K8s Operator/GitOps 패턴으로 OpenSearch 클러스터 구성하기

고영현 (Marcel Yeonghyeon Ko)

@SK hynix | AI/Data Engineer

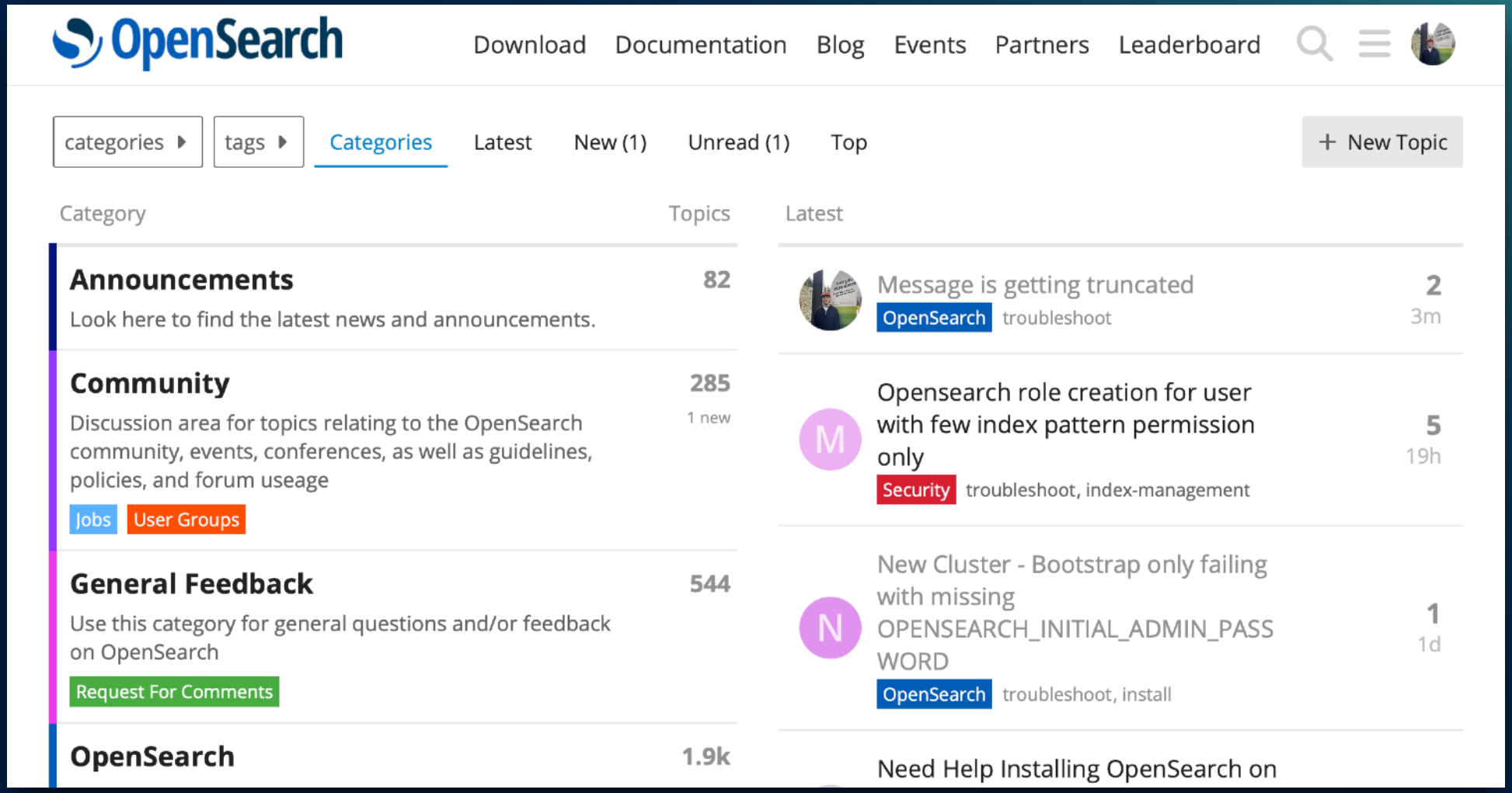
Introduction

고영현 (Marcel Yeonghyeon Ko)

- AI/Data Engineer @SK hynix
  - Search Engine & Performance Tuning (Elasticsearch, Spring Kafka)
  - RAG Pipeline (Crawling, KafkaConnect, OpenSearch)
  - Hybrid Cloud Platform 개발/운영 (Linux, Container, Kubernetes)
- OpenSource Contribution
  - Lucene, KafkaConnect, Elasticsearch, OpenSearch
- Community
  - AWSKRUG (AIEngineering, 플랫폼엔지니어링)
  - OpenSearch Forum/Community

OpenSearch

Introduction



OpenSearch

- OpenSearch Forum - https://forum.opensearch.com

Index

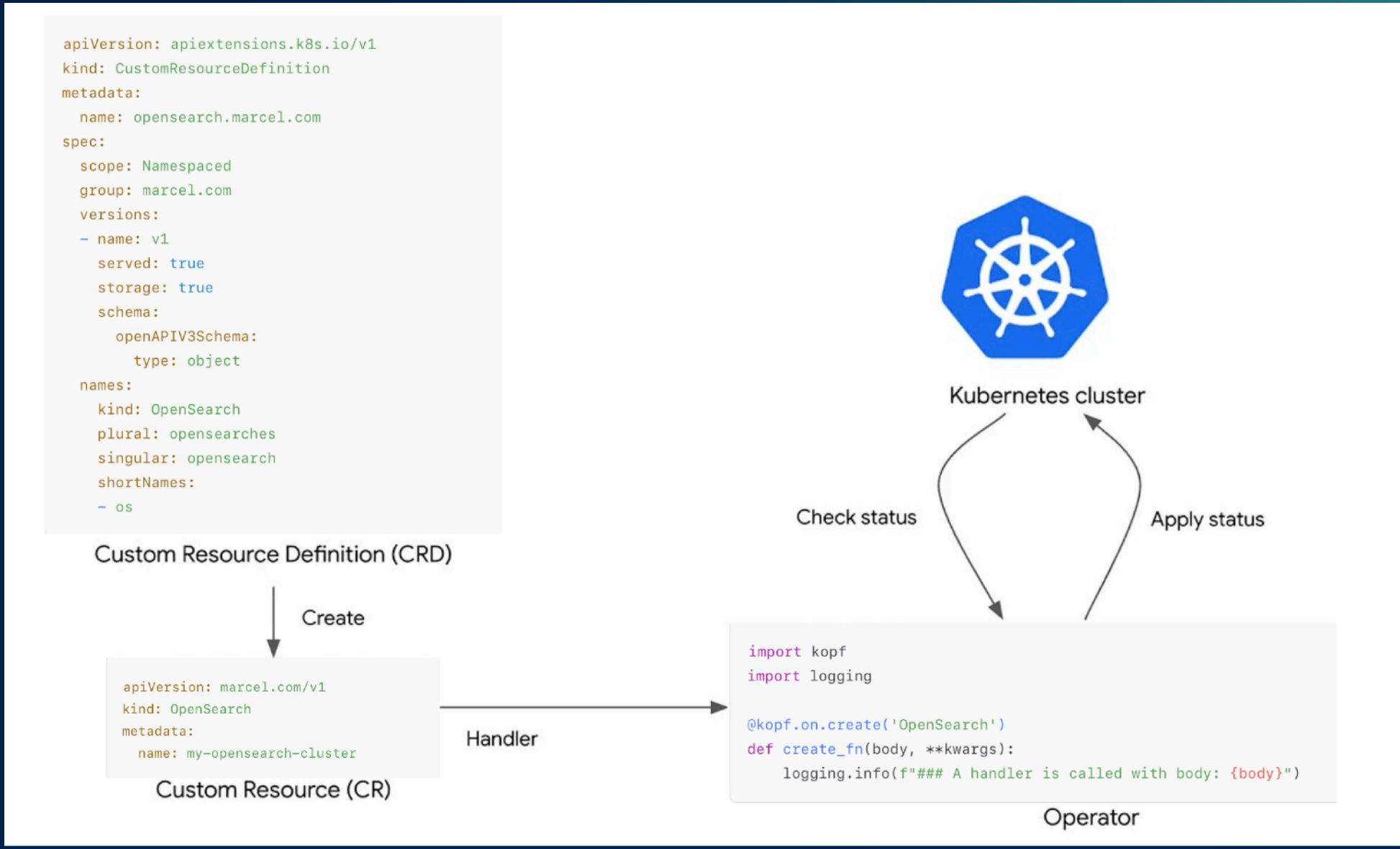
- Kubernetes Operator
- GitOps (feat. ArgoCD)
- OpenSearch 클러스터 아키텍처
- Transport/HTTP Layer의 TLS 인증 (feat. cert-manager)
- 클러스터/대시보드 Endpoint 외부 접근 (feat. Ingress)

1. Kubernetes Operator

K8s Operator/GitOps 패턴으로 OpenSearch 클러스터 구성하기

OpenSearch

1. Kubernetes Operator - **Flow**



1. Kubernetes Operator - **5W1H**

- **Who** - Managed Service 관리자 (ex. AWS 등 Cloud Service Provider)
- **What** - Container 기반 프로세스(CRD)와 운영관리 컨트롤러(Operator)
- **Where** - Kubernetes Cluster (CRD 버전 호환)
- **When** - 복잡한 리소스 생태계 관리, 모니터링/자동 복구
- **Why** - Declarative Setting/Deployment, Scalability, Overview
- **How** - Operator의 reconciler가 CR 변화 감지

