

Screenshot of `kubectl get nodes`

```
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl get nodes
NAME                STATUS    ROLES    AGE   VERSION
kind-control-plane  Ready    control-plane  26d   v1.29.2
kind-worker          Ready    <none>      26d   v1.29.2
kind-worker2         Ready    <none>      26d   v1.29.2
```

Screenshot of passing unit tests

```
go: downloading k8s.io/cloud-provider v0.27.1
go: downloading k8s.io/dynamic-resource-allocation v0.27.1
go: downloading k8s.io/csi-translation-lib v0.27.1
go: downloading github.com/josharian/intern v1.0.0
go: downloading github.com/go-logr/stdr v1.2.2
go: downloading github.com/asaskevich/govalidator v0.0.0-20190424111038-f61b66f89f4a
go: downloading github.com/mitchellh/mapstructure v1.4.1
go: downloading go.etcd.io/etcd/api/v3 v3.5.7
go: downloading github.com/cenkalti/backoff/v4 v4.1.3
go: downloading go.uber.org/atomic v1.7.0
go: downloading go.uber.org/multierr v1.6.0
go: downloading github.com/antlr/antlr4/runtime/Go/antlr v1.4.10
go: downloading github.com/grpc-ecosystem/grpc-gateway/v2 v2.7.0
go: downloading github.com/docker/distribution v2.8.1+incompatible
go: downloading k8s.io/controller-manager v0.27.1
go: downloading github.com/opencontainers/selinux v1.10.0
go: downloading k8s.io/mount-utils v0.27.1
go: downloading github.com/coreos/go-semver v0.3.0
go: downloading github.com/opencontainers/go-digest v1.0.0
go: downloading github.com/moby/sys/mountinfo v0.6.2
root@c2e0a8c109a0:/go/src/app# go test -v ./...
?      my-scheduler-plugins/cmd/scheduler      [no test files]
=== RUN   TestCustomScheduler_PreFilter
=== RUN   TestCustomScheduler_PreFilter/pod_is_accepted
finish adding2024/05/28 08:06:35 Pod  is in Prefilter phase.
=== RUN   TestCustomScheduler_PreFilter/pod_is_just_accepted
finish adding2024/05/28 08:06:35 Pod  is in Prefilter phase.
=== RUN   TestCustomScheduler_PreFilter/pod_is_rejected
finish adding2024/05/28 08:06:35 Pod  is in Prefilter phase.
--- PASS: TestCustomScheduler_PreFilter (0.01s)
    --- PASS: TestCustomScheduler_PreFilter/pod_is_accepted (0.01s)
    --- PASS: TestCustomScheduler_PreFilter/pod_is_just_accepted (0.00s)
    --- PASS: TestCustomScheduler_PreFilter/pod_is_rejected (0.00s)
=== RUN   TestCustomScheduler_Score
=== RUN   TestCustomScheduler_Score/least_mode
2024/05/28 08:06:35 Custom scheduler runs with the mode: Least.
2024/05/28 08:06:35 Pod  is in Score phase. Calculate the score of Node m1.
2024/05/28 08:06:35 Pod  is in Score phase. Calculate the score of Node m2.
=== RUN   TestCustomScheduler_Score/most_mode
2024/05/28 08:06:35 Custom scheduler runs with the mode: Least.
2024/05/28 08:06:35 Pod  is in Score phase. Calculate the score of Node m1.
2024/05/28 08:06:35 Pod  is in Score phase. Calculate the score of Node m2.
--- PASS: TestCustomScheduler_Score (0.00s)
    --- PASS: TestCustomScheduler_Score/least_mode (0.00s)
    --- PASS: TestCustomScheduler_Score/most_mode (0.00s)
=== RUN   TestCustomScheduler_NormalizeScore
=== RUN   TestCustomScheduler_NormalizeScore/scores_in_range
=== RUN   TestCustomScheduler_NormalizeScore/scores_out_of_range
=== RUN   TestCustomScheduler_NormalizeScore/negative_score
--- PASS: TestCustomScheduler_NormalizeScore (0.00s)
    --- PASS: TestCustomScheduler_NormalizeScore/scores_in_range (0.00s)
    --- PASS: TestCustomScheduler_NormalizeScore/scores_out_of_range (0.00s)
    --- PASS: TestCustomScheduler_NormalizeScore/negative_score (0.00s)
PASS
ok      my-scheduler-plugins/pkg/plugins      0.162s
root@c2e0a8c109a0:/go/src/app#
```

