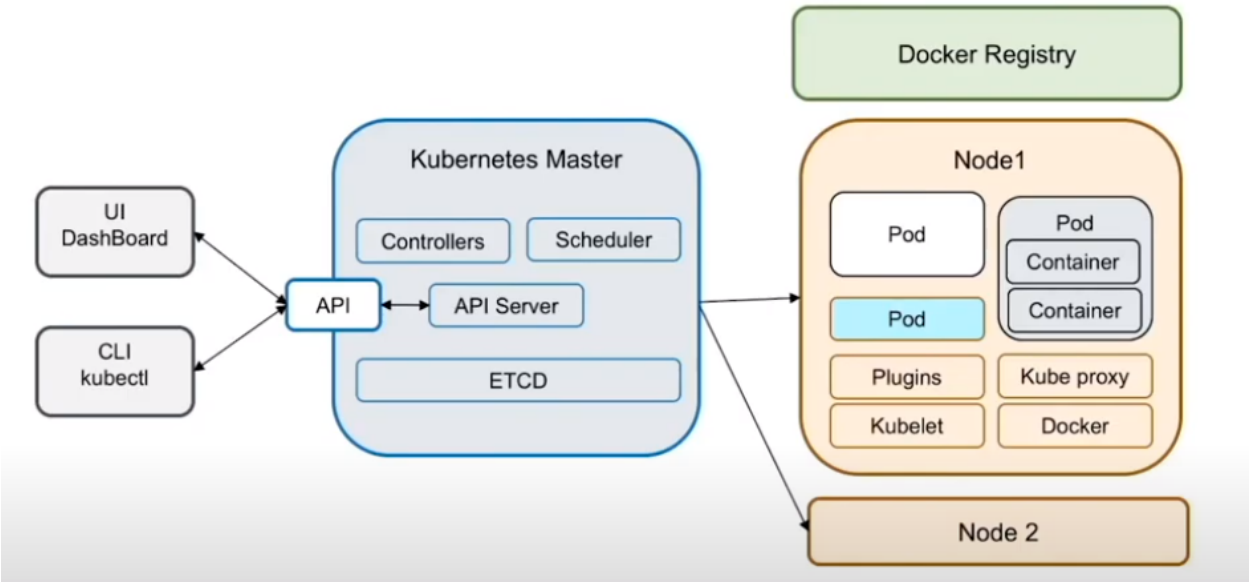
Flink Concepts

Kuberneties Architehture

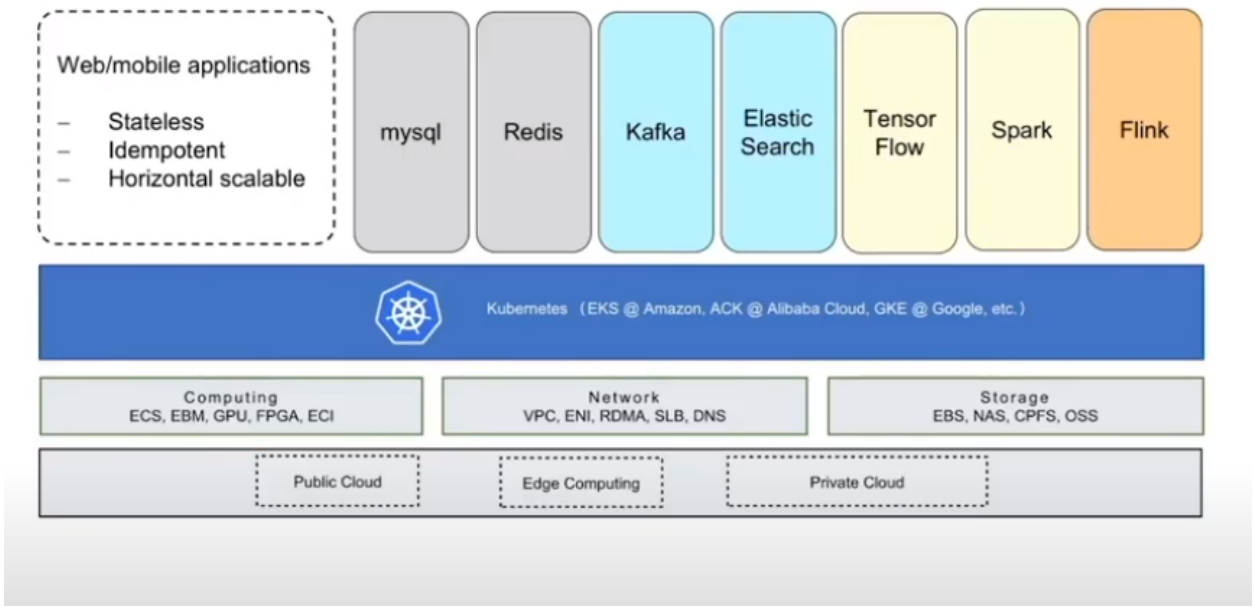


Config map : is dictionary of configuration settings .This dictionary consist of key value pair of strings

Service : Is an abstract way to expose and application running on a set of pods as a network service