1. Kubernetes Operator - **Reconciliation**

- Operator가 K8s 클러스터를 바라보고 있다가, CR 상의 변경(생성/수정/삭제)이 감지되면 object/resource/spec을 **reconcile**

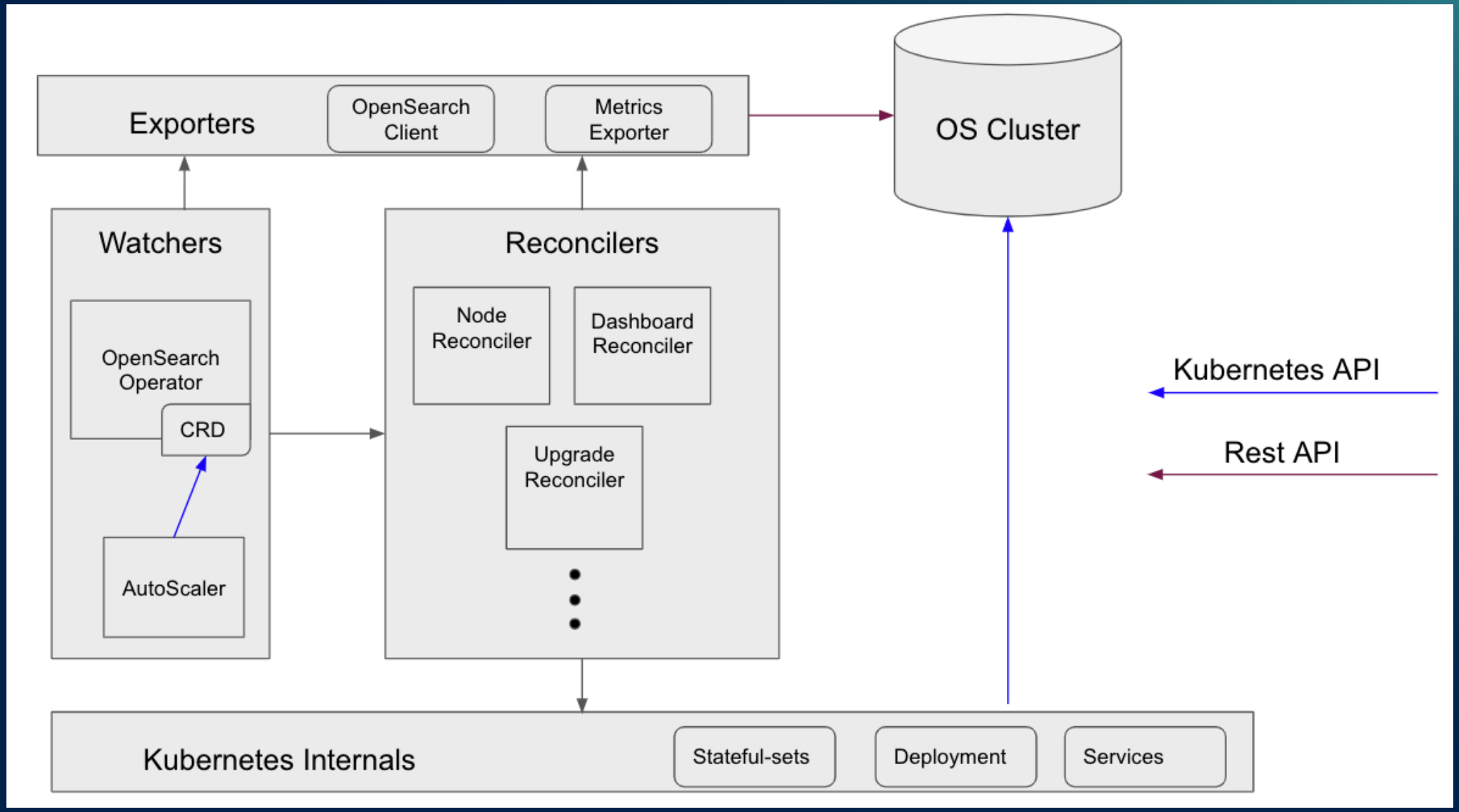
Reconciler	Description
Cluster	mapped to **nodePools**
Dashboards	mapped to **dashboards**
Configuration	In charge of OpenSearch **configuration**
TLS	mapped to security.**tls** (Security Plugin)
Securityconfig	mapped to security.**config** with users/roles (securityadmin.sh 실행)

1. Kubernetes Operator - **CRD**

- OpenSearchCluster CR의 spec

Name	Description
general	OpenSearch general configuration
bootstrap	Bootstrap pod configuration
dashboards	OpenSearch Dashboards configuration
security	Defined security reconciler configuration
nodePools[ ]	Each nodePool represents a group of nodes with the same roles/resources. Deployed as a Kubernetes StatefulSet.

1. Kubernetes Operator - **Architecture**



2. GitOps (feat. ArgoCD)

K8s Operator/GitOps 패턴으로 OpenSearch 클러스터 구성하기

OpenSearch

2. GitOps (feat. ArgoCD) : **Paradigm**

Manual Deployment

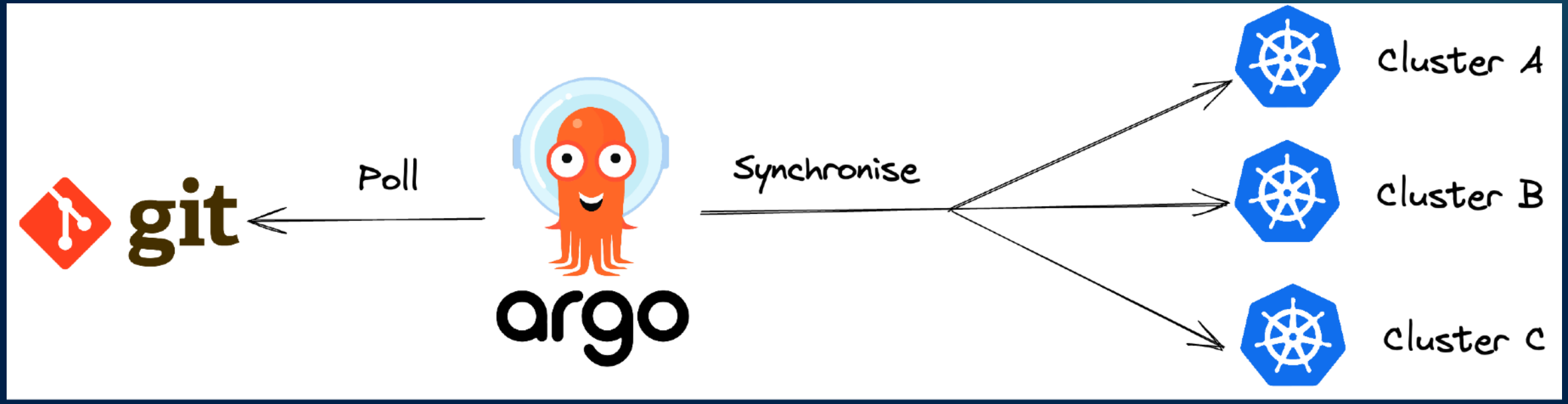
- BC (**B**efore **C**I/CD)

A-1	Local 환경에서 테스트가 끝나면 배포용 Branch에 머지된 코드를 Git의 Remote 저장소에 git push
A-2	배포 대상 인스턴스에 접속하여 git pull 후 build 진행 (ex. Java 애플리케이션의 경우 gradlew build -> java -jar {path})
B	Local 환경의 env 변경 및 build 후, FTP/SCP 등으로 인스턴스에 전송 및 배포

