Nordea

GitOps or a Journey to a Production Incident and Swiftly Back Again

Kubernetes Community Days Denmark

Pavol Hronsky, Head of Container Automation

14.11.2023

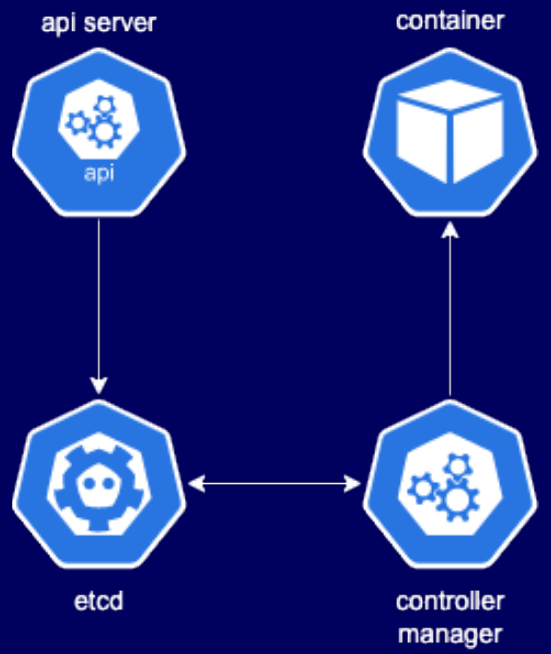
Chapter 1

The Prologue

https://www.freepik.com/free-photo/top-view-christmas-cookies-with-wooden-background_5795814.htm
Image by **Freepik**

Kubernetes' Reconciliation Loop

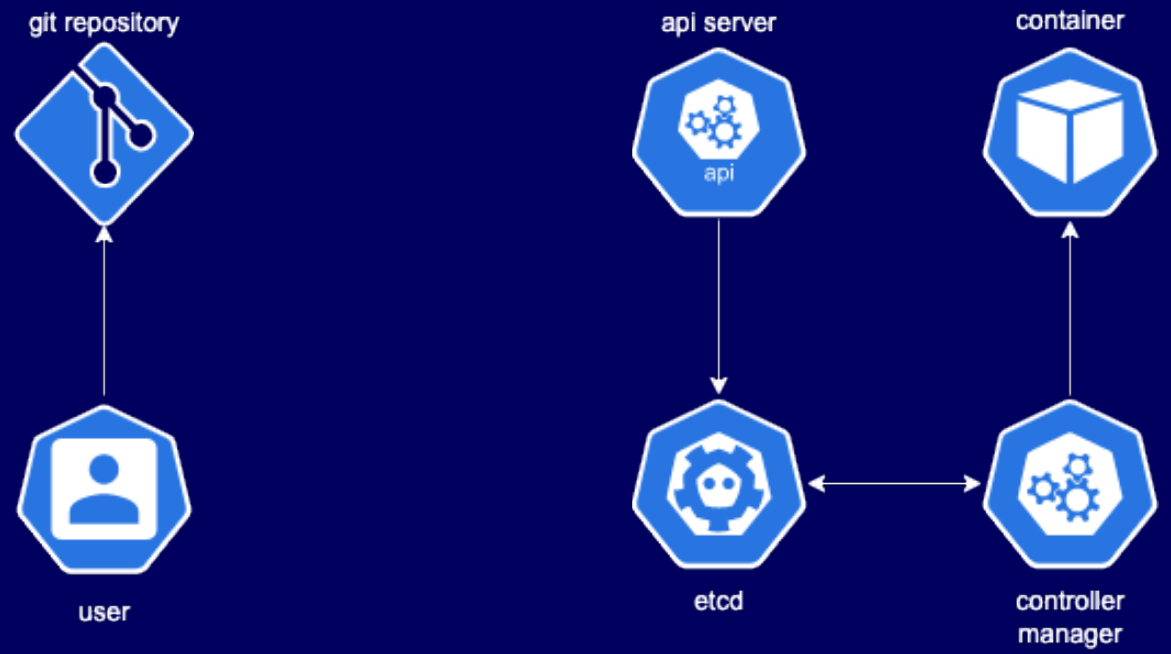
A continuous process which compares the actual state of the system to the desired state, and any differences are reconciled or corrected to bring the system back to the desired state.



