

第

3

节

K8s资源请求和限额(Request/Limit)

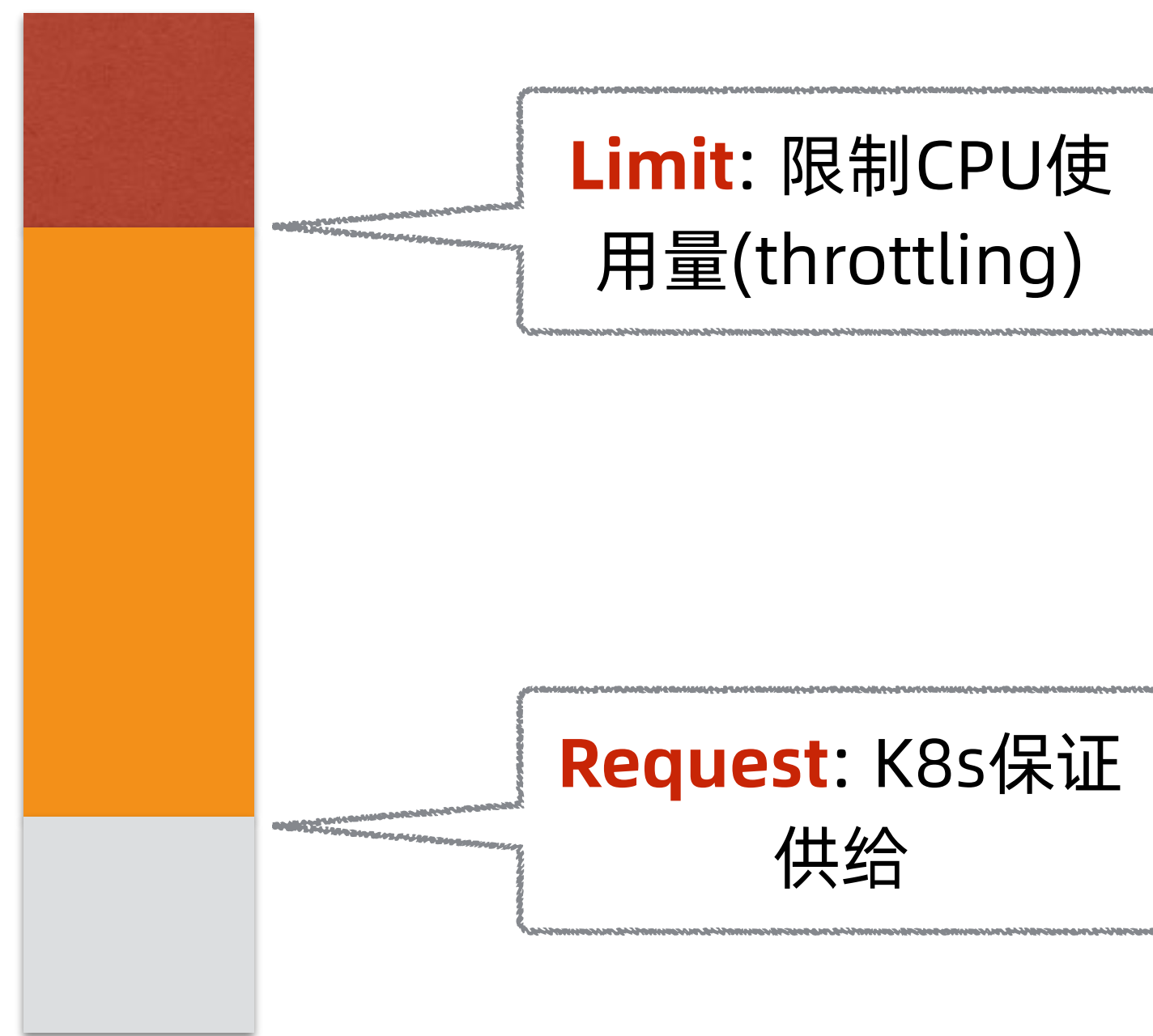
本课内容

- K8s资源请求和限额
 - Request
 - Limit
- Request/Limit演示

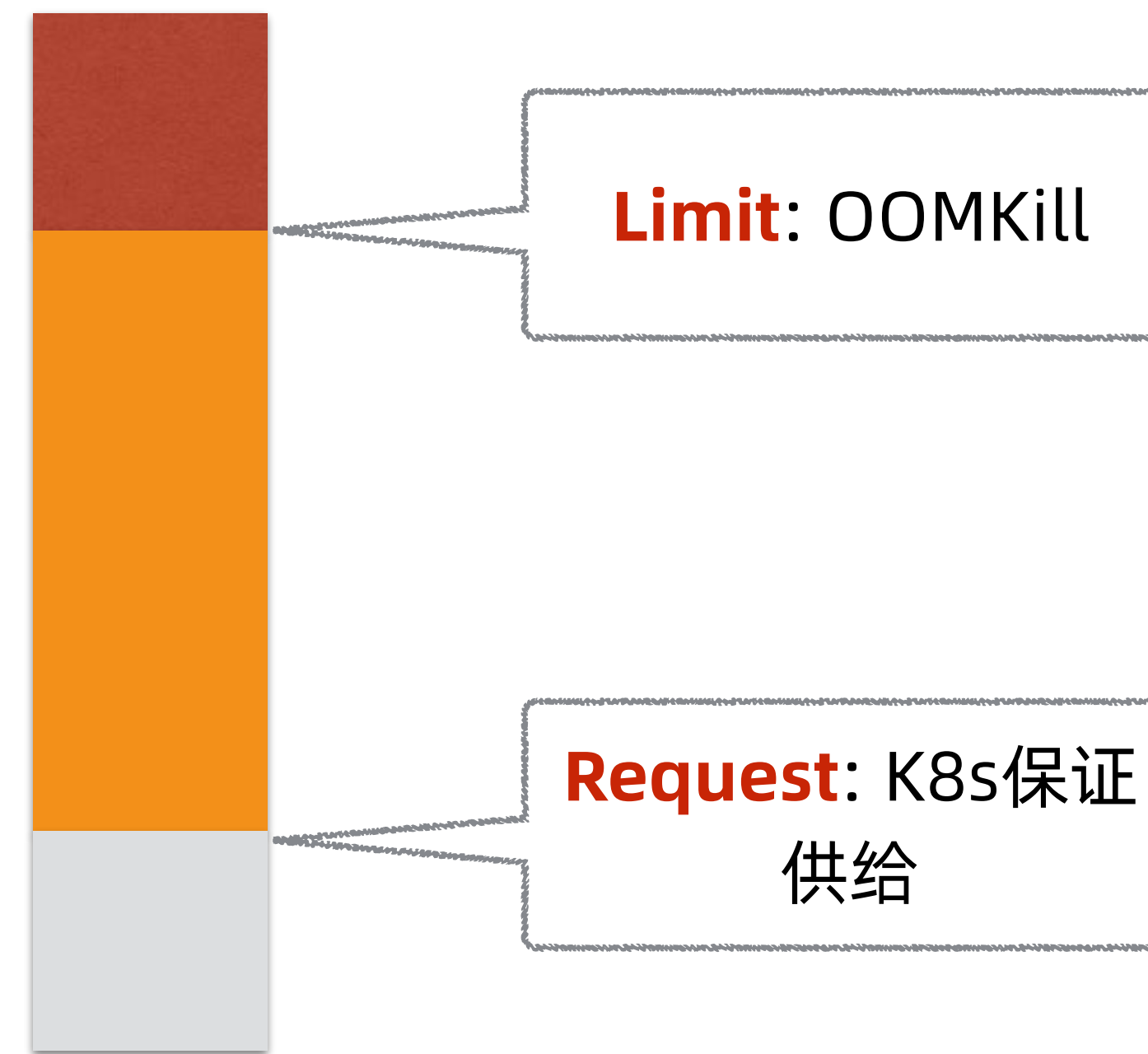


Request/Limit

CPU



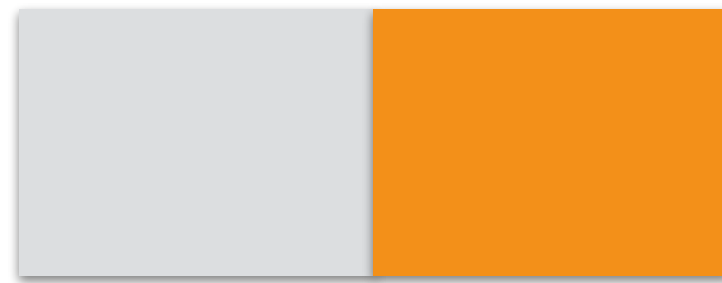
Memory



```
resources:
  requests:
    memory: "128Mi"
    cpu: "200m"
  limits:
    memory: "512Mi"
```

三种设置

Guaranteed



Request = Limit

Burstable



Request < Limit

Best Effort



No Setting

查看节点资源详情



```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
→ local git:(master) kubectl get no
NAME          STATUS    ROLES    AGE   VERSION
docker-desktop Ready    master   114d  v1.14.6
→ local git:(master) kubectl describe no docker-desktop