Node resources

```
Addresses:
  InternalIP: 172.19.0.2
  Hostname: kind-worker
Capacity:
  cpu: 6
  ephemeral-storage: 50620216Ki
  hugepages-2Mi: 0
  memory: 6063048Ki
  pods: 110
Allocatable:
  cpu: 6
  ephemeral-storage: 50620216Ki
  hugepages-2Mi: 0
  memory: 1344456Ki
  pods: 110
System Info:
```

```
Addresses:
  InternalIP: 172.19.0.4
  Hostname: kind-worker2
Capacity:
  cpu: 6
  ephemeral-storage: 50620216Ki
  hugepages-2Mi: 0
  memory: 6063048Ki
  pods: 110
Allocatable:
  cpu: 6
  ephemeral-storage: 50620216Ki
  hugepages-2Mi: 0
  memory: 1868744Ki
  pods: 110
System Info:
```

6 scenarios

1. pod-spec1.yaml with Least mode

```
1  apiVersion: v1
2  kind: Pod
3  metadata:
4    name: nginx
5    labels:
6      podGroup: "A"
7      minAvailable: "2"
8  spec:
9    schedulerName: my-scheduler
10   containers:
11   - name: nginx
12     image: nginx:1.14.2
13     ports:
14     - containerPort: 80
15     resources:
16       requests:
17         memory: "100Mi"
18       limits:
19         memory: "100Mi"
```

```
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl get po -o wide
NAME                                READY   STATUS    RESTARTS   AGE   IP            NODE             NOMINATED NODE   READINESS GATES
mnist-inference-599dc8b56-ftlbq     1/1     Running   1 (18d ago)  25d   10.244.1.2    kind-worker      <none>           <none>
mnist-inference-599dc8b56-sfqx5     1/1     Running   1 (18d ago)  25d   10.244.2.3    kind-worker2     <none>           <none>
mnist-train                          0/1     Completed 0           25d   <none>        kind-worker2     <none>           <none>
my-scheduler-69cfc986c7-6692c       1/1     Running   0           8m29s  10.244.2.5    kind-worker2     <none>           <none>
nginx                                0/1     Pending   0           3m33s  <none>        <none>           <none>           <none>
```

```
I0528 12:11:44.953921    1 log.go:194] Pod nginx is in Prefilter phase.
I0528 12:11:44.968534    1 schedule_one.go:867] "Unable to schedule pod; no fit; waiting" pod="default/nginx" err="0/3 nodes are available: preemption: 0/3 nodes are available: 3 No preemption victims found for incoming pod..."
```

★因為minAvailable改成2，所以只創一個pod的情況下會過不了prefilter

2. pod-spec1-2.yaml with **Least** mode

```
1  apiVersion: v1
2  kind: Pod
3  metadata:
4    name: nginx
5    labels:
6      podGroup: "A"
7    minAvailable: "2"
8  spec:
9    schedulerName: my-scheduler
10   containers:
11   - name: nginx
12     image: nginx:1.14.2
13     ports:
14     - containerPort: 80
15     resources:
16       requests:
17         memory: "100Mi"
18       limits:
19         memory: "100Mi"
20 ---
21  apiVersion: v1
22  kind: Pod
23  metadata:
24    name: nginx2
25    labels:
26      podGroup: "A"
27    minAvailable: "2"
28  spec:
29    schedulerName: my-scheduler
30   containers:
31   - name: nginx2
32     image: nginx:1.14.2
33     ports:
34     - containerPort: 80
35     resources:
36       requests:
37         memory: "100Mi"
38       limits:
39         memory: "100Mi"
40
```

```
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl delete po nginx
pod "nginx" deleted
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl create -f pod-spec1-2.yaml
pod/nginx created
pod/nginx2 created
```

```
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl get po -o wide
NAME                                READY   STATUS    RESTARTS   AGE   IP            NODE             NOMINATED NODE   READINESS GATES
mnist-inference-599dc8b56-ftlbq     1/1     Running   1 (18d ago)  25d   10.244.1.2    kind-worker      <none>            <none>
mnist-inference-599dc8b56-sfqx5     1/1     Running   1 (18d ago)  25d   10.244.2.3    kind-worker2     <none>            <none>
mnist-train                          0/1     Completed   0           25d   <none>        kind-worker2     <none>            <none>
my-scheduler-69cfc986c7-6692c       1/1     Running   0           34m   10.244.2.5    kind-worker2     <none>            <none>
nginx                                1/1     Running   0           19m   10.244.1.5    kind-worker      <none>            <none>
nginx2                              1/1     Running   0           19m   10.244.1.4    kind-worker      <none>            <none>
```