Nordea

Kubernetes' Reconciliation Loop

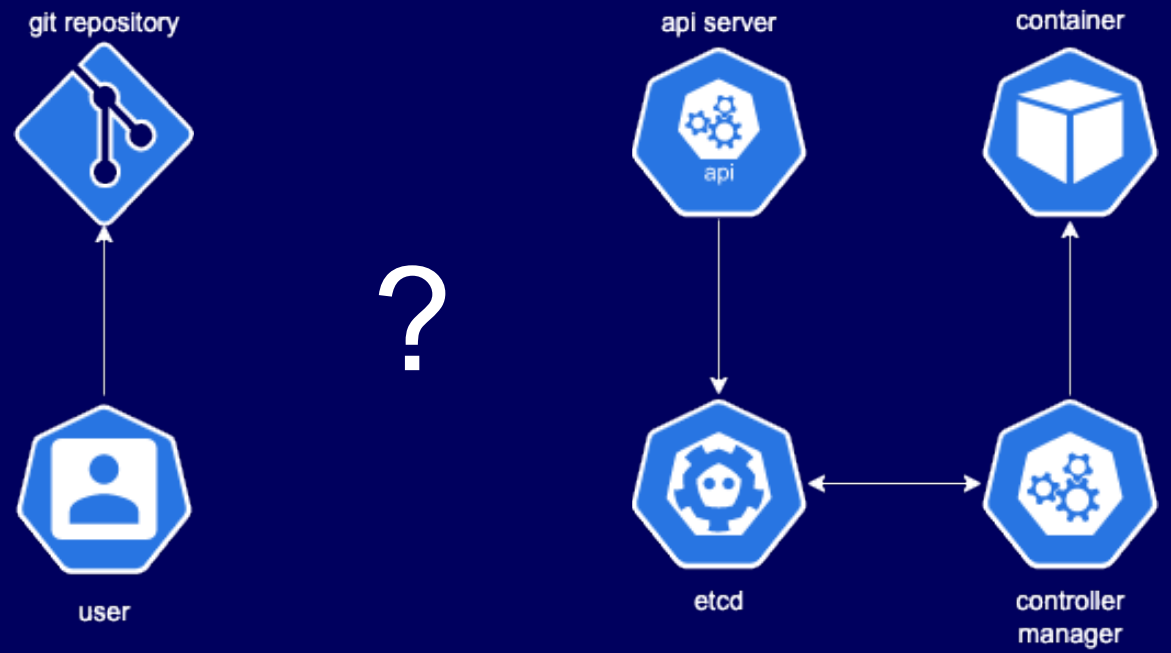
A continuous process which compares the actual state of the system to the desired state, and any differences are reconciled or corrected to bring the system back to the desired state.



Nordea

Kubernetes' Reconciliation Loop

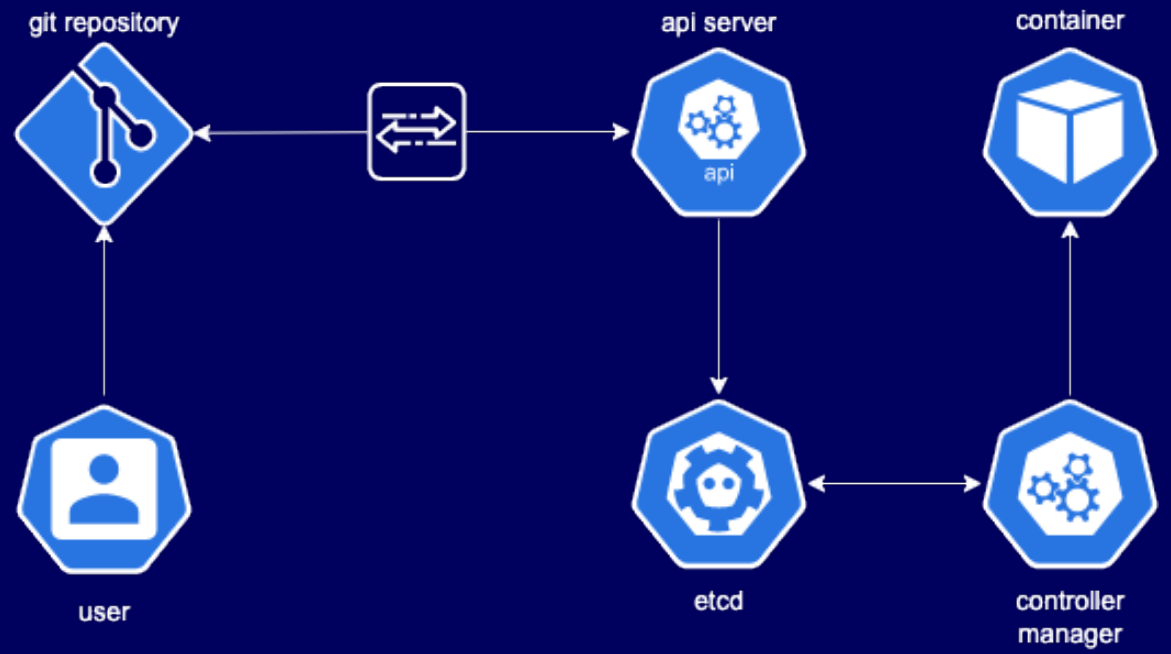
A continuous process which compares the actual state of the system to the desired state, and any differences are reconciled or corrected to bring the system back to the desired state.



