МЕГА СЛЁРМ

+

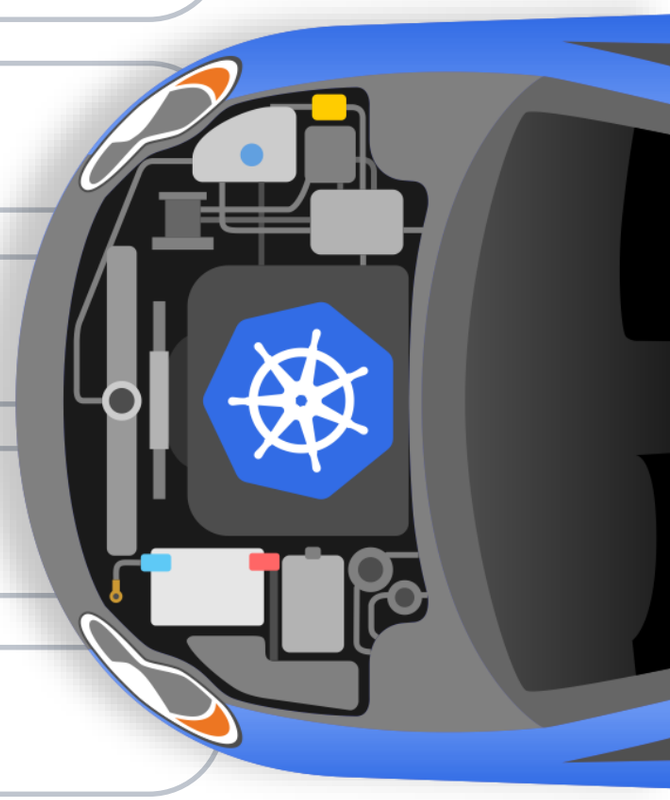
Southbridge

Kubernetes под капотом

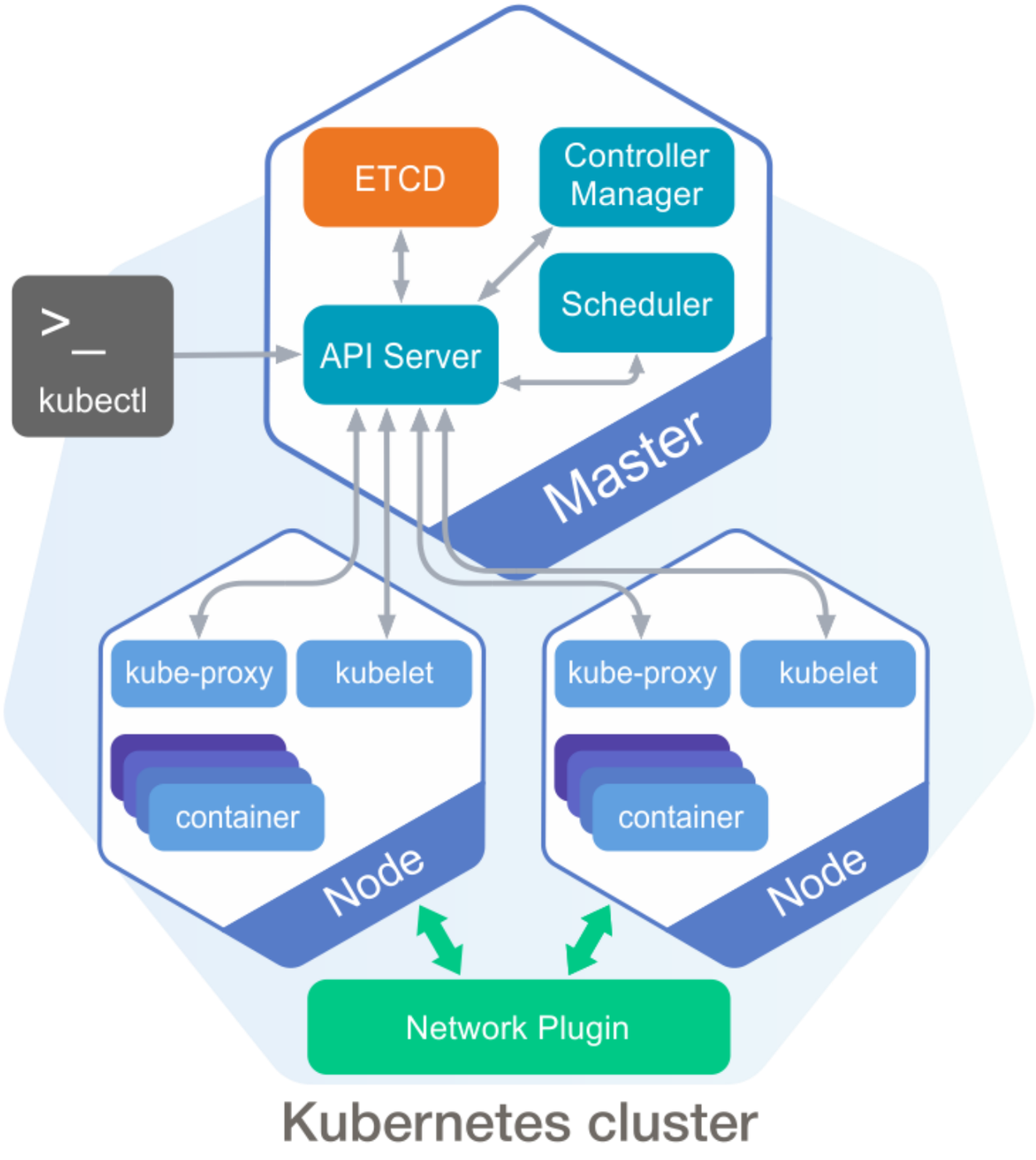
