

MODULE 09 · 심화

분산 클라우드 — 데이터가 있는 곳으로 클라우드가 온다

모듈 0에서 우리는 "데이터가 밖으로 나간다"와 "폐쇄망에서는 접속이 안 된다"는 두 벽 때문에 로컬 AI를 선택했습니다. 그런데 클라우드 업계도 같은 문제를 알고 있고, 그들의 답이 **분산 클라우드**입니다. 우리가 하는 로컬 sLM과 어디가 닮았고 어디가 다른지, 이 지형을 알면 인프라 논의에서 길을 잃지 않습니다. (시장 상황은 빠르게 변합니다 — 이 내용은 2026년 상반기 기준입니다.)

01 클라우드의 딜레마 — 편리함과 규제 사이

퍼블릭 클라우드(AWS, Azure, Google Cloud처럼 누구나 빌려 쓰는 원격 컴퓨팅)는 편리하지만, 규제가 있는 조직에는 근본적인 고민이 있습니다. 데이터가 어느 나라의 어느 서버에 저장되는지, 누구의 법이 적용되는지를 조직이 완전히 통제할 수 없다는 것입니다. 이것을 **데이터 주권(data sovereignty)** 문제라고 부릅니다 — 데이터가 저장·처리되는 위치와 조건을 법적으로 통제할 수 있어야 한다는 요구입니다.

그래서 클라우드 업계가 내놓은 답이 방향을 뒤집는 것입니다. **데이터를 클라우드로 보내는 대신, 클라우드를 데이터가 있는 곳으로 가져온다.** 어디서 들어본 논리죠? 모듈 0에서 우리가 모델을 우리 컴퓨터 안으로 가져온 것과 정확히 같은 발상입니다. 규모만 다를 뿐입니다.

02 헛갈리는 네 형제 — 분산·멀티·하이브리드·엣지

비슷해 보이는 용어 네 개를 먼저 구분합니다. 핵심 질문은 두 가지입니다: **어디에 있는가**, 그리고 **누가 운영을 책임지는가**.

개념	무엇	운영 책임
분산 클라우드	한 클라우드 회사의 서비스를 고객의 데이터센터·지역 거점 등 여러 위치에 배치	클라우드 회사가 계속 책임 (업데이트·보안 포함)

멀티클라우드	AWS·Azure·Google 등 여러 회사를 병행 사용	조직이 각각 따로 관리
하이브리드 클라우드	자체 전산실(온프레미스) + 퍼블릭 클라우드를 연결해 사용	조직이 연결·운영 관리
엣지 클라우드	공장·병원·현장처럼 지연시간이 치명적인 곳 바로 옆에 소형 배치	배치 형태에 따라 다름

💡 비유

은행에 비유하면 — 여러 은행에 계좌를 나눠 만드는 것이 멀티클라우드, 집 금고와 은행을 같이 쓰는 것이 하이브리드입니다. 분산 클라우드는 **은행이 우리 동네에 지점을 내주는 것**입니다. 금고(데이터)는 동네를 떠나지 않지만, 지점 운영·보안·시스템 관리는 여전히 은행 본사가 책임집니다.

시장조사기관 Gartner는 분산 클라우드를 "퍼블릭 클라우드 서비스를 여러 물리적 위치에 분산 배치하되, 운영·거버넌스·업데이트는 원 제공자가 책임지는 방식"으로 정의하고, 최근에는 온프레미스·클라우드·엣지를 통합 설계하는 **분산 하이브리드 인프라**라는 더 큰 틀로 묶고 있습니다. 흐름의 요지는 하나입니다 — "무조건 클라우드로"에서 "**데이터와 규제에 맞는 위치로**"로의 전환.