Nordea

Kubernetes' Reconciliation Loop

A continuous process which compares the actual state of the system to the desired state, and any differences are reconciled or corrected to bring the system back to the desired state.



GitOps

The core idea of GitOps is having a Git repository that always contains declarative descriptions of the infrastructure currently desired in the production environment and an automated process to make the production environment match the described state in the repository.

- https://www.gitops.tech/#what-is-gitops

Nordea

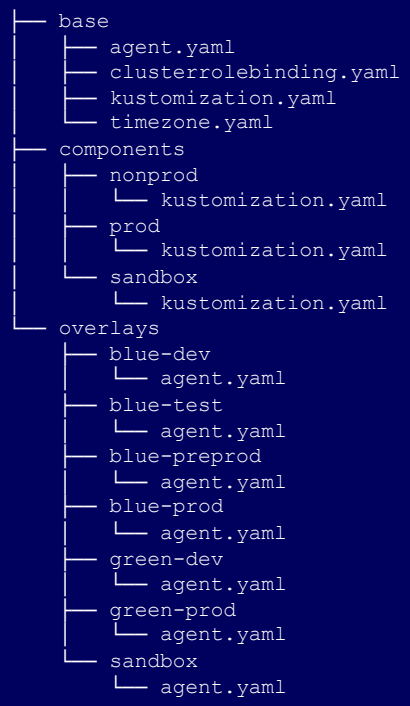
Our Setup

- We are currently responsible for 20 clusters…
  - … running almost 30k containers, …
  - … out of which more than 7k is in production.
- We are a platform team in charge of:
  - container orchestration,
  - application performance monitoring,
  - container runtime security,
  - log aggregation,
  - compliance policies, …
- Our setup is resilient as we are using 5 different environments:
  - sandbox,
  - development,
  - test,
  - preproduction,
  - production.
- All configuration is templated Kustomize and version controled. Let me show you the structure…

Our Setup

- We are currently responsible for 20 clusters…
  - … running almost 30k containers, …
  - … out of which more than 7k is in production.
- We are a platform team in charge of:
  - container orchestration,
  - application performance monitoring,
  - container runtime security,
  - log aggregation,
  - compliance policies, …
- Our setup is resilient as we are using 5 different environments:
  - sandbox,
  - development,
  - test,
  - preproduction,
  - production.
- All configuration is templated Kustomize and version controled. Let me show you the structure…

```
kubectl kustomize -h
```



Nordea

Chapter 2

The Incident

https://www.freepik.com/free-photo/businessman-video-conference-call-while-working-from-home-during-coronavirus-pandemic_18463004.htm
Image by **rawpixel.com** on **Freepik**

https://www.freepik.com/free-photo/man-with-glasses-student-computer-science-class-person-uses-computer_11183717.htm
Image by **prostooleh** on **Freepik**

```
% curl www.google.com

curl: (6) Could not resolve host: www.google.com
```

Any ideas what was going on?