2. GitOps (feat. ArgoCD) : **Paradigm**

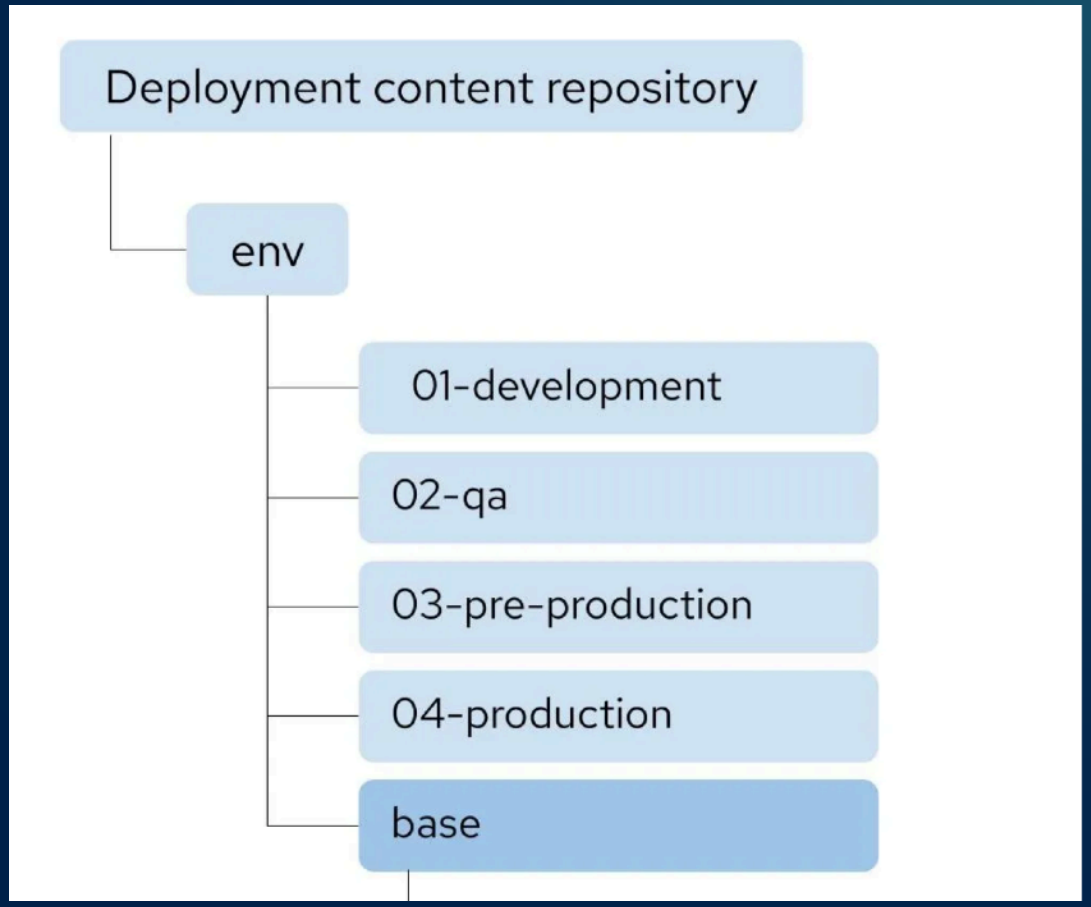
Declarative CI/CD와 Git Workflow

- AD (**A**fter ArgoC**D**)



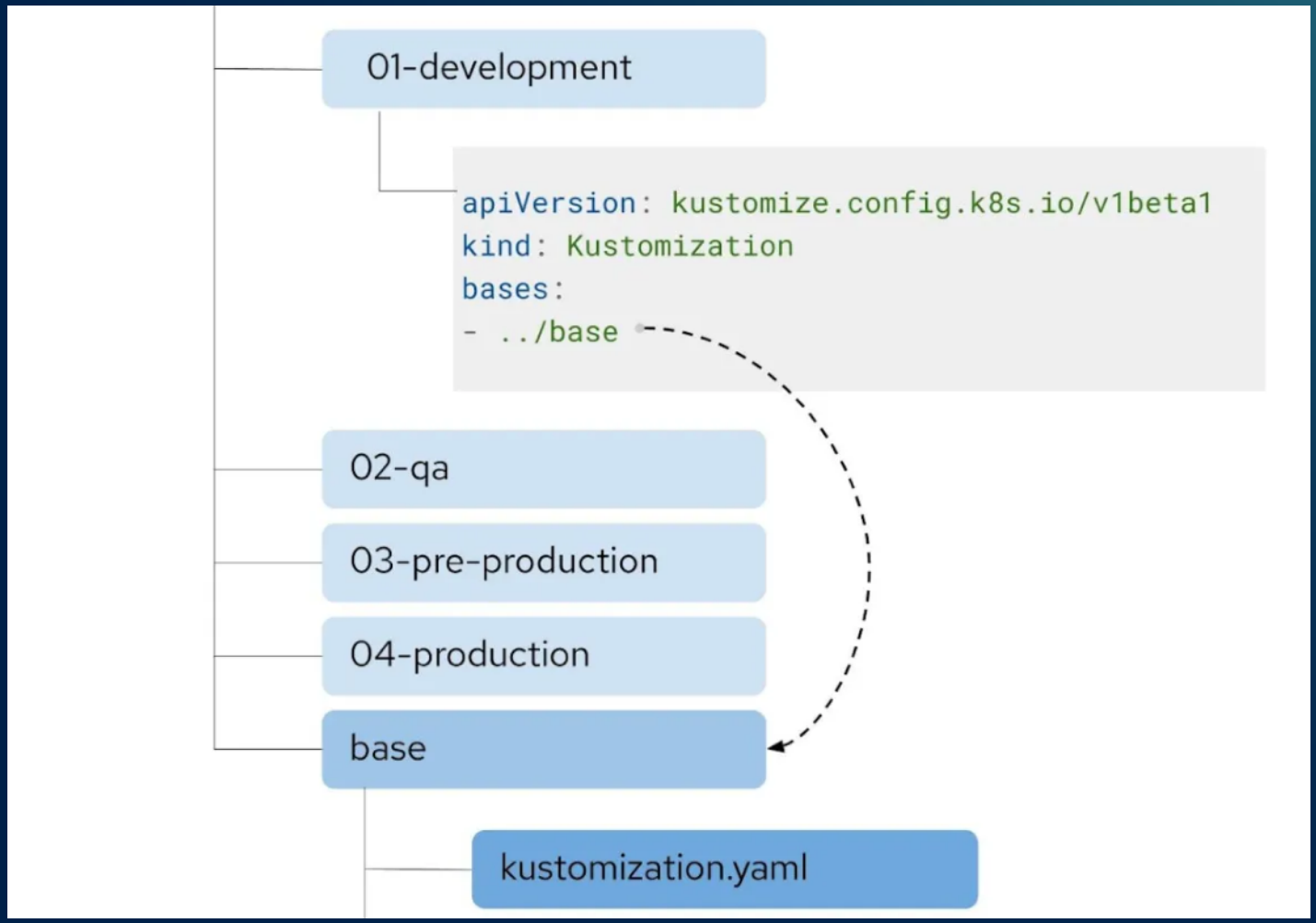
2. GitOps (feat. ArgoCD) : **Repo Structure**

- Git Repository 구조를 아래와 같이 base (주로 공식 helm chart)와 배포 환경(dev/qa/stg/prd)으로 구분하는 것이 일반적



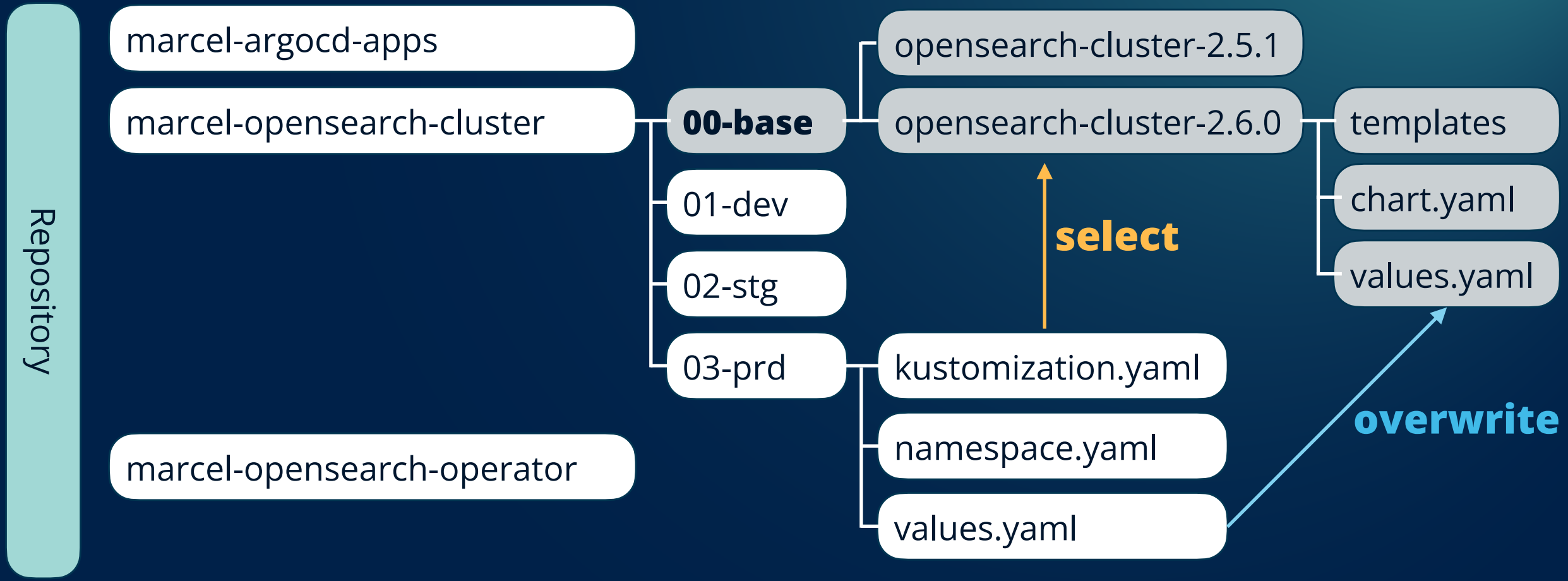
2. GitOps (feat. ArgoCD) : **Repo Structure**