Name:          docker-desktop
Roles:         master
Labels:        beta.kubernetes.io/arch=amd64
               beta.kubernetes.io/os=linux
               kubernetes.io/arch=amd64
               kubernetes.io/hostname=docker-desktop
               kubernetes.io/os=linux
               node-role.kubernetes.io/master=
Annotations:   kubeadm.alpha.kubernetes.io/cri-socket: /var/run/dockerhim.sock
               node.alpha.kubernetes.io/ttl: 0
               volumes.kubernetes.io/controller-managed-attach-detach: true
CreationTimestamp: Mon, 09 Sep 2019 19:18:46 +0800
```

```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
Addresses:
  InternalIP: 192.168.65.3
  Hostname:   docker-desktop
Capacity:
  cpu:          4
  ephemeral-storage: 61255492Ki
  hugepages-1Gi: 0
  hugepages-2Mi: 0
  memory:       4042204Ki
  pods:         110
Allocatable:
  cpu:          4
  ephemeral-storage: 56453061334
  hugepages-1Gi: 0
  hugepages-2Mi: 0
  memory:       3939804Ki
  pods:         110
System Info:
  Machine ID:      8cc60ad8-9f9d-4ef3-9895-090c13014c81
  System UUID:     140E4ED9-0000-0000-AD81-8B5ABEFC50B0
  Boot ID:         4bf122cc-94ad-4242-aa6e-4f8dfcee432b
  Kernel Version:  4.9.184-linuxkit
  OS Image:         Docker Desktop
  Operating System: linux
  Architecture:    amd64
  Container Runtime Version: docker://19.3.2
  Kubelet Version:  v1.14.6
  Kube-Proxy Version: v1.14.6
```

