
Pusan National University

Computer Science and
Engineering Technical Report

2025-09

멀티 클라우드 인프라 기반 연합학습 환경 구축 플랫폼 개발



저자1 202055595 전진혁

저자2 202155526 김민경

저자3 201924474 박재일

지도교수 염근혁

목 차

1. 서론	4
1.1. 연구 배경	4
1.2. 기존 문제점	5
1.2.1. 클라우드 벤더 종속성(Cloud Vendor Lock-in)	5
1.2.2. 지역 시간 및 비용 최적화 부재	5
1.2.3. 보안 취약성 및 데이터 프라이버시 문제	5
1.2.4. 동적 태스크 오케스트레이션 부재	5
1.3. 연구 목표	6
2. 연구 배경	6
2.1. 요구사항 분석	6
2.1.1. 시스템 요구사항	6
2.2. 선행기술 조사	9
2.2.1. AWS	9
2.2.2. GCP	10
2.2.3. OpenStack	11
2.2.4. Prometheus	12
2.2.5. Grafana	14
2.2.6. Terraform	15
2.2.7. Flower	16
2.2.8. MLflow	17
2.2.9. Docker	18
3. 연구 내용	19
3.1. 설계 상세화	19
3.1.1. 유스케이스 다이어그램	19
3.1.2. 유스케이스 명세	20
3.1.3. 클래스 다이어그램	33
3.1.4. 시퀀스 다이어그램	40
3.2. 시스템 구성	57
3.2.1. 시스템 아키텍처	57
3.2.2. 개발환경	58
3.3. 멀티 클라우드 인프라 기반 연합학습 환경 구축 지원 플랫폼	59
3.3.1. 사용자 인증	59
3.3.2. 클라우드 인증 정보 관리	59
3.3.3. 연합학습 집계자 관리	60

3.3.4. 연합학습 참여자(클러스터) 관리	64
3.3.5. 연합학습 관리	66
3.3.6. 글로벌 모델 관리	68
4. 연구 결과 분석 및 평가	69
4.1. 사례 연구 - 코로나19 진단 연합학습 모델 생성	69
4.1.1 환경 설정	69
4.1.2. 데이터셋 및 모델 구조	70
4.2. 평가	70
4.2.1. 연합학습 집계자 최적화 평가	70
4.2.2. 연합학습 동적 태스크 오케스트레이션 평가	71
4.2.3 유사 시스템 비교 평가	73
5. 결론 및 향후 연구 방향	75
6. 구성원 별 역할	77
7. 참고 문헌	78

1. 서론

1.1. 연구 배경

연합학습(Federated Learning)[1]은 참여자(Client)와 집계자(Aggregator)로 구성된 분산학습[2] 기술이다. 연합학습은 기존의 중앙집중형 기계학습(Machine Learning)과 달리 데이터를 중앙 서버로 모으지 않고 각 참여자의 로컬 환경에서 모델을 학습한 후 모델의 파라미터만 집계자에게 전송하고, 원본 데이터는 로컬 환경에 유지한다[3]. 이러한 연합학습의 특성은 데이터 유출 위험을 감소시켜, 민감한 데이터를 다루는 의료, 제조, 스마트시티 등의 분야에서 활발하게 활용되고 있다.

멀티 클라우드(Multi-Cloud)[4]는 이형의 클라우드 플랫폼(퍼블릭 클라우드, 프라이빗 클라우드)을 통합하여 단일 클라우드 플랫폼처럼 활용할 수 있는 기술이다. 멀티 클라우드와 연합학습을 결합한 플랫폼은 개별 클라우드에서 획득할 수 있는 장점(비용 효율성, 접근성, 보안성 등)을 적용하여 구축할 수 있다[5]. 예를 들어, 개인의 민감한 의료 정보와 같이 강력한 보안성을 요구하는 경우 데이터를 프라이빗 클라우드에 저장하여 보안을 강화하고, 다수의 데이터를 집계했을 때 가치를 발휘하는 질병 통계와 같은 정보는 퍼블릭 클라우드에 저장하여 접근성과 활용도를 높이는 전략을 수립할 수 있다.

하지만 대부분의 기존 연합학습 플랫폼은 단일 클라우드 환경을 기반으로 연합학습에 필요한 환경을 구축하므로 특정 클라우드에 종속되어 비용 및 리소스, 낭비가 발생할 수 있으며, 클라우드 플랫폼별 가용 리전의 물리적 위치에 따라 지연 시간(latency)이 증가하여 실시간 응답성이나 학습 속도에 영향을 미칠 수 있다. 또한 퍼블릭 클라우드 상에서 연합학습을 수행할 경우, 보안적 조치가 요구되는 데이터가 외부 인프라에 저장되어 보안 위협에 노출될 가능성이 높아진다.

이에 본 연구에서는 멀티 클라우드 환경에서 효율적인 연합학습 환경 구성을 위한 플랫폼을 제안한다. 본 연구에서 제안하는 플랫폼은 1) 클라우드 별 비용 및 지연 시간을 고려한 멀티 클라우드 기반 연합학습 환경 구축, 2) 연합학습 집계자 - 연합학습 참여자 계층 기반의 멀티 클라우드 지원 연합학습 방법 도출, 3) 연합학습 참여자 모니터링을 통한 연합학습 동적 태스크 오케스트레이션 기술을 제시한다.

또한, 본 연구의 사례 연구는 폐 이미지를 이용한 코로나19 진단 CNN 모델 학습을 수행한다. 폐 이미지를 비롯한 의료 데이터는 개인의 임상 정보를 포함하며 학습 시 프라이버시 보호가 요구된다. 이를 통해 데이터 공유 시 법적, 윤리적 제약이 존재하는 현실 상황에서 멀티 클라우드 기반 연합학습의 실효성을 입증할 수 있을 것으로 사료된다.

1.2. 기존 문제점

1.2.1. 클라우드 벤더 종속성(Cloud Vendor Lock-in)

AWS(Amazone Web Services)[6], Azure[7], GCP(Google Cloud Platform)[8] 등 퍼블릭 클라우드 플랫폼은 각기 고유한 API, 서비스 구조 및 관리 도구를 제공한다. 이들은 자사 생태계 내에서의 통합과 확장을 용이하게 설계되어 있으나, 플랫폼 간 호환성이 제한적이라는 문제가 존재한다. 이러한 클라우드 벤더 종속성은 비용 최적화 기회를 놓치게 하고, 서로 다른 클라우드 플랫폼 간의 자원을 통합적으로 활용하기 어렵게 한다.

1.2.2. 지연 시간 및 비용 최적화 부재

IBM Federated Learning[9], AWS Sage Maker[10], GCP Vertex AI[11]와 같은 클라우드 기반 연합학습 지원 플랫폼은 일반적으로 클라우드 리전 간 비용 차이와 연합학습 참여자와의 네트워크 지연을 고려한 최적화 기능을 제공하지 않는다. 이로 인해 연합학습 집계자와 연합학습 참여자가 물리적으로 멀리 떨어져 있는 경우, 네트워크 지연 시간이 길어져 모델의 학습 속도가 저하될 수 있다. 또한 클라우드별로 제공되는 가용 리전마다 비용 정책이 상이하므로 비용 효율적인 연합학습 환경을 구성하기 어렵다.

1.2.3. 보안 취약성 및 데이터 프라이버시 문제

AWS Sage Maker, GCP Vertex AI 시스템들은 프라이빗-퍼블릭 클라우드 간 계층적 구조를 통한 데이터 보안 강화 메커니즘이 부재하다. 대부분의 플랫폼은 모든 연합학습 프로세스를 단일 클라우드 환경에서 처리하므로, 보안적 조치가 필요한 데이터가 퍼블릭 클라우드에 저장된다는 문제점이 존재한다. 이로 인해 퍼블릭 클라우드 서비스 제공자(CSP)에 데이터에 대한 통제권을 위임하므로, 데이터 유출 및 보안 위협이 증가한다.

1.2.4. 동적 태스크 오케스트레이션 부재

대표적인 연합학습 프레임워크인 FedML[12]은 클라우드 환경에서 분산 학습을 지원하는 기능을 제공한다. 그러나 클라우드 내에서 운용되는 가상머신의 CPU, GPU, 메모리, 디스크 등 자원 상태를 실시간으로 고려하여 학습 작업을 배분하거나 장애 발생 시 동적으로 재구성하는 기능은 제한적이다. 이러한 한계로 인해 특정 가상머신에 과부하가 집중되거나, 자원 부족 및 비정상 상태로 인한 학습 중단, 참여자 이탈, 연합학습 실패와 같은 문제가 발생할 수 있다. 따라서 안정적인 연합학습 수행을 위해서는 동적 태스크 오케스트레이션을 지원하는 메커니즘이 필수적이다.

1.3. 연구 목표

- ① 클라우드 별 비용 및 지연 시간을 고려한 멀티 클라우드 기반 연합학습 환경 구축
- ② 연합학습 집계자 - 연합학습 참여자 계층 기반의 멀티 클라우드 지원 연합학습 방법 도출
- ③ 연합학습 참여자 모니터링을 통한 동적 태스크 오케스트레이션 기술 구현

2. 연구 배경

2.1. 요구사항 분석

서론에 제시한 연구 배경, 기존 문제점, 연구 목표를 바탕으로 제안하는 시스템의 요구사항을 분석하였다.

2.1.1. 시스템 요구사항

① 이기종 클라우드 연합학습 기능

1. 시스템은 이기종 클라우드(Public Cloud, Private Cloud) 플랫폼을 활용하여 연합학습을 수행할 수 있어야 한다.
2. 시스템은 다양한 클라우드(AWS, GCP) 서비스 제공자의 인증 정보를 관리할 수 있어야 한다.

② 연합학습 수행

1. 시스템은 연합학습 수행에 대한 로그를 출력할 수 있어야 한다.
2. 시스템은 연합학습 완료 후 최종 모델을 저장해야 한다.

③ 최적 모델 선정

1. 시스템은 모델의 성능 지표인 Accuracy, Recall, Precision, F1-score를 저장할 수 있어야 한다.
2. 시스템은 사용자가 정의한 성능 지표의 기준을 충족한 경우 평가 결과가 가장 좋은 모델을 선정할 수 있어야 한다.

④ 연합학습 집계자 배포 최적화

1. 시스템은 클라우드 가상머신 운용 비용과 연합학습 참여자와의 지연 시간을 종합적으로 고려하여 최적의 집계자 배포 위치를 추천해야 한다.

⑤ 클라우드 플랫폼 계층화를 통한 연합학습 구축

1. 시스템은 프라이빗 클라우드의 가상머신을 연합학습 참여자로 지정하여 민감한 데이터를 격리할 수 있어야 한다.
2. 시스템은 퍼블릭 클라우드의 가상머신을 연합학습 집계자로 지정하여 모델 집계 작업만을 수행하도록 해야한다.

⑥ 연합학습 참여자 관리

1. 시스템은 연합학습 참여자를 등록시킬 수 있어야 한다.
2. 시스템은 연합학습 참여자와의 연결 상태를 확인할 수 있어야 한다.
3. 시스템은 연합학습 참여자의 연결 상태가 정상인 참여자만 연합학습에 참여시킬 수 있어야 한다.

⑦ 연합학습 참여자 가상머신 상태 기반 동적 태스크 오케스트레이션

1. 시스템은 연합학습 참여자 내 가상머신의 자원(CPU, GPU, 메모리) 상태(사용량 및 가용량)를 실시간으로 모니터링 해야 한다.
2. 시스템은 연합학습 참여자 내 가상머신의 자원 상태에 따라 연합학습 작업을 동적으로 할당해야 한다.

⑧ 모델 정보 대시보드

1. 시스템은 연합학습 집계자의 모델 학습 진행 상황을 제공해야 한다.
2. 시스템은 연합학습 집계자의 모델 성능 지표를 실시간으로 표시해야 한다.

2.1.2. 사용자 요구사항

① 연합학습 참여자 선택

1. 사용자는 연합학습에 참여 중인 모든 연합학습 참여자의 지리적 위치를 확인할 수 있어야 한다.
2. 사용자는 웹 인터페이스를 통해 연합학습에 참여할 연합학습 참여자를 선택할 수 있어야 한다.

② 연합학습 집계자 설정 및 배포

1. 사용자는 웹 인터페이스를 통해 연합학습 집계자의 스펙(CPU, 메모리) 및 배포 리전을 선택할 수 있어야 한다.

③ 연합학습 집계자 모니터링

1. 사용자는 배포된 연합학습 집계자의 상태(실행 중, 중지됨, 오류 등)를 확인할 수 있어야 한다.
2. 사용자는 배포된 연합학습 집계자의 자원 상태(CPU, RAM 사용률 등)를 확인할 수 있어야 한다.
3. 사용자는 배포된 연합학습 집계자의 운영 비용을 확인할 수 있어야 한다.

④ 모델 선택 및 연합학습 실행

1. 사용자는 연합학습에 사용할 머신러닝 혹은 딥러닝 모델을 업로드할 수 있어야 한다.
2. 사용자는 모델의 하이퍼파라미터(학습률, 에포크 등)를 설정할 수 있어야 한다.
3. 사용자는 최적 모델 다운로드의 기준이 되는 모델 성능 지표에 대한 기준값을 설정할 수 있어야 한다.
4. 사용자는 선택한 모델과 설정으로 연합학습 작업을 시작할 수 있어야 한다.

⑤ 연합학습 모니터링

1. 사용자는 연합학습의 현재 라운드 진행 상황을 실시간으로 확인할 수 있어야 한다.
2. 사용자는 각 라운드별 연합학습 모델의 정확도, 손실 등의 성능 지표를 확인할 수 있어야 한다.
3. 사용자는 연합학습 결과로 생성된 최종 모델의 성능 평가 지표를 확인할 수 있어야 한다.
4. 사용자는 연합학습 과정에서 발생한 오류 및 경고 메시지를 확인할 수 있어야 한다.

⑥ 연합학습 참여자 가상머신 자원 모니터링

1. 사용자는 연합학습 참여자 내 가상머신의 자원 상태(CPU, GPU, RAM 사용률)를 실시간으로 확인할 수 있어야 한다.

⑦ 연합학습 모델 관리

1. 사용자는 완료된 연합학습 작업의 결과 모델 중 연합학습 생성 시 선택한 평가지표가 가장 높은 모델을 다운로드할 수 있어야 한다.
2. 사용자는 연합학습 수행 이력을 조회할 수 있어야 한다.

2.2. 선행기술 조사

2.2.1. AWS

AWS는 클라우드 기반 인프라와 서비스를 구축, 배포 및 관리할 수 있도록 설계된 클라우드 컴퓨팅 플랫폼이다. 컴퓨팅, 스토리지, 데이터베이스, 네트워킹, AI 등 다양한 서비스를 제공하며, 높은 가용성, 확장성, 비용 효율성을 갖추고 있다. AWS의 아키텍처는 그림 1과 같으며, 주요 구성요소 설명은 아래와 같다.

- ① EC2: EC2(Elastic Compute Cloud)는 가상 서버 인스턴스를 제공하는 서비스이다. 사용자는 이를 통해 컴퓨팅 리소스를 필요에 따라 유연하게 할당하고 사용할 수 있다.
- ② VPC: VPC(Virtual Private Cloud)는 사용자가 AWS 클라우드 내에 개인적인 가상 네트워크를 생성할 수 있게 해주는 서비스이다. 사용자는 자신의 IP 주소 범위를 선택하고, 서브넷등을 구성하여 물리적으로 격리된 것과 유사한 환경을 만들 수 있다.
- ③ IAM: IAM(Identity and Access Management)은 AWS 리소스에 대한 액세스를 보안적으로 관리할 수 있게 해주는 서비스이다. 사용자, 그룹, 역할을 만들어 특정 AWS 리소스에 대한 액세스 권한을 세밀하게 제어할 수 있다.
- ④ ECS: ECS(Elastic Container Service)는 AWS에서 제공하는 컨테이너 관리 서비스로, 컨테이너를 쉽게 배포, 관리 및 스케일링할 수 있게 해준다.

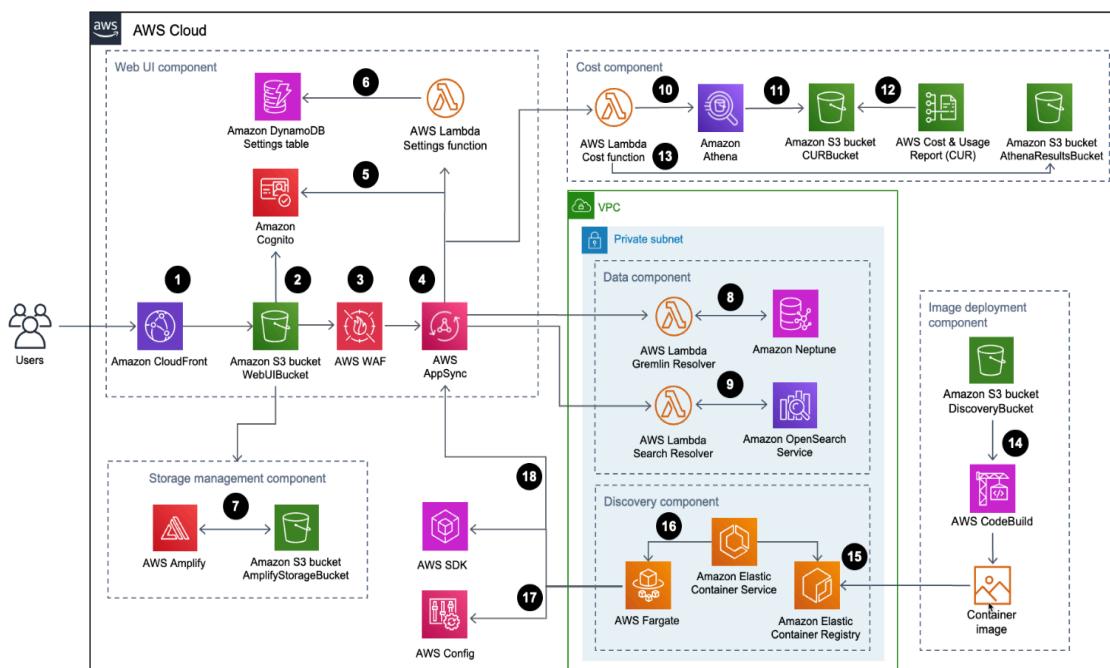


그림 1. AWS Architecture

2.2.2. GCP

GCP는 Google이 제공하는 클라우드 컴퓨팅 플랫폼으로, 컴퓨팅, 스토리지, 데이터베이스, 머신러닝, 네트워킹 등 다양한 클라우드 서비스를 제공한다. 그림 2는 GCP의 아키텍처이며 구성요소에 대한 설명은 다음과 같다.

- ① Compute Engine: Compute Engine은 GCP의 IaaS(Infrastructure as a Service) 솔루션으로, 가상 머신 인스턴스를 제공한다. 사용자는 다양한 머신 타입, 운영 체제, 디스크 옵션을 선택하여 워크로드에 맞는 가상 머신을 구성할 수 있다.
- ② Cloud Storage: Cloud Storage는 GCP의 객체 스토리지 서비스로, 대용량 데이터를 안전하게 저장하고 액세스할 수 있게 해준다. 데이터의 중요도와 액세스 빈도에 따라 다양한 스토리지 클래스를 제공하여 비용 효율적인 데이터 관리가 가능하다.
- ③ Cloud VPC: Cloud VPC는 GCP 내에서 사용자가 자신만의 가상 네트워크를 구성할 수 있게 해주는 서비스이다. IP 주소 범위, 방화벽 규칙 등의 네트워크 환경을 구축할 수 있다.
- ④ Cloud IAM: Cloud IAM은 GCP 리소스에 대한 접근 권한을 세밀하게 제어할 수 있는 서비스이다. 사용자, 그룹, 서비스 계정에 대한 권한을 관리할 수 있다.
- ⑤ GKE: GKE(Google Kubernetes Engine)는 컨테이너화된 애플리케이션을 배포, 관리, 스케일링할 수 있는 관리형 Kubernetes 서비스이다.

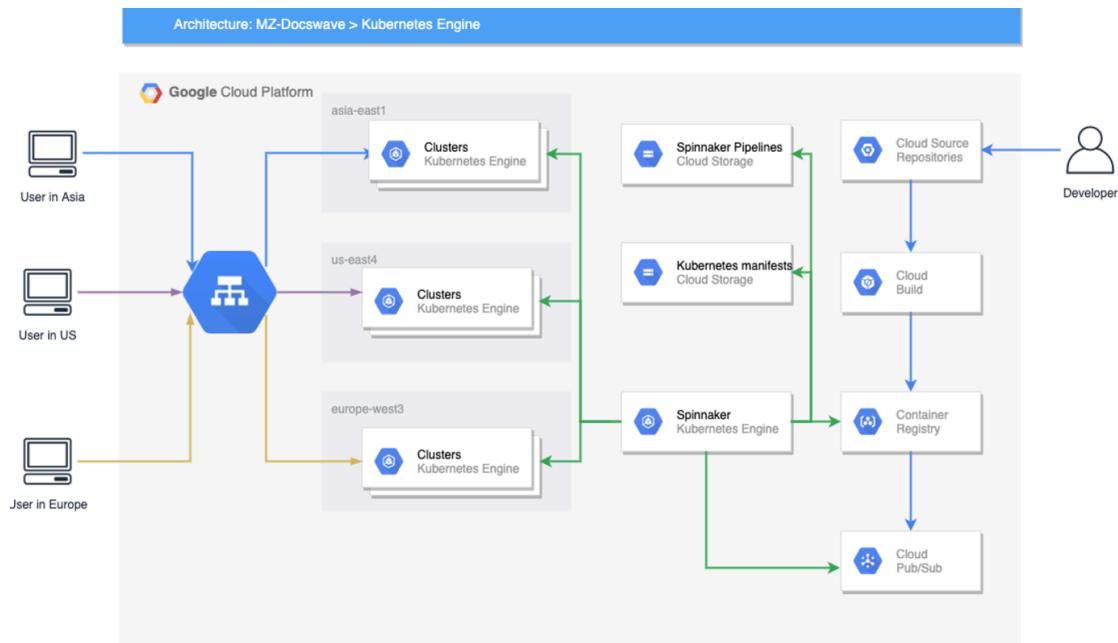


그림 2. GCP Architecture

2.2.3. OpenStack

OpenStack[13]은 기업과 조직이 프라이빗 또는 퍼블릭 클라우드 인프라를 생성, 배포 및 관리할 수 있도록 설계된 오픈 소스 클라우드 컴퓨팅 플랫폼이다. 그림 3은 OpenStack의 아키텍처이며 구성 요소의 설명은 다음과 같다.

- ① Nova: Nova는 OpenStack의 컴퓨팅 서비스로, 가상머신 인스턴스의 프로비저닝과 관리를 담당한다. Nova를 사용하면 사용자는 가상머신 인스턴스를 생성, 시작, 정지, 종료하는 등의 작업을 수행할 수 있으며, 리소스의 가용성과 성능을 고려하여 자동으로 배치 및 관리한다.
- ② Neutron: Neutron은 OpenStack의 네트워크 서비스로, 가상 네트워크 및 네트워크 리소스를 프로비저닝하고 관리하는 기능을 제공한다. 사용자는 가상 네트워크, 서브넷, 라우터, 로드 밸런서 등을 생성하여 가상머신 인스턴스 간의 통신과 네트워크 연결을 관리할 수 있다.
- ③ Cinder: Cinder는 Nova 서비스가 제공하는 인스턴스에 지속적으로 사용이 가능한 블록 스토리지 장치를 제공한다.
- ④ Keystone: Keystone은 OpenStack의 식별, 인증, 권한 부여 서비스이다. Keystone은 OpenStack 클라우드 환경에서 사용자, 서비스, 역할 등의 식별 정보를 관리하고 보안 인증 및 권한 부여를 처리한다.
- ⑤ Glance: Glance는 OpenStack의 이미지 서비스로, 가상머신 및 컨테이너 등 가상화 환경에서 사용되는 이미지 관리를 담당한다.
- ⑥ Horizon: Horizon은 OpenStack의 웹 기반 대시보드 인터페이스이다. Horizon은 OpenStack 클라우드 환경을 관리하고 모니터링할 수 있는 사용자 인터페이스를 제공한다.
- ⑦ Heat: Heat는 OpenStack의 오케스트레이션 서비스로, 클라우드 환경에서 인프라 리소스를 자동으로 프로비저닝하고 관리하는 기능을 제공한다.

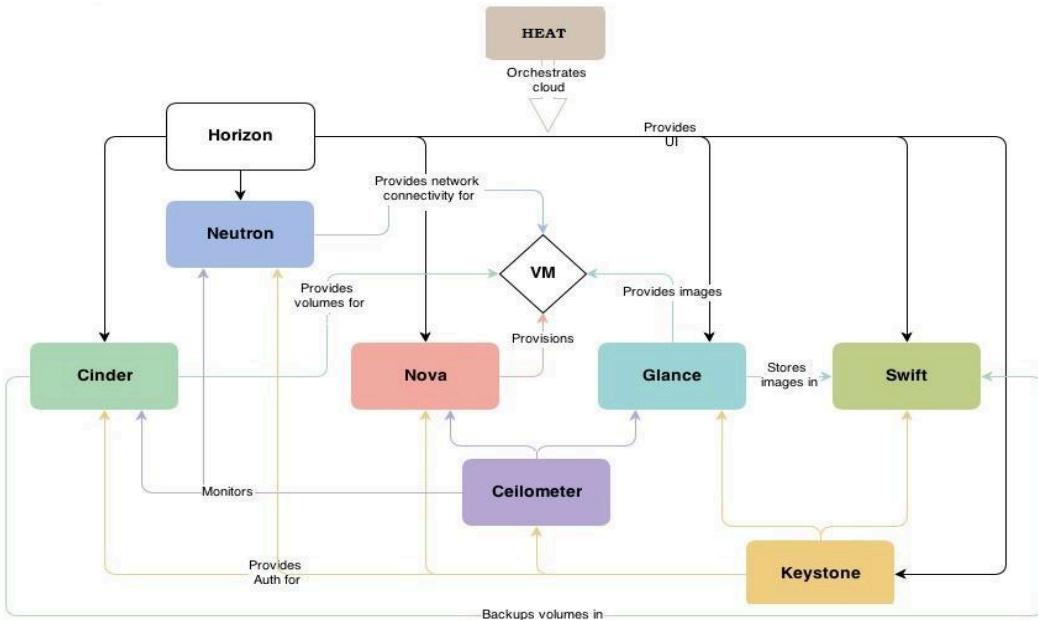


그림 3. OpenStack Architecture

2.2.4. Prometheus

Prometheus[14]는 오픈 소스 시스템 모니터링 및 알림 툴킷이다. 시계열 데이터베이스를 기반으로 메트릭을 수집하고 저장하며, 쿼리 언어(PromQL)를 통해 데이터를 분석하고 시각화할 수 있다. Prometheus는 서비스 디스커버리(Service Discovery) 기능을 통해 모니터링 대상을 자동으로 발견하고, 다양한 알림 규칙을 설정하여 시스템 이상을 감지할 수 있다.

- ① Prometheus server: Prometheus Server 자체는 각 모니터링 대상에서 Metrics(메트릭스)를 수집하거나, 수집한 Metrics에 대해 쿼리를 실행하여 Metrics 정보를 참조하거나, 자동으로 내부적으로 정기적으로 쿼리를 실행하여 경고를 관리하거나 하는 것을 담당한다.
- ② Service discovery: Service Discovery는 모니터링되는 정보를 자동으로 받아오는 구조이다. 이를 사용함으로써, 클라우드 플랫폼 또는 특정 소프트웨어 등의 해당 API를 주기적으로 호출하여, 거기에 등록된 인스턴스 정보를 수집한다.
- ③ Exporter: Exporter는 감시 에이전트이다. Exporter는 모니터링 대상에서 Metrics를 수집하여 Prometheus에 공개한다. Exporter를 둘으로써, 대상의 시스템으로부터 Metrics 정보를 받아올 수 있다. 받아온 Metrics 정보를 Prometheus가 읽을 수 있는 형태로 변환해 주는 기능을 제공한다.

- ④ Alert Manager: 설정된 Rule에 따른 알림 Notification을 담당한다. Slack 등에 연동해서 알림 시스템 구축이 가능하다.
- ⑤ PromQL: Prometheus Query Language의 약자이다. Metrics 라벨로 필터링 할 수 있는 기능을 제공한다. 예를 들어, 인스턴스의 이름, IP 주소, 클라우드 플랫폼의 리전 등을 라벨로 사용할 수 있다.

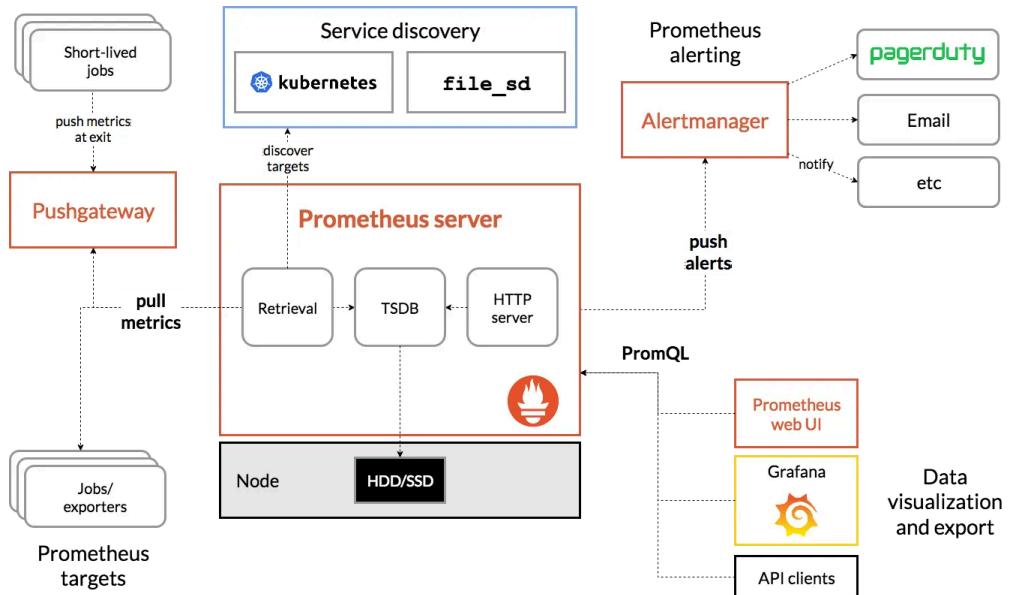


그림 4. Prometheus Architecture

2.2.5. Grafana

Grafana[15]는 메트릭 데이터 시각화 및 분석을 위한 대시보드 플랫폼이다. Prometheus, Elasticsearch 등 데이터 소스와 연동되어 시스템 메트릭을 실시간으로 모니터링할 수 있다.