★因為kind-worker的memory比kind-worker2的memory小，所以在Least mode底下會優先選擇kind-worker

3. pod-spec1-2.yaml with **Most** mode

```
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl get po -o wide
NAME                                READY   STATUS    RESTARTS   AGE   IP            NODE             NOMINATED NODE   READINESS GATES
mnist-inference-599dc8b56-ftlbq     1/1     Running   1 (18d ago)  25d   10.244.1.2    kind-worker      <none>           <none>
mnist-inference-599dc8b56-sfqx5     1/1     Running   1 (18d ago)  25d   10.244.2.3    kind-worker2     <none>           <none>
mnist-train                          0/1     Completed 0           26d   <none>        kind-worker2     <none>           <none>
my-scheduler-69cfc986c7-zhx6g       1/1     Running   0           7m8s   10.244.2.6    kind-worker2     <none>           <none>
nginx                                1/1     Running   0           5m34s  10.244.2.8    kind-worker2     <none>           <none>
nginx2                               1/1     Running   0           5m34s  10.244.2.7    kind-worker2     <none>           <none>
```

```
ms: kubelet for scheduling pod:
I0528 11:56:14.674825    1 log.go:194] Pod nginx2 is in Prefilter phase.
I0528 11:56:14.675817    1 log.go:194] Pod nginx2 is in Score phase. Calculate the score of Node kind-worker.
I0528 11:56:14.675840    1 log.go:194] Pod nginx2 is in Score phase. Calculate the score of Node kind-worker2.
I0528 11:56:14.724823    1 schedule_one.go:252] "Successfully bound pod to node" pod="default/nginx2" node="kind-worker2" evaluatedNodes=3 feasibleNodes=2
I0528 12:01:16.987000    1 log.go:194] Pod nginx is in Prefilter phase.
I0528 12:01:16.988044    1 log.go:194] Pod nginx is in Score phase. Calculate the score of Node kind-worker2.
I0528 12:01:16.988089    1 log.go:194] Pod nginx is in Score phase. Calculate the score of Node kind-worker.
I0528 12:01:17.005471    1 schedule_one.go:252] "Successfully bound pod to node" pod="default/nginx" node="kind-worker2" evaluatedNodes=3 feasibleNodes=2
```

★因為kind-worker2的memory比kind-worker的memory大，所以在Most mode底下會優先選擇kind-worker2

4. pod-spec1-2.yaml with **Least** mode and one additional replication for scheduler

```
I0528 12:26:11.306663    1 log.go:194] Pod nginx2 is in Prefilter phase.
I0528 12:26:11.307994    1 log.go:194] Pod nginx2 is in Score phase. Calculate the score of Node kind-worker2.
I0528 12:26:11.308047    1 log.go:194] Pod nginx2 is in Score phase. Calculate the score of Node kind-worker.
I0528 12:26:11.327029    1 schedule_one.go:252] "Successfully bound pod to node" pod="default/nginx2" node="kind-worker" evaluatedNodes=3 feasibleNodes=2
```

其中一個scheduler先把nginx2指定給kind-worker

```
I0528 12:26:11.308244    1 log.go:194] Pod nginx2 is in Prefilter phase.
I0528 12:26:11.310667    1 log.go:194] Pod nginx2 is in Score phase. Calculate the score of Node kind-worker2.
I0528 12:26:11.310958    1 log.go:194] Pod nginx2 is in Score phase. Calculate the score of Node kind-worker.
E0528 12:26:11.354469    1 framework.go:1100] "Failed running Bind plugin" err="Operation cannot be fulfilled on pods/binding \"nginx2\": pod nginx2 is already assigned to node \"kind-worker\" plug
in=\"DefaultBinder\" pod=\"default/nginx2\" node=\"kind-worker\"
I0528 12:26:11.354541    1 schedule_one.go:827] "Failed to bind pod" pod="default/nginx2"
I0528 12:26:11.354571    1 schedule_one.go:283] "Scheduler cache ForgetPod failed" err="pod 22593e4e-b46d-46dd-9659-892b78693d72(default/nginx2) wasn't assumed so cannot be forgotten"
E0528 12:26:11.354617    1 schedule_one.go:883] "Error scheduling pod; retrying" err="running Bind plugin \"DefaultBinder\": Operation cannot be fulfilled on pods/binding \"nginx2\": pod nginx2 is a
lready assigned to node \"kind-worker\" pod=\"default/nginx2\"
I0528 12:26:11.354654    1 schedule_one.go:895] "Pod has been assigned to node. Abort adding it back to queue." pod="default/nginx2" node="kind-worker"
I0528 12:31:38.178019    1 log.go:194] Pod nginx is in Prefilter phase.
I0528 12:31:38.178700    1 log.go:194] Pod nginx is in Score phase. Calculate the score of Node kind-worker.
I0528 12:31:38.178753    1 log.go:194] Pod nginx is in Score phase. Calculate the score of Node kind-worker2.
I0528 12:31:38.225744    1 schedule_one.go:252] "Successfully bound pod to node" pod="default/nginx" node="kind-worker" evaluatedNodes=3 feasibleNodes=2
```