Pod : Smallest deployable unit ,consist of one or more containers

Deployment : is a higher-level abstraction to manage set of identical pods



Flink's Runtime Building Blocks

Dispatcher

- Touch-point for job submissions
- Spawns JobManagers

ResourceManager

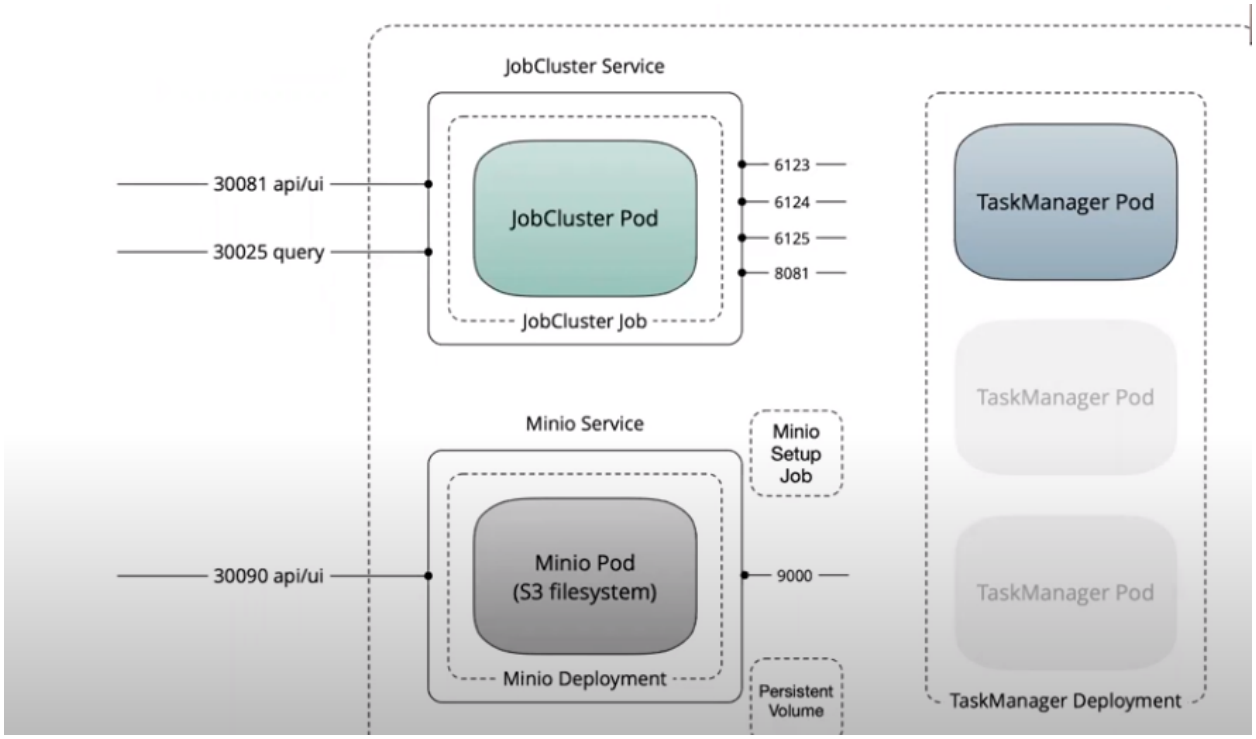
- Cluster framework-specific
- Manages available TaskManagers
- Acquires / releases resources