- base에 저장된 yaml을 바탕으로 각 배포 환경에 맞추어 overwrite



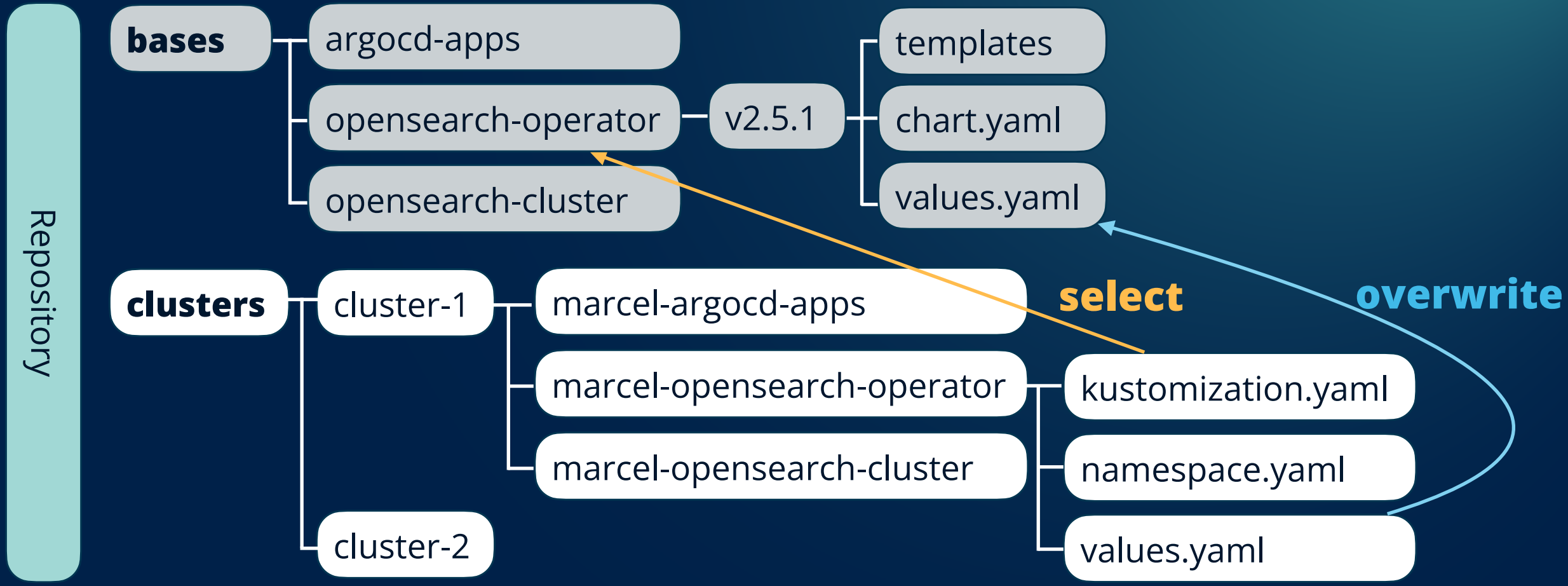
2. GitOps (feat. ArgoCD) : **Repo Structure**

Best Practice (1) - Single Cluster



2. GitOps (feat. ArgoCD) : **Repo Structure**

Best Practice (2) - Multi Clusters



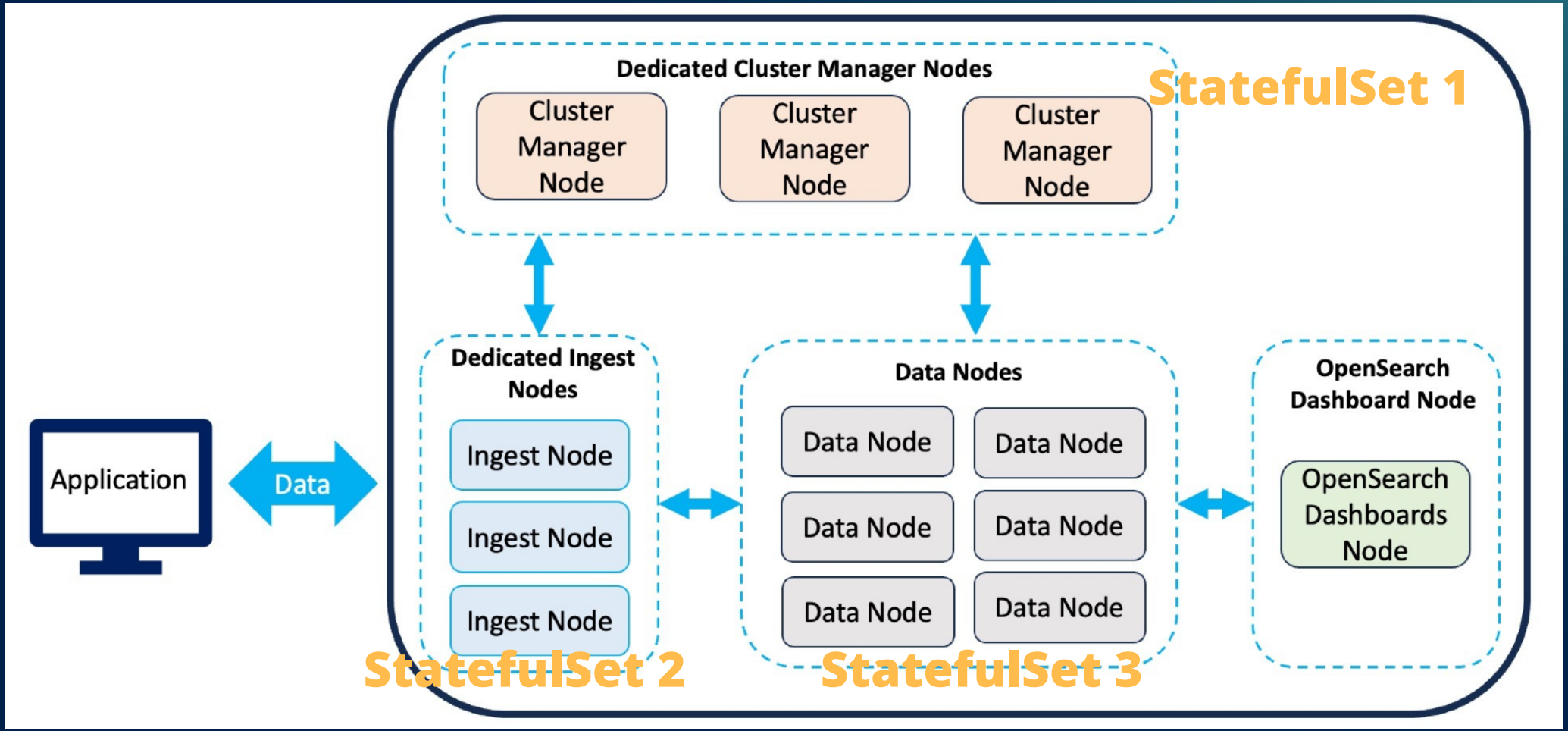
3. OpenSearch 클러스터 아키텍처

K8s Operator/GitOps 패턴으로 OpenSearch 클러스터 구성하기

OpenSearch

3. OpenSearch 클러스터 아키텍처 - **Node**

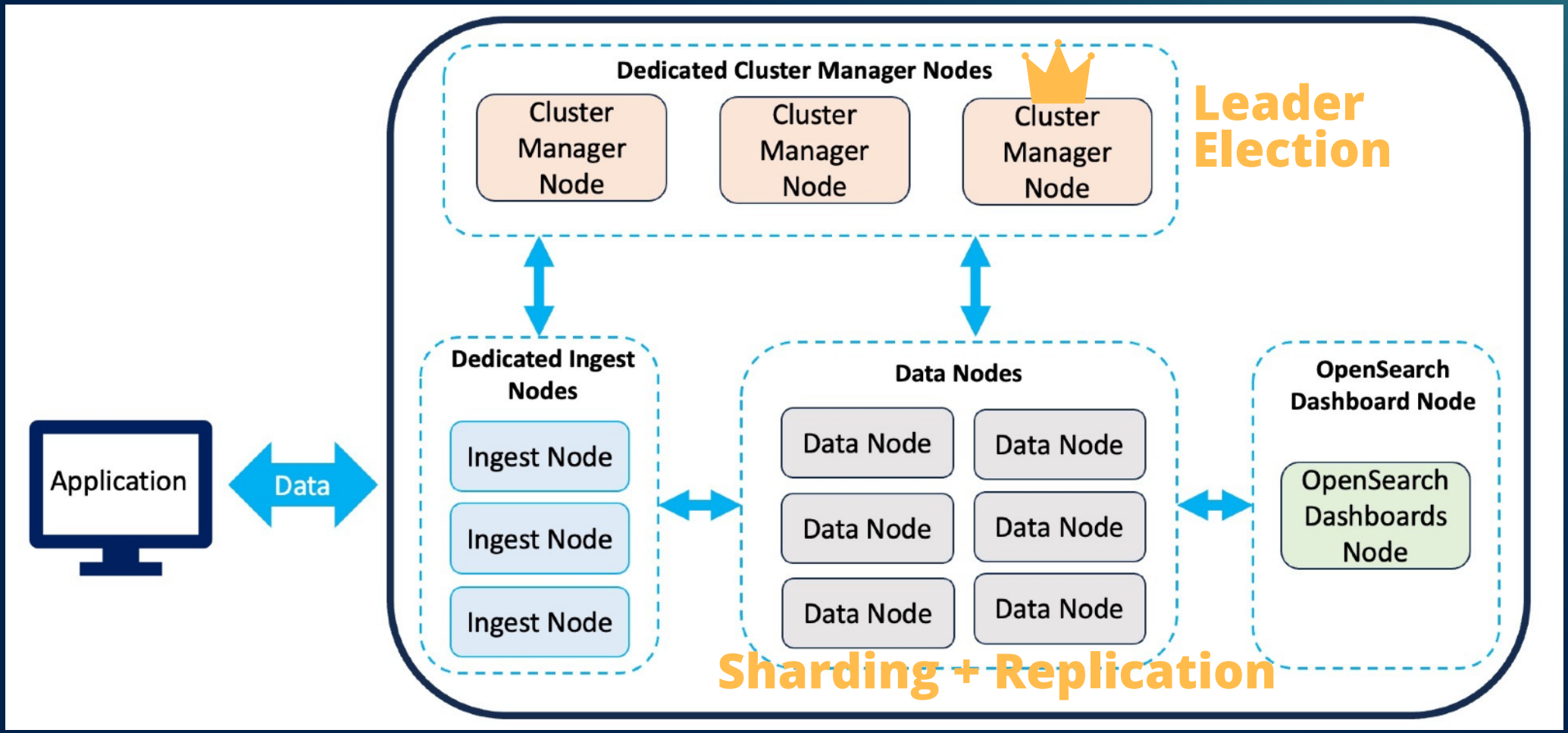
- OpenSearch : Kubernetes = **Nodes** : **StatefulSet**