- ① Distributor: Distributor는 Jaeger, OTLP 등 다양한 포맷의 트레이스 데이터를 수신하는 역할을 한다. 받은 트레이스는 traceID를 기준으로 해시하여 여러 Ingester로 분산시킨다.
- ② Ingester: Ingester는 실시간으로 들어오는 트레이스 데이터를 메모리에 저장하고, 일정 시간이나 조건에 따라 블록 단위로 압축해 오브젝트 스토리지에 업로드한다.
- ③ Query Frontend: Query Frontend는 사용자의 쿼리를 받아 병렬 처리를 위해 분산시키고, 캐싱을 통해 응답 속도를 높인다.
- ④ Querier: Querier는 쿼리 실행을 실제로 담당하는 컴포넌트이다. Ingester와 오브젝트 스토리지에서 데이터를 조회해 사용자에게 전달한다.
- ⑤ Compactor: Compactor는 오브젝트 스토리지에 저장된 블록들을 주기적으로 압축하고 병합하는 역할을 한다.
- ⑥ Metrics generator: Metrics Generator는 트레이스 데이터를 분석해 자연 시간, 오류율 같은 메트릭을 생성한다.

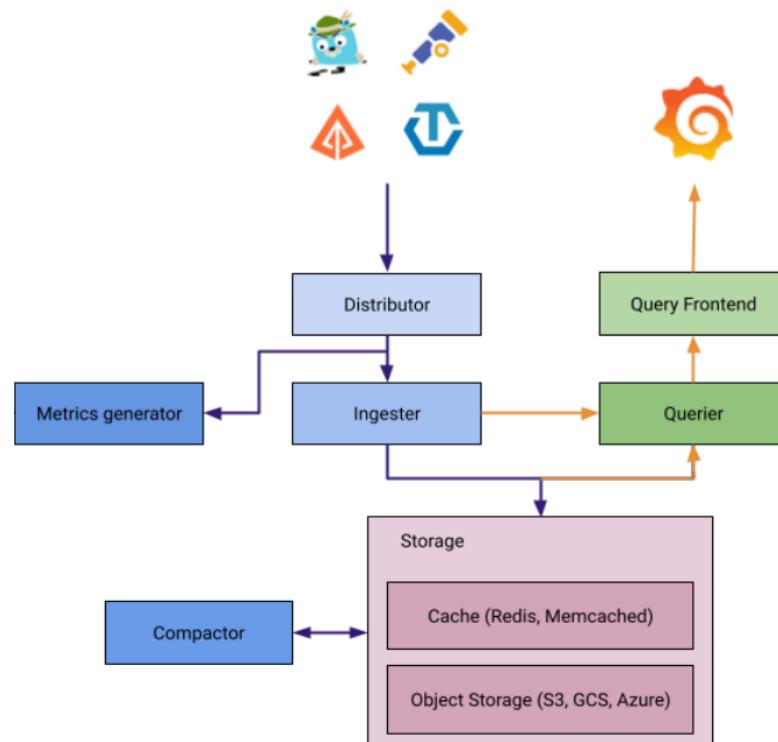


그림 5. Grafana Architecture

2.2.6. Terraform

Terraform[16]은 HashiCorp에서 개발한 Infrastructure as Code(IaC) 도구로, 클라우드 리소스를 코드로 관리 및 배포할 수 있게 해준다. 선언적인 구성 파일을 통해 인프라를 정의하고, 배포할 수 있다. Terraform은 AWS, GCP 등 다양한 클라우드 제공업체와 통합되어 있으며, 모듈화 및 재사용성을 통해 인프라 관리의 효율성을 높인다. 그림 5는 Terraform의 아키텍처이며, 아키텍처를 구성하는 컴포넌트의 설명은 아래와 같다.

- ① Terraform Core: Terraform CLI라고도 하는 Terraform Core는 Terraform 사용자를 위한 기본 인터페이스 역할을 한다.
- ② Terraform Provider: Terraform provider는 테라폼이 다양한 범주의 서비스 및 리소스와 통신을 가능하게끔 해주는 모듈이다. 각 Provider는 특정 서비스 내에서 Terraform이 관리할 수 있는 리소스에 대한 정의 및 Terraform configuration 내용을 해당 서비스에 대한 특정 API 호출로 변환하는 역할을 가지고 있다.
- ③ State File: state file은 테라폼의 기능 중 필수 요소로, 테라폼이 관리하는 리소스 정보뿐만 아니라, 리소스에 대한 현재 상태, 종속성 등을 저장하는 JSON 파일이다. Terraform은 State file을 활용하여 테라폼 설정이 적용될 때에 인프라 리소스에 반영되어야 할 변경사항을 결정한다.
- ④ Terraform Provisioners: 새로 생성된 리소스나 인스턴스에 대해 스크립트나 명령을 실행할 수 있도록 하는 기능이다. 여기에 사용되는 스크립트는 인프라 설정 및 구성, 소프트웨어 설치, 테스트 실행, 기타 필요한 작업 수행 등의 다양한 용도로 사용될 수 있다.

Terraform Architecture

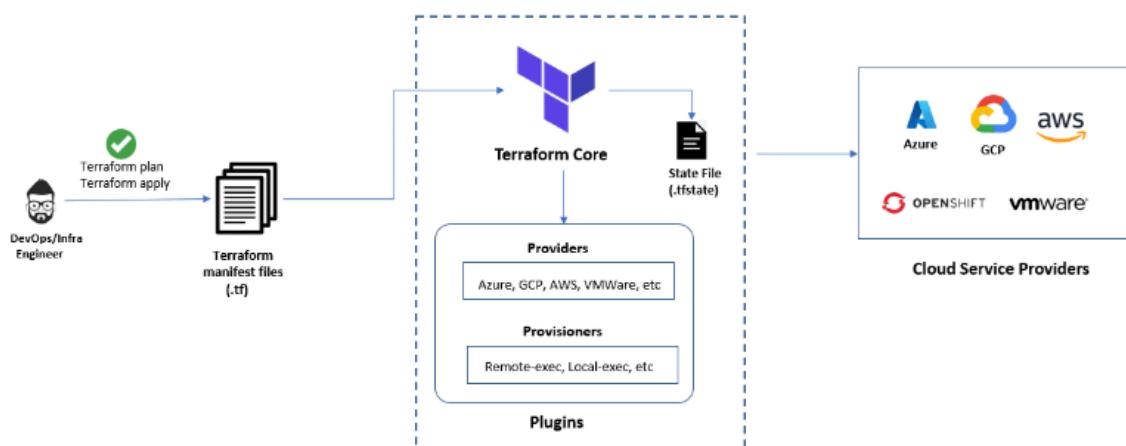


그림 6. Terraform 아키텍처

2.2.7. Flower

Flower[17]는 연합학습을 위한 오픈 소스 프레임워크이다. Client-Server 아키텍처를 기반으로 하며, 각 Client는 로컬 데이터로 모델을 훈련하고 모델 업데이트만 Server에 전송한다. Server는 이러한 업데이트를 집계하여 글로벌 모델을 개선한다.

그림 6은 Flower의 아키텍처를 나타낸 것으로, 서버 측 구성 요소와 클라이언트 측 구성 요소로 나누며, 각각 다음과 같이 구성된다.

① Flower Server

Flower Sever는 SuperLink와 ServerApp으로 구성된다. SuperLink는 클라이언트에게 연합학습 작업 지시를 전달하고 클라이언트로부터 작업 결과를 수신한다. ServerApp은 짧은 시간 실행되는 프로젝트 전용 서버 코드로, 클라이언트 선택, 클라이언트 설정, 결과 집계 전략 등 연합학습 서버 측 로직을 정의한다.

② Flower Client

Flower Client는 SuperNode와 ClientApp으로 구성된다. SuperNode는 Server의 SuperLink에 연결되어 연합학습 작업 요청을 수신하고, 로컬 모델 학습을 수행한 후 결과를 반환하는 프로세스이다. ClientApp은 짧은 시간 실행되는 프로젝트 전용 클라이언트 코드로, 로컬 모델 학습, 평가, 전처리 및 후처리 등 클라이언트 측 연합학습 로직을 정의한다.

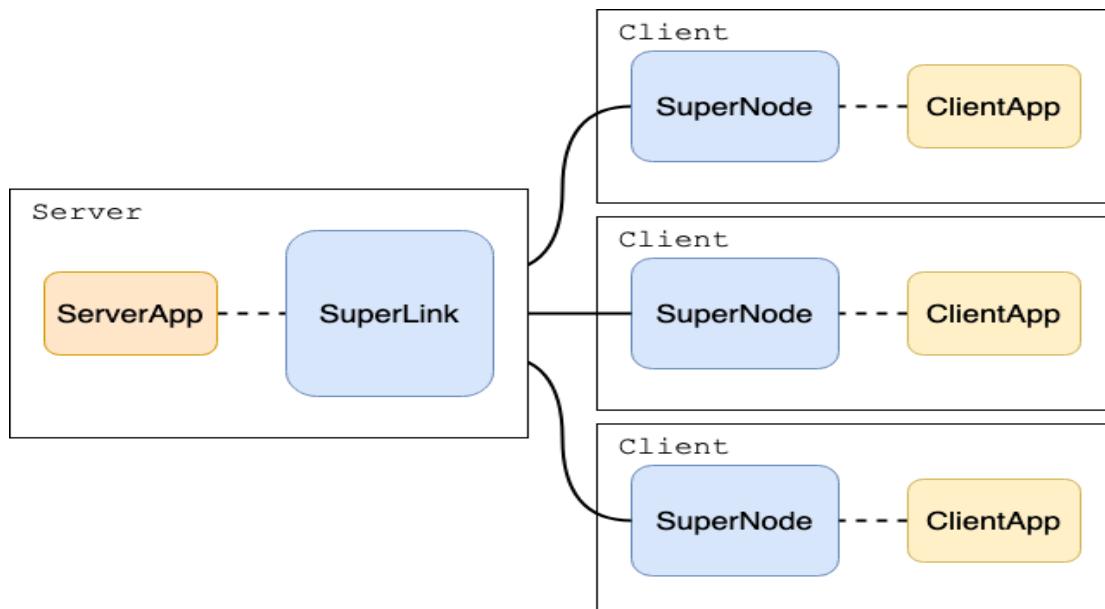


그림 7. Flower Architecture

2.2.8. MLflow

MLflow[18]는 머신러닝 라이프사이클 전반을 효과적으로 관리하기 위해 개발된 오픈 소스 플랫폼이다. MLflow는 4개의 모듈로 구성되어 있으며 연구 단계에서의 실험 추적부터 코드 및 환경의 패키징, 학습된 모델의 저장과 재사용, 다양한 환경으로의 배포에 이르기까지 머신러닝 프로젝트의 전체 과정을 지원한다.

- ① MLflow Tracking: MLflow Tracking은 머신러닝 실험의 매개변수, 코드 버전, 실행 메트릭, 아티팩트(모델, 로그 등)을 기록하고 관리할 수 있는 모듈이다. UI 대시보드를 제공하여 다양한 실험 결과를 비교할 수 있으며 다양한 하이퍼파라미터(learning rate, batch size, epoch 등) 조합에 따른 모델 평가 지표를 시각적으로 비교할 수 있다.
- ② MLflow Projects: MLflow Projects는 코드와 환경(conda 환경, Docker 컨테이너 등)을 재현 가능한 형태로 패키징하여, 어떤 환경에서도 동일한 방식으로 실행될 수 있도록 표준화된 형식을 제공한다. 이를 통해 연구 환경과 운영 환경 간의 불일치를 줄이고, 팀원 간 코드 공유와 모델 재사용성을 높일 수 있다.
- ③ MLflow Models: MLflow Models는 학습된 모델을 표준화된 형식으로 저장하고, 여러 가지 모델 서빙 환경에 손쉽게 배포할 수 있도록 지원한다. 모델은 특정 머신러닝 프레임워크(e.g. PyTorch[19], TensorFlow[20] 등)를 지원하는 형식으로 저장되며 REST API, Docker 컨테이너 등 다양한 배포 환경으로 서빙이 가능하다. 이를 통해 모델의 이동성과 호환성을 높여, 연구 단계에서 운영 단계로 손쉽게 이동시킬 수 있다.
- ④ Model Registry: Model Registry는 모델의 버전 관리, 스테이징, 프로덕션 배포를 중앙에서 관리할 수 있는 레지스트리 기능을 제공한다. 이를 통해 협업 환경에서 모델의 수명주기(개발 → 테스트 → 운영)를 체계적으로 관리할 수 있으며, 운영 환경에 어떤 버전의 모델이 배포되어 있는지 추적할 수 있다.

2.2.9. Docker

Docker[21]는 컨테이너 기반 가상화 플랫폼으로, 애플리케이션과 그 종속성을 경량화된 컨테이너로 패키징하여 어떤 환경에서든 일관되게 실행할 수 있도록 한다. Docker는 개발, 테스트, 배포 과정에서의 환경 차이로 인한 문제를 해결하고, 애플리케이션 배포와 확장을 용이하게 한다.

- ① Docker Engine: Docker의 핵심 런타임으로, 컨테이너의 생성, 실행, 관리를 담당한다. Linux 커널의 cgroups와 namespaces 기능을 활용하여 프로세스 격리를 제공한다.
- ② Docker Image: 컨테이너를 생성하기 위한 템플릿으로, 애플리케이션 코드, 런타임 라이브러리, 환경 변수 등을 포함한다.
- ③ Docker Container: 이미지를 기반으로 생성된 실행 가능한 인스턴스로, 독립된 파일시스템, 네트워크, 프로세스 공간을 갖는다.
- ④ Docker Registry: Docker Hub와 같은 중앙 저장소에서 이미지를 저장, 공유, 배포할 수 있다. 프라이빗 레지스트리를 구축하여 조직 내부에서 이미지를 관리할 수도 있다.
- ⑤ Docker Compose: 다중 컨테이너 애플리케이션을 정의하고 실행할 수 있는 도구로, YAML 파일을 통해 여러 서비스의 구성을 선언적으로 관리할 수 있다.

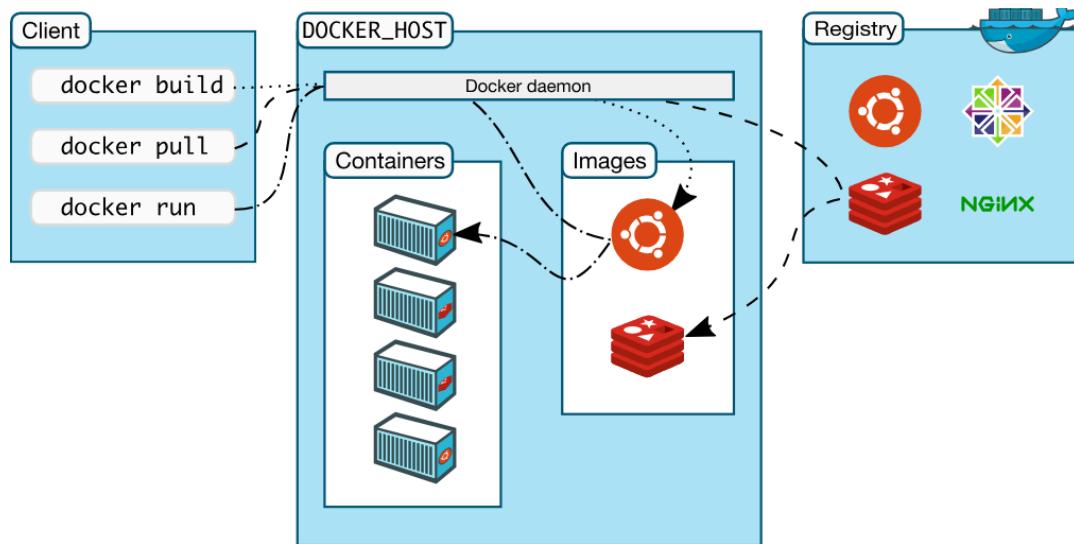


그림 8. Docker Architecture

3. 연구 내용

3.1. 설계 상세화

3.1.1. 유스케이스 다이어그램

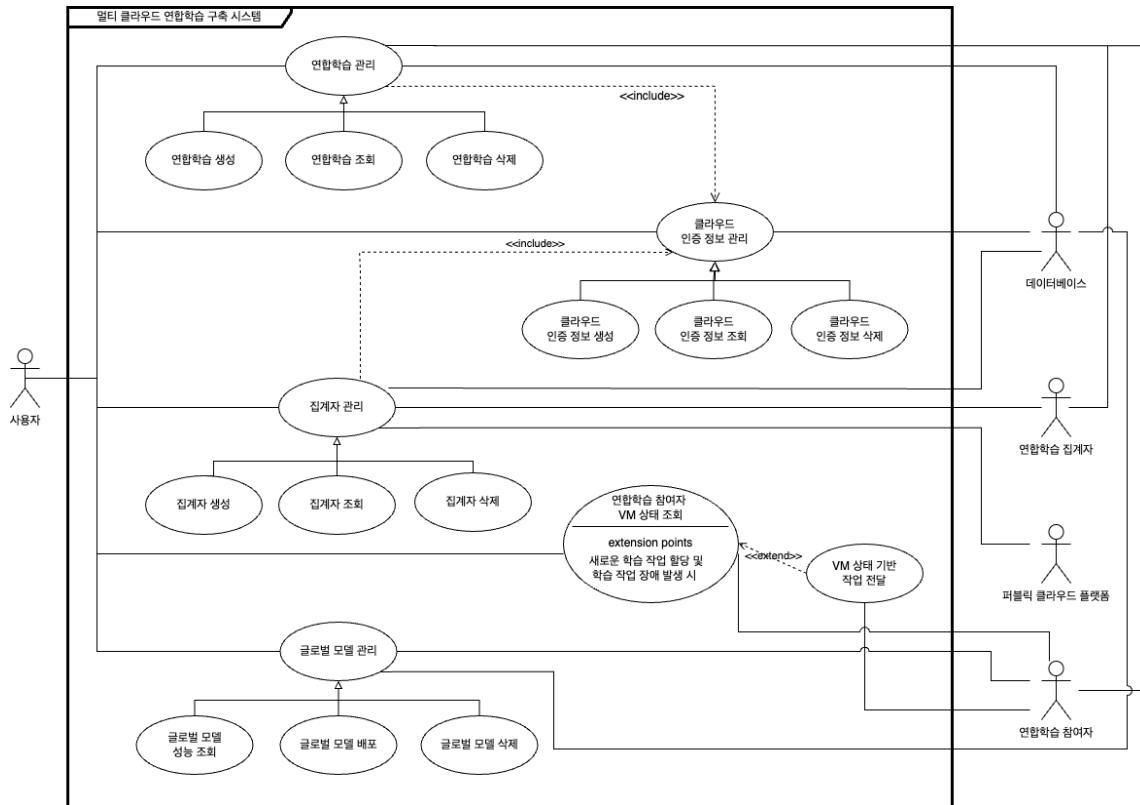


그림 9. 멀티 클라우드 기반 연합학습 시스템 유스케이스 다이어그램

표 1 Actor 목록

액터	설명
사용자	시스템의 기능에 접근하는 사용자
퍼블릭 클라우드 플랫폼	연합학습 집계자를 배포하는 퍼블릭 클라우드(AWS, GCP)
연합학습 집계자	퍼블릭 클라우드 상에 배포되어 연합학습 참여자로부터 모델 파라미터를 제공받아 글로벌 모델을 생성하는 시스템
연합학습 참여자	프라이빗 클라우드로 구성되어 개별 학습 데이터를 사용해 로컬 학습을 수행하고, 학습된 모델 파라미터를 연합학습 집계자로 전송하는 시스템
데이터베이스	사용자 정보, 연합학습 이력, 참여자 정보, 연합학습 결과 등을 저장하는 시스템

3.1.2. 유스케이스 명세

① 연합학습 생성

유스케이스명	연합학습 생성
개요	특정 모델을 학습하기 위해 연합학습 작업을 생성하는 기능
관련 액터	사용자, 데이터베이스, 연합학습 참여자, 연합학습 집계자
선행 조건	<ul style="list-style-type: none"> ● 사용자는 시스템에 로그인된 상태여야 한다. ● 클라우드 인증 정보가 등록된 상태여야 한다. ● 연합학습 참여자가 등록되어 있는 상태여야 한다.
기본 시나리오	<ol style="list-style-type: none"> 1. 사용자는 연합학습 관리 메뉴를 선택한다. 2. 시스템은 기존 연합학습 목록과 연합학습 생성 버튼을 출력한다. 3. 사용자는 연합학습 생성 버튼을 선택한다. 4. 시스템은 연합학습 설정 값 입력 폼(라운드 수, 모델 업로드, 참여자 선택, 모델 평가 기준)을 출력한다. 5. 사용자는 설정 값을 입력하고 생성 버튼을 선택한다. 6. 시스템은 라운드 수가 0보다 큰지 검증한다. 7. 시스템은 모델 파일이 유효한 확장자를 가지는지 검증한다. 8. 시스템은 참여자 수가 0보다 큰지 검증한다. 9. 시스템은 연합학습 집계자를 생성한다. 10. 시스템은 연합학습 집계자 및 연합학습 참여자 가상머신에 환경 설정 및 연합학습 수행 명령을 내린다. 11. 연합학습 참여자와 연합학습 집계자는 연합학습을 수행한다. 12. 연합학습 정보는 데이터베이스에 저장되어 모니터링 대시보드에 반영된다.
대안 시나리오	<p>A1: 연합학습 설정 값 유효성 검증 실패 시</p> <p>6.~9.에서 설정 값 유효성 검증에 실패했을 때</p> <ol style="list-style-type: none"> 1. 시스템은 연합학습 생성 실패 메시지와 오류 내용을 출력한다. 2. 기본 시나리오 2.로 돌아간다. <p>A2: 연합학습 집계자 생성 실패 시</p> <p>5.에서 연합학습 집계자 생성에 실패했을 때</p> <ol style="list-style-type: none"> 1. 시스템은 집계자 생성에 실패하였다는 메시지를 출력한다. 2. 기본 시나리오 2.로 돌아간다.
후행 조건	<ul style="list-style-type: none"> ● 연합학습 참여자와 연합학습 집계자가 연계되어 연합학습을 수행한다. ● 생성한 연합학습을 연합학습 모니터링 대시보드에서 확인할 수 있다.

② 연합학습 조회

유스케이스명	연합학습 조회
개요	실행 중이거나 실행되었던 연합학습의 상태 및 결과를 조회하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.● 실행 중이거나 실행되었던 연합학습이 존재해야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 연합학습 관리 메뉴를 선택한다.2. 시스템은 기존 연합학습 목록과 연합학습 생성 버튼을 출력한다.3. 사용자는 조회할 연합학습을 선택한다.4. 시스템은 데이터베이스에서 선택한 연합학습에 대한 정보(모델 정보, 평가 결과, 걸린 시간, 참여자 정보)를 불러온다.5. 시스템은 사용자가 선택한 연합학습의 정보(모델 정보, 평가 결과, 걸린 시간, 참여자 정보)를 출력한다.
후행 조건	<ul style="list-style-type: none">● 없음

③ 연합학습 삭제

유스케이스명	연합학습 삭제
개요	실행되었던 연합학습의 기록을 삭제하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none"> ● 사용자는 시스템에 로그인된 상태여야 한다. ● 실행되었던 연합학습이 존재해야 한다.
기본 시나리오	<ol style="list-style-type: none"> 1. 사용자는 연합학습 관리 메뉴를 선택한다. 2. 시스템은 기존 연합학습 목록과 연합학습 생성 버튼을 출력한다. 3. 사용자는 삭제할 연합학습 이력을 선택한다. 4. 시스템은 사용자가 선택한 연합학습의 정보 및 삭제 버튼을 출력한다. 5. 사용자는 삭제 버튼을 선택한다. 6. 시스템은 삭제 재확인 메시지 및 버튼을 출력한다. 7. 사용자는 삭제 재확인 버튼을 선택한다. 8. 시스템은 데이터베이스에서 해당 연합학습 정보를 삭제한다. 9. 시스템은 연합학습 이력이 성공적으로 삭제되었다는 메시지를 출력한다.
대안 시나리오	<p>A1: 사용자가 삭제 취소 시</p> <p>6.에서 사용자가 연합학습 삭제를 취소한 경우</p> <ol style="list-style-type: none"> 1. 기본 시나리오 4.로 돌아간다. <p>A2:연합학습 이력 삭제 실패 시</p> <p>8.에서 연합학습 이력 삭제에 실패했을 때</p> <ol style="list-style-type: none"> 1. 시스템은 연합학습 삭제에 실패하였다는 메시지를 출력한다. 2. 기본 시나리오 2.로 돌아간다.
후행 조건	<ul style="list-style-type: none"> ● 데이터베이스에서 연합학습 이력이 삭제된다.

④ 클라우드 인증 정보 등록

유스케이스명	클라우드 인증 정보 등록
개요	다양한 클라우드 서비스 제공자(AWS, GCP)의 인증 정보를 등록하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 클라우드 인증 정보 관리 메뉴를 선택한다.2. 시스템은 등록된 인증 정보 목록과 새 인증 정보 등록 버튼을 출력한다.3. 사용자는 새 인증 정보 등록 버튼을 클릭한다.4. 시스템은 등록 가능한 클라우드 종류 목록을 출력한다.5. 사용자는 등록할 인증 정보의 클라우드 플랫폼을 선택한다.6. 사용자는 선택한 클라우드 인증 정보 등록에 필요한 정보를 입력하고 저장 버튼을 클릭한다.7. 시스템은 입력된 인증 정보를 기반으로 API 호출을 통해 유효성 검증을 수행한다.8. 시스템은 인증 정보를 암호화하여 데이터베이스에 저장한다.9. 시스템은 사용자에게 등록 성공 메시지를 출력한다.
대안 시나리오	<p>A1: 입력된 인증 정보 유효성 검증 실패 시</p> <ol style="list-style-type: none">7. 에서 입력된 인증 정보 유효성 검증에 실패했을 때1. 시스템은 잘못된 인증 정보라는 메시지를 출력한다.2. 기본 시나리오 2.로 돌아간다.
후행 조건	<ul style="list-style-type: none">● 데이터베이스에서 클라우드 인증 정보가 삭제된다.

⑤ 클라우드 인증 정보 조회

유스케이스명	클라우드 인증 정보 조회
개요	다양한 클라우드 서비스 제공자(AWS, Azure, GCP 등)의 인증 정보를 조회하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.● 데이터베이스에 등록된 인증 정보가 존재해야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 클라우드 인증 정보 관리 메뉴를 선택한다.2. 시스템은 데이터베이스에 등록된 클라우드 인증 정보 목록을 출력한다.3. 시스템은 데이터베이스에 저장된 클라우드 인증 정보를 조회한다.4. 사용자는 특정 클라우드 인증 정보 목록을 선택한다.5. 시스템은 해당 클라우드 인증 정보를 데이터베이스에서 조회한다.6. 시스템은 사용자가 선택한 클라우드 인증 정보에 대해 클라우드 종류, 인증 정보 이름, 등록 시기를 출력한다.
후행 조건	<ul style="list-style-type: none">● 없음

⑥ 클라우드 인증 정보 삭제

유스케이스명	클라우드 인증 정보 삭제
개요	다양한 클라우드 서비스 제공자(AWS, Azure, GCP 등)의 인증 정보를 삭제하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.● 데이터베이스에 등록된 클라우드 인증 정보가 존재해야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 클라우드 인증 정보 관리 메뉴를 선택한다.2. 시스템은 데이터베이스에 등록된 클라우드 인증 정보 목록을 출력한다.3. 사용자는 특정 클라우드 인증 정보를 클릭한다.4. 시스템은 특정 클라우드에 인증 정보 내용과 삭제 버튼을 출력한다.5. 사용자는 삭제하고 싶은 클라우드 인증 정보의 삭제 버튼을 클릭한다.6. 시스템은 삭제 재확인 메시지와 버튼을 출력한다.7. 사용자는 삭제 재확인 버튼을 클릭한다.8. 시스템은 해당 인증 정보를 데이터베이스에서 삭제한다.9. 시스템은 성공적으로 삭제되었다는 메시지를 출력한다.
대안 시나리오	<p>A1: 사용자가 삭제 취소 시</p> <ol style="list-style-type: none">6. 에서 사용자가 클라우드 인증 정보 삭제를 취소한 경우1. 기본 시나리오 4.로 돌아간다. <p>A2: 클라우드 인증 정보 삭제 실패 시</p> <ol style="list-style-type: none">8. 에서 클라우드 인증 정보 삭제에 실패했을 때1. 시스템은 삭제에 실패하였다는 메시지와 오류 메시지를 출력한다.2. 기본 시나리오 2.로 돌아간다.
후행 조건	<ul style="list-style-type: none">● 클라우드 인증 정보가 데이터베이스에서 삭제된다.

