

Auto Scaling trên với Kubernetes

Agenda

1. Horizontal Pod Autoscaling
2. Vertical Pod Autoscaler
3. Cluster Autoscaler

Autoscale

Autoscale là một tính năng của điện toán đám mây giúp tự động điều chỉnh số lượng tài nguyên máy chủ theo nhu cầu sử dụng. Autoscale có thể được sử dụng cho nhiều loại ứng dụng và dịch vụ, giúp tối ưu hóa chi phí, hiệu suất và tính sẵn sàng.

Dưới đây là một số lợi ích cụ thể của việc sử dụng autoscale:

- **Tối ưu hóa chi phí:** Autoscale giúp bạn chỉ cần trả tiền cho những gì bạn sử dụng. Khi lưu lượng truy cập thấp, bạn có thể giảm số lượng máy chủ để tiết kiệm chi phí. Khi lưu lượng truy cập tăng cao, bạn có thể tăng số lượng máy chủ để đáp ứng nhu cầu.
- **Tăng hiệu suất:** Autoscale giúp đảm bảo rằng bạn luôn có đủ tài nguyên máy chủ để đáp ứng nhu cầu sử dụng. Điều này giúp tránh tình trạng quá tải tài nguyên, dẫn đến chậm trễ hoặc ngừng hoạt động của ứng dụng.
- **Tăng tính sẵn sàng:** Autoscale giúp tăng tính sẵn sàng của ứng dụng bằng cách tự động khởi chạy máy chủ mới khi máy chủ hiện tại gặp sự cố.

Autoscale trong Kubernetes

Autoscale trong Kubernetes là quá trình tự động điều chỉnh số lượng pod hoặc node trong một cụm Kubernetes dựa trên yêu cầu tài nguyên của các ứng dụng đang chạy trên đó.

Horizontal Pod Autoscaling (HPA) tự động mở rộng quy mô số lượng pod

Vertical Pod Autoscaling (VPA) là một bộ điều khiển Kubernetes khác tự động điều chỉnh limit và request tài nguyên của pod dựa trên mức sử dụng tài nguyên thực tế của chúng.

Cluster Autoscaler là một công cụ riêng biệt quản lý số lượng node trong một cụm Kubernetes. Nó giám sát mức sử dụng tài nguyên của cụm và thêm hoặc xóa node khi cần thiết để duy trì mức sẵn có tài nguyên mong muốn.