https://www.instaclustr.com/blog/dedicated-ingest-nodes-public-preview-release-for-opensearch-on-instaclustr-by-netapp/

3. OpenSearch 클러스터 아키텍처 - **Node**

- **Quorum** & **Shard Replication**으로 High Availability 보장



https://www.instaclustr.com/blog/dedicated-ingest-nodes-public-preview-release-for-opensearch-on-instaclustr-by-netapp/

3. OpenSearch 클러스터 아키텍처 - **Node**

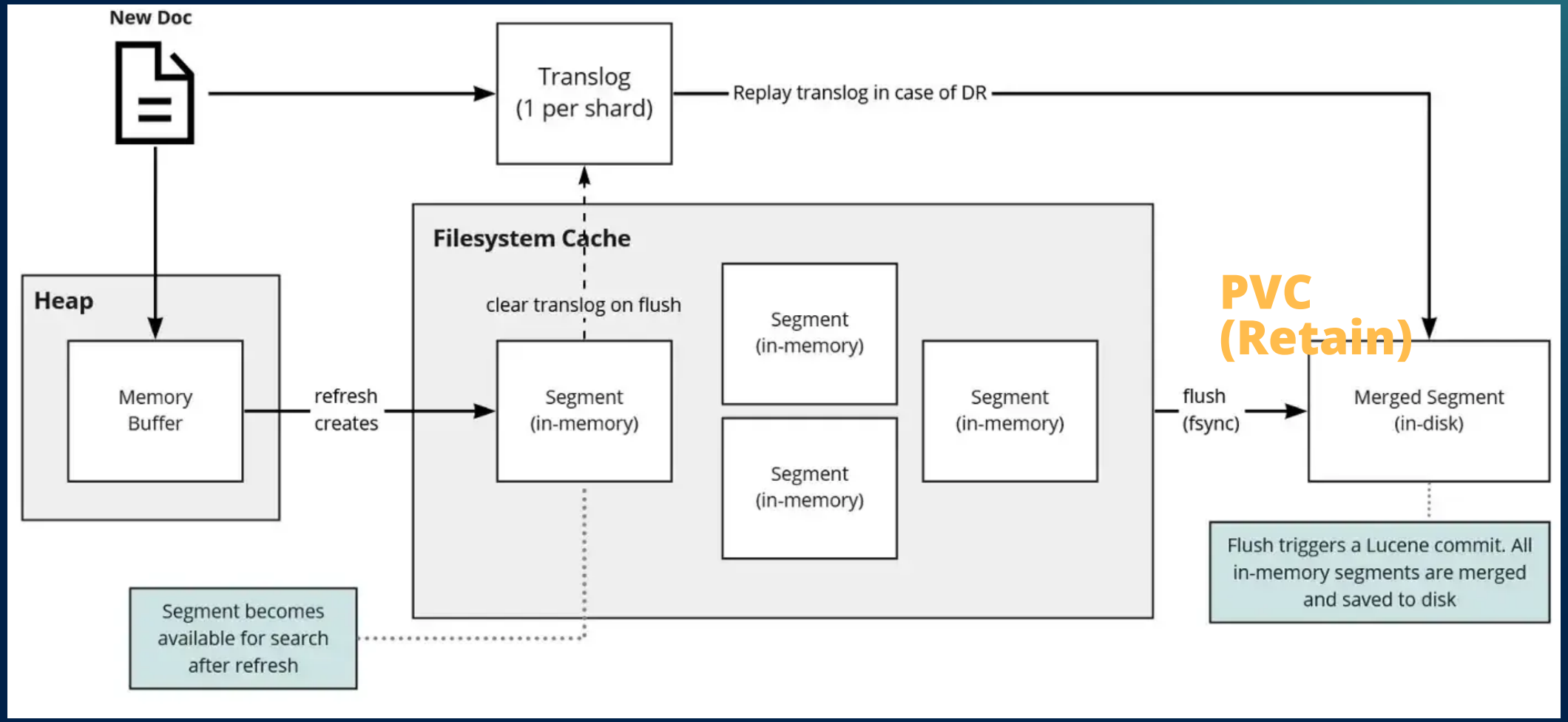
```yaml
spec:
  nodePools:
  - component: master
    replicas: 3
    diskSize: "10Gi"
    resources:
      requests:
        memory: "2Gi"
        cpu: "100m"
      limits:
        memory: "2Gi"
        cpu: "1"
    roles:
      - "cluster_manager"
  - component: data
    replicas: 2
    diskSize: "500Gi"
    resources:
      requests:
        memory: "8Gi"
        cpu: "100m"
      limits:
        memory: "8Gi"
        cpu: "4"
    roles:
      - "data"
```

StatefulSet - master

StatefulSet - data

3. OpenSearch 클러스터 아키텍처 - **Segment**

- OpenSearch : Kubernetes = **Segments** : **PVCs**



https://blog.devgenius.io/elasticsearch-solution-to-searching-71116220c82f

3. OpenSearch 클러스터 아키텍처 - **Segment**

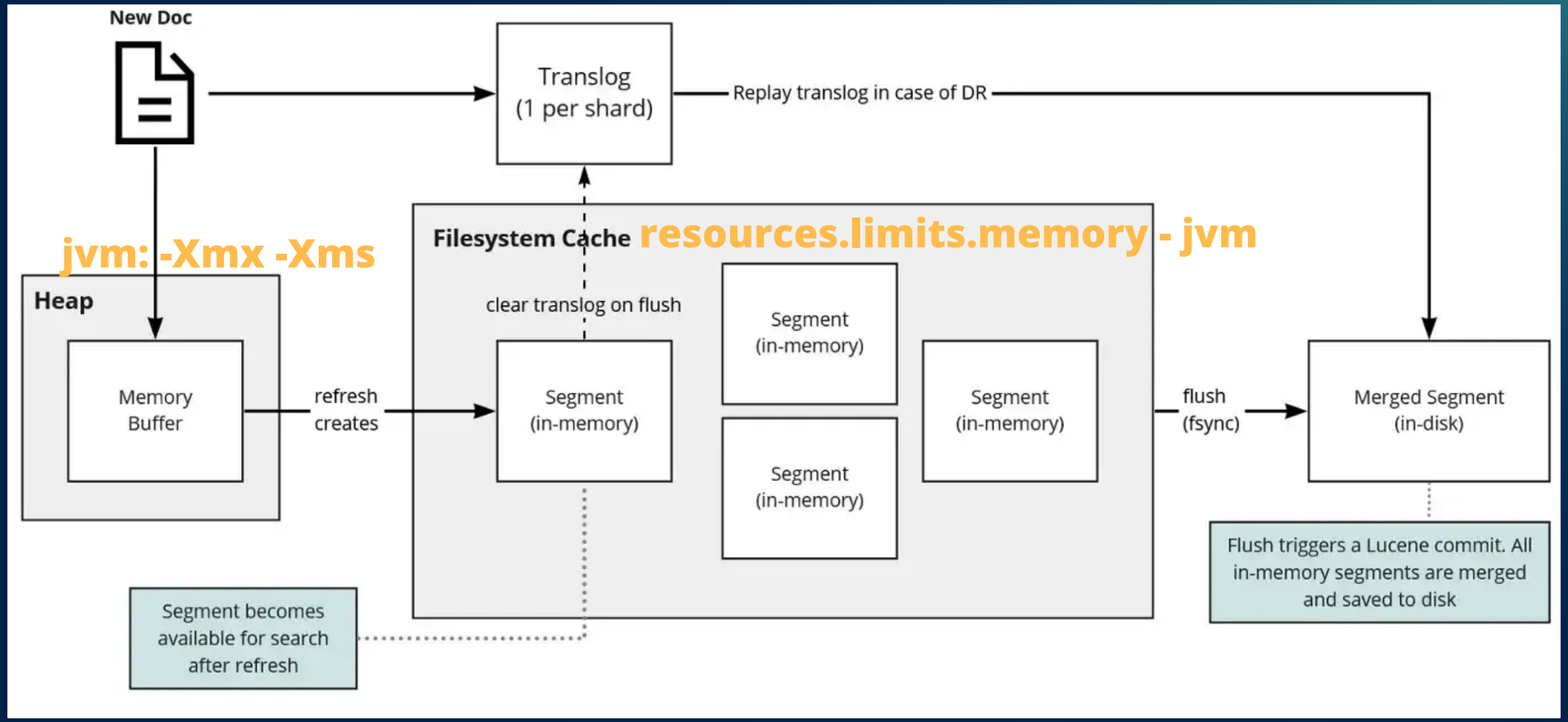
- OpenSearch : Kubernetes = **Segments** : **PVCs**

```
nodePools:
- component: standalone
  replicas: 3
  diskSize: 300
  roles:
    - "data"
    - "cluster_manager"
  persistence:
    pvc:
      storageClass: marcel-sc-vmware-vsphere
      accessModes:
      - ReadWriteOnce
```