⑦ 연합학습 집계자 생성

유스케이스명	연합학습 집계자 생성
개요	새로운 연합학습 집계자를 생성하여 연합학습에 배치하는 기능
관련 액터	사용자, 데이터베이스, 퍼블릭 클라우드 플랫폼, 연합학습 집계자
선행 조건	<ul style="list-style-type: none">사용자는 시스템에 로그인된 상태여야 한다.시스템에 클라우드 인증 정보가 등록되어있어야 한다.사용자가 연합학습 생성을 위해 참여자를 선택한 상태여야 한다.
기본 시나리오	<ol style="list-style-type: none">시스템은 연합학습 집계자 생성 버튼을 출력한다.시스템은 지연 시간 및 비용을 최적화하고 집계자를 배포할 리전을 결정한다.시스템은 최적화한 지연 시간 및 비용, 결정된 리전을 출력한다.사용자는 배포할 리전을 확인하고 연합학습 집계자 생성 버튼을 선택한다.시스템은 퍼블릭 클라우드 플랫폼에 집계자 생성 요청을 보낸다.퍼블릭 클라우드 플랫폼은 결정된 리전에 연합학습 집계자를 생성한다.시스템은 연합학습 집계자 정보를 데이터베이스에 저장한다.시스템은 연합학습 집계자 생성 완료 메시지를 출력한다.
대안 시나리오	A1: 퍼블릭 클라우드 플랫폼에서 연합학습 집계자 생성 실패 시 6. 에서 연합학습 집계자 생성에 실패했을 때 <ol style="list-style-type: none">시스템은 집계자 생성에 실패하였다는 메시지와 오류 메시지를 출력한다.기본 시나리오 1.로 돌아간다.
후행 조건	<ul style="list-style-type: none">비용 및 지연 시간이 최적화된 집계자가 생성된다.생성된 연합학습 집계자가 연합학습에 배치된다.

⑧ 연합학습 집계자 조회

유스케이스명	연합학습 집계자 조회
개요	연합학습에 배치된 집계자를 조회하는 기능
관련 액터	사용자, 데이터베이스, 퍼블릭 클라우드 플랫폼, 연합학습 집계자
선행조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.● 시스템에 클라우드 인증 정보가 등록되어있어야 한다.● 시스템은 집계자를 생성한 상태여야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 집계자 관리 메뉴를 선택한다.2. 시스템은 등록된 집계자 목록과 집계자 삭제 버튼을 출력한다.3. 사용자는 조회할 집계자를 선택한다.4. 시스템은 선택한 집계자의 정보(관련 연합학습, 하드웨어 정보, 연결된 클라이언트)를 데이터베이스에서 조회한다.5. 데이터베이스는 집계자의 정보를 시스템에 전달한다.6. 시스템은 퍼블릭 클라우드 플랫폼에 집계자의 정보(고유 식별자, 네트워크 구성, 비용 정보 등)를 요청한다.7. 퍼블릭 클라우드 플랫폼은 집계자의 정보를 시스템에 전달한다.8. 시스템은 집계자의 자원 사용률 및 정보를 출력한다.
후행 조건	<ul style="list-style-type: none">● 없음

⑨ 집계자 삭제

유스케이스명	연합학습 집계자 삭제
개요	연합학습에 배치된 집계자를 삭제하는 기능
관련 액터	사용자, 데이터베이스, 퍼블릭 클라우드 플랫폼, 연합학습 집계자
선행 조건	<ul style="list-style-type: none"> 사용자는 시스템에 로그인된 상태여야 한다. 시스템에 클라우드 인증 정보가 등록되어있어야 한다. 해당 집계자를 사용한 연합학습이 종료된 상태여야한다.
기본 시나리오	<ol style="list-style-type: none"> 사용자는 집계자 관리 메뉴를 선택한다. 시스템은 등록된 집계자 목록과 집계자 삭제 버튼을 출력한다. 사용자는 삭제하고자 하는 집계자의 삭제 버튼을 선택한다. 시스템은 사용자에게 삭제 확인 메시지를 출력한다. 사용자가 삭제를 확정한다. 시스템은 퍼블릭 클라우드 플랫폼에 해당 집계자 삭제 요청을 보낸다. 퍼블릭 클라우드 플랫폼은 집계자를 삭제한다. 퍼블릭 클라우드 플랫폼은 집계자에 활용된 네트워크 및 키페어 구성 요소를 삭제한다. 시스템은 데이터베이스에서 해당 집계자 정보를 삭제한다. 시스템은 집계자가 성공적으로 삭제되었다는 메시지를 출력한다.
대안 시나리오	<p>A1: 사용자가 집계자 삭제 요청 취소 시</p> <ol style="list-style-type: none"> 에서 사용자가 집계자 삭제 요청을 취소했을 시 기본 시나리오 2.로 돌아간다. <p>A2: 집계자 삭제 실패 시</p> <ol style="list-style-type: none"> 에서 집계자 삭제에 실패했을 시 시스템은 삭제에 실패하였다는 오류 메시지를 출력한다. 기본 시나리오 2.로 돌아간다.
후행 조건	<ul style="list-style-type: none"> 집계자가 데이터베이스에서 삭제된다.

⑩ 연합학습 참여자 동적 태스크 오케스트레이션

유스케이스명	연합학습 참여자 동적 태스크 오케스트레이션
개요	연합학습 참여자 가상머신의 상태를 기반으로 적절한 가상머신에 학습 작업 할당 및 이관하는 기능
관련 액터	연합학습 참여자
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.● 연합학습이 생성된 상태여야 한다.● 연합학습 집계자와 연합학습 참여자의 가상머신이 생성되어 있어야 한다.● 각 연합학습 참여자의 가상머신 상태 정보가 모니터링되고 있어야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 시스템은 연합학습 참여자에게 연합학습 작업을 할당한다.2. 시스템은 각 연합학습 참여자 가상머신의 리소스 상태(CPU, RAM 사용량, Disk)를 확인한다.3. 시스템은 가상머신의 리소스가 기준치 이상/이하인지 평가하여 작업을 할당할 후보군을 도출한다.4. 시스템은 도출한 후보군을 대상으로 원형 큐 알고리즘을 이용해 가상머신을 선택한다.5. 시스템은 해당 가상머신에 연합학습 작업을 전달한다.6. 작업이 완료되면 시스템은 결과를 수집하고 다음 작업 할당을 준비한다.
대안 시나리오	<p>A1: 가상머신에 장애가 발생한 경우</p> <p>5. 에서 가상머신에 장애가 발생했을 경우</p> <ol style="list-style-type: none">1. 시스템은 해당 가상머신에 할당된 작업을 중단한다.2. 시스템은 다른 가상머신에 해당 작업을 재할당한다.3. 기본 시나리오 1.로 돌아간다.
후행 조건	<ul style="list-style-type: none">● 가상머신의 자원 상태에 최적화된 작업 할당이 이루어진다.● 장애 발생 시 작업이 다른 가상머신을 이용해 연합학습이 재시도된다.

⑪ 글로벌 모델 성능 조회

유스케이스명	글로벌 모델 성능 조회
개요	연합학습으로 생성된 글로벌 모델의 성능 지표를 조회하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none">사용자는 시스템에 로그인된 상태여야 한다.학습이 완료된 글로벌 모델이 최소 1개 이상 존재하여야 한다.
기본 시나리오	<ol style="list-style-type: none">사용자는 글로벌 모델 관리 메뉴를 선택한다.시스템은 글로벌 모델 목록을 출력한다.사용자는 성능을 조회하고자 하는 글로벌 모델을 선택한다.시스템은 데이터베이스에서 선택한 글로벌 모델에 대한 정보를 조회한다.시스템은 선택된 모델의 성능 지표(Accuracy, Recall, Precision, F1-score)를 꺾은선 그래프 형태로 출력한다.시스템은 해당 모델 학습을 수행한 연합학습 정보와 참여자 정보자 정보를 출력한다.
후행 조건	<ul style="list-style-type: none">없음

⑫ 글로벌 모델 배포

유스케이스명	글로벌 모델 배포
개요	연합학습으로 생성된 글로벌 모델을 연합학습 참여자에 배포하는 기능
관련 액터	사용자, 데이터베이스, 연합학습 참여자
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.● 연합학습이 완료되어 배포 가능한 글로벌 모델이 생성되어 있어야 한다.● 클라우드 인증 정보가 등록되어 있어야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 글로벌 모델 관리 메뉴를 선택한다.2. 시스템은 글로벌 모델 목록을 출력한다.3. 사용자는 배포하고자 하는 글로벌 모델을 선택한다.4. 시스템은 선택된 모델의 정보와 함께 배포 버튼을 출력한다.5. 사용자는 배포 버튼을 선택한다.6. 시스템은 선택된 글로벌 모델을 해당 모델의 연합학습에 참여한 참여자에게 배포한다.7. 시스템은 배포 상태와 진행 상황을 출력한다.8. 배포가 완료되면 시스템은 배포 성공 메시지를 출력한다.
대안 시나리오	<p>A1: 글로벌 모델 배포 실패 시</p> <ol style="list-style-type: none">8. 에서 글로벌 모델 배포에 실패했을 때 <ol style="list-style-type: none">1. 시스템은 배포 실패 메시지와 에러 메시지를 출력한다.2. 시스템은 글로벌 모델을 배포받지 못한 연합학습 참여자 목록을 출력한다.3. 기본 시나리오 4.로 돌아간다.
후행 조건	<ul style="list-style-type: none">● 글로벌 모델이 연합학습 참여자에게 배포된다.● 배포 정보가 데이터베이스에 저장된다.

⑯ 글로벌 모델 삭제

유스케이스명	글로벌 모델 삭제
개요	학습이 완료된 글로벌 모델을 데이터베이스에서 삭제하는 기능
관련 액터	사용자, 데이터베이스
선행 조건	<ul style="list-style-type: none">● 사용자는 시스템에 로그인된 상태여야 한다.
기본 시나리오	<ol style="list-style-type: none">1. 사용자는 글로벌 모델 관리 메뉴를 선택한다.2. 시스템은 글로벌 모델 목록을 출력한다.3. 사용자는 삭제하고자 하는 글로벌 모델을 선택한다.4. 시스템은 선택된 모델의 정보와 함께 삭제 버튼을 출력한다.5. 사용자는 삭제 버튼을 선택한다.6. 시스템은 삭제 재확인 메시지 및 버튼을 출력한다.7. 사용자는 삭제 재확인 버튼을 선택한다.8. 시스템은 데이터베이스에서 해당 모델을 삭제한다.9. 시스템은 삭제 완료 메시지를 출력한다.
대안 시나리오	<p>A1: 사용자가 삭제 취소 시</p> <ol style="list-style-type: none">7. 에서 사용자가 글로벌 모델 삭제를 취소한 경우 <p>A2: 모델 삭제 실패 시</p> <ol style="list-style-type: none">8. 에서 모델 삭제에 실패한 경우 <ol style="list-style-type: none">1. 시스템은 삭제 실패 메시지와 함께 실패 원인을 출력한다.2. 기본 시나리오 4.로 돌아간다.
후행 조건	<ul style="list-style-type: none">● 데이터베이스에서 해당 글로벌 모델에 대한 정보가 삭제된다.