03 실제 제품들 — 폐쇄망까지 내려온 클라우드

- ✓ **AWS Outposts** — AWS의 서버 랙을 고객 전산실에 직접 설치하고, 클라우드와 동일한 방식으로 쓰게 해줍니다. 가장 성숙한 사례로, 의료·금융·방위처럼 "데이터는 내부에, 도구는 클라우드처럼"이 필요한 산업이 주 고객입니다.
- ✓ **Microsoft Azure Arc / Azure Local** — Azure의 관리 화면 하나로 온프레미스와 타사 클라우드 자원까지 함께 관리하는 접근입니다. Azure Local은 **인터넷이 단절된 환경**에서의 운영을 지원하는 방향으로 진화하고 있습니다.
- ✓ **Google Distributed Cloud** — 연결형·엣지형에 더해 **에어갭형(air-gapped, 외부망 완전 차단)** 배포를 공식 제공합니다. "폐쇄망에서는 클라우드를 못 쓴다"는 통념이 제품 수준에서 깨지고 있는 것입니다.

우리 연구회에선

"에어갭형 클라우드"라는 말에 주목하세요. 우리 업무망 같은 폐쇄 환경을 클라우드 회사들이 이제 정식 시장으로 취급한다는 뜻입니다. 다만 이런 제품은 도입 비용과 규모가 기관 단위의 결정이고, 우리 연구회가 하는 로컬 sLM은 **같은 원리를 장비 한 대 규모에서 지금 바로 실험하는 것**입니다.

04 AI와 만나면 — 추론이 데이터 곁으로 이동한다

모듈 0에서 본 흐름을 기억하세요 — AI 연산의 무게중심은 학습에서 **추론**으로 이동하고 있습니다. 그리고 추론은 데이터가 만들어지는 현장 가까이에서 돌수록 유리합니다: 응답이 빠르고(지연시간), 회선 비용이 줄고(대역폭), 데이터가 밖으로 나가지 않습니다(유출 위험). 그래서 AI 추론을 중앙 클라우드가 아닌 현장 근처에서 실행하는 **엣지 AI** 시장이 빠르게 크고 있습니다(시장조사 전망 기준, 2025~2030년 사이 약 3배 성장 예측).

여기에 모듈 7에서 본 **소버린 AI**(AI의 모델·데이터·인프라·운업을 자국/자기 조직의 통제 아래 두려는 흐름)가 겹칩니다. 국가 단위로는 자국 데이터센터와 자국 반도체로 AI 인프라를 세우려 하고, 조직 단위로는 폐쇄망 안에서 AI를 돌리려 합니다. 2025년에는 기업들이 클라우드 관리형 서비스 대신 **자체 데이터센터에서 컨테이너 인프라(Kubernetes)를 직접 운영하는 회귀 현상**이 보고되기도 했습니다 — 기술 독립성과 데이터 통제를 편의보다 앞세우기 시작한 신호입니다.

05 한국 지형 — K-클라우드, N2SF, 국가 AI 컴퓨팅 센터

- ✓ **K-클라우드 프로젝트** — 국산 AI 반도체(모듈 7의 K-NPU)를 개발하고 그것으로 국내 클라우드·데이터 센터를 구축하는 국가 프로젝트입니다. 국내 클라우드 3사(네이버·NHN·KT)와 NPU 기업들(리벨리온·퓨리오사 등)이 참여합니다.
- ✓ **공공 클라우드 정책 전환(N2SF)** — 2026년부터 공공 데이터를 중요도에 따라 등급을 나눠, 등급별로 다른 보안 수준의 클라우드 이용을 허용하는 방향으로 바뀝니다. "전부 최고 등급 폐쇄망" 일률 규제에서 **데이터 등급별 차등 배치**로 — 분산 클라우드의 사고방식이 정책에 들어온 것입니다.
- ✓ **국가 AI 컴퓨팅 센터** — 전남 해남 솔라시도에 대규모 AI 데이터센터가 추진 중이고(2027년 개설 목표), 국산 NPU 비중을 단계적으로 높이는 계획과 국산 NPU 성능 평가 기준(K-Perf) 공개가 이어지고 있습니다. 공공·국방 분야부터 국산 칩 실증을 확대하는 흐름입니다.