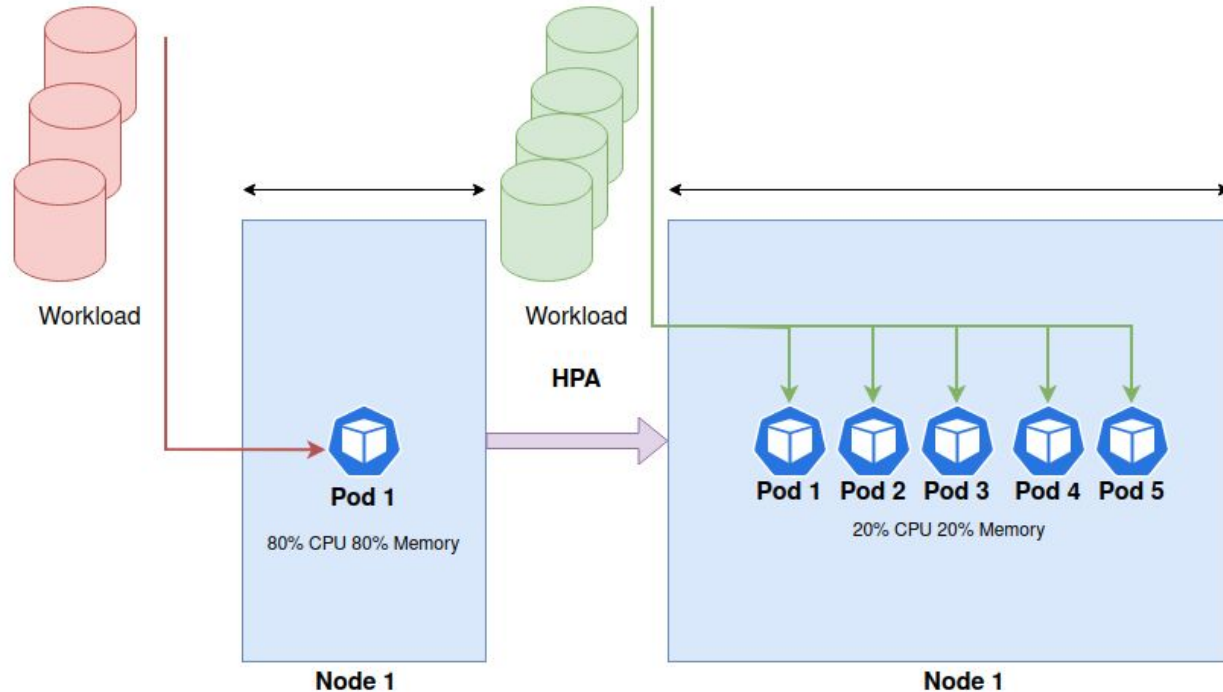
Horizontal Pod Autoscaling

Horizontal pod autoscaling được cài đặt sẵn ở trong cụm K8s có thể được sử dụng để scale up/down pod cho các đối tượng ReplicaSet hoặc Deployment.

Điều kiện tiên quyết

Dù HPA được cài đặt mặc định sẵn trong K8s tuy nhiên để có thể lấy được thông tin cho việc giám sát (CPU, Memory) thì phải cài đặt thêm metrics server tại đường dẫn: <https://github.com/kubernetes-sigs/metrics-server>

Horizontal Pod Autoscaling



Horizontal Pod Autoscaling

Nguyên lý hoạt động:

Mỗi HPA sẽ có một điều kiện riêng để kiểm tra, nếu điều kiện đó đạt thì sẽ tiến hành scale up và khi điều kiện không còn thỏa mãn sẽ scale down.

Số pod cần thiết sẽ được tính bằng công thức:

$$\text{số pod cần thiết} = \text{ceil}[\text{số pod hiện tại nhân} * (\text{giá trị điều kiện hiện tại} / \text{giá trị điều kiện mong đợi.})]$$

ceil: làm tròn

Horizontal Pod Autoscaling

Triển khai ứng dụng và HPA.

Tải manifest sử dụng lệnh:

Truy cập vào thư mục Day2 -> Afternoon -> HPA

Horizontal Pod Autoscaling

Triển khai ứng dụng test sử dụng lệnh: **kubectl apply -f php.yaml**

Kiểm tra sử dụng lệnh **kubectl get deployment php-apache** và **kubectl get svc php-apache**

```
quanlm@quanlm-desktop:~/Work/workshop-bke$ kubectl apply -f Day2/2.Afternoon/1.\ HPA/php.yaml
deployment.apps/php-apache created
service/php-apache created
quanlm@quanlm-desktop:~/Work/workshop-bke$ kubectl get deployment php-apache
NAME          READY   UP-TO-DATE   AVAILABLE   AGE
php-apache    1/1     1            1           51s
quanlm@quanlm-desktop:~/Work/workshop-bke$ kubectl get svc php-apache
NAME          TYPE          CLUSTER-IP    EXTERNAL-IP   PORT(S)    AGE
php-apache    ClusterIP     10.93.134.118 <none>        80/TCP     56s
```

Horizontal Pod Autoscaling

Triển khai HPA sử dụng lệnh: **kubectl apply -f hpa.yaml**

Kiểm tra sử dụng lệnh **kubectl get hpa**

```
quanlm@quanlm-desktop:~/Work/workshop-bke$ kubectl apply -f Day2/2.Afternoon/1.\ HPA/hpa.yaml
horizontalpodautoscaler.autoscaling/http-app created
quanlm@quanlm-desktop:~/Work/workshop-bke$ kubectl get hpa
```

NAME	REFERENCE	TARGETS	MINPODS	MAXPODS	REPLICAS	AGE
http-app	Deployment/php-apache	0%/70%	1	10	1	46s

Horizontal Pod Autoscaling

Tạo workload để kiểm tra cách hoạt động của HPA:

```
kubectrl run -i --tty load-generator --rm --image=busybox:1.28 --restart=Never -- /bin/sh -c "while sleep 0.01; do wget -q -O-http://php-apache; done"
```

Horizontal Pod Autoscaling

Theo dõi HPA sử dụng lệnh: **kubectl get hpa php-apache --watch**

```
quanlm@quanlm-desktop:~$ kubectl get hpa http-app --watch
```

NAME	REFERENCE	TARGETS	MINPODS	MAXPODS	REPLICAS	AGE
http-app	Deployment/php-apache	0%/70%	1	10	1	3m42s
http-app	Deployment/php-apache	176%/70%	1	10	1	4m31s
http-app	Deployment/php-apache	250%/70%	1	10	3	4m46s
http-app	Deployment/php-apache	228%/70%	1	10	4	5m2s
http-app	Deployment/php-apache	157%/70%	1	10	4	5m17s
http-app	Deployment/php-apache	89%/70%	1	10	5	5m32s
http-app	Deployment/php-apache	53%/70%	1	10	5	5m47s
http-app	Deployment/php-apache	60%/70%	1	10	5	6m2s
http-app	Deployment/php-apache	59%/70%	1	10	5	6m17s
http-app	Deployment/php-apache	60%/70%	1	10	5	6m47s
http-app	Deployment/php-apache	60%/70%	1	10	5	7m3s