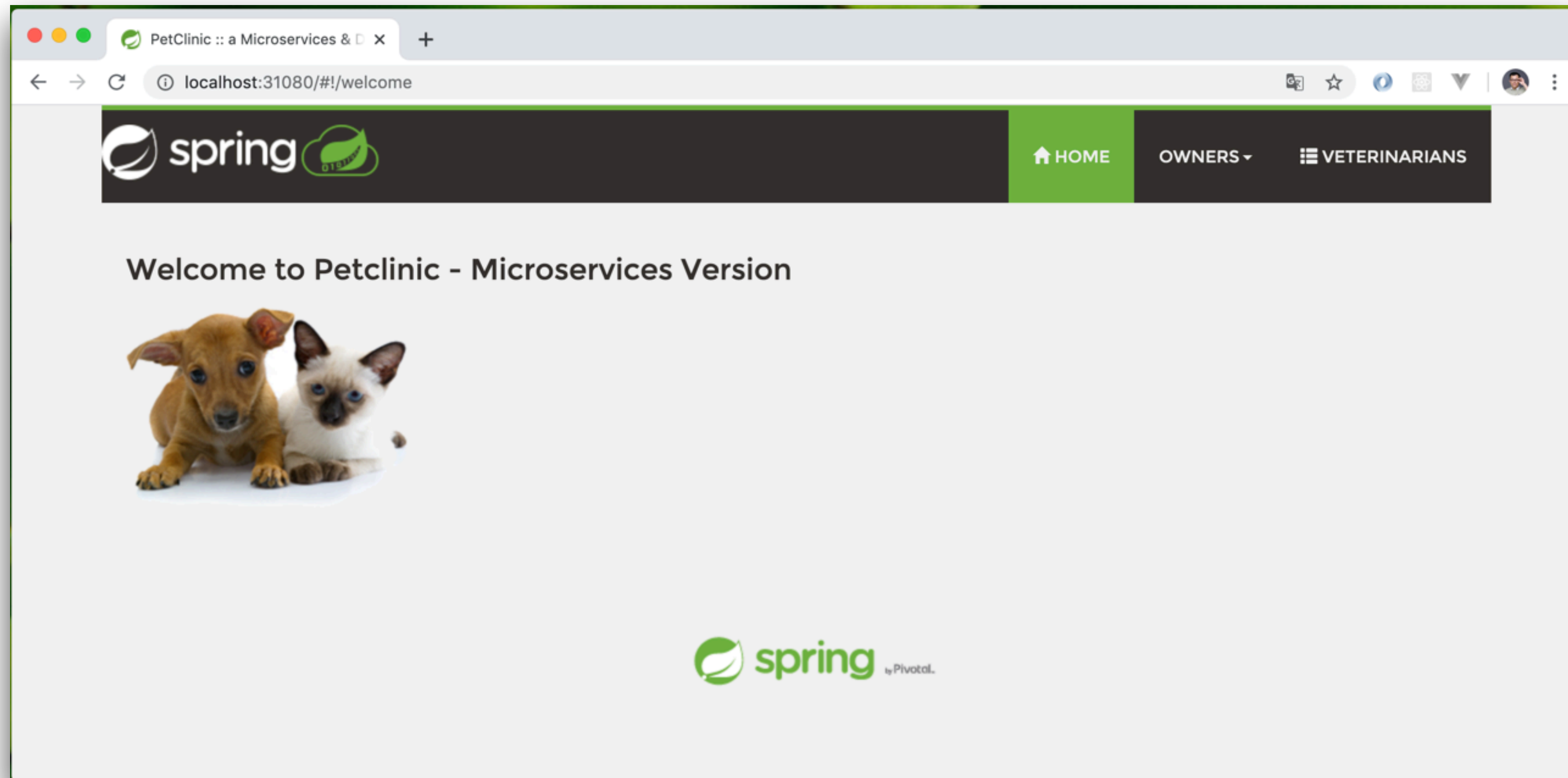
customers-svc.yml添加资源申请和限额

```
1  apiVersion: apps/v1
2  kind: Deployment
3  metadata:
4    name: customers
5  spec:
6    selector:
7      matchLabels:
8        app: customers
9    replicas: 1
10   template:
11     metadata:
12       labels:
13         app: customers
14     spec:
15       containers:
16         - name: customers
17           image: spring2go/spring-petclinic-customers-service:1.0.0.RELEASE
18           resources:
19             requests:
20               memory: "128Mi"
21               cpu: "200m"
22             limits:
23               memory: "512Mi"
24           env:
25             - name: SERVER_PORT
26               value: "8080"
```