Nordea

```
% kubectl get pods -n kube-system | grep coredns

NAME                       READY   STATUS             RESTARTS          AGE
coredns-5f798ff7bc-gnchw   1/1     CrashLoopBackOff   188 (28s ago)     29d
coredns-5f798ff7bc-s6qtq   1/1     CrashLoopBackOff   196 (26s ago)     29d
coredns-5f798ff7bc-sw6wm   1/1     CrashLoopBackOff   147 (26s ago)     29d
coredns-5f798ff7bc-zrrcw   1/1     CrashLoopBackOff   189 (29s ago)     29d
```


Nordea

Chapter 3

The Hunt

Kubernetes Audit Logs

Kubernetes *auditing* provides a security-relevant, chronological set of records documenting the sequence of actions in a cluster. The cluster audits the activities generated by users, by applications that use the Kubernetes API, and by the control plane itself.

- https://kubernetes.io/docs/tasks/debug/debug-cluster/audit/

Auditing allows cluster administrators to answer the following questions:

- what happened?
- when did it happen?
- who initiated it?
- on what did it happen?
- where was it observed?
- from where was it initiated?
- to where was it going?

Nordea

Kubernetes Audit Logs

Kubernetes *auditing* provides a security-relevant, chronological set of records documenting the sequence of actions in a cluster. The cluster audits the activities generated by users, by applications that use the Kubernetes API, and by the control plane itself.

- https://kubernetes.io/docs/tasks/debug/debug-cluster/audit/

Audit Policy

Audit policy defines rules about what events should be recorded and what data they should include. The defined audit levels are:

- None
- Metadata
- Request
- RequestResponse

- https://kubernetes.io/docs/tasks/debug/debug-cluster/audit/#audit-policy

Auditing allows cluster administrators to answer the following questions:

- what happened?
- when did it happen?
- who initiated it?
- on what did it happen?
- where was it observed?
- from where was it initiated?
- to where was it going?

Nordea

```
>> index="audit_logs" source="blue_prod"
```

```
<< Logs found: 58,033
```


Nordea

```
>> index="audit_logs" source="blue_prod”
   audit.verb IN (delete, create, patch, update)


<< Logs found: 7,672
```


Nordea

```
>> index="audit_logs" source="blue_prod”

   audit.verb IN (delete, create, patch, update)

   audit.objectRef.resource IN (configmaps, secrets, deployments,
     daemonsets, jobs, cronjobs, persistentvolumeclaims, persistentvolumes)

<< Logs found: 10
```


Nordea

Results

Those 10 log messages originated from only 3 unique namespaces:

- alpha-prod
- gamma-prod
- monitoring-agents

We spot the one!

Those 10 log messages originated from only 3 unique namespaces:

- alpha-prod
- gamma-prod
- monitoring-agents


```
{
  "audit": {
    "annotations": { --- },
    "auditID": "xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx",
    "level": "Metadata",
    "metadata": {---},
    "objectRef": {
      "apiGroup": "apps",
      "apiVersion": "v1",
      "name": "monitoring-agent",
      "namespace": "monitoring-agents",
      "resource": "daemonsets"
    },
    "requestReceivedTimestamp": "2023-xx-xxTxx:xx:xx.xxxxxxZ",
    "requestURI": "/apis/apps/v1/namespaces/monitoring-agents/daemonsets/monitoring-
agent?fieldManager=gitops-controller",
    "responseStatus": {
      "code": 200,
      "metadata": {}
    },
    "sourceIPs": [ --- ],
    "stage": "ResponseComplete",
    "stageTimestamp": "2023-xx-xxTxx:xx:xx.xxxxxxZ",
    "timestamp": null,
    "user": { --- },
    "userAgent": "gitops-application-controller/v0.0.0 (linux/amd64)
kubernetes/$Format",
    "verb": "patch"
  },
  "level": "info",
  "msg": "audit",
  "time": "2023-xx-xxTxx:xx:xxZ"
}
```