Horizontal Pod Autoscaling

Kiểm tra số lượng pod sử dụng lệnh **kubectl get pod | grep php**

```
^Cquanlm@quanlm-desktop:~$ kubectl get pod | grep php
php-apache-598b474864-62j8h      1/1      Running    0          2m15s
php-apache-598b474864-96bfg      1/1      Running    0          3m1s
php-apache-598b474864-jl5zf      1/1      Running    0          2m46s
php-apache-598b474864-jxdpn      1/1      Running    0          3m1s
php-apache-598b474864-pdlml      1/1      Running    0          11m
```

Horizontal Pod Autoscaling

Ngắt workload sử dụng ctrl + C để exit

[illegible]

Horizontal Pod Autoscaling

Kiểm tra HPA:

```
quanlm@quanlm-desktop:~$ kubectl get hpa
```

NAME	REFERENCE	TARGETS	MINPODS	MAXPODS	REPLICAS	AGE
http-app	Deployment/php-apache	0%/70%	1	10	1	67m

Horizontal Pod Autoscaling

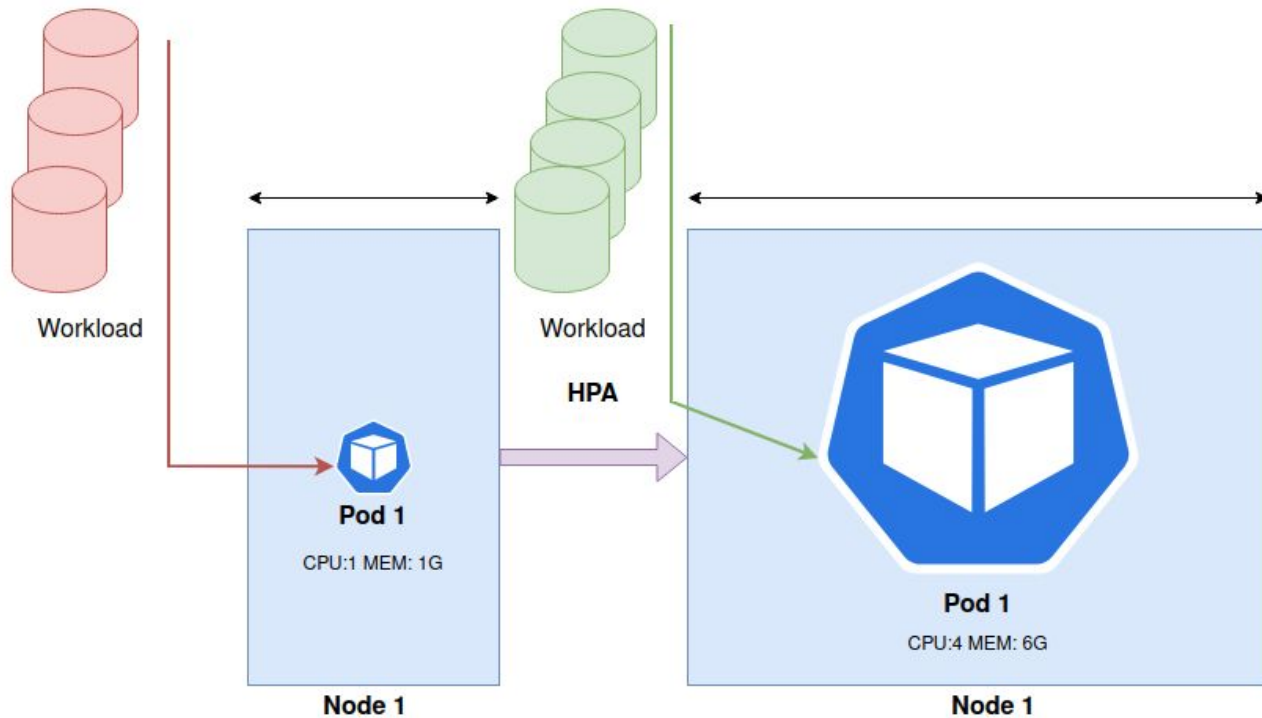
Kiểm tra số lượng pod

```
quanlm@quanlm-desktop:~$ kubectl get pod | grep php  
php-apache-598b474864-96bfg    1/1    Running    0    63m
```


