

하이브리드 클라우드 환경에서 자동 백업 시스템 기반

워크로드 이전기능 설계

권무성, 윤지혜, 김영한*

*숭실대학교

juniyo97@dcn.ssu.ac.kr, dbswlgp99@dcn.ssu.ac.kr, younghak@ssu.ac.kr

Design of workload migration function based on automatic backup system in hybrid cloud environment

Kwon Mu Seong, Yun Ji Hye, Kim Young Han*

*Soongsil Univ.

요약

클라우드 환경은 다양한 이유로 인해 문제 상황이 발생할 수 있으며, 그로 인해 워커노드가 중단되었을 경우 다른 클라우드로의 복원이 요구될 수 있다. 이때 워크로드 복원에 필요한 데이터에는 여러 종류가 있으며, 데이터의 종류에 따라 백업 방식이 달라지게 된다. 이에 본 논문에서는 클러스터에서 하나의 워커노드가 중단된 경우를 가정하여 그때 복원을 위해 필요한 백업 데이터와 백업 방법에 대해 논의하고자 한다.

I. 서론

클라우드 환경은 데이터 센터의 화재, 자연 재해 등 다양한 이유로 인해 문제가 발생할 수 있다. 이때 쿠버네티스 클러스터의 Control Plane 과 Etcd 등은 체인화 되어 HA(High Availability)로 구성되어 있을 경우 문제없이 동작할 수 있지만, 워커노드는 복제된 형태로 존재하지 않아 그 워커노드에서 실행되던 워크로드들이 중단될 수 있다. 이때 기본적으로 Control Plane 은 중단된 Pod 들을 다시 실행하기 위해 다른 워커노드로의 배포를 시도하지만, 클러스터에 리소스가 부족할 경우 배포에 실패하게 된다. 따라서 그러한 경우 다른 클라우드를 활용하여 워크로드의 복원을 수행해야 한다. 그리고 이때 복원을 위해서는 인프라와 리소스에 대한 백업 작업이 선행되어 있어야 하고, 그 백업된 데이터를 통해 워크로드를 기존과 동일하게 실행할 수 있어야 한다. 그러나 백업을 수행할 때 노드 그대로를 이미지화하여 백업하면 마스터와의 종속성으로 인해 워커노드가 정상적으로 동작할 수 없기 때문에, 각 백업 데이터의 유형에 맞는 방법을 사용하여 백업을 수행해야 한다. 이에 본 논문에서는 쿠버네티스 클러스터에서 하나의 워커노드가 중단된 경우를 가정하여 그때 복원을 위해

필요한 백업 데이터와 백업 방법에 대해 다루고자 한다.

II. 본론

본론에서는 어떤 데이터들을 어떻게 백업 하는지에 대해 백업하는 방식과 백업하는 대상으로 세분화하여 분석한 내용을 서술하고, 자동 백업을 수행하여 워크로드를 이전하기 위한 기능을 제안한다.

백업된 리소스를 복원하기 위해서는 크게 인프라와 리소스의 복원 작업이 필요한데, 인프라는 초기에 워커노드 생성 시 OS, Container Runtime 등의 소프트웨어를 포함하는 VM 이미지를 생성하여 remote 스토리지에 저장해 놓으면 추후 해당 이미지를 사용하여 워커노드의 인프라를 복원할 수 있다. 이와 달리 리소스에 관련된 데이터들은 클러스터 상에서 애플리케이션을 운영하면서 지속적으로 변화하는 데이터이고, 각 데이터의 종류에 따라 저장되는 위치가 다르기 때문에 적합한 방식을 이용하여 주기적으로 백업해 주어야 한다.

1. 백업 데이터 유형에 따른 백업 방식

쿠버네티스 클러스터에서 실행되는 워크로드의 복원이 가능하도록 하기 위해 백업이 요구되는 데이터는 다음과 같다. 모든 백업된 데이터는 remote

스토리지에 저장되며, 데이터의 종류에 따라 백업 방식이 달라진다.

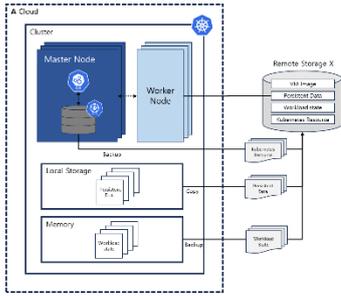


그림 1. 백업 데이터 유형에 따른 백업 방식

1.1 쿠버네티스 리소스

워커노드에서 실행되는 리소스에는 Pod, Service 등과 같은 기본적인 쿠버네티스 리소스가 있으며 이 리소스들에 대한 데이터들은 주기적인 백업을 필요로 한다. 쿠버네티스 리소스 데이터에 대한 백업은 Etcd 데이터베이스를 통해 수행하며, 추출한 데이터는 remote 스토리지에 백업해 두어야 한다. 이때 실행되고 있던 워크로드가 stateless 한 워크로드일 경우에는 복원을 위해 추가적인 데이터를 필요로 하지 않지만, stateful 한 워크로드일 경우에는 워크로드 실행에 필요한 다른 데이터들도 함께 백업해 주어야 한다.