3.1.3. 클래스 다이어그램

① 사용자 정보 관리

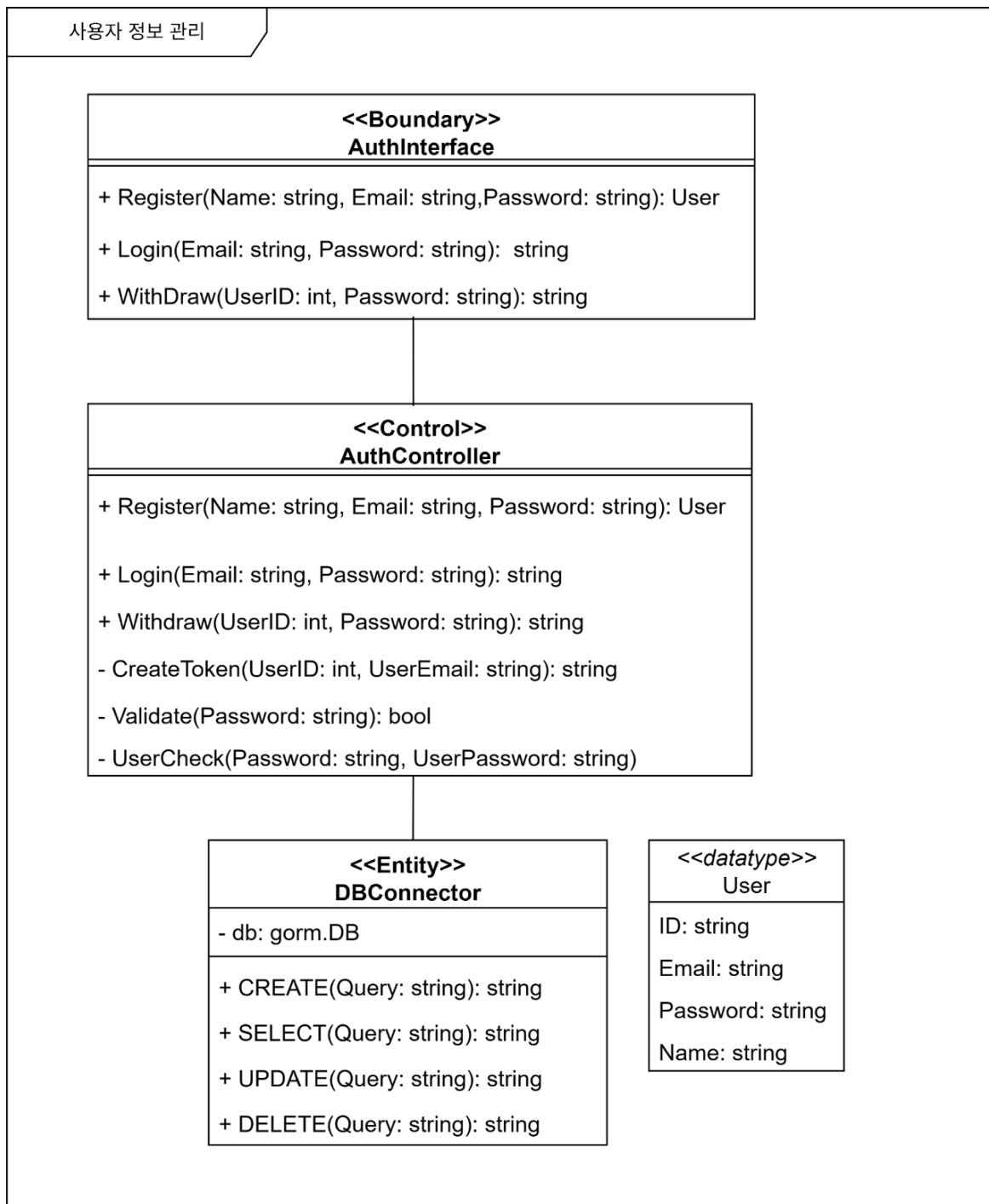


그림 10. 사용자 정보 관리 클래스 다이어그램

② 클라우드 인증 정보 관리

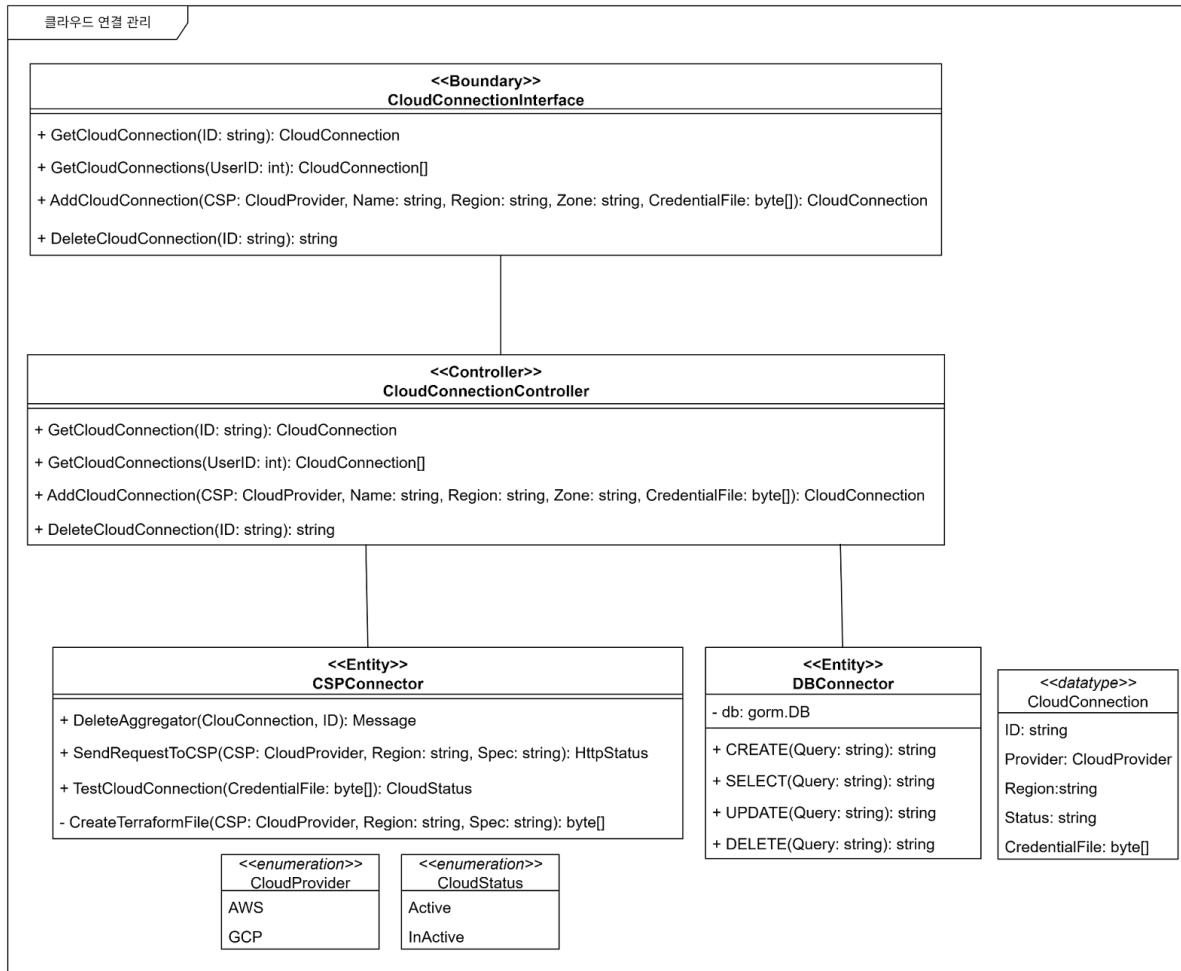


그림 11. 클라우드 인증 정보 관리 클래스 다이어그램

③ 연합학습 집계자 관리

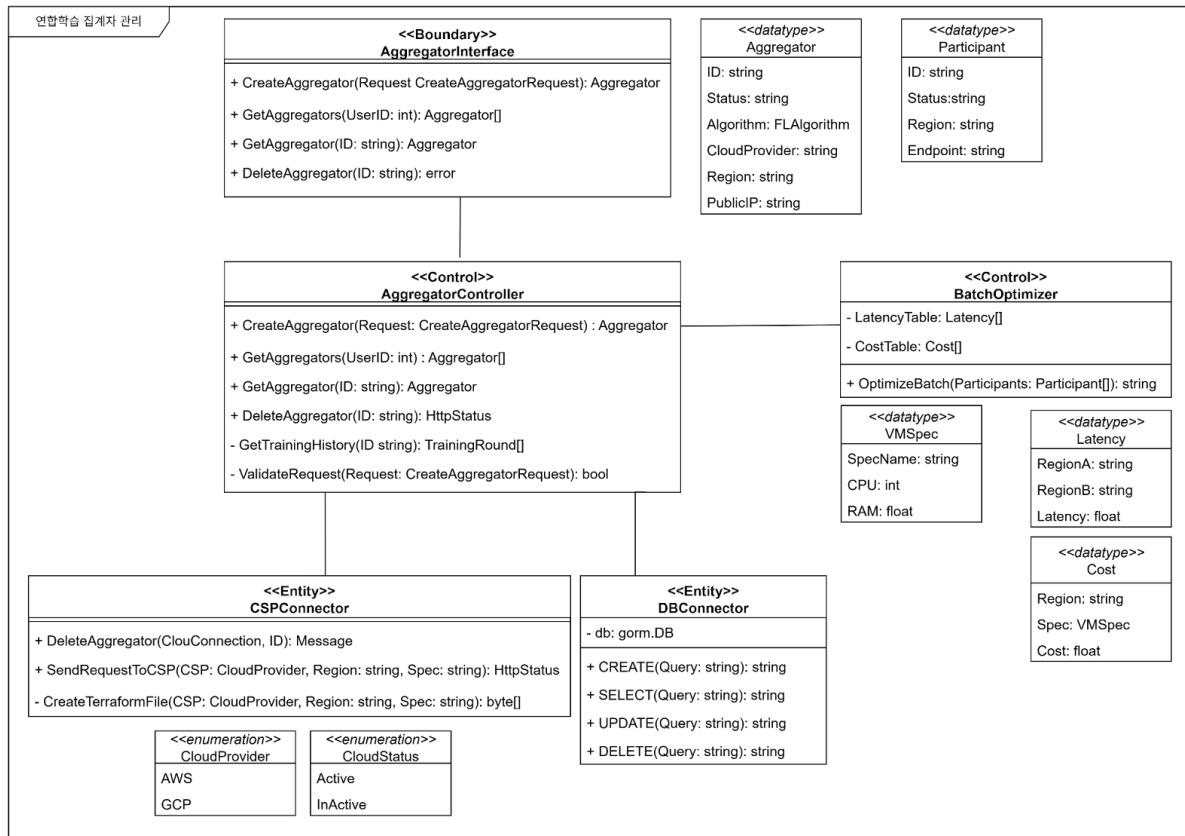


그림 12. 연합학습 집계자 관리 클래스 다이어그램

④ 연합학습 참여자 관리

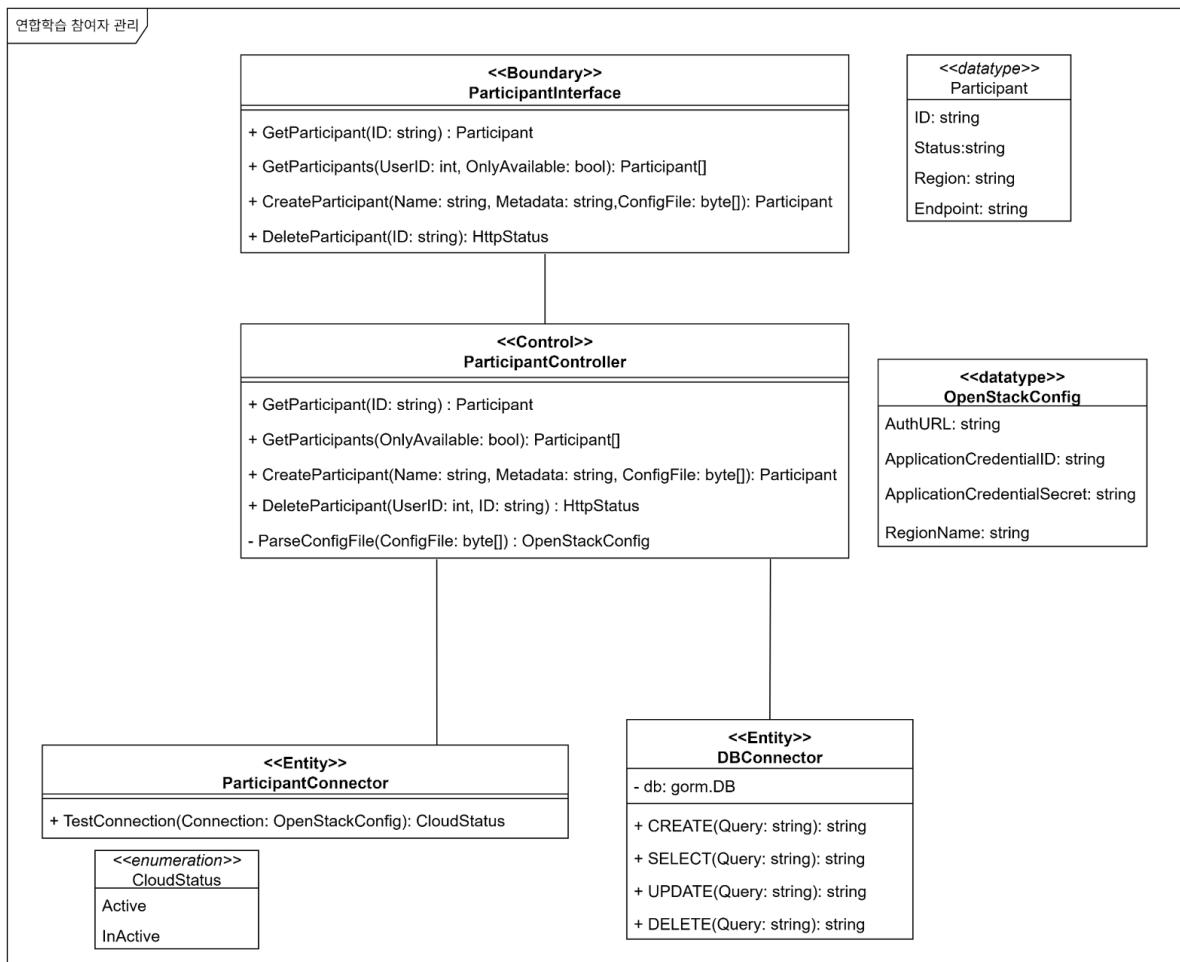


그림 13. 연합학습 참여자 관리 클래스 다이어그램

⑤ 연합학습 관리

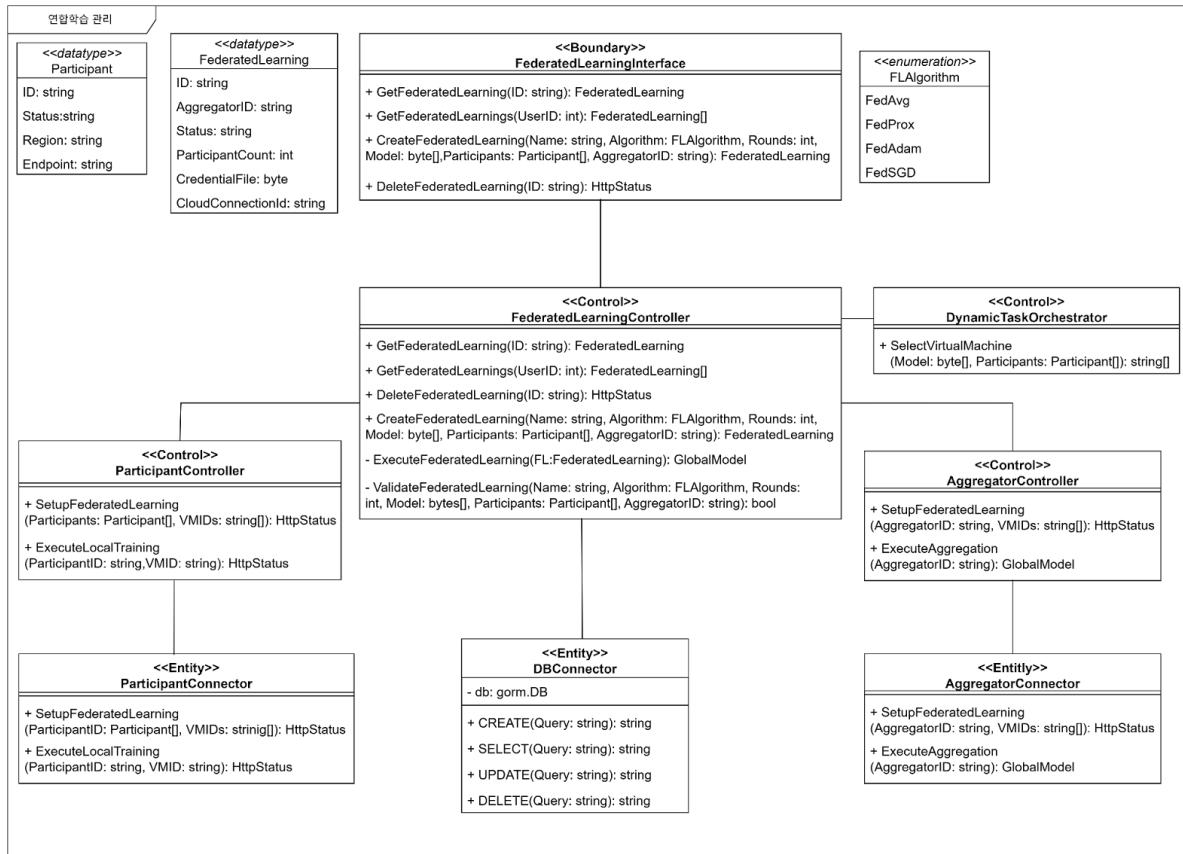


그림 14. 연합학습 관리 클래스 다이어그램

⑥ 글로벌 모델 관리

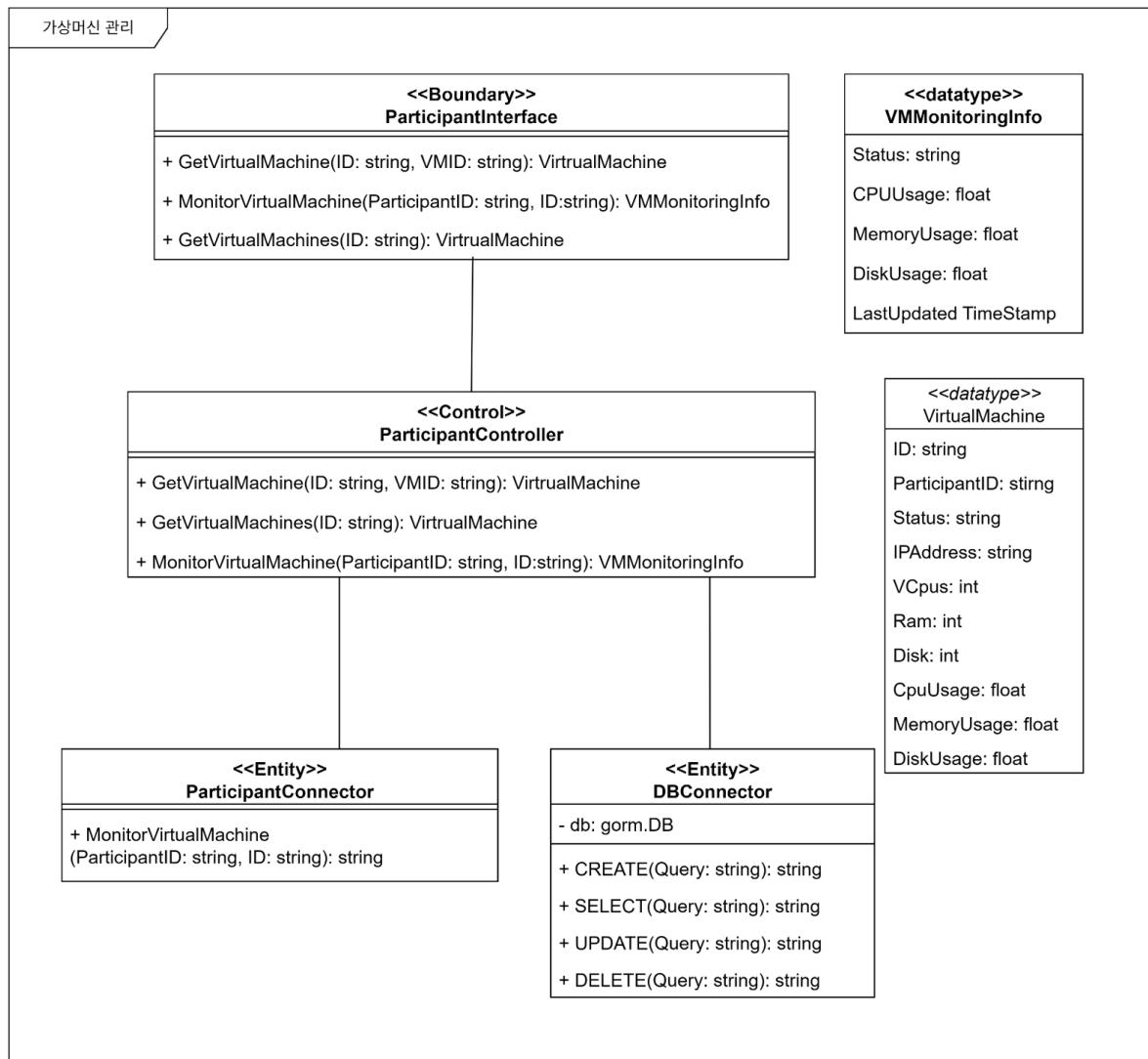


그림 15. 글로벌 모델 관리 클래스 다이어그램

⑦ 가상머신 관리 클래스 다이어그램

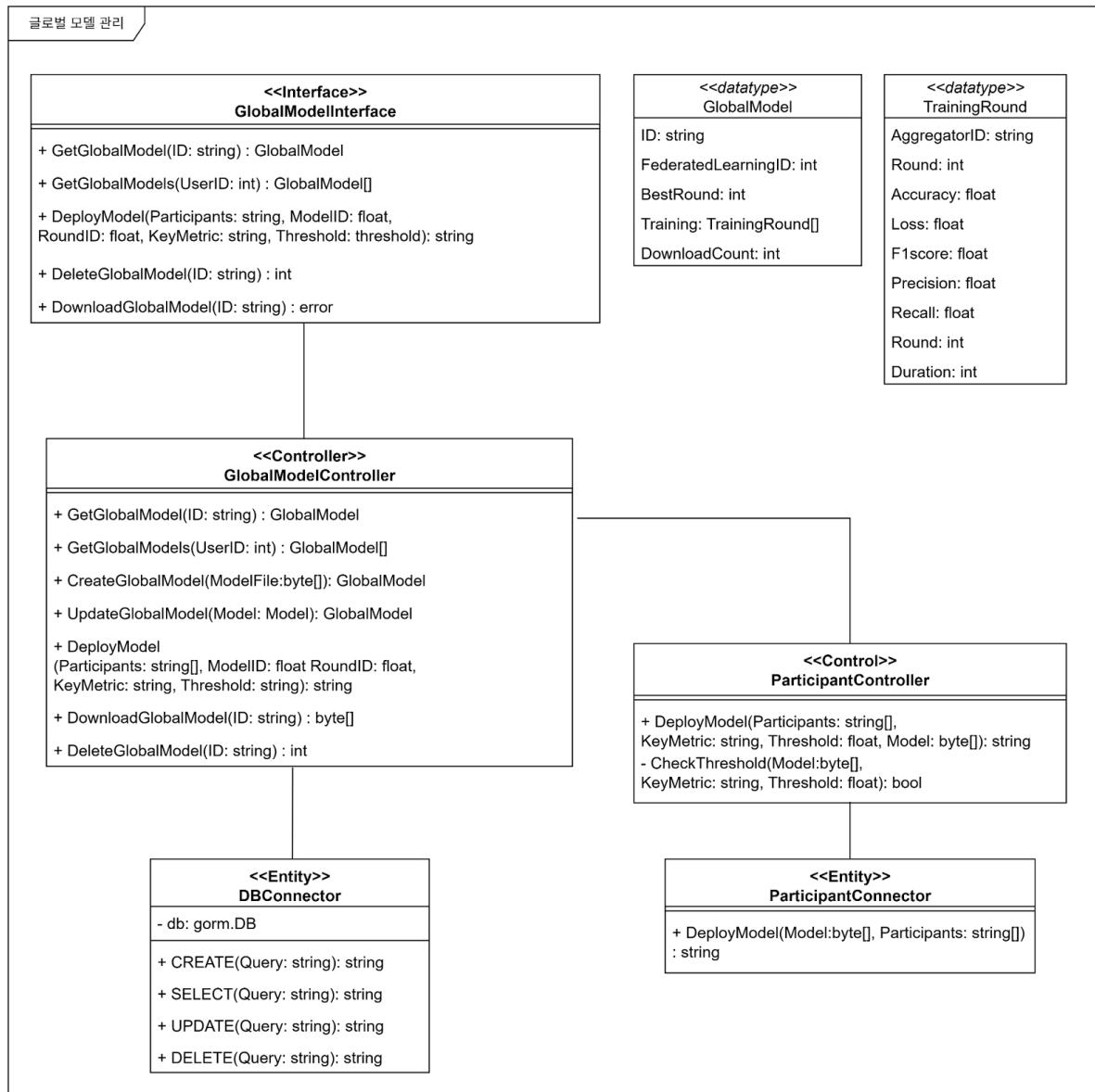


그림 16. 가상머신 관리 클래스 다이어그램

3.1.4. 시퀀스 다이어그램

① 사용자 인증 관리

1) 로그인

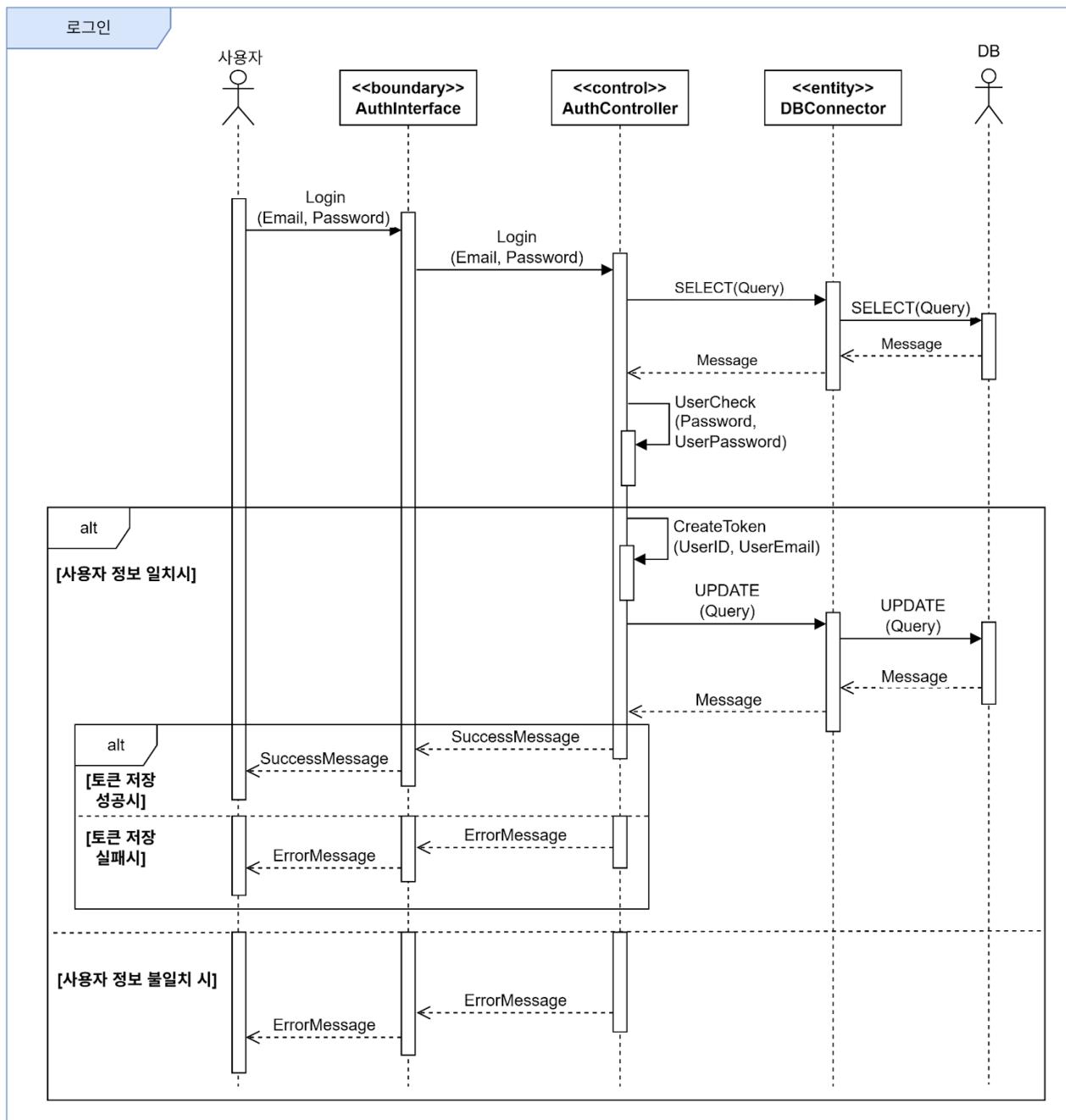


그림 17. 로그인 시퀀스 다이어그램

2) 회원가입

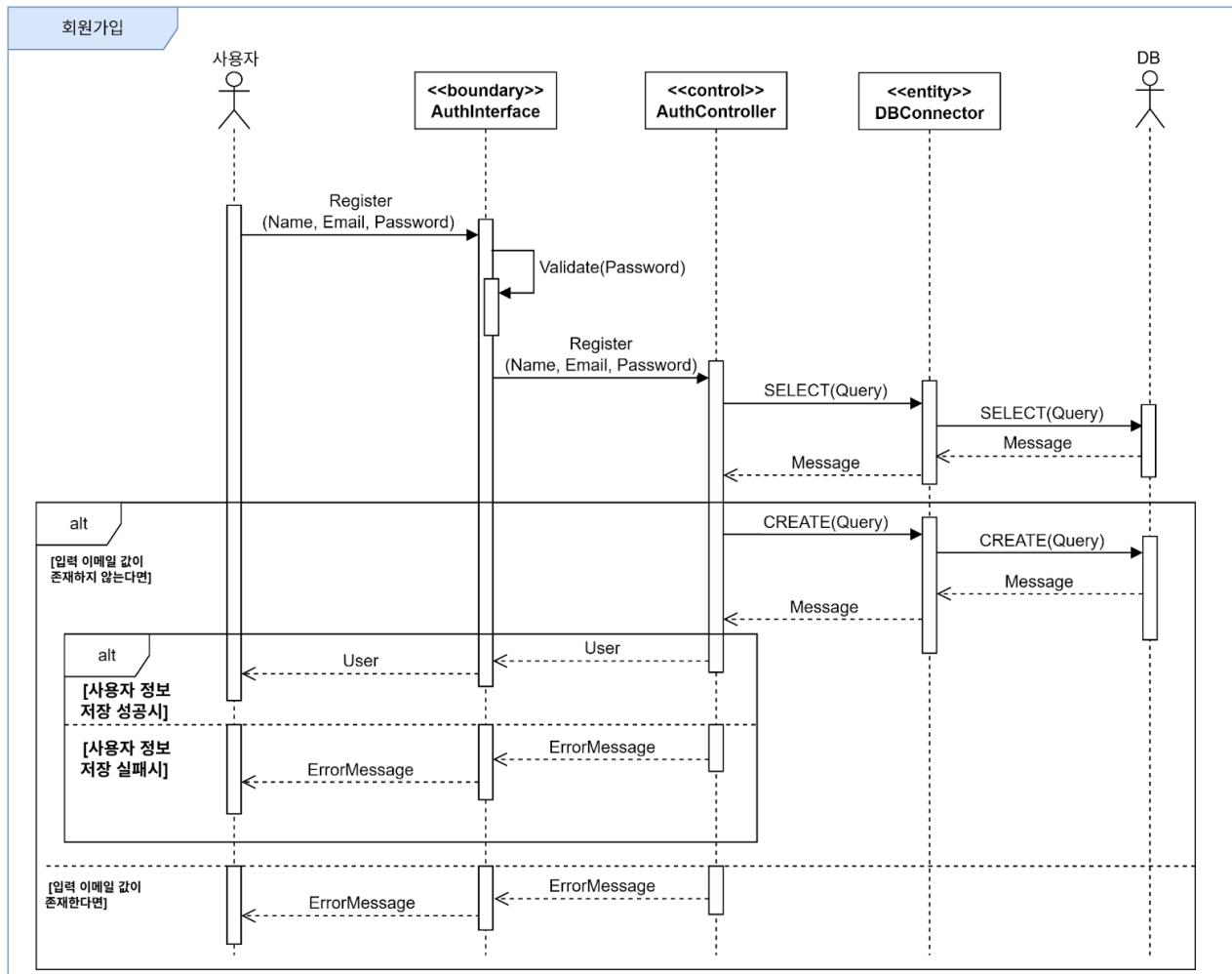


그림 18. 회원가입 시퀀스 디어그램

3) 회원탈퇴

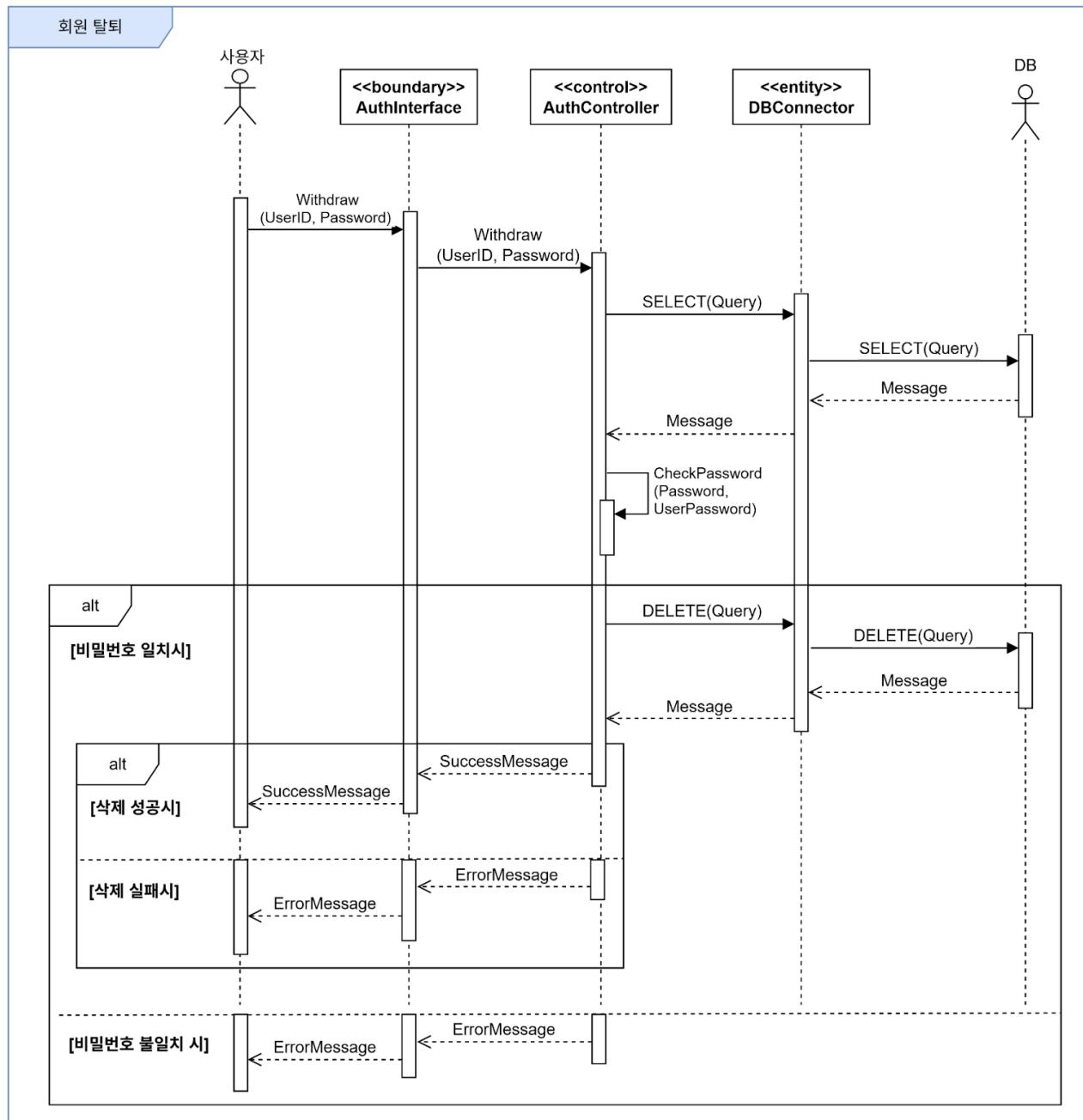


그림 19. 회원 탈퇴 시퀀스 다이어그램

② 클라우드 인증 정보 관리

1) 클라우드 인증 정보 등록

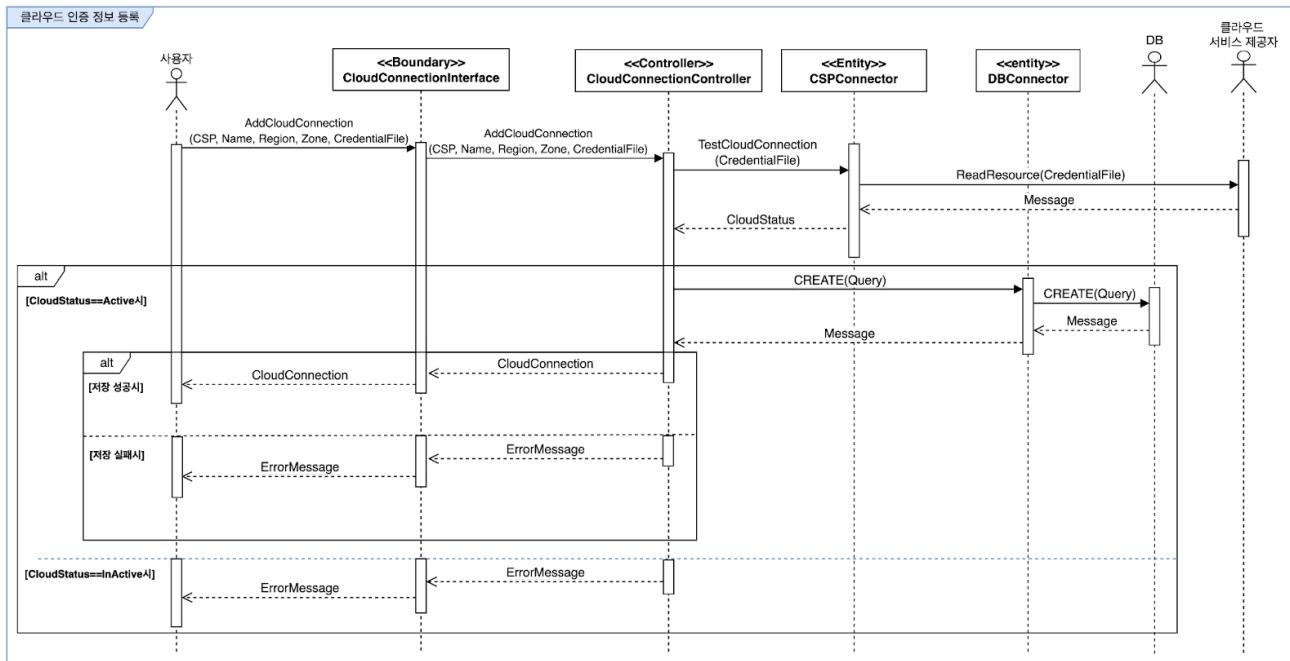


그림 20. 클라우드 인증 정보 등록 시퀀스 다이어그램

2) 클라우드 인증 정보 상세 조회

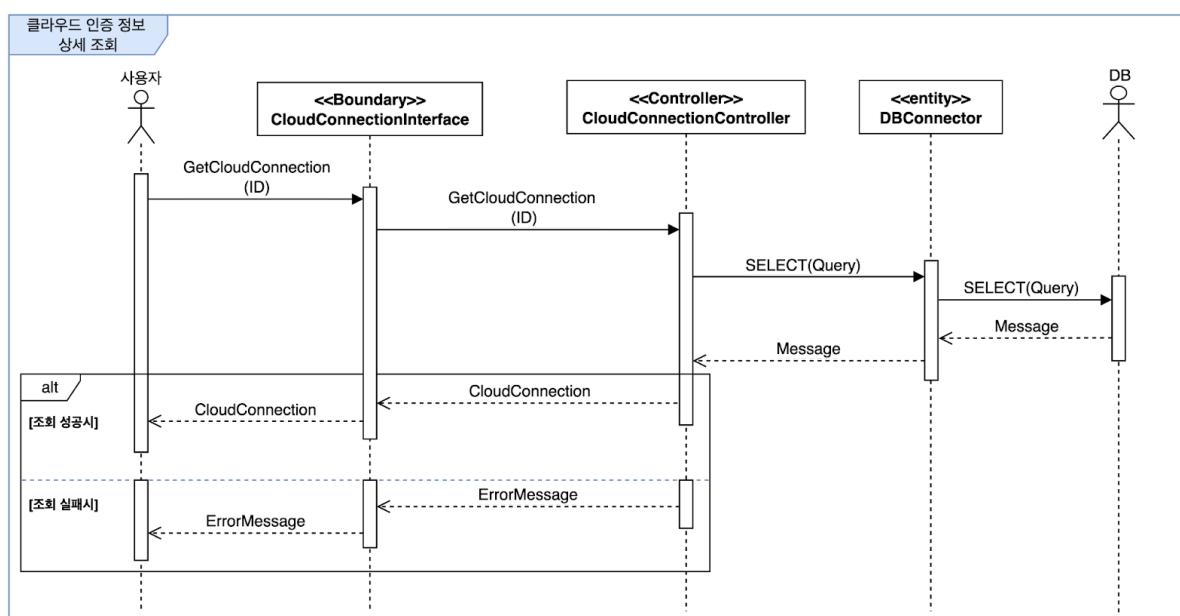


그림 21. 클라우드 인증 정보 상세 조회 시퀀스 다이어그램

3) 클라우드 인증 정보 전체 조회

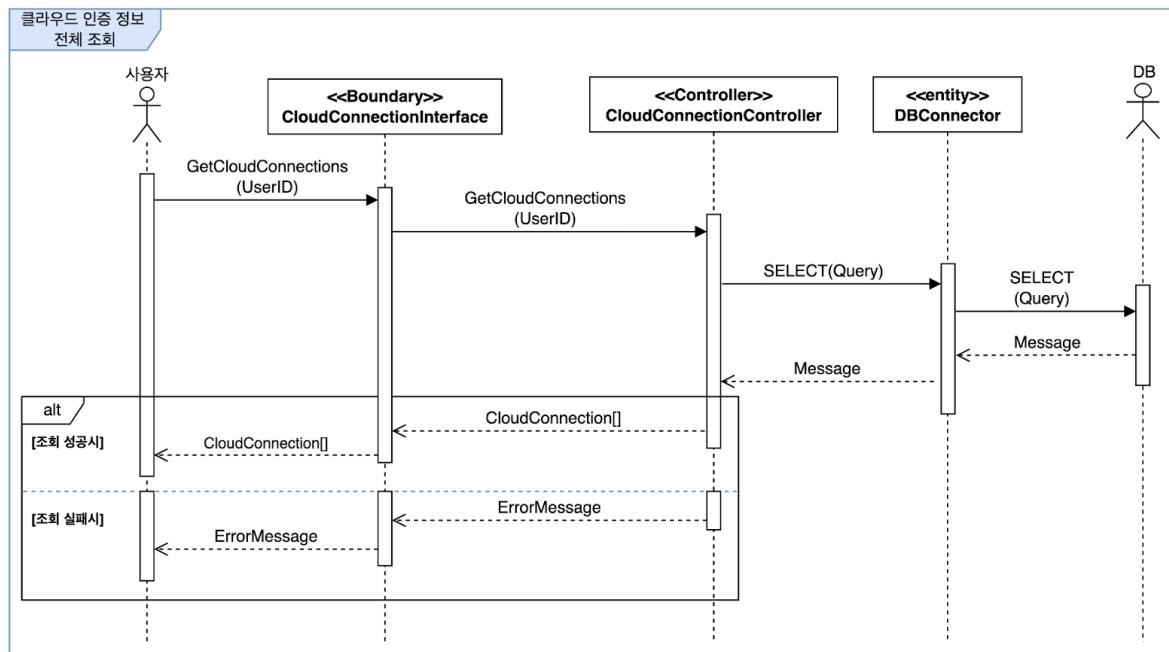


그림 22. 클라우드 인증 정보 전체 조회 시퀀스 다이어그램

4) 클라우드 인증 정보 삭제

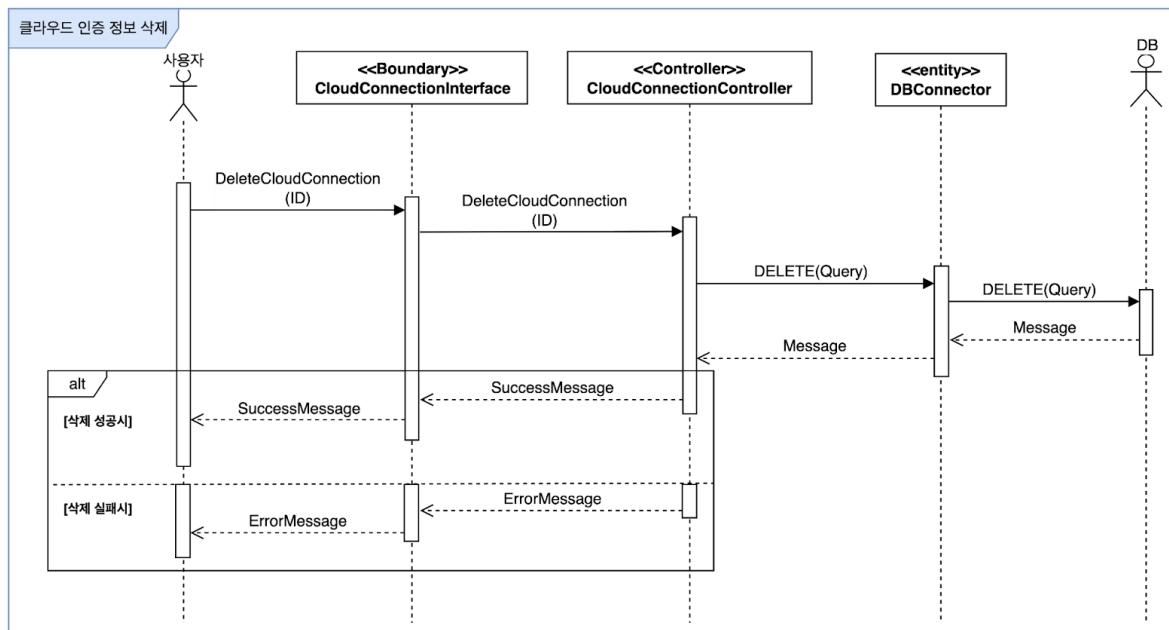