另一個scheduler再把nginx2拿來指定給kind-worker時發現已經指定過了，所以出現error。
之後再把nginx指定給kind-worker。

5. pod-spec1-3.yaml with **Least** mode with request memory and limit memory exceed node allocable memory

```
1  apiVersion: v1
2  kind: Pod
3  metadata:
4    name: nginx
5    labels:
6      podGroup: "A"
7      minAvailable: "2"
8  spec:
9    schedulerName: my-scheduler
10   containers:
11   - name: nginx
12     image: nginx:1.14.2
13     ports:
14     - containerPort: 80
15     resources:
16       requests:
17         memory: "2000Mi"
18       limits:
19         memory: "2000Mi"
20 ---
21  apiVersion: v1
22  kind: Pod
23  metadata:
24    name: nginx2
25    labels:
26      podGroup: "A"
27      minAvailable: "2"
28  spec:
29    schedulerName: my-scheduler
30   containers:
31   - name: nginx2
32     image: nginx:1.14.2
33     ports:
34     - containerPort: 80
35     resources:
36       requests:
37         memory: "2000Mi"
38       limits:
39         memory: "2000Mi"
40
```

```
I0528 12:58:21.387727 1 log.go:194] Pod nginx2 is in Prefilter phase.
I0528 12:58:21.388130 1 schedule_one.go:867] "Unable to schedule pod; no fit; waiting" pod="default/nginx2" err="0/3 nodes are available: 1 node(s) had untolerated taint {node-role.kubernetes.io/control-plane: }, 2 Insufficient memory. preemption: 0/3 nodes are available: 1 Preemption is not helpful for scheduling, 2 No preemption victims found for incoming pod.."
student@ds-student20:~/NTHU-Scheduler-Plugins$ kubectl get po -o wide
```

NAME	READY	STATUS	RESTARTS	AGE	IP	NODE	NOMINATED NODE	READINESS GATES
mnist-inference-599dc8b56-ftlbq	1/1	Running	1 (18d ago)	26d	10.244.1.2	kind-worker	<none>	<none>
mnist-inference-599dc8b56-sfqx5	1/1	Running	1 (18d ago)	26d	10.244.2.3	kind-worker2	<none>	<none>
mnist-train	0/1	Completed	0	26d	<none>	kind-worker2	<none>	<none>
my-scheduler-69cfc986c7-v86pm	1/1	Running	0	97s	10.244.2.11	kind-worker2	<none>	<none>
nginx	0/1	Pending	0	33s	<none>	<none>	<none>	<none>
nginx2	0/1	Pending	0	33s	<none>	<none>	<none>	<none>

★pod的request memory超過node的allocable memory會無法配置

6. pod-spec1-4.yaml with **Least** mode

```
1  apiVersion: v1
2  kind: Pod
3  metadata:
4    name: nginx
5    labels:
6      podGroup: "A"
7      minAvailable: "1"
8  spec:
9    schedulerName: my-scheduler
10   containers:
11   - name: nginx
12     image: nginx:1.14.2
13     ports:
14     - containerPort: 80
15     resources:
16       requests:
17         memory: "100Mi"
18       limits:
19         memory: "100Mi"
```

```
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl create -f pod-spec1-4.yaml
pod/nginx created
student@ds-student20:~/NTHU-Scheduler-Plugin$ kubectl get po
NAME                                READY   STATUS    RESTARTS   AGE
mnist-inference-599dc8b56-ftlbq     1/1     Running   2 (6d1h ago)  39d
mnist-inference-599dc8b56-sfqx5     1/1     Running   2 (6d1h ago)  39d
mnist-train                          0/1     Completed 0             39d
my-scheduler-69cfc986c7-jrppl       1/1     Running   0             39s
nginx                                1/1     Running   0             20s
```

★和pod-example一樣設定minAvailable為1，且request memory小於node的allocable memory，所以過得了prefilter