1.2 Persistent 데이터

Persistent 데이터는 remote 스토리지에 저장되는 데이터와 local 에 저장되는 데이터로 나눌 수 있다. Remote 스토리지에 저장되는 데이터는 별도의 백업이 요구되지 않는 반면, local 에 저장되는 데이터는 주기적으로 remote 스토리지에 백업해 주어야 한다. 이때는 단순히 데이터를 복사하는 것만으로 백업을 수행할 수 있다.

1.3 워크로드 상태 정보

메모리상에 존재하는 워크로드의 상태 정보 또한 remote 스토리지로의 주기적인 백업이 요구되는 데이터이다. 이 상태 정보를 백업함으로써 워크로드 복원 시 이전에 실행 중이던 상태 그대로 완전한 복원을 수행할 수 있다.

위와 같은 데이터의 백업은 워크로드의 자동 백업을 지원하는 오픈소스 프로젝트인 Velero[1] 및 Checkpoint 동작을 통해 워크로드의 스냅샷을 저장하는 CRUI[2] 등을 통해 수행할 수 있다.

2. 복원/이전

백업된 데이터를 활용하여 워크로드를 복원하기 위해서는 먼저 워커노드가 준비되어 있어야 한다. 이를 위해 복원을 수행하는 클라우드에는 이미 독립적인 쿠버네티스 클러스터가 구축되어 있는 환경을 가정한다. 워커노드를 등록하기 위해 OS, Container Runtime 등의 소프트웨어가 포함되어 있는 VM

이미지를 사용하여 VM 을 생성하고, Control Plane 에 연결한다. 그리고 remote 스토리지에 저장된 백업 데이터를 이용하여 리소스의 복원을 수행한다. 그러나 이때 백업 데이터가 저장되어 있는 remote 스토리지가 복원을 수행하고자 하는 클라우드의 스토리지가 아닐 경우, 데이터에 접근이 가능하도록 부여해 주어야 한다. 그러면 remote 스토리지에 존재하는 쿠버네티스 리소스, 워크로드 실행 상태 정보, 가능해진다. 여기서 쿠버네티스 리소스와 워크로드 실행 상태 정보, local 에서 백업된 persistent 데이터는 클러스터로 불러와서 복원해야 하고, 나머지 persistent 데이터는 remote 스토리지에 있는 데이터를 사용하면 된다.

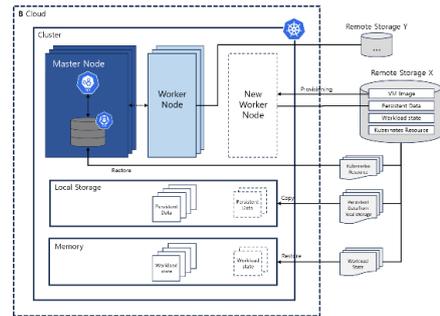


그림 2. 퍼블릭 클라우드로의 워크로드 복원 방법

이와 같은 과정을 통해 중단된 워커노드의 워크로드를 모두 복원하여 퍼블릭 클라우드로의 이전을 수행할 수 있다.

III. 결론

본 논문에서는 자동 백업 및 복원 시스템 기반으로 하여 워크로드를 다른 클라우드로 이전하기 위한 방법에 대해 분석하였다. 쿠버네티스 클러스터에서 하나의 워커노드가 중단된 경우를 가정하여 그때 복원을 위해 필요한 백업 데이터와 백업 방법에 대해 다루었으며 복원 과정에서 요구되는 여러 절차에 대한 수행 방안을 제시하였다. 이를 통해 필요에 따라 워커노드를 다른 클라우드로 이전했다가 이후 다시 복원하는 등 클라우드를 유연하게 활용할 수 있다.

ACKNOWLEDGMENT

이 논문은 2024 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00946, 하이브리드 클라우드 환경에서의 고속, 자동 서비스 복구 및 이전 소프트웨어 개발)

참 고 문 헌

- [1] 강문준 and 김영한, "쿠버네티스 기반 클라우드 시스템에서의 백업 및 복원," 한국통신학회 학술대회논문집, pp. 1337-1338, 2023
- [2] CRUI-mainpage, https://criu.org/Main_Page, 2024