Nordea

Map it to a commit

```
{
  "audit": {
    "annotations": { --- },
    "auditID": "xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx",
    "level": "Metadata",
    "metadata": {---},
    "objectRef": {
      "apiGroup": "apps",
      "apiVersion": "v1",
      "name": "monitoring-agent",
      "namespace": "monitoring-agents",
      "resource": "daemonsets"
    },
    "requestReceivedTimestamp": "2023-xx-xxTxx:xx:xx.xxxxxxZ",
    "requestURI": "/apis/apps/v1/namespaces/monitoring-agents/daemonsets/monitoring-
agent?fieldManager=gitops-controller",
    "responseStatus": {
      "code": 200,
      "metadata": {}
    },
    "sourceIPs": [ --- ],
    "stage": "ResponseComplete",
    "stageTimestamp": "2023-xx-xxTxx:xx:xx.xxxxxxZ",
    "timestamp": null,
    "user": { --- },
    "userAgent": "gitops-application-controller/v0.0.0 (linux/amd64)
kubernetes/$Format",
    "verb": "patch"
  },
  "level": "info",
  "msg": "audit",
  "time": "2023-xx-xxTxx:xx:xxZ"
}
```

```
% git log --after "2023-xx-xx" --before "2023-xx-yy"

commit d3b0ab29119f6a6af9759369a9f769074175086f
Merge: ca0b3b6 22fcb8d
Author: Xxxxx, Xxxxx <xxxxx.xxxxx@nordea.com>
Date: DoW MMM DD xx:xx:xx xxxx +0200

    Pull request #141: Added prom service discovery

    Merge in INFRA/cluster-bootstrap from feature/xxxx to master

    * commit '22fcb8d8cf9a106c7ce377507ccc1b4c9979cd76':
    Added monitoring-federation

commit 22fcb8d8cf9a106c7ce377507ccc1b4c9979cd76
Author: Xxxxx Xxxxx <xxxxx.xxxxx@nordea.com>
Date: DoW MMM DD xx:xx:xx xxxx +0200

    Added prom service discovery
```


Nordea

Single line changed…

```
% git log --after "2023-xx-xx" --before "2023-xx-yy"

commit d3b0ab29119f6a6af9759369a9f769074175086f
Merge: ca0b3b6 22fcb8d
Author: Xxxxx, Xxxxx <xxxxx.xxxxx@nordea.com>
Date: DoW MMM DD xx:xx:xx xxxx +0200

    Pull request #141: Added prom service discovery

    Merge in INFRA/cluster-bootstrap from feature/xxxx to master

    * commit '22fcb8d8cf9a106c7ce377507ccc1b4c9979cd76':
    Added monitoring-federation

commit 22fcb8d8cf9a106c7ce377507ccc1b4c9979cd76
Author: Xxxxx Xxxxx <xxxxx.xxxxx@nordea.com>
Date: DoW MMM DD xx:xx:xx xxxx +0200

    Added prom service discovery
```

services/monitoring/overlays/blue-prod/agent.yaml

```
14  14      k8s_extra_resources:
15  15        include:
16  16          - services
17  17          - resourcequotas
18  18          - persistentvolumes
19  19          - persistentvolumeclaims
20  20          - horizontalpodautoscalers
21  21      prometheus:x
22  22        enabled: true
    23 +      prom_service_discovery: true
23  24      metrics_excess_log: true
24  25      metrics_filter:
25  26        - include: "kube_workload_*"
26  27        - include: "kubelet_volume_*"
27  28        - include: "kubelet_volume_stats_used_bytes"
28  29        - exclude: "kubelet_running_container_count"
29  30        - exclude: "kubelet_*"
30  31        - exclude: "volume_*"
31  32        - exclude: "storage_*"
32  33        - exclude: "storage_operation_errors_total*"
```


Nordea

Single line changed…

```
% git log --after "2023-xx-xx" --before "2023-xx-yy"

commit d3b0ab29119f6a6af9759369a9f769074175086f
Merge: ca0b3b6 22fcb8d
Author: Xxxxx, Xxxxx <xxxxx.xxxxx@nordea.com>
Date: DoW MMM DD xx:xx:xx xxxx +0200

    Pull request #141: Added prom service discovery

    Merge in INFRA/cluster-bootstrap from feature/xxxx to master

    * commit '22fcb8d8cf9a106c7ce377507ccc1b4c9979cd76':
    Added monitoring-federation

commit 22fcb8d8cf9a106c7ce377507ccc1b4c9979cd76
Author: Xxxxx Xxxxx <xxxxx.xxxxx@nordea.com>
Date: DoW MMM DD xx:xx:xx xxxx +0200

    Added prom service discovery
```

services/monitoring/overlays/blue-prod/agent.yaml

```
14  14      k8s_extra_resources:
15  15        include:
16  16          - services
17  17          - resourcequotas
18  18          - persistentvolumes
19  19          - persistentvolumeclaims
20  20          - horizontalpodautoscalers
21  21      prometheus:x
22  22        enabled: true
    23 +      prom_service_discovery: true
23  24      metrics_excess_log: true
24  25      metrics_filter:
25  26        - include: "kube_workload_*"
26  27        - include: "kubelet_volume_*"
27  28        - include: "kubelet_volume_stats_used_bytes"
28  29        - exclude: "kubelet_running_container_count"
29  30        - exclude: "kubelet_*"
30  31        - exclude: "volume_*"
31  32        - exclude: "storage_*"
32  33        - exclude: "storage_operation_errors_total*"
```

