



# 아파치 기류 NetApp Solutions

NetApp  
May 10, 2024

# 목차

아파치 기류 .....	1
Apache Airflow Deployment .....	1
NetApp DataOps 툴킷을 공기 흐름과 함께 사용하십시오 .....	4

# 아파치 기류

## Apache Airflow Deployment

이 섹션에서는 Kubernetes 클러스터에 공기 흐름을 구축하기 위해 완료해야 하는 작업에 대해 설명합니다.



Kubernetes 이외의 플랫폼에 공기 흐름을 배포할 수 있습니다. Kubernetes가 아닌 다른 플랫폼에 공기 흐름을 배포하는 것은 이 솔루션의 범위를 벗어납니다.

### 필수 구성 요소

이 섹션에 요약된 배포 연습을 수행하기 전에 이미 다음 작업을 수행했다고 가정합니다.

1. Kubernetes 클러스터 작업이 이미 진행 중입니다.
2. Kubernetes 클러스터에 NetApp Astra Trident를 이미 설치하고 구성했습니다. Astra Trident에 대한 자세한 내용은 [참조하십시오 "Astra Trident 문서"](#).

### 제어 장치를 설치합니다

Kubernetes의 유명 패키지 매니저인 Helm을 사용하여 공기 흐름을 구축합니다. 공기 흐름을 배치하기 전에 배포 점프 호스트에 Helm을 설치해야 합니다. 배포 점프 호스트에 Helm을 설치하려면 [에 따르면 "설치 지침" 공식 Helm 문서](#).

### 기본 Kubernetes StorageClass를 설정합니다

공기 흐름을 구축하기 전에 Kubernetes 클러스터 내에서 기본 StorageClass를 지정해야 합니다. Airflow 배포 프로세스는 기본 StorageClass를 사용하여 새 영구 볼륨의 프로비저닝을 시도합니다. 기본 StorageClass로 지정된 StorageClass가 없으면 배포가 실패합니다. 클러스터 내에서 기본 StorageClass를 지정하려면 [에 설명된 지침](#)을 따릅니다 ["Kubeflow 구축"](#) 섹션을 참조하십시오. 클러스터 내에서 기본 StorageClass를 이미 지정한 경우에는 이 단계를 건너뛸 수 있습니다.

### Helm을 사용하여 공기 흐름을 전개하십시오

Helm을 사용하여 Kubernetes 클러스터에 공기 흐름을 배포하려면 배포 점프 호스트에서 다음 작업을 수행하십시오.

1. [에 따라 헬름으로 공기 흐름을 전개하십시오 "배포 지침"](#) Artifact Hub의 공식 공기 흐름 도표입니다. 다음 명령 예는 Helm을 사용한 공기 흐름의 배치를 보여줍니다. 환경과 원하는 구성에 따라 필요에 따라 'custom-values.yaml' 파일에서 값을 수정, 추가 및/또는 제거합니다.

```
$ cat << EOF > custom-values.yaml
#####
# Airflow - Common Configs
#####
airflow:
  ## the airflow executor type to use
  ##
```

```

    executor: "CeleryExecutor"
    ## environment variables for the web/scheduler/worker Pods (for
airflow configs)
    ##
    #
#####
# Airflow - WebUI Configs
#####
web:
    ## configs for the Service of the web Pods
    ##
    service:
        type: NodePort
#####
# Airflow - Logs Configs
#####
logs:
    persistence:
        enabled: true
#####
# Airflow - DAGs Configs
#####
dags:
    ## configs for the DAG git repository & sync container
    ##
    gitSync:
        enabled: true
        ## url of the git repository
        ##
        repo: "git@github.com:mboglesby/airflow-dev.git"
        ## the branch/tag/sha1 which we clone
        ##
        branch: master
        revision: HEAD
        ## the name of a pre-created secret containing files for ~/.ssh/
        ##
        ## NOTE:
        ## - this is ONLY RELEVANT for SSH git repos
        ## - the secret commonly includes files: id_rsa, id_rsa.pub,
known_hosts
        ## - known_hosts is NOT NEEDED if `git.sshKeyscan` is true
        ##
        sshSecret: "airflow-ssh-git-secret"
        ## the name of the private key file in your `git.secret`
        ##
        ## NOTE:

