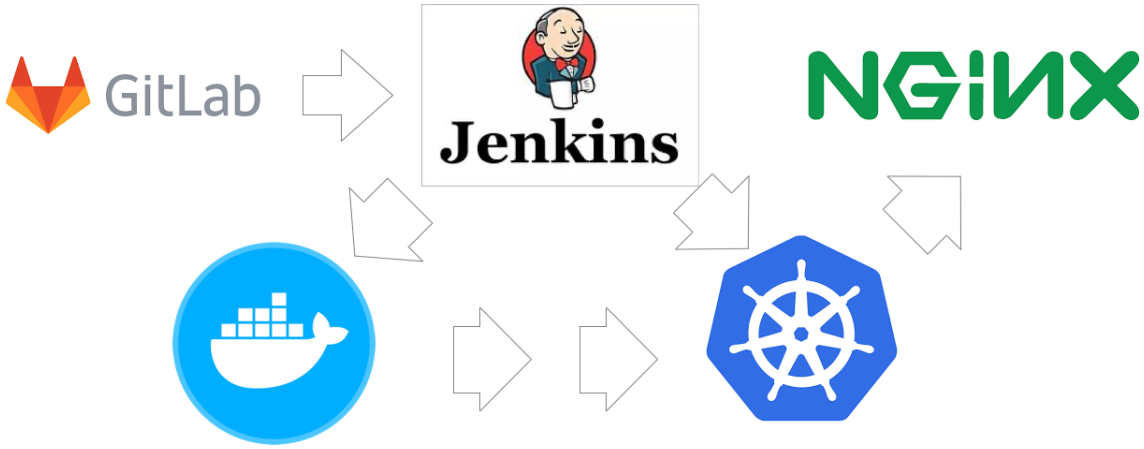
讲师(老司机)

容器虚拟化技术和自动化部署



k8s高可用-Sealos

Sealos

`Sealos` 是一个 Go 语言开发的简单干净且轻量的 `Kubernetes` 集群部署工具，`Sealos` 能很好的支持在生产环境中部署高可用的 `Kubernetes` 集群。

Sealos 特性与优势

1. 支持离线安装，工具与部署资源包分离，方便不同版本间快速升级。
2. 证书有效期默认延期至 99 年。
3. 工具使用非常简单。
4. 支持使用自定义配置文件，可灵活完成集群环境定制。
5. 使用内核进行本地负载，稳定性极高，故障排查也极其简单。
6. 最主要的优势是不需要翻墙出去！！！

为什么不用 KeepAlived 和 HAProxy 实现集群高可用

无论是通过 `KeepAlived` 还是 `HAProxy` 进行高可用集群调度都会存在以下一些劣势。

1. 软件源不一致可能导致容器中安装的软件版本也不一致，进而会引起相应检查脚本不生效等故障。
2. 可能因为系统依赖库问题，在某些特定环境下就直接无法完成安装。
3. 只依靠检测 `HAProxy` 进程是否存活是无法保证集群高可用的，正确的检测方式应该是判断 `ApiServer` 是否 `healthz` 状态。

4. `Keepalived` 可能存在 `Cpu` 占满的情况。

本地负载为什么不使用 Envoy 或者 Nginx 实现

`Sealos` 高可用实现是通过本地负载方式完成的。本地负载实现方式有多种，比如：`IPVS`、`Envoy`、`Nginx` 等，而 `Sealos` 采用的是通过内核 `IPVS` 来实现的。

> 本地负载：在每个 `Node` 节点上都启动一个负载均衡，同时监听集群中的多个 `Master` 节点。

`Sealos` 选择通过内核 `IPVS` 来实现主要有以下几个原因：

- 如果使用 `Envoy` 等需要在每个节点上都跑一个进程，消耗更多资源。虽然 `IPVS` 实际上也会多跑一个 `lvscare` 进程，但是 `lvscare` 只是负责管理 `IPVS` 规则，原理和 `Kube-Proxy` 类似。真正的流量直接从内核层面走，不需要把数据包先走到用户态中去处理。
- 使用 `Envoy` 存在启动优先级的问题，比如：Join 集群时，如果负载均衡没有建立，Kubelet 就会启动失败。使用 `IPVS` 则不会存在这样的问题，因为我们可以在 Join 集群前先建立好转发规则。

为什么要定制 Kubeadm

- 解决默认证书有效期只有一年的问题。
- 更方便的实现本地负载。
- 核心的功能均集成到 Kubeadm 中了，Sealos 只管分发和执行上层命令，相对就更轻量了。

Sealos 执行流程

1. 通过 `SFTP` 或者 `Wget` 命令把离线安装包拷贝到目标机器上，包括所有 `Master` 和 `Node` 节点。
2. 在 `Master 0` 节点上执行 `kubeadm init` 命令。
3. 在其它 `Master` 节点上执行 `kubeadm join` 命令并设置控制面。这个过程中多个 `Master` 节点上的 `Etcd` 会自动组成一个 `Etcd` 集群，并启动相应控制组件。
4. 所有 `Node` 节点都加入到集群中，这个过程中会在 `Node` 节点上进行 `IPVS` 转发规则和 `/etc/hosts` 配置。

`Node` 节点对 `ApiServer` 的访问均是通过域名进行的。因为 `Node` 节点需要通过 `虚拟 IP` 连接到多个 `Master` 上，但是每个 `Node` 节点的 `Kubelet` 与 `Kube-Proxy` 访问 `ApiServer` 的地址是不同的，所以这里使用域名来解析每个节点上 `ApiServer` 不同的 `IP` 地址。

部署准备

sealos官网

```
官网地址:
https://sealyun.com/

企业级应用的集群离线包需要付费，我们只是学习，使用作者提供的免费离线包:
http://store.lameleg.com/
```