slurm.io

План

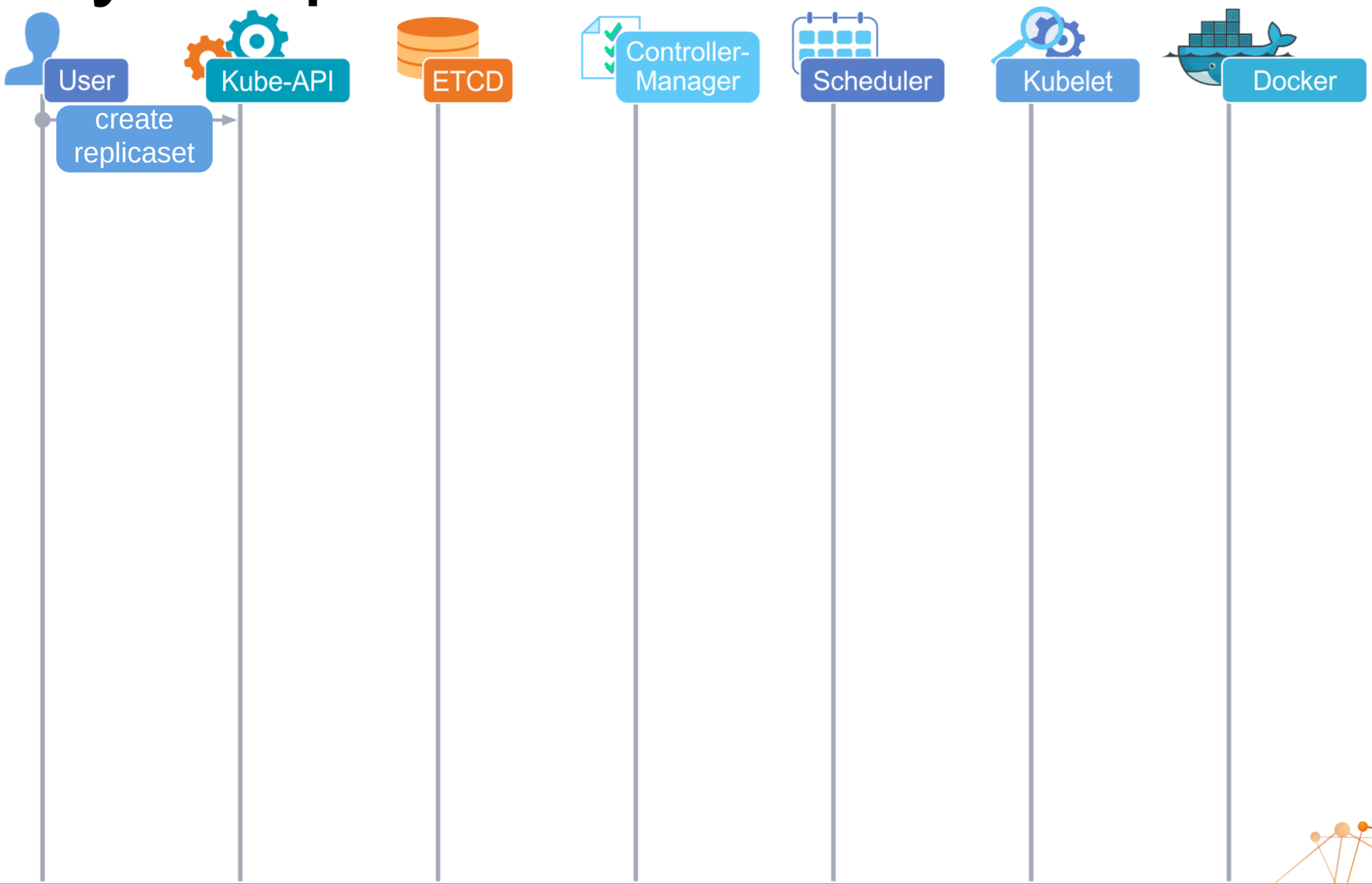
- Структура Kubernetes
- Scheduler
- Контроллеры
- CRD
- Операторы
- Operator SDK