Vertical Pod Autoscaler

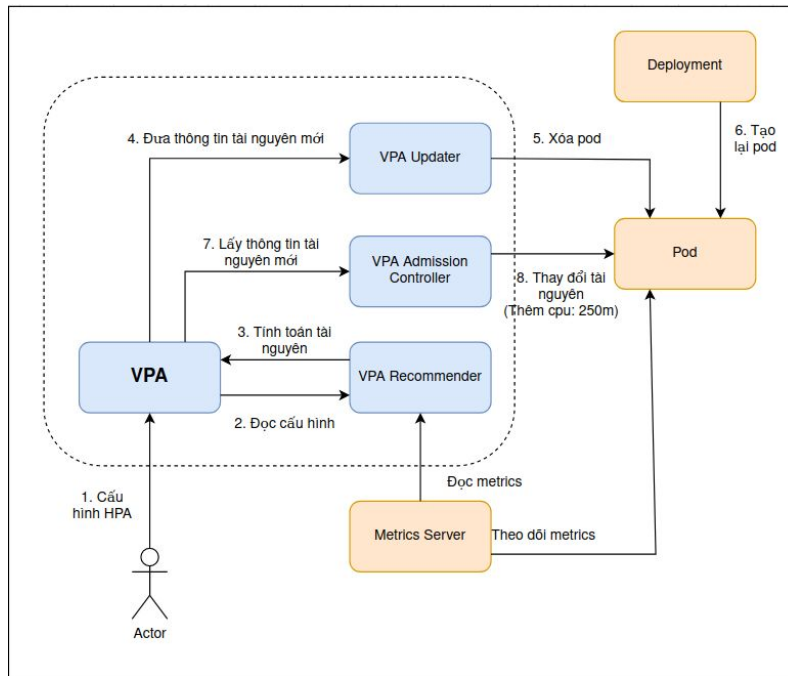
Vertical pod autoscaler, là một công cụ trong Kubernetes giúp tự động điều chỉnh tài nguyên CPU và bộ nhớ cho pod để phù hợp với nhu cầu của ứng dụng.

Vertical Pod Autoscaler



Vertical Pod Autoscaler

Nguyên lý hoạt động



Vertical Pod Autoscaler

1 số lưu ý:

- Không sử dụng VPA và HPA đồng thời cùng metrics trên một tài nguyên
- VPA có thể sẽ khuyến nghị sử dụng nhiều tài nguyên hơn giới hạn của node

Vertical Pod Autoscaler

Triển khai VPA

Tải về sử dụng lệnh

git clone <https://github.com/kubernetes/autoscaler.git>

truy cập thư mục

cd autoscaler/vertical-pod-autoscaler

Vertical Pod Autoscaler

Triển khai VPA

cài đặt sử dụng lệnh **./hack/vpa-up.sh**

```
quanl@quanlm-desktop: ~/Work/workshop-bke/Day2/2. Afternoon/2. VPA/autoscaler/vertical-pod-autoscaler$ ./hack/vpa-up.sh
customresourcedefinition.apixtensions.k8s.io/verticalpodautoscalercheckpoints.autoscaling.k8s.io created
customresourcedefinition.apixtensions.k8s.io/verticalpodautoscalers.autoscaling.k8s.io created
clusterrole.rbac.authorization.k8s.io/system:metrics-reader created
clusterrole.rbac.authorization.k8s.io/system:vpa-actor created
clusterrole.rbac.authorization.k8s.io/system:vpa-status-actor created
clusterrole.rbac.authorization.k8s.io/system:vpa-checkpoint-actor created
clusterrole.rbac.authorization.k8s.io/system:evictioner created
clusterrolebinding.rbac.authorization.k8s.io/system:metrics-reader created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-actor created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-status-actor created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-checkpoint-actor created
clusterrole.rbac.authorization.k8s.io/system:vpa-target-reader created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-target-reader-binding created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-evictioner-binding created
serviceaccount/vpa-admission-controller created
serviceaccount/vpa-recommender created
serviceaccount/vpa-updater created
clusterrole.rbac.authorization.k8s.io/system:vpa-admission-controller created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-admission-controller created
clusterrole.rbac.authorization.k8s.io/system:vpa-status-reader created
clusterrolebinding.rbac.authorization.k8s.io/system:vpa-status-reader-binding created
deployment.apps/vpa-updater created
deployment.apps/vpa-recommender created
Generating certs for the VPA Admission Controller in /tmp/vpa-certs.
Generating RSA private key, 2048 bit long modulus (2 primes)
.....+++++
.....+++++
e is 65537 (0x010001)
Generating RSA private key, 2048 bit long modulus (2 primes)
.....+++++
.....+++++
e is 65537 (0x010001)
Signature ok
subject=CN = vpa-webhook.kube-system.svc
Getting CA Private Key
Uploading certs to the cluster.
secret/vpa-tls-certs created
Deleting /tmp/vpa-certs.
deployment.apps/vpa-admission-controller created
service/vpa-webhook created
```