It took less than 5min to revert the change and get cluster back to the normal condition.

Nordea

Chapter 4

The Investigation

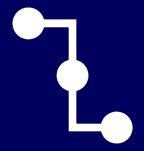
What had really happened?

Nordea

A series of unfortunate events happened

Nordea

A series of unfortunate events happened

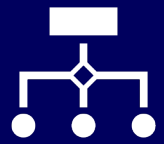


The change was previously tested on sandbox, development and test environments without visible issues.

Nordea

A series of unfortunate events happened

The change was previously tested on sandbox, development and test environments without visible issues.



The production config change was bundled into the same merge request as preproduction one.

Nordea

A series of unfortunate events happened

The change was previously tested on sandbox, development and test environments without visible issues.

The production config change was bundled into the same merge request as preproduction one.

Others approving the merge request did not fully understand the impact.

Nordea

A series of unfortunate events happened

The change was previously tested on sandbox, development and test environments without visible issues.

The production config change was bundled into the same merge request as preproduction one.

Others approving the merge request did not fully understand the impact.

The colleague making the change worked with a wrong assumption about how a synchronization is set up.

Nordea

Chapter 5

The Learning

1. Use testing environments

Nordea

2. One change per environment

Nordea

3. Four eyes can see more than two

Nordea

4. When making larger changes, use synchronization window

Nordea

5. Beware of actual changes while using templating

base/kustomization.yaml

```
1  1     apiVersion: kustomize.config.k8s.io/v1beta1
2  2     kind: Kustomization
3  3
4  4     resources:
5  5     - deployment.yaml
   6 +   - role.yaml
```