JobManager

- One per job
- Schedules job in terms of "task slots"
- Monitors task execution
- Coordinates checkpointing

TaskManager

- Registers with ResourceManager
- Provides "task slots"
- Assigned tasks by JobManager(s)



1: Build a docker image

Dockerfile

```
ADD $flink_dist $FLINK_INSTALL_PATH
ADD $job_jar $FLINK_INSTALL_PATH/job.jar
. . .
COPY docker/flink/flink-conf.yaml $FLINK_HOME/conf
COPY docker/flink/log4j-console.properties $FLINK_HOME/conf
COPY docker/flink/docker-entrypoint.sh /
. . .
ENTRYPOINT ["/docker-entrypoint.sh"]
```

docker-entrypoint.sh

```
. . .
JOB_CLUSTER="job-cluster"
TASK_MANAGER="task-manager"

CMD="$1"
shift;

if [ "${CMD}" == "${JOB_CLUSTER}" -o "${CMD}" == "${TASK_MANAGER}" ]; then
    if [ "${CMD}" == "${TASK_MANAGER}" ]; then
        exec $FLINK_HOME/bin/taskmanager.sh start-foreground "$@"
    else
        exec $FLINK_HOME/bin/standalone-job.sh start-foreground "$@"
    fi
fi

exec "$@"
```

2: K8s manifests

task-manager-deployment.yaml.template

```
apiVersion: extensions/v1beta1
kind: Deployment
metadata:
  name: flink-task-manager
spec:
  replicas: ${FLINK_NUM_OF_TASKMANAGERS}
  template:
    metadata:
      labels:
        app: flink
        component: task-manager
    spec:
      containers:
      - name: flink-task-manager
        image: ${FLINK_IMAGE_NAME}
        imagePullPolicy: Never
        args: ["task-manager",
               "-Djobmanager.rpc.address=flink-job-cluster"]
```

job-cluster-job.yaml.template

```
apiVersion: batch/v1
kind: Job
metadata:
  name: flink-job-cluster
spec:
  template:
    metadata:
      labels:
        app: flink
        component: job-cluster
    spec:
      restartPolicy: OnFailure
      containers:
      - name: flink-job-cluster
        image: ${FLINK_IMAGE_NAME}
        imagePullPolicy: Never
        args: ["job-cluster",
               "-Djobmanager.rpc.address=flink-job-cluster",
               "-Dblob.server.port=6124",
               "-Dqueryable-state.server.ports=6125"]
        ports:
        - containerPort: 6123
          name: rpc
        - containerPort: 6124
          name: blob
        - containerPort: 6125
          name: query
        - containerPort: 8081
          name: ui
```