그림 23. 클라우드 인증 정보 삭제 시퀀스 다이어그램

③ 연합학습 집계자 관리

1) 연합학습 집계자 생성

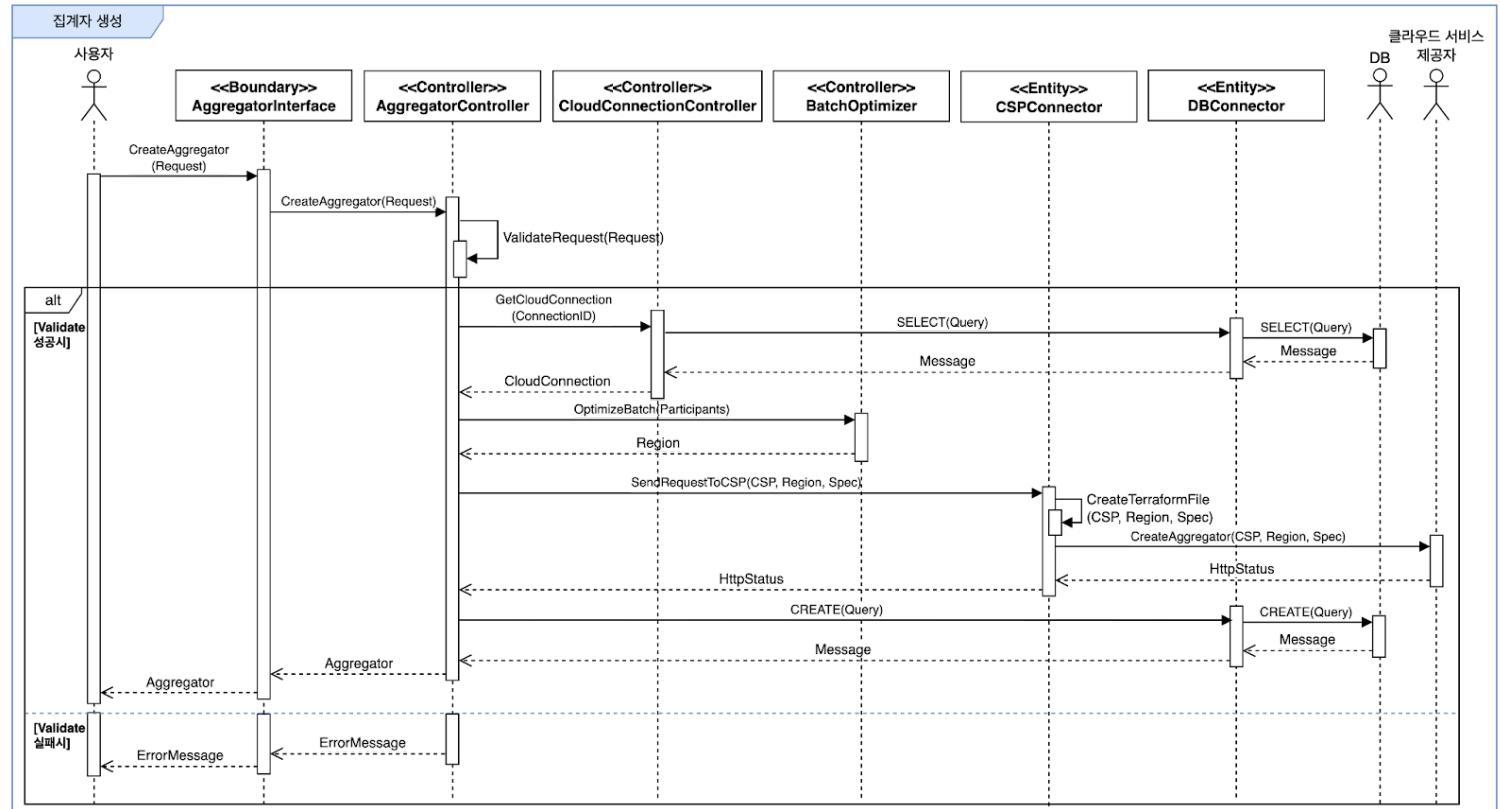


그림 24. 연합학습 집계자 생성 시퀀스 다이어그램

2) 연합학습 집계자 상세 조회

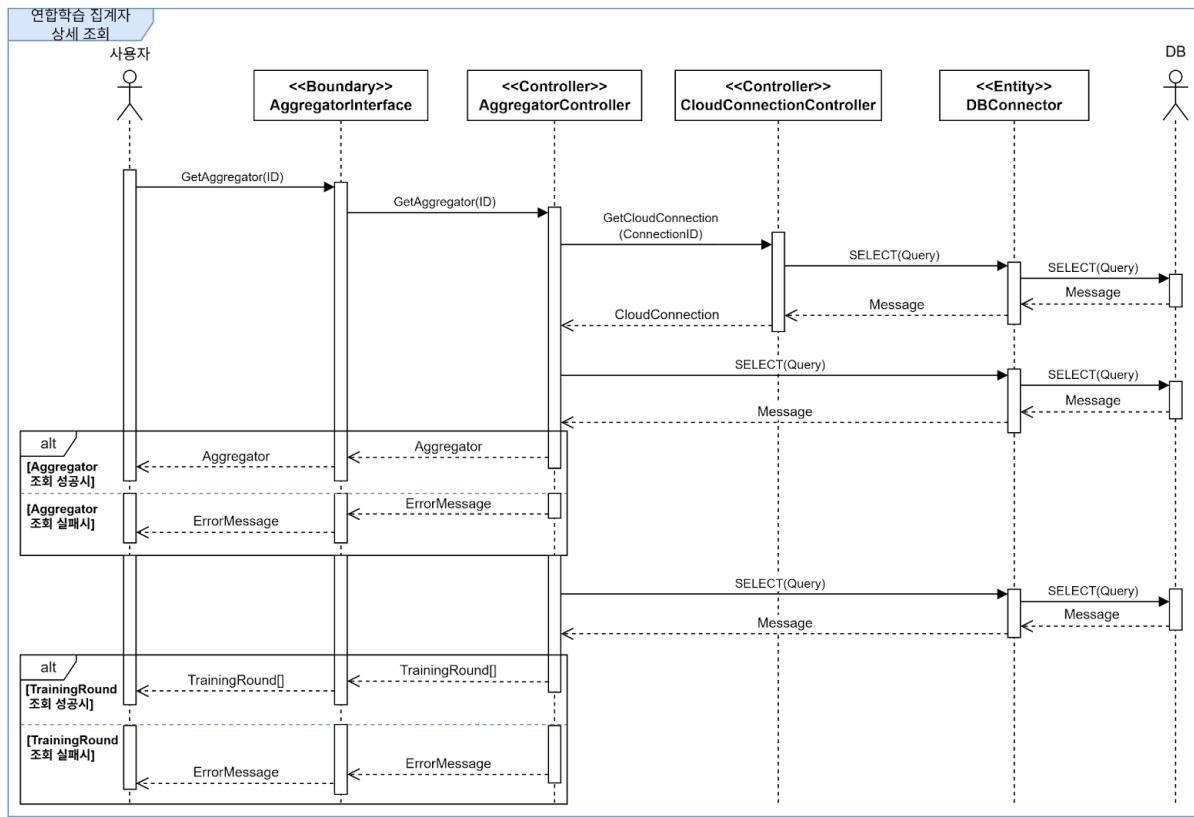


그림 25. 연합학습 집계자 상세 조회 시퀀스 다이어그램

3) 연합학습 집계자 전체 조회

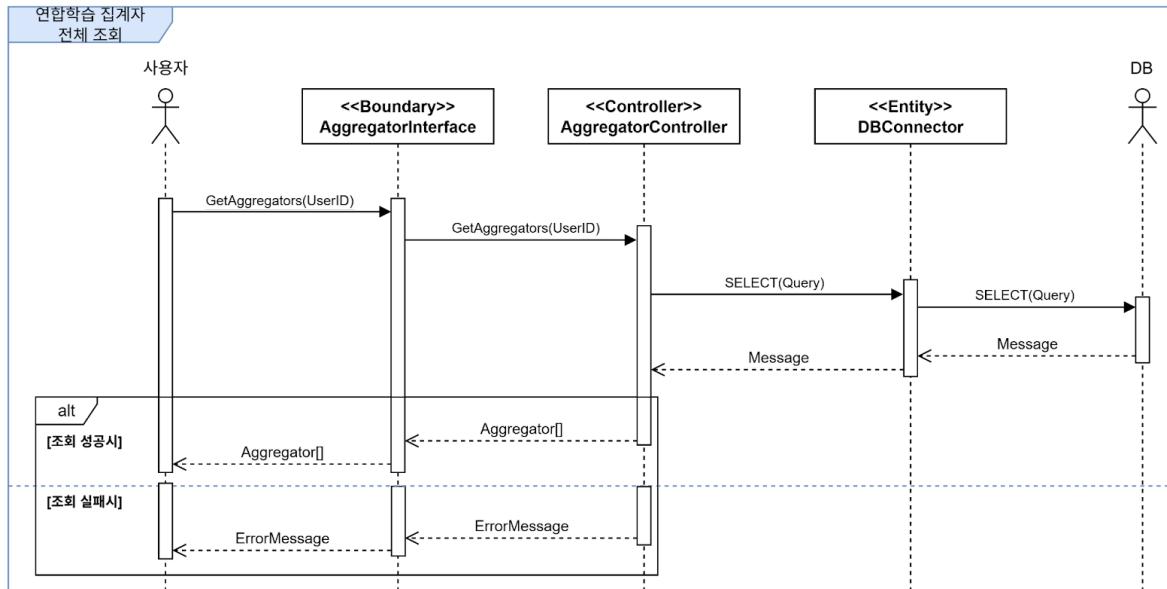


그림 26. 연합학습 집계자 전체 조회 시퀀스 다이어그램

4) 연합학습 집계자 삭제

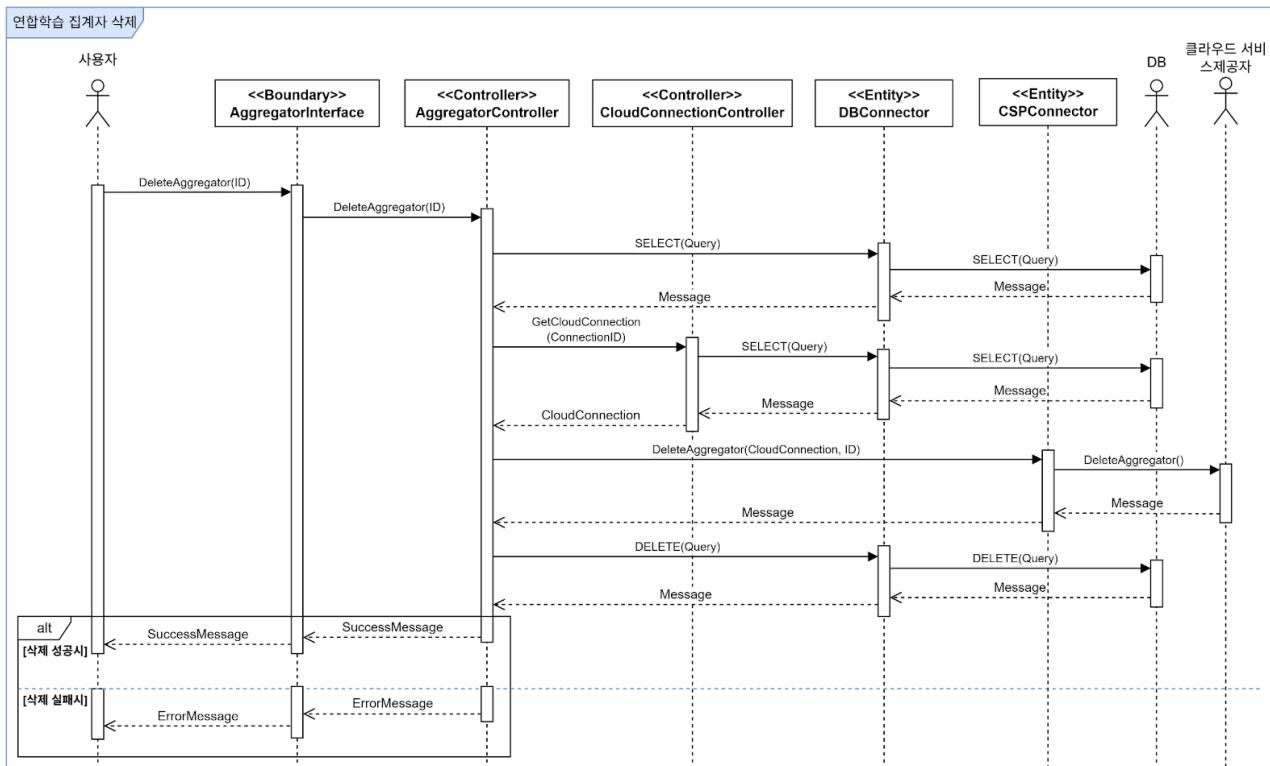


그림 27. 연합학습 집계자 삭제 시퀀스 다이어그램

④ 연합학습 참여자 관리

1) 연합학습 참여자 등록

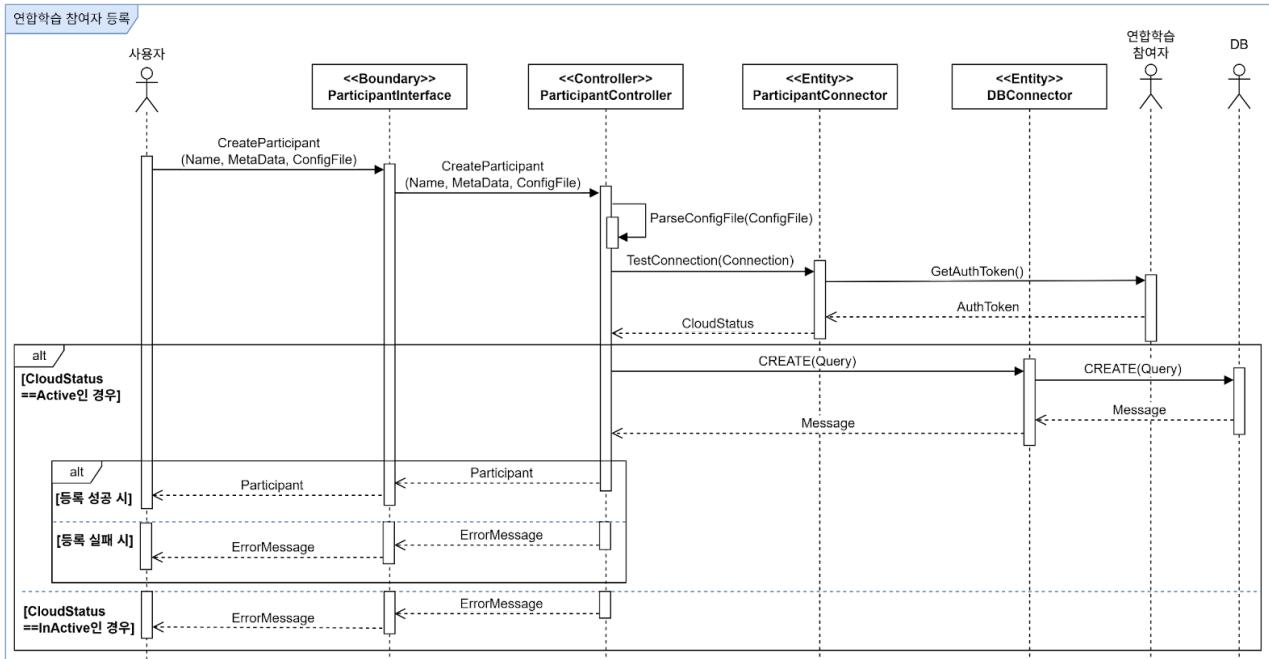


그림 28. 연합학습 참여자 등록 시퀀스 다이어그램

2) 연합학습 참여자 상세 조회

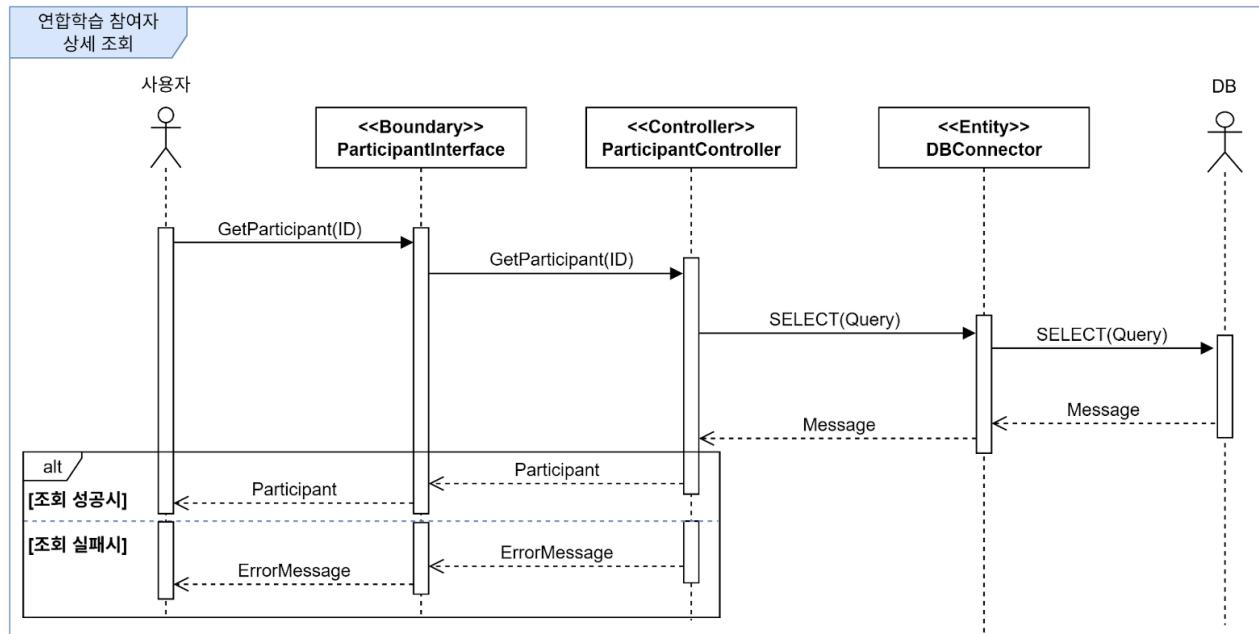


그림 29. 연합학습 참여자 상세 조회 시퀀스 다이어그램

3) 연합학습 참여자 전체 조회

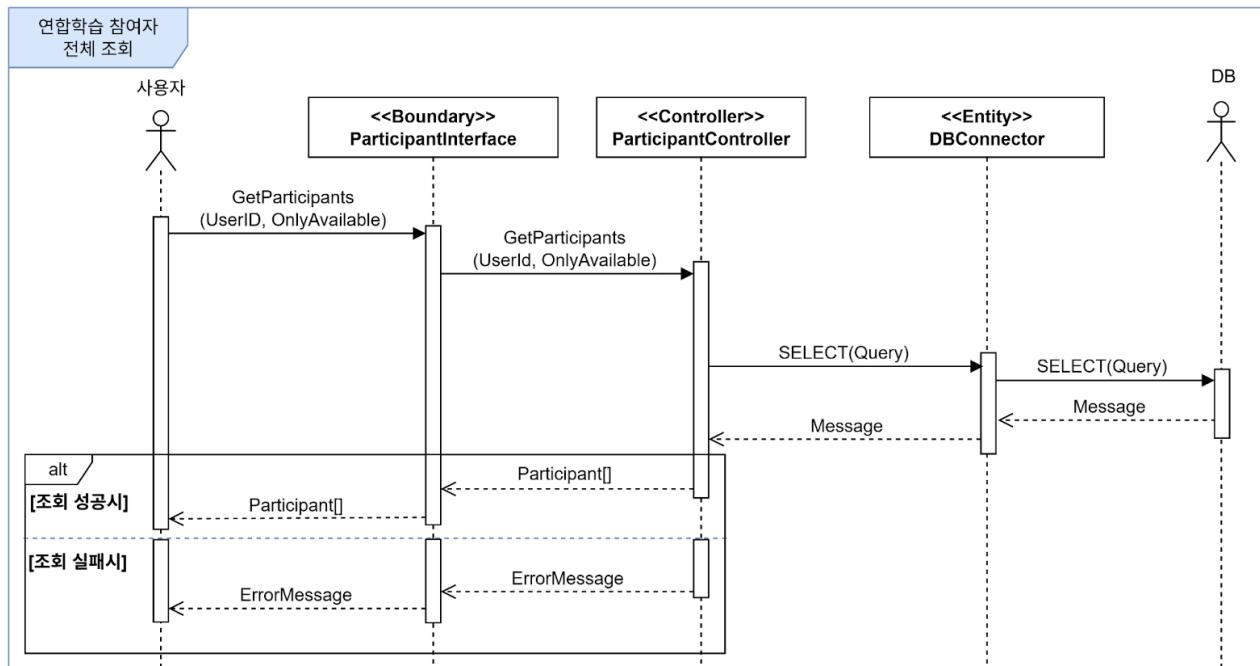


그림 30. 연합학습 참여자 전체 조회 시퀀스 다이어그램

4) 연합학습 참여자 삭제

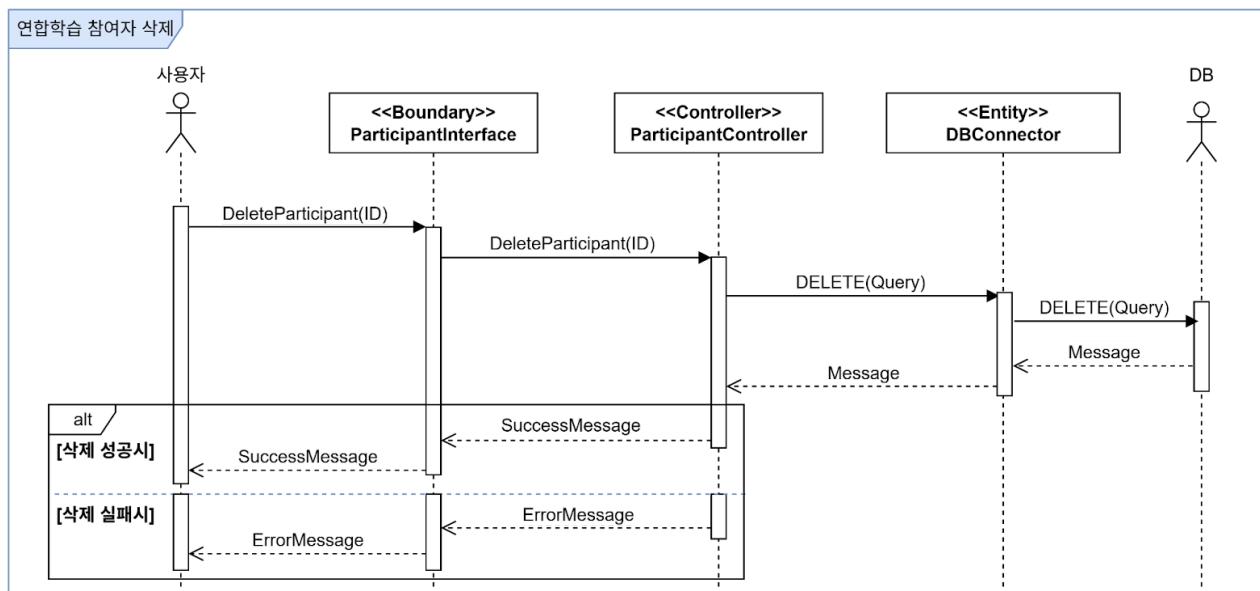


그림 31. 연합학습 참여자 삭제 시퀀스 다이어그램

1) 연합학습 생성

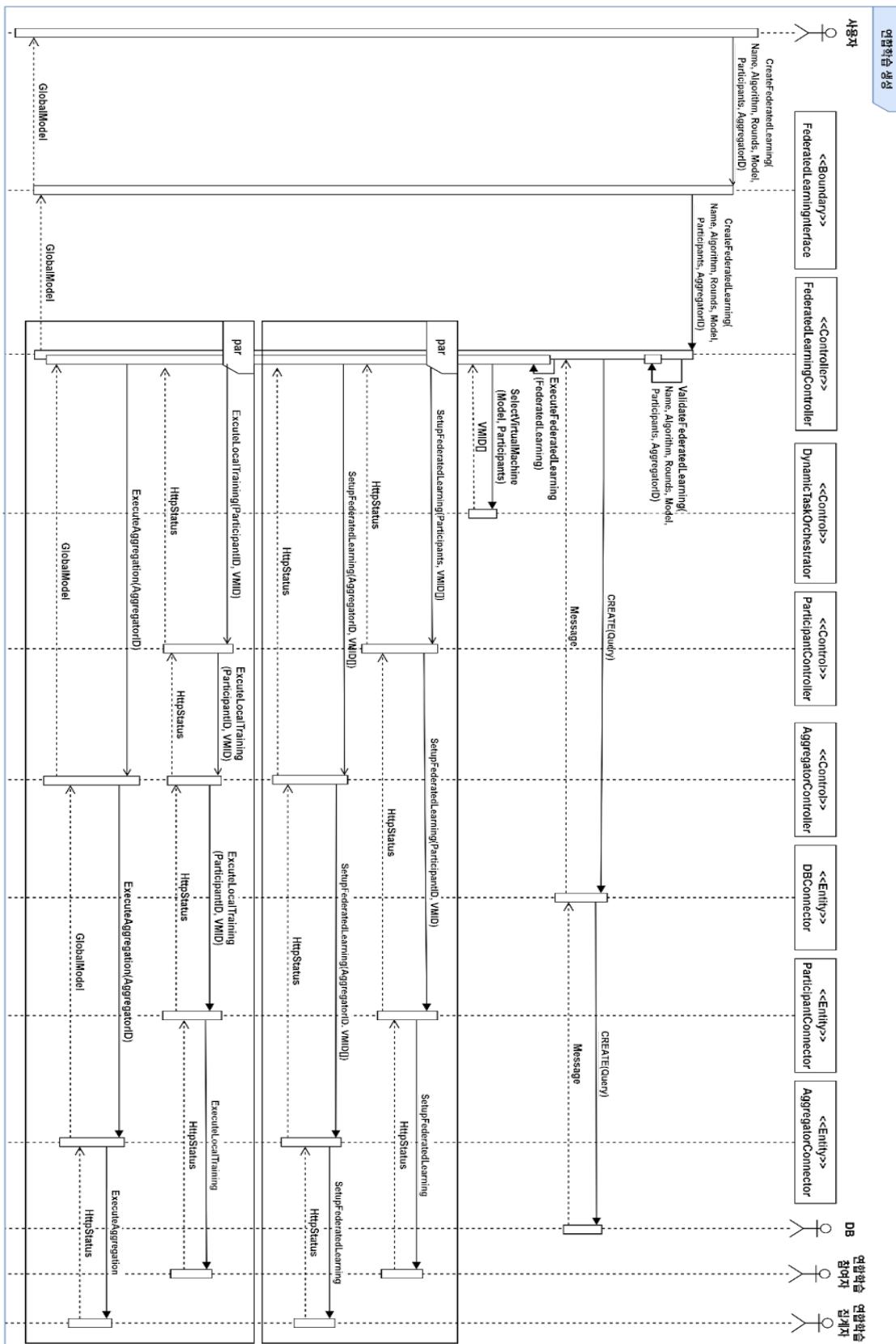


그림 32. 연합학습 생성 시퀀스 다이어그램

2) 연합학습 상세 조회

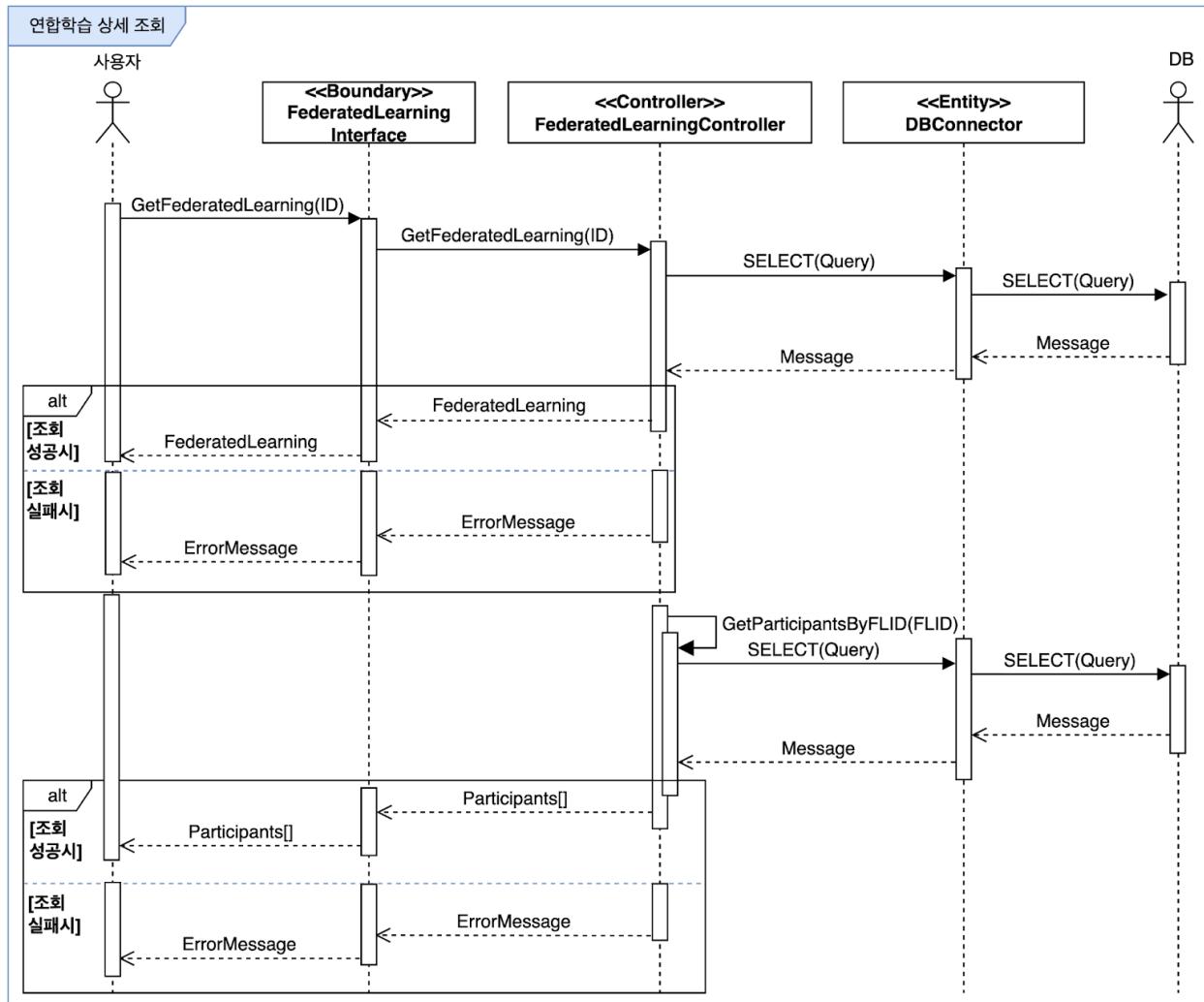


그림 33. 연합학습 상세 조회 시퀀스 다이어그램

3) 연합학습 전체 조회

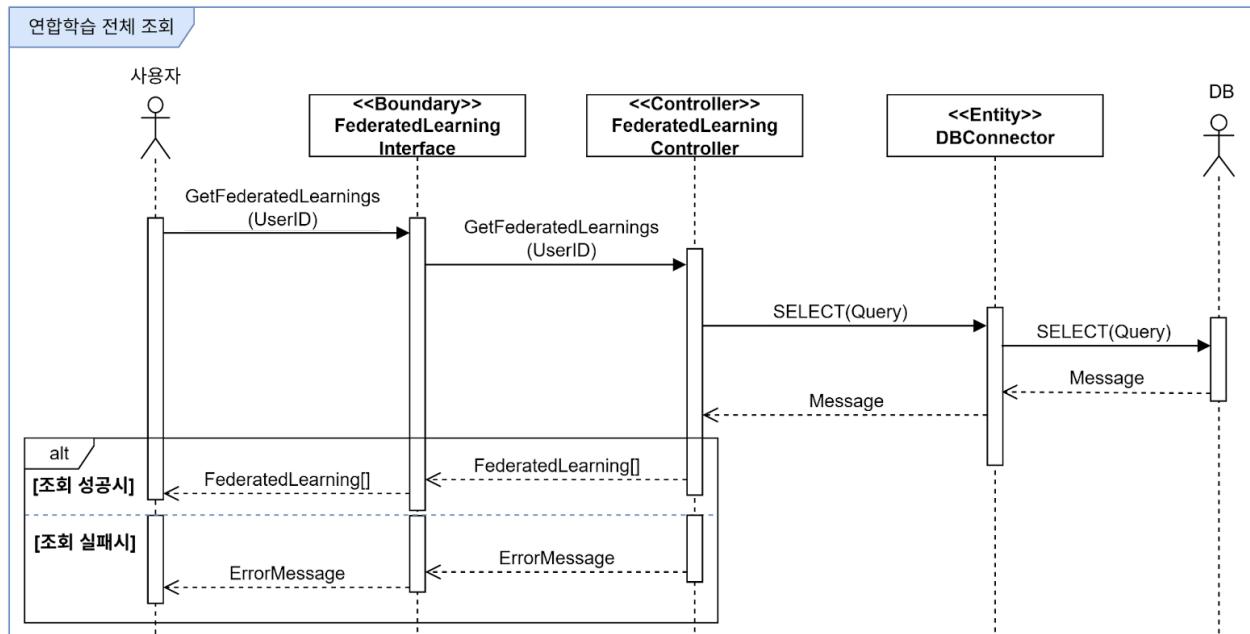


그림 34. 연합학습 전체 조회 시퀀스 다이어그램

4) 연합학습 삭제

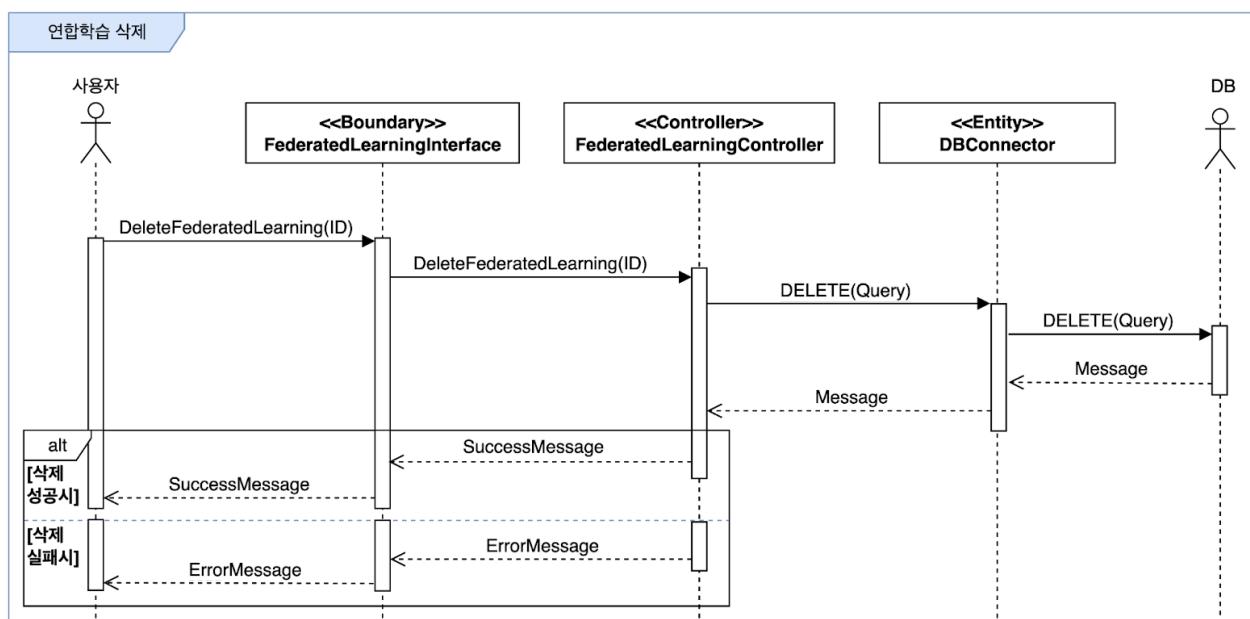


그림 35. 연합학습 삭제 시퀀스 다이어그램

⑥ 글로벌 모델 관리

1) 글로벌 모델 생성

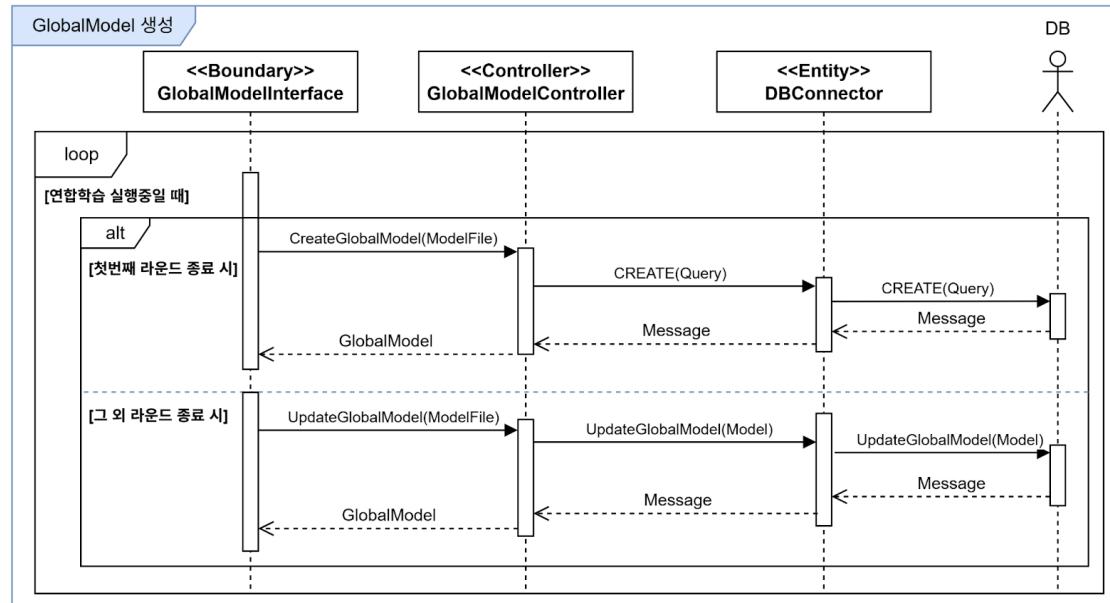


그림 36. 글로벌 모델 생성 시퀀스 다이어그램

2) 글로벌 모델 배포