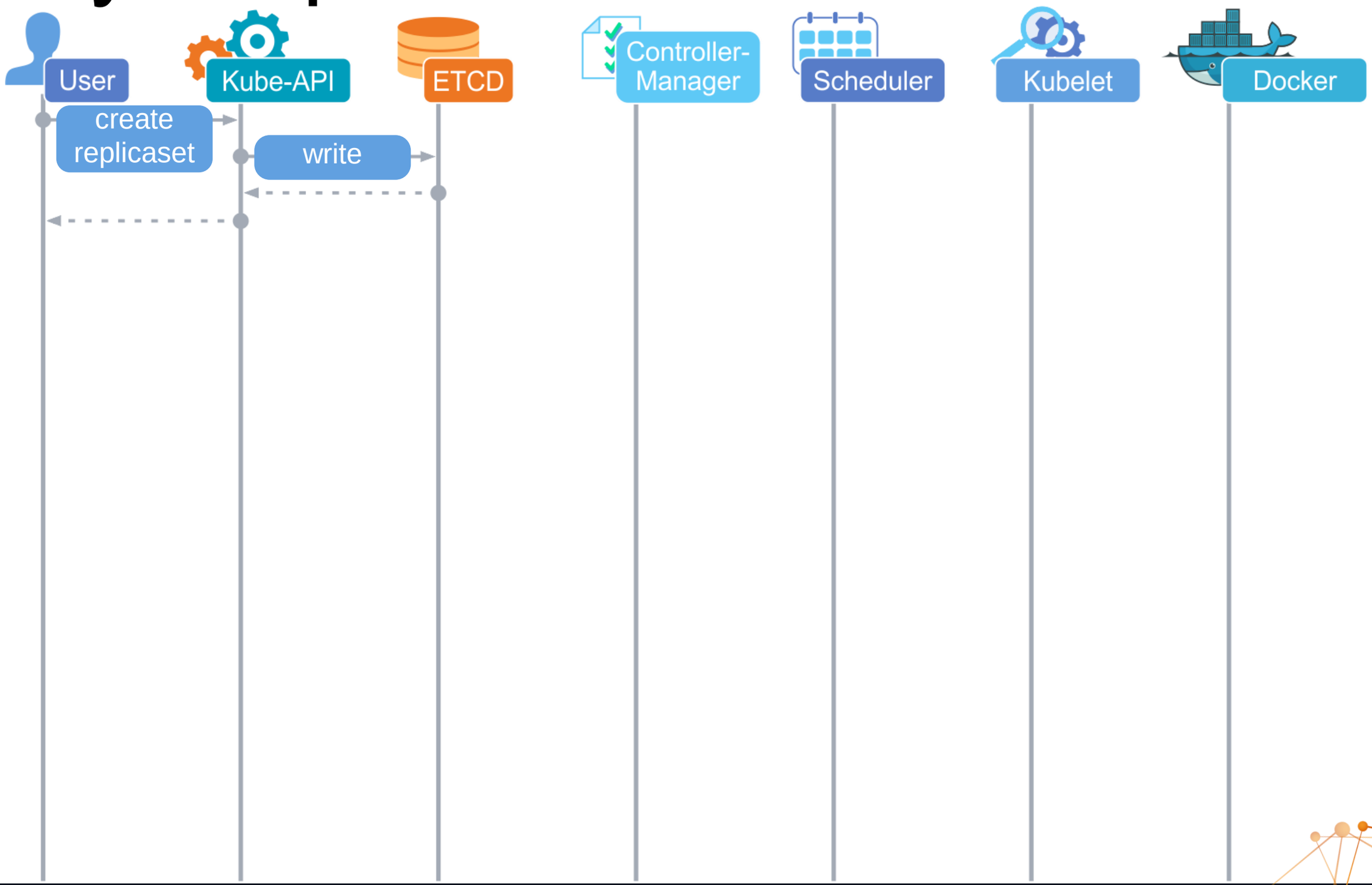
Kubernetes издалека



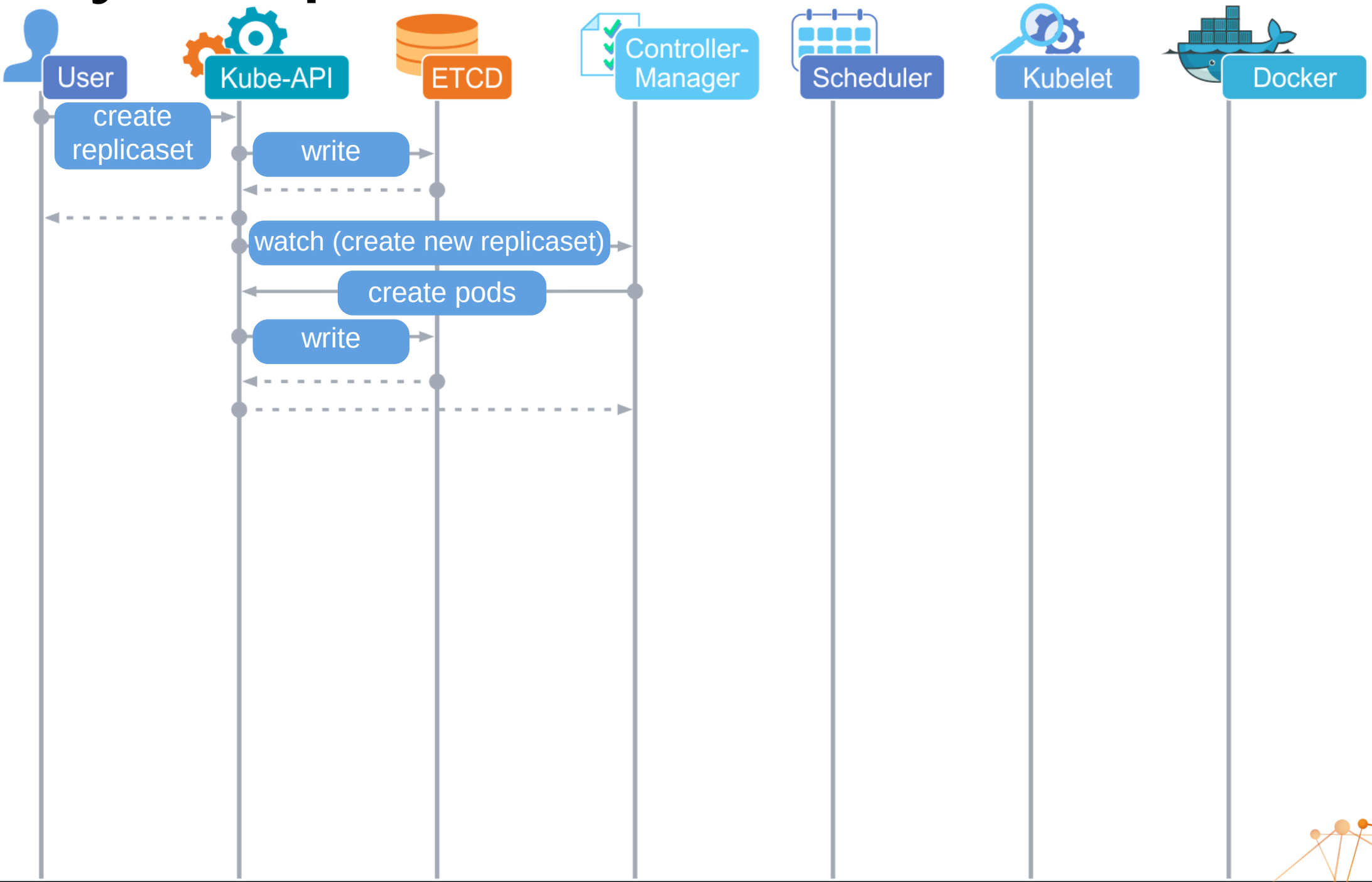
Запуск приложения



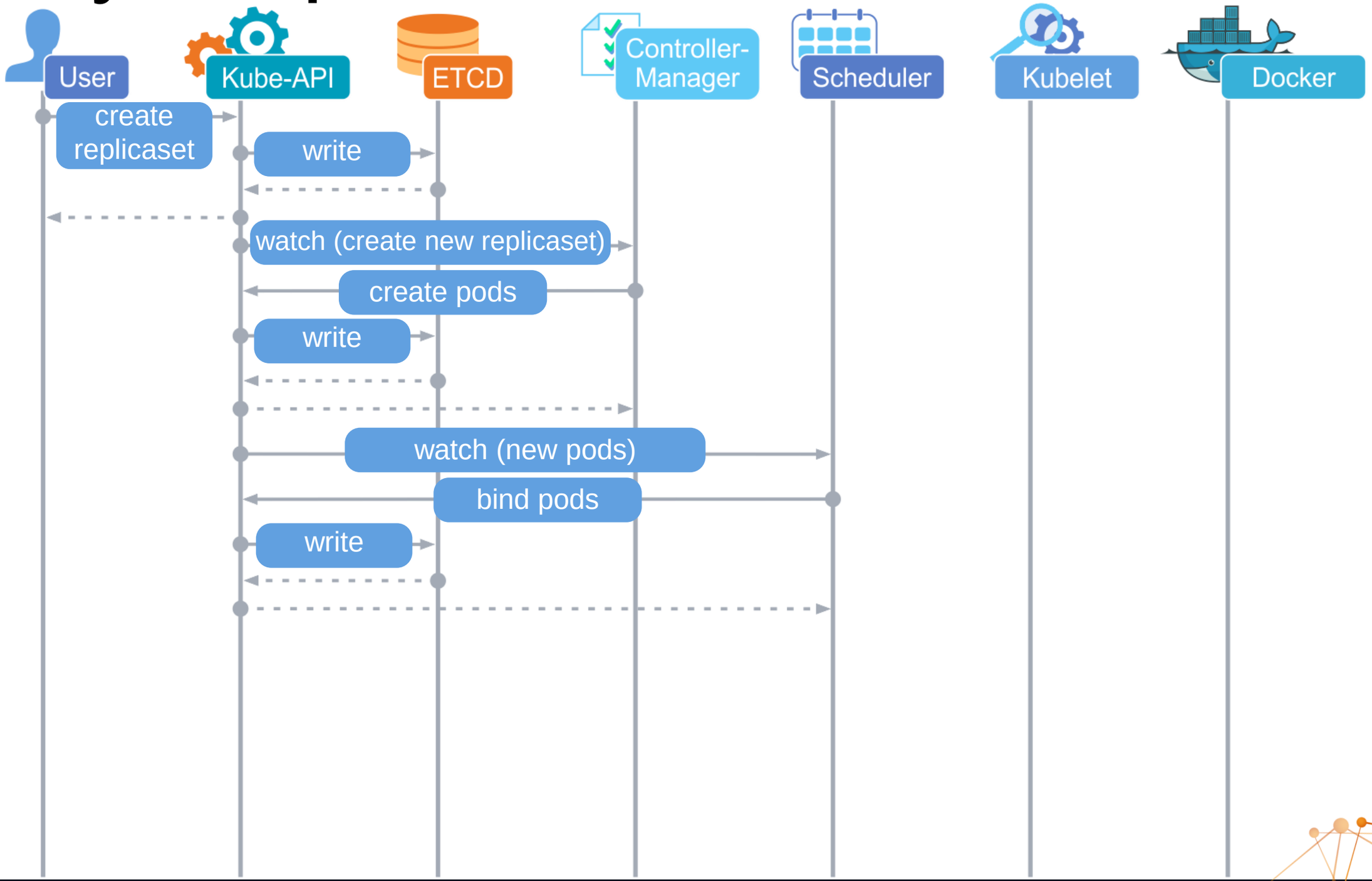
Запуск приложения



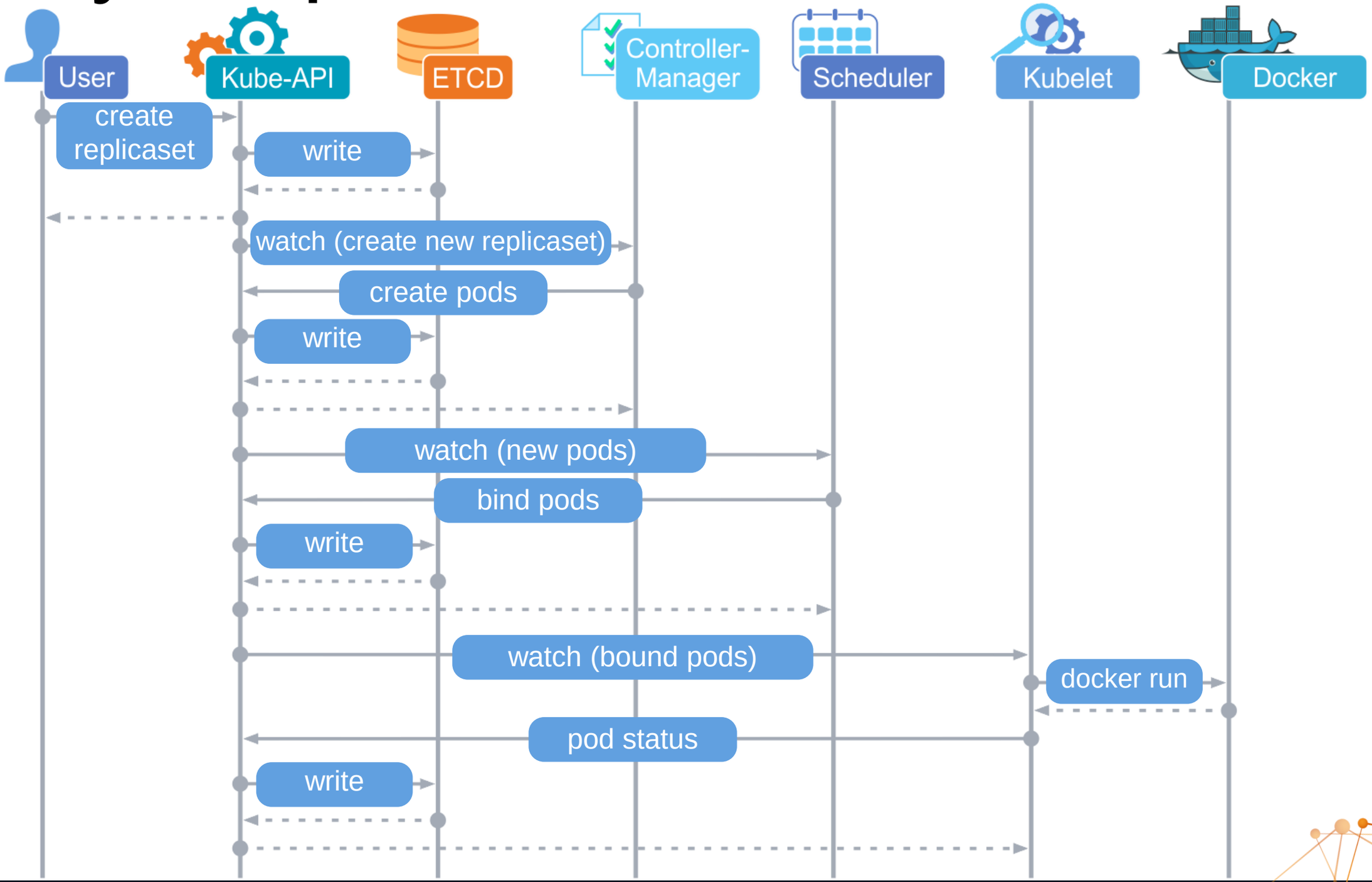
Запуск приложения



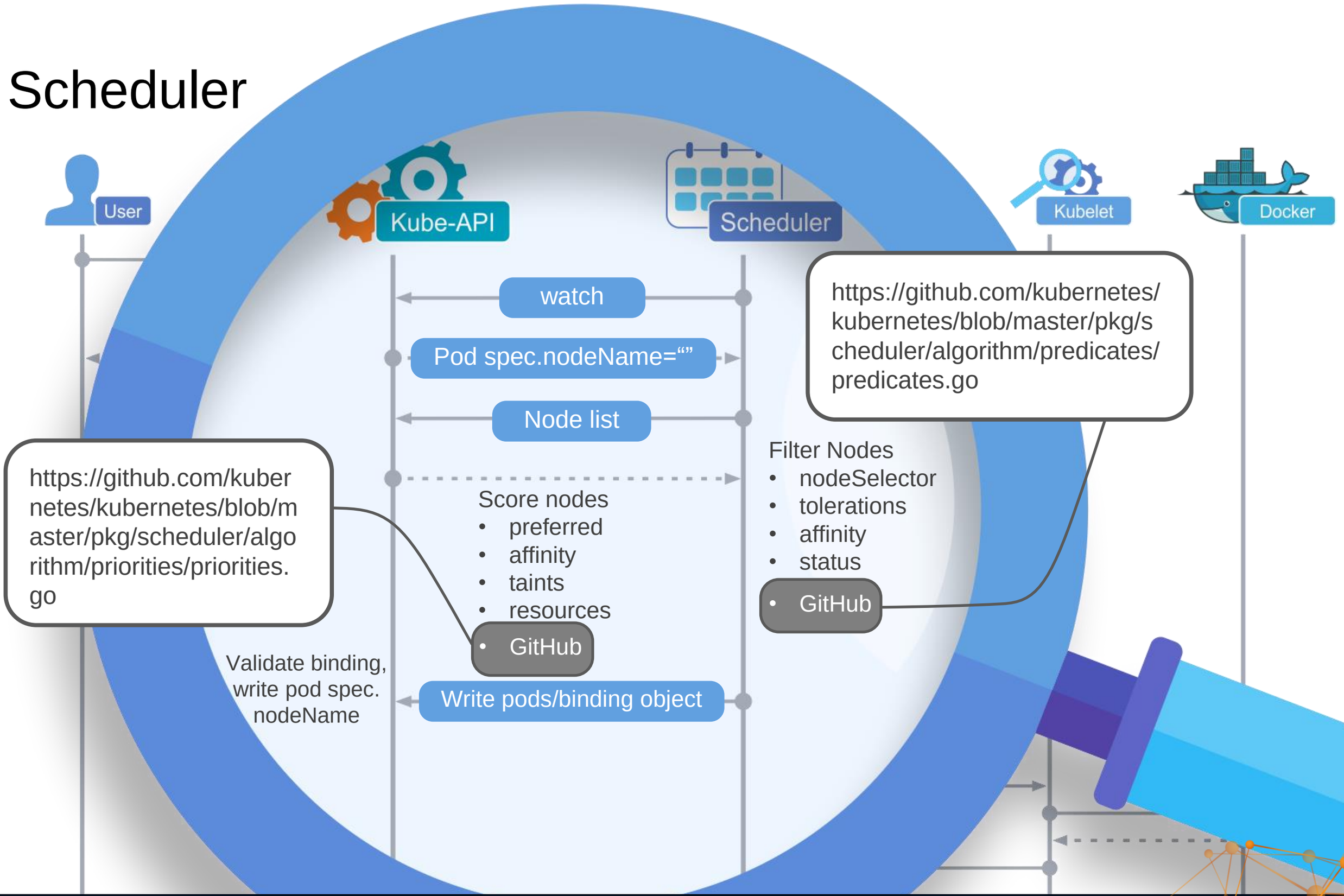
Запуск приложения



Запуск приложения



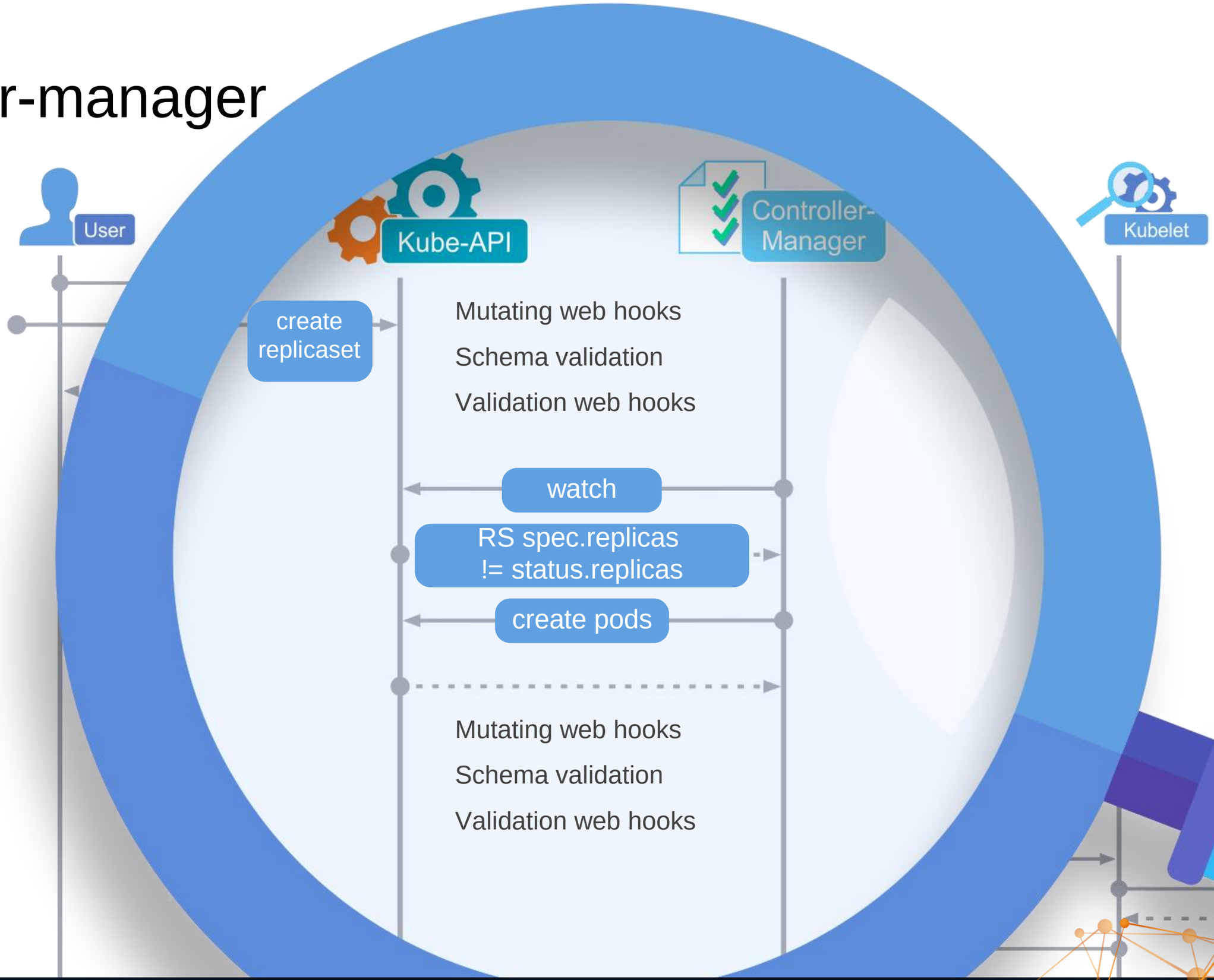
Scheduler


User
Kube-API
Scheduler
Kubelet
Docker