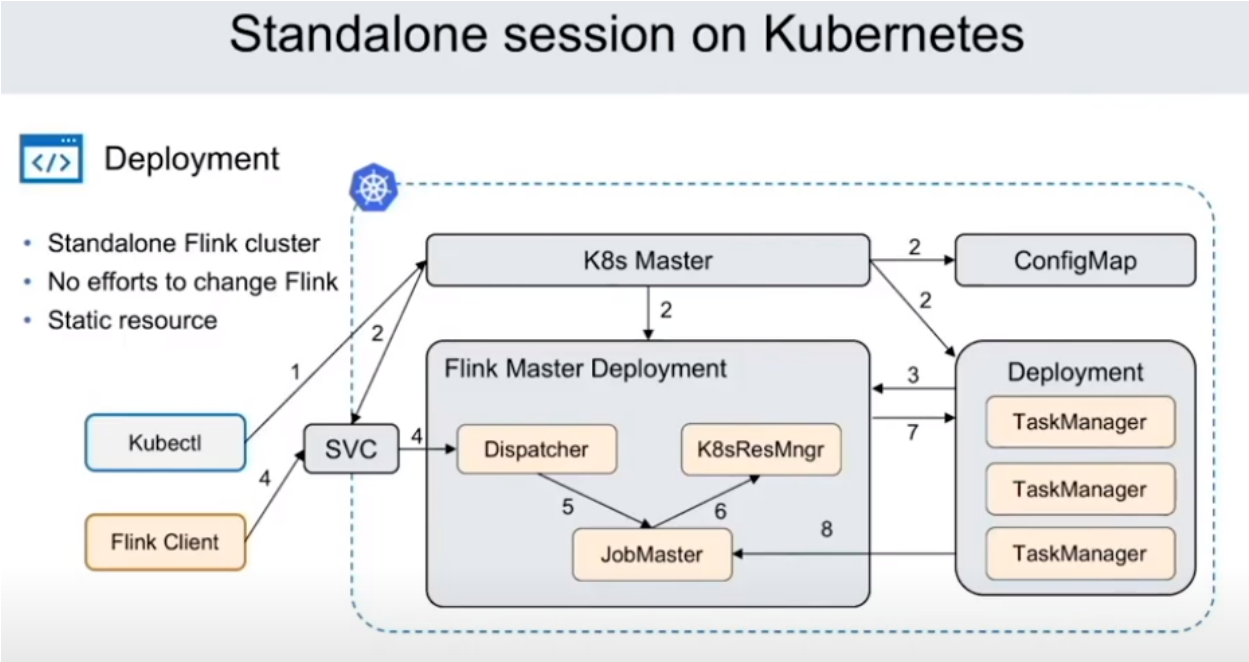

job-cluster-service.yaml

```
apiVersion: v1
kind: Service
metadata:
  name: flink-job-cluster
  labels:
    app: flink
    component: job-cluster
spec:
  ports:
    - name: rpc
      port: 6123
    - name: blob
      port: 6124
    - name: query
      port: 6125
      nodePort: 30025
    - name: ui
      port: 8081
      nodePort: 30081
  type: NodePort
  selector:
    app: flink
    component: job-cluster
```