发布Petclinic微服务版

```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
→ local git:(master) x kubectl apply -f .
deployment.apps/customers created
service/customers created
deployment.apps/gateway created
service/gateway created
deployment.apps/vets created
service/vets created
deployment.apps/visits created
service/visits created
deployment.apps/web created
service/web created
→ local git:(master) x kubectl get po
NAME                                READY   STATUS    RESTARTS   AGE
customers-858cb7bbd-fjx7q           1/1     Running   0           15s
gateway-558d7f976c-kbmft            1/1     Running   0           15s
vets-8bf75546c-mf566                1/1     Running   0           15s
visits-57f4598df9-pwclr             1/1     Running   0           15s
web-6d46d5c849-8j76g                1/1     Running   0           15s
→ local git:(master) x
```

浏览器校验成功



查看节点详情

```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)

Non-terminated Pods:      (17 in total)
  Namespace              Name                                CPU Requests  CPU Limits  Memory Requests  Memory Limits  AGE
  -----              -
  default               customers-858cb7bbd-fjx7q          200m (5%)    0 (0%)     128Mi (3%)     512Mi (13%)   90s
  default               gateway-558d7f976c-kbmft          0 (0%)       0 (0%)     0 (0%)         0 (0%)        90s
  default               vets-8bf75546c-mf566              0 (0%)       0 (0%)     0 (0%)         0 (0%)        90s
  default               visits-57f4598df9-pwclr           0 (0%)       0 (0%)     0 (0%)         0 (0%)        90s
  default               web-6d46d5c849-8j76g              0 (0%)       0 (0%)     0 (0%)         0 (0%)        90s
  docker                compose-6c67d745f6-6brc7           0 (0%)       0 (0%)     0 (0%)         0 (0%)        114d
  docker                compose-api-57ff65b8c7-777m4       0 (0%)       0 (0%)     0 (0%)         0 (0%)        114d
  kube-system           coredns-584795fc57-hh1r6          100m (2%)    0 (0%)     70Mi (1%)      170Mi (4%)    114d
  kube-system           coredns-584795fc57-nv8wq          100m (2%)    0 (0%)     70Mi (1%)      170Mi (4%)    114d
  kube-system           etcd-docker-desktop               0 (0%)       0 (0%)     0 (0%)         0 (0%)        114d
  kube-system           kube-apiserver-docker-desktop      250m (6%)    0 (0%)     0 (0%)         0 (0%)        114d
  kube-system           kube-controller-manager-docker-desktop 200m (5%)    0 (0%)     0 (0%)         0 (0%)        114d
  kube-system           kube-proxy-vv5sp                  0 (0%)       0 (0%)     0 (0%)         0 (0%)        114d
  kube-system           kube-scheduler-docker-desktop      100m (2%)    0 (0%)     0 (0%)         0 (0%)        114d
  kube-system           metrics-server-564fbf75b5-xcts7    0 (0%)       0 (0%)     0 (0%)         0 (0%)        7h49m
  kubernetes-dashboard  dashboard-metrics-scraper-69fcc6d9df-vtsm6 0 (0%)       0 (0%)     0 (0%)         0 (0%)        10d
  kubernetes-dashboard  kubernetes-dashboard-6d75768647-vbdfv    0 (0%)       0 (0%)     0 (0%)         0 (0%)        10d

Allocated resources:
  (Total limits may be over 100 percent, i.e., overcommitted.)
  Resource           Requests    Limits
  -----
  cpu                 950m (23%) 0 (0%)
  memory              268Mi (6%) 852Mi (22%)
  ephemeral-storage   0 (0%)     0 (0%)

Events: <none>
→ local git:(master) x
```

设置内存超限

```
1  apiVersion: apps/v1
2  kind: Deployment
3  metadata:
4    name: customers
5  spec:
6    selector:
7      matchLabels:
8        app: customers
9    replicas: 5
10   template:
11     metadata:
12       labels:
13         app: customers
14     spec:
15       containers:
16         - name: customers
17           image: spring2go/spring-petclinic-customers-service:1.0.0.RELEASE
18           resources:
19             requests:
20               memory: "1024Mi"
21               cpu: "200m"
22             limits:
23               memory: "1024Mi"
24       env:
25         - name: SERVER_PORT
26           value: "8080"
```