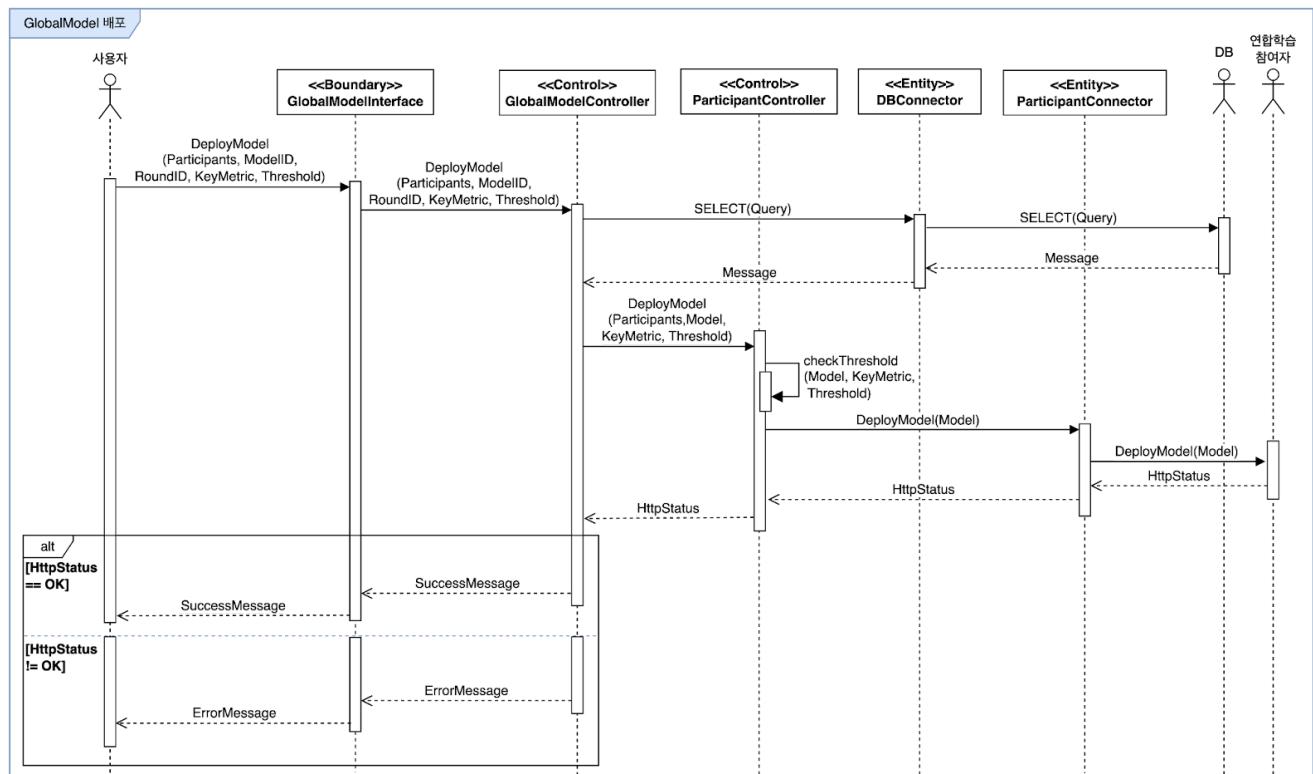


그림 37. 글로벌 모델 배포 시퀀스 다이어그램

3) 글로벌 모델 상세 조회

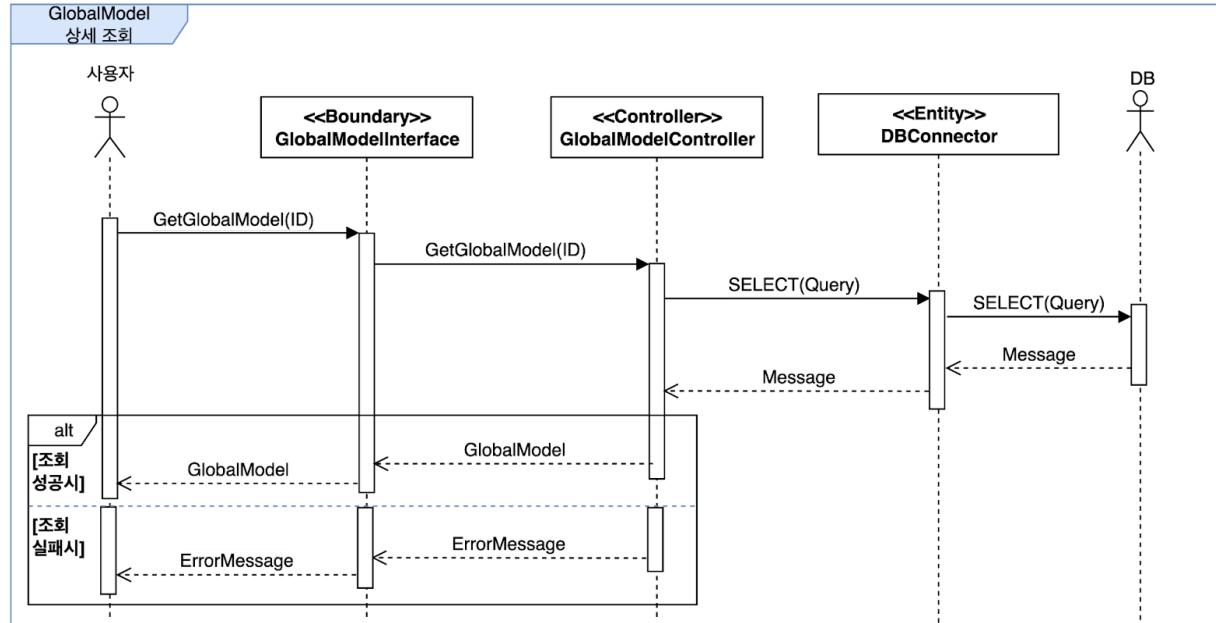


그림 38. 글로벌 모델 상세 조회 시퀀스 다이어그램

4) 글로벌 모델 전체 조회

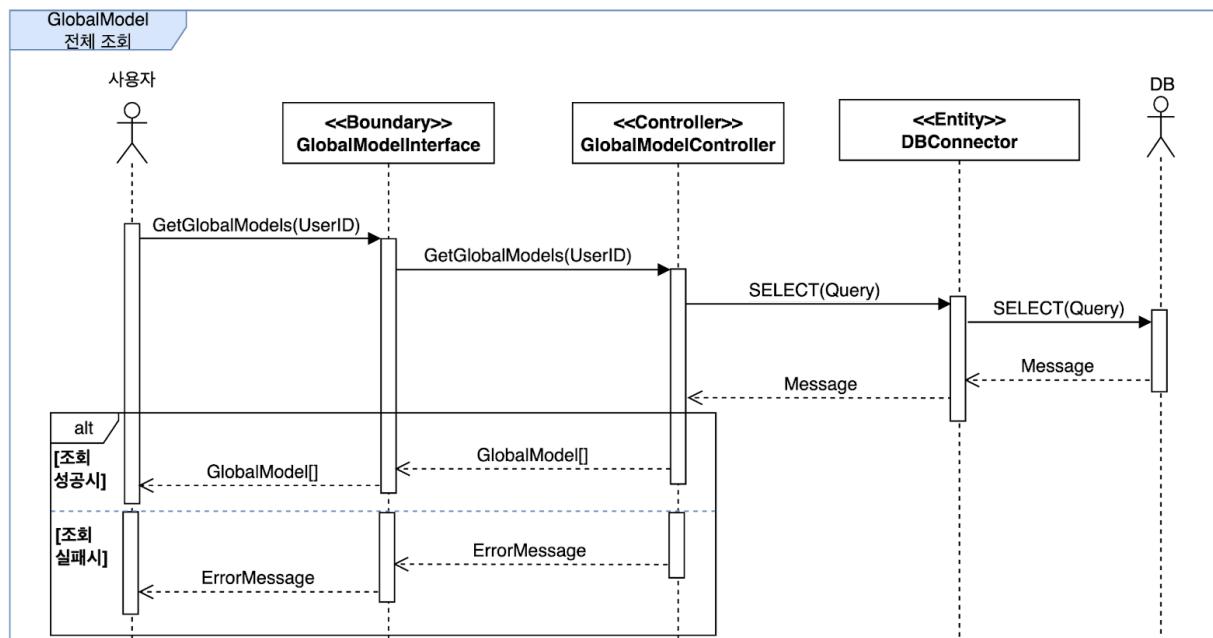


그림 39. 글로벌 모델 전체 조회 시퀀스 다이어그램

5) 글로벌 모델 삭제

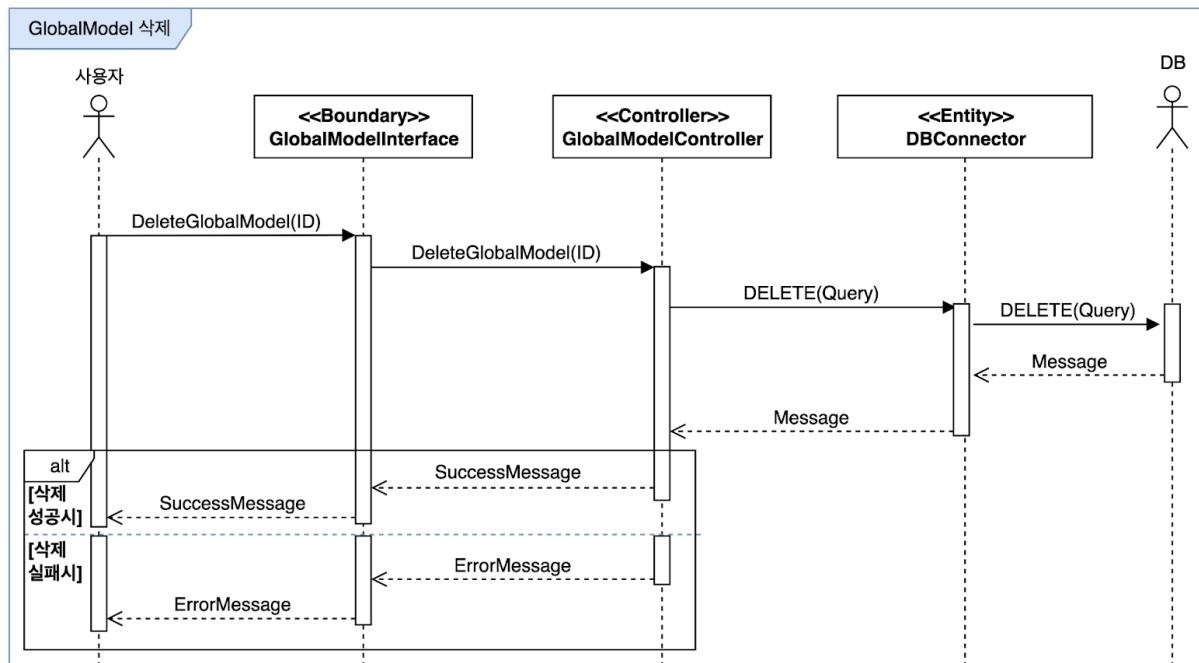


그림 40. 글로벌 모델 삭제 시퀀스 다이어그램

6) 글로벌 모델 다운로드

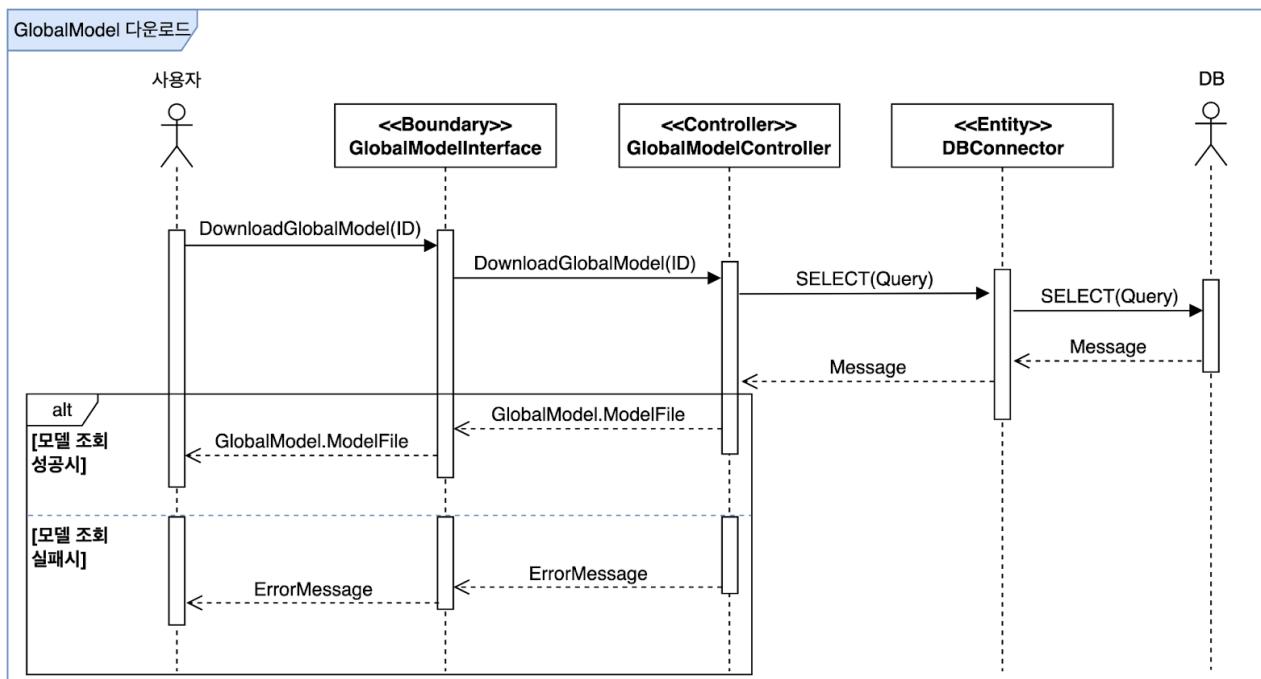


그림 41. 글로벌 모델 다운로드 시퀀스 다이어그램

⑦ 가상머신 관리

1) 가상머신 모니터링

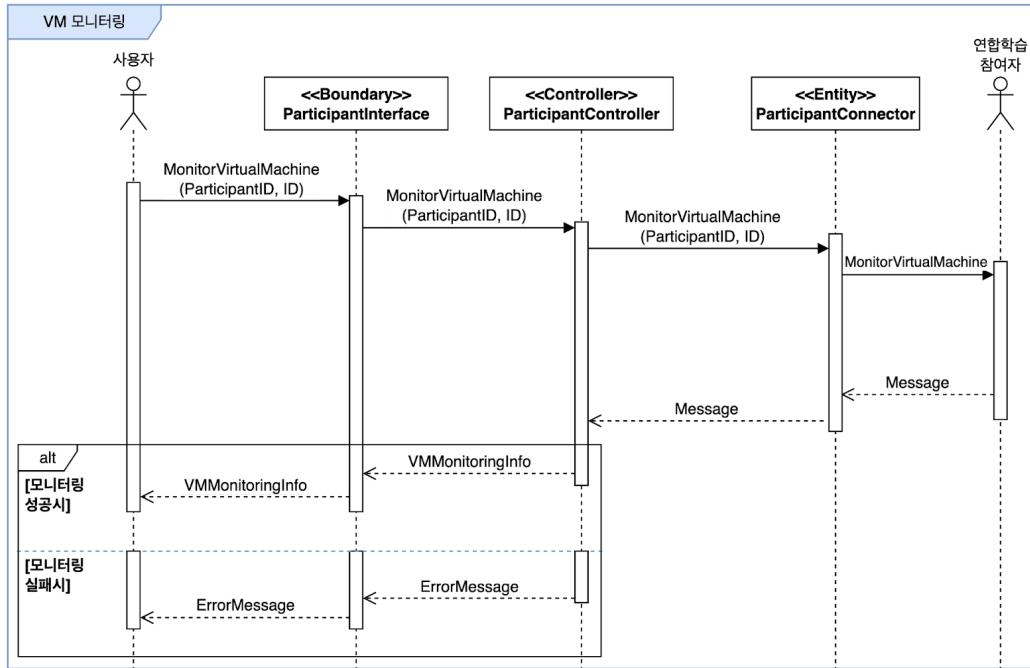


그림 42. 가상머신 모니터링 시퀀스 다이어그램

2) 가상머신 전체 조회

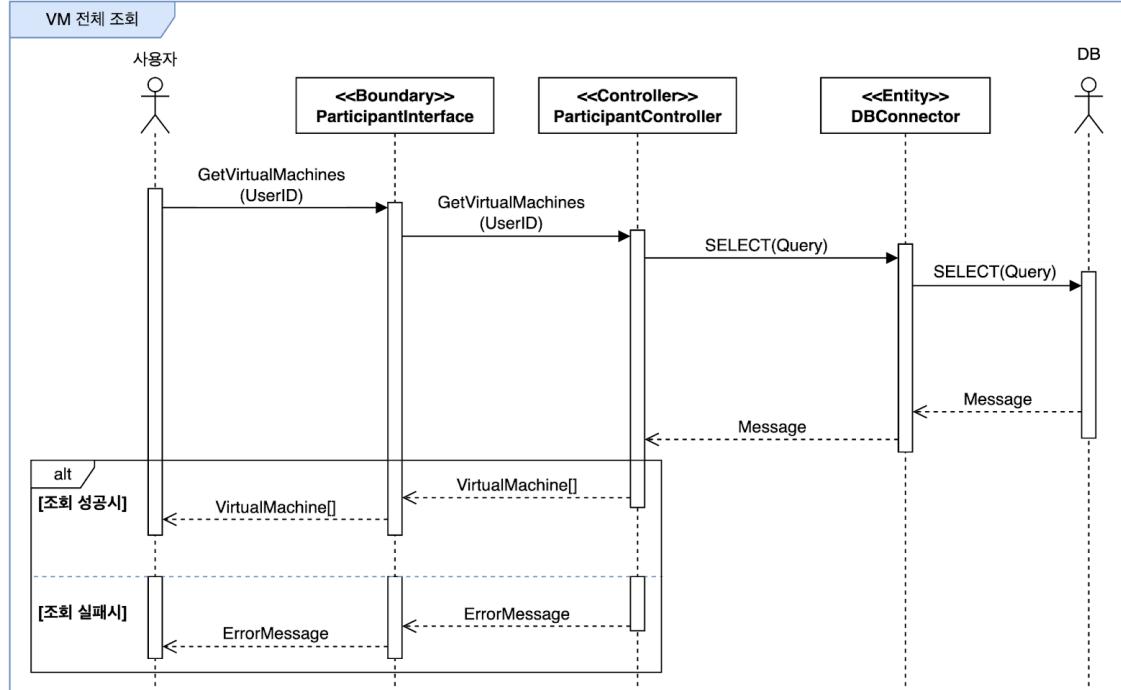


그림 43. 가상머신 전체 조회 시퀀스 다이어그램

3.2. 시스템 구성

3.2.1. 시스템 아키텍처

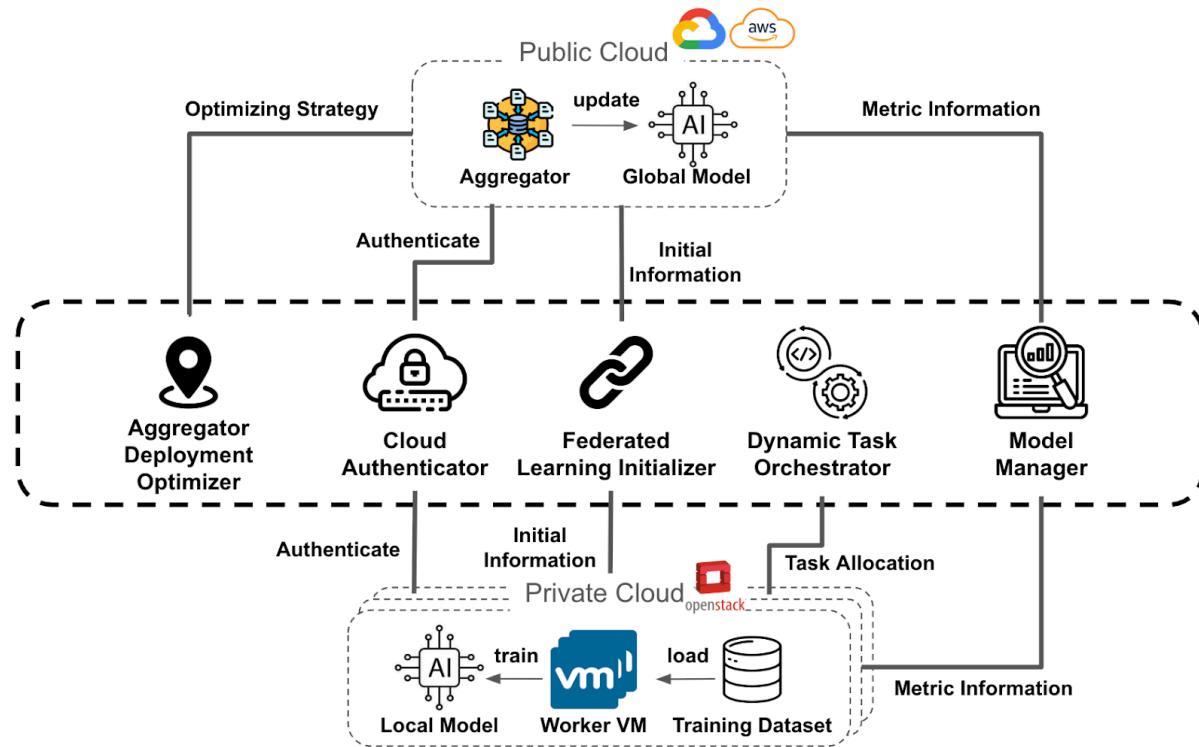


그림 44. 시스템 아키텍처

본 연구에서는 퍼블릭 클라우드와 프라이빗 클라우드의 기능을 명확히 구분하여 연합학습을 수행할 수 있는 멀티 클라우드 기반 연합학습 시스템을 구축한다.

그림 44는 제안하는 시스템의 아키텍처를 나타낸 것이다. 퍼블릭 클라우드에는 연합학습 집계자와 글로벌 모델이 배치되어, 모델 파라미터 집계와 글로벌 모델 업데이트를 담당한다. 반면 프라이빗 클라우드(OpenStack 기반)에는 로컬 모델, 연합학습 참여자 가상머신, 학습 데이터셋이 배치되어, 각 연합학습 참여자는 로컬 환경에서 개별적으로 모델 학습을 수행한다. 학습을 수행하며, 연합학습 참여자 가상머신은 로컬 모델의 파라미터를 퍼블릭 클라우드의 집계자에게 전송하고, 집계자는 전송받은 파라미터를 통합하여 글로벌 모델을 개선한다. 이와 같은 계층 구조를 통해 데이터는 프라이빗 클라우드 내부에 안전하게 유지되면서도, 퍼블릭 클라우드의 접근성과 확장성을 활용한 효율적인 연합학습 수행이 가능하다.

나아가, 제안 시스템은 멀티 클라우드 기반 연합학습 환경을 구축하고 효율적인

연합학습을 지원하기 위해 다음과 같은 핵심 모듈들을 포함한다.