Vertical Pod Autoscaler

Kiểm tra cài đặt sử dụng lệnh

`kubectl get pods -n kube-system`

```
quanlm@quanlm-desktop:~$ kubectl get pods -n kube-system
```

NAME	READY	STATUS	RESTARTS	AGE
coredns-d67cc9d8f-6srkh	1/1	Running	0	24h
coredns-d67cc9d8f-g54sv	1/1	Running	0	24h
csi-bizflycloud-nodeplugin-rq8tw	2/2	Running	0	24h
csi-bizflycloud-nodeplugin-z4469	2/2	Running	0	24h
kube-router-ffstp	1/1	Running	0	24h
kube-router-xwzs9	1/1	Running	0	24h
metrics-server-5d79fcb458-6n2tl	1/1	Running	0	24h
metrics-server-5d79fcb458-tns2n	1/1	Running	0	24h
vpa-admission-controller-7467db745-5f5qg	1/1	Running	0	73s
vpa-recommender-597b7c765d-vzlcfc	1/1	Running	0	74s
vpa-updater-884d4d7d9-zbhtb	1/1	Running	0	74s

Vertical Pod Autoscaler

Triển khai ứng dụng demo

kubectl apply -f examples/hamster.yaml

```
quanlm@quanlm-desktop: ~/Work/workshop-bke/Day2/2.Afternoon/2. VPA/autoscaler/vertical-pod-autoscaler$ kubectl apply -f examples/hamster.yaml

verticalpodautoscaler.autoscaling.k8s.io/hamster-vpa created
deployment.apps/hamster created
```


Vertical Pod Autoscaler

Kiểm tra VPA sử dụng lệnh **kubectl get vpa**

```
quanlm@quanlm-desktop:~$ kubectl get vpa
```

NAME	MODE	CPU	MEM	PROVIDED	AGE
hamster-vpa	Auto	587m	262144k	True	6m2s

Vertical Pod Autoscaler

Kiểm tra Recommendation sử dụng lệnh **kubectl describe vpa hamster-vpa**

```
Status:
Conditions:
  Last Transition Time: 2023-12-06T03:41:24Z
  Status:              True
  Type:                RecommendationProvided
Recommendation:
  Container Recommendations:
    Container Name: hamster
    Lower Bound:
      Cpu:    358m
      Memory: 262144k
    Target:
      Cpu:    587m
      Memory: 262144k
    Uncapped Target:
      Cpu:    587m
      Memory: 262144k
    Upper Bound:
      Cpu:    1
      Memory: 500Mi
  Events: <none>
quanlm@quanlm-desktop:~$ |
```

Vertical Pod Autoscaler

Kiểm tra pod sử dụng lệnh: **kubectl describe pod hamster-c6967774f-t6dxv**

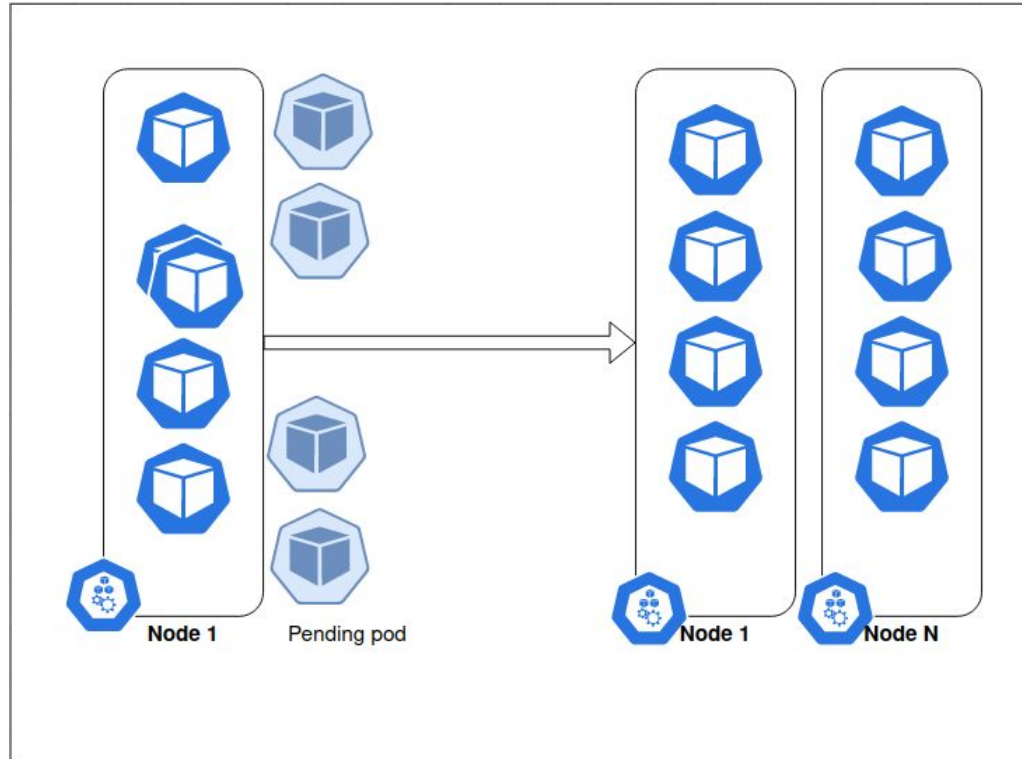
Tài nguyên đã giống với recommend

```
Restart Count: 0
Requests:
  cpu:          587m
  memory:       262144k
Environment:    <none>
Mounts:
```

