

멀티 클라우드 가상머신의 리전 배치 최적화를 지원하는 추천 기법

진진혁¹⁰, 정수민², 박준석³, 엄근혁^{4*}교신저자

¹부산대학교 정보컴퓨터공학부 학부생

²부산대학교 정보융합공학과 박사과정

³부산대학교 지능물류빅데이터연구소 연구교수

^{4*}부산대학교 정보컴퓨터공학부 교수

{asftoy, sumin2708, pjs50, yeom}@pusan.ac.kr

Recommendation Approach to Support Region Optimal Deployment of Multi-cloud Virtual Machines

Jinhyeok Jeon¹⁰, Sumin Jeong², Joonseok Park³, Keunhyuk Yeom^{1*}

¹School of Computer Science and Engineering, Pusan National University

²Dept. of Information Convergence Engineering, Pusan National University

³Research Institute of Intelligent Logistics Big Data, Pusan National University

요 약

멀티 클라우드 서비스 제공을 위해 가상머신 환경 구축 시 서비스의 성능, 비용 등을 최적화하면서 가상머신의 적절한 배치 위치를 결정하는 것은 중요한 연구 이슈 중 하나이다. 따라서 본 논문에서는 가상머신 배포 시 비용과 성능에 기반한 가상머신 배치 위치 최적화 및 추천 기법을 제시한다. 제안하는 방법은 NSGA-II(Non dominated Sorting Genetic Algorithm-II)를 활용한 다목적 최적화를 수행하고, 가중합 기반 점수 산출을 통해 최적 배치 위치를 추천한다. 멀티 클라우드의 14개 리전에서 4개의 가상머신 배치 실험을 수행한 결과, 총 1,716개의 배치 경우의 수 중 2개의 최적화 된 배치 방법을 획득하였다. 본 논문의 방법은 멀티 클라우드 환경에서 효율적인 가상머신 배치 및 마이그레이션을 통한 재배치 기법의 기반 기술로 활용될 수 있을 것이다.

1. 서론

최근 클라우드 환경은 벤더 종속성 탈피를 위해 멀티 클라우드 환경을 도입하고 있으며, 멀티 클라우드 환경에서 가상머신(Virtual Machine)을 배치하는 방법으로 멀티 리전 전략과 마이크로서비스 아키텍처[1]를 활용할 수 있다.

멀티 리전 전략은 여러 리전(Region, 물리적인 데이터 센터를 운영하는 지역)에 가상머신을 분산 배포하는 방법이다. 마이크로서비스 아키텍처는 애플리케이션을 구성하는 개별 서비스를 분리하고, 각 서비스를 RESTful API와 같은 약결합 기술을 통해 연결하는 방법이다.

그러나, 멀티 클라우드 환경에서 멀티 리전 전략과 마이크로서비스 아키텍처를 활용한 가상머신 배치 시, 성능 감소, 비용(price) 증가 등의 서비스 품질이 저하될 수 있다. 따라서, 본 논문에서는 멀티 클라우드 리전 별 운영 비용 및 가상머신 간의 RTT(Round Trip Time) 성능을 활용한 멀티 클라우드 가상머신 최적 배치 탐색 기법을 제안한다.

2. 관련 연구

X. He 등[1]은 클라우드 환경에서 마이크로서비스 시스템의 평균 응답 시간 감소를 위해 탐욕 알고리즘 기반 배치 최적화를 제안하였다. 또한, J. Dognai 등[2]은 포그 클라우드 환경에서 NSGA-II(Non-dominated Sorting Genetic Algorithm-II)를 활용한 컨테이너 서비스 배치 최적화를 연구하였다. 반면, 본 연구는 멀티 클라우드 환경에서의 가상머신 배치 최적화를 위해 유전 알고리즘 중 하나인 NSGA-II 기반 다목적 최적화를 수행하였다는 차이가 있다.

3. 멀티 클라우드 가상머신 배치 최적화 기법

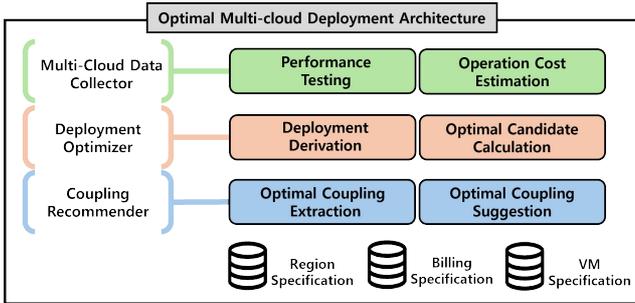
<표 1> CSP 리전별 가상머신 운영 비용 및 RTT

CSP	Region	Spec	price (\$)	RTT (ms)
AWS	eu-central-1	t2a.large (AWS),	2.004	10.313
	eu-west-1		2.189	37.576
GCP	europa-central2	e2-standard-2 (GCP)	2.072	30.273
	europa-west1		1.770	19.952

<표 1>에서 t2a.large와 e2-standard-2는 동일한 가상머신 하드웨어를 나타내며, 동일 환경에서 리전에 따른 비용과 RTT 성능이 다른 것을 확인할 수 있다. 따라서, 멀티 클라우드 가상머신 배치 시 비용과 RTT의 최적화가 필요함을 알 수 있다.