csi-driver & k8s version 호환 시
PVC volume expansion도 가능

3. OpenSearch 클러스터 아키텍처 - **Segment**

- K8s/Linux 환경에서의 JVM Heap & Page Cache



https://blog.devgenius.io/elasticsearch-solution-to-searching-71116220c82f

3. OpenSearch 클러스터 아키텍처 - **Segment**

- K8s/Linux 환경에서의 JVM Heap & Page Cache

```
spec:
    nodePools:
      - component: nodes
        replicas: 3
        diskSize: "10Gi"
        jvm: -Xmx1024M -Xms1024M
        resources:
          requests:
            memory: "2Gi"
            cpu: "500m"
          limits:
            memory: "2Gi"
            cpu: "500m"
        roles:
          - "data"
```

2Gi - 1024M ~= 1123M을
커널 시스템이 Cache로 활용
*** mmap 방식**

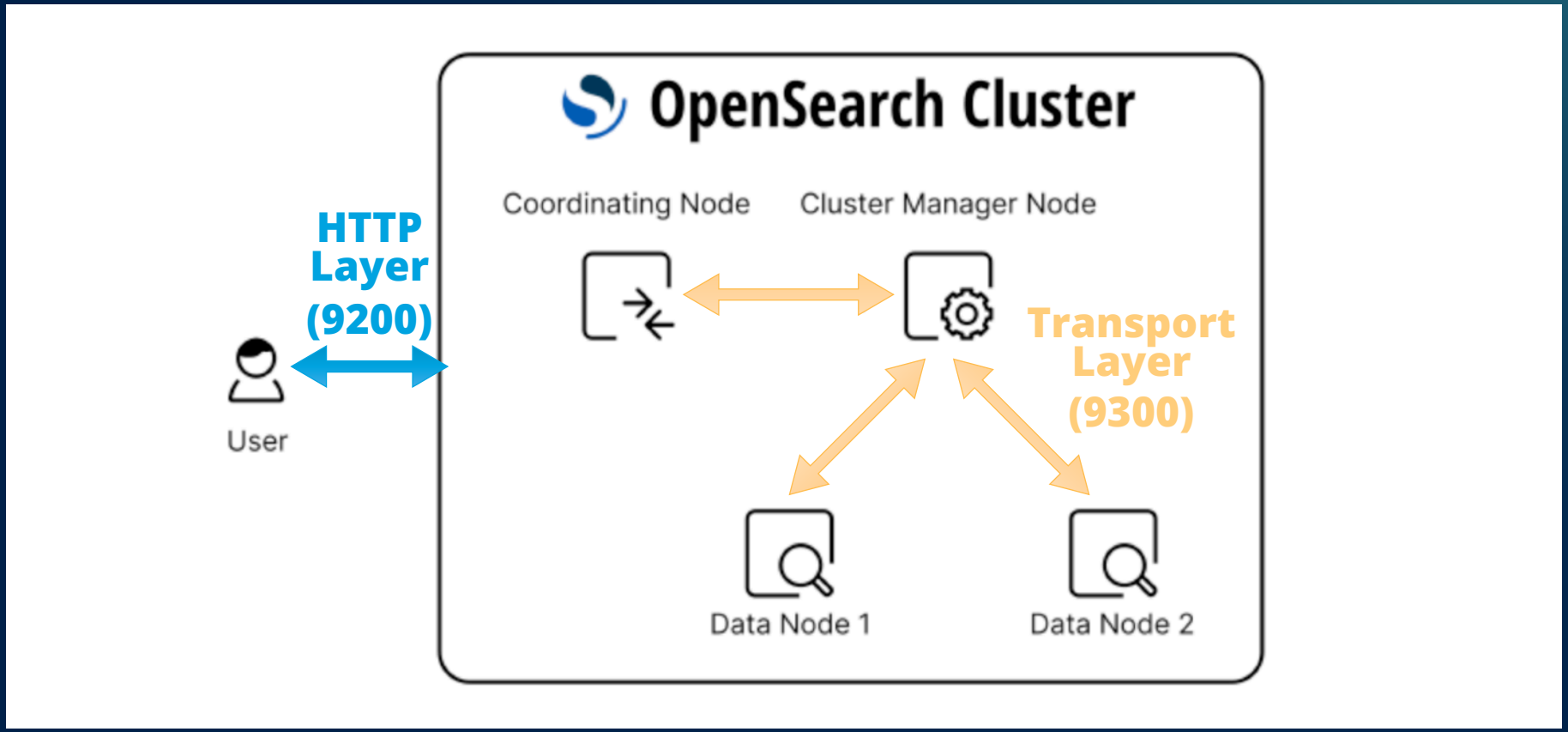
4. Transport/HTTP Layer의 TLS 인증 (feat. cert-manager)

K8s Operator/GitOps 패턴으로 OpenSearch 클러스터 구성하기

OpenSearch

4. Transport/HTTP Layer의 TLS 인증

- OpenSearch는 HTTP/Transport Layer 형태로 Network 통신



4. Transport/HTTP Layer의 TLS 인증

- OpenSearch K8s Operator에는 2가지 방식의 TLS 인증 체계 존재
  A. Operator-generated CA/Certificates
  B. Self-signed CA/Certificates

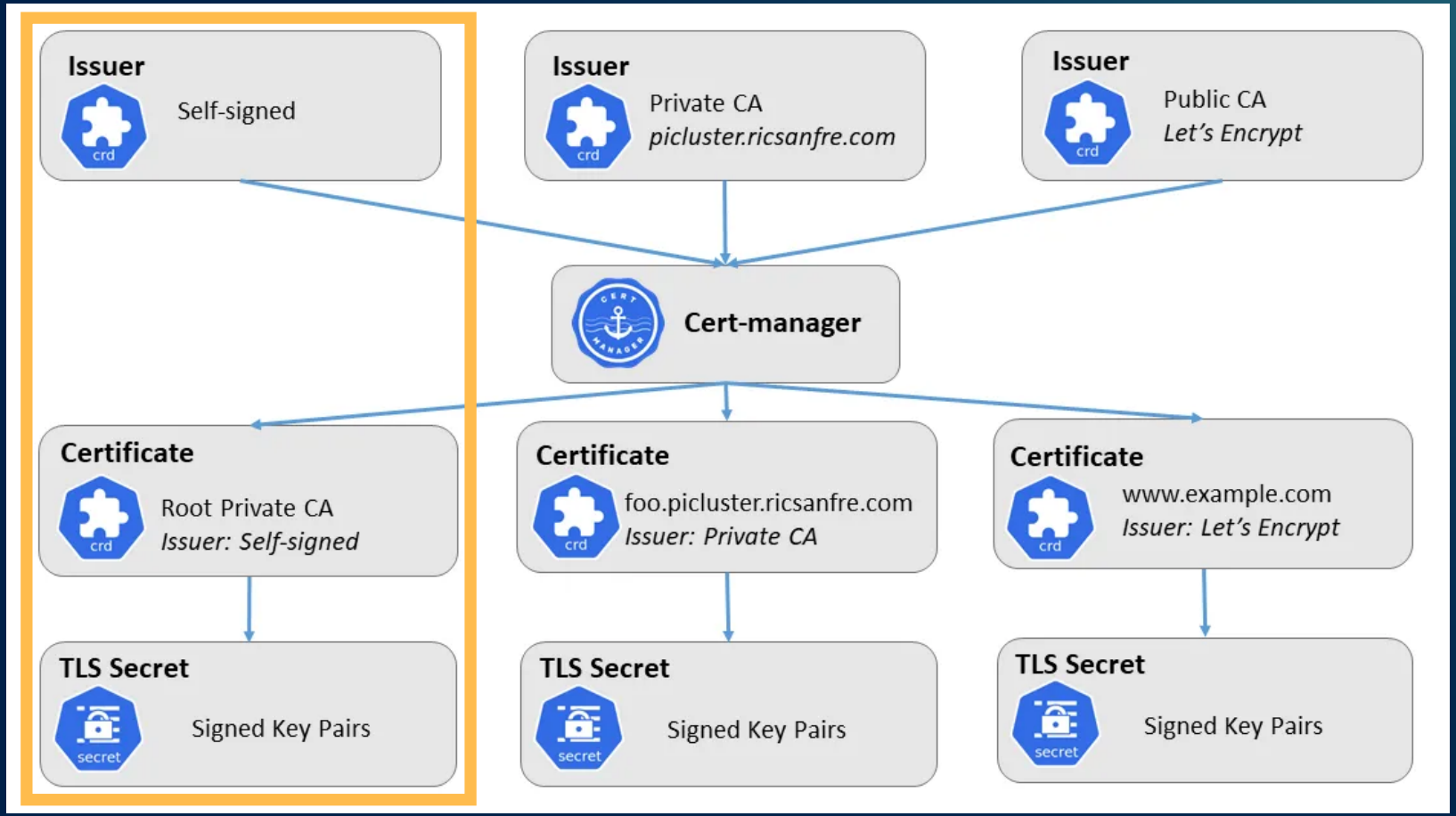
Depending on your requirements, the Operator offers two ways of managing TLS certificates. You can either supply your own certificates, or the Operator will generate its own CA and sign certificates for all nodes using that CA. The second option is recommended, unless you want to directly expose your OpenSearch cluster outside your Kubernetes cluster, or your organization has rules about using self-signed certificates for internal communication.