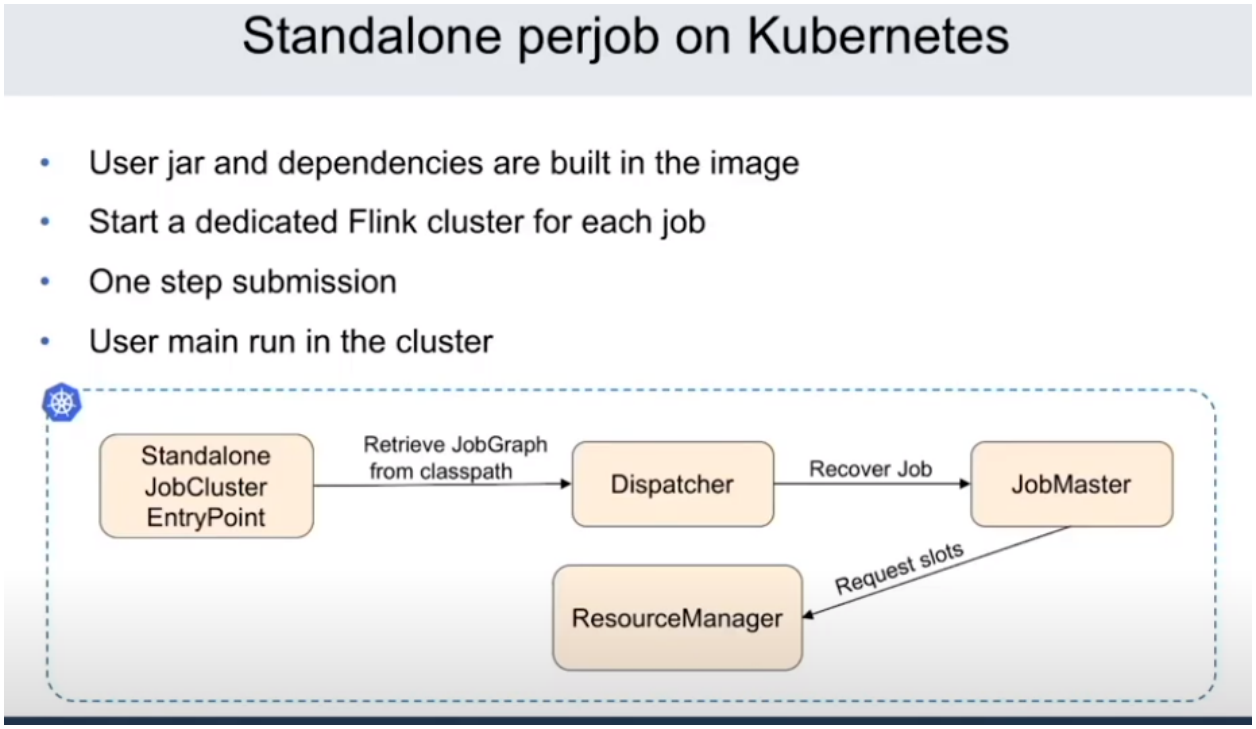
internal ports

external ports

./docker/flink/build.sh --job-jar target/streaming-job-*.jar --from-archive ~/flink-1.14.3-bin-scala_2.11.tgz --image-name streaming.job:latest



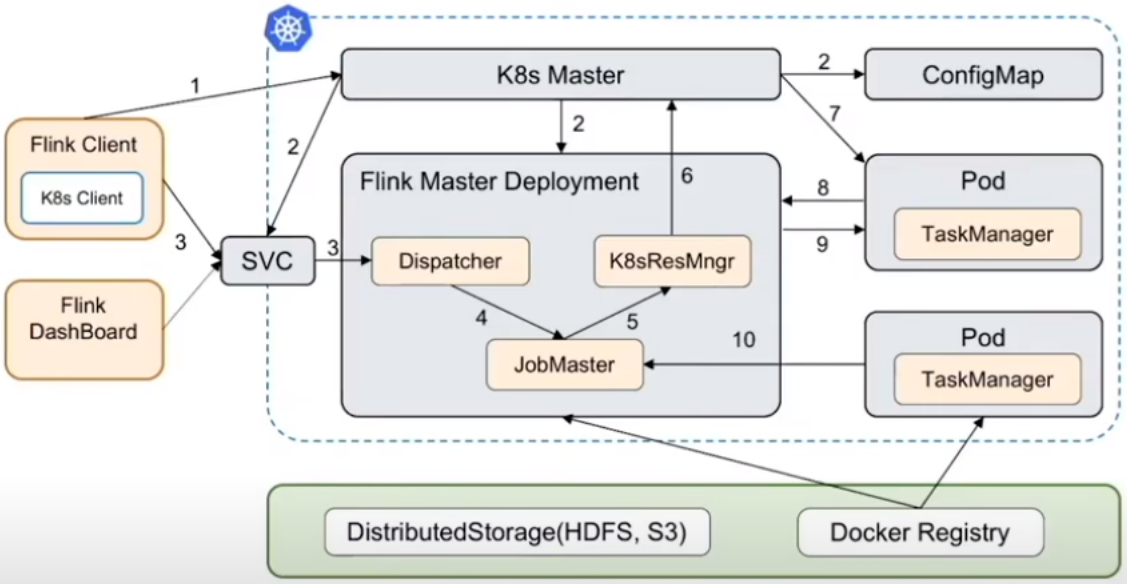
Manually requires to scale task managers