Nordea

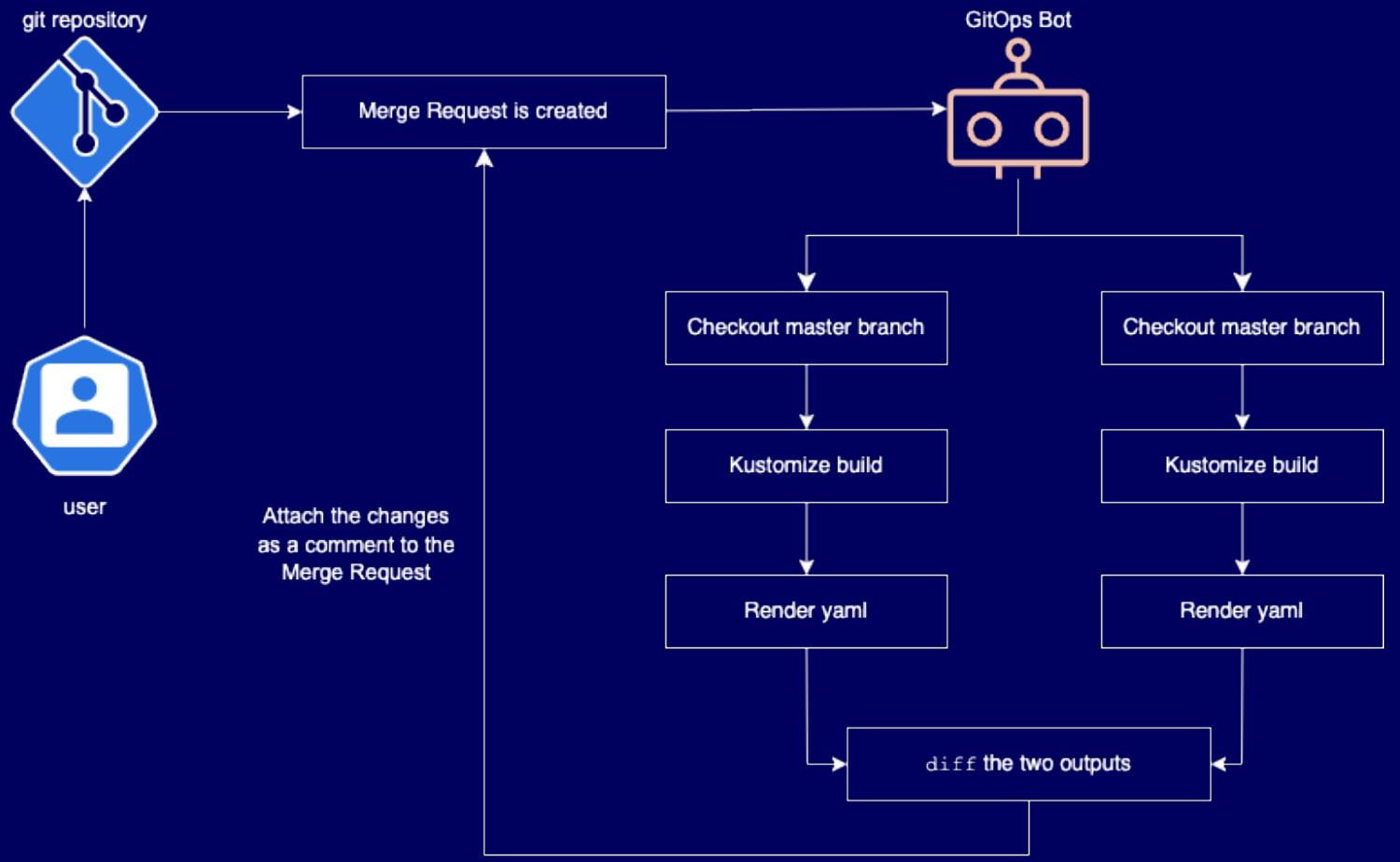
base/kustomization.yaml

```
1  1      apiVersion: kustomize.config.k8s.io/v1beta1
2  2      kind: Kustomization
3  3
4  4      resources:
5  5      - deployment.yaml
   6 +    - role.yaml
```

base/role.yaml

```
 1   1      apiVersion: apps/v1
 2   2      kind: Deployment
 3   3      metadata:
 4   4        name: foo
 5   5        namespace: bar
 6   6      spec:
 7   7        replicas: 1
 8   8        template:
 9   9          spec:
10  10            containers:
11  11            - image: nginx:1.14.4
12  12              name: nginx
13  13              ports:
14  14              - containerPort: 80
    15 +     ---
    16 +     kind: Role
    17 +     metadata:
    18 +     name: pod-reader
    19 +     namespace: default
    20 +     rules:
    21 +     - apiGroups:
    22 +     - ''
    23 +     resources:
    24 +     - pods
    25 +     verbs:
    26 +     - get
    27 +     - watch
    28 +     - list
```


Nordea


git repository
Merge Request is created
GitOps Bot
Checkout master branch
Checkout master branch
Kustomize build
Kustomize build
Render yaml
Render yaml
diff the two outputs
user
Attach the changes as a comment to the Merge Request

Nordea

To summarize…

0. Audit logs are your friend

Nordea

To summarize…

0. Audit logs are your friend
1. Use more then one environment for testing

Nordea

To summarize…

0. Audit logs are your friend
1. Use more then one environment for testing
2. When making a git change always target only single environment

Nordea

To summarize…

0. Audit logs are your friend
1. Use more then one environment for testing
2. When making a git change always target only single environment
3. Make a good use of four eyes principle

Nordea

To summarize…

0. Audit logs are your friend
1. Use more then one environment for testing
2. When making a git change always target only single environment
3. Make a good use of four eyes principle
4. Utilize / configure a predictable synchronization window

Nordea

To summarize…

0. Audit logs are your friend
1. Use more then one environment for testing
2. When making a git change always target only single environment
3. Make a good use of four eyes principle
4. Utilize / configure a predictable synchronization window
5. When templating simple `git diff` might not enough

Nordea

To summarize…

0. Audit logs are your friend
1. Use more then one environment for testing
2. When making a git change always target only single environment
3. Make a good use of four eyes principle
4. Utilize / configure a predictable synchronization window
5. When templating simple `git diff` might not enough

Nordea

Nordea

Thank you!

Any questions?