> ⚠️ **Clusters with operator-generated certificates will stop working after 1 year**: Make sure you have tested certificate renewals in your cluster before putting it in production!

https://github.com/opensearch-project/opensearch-k8s-operator/edit/main/docs/userguide/main.md#tls

4. Transport/HTTP Layer의 TLS 인증

- **cert-manager**로 self-signed CA/Certificates를 만들자



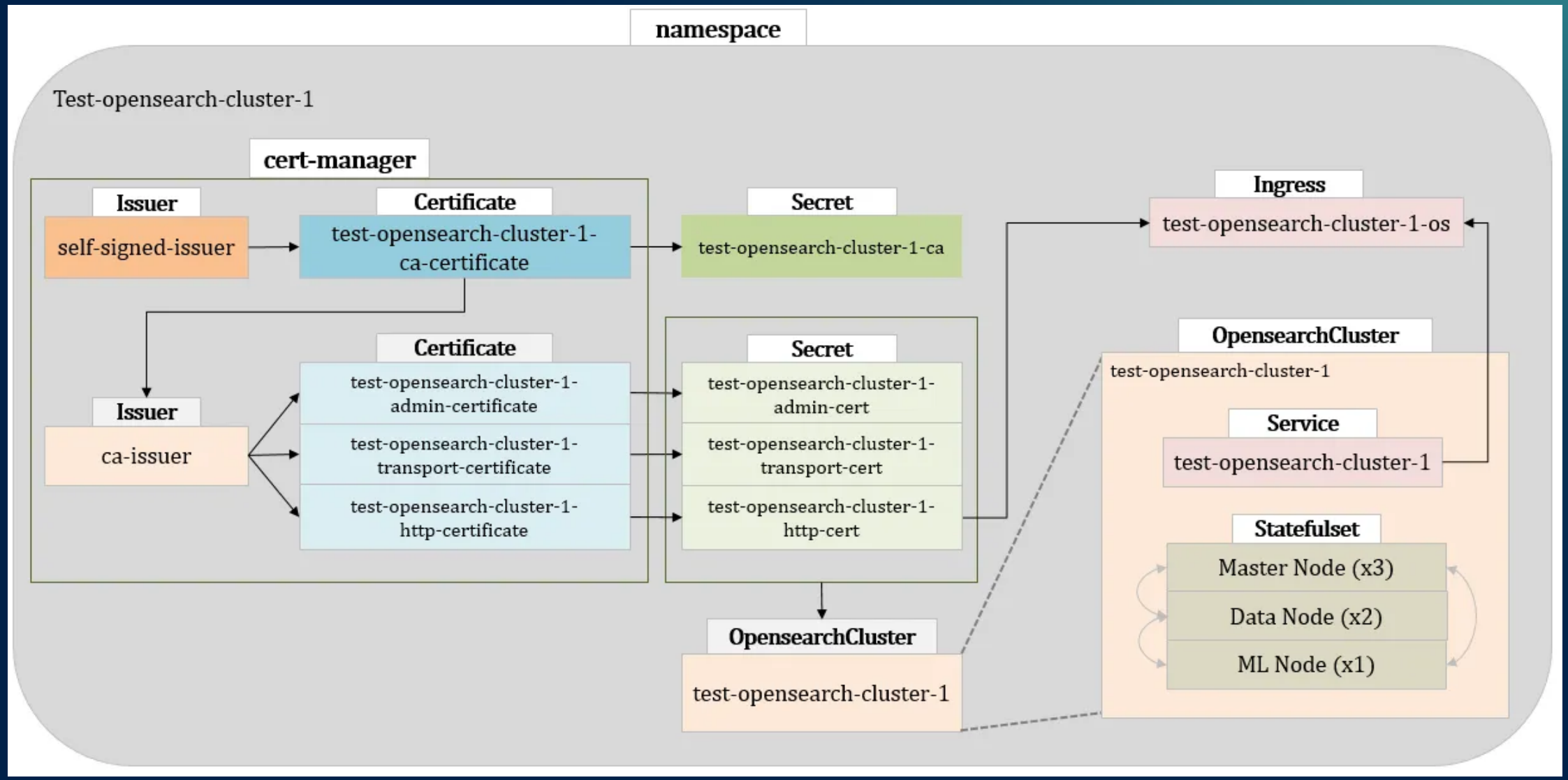
4. Transport/HTTP Layer의 TLS 인증

- **Issuer** : 클러스터 내에서의 CA 역할을 담당
- **Certificate** : Issuer를 통해 생성된 인증서
- **Secret** : CA 인증서를 비롯한 TLS 인증서와 키파일을 포함

```
# transport/http 계층의 tls auto-generate 방식도 admin 인증서는 별도로 생성함.
test-opensearch-cluster-1-admin-cert       kubernetes.io/tls   3      26d
test-opensearch-cluster-1-admin-password   Opaque              2      26d
test-opensearch-cluster-1-ca               Opaque              2      26d
test-opensearch-cluster-1-http-cert        kubernetes.io/tls   3      26d
test-opensearch-cluster-1-transport-cert   kubernetes.io/tls   3      26d
```

https://medium.com/@dk02315/customizing-tls-for-opensearch-cluster-using-cert-manager-fb7c212e8ded

4. Transport/HTTP Layer의 TLS 인증



https://medium.com/@dk02315/customizing-tls-for-opensearch-cluster-using-cert-manager-fb7c212e8ded

5. 클러스터/대시보드 Endpoint 외부 접근 (feat. Ingress)

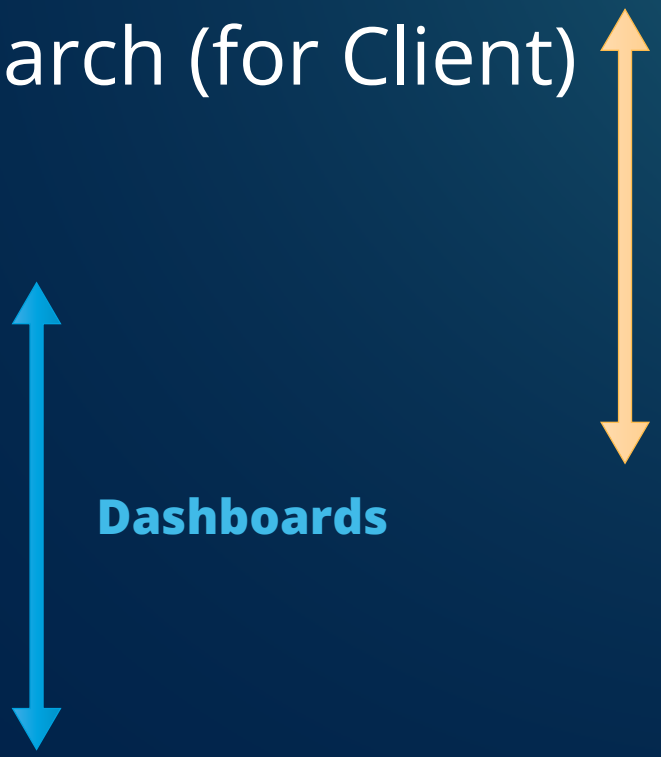
K8s Operator/GitOps 패턴으로 OpenSearch 클러스터 구성하기

OpenSearch

5. 클러스터/대시보드 Endpoint 외부 접근

- Why are Endpoints important for accessing OpenSearch?