우리 연구회에선

모듈 7의 결론이 여기서도 통합니다 — 어떤 인프라가 오든 "그 위에서 어떤 모델이 우리 업무 요구를 만족하는가"를 측정할 수 있는 사람이 필요합니다. 골든셋과 실측(모듈 6)은 장비가 GPU든 NPU든, 클라우드든 폐쇄망이든 똑같이 쓰이는 능력입니다.

06 비판적 시각 — 분산돼도 제어판은 중앙에 있다

장밋빛 전망만 볼 일은 아닙니다. 2025년에는 대형 클라우드 장애가 유난히 많았습니다 — 한 대형 클라우드의 특정 리전 장애가 15시간 동안 1,000여 개 서비스를 멈춰 세운 사건을 포함해, 주요 사고들의 공통점은 데이터가 있는 계층이 아니라 제어층(control plane, 전체를 지휘하는 중앙 관리 계층)에서 터졌다는 것입니다.

- ✓ **인프라는 분산돼도 제어는 중앙** — 서버가 우리나라에 있어도, 그것을 지휘하는 제어판이 해외 리전에 있으면 그쪽 장애에 같이 묶입니다. 분산 클라우드를 검토할 때 "데이터 위치"만이 아니라 "제어층 위치와 단절 시 동작"을 물어야 하는 이유입니다.
- ✓ **복잡성은 공짜가 아니다** — 여러 위치·여러 클라우드를 섞을수록 운영 난이도가 올라갑니다. 멀티클라우드 투자 조직의 상당수가 기대한 성과를 얻지 못할 것이라는 전망(Gartner)도 같은 맥락입니다.

비유

동네 지점(분산 배치)이 생겨도 전산 시스템이 본사에 있으면, 본사 전산이 멈추는 날엔 지점 창구도 멈춥니다. "지점이 있으니 안심"이 아니라 "본사가 끊겨도 지점이 혼자 돌아가는가"를 물어야 합니다 — 에어갭형이 따로 존재하는 이유이고, 우리가 로컬 모델로 연습하는 "단절돼도 돌아간다"는 감각이 바로 그 질문입니다.

07 우리에게 의미 — 로컬 sLM은 분산의 최소 단위

이 지형도에서 가져갈 결론은 세 가지입니다.

- ✓ **우리 방향은 업계 흐름과 같은 편입니다.** "데이터 곁에서 추론한다"는 원칙을 클라우드 회사는 데이터 센터 규모로, 우리는 장비 한 대 규모로 실행하고 있을 뿐입니다.

- ✓ **정책이 등급별 차등 배치로 갑니다.** 모든 것을 폐쇄망에 두는 시대에서, 데이터 중요도별로 위치를 고르는 시대로 — "이 업무의 데이터는 어느 등급인가"를 판단하는 감각이 실무 역량이 됩니다.
- ✓ **단절 내성이 평가 기준입니다.** 클라우드든 로컬이든, "연결이 끊겨도 핵심 업무가 도는가"를 물으세요. 우리의 로컬 sLM 실습이 그 감각의 훈련장입니다.

↗ 심화 학습

용어의 원 출처 — **Gartner: Distributed Cloud 정의** · 영어. 분산 클라우드의 공식 정의.

데이터 주권 개념 정리 — **NetApp: 데이터 주권이란** · 한국어. 규제 관점의 기초 설명.

CHECK

이해했는지 확인해 봅시다

틀려도 괜찮습니다 — 오답을 고르면 왜 아닌지 설명이 나오고, 정답을 찾을 때까지 다시 고를 수 있습니다. 점수는 첫 시도 기준입니다.

문제 1 / 3

Q1. 분산 클라우드가 하이브리드·멀티클라우드와 구분되는 핵심은?

- ① 여러 클라우드 회사를 동시에 쓴다는 점
- ② 서비스가 여러 위치에 배치되어도 운영·보안 책임은 원 클라우드 제공자에게 남는다는 점
- ③ 자체 전산실 장비와 클라우드를 네트워크로 연결한다는 점
- ④ 인터넷 없이도 동작하도록 설계된 유일한 방식이라는 점

[← 이전](#)

모듈 8 · 추론 인프라의 최전선

[다음 →](#)

모듈 10 · 로컬 서빙의 다음 카드