再次发布Petclinic微服务版

```
1. william@jskil: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
ephemeral-storage 0 (0%)    0 (0%)
Events:             <none>
→ local git:(master) x kubectl apply -f .
deployment.apps/customers configured
service/customers unchanged
deployment.apps/gateway unchanged
service/gateway unchanged
deployment.apps/vets unchanged
service/vets unchanged
deployment.apps/visits unchanged
service/visits unchanged
deployment.apps/web unchanged
service/web unchanged
→ local git:(master) x kubectl get po
NAME                                READY    STATUS    RESTARTS   AGE
customers-5f4fd566db-b2pcl          0/1     Pending   0           7s
customers-5f4fd566db-c8tgk          1/1     Running   0           11s
customers-5f4fd566db-jxx8r          1/1     Running   0           11s
customers-5f4fd566db-pvfvn          0/1     Pending   0           6s
customers-5f4fd566db-z2pg2          1/1     Running   0           11s
customers-858cb7bbd-72wln           1/1     Running   0           11s
customers-858cb7bbd-fjx7q           1/1     Running   0           11m
customers-858cb7bbd-gx75d           0/1     Terminating 0           11s
customers-858cb7bbd-nwghz           0/1     Terminating 0           11s
gateway-558d7f976c-kbmft           1/1     Running   0           11m
vets-8bf75546c-mf566               1/1     Running   0           11m
visits-57f4598df9-pwclr            1/1     Running   0           11m
web-6d46d5c849-8j76g               1/1     Running   0           11m
→ local git:(master) x
```

```
1. william@jskil: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
service/web unchanged
→ local git:(master) x kubectl get po
NAME                                READY    STATUS    RESTARTS   AGE
customers-5f4fd566db-b2pcl          0/1     Pending   0           7s
customers-5f4fd566db-c8tgk          1/1     Running   0           11s
customers-5f4fd566db-jxx8r          1/1     Running   0           11s
customers-5f4fd566db-pvfvn          0/1     Pending   0           6s
customers-5f4fd566db-z2pg2          1/1     Running   0           11s
customers-858cb7bbd-72wln           1/1     Running   0           11s
customers-858cb7bbd-fjx7q           1/1     Running   0           11m
customers-858cb7bbd-gx75d           0/1     Terminating 0           11s
customers-858cb7bbd-nwghz           0/1     Terminating 0           11s
gateway-558d7f976c-kbmft           1/1     Running   0           11m
vets-8bf75546c-mf566               1/1     Running   0           11m
visits-57f4598df9-pwclr            1/1     Running   0           11m
web-6d46d5c849-8j76g               1/1     Running   0           11m
→ local git:(master) x kubectl get po
NAME                                READY    STATUS    RESTARTS   AGE
customers-5f4fd566db-b2pcl          0/1     Pending   0           64s
customers-5f4fd566db-c8tgk          1/1     Running   0           68s
customers-5f4fd566db-jxx8r          1/1     Running   0           68s
customers-5f4fd566db-pvfvn          0/1     Pending   0           63s
customers-5f4fd566db-z2pg2          1/1     Running   0           68s
customers-858cb7bbd-fjx7q           1/1     Running   0           12m
gateway-558d7f976c-kbmft           1/1     Running   0           12m
vets-8bf75546c-mf566               1/1     Running   0           12m
visits-57f4598df9-pwclr            1/1     Running   0           12m
web-6d46d5c849-8j76g               1/1     Running   0           12m
→ local git:(master) x
```

限制内存过小

```
1  apiVersion: apps/v1
2  kind: Deployment
3  metadata:
4    name: customers
5  spec:
6    selector:
7      matchLabels:
8        app: customers
9    replicas: 1
10   template:
11     metadata:
12       labels:
13         app: customers
14     spec:
15       containers:
16         - name: customers
17           image: spring2go/spring-petclinic-customers-service:1.0.0.RELEASE
18           resources:
19             requests:
20               memory: "20Mi"
21               cpu: "200m"
22             limits:
23               memory: "20Mi"
24           env:
25             - name: SERVER_PORT
26               value: "8080"
```


