



NetApp AIPOd Mini for ERAG - 인프라 준비 체크리스트

NetApp artificial intelligence solutions

NetApp
February 12, 2026

목차

NetApp AIPOd Mini for ERAG - 인프라 준비 체크리스트	1
비즈니스 및 사용 사례 준비 상태	1
기술 및 인프라 요구사항	1
소프트웨어 및 에코시스템 정렬	1
데이터 거버넌스 및 보안	1
운영 및 지원 고려 사항	2
상업 및 GTM 정렬	2

NetApp AI Pod Mini for ERAG - 인프라 준비 체크리스트

이 문서는 NetApp AI Pod Mini for Enterprise RAG의 포괄적인 인프라 준비 체크리스트를 설명하며, 배포 전 참고 자료로 활용할 수 있습니다.

비즈니스 및 사용 사례 준비 상태

- 해당 솔루션은 사업 부문별 성과(예: 생산성, 고객 서비스, 법률, 의료, 제조, 공공 부문)와 부합합니까?
- LLM 워크로드에 필요한 첫 토큰 획득 시간(TTFT) 및 지연 시간을 예측하십니까?
- 예상 동시성/사용자 부하를 알고 계십니까(예: RAG의 경우 2-워커 노드당 32명의 동시 사용자)?
- 주요 AI/GenAI 워크로드(RAG, 추론, 미세 조정, 부서별 LLM, 벡터 DB 통합)를 파악하십니까?
- 비용/성능 균형을 위해 CPU 기반 AI 옵션(OPEA, Intel Xeon)과 GPU 대안을 평가하고 계십니까?

기술 및 인프라 요구사항

- 데이터 파이프라인(데이터 준비, ETL, 벡터 DB로의 안전한 수집)이 준비되었습니까?
- 고가용성, 이중화 또는 DR 기능이 필요하십니까?
- Ubuntu Linux / Kubernetes / Red Hat OpenShift AI 스택 지원을 활용하고 계십니까?
- 네트워크 성능(워크로드에 따라 25~100GbE)을 검증하십니까?
- Kubernetes 영구 저장소를 위해 NetApp ONTAP + Trident CSI 드라이버로 스토리지가 프로비저닝되었습니까?
- 최소 3개의 컴퓨팅 노드(워커 2개, 컨트롤 플레인 1개)의 크기가 올바르게 지정되었습니까?

소프트웨어 및 에코시스템 정렬

- 컨테이너화된 앱이 제공된 Kubernetes 및 Helm 차트와 호환됩니까?
- 배포 예정인 벡터 데이터베이스는 무엇입니까(예: Milvus, pgvector)?
- 검색 증강 생성(RAG)을 위해 OPEA(Open Platform for Enterprise AI) 사전 통합이 필요하신가요?
- 하이브리드 클라우드 옵션(Cloud Volumes ONTAP, FSxN, Anthos, Azure Arc)을 활용하고 계십니까?
- 파트너 ISV 통합(ESRI, 의료 PACS, 금융/법률 ISV)이 필요하십니까?

데이터 거버넌스 및 보안

- Kubernetes에서 역할 기반 액세스 제어(RBAC)를 활성화하십니까?
- 데이터 보호 및 백업 계획(SnapMirror, SnapCenter, 랜섬웨어 방지)이 있습니까?
- 데이터 규정 준수 요구 사항(HIPAA, GDPR, FedRAMP, CJIS)을 파악하십니까?
- 프라이빗 AI 배포(에어갭, 온프레미스, 보안 영역)가 필요하십니까?

운영 및 지원 고려 사항

- 관리자는 Kubernetes, Trident CSI 및 OPEA 스택 배포에 대한 교육을 받았거나 사용할 수 있습니까?
- 멀티 테넌시(multi-tenancy)(부서, SLED 기관, 사업부)에 대한 지원이 필요하십니까?
- ONTAP System Manager, Cloud Insights, Prometheus/Grafana 등을 활용한 모니터링 및 관찰 가능성 확보 계획이 있습니까?
- 2일차 운영은 누가 담당하게 될까요(고객 IT, 파트너, 관리 서비스 제공업체)?