```

```

    ## - this is ONLY RELEVANT for PRIVATE SSH git repos
    ##
    sshSecretKey: id_rsa
    ## the git sync interval in seconds
    ##
    syncWait: 60
EOF
$ helm install airflow airflow-stable/airflow -n airflow --version 8.0.8
--values ./custom-values.yaml
...
Congratulations. You have just deployed Apache Airflow!
1. Get the Airflow Service URL by running these commands:
    export NODE_PORT=$(kubectl get --namespace airflow -o
    jsonpath="{.spec.ports[0].nodePort}" services airflow-web)
    export NODE_IP=$(kubectl get nodes --namespace airflow -o
    jsonpath="{.items[0].status.addresses[0].address}")
    echo http://$NODE_IP:$NODE_PORT/
2. Open Airflow in your web browser

```

2. 모든 공기 흐름 포드가 실행 중인지 확인합니다. 모든 Pod를 시작하는 데 몇 분 정도 걸릴 수 있습니다.

```

$ kubectl -n airflow get pod

```

NAME	READY	STATUS	RESTARTS	AGE
airflow-flower-b5656d44f-h8qjk	1/1	Running	0	2h
airflow-postgresql-0	1/1	Running	0	2h
airflow-redis-master-0	1/1	Running	0	2h
airflow-scheduler-9d95fcd9-clf4b	2/2	Running	2	2h
airflow-web-59c94db9c5-z7rg4	1/1	Running	0	2h
airflow-worker-0	2/2	Running	2	2h

3. 1단계에서 제어 장치를 사용하여 공기 흐름을 배포할 때 콘솔에 인쇄된 지침에 따라 공기 흐름 웹 서비스 URL을 얻습니다.

```

$ export NODE_PORT=$(kubectl get --namespace airflow -o
    jsonpath="{.spec.ports[0].nodePort}" services airflow-web)
$ export NODE_IP=$(kubectl get nodes --namespace airflow -o
    jsonpath="{.items[0].status.addresses[0].address}")
$ echo http://$NODE_IP:$NODE_PORT/

```

4. Airflow 웹 서비스에 액세스할 수 있는지 확인합니다.

	DAG	Schedule	Owner	Recent Tasks	Last Run	DAG Runs	Links
	ai_training_run	None	NetApp				
	create_data_scientist_workspace	None	NetApp				
	example_bash_operator	0 0 * * *	Airflow				
	example_branch_dop_operator_v3	* * * * *	Airflow				
	example_branch_operator	@daily	Airflow				
	example_complex	None	airflow				
	example_external_task_marker_child	None	airflow				
	example_external_task_marker_parent	None	airflow				
	example_http_operator	1 day, 0:00:00	Airflow				
	example_kubernetes_executor_config	None	Airflow				
	example_nested_branch_dag	@daily	airflow				
	example_passing_params_via_test_command	* * * * *	airflow				
	example_pig_operator	None	Airflow				
	example_python_operator	None	Airflow				
	example_short_circuit_operator	1 day, 0:00:00	Airflow				
	example_skip_dag	1 day, 0:00:00	Airflow				

## NetApp DataOps 툴킷을 공기 흐름과 함께 사용하십시오

를 클릭합니다 "Kubernetes용 NetApp DataOps 툴킷" 공기 흐름과 함께 사용할 수 있습니다. NetApp DataOps 툴킷을 공기 흐름과 함께 사용하면 스냅샷 및 클론 생성과 같은 NetApp 데이터 관리 작업을 공기 흐름으로 오케스트레이션되는 자동 워크플로우에 통합할 수 있습니다.

을 참조하십시오 "공기 흐름의 예" 섹션에서 NetApp DataOps 툴킷 사용에 대한 자세한 내용을 확인할 수 있습니다.

## 저작권 정보

Copyright © 2024 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.