watch
Pod spec.nodeName=“”
Node list
Score nodes
• preferred
• affinity
• taints
• resources
• GitHub
Filter Nodes
• nodeSelector
• tolerations
• affinity
• status
• GitHub
Validate binding, write pod spec. nodeName
Write pods/binding object
https://github.com/kubernetes/kubernetes/blob/master/pkg/scheduler/algorithm/predicates/predicates.go
https://github.com/kubernetes/kubernetes/blob/master/pkg/scheduler/algorithm/priorities/priorities.go

Scheduler

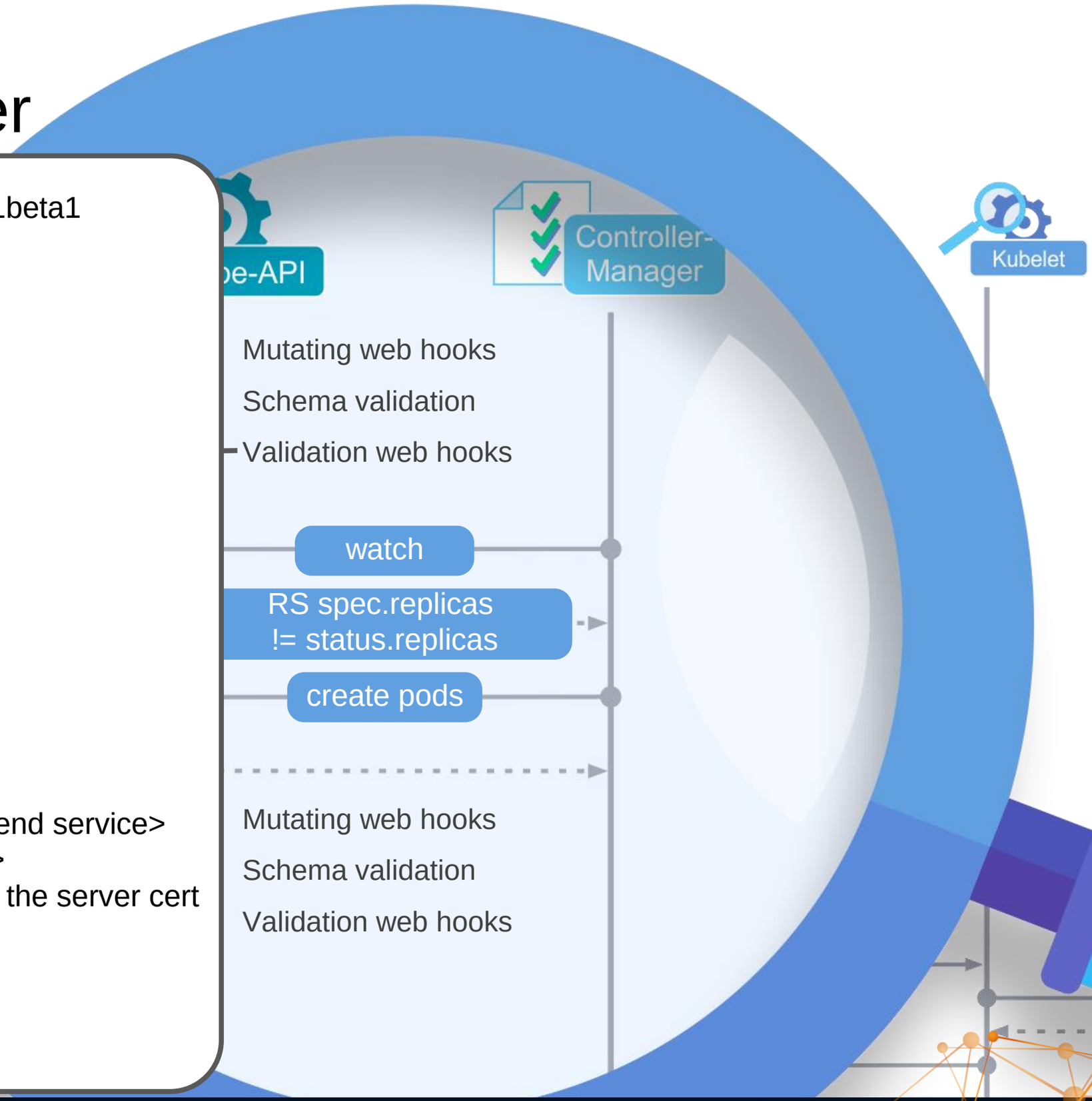
```
while true;

do
  for PODNAME in $(kubectl --server $SERVER get pods -o json | jq '.items[] |
select(.spec.nodeName == null) | .metadata.name' | tr -d '"');

  do

    NODES=($(kubectl --server $SERVER get nodes -o json | jq '.items[].metadata.name' |
tr -d '"'))
    NUMNODES=${#NODES[@]}
    CHOSEN=${NODES[$[$RANDOM % $NUMNODES]]}

    curl --header "Content-Type:application/json" --request POST --data '
      {
        "apiVersion":"v1",
        "kind": "Binding",
         "metadata":
         {
           "name": "'$PODNAME'"
         },
         "target":
         {
           "apiVersion": "v1",
           "kind": "Node",
           "name": "'$CHOSEN'"
         }
      }
    ' http://$SERVER/api/v1/namespaces/default/pods/$PODNAME/binding/

  done

done
```