상업 및 GTM 정렬

- 부서별 AI 도입부터 전사적 AI 확장까지 단계별 로드맵이 있습니까?
- 다년간 예상 재무제표(TCO, ARR, 마진 상승률)를 보유하고 계신가요?
- 라이선스 업리프트 시나리오(벡터 DB, ISV 소프트웨어, AI ops 툴)가 명확하십니까?
- 파트너 인센티브(총판 마진, OEM/Intel 공동 투자)를 검토해 보셨습니까?
- 구매가 예산 주기(CapEx vs OpEx, 소비 모델)와 일치하십니까?
- 사이징 및 구축을 위한 서비스 파트너(Arrow, WWT, Presidio 등)가 있습니까?

저작권 정보

Copyright © 2026 NetApp, Inc. All Rights Reserved. 미국에서 인쇄된 본 문서의 어떠한 부분도 저작권 소유자의 사전 서면 승인 없이는 어떠한 형식이나 수단(복사, 녹음, 녹화 또는 전자 검색 시스템에 저장하는 것을 비롯한 그래픽, 전자적 또는 기계적 방법)으로도 복제될 수 없습니다.

NetApp이 저작권을 가진 자료에 있는 소프트웨어에는 아래의 라이선스와 고지사항이 적용됩니다.

본 소프트웨어는 NetApp에 의해 '있는 그대로' 제공되며 상품성 및 특정 목적에의 적합성에 대한 명시적 또는 묵시적 보증을 포함하여(이에 제한되지 않음) 어떠한 보증도 하지 않습니다. NetApp은 대체품 또는 대체 서비스의 조달, 사용 불능, 데이터 손실, 이익 손실, 영업 중단을 포함하여(이에 국한되지 않음), 이 소프트웨어의 사용으로 인해 발생하는 모든 직접 및 간접 손해, 우발적 손해, 특별 손해, 징벌적 손해, 결과적 손해의 발생에 대하여 그 발생 이유, 책임론, 계약 여부, 엄격한 책임, 불법 행위(과실 또는 그렇지 않은 경우)와 관계없이 어떠한 책임도 지지 않으며, 이와 같은 손실의 발생 가능성이 통지되었다 하더라도 마찬가지입니다.

NetApp은 본 문서에 설명된 제품을 언제든지 예고 없이 변경할 권리를 보유합니다. NetApp은 NetApp의 명시적인 서면 동의를 받은 경우를 제외하고 본 문서에 설명된 제품을 사용하여 발생하는 어떠한 문제에도 책임을 지지 않습니다. 본 제품의 사용 또는 구매의 경우 NetApp에서는 어떠한 특허권, 상표권 또는 기타 지적 재산권이 적용되는 라이선스도 제공하지 않습니다.

본 설명서에 설명된 제품은 하나 이상의 미국 특허, 해외 특허 또는 출원 중인 특허로 보호됩니다.

제한적 권리 표시: 정부에 의한 사용, 복제 또는 공개에는 DFARS 252.227-7013(2014년 2월) 및 FAR 52.227-19(2007년 12월)의 기술 데이터-비상업적 품목에 대한 권리(Rights in Technical Data -Noncommercial Items) 조항의 하위 조항 (b)(3)에 설명된 제한사항이 적용됩니다.

여기에 포함된 데이터는 상업용 제품 및/또는 상업용 서비스(FAR 2.101에 정의)에 해당하며 NetApp, Inc.의 독점 자산입니다. 본 계약에 따라 제공되는 모든 NetApp 기술 데이터 및 컴퓨터 소프트웨어는 본질적으로 상업용이며 개인 비용만으로 개발되었습니다. 미국 정부는 데이터가 제공된 미국 계약과 관련하여 해당 계약을 지원하는 데에만 데이터에 대한 전 세계적으로 비독점적이고 양도할 수 없으며 재사용이 불가능하며 취소 불가능한 라이선스를 제한적으로 가집니다. 여기에 제공된 경우를 제외하고 NetApp, Inc.의 사전 서면 승인 없이는 이 데이터를 사용, 공개, 재생산, 수정, 수행 또는 표시할 수 없습니다. 미국 국방부에 대한 정부 라이선스는 DFARS 조항 252.227-7015(b)(2014년 2월)에 명시된 권한으로 제한됩니다.

상표 정보

NETAPP, NETAPP 로고 및 <http://www.netapp.com/TM>에 나열된 마크는 NetApp, Inc.의 상표입니다. 기타 회사 및 제품 이름은 해당 소유자의 상표일 수 있습니다.