Cluster Autoscaler

Cluster Autoscaler sẽ thêm hoặc xóa các node trong cụm để đảm bảo rằng có đủ tài nguyên để đáp ứng nhu cầu của các ứng dụng đang chạy trong cụm.

Cluster Autoscaler



Cluster Autoscaler

Nguyên lý hoạt động

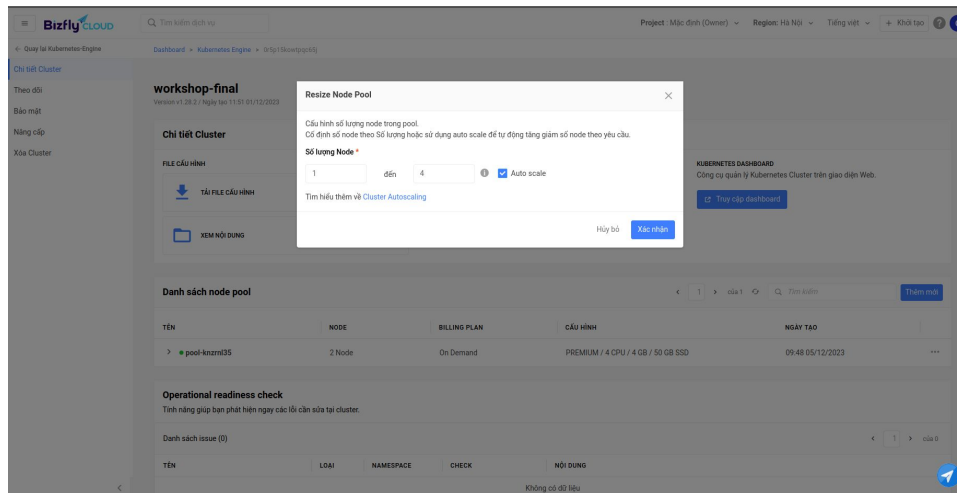
Cluster Autoscaler sẽ theo dõi các pod và node, trong trường hợp k thể schedule được pod do node thiếu tài nguyên sẽ thực hiện tăng số lượng node

Cluster Autoscaler

Lưu ý:

Cluster Autoscaler được quản lý bởi Bizfly Cloud

Khách hàng chỉ cần cấu hình lượng node tối thiểu - tối đa ở trên dashboard



Cluster Autoscaler

Truy cập thư mục Day2 -> 2. Afternoon -> 3. CA

Chạy lệnh **kubectl apply -f php.yaml** để triển khai ứng dụng demo

```
quanlm@quanlm-desktop: ~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get pod | grep php  
php-58b697f5f4-f98wg      1/1      Running    0          72s
```


Cluster Autoscaler

Tăng số lượng pod của deployment lên sử dụng lệnh **kubectl scale --replicas=10 deployment php**

```
deployment.apps/php scaled
● quanlm@quanlm-desktop: ~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get deployment
```

NAME	READY	UP-TO-DATE	AVAILABLE	AGE
external-dns	1/1	1	1	4d1h
hamster	2/2	2	2	39m
php	4/10	10	4	3m12s
wordpress	3/3	3	3	4d2h