Kuberneties Native means

Self Container
  Embeded K8 client inside
  Do not need external tools to start/stop flink cluster
Flink Client natively contact Kuerneties api server to create JoManager
Flink resource Manager natively contact Kubernetes to create Task Manager Pod on Demand

Native Kubernetes session



Session VS Perjob

- Where the user main code is executed?
  - Session: Client
  - Perjob: Cluster
- How the job graph and user jars are distributed?
  - Session: Upload via rest client and localized by Flink distributed cache
  - Perjob: Built-in the image or downloaded by init container
- Isolation between different jobs
- Cluster lifecycle
  - Session: Manually start stop
  - Perjob: Bound to the only one job

Session

```
wangyang-pc:flink-master danrtsey.wy$ ./bin/kubernetes-session.sh \
> -Dkubernetes.cluster-id=flink-native-k8s-session-1 \
> -Dkubernetes.container.image=registry.cn-beijing.aliyuncs.com/streamcompute/flink:k8s-ff-sf \
> -Djobmanager.heap.size=4096m \
> -Dtaskmanager.memory.process.size=4096m \
> -Dtaskmanager.numberOfTaskSlots=4 \
> -Dkubernetes.jobmanager.cpu=1 -Dkubernetes.taskmanager.cpu=2 \
> -Dresourcemanager.taskmanager-timeout=5000 \
> -Dkubernetes.container-start-command-template="%java% %classpath% %jvmmem% %jvmopts% %logging% %class% %args%"
```

```
wangyang-pc:flink-master danrtsey.wy$ ./bin/flink run -d -p 10 -e kubernetes-session -Dkubernetes.cluster-id=flink-nati
ve-k8s-session-1 examples/streaming/WindowJoin.jar
```

Per job session

```
wangyang-pc:flink-master danrtsey.wy$ ./bin/flink run -d -R -p 10 -e kubernetes-per-job \
> -Dkubernetes.cluster-id=flink-native-k8s-per-job-1 \
> -Dkubernetes.container.image=registry.cn-beijing.aliyuncs.com/streamcompute/flink:k8s-ff-sf \
> -Djobmanager.heap.size=4096m -Dtaskmanager.memory.process.size=4096m \
> -Dkubernetes.jobmanager.cpu=1 -Dkubernetes.taskmanager.cpu=2 \
> -Dtaskmanager.numberOfTaskSlots=4 \
> -Dkubernetes.container-start-command-template="%java% %classpath% %jvmmem% %jvmopts% %logging% %class% %args%" \
> file:///opt/flink/examples/streaming/WindowJoin.jar
2020-04-06 18:42:43,160 INFO  org.apache.flink.kubernetes.utils.KubernetesUtils           [] - Kubernetes deployment r
equires a fixed port. Configuration blob.server.port will be set to 6124
2020-04-06 18:42:43,160 INFO  org.apache.flink.kubernetes.utils.KubernetesUtils           [] - Kubernetes deployment r
equires a fixed port. Configuration taskmanager.rpc.port will be set to 6122
2020-04-06 18:42:44,979 INFO  org.apache.flink.kubernetes.KubernetesClusterDescriptor     [] - Create flink per-job cl
uster flink-native-k8s-per-job-1 successfully, JobManager Web Interface: http://11.164.91.5:30641
wangyang-pc:flink-master danrtsey.wy$ ll /opt/flink/examples/streaming/WindowJoin.jar
ls: /opt/flink/examples/streaming/WindowJoin.jar: No such file or directory
wangyang-pc:flink-master danrtsey.wy$
wangyang-pc:flink-master danrtsey.wy$
```