3.1 제안 아키텍처

본 논문에서 제안하는 멀티 클라우드 가상머신 배치 추천 아키텍처는 (그림 1)과 같다.

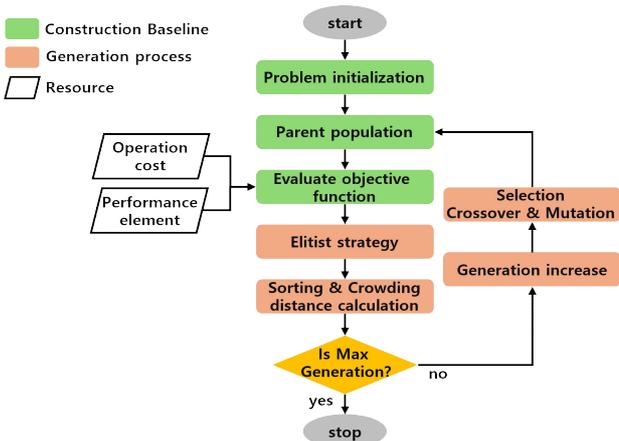


(그림 1) 멀티 클라우드 가상머신 배치 최적화 아키텍처

멀티 클라우드 가상머신 배치 최적화 아키텍처는 비용 및 성능 데이터를 수집하는 Multi-Cloud Data Collector, 수집된 정보를 바탕으로 다목적 최적화를 진행하여 최적 배치 후보를 도출하는 Deployment Optimizer, 가중합 기반 점수를 통해 최적 배치 방법을 추천하는 Coupling Recommender로 구성된다.

3.2 최적화 수행 알고리즘

본 논문의 다목적 최적화는 비용과 성능 요소를 적용한 NSGA-II를 활용하며 (그림 2)와 같다.



(그림 2) 제안하는 방법의 NSGA-II 수행 절차

이후, 도출된 최적 배치 후보에서 가중합 기반 점수를 산출하고 최적 배치 추천을 수행한다.

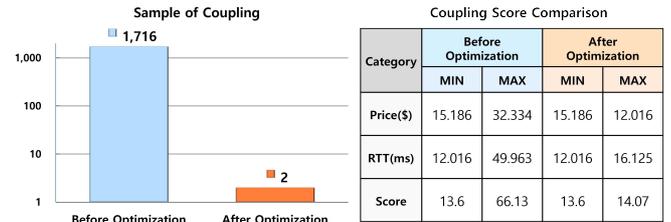
4. 실험 및 평가

(그림 1) Coupling Recommender의 최적 배치 추천 시 활용하는 점수는 식 (1)을 통해 산출한다.

$$S_{total} = w_{price} \cdot S_{price} + w_{rtt} \cdot S_{rtt} \quad (w_{price} + w_{rtt} = 1)$$

식(1)에서 S_{total} 은 가중합 결과를 나타내며, S_{price} , S_{rtt} 는 각각 수집된 price와 RTT를 나타낸다. 또한 w_{price} , w_{rtt} 는 price, RTT의 가중치를 나타내며, 각 가중치의 합은 항상 1을 만족한다.

(그림 3)은 AWS와 GCP의 14개 리전을 대상으로 4개의 가상머신 배치 최적화를 수행한 결과를 나타낸 것이다.



(그림 3) 배치 최적화 수행 결과

배치 최적화 수행 결과로 1,716개의 배치 경우의 수가 2개로 줄어드는 것을 확인하였다. 또한, 점수 산출 결과물에서 확인할 수 있듯이 최소 점수에 수렴하는 결과를 획득하였다.

5. 결론

본 논문에서는 멀티 클라우드 환경에서 비용과 RTT를 측정하여 가상머신 배치를 최적화하는 기법을 제시하였다. AWS와 GCP의 리전에 대해 배치 최적화 실험을 수행하여 전체 배치 경우의 수 1,716개 중 2개의 최적화된 배치 방법을 도출하였다. 본 논문의 가상머신 배치 최적화 기법은 멀티 클라우드 환경에서 성능 기반 마이그레이션 기술에 활용될 수 있을 것이다.

사사

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 융합보안핵심인재양성사업의 연구 결과로 수행되었음 (IITP-2024-2022-0-01201)

참고문헌

[1] X. He, H. Xu, X. Xu, Y. Chen, and Z. Wang, "An Efficient Algorithm for Microservice Placement in Cloud-Edge Collaborative Computing Environment," IEEE Transactions on Service Computing, Vol. 14, No. 8, pp.1-16, 2024.
 [2] J. Dognai, A. Yazdanpanah, A. Zare, and F. Khunjush, "A two-tier multi-objective service placement in container-based fog-cloud computing platforms," Cluster Computing, Vol. 27, pp. 4491-4514, 2024.