✓ REST API : _index, _search (for Client)
✓ Cluster HealthCheck
✓ Index Management
✓ Security
✓ DevTools
✓ Discover / Visualize



https://playground.opensearch.org/auth/anonymous?nextUrl=/app/home#/

5. 클러스터/대시보드 Endpoint 외부 접근

```yaml
security:
  tls:
    transport:
      generate: false
      perNode: false
      secret:
        name: test-opensearch-cluster-1-transport-cert
      caSecret:
        name: test-opensearch-cluster-1-ca
      nodesDn: ["CN=test-opensearch-cluster-1,OU=test-opensearch-cluster-1"]
      adminDn: ["CN=admin,OU=test-opensearch-cluster-1"]
    http:
      generate: false
      secret:
        name: test-opensearch-cluster-1-http-cert
  config:
    adminSecret:
      name: test-opensearch-cluster-1-admin-cert
    adminCredentialsSecret:
      name: admin-credentials-secret
    securityConfigSecret:
      name: securityconfig-secret
```

5. 클러스터/대시보드 Endpoint 외부 접근

[Cluster]

```yaml
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  annotations:
    ingress.kubernetes.io/backend-protocol: HTTPS
  name: test-opensearch-cluster-1-os
  namespace: test-opensearch-cluster-1
spec:
  ingressClassName: nginx
  rules:
  - host: test-opensearch-cluster-1-os.marcel.com
    http:
      paths:
      - backend:
          service:
            name: test-opensearch-cluster-1
            port:
              number: 9200
        path: /
        pathType: ImplementationSpecific
  tls:
  - hosts:
    - test-opensearch-cluster-1-os.marcel.com
    secretName: test-opensearch-cluster-1-http-cert
```

[Dashboards]

```yaml
apiVersion: networking.k8s.io/v1
kind: Ingress
metadata:
  name: test-opensearch-cluster-1-osd
  namespace: test-opensearch-cluster-1
spec:
  ingressClassName: nginx
  rules:
  - host: test-opensearch-cluster-1-osd.marcel.com
    http:
      paths:
      - backend:
          service:
            name: test-opensearch-cluster-1-dashboards
            port:
              number: 5601
        path: /
        pathType: ImplementationSpecific
```

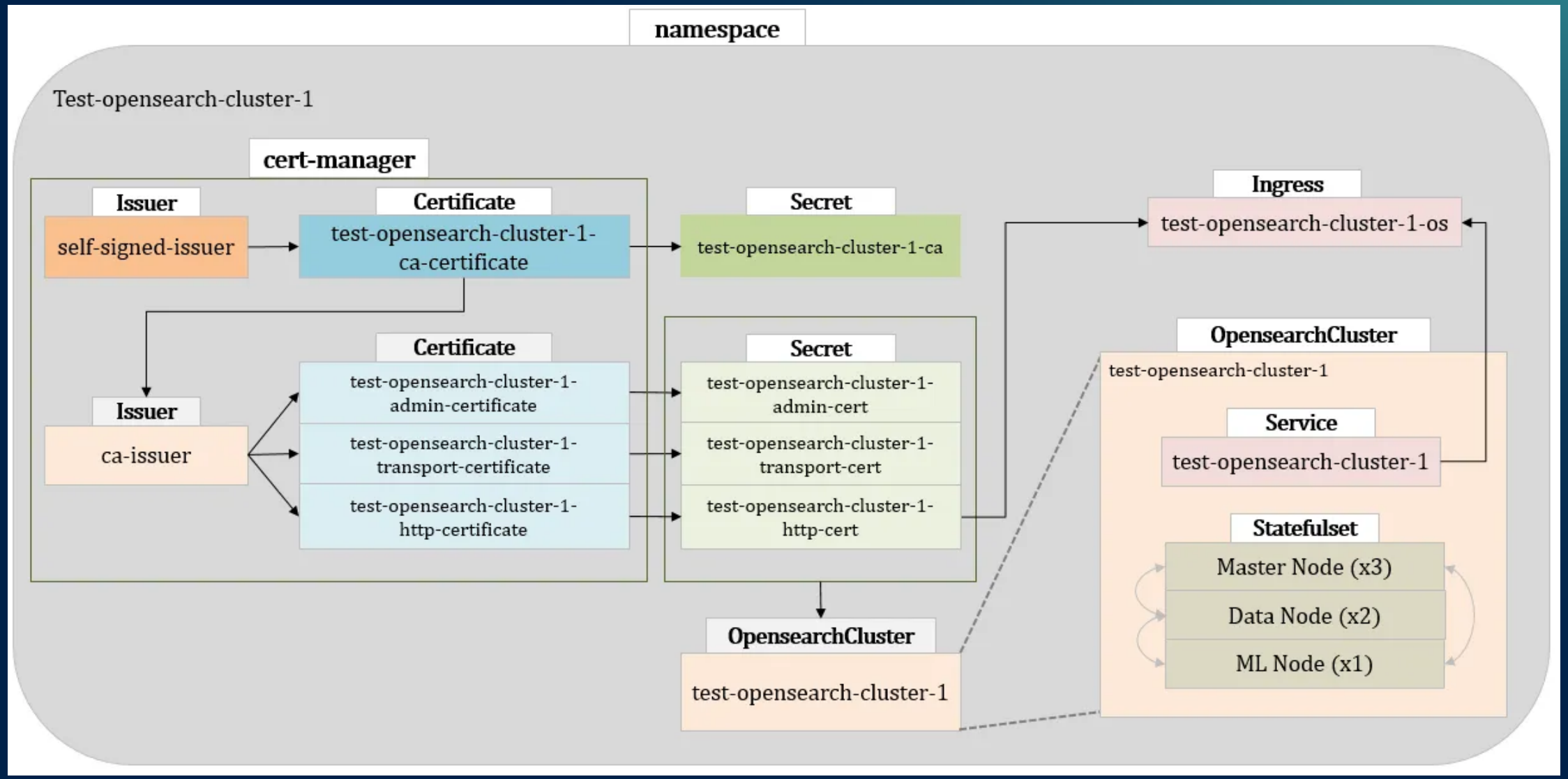
5. 클러스터/대시보드 Endpoint 외부 접근

```
curl -ku admin:test -v https://test-opensearch-cluster-1-os.marcel.com/

*   Trying 10.100.200.32:443...
* TCP_NODELAY set
* Connected to test-opensearch-cluster-1-os.marcel.com (10.100.200.32) port 443 (#0)
* ALPN, offering h2
* ALPN, offering http/1.1
* successfully set certificate verify locations:
*   CAfile: /etc/ssl/certs/ca-certificates.crt
  CApath: /etc/ssl/certs
* TLSv1.3 (OUT), TLS handshake, Client hello (1):
* TLSv1.3 (IN), TLS handshake, Server hello (2):
* TLSv1.3 (IN), TLS handshake, Encrypted Extensions (8):
* TLSv1.3 (IN), TLS handshake, Certificate (11):
* TLSv1.3 (IN), TLS handshake, CERT verify (15):
* TLSv1.3 (IN), TLS handshake, Finished (20):
* TLSv1.3 (OUT), TLS change cipher, Change cipher spec (1):
* TLSv1.3 (OUT), TLS handshake, Finished (20):
* SSL connection using TLSv1.3 / TLS_AES_256_GCM_SHA384
...
* TLSv1.3 (IN), TLS handshake, Newsession Ticket (4):
* old SSL session ID is stale, removing
* Connection state changed (MAX_CONCURRENT_STREAMS == 128)!
< HTTP/2 200
< date: Sun, 05 Jan 2025 07:37:26 GMT
< content-type: application/json; charset=UTF-8
< content-length: 599
< strict-transport-security: max-age=15724800; includeSubDomains
<
```


```
{
  "name" : "test-opensearch-cluster-1-data-0",
  "cluster_name" : "test-opensearch-cluster-1",
  "cluster_uuid" : "cDYu-y5bSfGWD-pQPP1VQQ",
  "version" : {
    "distribution" : "opensearch",
    "number" : "2.17.1",
    "build_type" : "tar",
    "build_hash" : "61dbcd0795c9bfe9b81e5762175414bc38bbcadf",
    "build_date" : "2024-06-20T03:26:49.193630411Z",
    "build_snapshot" : false,
    "lucene_version" : "9.11.0",
    "minimum_wire_compatibility_version" : "7.10.0",
    "minimum_index_compatibility_version" : "7.0.0"
  },
  "tagline" : "The OpenSearch Project: https://opensearch.org/"
}
```

5. 클러스터/대시보드 Endpoint 외부 접근



OpenSearch Cluster 운영 Know-how

✓ Managed Service에서 사용자들을 어디까지 믿을 것인가?

- 사용자를 불신할수록 서비스 자유도가 낮아지는 trade-off

  -> Guideline/reference 및 UI/UX로 어느 정도 극복 가능

1) drainDataNodes (cluster.routing.allocation.exclude)

2) JVM Heap을 항상 limited.memory의 50%로 할당할지

3) internal_user & roles 권한 세분화

OpenSearch Cluster 운영 Know-how

✓ Air-Gapped 환경을 위한 Tip

- serviceMonitor는 OpenSearch Cluster CR 3.0.0 부터 prometheus-exporter.zip 자동 설치 기능이 존재
- But Air-Gapped에서는 방화벽 에러만 로그에 찍힐 뿐...
- Docker Image에 애초에 필요한 plugin을 COPY layer로 넣자
- 비슷한 방식으로 ML Model 배포 가능 (혹은 CDN 이용)

Wrap-up

- ✓ OpenSearch용 K8s Operator
- ✓ GitOps 패턴의 OpenSearch 클러스터 CI/CD
- ✓ Kubernetes 생태계 속의 OpenSearch
- ✓ OpenSearch Network - TLS in HTTP/Transport Layer
- ✓ Cluster/Dashboards Endpoint

Q&A

OpenSearch