Controller-manager

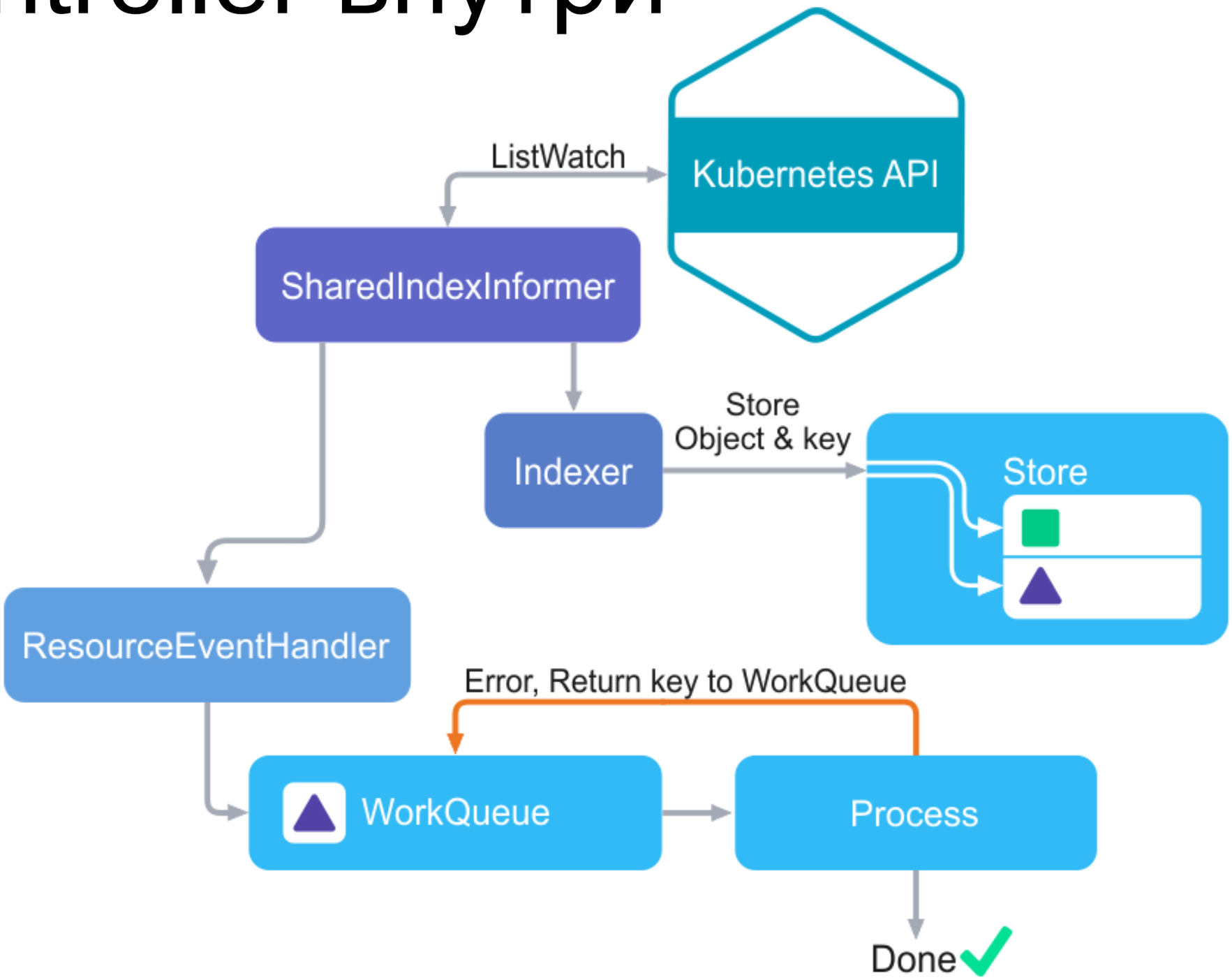


Controller-manager

```
apiVersion: admissionregistration.k8s.io/v1beta1
kind: ValidatingWebhookConfiguration
metadata:
  name: test
webhooks:
- name: test.southbridge.io
  rules:
  - apiGroups:
    - ""
    apiVersions:
    - v1/apps
    operations:
    - CREATE
    resources:
    - replicates
 clientConfig:
    service:
      namespace: <namespace of the front-end service>
      name: <name of the front-end service>
    caBundle: <pem encoded ca cert, signs the server cert
used by the webhook>
  admissionReviewVersions:
  - v1beta1
  timeoutSeconds: 1
```