Cluster Autoscaler

Kiểm tra pod, ta thấy có nhiều pod bị pending do thiếu tài nguyên

```
worupress mysql &
quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get pod | grep php
php-58b697f5f4-2hrl9      0/1      Pending    0          27s
php-58b697f5f4-4bsng      0/1      Pending    0          26s
php-58b697f5f4-4mdhr      0/1      Pending    0          27s
php-58b697f5f4-92zpn      1/1      Running    0          27s
php-58b697f5f4-dl9nk      0/1      Pending    0          27s
php-58b697f5f4-f98wg      1/1      Running    0          3m31s
php-58b697f5f4-ffbr9      0/1      Pending    0          26s
php-58b697f5f4-gfznt      1/1      Running    0          27s
php-58b697f5f4-vqbnd      0/1      Pending    0          27s
php-58b697f5f4-x4qrn      1/1      Running    0          27s
```

Cluster Autoscaler

Kiểm tra status của Cluster Autoscaler sử dụng lệnh **kubectl describe configmap cluster-autoscaler-status -n kube-system**

```
Annotations: cluster-autoscaler.kubernetes.io/last-updated: 2023-12-06 04:21:06.758150162 +0000 UTC
Data
=====
status:
----
Cluster-autoscaler status at 2023-12-06 04:21:06.758150162 +0000 UTC:
Cluster-wide:
  Health: Healthy (ready=2 unready=0 notStarted=0 longNotStarted=0 registered=2 longUnregistered=0)
           LastProbeTime: 2023-12-06 04:21:02.506364372 +0000 UTC m=+430041.545570126
           LastTransitionTime: 2023-12-01 04:54:17.894587224 +0000 UTC m=+36.933792979
  ScaleUp: InProgress (ready=2 registered=2)
           LastProbeTime: 2023-12-06 04:21:02.506364372 +0000 UTC m=+430041.545570126
           LastTransitionTime: 2023-12-06 04:20:22.191302942 +0000 UTC m=+430001.230508700
  ScaleDown: NoCandidates (candidates=0)
           LastProbeTime: 2023-12-06 04:21:02.506364372 +0000 UTC m=+430041.545570126
           LastTransitionTime: 2023-12-01 04:54:17.894587224 +0000 UTC m=+36.933792979

NodeGroups:
  Name: 656e8f77601d441cd41cccb3
  Health: Healthy (ready=2 unready=0 notStarted=0 longNotStarted=0 registered=2 longUnregistered=0 cloudProviderTarget=4 (m
inSize=2, maxSize=4))
           LastProbeTime: 2023-12-06 04:21:02.506364372 +0000 UTC m=+430041.545570126
           LastTransitionTime: 2023-12-06 04:12:59.324568317 +0000 UTC m=+429558.363774145
  ScaleUp: InProgress (ready=2 cloudProviderTarget=4)
           LastProbeTime: 2023-12-06 04:21:02.506364372 +0000 UTC m=+430041.545570126
           LastTransitionTime: 2023-12-06 04:20:22.191302942 +0000 UTC m=+430001.230508700
  ScaleDown: NoCandidates (candidates=0)
           LastProbeTime: 2023-12-06 04:21:02.506364372 +0000 UTC m=+430041.545570126
           LastTransitionTime: 2023-12-06 04:12:59.324568317 +0000 UTC m=+429558.363774145

BinaryData
=====

Events:
  Type    Reason      Age   From           Message
  ----    -
  Normal  ScaledUpGroup 48s   cluster-autoscaler Scale-up: setting group 656e8f77601d441cd41cccb3 size to 4 instead of 2 (max: 4)
  Normal  ScaledUpGroup 48s   cluster-autoscaler Scale-up: group 656e8f77601d441cd41cccb3 size set to 4 instead of 2 (max: 4)
```

Cluster Autoscaler

Kiểm tra ở trên dashboard

Danh sách node pool					
<div>< 1 > của 1 <input type="text" value="Tìm kiếm"/> <button>Thêm mới</button></div>					
TÊN	NODE	BILLING PLAN	CẤU HÌNH	NGÀY TẠO	
▼ ● pool-knznrl35	2 - 4 (4) Nodes	On Demand	PREMIUM / 4 CPU / 4 GB / 50 GB SSD	09:48 05/12/2023	***
● pool-knznrl35-0r5p15kowtpqc65j-n...	123.30.48.86 ▼				***
● pool-knznrl35-0r5p15kowtpqc65j-n...	123.30.48.104 ▼				***
● pool-knznrl35-0r5p15kowtpqc65j-n...	123.30.48.118 ▼				***
● pool-knznrl35-0r5p15kowtpqc65j-n...	123.30.48.117 ▼				***

Cluster Autoscaler

Sử dụng lệnh để kiểm tra nodes trong cụm Kubernetes sử dụng lệnh **kubectl get nodes**

```
quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get nodes
```

