002.
 [5] W. Cirne, F. Brasileiro, J. Sauve, N. Andrade, D. Paranhos, E. Santos-Neto, and R. Medeiros, "Grid Computing for Bag of Tasks Applications," Proceedings of the 3rd IFIP Conference on E-Commerce, E-Business and E-Government, September 2003.