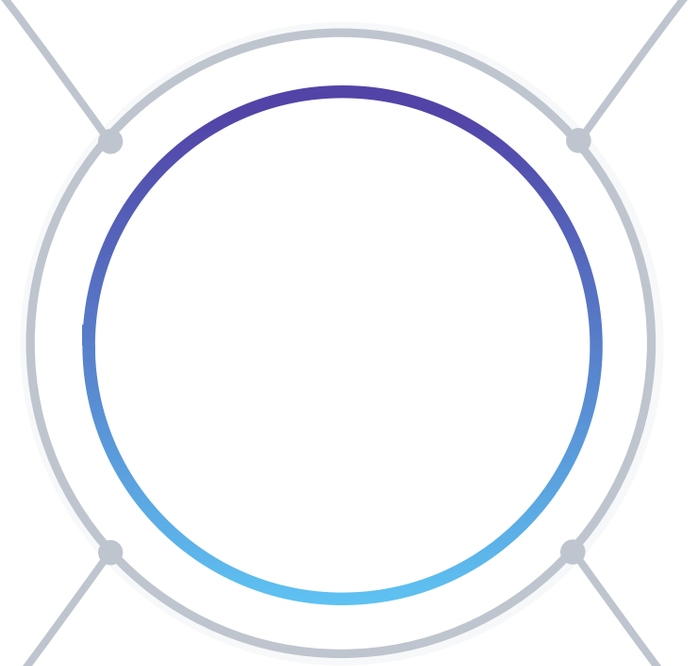


Controller внутри



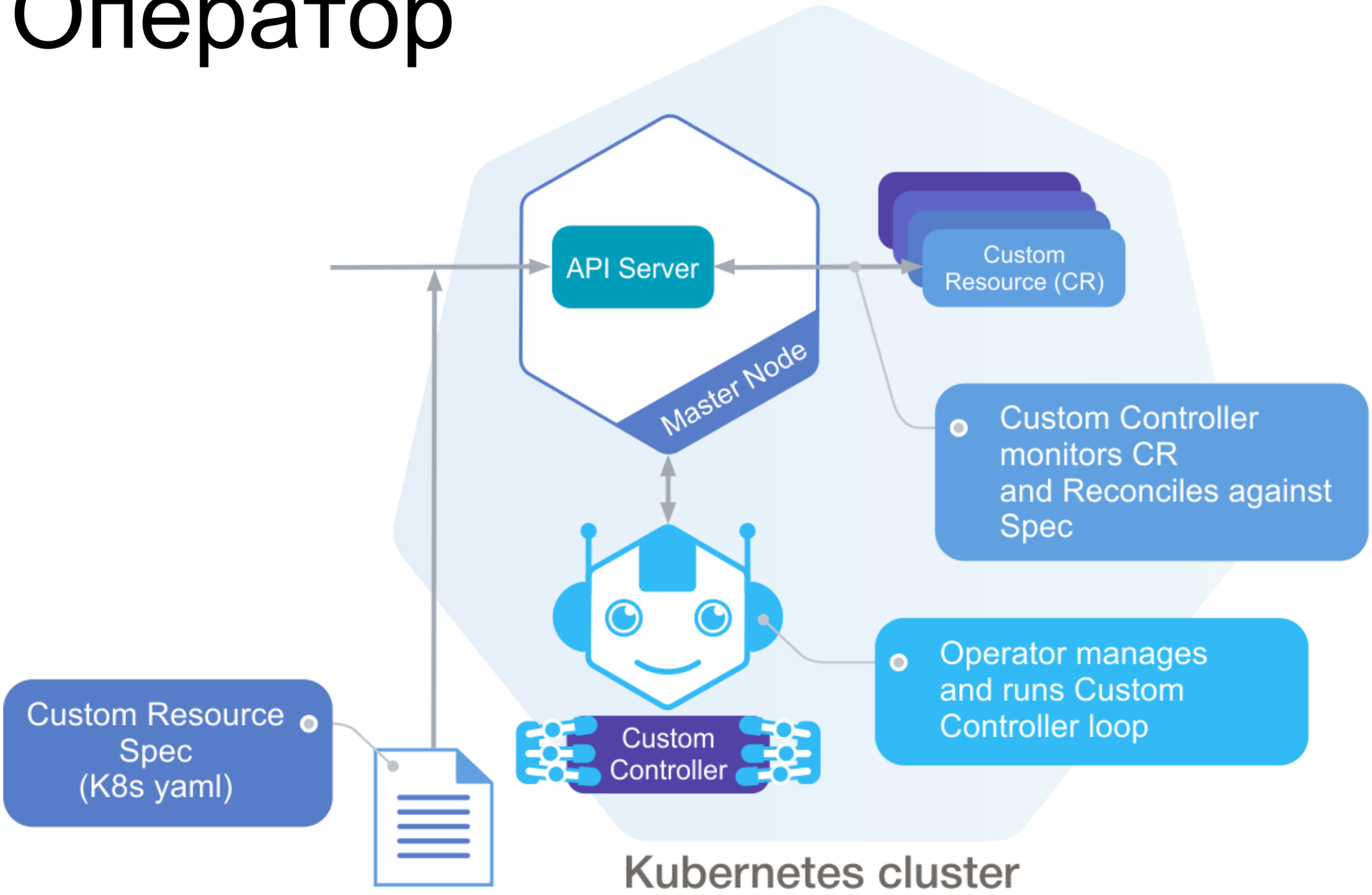
Custom Resource Definition

- Позволяет создавать новые объекты в кластере
- Настраиваемый вывод в kubectl get



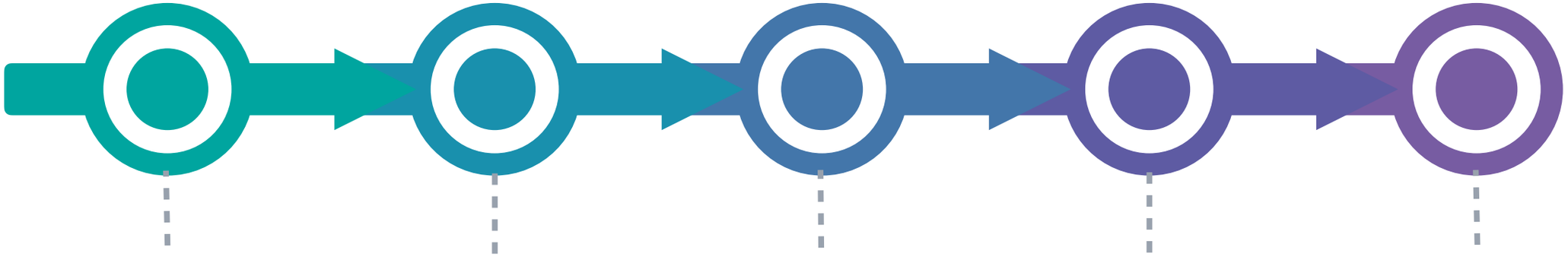
- Настраиваемая валидация
- Обрабатывается операторами

Оператор


API Server
Master Node
Custom Resource (CR)
Custom Controller monitors CR and Reconciles against Spec
Operator manages and runs Custom Controller loop
Custom Controller
Custom Resource Spec (K8s yaml)
Kubernetes cluster

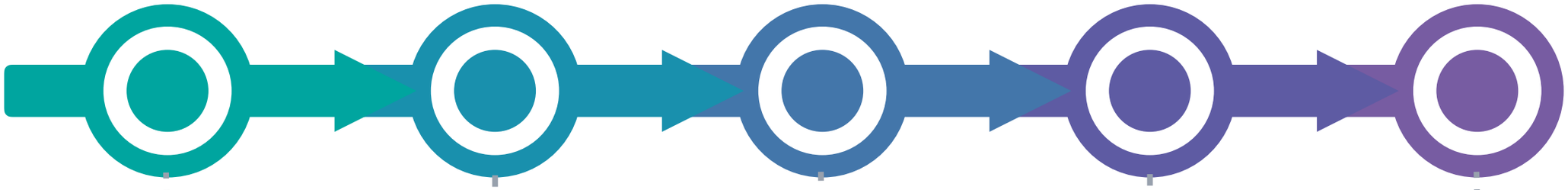
Operator SDK

https://github.com/operator-framework/operator-sdk



1. Создаем новый проект при помощи SDK Command Line Interface (CLI)
2. Создаем CRD
3. Пишем логику обработки CRD
4. С помощью SDK CLI собираем оператор
5. Деплоим в кластер

Operator SDK



1. Создаем новый проект при помощи SDK Command Line Interface (CLI)
2. Создаем CRD
3. Пишем логику обработки CRD
4. С помощью SDK CLI собираем оператор
5. Деплоим в кластер

МЕГА СЛЁРМ

+

Southbridge

Перерыв

slurm.io