安装文档

```
官网中文安装手册：
https://github.com/fanux/sealos
```

sealos下载

```
可以下载sealos不同版本的二进制文件
https://github.com/fanux/sealos/releases
```

K8S离线包

离线包包含所有二进制文件配置文件和镜像

```
非免费V1.16.0版本下载地址
https://sealyun.oss-cn-beijing.aliyuncs.com/cf6bece970f6dab3d8dc8bc5b588cc18-1.16.0/kube1.16.0.tar.gz

免费V1.17.0版本下载地址
https://sealyun.oss-cn-beijing.aliyuncs.com/413bd3624b2fb9e466601594b4f72072-1.17.0/kube1.17.0.tar.gz

非免费V1.18.0版本下载地址：
https://sealyun.oss-cn-beijing.aliyuncs.com/7b6af025d4884fdd5cd51a674994359c-1.18.0/kube1.18.0.tar.gz

免费V1.18.1版本下载地址：
https://sealyun.oss-cn-beijing.aliyuncs.com/7b6af025d4884fdd5cd51a674994359c-1.18.0/kube1.18.0.tar.gz

非免费V1.18.2版本下载地址：
https://sealyun.oss-cn-beijing.aliyuncs.com/9a8299ea8016abe32e1564a44d5162e4-1.18.2/kube1.18.2.tar.gz

```

Sealos 常用参数说明

```
--master   Master 节点服务器地址列表
--node     Node 节点服务器地址列表
--user     服务器 SSH 用户名
--passwd   服务器 SSH 用户密码
--pkg-url  离线包所在位置，可以是本地目录，也可以是一个 HTTP 地址
--version  指定需要部署的 Kubernetes 版本
--pk       指定 SSH 私钥所在位置，默认为 /root/.ssh/id_rsa

Other flags:

 --kubeadm-config string   kubeadm-config.yaml 用于指定自定义 kubeadm 配置文件
 --vip string              virtual ip (default "10.103.97.2") 本地负载时虚拟 IP ，不推荐修改，集群外不可访问
```

部署多主节点的高可用集群

节点信息

服务器用户名：root，服务器密码：123456

主机名	IP地址
sealos-master01	192.168.198.121
sealos-master02	192.168.198.122
sealos-master03	192.168.198.123
sealos-node01	192.168.198.124
sealos-node02	192.168.198.125
sealos-node03	192.168.198.126

安装相关环境依赖

```
通过 Sealos 进行 Kubernetes 集群部署，你需要先准备好以下环境。

1.在所有要部署的机器上，先完成 Docker 的安装和启动。
2.下载 Kubernetes 离线安装包。
3.下载最新版本 Sealos。
4.对所有服务器进行时间同步。
5.系统支持：centos7.2以上，ubuntu16.04以上。内核推荐4.14以上。推荐配置：centos7.8
6.主机名不可重复
7.master节点CPU必须2C以上
8.请使用sealos 3.2.0以上版本
```

升级系统内核

```
rpm -Uvh http://www.elrepo.org/elrepo-release-7.0-3.el7.elrepo.noarch.rpm
yum --enablerepo=elrepo-kernel install -y kernel-lt
grep initrd16 /boot/grub2/grub.cfg
grub2-set-default 0

reboot
```

修改Cgroup Driver

```
修改/etc/docker/daemon.json，新增：

"exec-opts": ["native.cgroupdriver=systemd"]

重启docker服务：
systemctl daemon-reload
systemctl restart docker
查看修改后状态：
docker info | grep Cgroup
```

设置hostname

```
cat <<EOF >>/etc/hosts
192.168.198.121 sealos-master01
192.168.198.122 sealos-master02
192.168.198.123 sealos-master03
192.168.198.124 sealos-node01
192.168.198.125 sealos-node02
192.168.198.126 sealos-node03
EOF
```

初始化安装

多master

```
将sealos二进制文件上传sealos-master01:/data
将kubernates离线安装包上传sealos-master01:/data
cd /data

授权并移动到/usr/bin目录中
chmod +x sealos && mv sealos /usr/bin

多master HA:
sealos init \
    --master 192.168.198.121 \
    --master 192.168.198.122 \
    --master 192.168.198.123 \
    --node 192.168.198.124 \
```

```
    --node 192.168.198.125 \
    --node 192.168.198.126 \
    --user root \
    --passwd 123456 \
    --version v1.17.0 \
    --pkg-url /data/kube1.17.0.tar.gz

单master多node:
sealos init --master 192.168.198.121 \
    --node 192.168.198.124 \
    --node 192.168.198.125 \
    --node 192.168.198.126 \
    --user root \
    --passwd 123456 \
    --version v1.17.0 \
    --pkg-url /data/kube1.17.0.tar.gz
```

kubectl命令自动补全

sealos默认已经帮我们安装命令补全功能。

```
echo "source <(kubectl completion bash)" >> ~/.bash_profile
source ~/.bash_profile
```

master节点操作

增加master

```
sealos join --master 192.168.198.127 --master 192.168.198.128

或者多个连续IP
sealos join --master 192.168.198.127-192.168.198.128
```

删除指定master节点

```
sealos clean --master 192.168.198.122 --master 192.168.198.123

或者多个连续IP
sealos clean --master 192.168.198.122-192.168.198.123
```

node节点操作

增加node

```
sealos join --node 192.168.198.127 --node 192.168.198.128

或者多个连续IP
sealos join --node 192.168.198.127-192.168.198.128
```

删除指定node节点

```
sealos clean --node 192.168.198.125 --node 192.168.198.126

或者多个连续IP
sealos clean --node 192.168.198.125-192.168.198.126
```

清理集群

```
sealos clean --all -f
```