NAME	STATUS	ROLES	AGE	VERSION
pool-knznrl35-0r5p15kowtpqc65j-node-gw5cnyvn	Ready	<none>	25h	v1.28.2
pool-knznrl35-0r5p15kowtpqc65j-node-iryp6no4	Ready	<none>	2m	v1.28.2
pool-knznrl35-0r5p15kowtpqc65j-node-nxdbqsw9	Ready	<none>	2m	v1.28.2
pool-knznrl35-0r5p15kowtpqc65j-node-xewmluyc	Ready	<none>	25h	v1.28.2

Cluster Autoscaler

Kiểm tra pod sử dụng lệnh **kubectl get pod | grep php**, tất cả pod đã lên

```
quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get pod | grep php
php-58b697f5f4-2hr19      1/1      Running    0          4m29s
php-58b697f5f4-4bsng      1/1      Running    0          4m28s
php-58b697f5f4-4mdhr      1/1      Running    0          4m29s
php-58b697f5f4-92zpn      1/1      Running    0          4m29s
php-58b697f5f4-dl9nk      1/1      Running    0          4m29s
php-58b697f5f4-f98wg      1/1      Running    0          7m33s
php-58b697f5f4-ffbr9      1/1      Running    0          4m28s
php-58b697f5f4-gfznt      1/1      Running    0          4m29s
php-58b697f5f4-vqbnf      1/1      Running    0          4m29s
php-58b697f5f4-x4qrn      1/1      Running    0          4m29s
```

Cluster Autoscaler

Ta sử dụng lệnh **kubectl scale --replicas=1 deployment php** để đưa deployment về lại 1 pod

```
quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get deployment
```

NAME	READY	UP-TO-DATE	AVAILABLE	AGE
external-dns	1/1	1	1	4d1h
hamster	2/2	2	2	43m
php	1/1	1	1	7m51s
wordpress	3/3	3	3	4d2h

Cluster Autoscaler

Kiểm tra pod chỉ còn 1 pod

```
quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get pod | grep php  
php-58b697f5f4-ffbr9      1/1      Running    0          5m15s
```


Cluster Autoscaler

Kiểm tra configmap cluster-autoscaler-status, có thể thấy thực hiện xóa bớt node thừa không dùng đến

```
Events:
  Type    Reason             Age   From              Message
  ----    -
  Normal  ScaledUpGroup      15m   cluster-autoscaler  Scale-up: setting group 656e8f77601d441cd41cccb3 size to 4 instead of 2 (max: 4)
  Normal  ScaledUpGroup      15m   cluster-autoscaler  Scale-up: group 656e8f77601d441cd41cccb3 size set to 4 instead of 2 (max: 4)
  Normal  ScaleDownEmpty     25s   cluster-autoscaler  Scale-down: removing empty node pool-knznrl35-0r5p15kowtpqc65j-node-iryp6no4
  Normal  ScaleDownEmpty     22s   cluster-autoscaler  Scale-down: empty node pool-knznrl35-0r5p15kowtpqc65j-node-iryp6no4 removed
```

Cluster Autoscaler

Ta có thể thấy từ 4 -> 3 node

```
quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get nodes
```

NAME	STATUS	ROLES	AGE	VERSION
pool-knznrl35-0r5p15kowtpqc65j-node-gw5cnyvn	Ready	<none>	25h	v1.28.2
pool-knznrl35-0r5p15kowtpqc65j-node-nxdbqsw9	Ready	<none>	13m	v1.28.2
pool-knznrl35-0r5p15kowtpqc65j-node-xewmluyc	Ready	<none>	25h	v1.28.2

Cluster Autoscaler

Sau 1 lúc sẽ trở về trạng thái còn 2 node

```
• quanlm@quanlm-desktop:~/Work/workshop-bke/Day2/2.Afternoon/3. CA$ kubectl get nodes
```

NAME	STATUS	ROLES	AGE	VERSION
pool-knznrl35-0r5p15kowtpqc65j-node-gw5cnyvn	Ready	<none>	25h	v1.28.2
pool-knznrl35-0r5p15kowtpqc65j-node-xewmluyc	Ready	<none>	25h	v1.28.2

So sánh

	HPA	VPA	Cluster Autoscaler
Mục đích	Tăng/giảm số lượng pod của replicas	Tăng lượng tài nguyên cấp cho pod	Tăng/giảm số lượng node khi có nhiều/ít pod
Quản lý	Đơn giản, có sẵn trong Kubernetes	Phức tạp, cần cài đặt thêm nhiều thành phần để quản lý	Đơn giản, Bizflycloud đã quản lý

Thank you !

Q & A

More information

Bizfly Kubernetes Engine

<https://bizflycloud.vn/kubernetes-engine>

Container Registry

<https://bizflycloud.vn/container-registry>



Bizfly^{CLOUD} ❤️
kubernetes