- Aggregator Deployment Optimizer: 사용자 요구사항 기반의 최적 연합학습 집계자 명세(클라우드 리전, 스펙) 추천 및 추천 명세 기반 연합학습 집계자 배포를 통한 클라우드 비용 및 지연시간 최적화
- Cloud Authenticator: 멀티 클라우드 연동을 퍼블릭 클라우드 인증 정보 관리 및 클라우드 인증 정보 기반의 연합학습 참여자 등록
- Federated Learning Initializer: 연합학습 집계자와 연합학습 참여자의 연합학습 수행을 위한 환경 설정(필수 라이브러리 설치, 환경 변수 설정, 학습 코드) 및 연합학습 수행 명령 전달
- Dynamic Task Orchestrator: 연합학습 참여자 내 가상머신의 자원 상태와 장애 여부를 기반으로 적절한 가상머신에 작업을 동적으로 할당하고, 실패 시 재시도를 통한 안정적인 학습 수행 보장
- Model Manager: 연합학습의 각 라운드별로 생성된 글로벌 모델의 관리를 수행. 글로벌 모델 저장 및 평가 지표 모니터링, 모델 파일 관리, 사용자 요구사항 기반 최적의 모델 다운로드 지원

3.2.2. 개발환경

표 2 개발 환경

개발환경	사용기술
소프트웨어 형상관리(SCM)	GitHub
컨테이너 기술	Docker
클라우드 플랫폼	AWS, GCP, OpenStack
모니터링	Prometheus, Grafana
Infra Provisioning 자동화(IaC)	Terraform
MLOps	MLFlow
Federated Learning Framework	Flower
DBMS	PostgreSQL
Server	gin-gonic
UI	Next.js
프로젝트 문서화 관리	Notion, Google Drive

3.3. 멀티 클라우드 인프라 기반 연합학습 환경 구축 지원 플랫폼

3.3.1. 사용자 인증

그림 45는 사용자 인증을 위한 로그인 화면을 나타낸 것이다. 플랫폼에 접근하는 사용자의 신원을 확인하고 권한을 부여함으로써, 사용자 간 분리된 연합학습 플랫폼을 제공하기 위해 사용자 인증 기능을 제공한다. 이를 통해 사용자는 자신이 생성 및 등록한 데이터에만 접근할 수 있다.

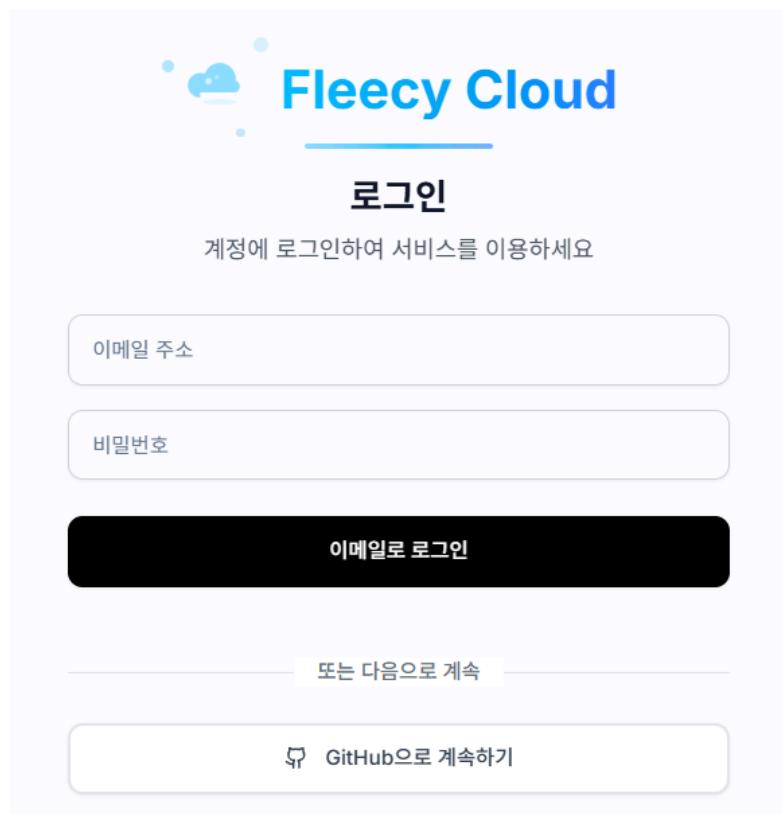


그림 45. 사용자 인증 화면

3.3.2. 클라우드 인증 정보 관리

클라우드 인증 정보 관리 기능은 멀티 클라우드 연동을 위해 다양한 클라우드 제공자(AWS, GCP)의 인증 정보를 등록 및 관리할 수 있도록 한다. 이를 통해 사용자는 멀티 클라우드 환경에서 각 클라우드의 기능을 활용하여 연합학습을 수행할 수 있다.

그림 46은 클라우드 인증 정보 등록 화면이다. 사용자는 클라우드 인증 정보 이름, 사용할 클라우드의 리전 이름, 영역 이름과 같은 기본 정보를 입력하고 클라우드 자격 증명 파일을

업로드 해야한다. 시스템은 업로드 된 자격 증명 파일의 유효성을 검증하기 위해 클라우드 리소스를 읽어오는 작업을 수행한 후 검증이 성공적으로 완료된 경우에만 인증 정보를 저장한다. AWS의 경우 CSV파일, GCP의 경우 Json 파일을 자격 증명 파일로 업로드해야 한다.



그림 46. 클라우드 인증 정보 등록 화면

3.3.3 연합학습 집계자 관리

연합학습 집계자는 각 연합학습 참여자의 학습 결과를 수집 및 통합하여 글로벌 모델을 갱신하는 핵심 역할을 수행한다. 본 연구는 이러한 연합학습 집계자 인스턴스를 효율적으로 배포하고, 인스턴스의 상태를 직관적으로 모니터링할 수 있는 기능을 제공한다. 이를 통해 사용자는 요구사항을 기반으로 멀티 클라우드 환경에서 비용 및 지연 시간이 최적화된 집계자를 생성할 수 있으며, 연합학습 집계자 관리를 수행할 수 있다.

① 연합학습 집계자 최적화

3.3.5.절의 ① 연합학습 생성 단계 중에 시스템은 사용자 요구사항을 기반으로 연합학습 집계자의 배포 명세(배포 리전, 스펙) 최적화를 수행한다. 최적화의 목표는 클라우드 비용 및 연합학습 참여자간 네트워크 지연 시간 최소화이며. 사용자 요구사항은 그림 47과 같이 최대 허용 비용, 최대 허용 지연시간 그리고 비용-지연시간 가중치(우선순위)이다.

비용과 지연 시간이라는 두 개의 상충될 수 있는 목표를 최적화하기 위해 본 연구는 유전 알고리즘 중 하나인 NSGA-II(Non-dominated Sorting Genetic Algorithm II)[22]를 활용한다. NSGA-II는 두 개 이상의 상충될 수 있는 목표를 동시에 고려하면서 목표간의 균형점을 찾는데 효과적인 알고리즘이다.

최적화의 수행 과정은 다음과 같다. 먼저 클라우드 비용과 네트워크 지연 시간을 기반으로 모든 집계자 배포 명세 후보군을 생성한 뒤 사용자의 요구사항인 최대 허용 비용과 최대 허용 지연시간을 적용하여 이를 초과하는 후보를 제외한다. 그 후 NSGA-II 알고리즘을 적용하여 두 목표간의 균형점의 집합인 파레토 프론트[23]를 도출한다. 알고리즘은 선택·교차·변이 연산을 반복하며 해의 품질과 다양성을 개선하고, 성능이 비슷한 해들이 한쪽에 몰리지 않도록 거리 값을 계산해 고르게 분산시킨다. 마지막으로 비용-지연시간 가중치를 반영해 비용과 지연 시간이 최적화된 집계자의 배포 명세를 추천한다.

그림 48은 사용자 요구사항을 기반으로 한 최적화 과정을 통해 도출된 집계자 최적화 결과를 보여준다. 시스템은 최적화 수행 후 사용자 요구사항을 반영한 추천 점수가 높은 순으로 배포 명세를 최대 20개까지 나열한다. 추천 점수 산출 공식은 수식 (1)과 같다. 최종적으로 사용자는 선호하는 집계자 배포 명세를 선택하여 집계자를 배포할 수 있다.

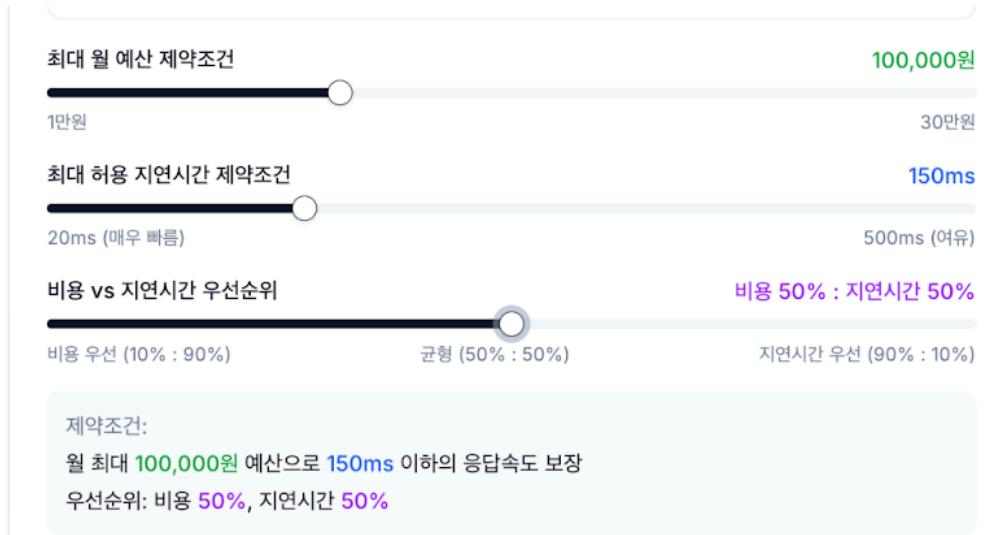


그림 47. 사용자 요구사항을 기반으로 한 집계자 최적화 설정 화면

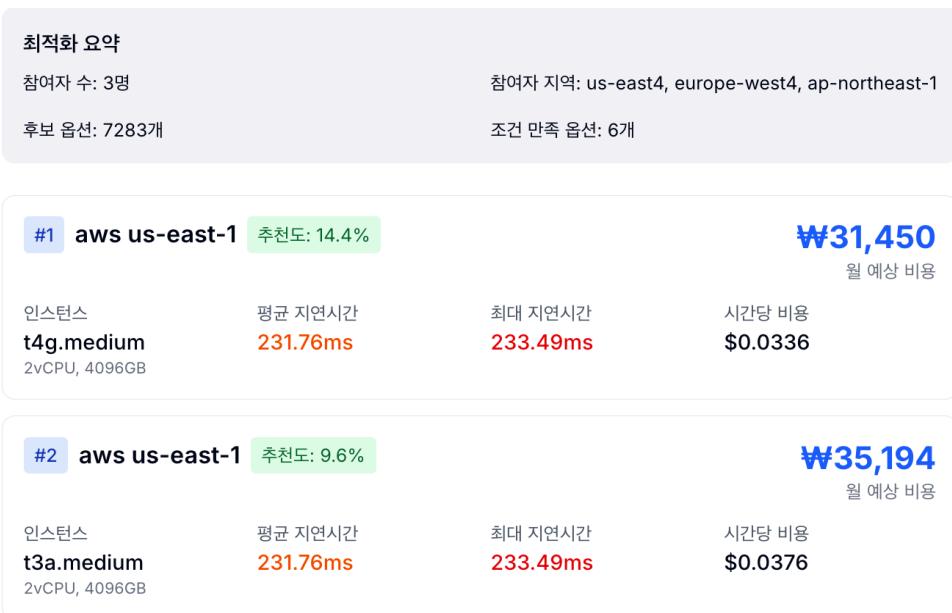


그림 48. 집계자 배포 최적화 결과 화면

$$Score = (1 - (w_c \times \frac{C_i}{C_{max}} + w_l \times \frac{L_i}{L_{max}})) \times 100 \quad (1)$$

w_c : 사용자 설정 비용 가중치

C_i : 파레토 프론트 i의 비용

w_l : 사용자 설정 클라우드 지연 시간

C_{max} : 사용자 설정 최대 허용 비용

가중치

L_i : 파레토 프론트 i의 지연 시간

L_{max} : 사용자 설정 최대 허용 지연 시간

② 연합학습 집계자 상세 모니터링

사용자는 연합학습 집계자를 실시간으로 모니터링할 수 있다. 이때 집계자의 기본 정보 및 하드웨어 사양과 더불어, 그림 49과 같이 현재 진행 중인 연합학습의 라운드 및 모델 성능 지표(Accuracy, Precision, Recall, F1-Score)를 실시간으로 확인할 수 있다. 또한, Prometheus를 통해 실시간으로 모니터링 되는 연합학습 집계자의 시스템 메트릭(CPU 사용률, 메모리 사용률 등)을 확인할 수 있으며, 연합학습 라운드가 진행됨에 따라 그림 50과 같이 연합학습 라운드 별 모델 성능 지표 변화를 꺾은선 그래프 형태로 확인할 수 있다.

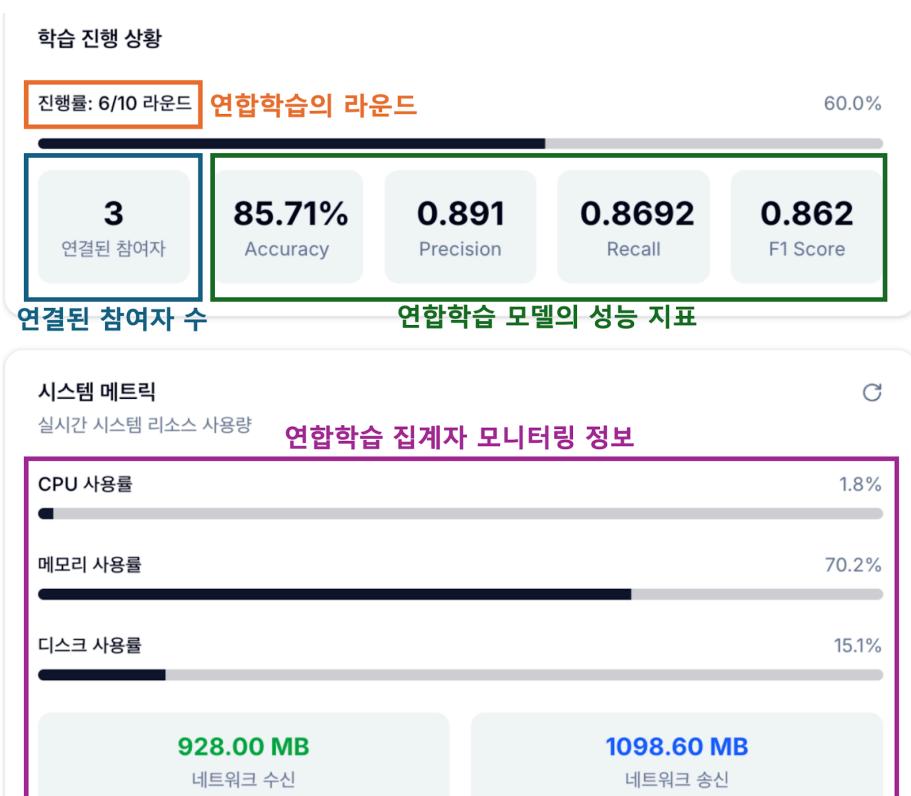


그림 49. 집계자 상세 보기 화면 - 학습 진행 상황 및 시스템 메트릭

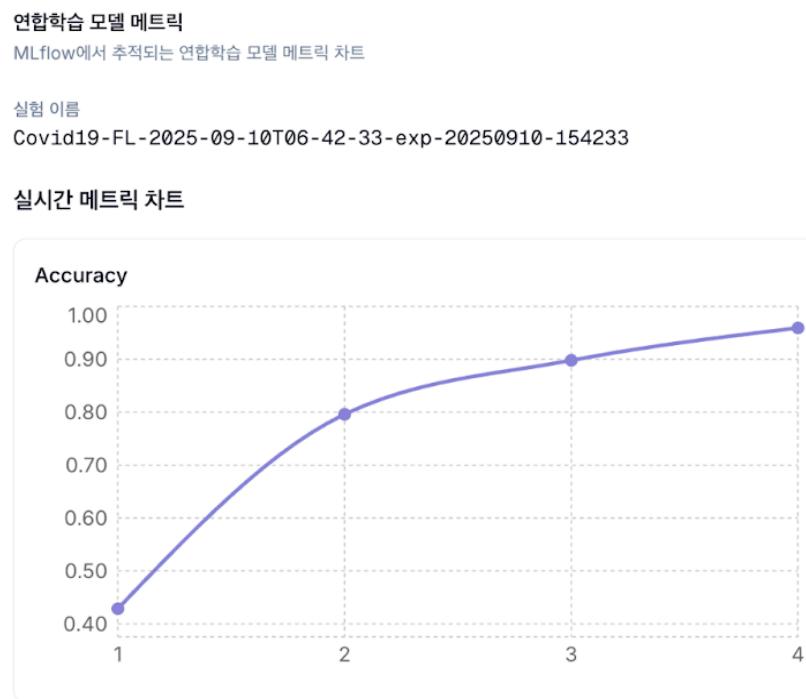


그림 50. 집계자 상세 보기 화면 - 연합학습 모델 메트릭

3.3.4. 연합학습 참여자(클러스터) 관리

본 기능은 프라이빗 클라우드 클러스터를 연합학습 참여자로 등록하고, 해당 참여자에 연합학습 작업 할당 및 모니터링할 수 있도록 지원한다. 이를 통해 사용자는 연합학습 참여자를 시스템에 등록하고 안정적인 학습을 수행할 수 있다.

① 연합학습 참여자 조회 및 등록

사용자는 각 연합학습 참여자의 기본 정보를 확인하고 신규 참여자를 추가하거나 기존 정보를 수정하는 등 기본적인 관리 작업을 수행할 수 있다.

참여자를 새로 추가할 경우, 이름 및 리전과 함께 OpenStack 기반 연합학습 참여자에 접근하기 위한 정보가 담긴 YAML 파일을 업로드한다. 시스템은 YAML 설정 파일을 기반으로 연합학습 참여자의 리소스 접근 검증을 수행하고, 검증에 성공한 경우에만 연합학습 참여자 추가를 완료한다.

② 연합학습 참여자 내 가상머신 모니터링

사용자는 그림 51과 같이 연합학습 참여자 내부의 가상머신의 실시간 리소스 상태를 모니터링 할 수 있다. 이 대시보드는 Grafana의 패널을 페이지 내에 임베딩한 것으로, CPU, 메모리 등의 현황을 제공하여 사용자가 연합학습 참여자 가상머신의 자원 상태를 파악할 수 있도록 돋는다.

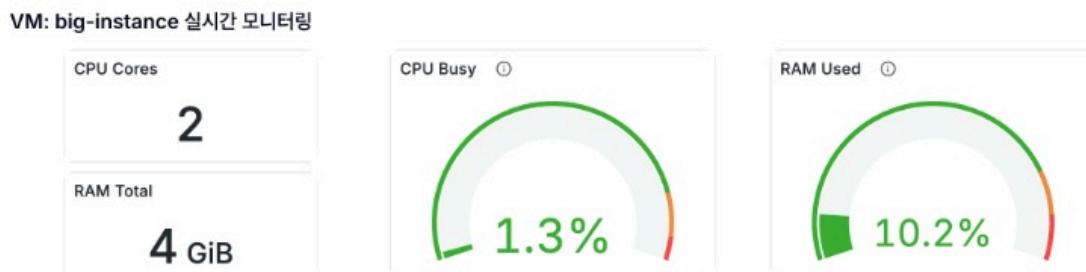


그림 51. 연합학습 참여자 가상머신 모니터링 화면

③ 연합학습 동적 태스크 오케스트레이션

연합학습 참여자는 기관 내부 프라이빗 클라우드 환경에서 여러 대의 가상머신을 운영한다. 따라서 연합학습 작업 수행 시, 연합학습 작업을 할당할 가상머신을 선택하는 과정이 필수적이다. 단순히 무작위로 가상머신을 선택한 경우, 성능이 낮아 학습을 수행할 수 없거나, 상태가 비정상적인 가상머신에 작업이 할당될 수 있다. 이는 연합학습 참여자가

연합학습에 참여하지 못하는 문제가 발생하거나 연합학습 실패로 이어질 수 있다.

이를 해결하기 위해 본 연구에서는 1) 자원 기반 필터링, 2) 원형 큐(Circular queue) 기반 가상머신 선택, 3) 실패 시 재시도로 이어지는 동적 태스크 오케스트레이션 알고리즘을 제안하였다. 알고리즘의 주요 절차는 다음과 같다.

- 1) 자원 기반 필터링: 모니터링되는 CPU, RAM, 디스크 용량, 동작 상태(Active, Inactive) 정보를 기반으로 미리 정의된 조건($vCPU \geq 2$, $RAM \geq 2GB$, $Disk \geq 10GB$, 상태=Active)을 만족하지 못하는 가상머신은 후보에서 제외한다.
- 2) 원형 큐 기반 작업 할당: 필터링을 통과한 가상머신들을 리스트로 구성하고, 순차적으로 가상머신에 작업을 분배한다.
- 3) 실패 시 재시도: 원형 큐 방식으로 작업이 할당된 가상머신이 실행 과정에서 오류 발생 혹은 자원 부족으로 인한 작업 불가가 발생할 경우 시스템은 자동으로 다른 가상머신에 작업을 재할당하고 연합학습 작업을 재수행한다. 이를 통해 개별 가상머신 장애로 인한 연합학습 참여자 이탈률을 최소화 하고, 전체 연합학습의 실패율을 줄일 수 있도록 한다. 참여자 이탈률이란 연합학습 수행 시 특정 참여자가 배정받은 가상머신에서 학습 작업을 정상적으로 실행하지 못하고 이탈한 비율을 말한다.

위 과정이 완료되면 그림 52와 같이 선택된 가상머신을 확인할 수 있다.

가상머신 목록

×

EU-Cluster 클러스터의 가상머신 목록

이름	상태	스펙 (CPU/RAM/Disk)	IP 주소
vm-001	ACTIVE Running	large CPU: 4 vCPU RAM: 8.0 GB Disk: 40 GB	34.97.203.78 private
vm-002	ACTIVE Running	medium CPU: 2 vCPU RAM: 4.0 GB Disk: 20 GB	34.15.453.54 private

그림 52. 동적 태스크 오케스트레이션 완료 화면

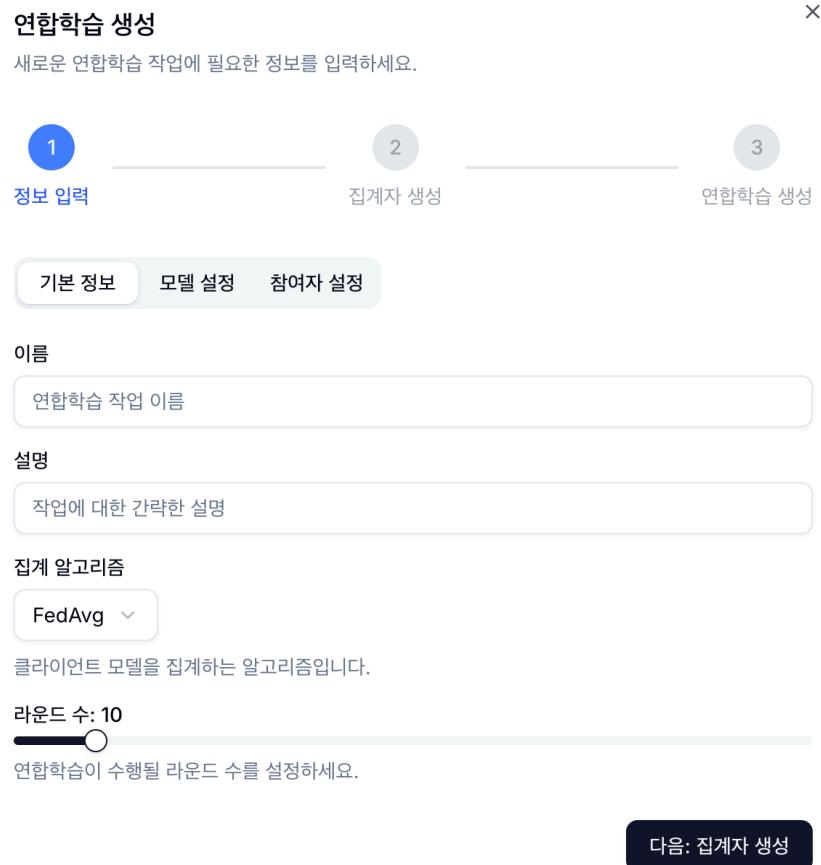


그림 53. 연합학습 기본 정보 입력

3.3.5. 연합학습 관리

① 연합학습 생성

연합학습 생성은 1) 연합학습 정보 입력, 2) 연합학습 집계자 배포, 3) 연합학습 실행의 3단계로 이루어진다.

- 1) 연합학습 정보 입력: 그림 53과 같이 기본 정보, 모델 설정, 참여자 설정을 시스템에 입력한다. 기본 정보 입력은 연합학습에 필요한 정보를 입력하는 과정이다. 사용자는 연합학습 작업의 이름, 설명, 집계 알고리즘(e.g. FedAvg, FedProx 등) 그리고 총 수행 라운드 수를 지정한다. 이어서 모델 설정 탭에서 연합학습 모델을 입력한다. 사용자는 모델 유형(이미지 분류, 자연어 처리 등)과 평가 기준(Accuracy, F1-Score, Precision, Recall)를 선택한 후 PyTorch 또는 TensorFlow 기반의 모델 정의 파일을 업로드한다. 다음으로 참여자 설정 탭에서 해당 연합학습 작업에 참여시킬

연합학습 참여자를 선택한다.

- 2) 연합학습 집계자 배포: 3.3.3. 절에서 제시한 ① 집계자 최적화 과정을 통해 선택한 집계자 배포 리전 및 스펙 선택 후에는 집계자 배포 단계로 넘어간다. 먼저 시스템은 사용자가 사전에 등록한 클라우드 인증 정보를 이용하여 클라우드 인증을 수행한다. 인증에 성공하였을 경우 IaC도구인 Terraform을 활용하여 집계자 생성에 필요한 리소스(VPC, Keypair, 보안그룹 등)를 생성한 후 집계자를 배포한다. 연합학습 집계자 배포가 정상적으로 완료되면 그림 54와 같이 집계자 정보와 연합학습 정보가 요약 표시되며, 배포 단계(자격증명 조회, 키페어 생성, Terraform 실행 등)의 진행률을 실시간으로 확인할 수 있다.

⌚ 배포 완료!

선택하신 aws 인스턴스로 집계자를 배포하고 있습니다

⚡ 선택된 배포 스펙
최적화에서 선택하신 인스턴스 구성

순위	#1	vCPU	2 코어
클라우드 제공자	aws	메모리	4.0GB
리전	ap-northeast-2	지연시간	12.33ms
인스턴스 타입	t4g.medium	월 예상 비용	₩38,938

연합학습 정보
배포할 연합학습 작업 정보

작업 이름	covid19
알고리즘	fedavg
라운드 수	5회
참여자 수	3명
모델 파일	testFL.py

⌚ 배포 진행 상황

집계자가 성공적으로 생성되었습니다!

배포 진행률 100%

[대시보드로 이동](#) [연합학습 시작](#)

그림 54. 연합학습 집계자 배포 화면

- 3) 연합학습 실행: 집계자 배포가 완료되면 연합학습의 구성을 최종 확인한다. 배포된 집계자 정보 및 연합학습 작업 정보를 확인할 수 있다. 또한 참여자 목록에서 3.3.4. 절에서 제시한 ③ 연합학습 동적 태스크 오케스트레이션 과정을 통해 선택된 연합학습 참여자의 가상머신을 확인할 수 있다. 이후 연합학습이 실행되면, 배포한 연합학습 집계자와 선택된 연합학습 참여자의 가상머신에 연합학습 환경 설정 및 연합학습 실행 코드를 배포한 후 연합학습 실행 명령을 내린다. 생성된 연합학습을 선택하면 상세 정보를 확인할 수 있고, 연합학습 실행 로그 모니터링 및 글로벌 모델 페이지에 접근할 수 있다. 그림 55는 “로그 확인” 버튼을 클릭한 후 연합학습 실행 로그를 모니터링하는 화면을 나타낸 것이다.

```
To view usage and all available options, run:  
$ flower-superlink --help  
  
Using 'start_server()' is deprecated.  
  
This is a deprecated feature. It will be removed entirely in future versions of Flower.  
-[92mINFO -[0m: Starting Flower server, config: num_rounds=10, no round_timeout  
-[92mINFO -[0m: Flower ECE: gRPC server running (10 rounds), SSL is disabled  
-[92mINFO -[0m: [INIT]  
-[92mINFO -[0m: Using initial global parameters provided by strategy  
-[92mINFO -[0m: Starting evaluation of initial global parameters  
-[92mINFO -[0m: Evaluation returned no results ('None')  
-[92mINFO -[0m:  
-[92mINFO -[0m: [ROUND 1]  
-[92mINFO -[0m: configure_fit: strategy sampled 1 clients (out of 1)  
-[92mINFO -[0m: aggregate_fit: received 1 results and 0 failures  
-[92mINFO -[0m: Sampling failed: number of available clients (1) is less than number of requested clients (2).  
-[92mINFO -[0m: configure_evaluate: no clients selected, skipping evaluation  
-[92mINFO -[0m:  
-[92mINFO -[0m: [ROUND 2]  
-[92mINFO -[0m: configure_fit: strategy sampled 1 clients (out of 1)
```

그림 55. 연합학습 로그 모니터링 화면

3.3.6. 글로벌 모델 관리

연합학습이 실행되면, 시스템은 연합학습 라운드별로 생성된 글로벌 모델의 저장을 수행하고 사용자는 글로벌 모델 평가지표 모니터링 및 사용자 요구사항 기반 최적의 성능을 가지는 모델을 다운로드 할 수 있다.

① 모델 저장

연합학습의 각 라운드가 끝날 때마다 갱신되는 글로벌 모델은 연합학습 집계자와 같은 클라우드 및 같은 리전의 객체 스토리지(AWS: S3, GCP: GCS)에 자동으로 저장된다. 이는 글로벌 모델의 버전 관리와 보존을 위함으로, 학습 인스턴스의 상태와 관계없이 언제든 특정 라운드의 모델을 안전하게 조회하고 접근할 수 있도록 보장한다.