再次发布Petclinic微服务版

```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
→ local git:(master) x kubectl apply -f .
deployment.apps/customers configured
service/customers unchanged
deployment.apps/gateway unchanged
service/gateway unchanged
deployment.apps/vets unchanged
service/vets unchanged
deployment.apps/visits unchanged
service/visits unchanged
deployment.apps/web unchanged
service/web unchanged
→ local git:(master) x kubectl get po
NAME                                READY   STATUS             RESTARTS   AGE
customers-5f4fd566db-c8tgk          0/1     Terminating       0           4m42s
customers-5f4fd566db-z2pg2          0/1     Terminating       0           4m42s
customers-7dbcb7cc8-t8xl7           0/1     OOMKilled           1           5s
customers-858cb7bbd-fjx7q           0/1     Terminating       0           16m
gateway-558d7f976c-kbmft            1/1     Running             0           16m
vets-8bf75546c-mf566                1/1     Running             0           16m
visits-57f4598df9-pwclr             1/1     Running             0           16m
web-6d46d5c849-8j76g               1/1     Running             0           16m
→ local git:(master) x
```

```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
customers-5f4fd566db-z2pg2          0/1     Terminating       0           4m42s
customers-7dbcb7cc8-t8xl7           0/1     OOMKilled           1           5s
customers-858cb7bbd-fjx7q           0/1     Terminating       0           16m
gateway-558d7f976c-kbmft            1/1     Running             0           16m
vets-8bf75546c-mf566                1/1     Running             0           16m
visits-57f4598df9-pwclr             1/1     Running             0           16m
web-6d46d5c849-8j76g               1/1     Running             0           16m
→ local git:(master) x kubectl get po
NAME                                READY   STATUS             RESTARTS   AGE
customers-7dbcb7cc8-t8xl7           0/1     OOMKilled           2           31s
gateway-558d7f976c-kbmft            1/1     Running             0           16m
vets-8bf75546c-mf566                1/1     Running             0           16m
visits-57f4598df9-pwclr             1/1     Running             0           16m
web-6d46d5c849-8j76g               1/1     Running             0           16m
→ local git:(master) x kubectl get po
NAME                                READY   STATUS             RESTARTS   AGE
customers-7dbcb7cc8-t8xl7           0/1     CrashLoopBackOff    2           37s
gateway-558d7f976c-kbmft            1/1     Running             0           16m
vets-8bf75546c-mf566                1/1     Running             0           16m
visits-57f4598df9-pwclr             1/1     Running             0           16m
web-6d46d5c849-8j76g               1/1     Running             0           16m
→ local git:(master) x
```

环境清理

```
1. william@jskill: ~/mydev/github/spring2go/spring-petclinic-msa/k8s/local (zsh)
→ local git:(master) x kubectl delete -f .
deployment.apps "customers" deleted
service "customers" deleted
deployment.apps "gateway" deleted
service "gateway" deleted
deployment.apps "vets" deleted
service "vets" deleted
deployment.apps "visits" deleted
service "visits" deleted
deployment.apps "web" deleted
service "web" deleted
→ local git:(master) x kubectl get all
NAME                                TYPE                CLUSTER-IP      EXTERNAL-IP      PORT(S)
AGE
service/kubernetes                  ClusterIP           10.96.0.1       <none>           443/TCP
2d6h
→ local git:(master) x
```


本课小结



- 资源申请Request, **K8s调度决策**, 如有资源则保证供给
- 资源限额Limit
 - **内存超额 > OOMKill**
 - CPU超额 > 限制使用量(throttling)
- Request <= Limit
- 管理员还可配置名字空间级别的配额(ResourceQuota/
LimitRange)