② 모델 평가지표 모니터링

그림 56은 전체적으로 실행한 연합학습의 개요를 보여주는 화면이다. 사용자가 설정한 총 라운드 중에서 현재까지 완료한 라운드 수를 확인할 수 있고, 연합학습을 생성 시 사용자가 설정한 모델 평가 지표를 기준으로 전체적인 개요를 확인할 수 있다.

③ 최적의 글로벌 모델 다운로드

시스템은 모든 연합학습 라운드가 완료된 후, 연합학습 생성 시 사용자가 설정한 평가 지표를 기준으로 가장 높은 성능을 기록한 라운드의 모델을 자동으로 선별한다. 연합학습이 종료된 후, 사용자는 그림 56의 “최고 성능 모델 다운로드” 버튼을 클릭하여 모델을 다운로드할 수 있다.

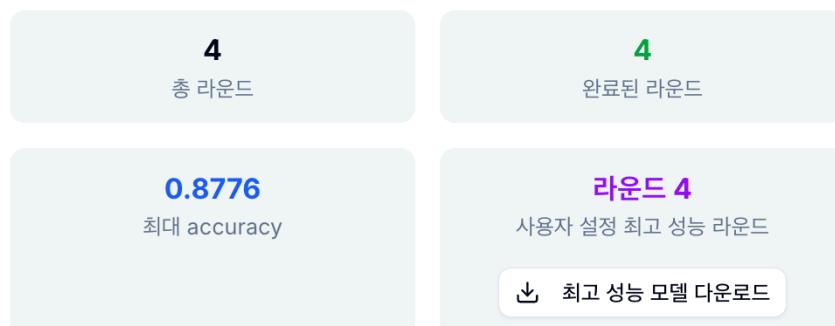


그림 56. 글로벌 모델 관리 개요

4. 연구 결과 분석 및 평가

4.1. 사례 연구 - 코로나19 진단 연합학습 모델 생성

본 사례 연구는 구축한 멀티 클라우드 인프라 기반 연합학습 환경 구축 플랫폼의 실질적 활용 가능성을 검증하기 위해, 폐 엑스레이 이미지를 기반으로 한 코로나19 진단 모델을 학습하는 과정을 다룬다. 제안 플랫폼은 퍼블릭 클라우드 - 프라이빗 클라우드 계층형 구조를 적용하여, 퍼블릭 클라우드에는 연합학습 집계자를, 프라이빗 클라우드에는 연합학습 참여자를 각각 배치하였다. 집계자는 모델 파라미터의 수집과 통계 작업을 담당하며, 참여자는 실제 학습 데이터를 보관하고 로컬 학습을 수행한다.

4.1.1 환경 설정

글로벌 의료 기관 협력 시나리오를 모사하기 위해, 프라이빗 클라우드 기반 연합학습 참여자는 유럽, 아시아, 미국 리전에 각각 배치되었다. 각 참여자는 폐 엑스레이 데이터를 개별적으로 보유하며, 이를 기반으로 로컬 학습을 수행하였다. 이러한 분산 배치는 실제

의료기관 간 데이터 공유가 제한되는 상황에서 학습 데이터 이동 없이 연합학습이 가능함을 보여준다.

4.1.2. 데이터셋 및 모델 구조

사례 연구에는 Kaggle의 Covid-19 Image Dataset[24]을 활용하였으며, 데이터의 예시는 그림 57과 같다. 데이터는 정상 환자, 코로나 19 확진 환자, 폐렴 환자로 구성되어 있다. 모델 구조는 일반적인 CNN 기반 진단 모델을 사용하였으며, 모델 크기는 약 26MB로 측정되었다.

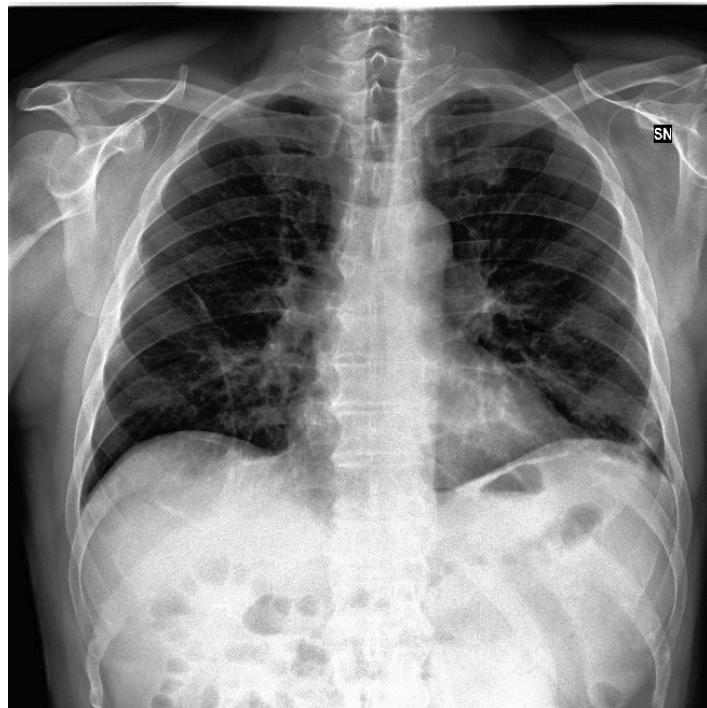


그림 57. Covid-19 Image Dataset 예시

4.2. 평가

4.2.1. 연합학습 집계자 최적화 평가

그림 58은 제안한 사용자 요구사항 기반 연합학습 집계자 최적화 기법의 결과를 나타낸다. 실험은 사용자 요구사항을 반영하여 비용-지연 시간 가중치를 변경하여 세 가지 시나리오를 진행하였다. 시나리오는 집계자 최적화를 수행하지 않은 최적화 전, 비용 우선 최적화(비용:지연 시간=9:1), 지연 시간 우선 최적화(비용:지연 시간=1:9)로 구성하였으며 동일한 연산 자원(2vCPU, 4GB RAM)을 기준으로 비교하였다.

실험 결과, 최적화 전 시나리오의 경우 라운드당 평균 학습 시간은 52.17s, 월간 클라우드 비용은 38,938 KRW으로 나타났다. 비용 우선 최적화 시나리오에서는 평균 학습 시간이 53s로 최적화 전과 큰 차이를 보이지 않으면서도 월간 비용이 23,026 KRW으로 약 41% 절감되었다. 반면, 지연 시간 우선 최적화 시나리오에서는 평균 학습 시간이 48.4s로 약 7% 단축되었으며, 월간 비용 또한 31,450KRW으로 약 19% 절감되는 효과가 확인되었다.

이와 같은 결과는 제안한 최적화 기법이 사용자 요구사항에 따라 효과적으로 적용 가능함을 보여준다. 비용을 중점적으로 고려하는 경우, 비용 효율성을 크게 향상시킬 수 있으며, 지연 시간을 우선하는 경우, 최적화 전보다 낮은 비용으로도 학습 속도를 크게 개선할 수 있음을 검증하였다. 이는 연합학습 집계자 배포 시 제안된 최적화 기법의 실질적 활용 가능성을 뒷받침한다.

Federated Learning Training Round Duration & Cost by Region/Instance

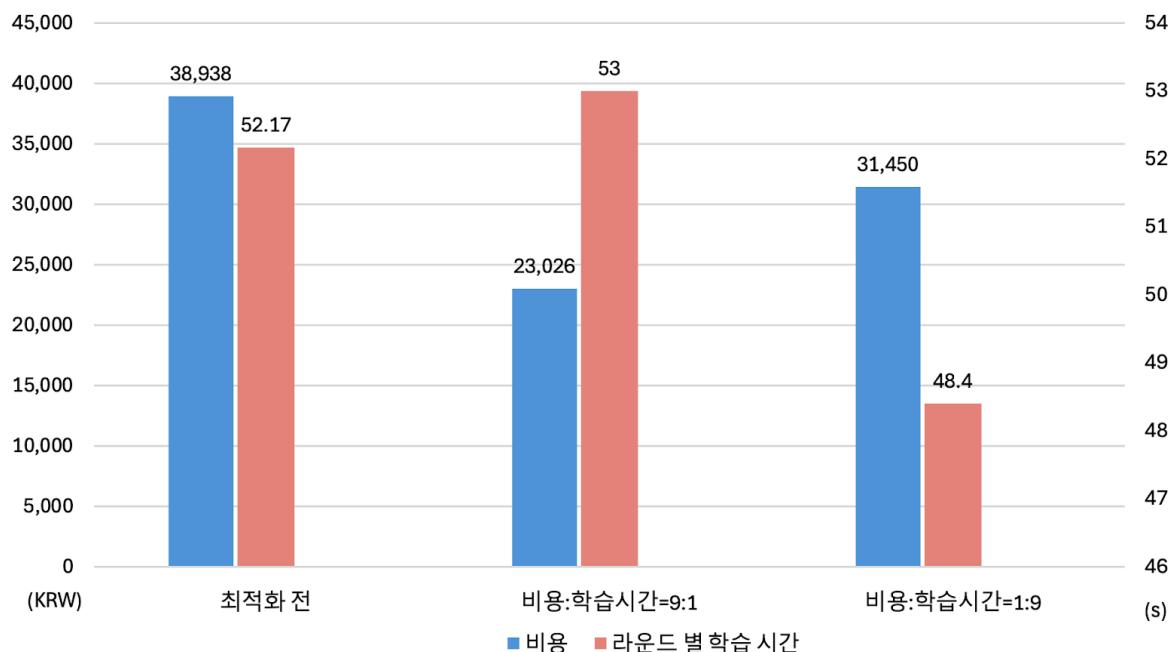


그림 58. 연합학습 집계자 위치 별 평균 라운드 수행 시간 및 월별 비용

4.2.2. 연합학습 동적 태스크 오케스트레이션 평가

본 사례 연구에서는 연합학습 동적 태스크 오케스트레이션 기법이 참여자 이탈률 및 연합학습 성공률에 미치는 효과를 검증하였다. 이를 위해 각 참여자가 프라이빗 클라우드 환경에서 5대의 가상머신을 운영하고 있으며, 총 20회의 연합학습을 수행하는 시나리오를 가정하였다. 검증은 3.3.4절에서 제시한 동적 태스크 오케스트레이션 알고리즘과 무작위

선택 알고리즘(가상머신 임의 선택 및 재시도 과정 없음)을 적용하여 수행하였다. 표 5는 5대의 가상머신 명세를 나타낸다. 이 중 vm-003과 vm-004는 Inactive 상태로 작업 수행이 불가능하며, vm-005는 학습을 수행하기에 연산 자원 및 디스크 크기가 충분하지 못하다. 따라서 실제 학습에 활용 가능한 가상머신 후보군은 vm-001과 vm-002로 제한된다.

20회 실험 결과, 무작위 선택 알고리즘은 가상머신 자원 상태를 고려하지 않고 연합학습 작업을 배정하여 참여자 이탈률이 60%, 성공률 40%에 그쳤다. 특히 비활성화된 vm-003, vm-004가 선택되면서 연합학습이 실패하는 사례가 다수 발생하였다.

반면, 제안한 동적 태스크 오케스트레이션 알고리즘은 자원 기반 필터링을 통해 최소 사양을 보장하고, 원형 큐 방식으로 가상머신을 균형 있게 활용하며, 장애 발생 및 자원 부족 시 재시도를 통해 작업을 즉시 다른 vm에 재할당하였다. 그 결과 참여자 이탈이 전혀 발생하지 않았고, 모든 실험에서 연합학습 작업이 성공하였다.

그림 59는 두 방법의 성능 차이를 시각적으로 비교한 결과이다. 무작위 선택 전략은 높은 이탈률과 낮은 연합학습 작업 성공률을 기록한 반면, 제안한 알고리즘은 100%의 연합학습 작업 성공률을 기록하며 연합학습 작업의 안정성을 확보하였다. 이는 본 연구의 동적 태스크 오케스트레이션이 연합학습 수행 시 장애 대응 및 안정적 학습 보장을 달성할 수 있음을 입증한 결과라고 할 수 있다.

표 5. 사례 연구의 연합학습 참여자 가상머신 구성 명세

VM ID	상태	vCPU	RAM(GB)	Disk(GB)
vm-001	Active	4	8	40
vm-002	Active	2	4	20
vm-003	Inactive	8	16	80
vm-004	Inactive	2	4	20
vm-005	Active	1	1	10

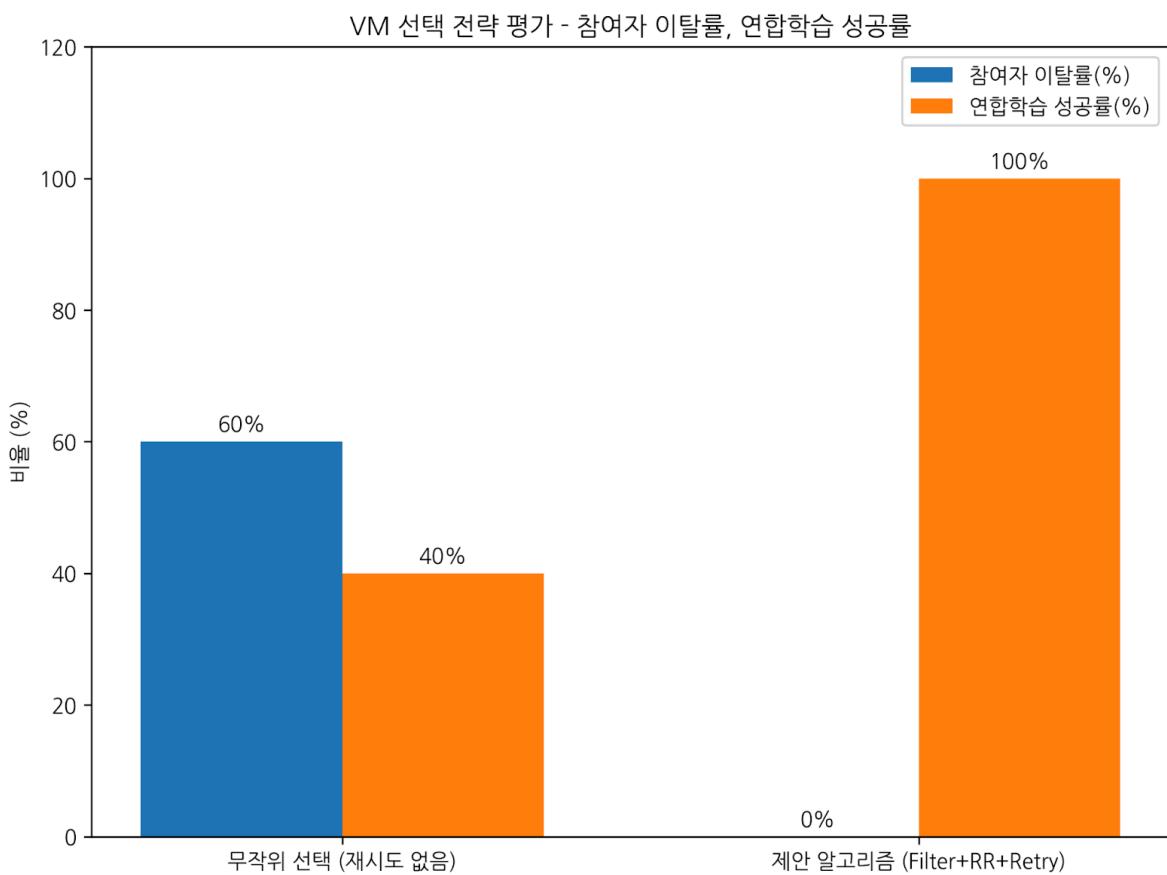


그림 59. 알고리즘에 따른 참여자 이탈률과 연합학습 성공률 그림

4.2.3 유사 시스템 비교 평가

본 절에서는 제안 시스템과 AWS SageMaker 기반 연합학습, GCP Vertex AI 기반 연합학습 및 오픈 소스 연합학습 프레임워크 및 플랫폼인 FedML을 비교 평가하였다. 평가 항목은 데이터 저장 위치, 데이터 신뢰 경계, 연합학습 집계자 최적화 지원 여부, 동적 태스크 오케스트레이션 지원 여부의 네 가지로 설정하였다.

데이터 저장 위치와 데이터 신뢰 경계는 시스템의 보안성과 신뢰성을 결정짓는 핵심 지표이므로 본 연구의 평가 항목으로 설정하였다. 데이터 저장 위치는 민감 데이터가 기관 내부에 유지되는지, 혹은 퍼블릭 클라우드 스토리지로 확장되는지에 따라 보안 위협 및 프라이버시 준수 여부가 달라진다. 의료 및 금융 도메인과 같이 데이터 프라이버시가 중요한 도메인에서는 데이터 저장 위치가 곧 시스템의 보안 수준을 판단하는 기준이 된다.

한편, 데이터 신뢰 경계(Trust Boundary)는 시스템 내외부를 구분하는 보안적 경계로, 공격 표면과 직결된다. 실제 STRIDE 기반 위협 분석 연구에 따르면, 사용자와 시스템 간 신뢰 경계에서 가장 많은 취약점이 발견되었으며, 이는 외부 인터넷 노출에 따른 Man-in-the-Middle, 세션 탈취 등 다양한 보안 위협과 밀접하게 연관되어 있다[25]. 따라서 데이터 저장 위치와 데이터 신뢰 경계는 연합학습 플랫폼의 보안성 평가에 있어 필수적으로 고려되어야 한다.

연합학습 집계자의 배포 위치 및 스펙 최적화는 연합학습 시스템의 비용 효율성과 학습 속도를 결정하는 중요한 요인이다. 집계자는 모든 연합학습 참여자로부터 모델 파라미터를 수신받고 통합하기 때문에, 집계자의 배포 위치가 네트워크 지연시간에 따른 학습 속도 직접적인 영향을 미친다[26]. 또한 동일한 연산 자원을 가지는 경우라도 인스턴스 유형에 따라 가격 차이가 발생하며, 이는 장기간 운영되는 학습 프로세스에서 상당한 비용 편차로 이어질 수 있다. 이에 따라 연합학습 집계자 최적화 지원 여부를 유사 시스템 비교 평가의 항목으로 설정하였다.

동적 태스크 오케스트레이션은 연합학습 참여자의 가상머신의 자원 상태에 따라 동적으로 작업을 할당하는 것을 의미한다. 자원 상태를 고려하지 않은 작업 할당의 경우 가상머신의 자원이 학습을 수행하기에 부족하거나, 가상머신의 장애로 인해 학습을 수행할 수 없는 상태가 발생할 수 있다. 따라서 동적 태스크 오케스트레이션 지원 여부는 연합학습의 안정성 확보에 중요한 요소이므로 평가 항목으로 설정하였다.

데이터 저장 위치와 데이터 신뢰 경계 항목의 경우, AWS SageMaker 기반 연합학습과 GCP Vertex AI 기반 연합학습은 모두 학습 데이터를 퍼블릭 클라우드 스토리지(AWS S3, GCP GCS 등)에 저장하는 구조를 채택한다. 이로 인해 데이터가 기관 외부 인프라로 확장되며, 신뢰 경계 또한 기관 내부에서 퍼블릭 클라우드까지 넓어진다. 반면 FedML은 참여자의 로컬 환경에 데이터를 보관하므로 외부 노출 위험을 상대적으로 줄일 수 있으며, 제안 시스템은 모든 학습 데이터를 기관 내부 프라이빗 클라우드에만 보관하므로 데이터 유출 위험을 근본적으로 차단한다. 이에 따라 신뢰 경계가 기관 내부에 유지된다.

연합학습 집계자 최적화와 동적 태스크 오케스트레이션 지원 여부의 경우, AWS

SageMaker 기반 연합학습과 GCP Vetex AI 기반 연합학습, FedML 모두 지원하지 않는다. 반면 제안 시스템은 3.3.3.절 ①에서 제시한 바와 같이 연합학습 집계자 최적화를 지원하고, 3.3.4.절 ③에서 제시한 바와 같이 동적 태스크 오케스트레이션을 지원한다.

표 4. 유사 시스템 비교 평가

구분	제안 시스템	AWS SageMaker 기반 연합학습	GCP Vertex AI 기반 연합학습	FedML
학습 데이터 저장 위치	프라이빗 클라우드 내부	AWS 내부 스토리지(S3 등)	GCP 내부 스토리지(GCS 등)	참여자 로컬 환경
데이터 신뢰 경계	기관 내부	기관 내부 및 AWS 클라우드	기관 내부 및 GCP 클라우드	참여자 로컬 환경
연합학습 집계자 최적화	지원 (3.3.3.절 ①)	미지원	미지원	미지원
동적 태스크 오케스트레이션	지원 (3.3.4. 절 ③)	미지원	미지원	미지원

5. 결론 및 향후 연구 방향

본 연구에서는 기존 클라우드 기반 연합학습 플랫폼이 가지는 세 가지 문제점, 1) 비용 및 지연시간 최적화 부재, 2) 보안 취약성 및 데이터 프라이버시 문제, 3) 동적 태스크 오케스트레이션 미지원 문제를 해결하기 위해 멀티 클라우드 인프라 기반의 연합학습 환경 구축 플랫폼을 제안하였다.

제안한 시스템은 퍼블릭 클라우드에 집계자를 배치하고, 프라이빗 클라우드에 연합학습 참여자를 배치하는 계층형 아키텍처로 구현하였으며, 이를 기반으로 사례 연구를 수행하였다. 사례 연구에서는 Kaggle의 Covid 19 Image Dataset를 활용하여 CNN 진단 모델을 생성하였다. 연합학습 참여자는 유럽, 아시아, 미국 리전에 배치하여 지리적으로 분산된 환경을 구축하여 글로벌 의료 기관 협력 시나리오를 모사하였다.

사례 연구 결과, 집계자 비용 우선 최적화 시나리오에서는 약 41%의 비용 절감, 집계자 지연시간 우선 최적화 시나리오에서는 약 7%의 학습 속도 향상과 19%의 비용 절감 효과가

확인되었다. 또한 동적 태스크 오케스트레이션 알고리즘을 적용한 경우, 가용하지 않은 VM 선택으로 인한 참여자 이탈 문제가 완전히 해소되었으며, 모든 학습 작업에서 100% 성공률을 달성하였다. 이러한 결과는 제안한 시스템이 데이터 보안성을 유지하면서도 비용·학습 속도 최적화와 안정적인 학습 수행을 달성할 수 있음을 실험적으로 입증한다.

향후에는 본 연구에서 수행한 비용 및 지연시간 최적화를 확장하여 에너지 효율, 네트워크 사용량 등 복합적인 지표를 반영하는 최적화 전략을 연구할 계획이다. 이를 위해 강화학습(Reinforcement Learning) 기반의 지능형 오케스트레이션 기법을 적용하여, 환경변화에 능동적으로 적응하는 자율적 스케줄링 방식을 마련할 계획이다.

아울러, 연합학습 참여자의 리소스 효율성 및 공정성(Fairness) 보장을 위한 방법을 연구할 계획이다. 이를 위해 각 참여자의 데이터 품질, 참여 빈도, 네트워크 상태 등을 종합적으로 고려한 다차원적 연합학습 스케줄링 알고리즘을 설계하고자 한다. 이러한 접근을 통해 특정 기관에 과부하가 집중되거나 자원 활용의 불균형이 발생하는 문제를 완화하고, 전체 학습 과정의 성능과 안정성을 동시에 강화하는 방안을 마련할 예정이다.

6. 구성원 별 역할

표 6 구성원 별 역할

학 번	성 명	구성원 별 역할
20205595	전진혁	<ul style="list-style-type: none">멀티 클라우드 연동: 멀티 클라우드 서비스 제공을 위한 퍼블릭 클라우드(AWS, GCP) 연동 기능 개발연합학습 관리 기능 개발: 연합학습 관리 API 및 파이프라인 설계 및 구현연합학습 참여자 모니터링: 연합학습 참여자 가상머신의 자원 사용량 실시간 분석 및 시각화 기능 구현연합학습 환경 구축: 퍼블릭 클라우드 - 프라이빗 클라우드 연합학습 수행을 위한 인프라 환경 구축연구 결과 분석: 유사 시스템 비교 평가, 비용·지연시간 최적화 및 동적 태스크 오케스트레이션 실험 수행 및 결과 분석
202155526	김민경	<ul style="list-style-type: none">멀티 클라우드 IaC 자동화: Terraform을 활용한 퍼블릭 클라우드 인프라 자동화 구축, 환경별 자격 증명 관리 시스템 및 동적 키 페어 핸들링 구현연합학습 집계자 모니터링: 연합학습 집계자 가상머신의 자원 사용량 실시간 분석 및 시각화 기능 구현동적 태스크 오케스트레이션: 프라이빗 클라우드 가상머신 환경 구축, 오픈스택 환경에서 테스트
201924474	박재일	<ul style="list-style-type: none">클라우드 비용 정보 및 지연 시간 측정: AWS, GCP의 리전별 인스턴스 비용 정보 수집 및 리전 간 지연 시간 측정연합학습 집계자 최적화 구현: 클라우드 비용 정보와 지연 시간을 통한 집계자 배포 최적화 구현MLOps 적용: MLflow를 통한 연합학습 모델 메트릭 측정, 모델 아티펙트 저장, 사용자 요구에 맞춘 모델 다운로드 지원

7. 참고 문헌

- [1] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, PMLR 54, pp. 1273-1282, Apr. 2017.
- [2] J. Dean, G. Corrado, R. Monga, K. Chen, M. Devin, Q. Le, M. Mao, M. Ranzato, A. Senior, P. Tucker, K. Yang, and A. Ng, "Large Scale Distributed Deep Networks," *Proceedings of the 25th International Conference on Neural Information Processing Systems (NeurIPS)*, pp. 1223-1231, Dec. 2012.
- [3] L. Yuan, et al., "Decentralized federated learning: A survey and perspective," IEEE Internet of Things Journal, vol. 11, no. 21, pp. 34617-34638, 2024.
- [4] J. Hong, T. Dreibholz, J. A. Schenkel, and J. A. Hu, "An Overview of Multi-cloud Computing," Web, Artificial Intelligence and Network Applications, pp. 1055-1068, 2019.
- [5] J. Alonso, et al., "Understanding the challenges and novel architectural models of multi-cloud native applications - a systematic literature review," Journal of Cloud Computing, vol. 12, no. 1, p. 6, 2023.
- [6] AWS, "클라우드 컴퓨팅 서비스 - Amazon Web Services(AWS)," [Online]. Available: <https://aws.amazon.com>. [Accessed: Sep. 11, 2025].
- [7] Microsoft, "Azure란? | Microsoft Azure," [Online]. Available: <https://azure.microsoft.com/ko-kr/resources/cloud-computing-dictionary/what-is-azure>. [Accessed: Sep. 11, 2025].
- [8] Google, "클라우드 컴퓨팅 서비스 | Google Cloud," [Online]. Available: <https://cloud.google.com>. [Accessed: Sep. 11, 2025].
- [9] IBM, "Welcome to IBM Federated Learning — ibm-federated-learning," [Online]. Available: <https://ibmfl-api-docs.res.ibm.com>. [Accessed: Sep. 11, 2025].
- [10] AWS, "Amazon SageMaker AI Documentation," [Online]. Available: <https://docs.aws.amazon.com/sagemaker>. [Accessed: Sep. 11, 2025].
- [11] Google, "Vertex AI 문서 | Google Cloud," [Online]. Available: <https://cloud.google.com/vertex-ai/docs>. [Accessed: Sep. 11, 2025].

- [12] FedML, “TensorOpera® Documentation,” [Online]. Available: <https://doc.fedml.ai/> [Accessed: Sep. 11, 2025].
- [13] J. Proudman, “Openstack Docs: 2025.1,” [Online]. Available: <https://docs.openstack.org>. [Accessed: Sep. 11, 2025].
- [14] Prometheus Authors, “Overview | Prometheus,” [Online]. Available: <https://prometheus.io/docs>. [Accessed: Sep. 13, 2025].
- [15] Grafana, “Technical documentation | Grafana Labs,” [Online]. Available: <https://grafana.com/docs>. [Accessed: Sep. 11, 2025].
- [16] Hashicorp, “Terraform | Terraform | HashiCorp Developer,” [Online]. Available: <https://developer.hashicorp.com/terraform>. [Accessed: Sep. 11, 2025].
- [17] Flower, “Flower Documentation,” [Online]. Available: <https://flower.dev/docs>. [Accessed: Sep. 11, 2025].
- [18] MLflow, “MLflow,” [Online]. Available: <https://mlflow.org>. [Accessed: Sep. 11, 2025].
- [19] PyTorch, “PyTorch documentation - PyTorch 2.8 documentation,” [Online]. Available: <https://pytorch.org/docs>. [Accessed: Sep. 11, 2025].
- [20] TensorFlow, “TensorFlow,” [Online]. Available: <https://www.tensorflow.org>. [Accessed: Sep. 11, 2025].
- [21] Docker, “Docker Docs,” [Online]. Available: <https://docs.docker.com>. [Accessed: Sep. 11, 2025].
- [22] J.Liu, and X.Chen, “An Improved NSGA-II Algorithm Based on Crowding Distance Elimination Strategy,” International Journal of Computational Intelligence Systems, Vol.12, No.2, pp.513-518, 2019.
- [23] Z.Osika, P.Koch, and T.Wagner, “What lies beyond the Pareto front? A survey on decision-support methods for multi-objective optimization,” in arXiv preprint, pp.1-9, 2023.
- [24] Kaggle, “Covid-19 Image Dataset,” [Online]. Available: <https://www.kaggle.com/datasets/pranavraikokte/covid19-image-dataset/data>. [Accessed: Sep. 11, 2025].

-
- [25] N. Gavric, A. Shalaginov, A. Andrushevich, A. Rumsch, and A. Paice, "Enhancing International Data Spaces Security: A STRIDE Framework Approach," *Preprints*, 2024.

[26] ALI-POUR, Amir, et al. Towards a distributed federated learning aggregation placement using particle swarm intelligence. arXiv preprint arXiv:2504.16227, 2025.